13. A method of securing a data signal comprising:

applying a data reduction technique to reduce the data signal into a reduced data signal;

subtracting said reduced data signal from the data signal to produce a remainder signal;

using a first cryptographic technique to encrypt the reduced data signal to produce an encrypted, reduced data signal;

using a second cryptographic technique to encrypt the remainder signal to produce an encrypted remainder signal; and

adding said encrypted, reduced data signal to said encrypted remainder signal to produce an output signal.

14. The method of claim 13 wherein said first and second cryptographic techniques are identical.

15. The method of claim 13 wherein at least one of said first and second cryptographic techniques is a watermarking technique.

16. The method of claim 15, wherein at least one of the watermarks is embedded using at least one key.

17. The method of claim 15, wherein at least one of the watermarks is embedded using a key pair.

18. The method of claim 13 wherein at least one of said first and second cryptographic techniques is a scrambling technique.

19. The method of claim 13 wherein one of said first and second cryptographic techniques is a watermarking technique and the other is a scrambling technique.

20. The method of claim 13 wherein said first and second cryptographic techniques are identical.

21. A system for securing a data signal comprising:

means to apply a data reduction technique to reduce the data signal into a reduced data signal;

means to subtract said reduced data signal from the data signal to produce a remainder signal;

means to apply a first cryptographic technique to encrypt the reduced data signal to produce an encrypted, reduced data signal;

means to apply a second cryptographic technique to encrypt the remainder signal to produce an encrypted remainder signal; and

means to add said encrypted, reduced data signal to said encrypted remainder signal to produce an output signal.

22. The system of claim 21 wherein said first and second cryptographic techniques are identical.

23. The system of claim 21 wherein at least one of said means to apply a first and second cryptographic technique utilizes a watermarking technique.

24. The system of claim 21 wherein at least one of said means to apply a first and second cryptographic technique utilizes a scrambling technique.

25. The system of claim 13 wherein said means to apply a first cryptographic technique is a means to apply a watermarking technique and said means to apply a second cryptographic technique is a means to apply a scrambling technique.

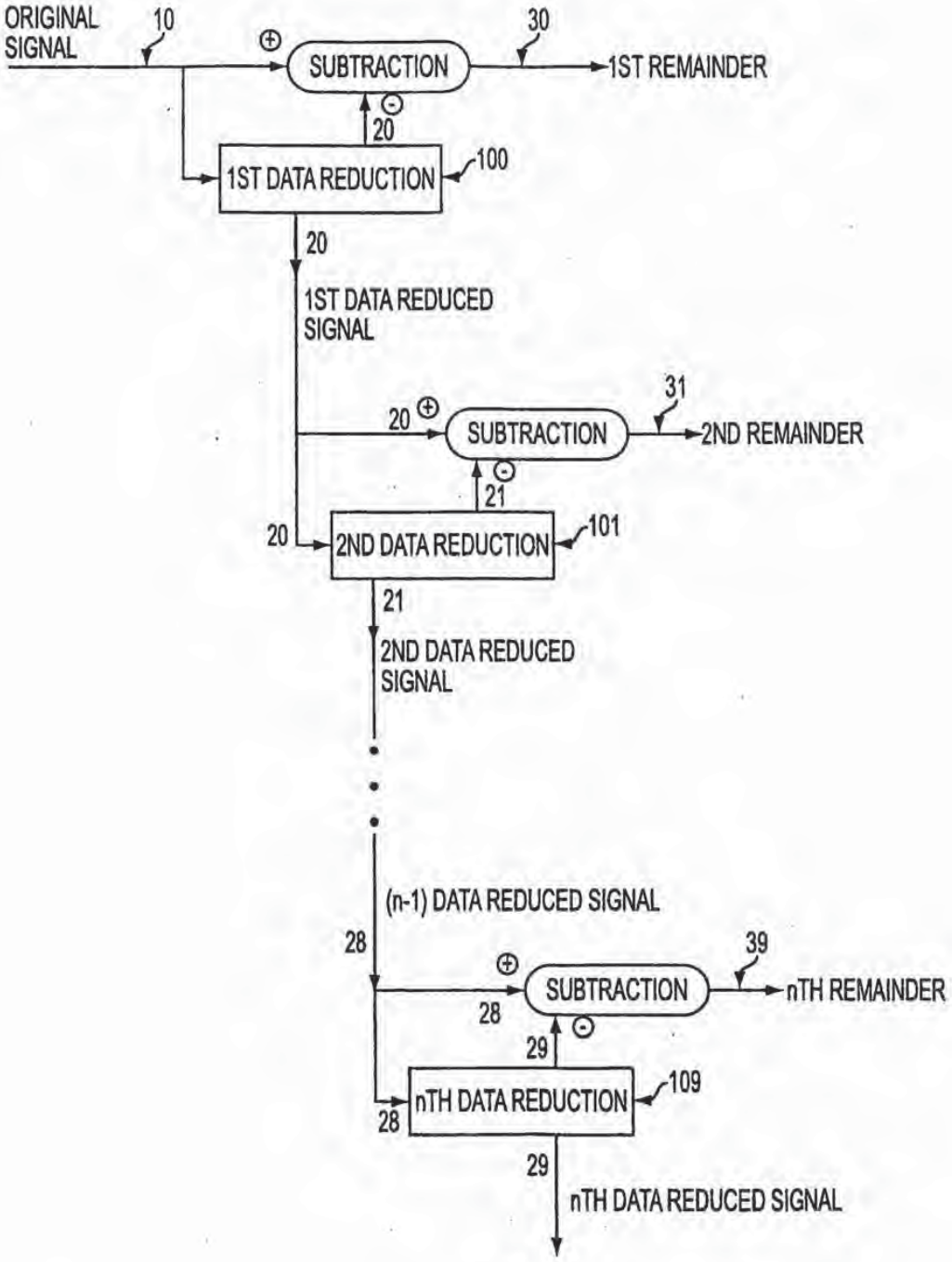


FIG. 1

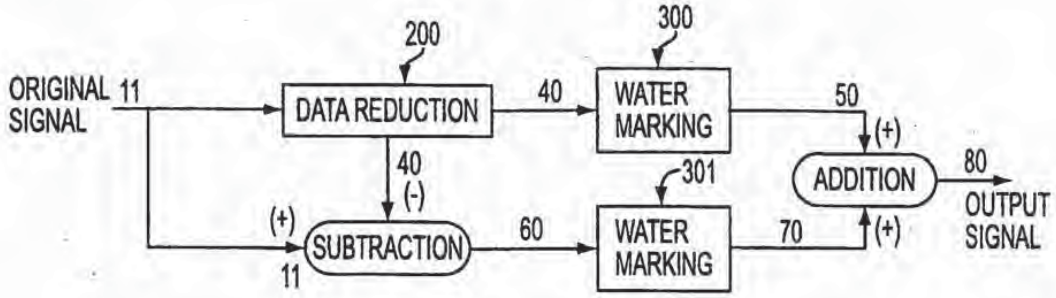


FIG. 2

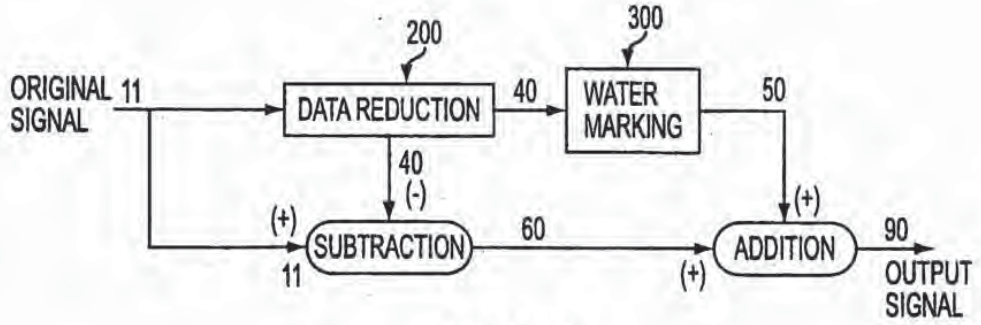


FIG. 3

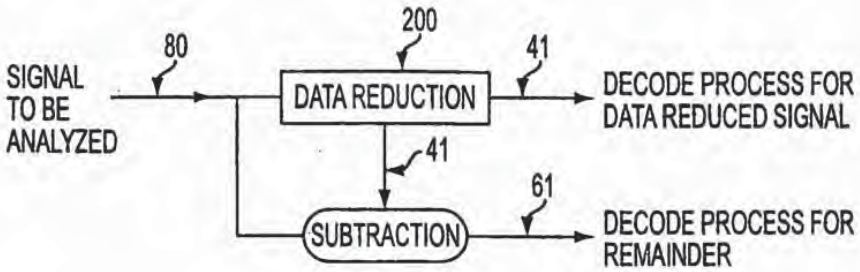

200
80
41
SIGNAL TO BE ANALYZED
DATA REDUCTION
DECODE PROCESS FOR DATA REDUCED SIGNAL
41
61
SUBTRACTION
DECODE PROCESS FOR REMAINDER

FIG. 4

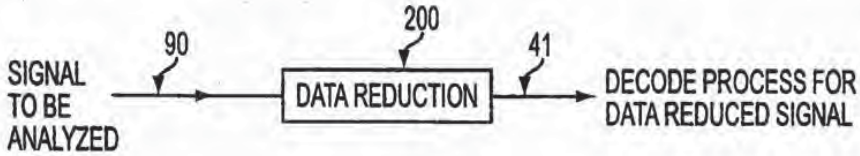

200
90
41
SIGNAL TO BE ANALYZED
DATA REDUCTION
DECODE PROCESS FOR DATA REDUCED SIGNAL

FIG. 5

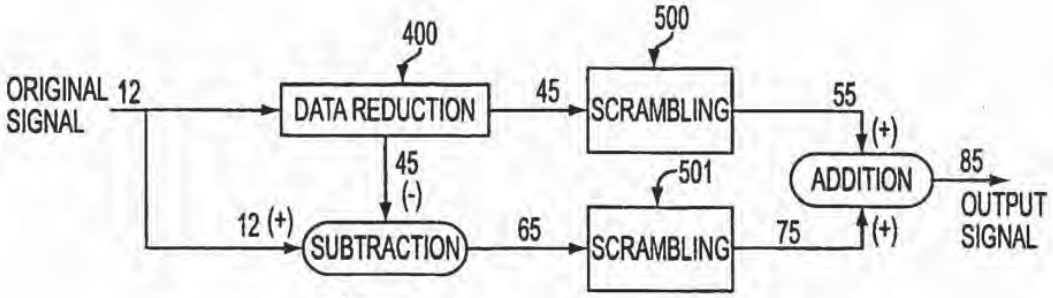


FIG. 6

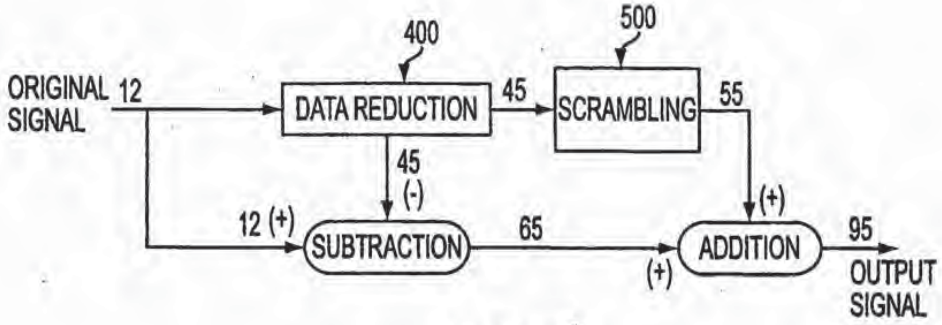


FIG. 7

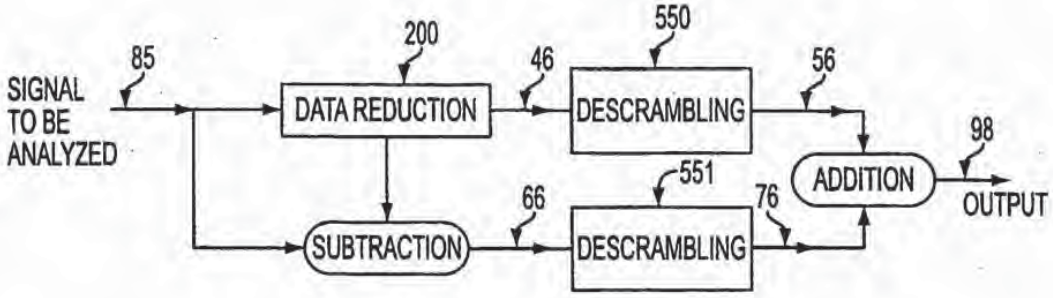

SIGNAL TO BE ANALYZED
85
200
DATA REDUCTION
46
550
DESCRAMBLING
56
98
ADDITION
OUTPUT
551
66
SUBTRACTION
DESCRAMBLING
76

FIG. 8

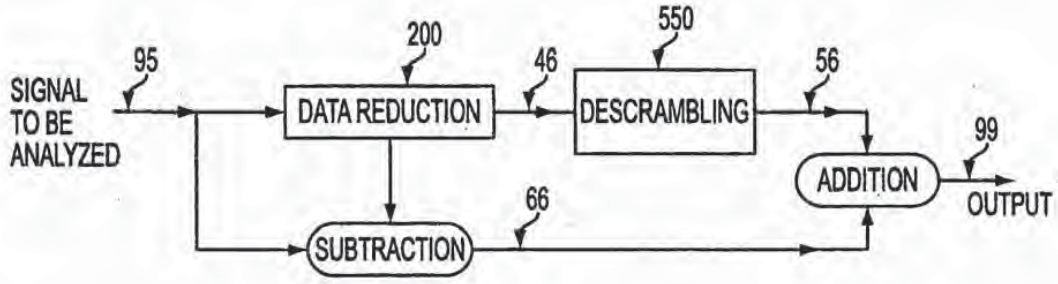

SIGNAL TO BE ANALYZED
95
200
DATA REDUCTION
46
550
DESCRAMBLING
56
99
ADDITION
OUTPUT
66
SUBTRACTION

FIG. 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/06522

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : H04N 7/167
US CL. : 713/176
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 380/200,206,207,237,238; 705/54; 704/216-218, 226-228, 500, 501, 503,504; 713/176; 360/49; 348/461, 462

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Watermark Digest: Art Unit 2767

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
IEEE, EAST, Internet, Dialog

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,E	US 6,061,793 A [TEWFIK et al.] 09 MAY 2000, Entire Document	1-25
X	US 5,809,139 A [GIROD et al.] 15 SEPTMBER 1998, Entire Document	1-25
X	US 5,848,155 A [COX] 08 DECEMBER 1998, Entire Document	1-25
A,P	US 5,889,868 A [MOSKOWITZ et al.] 30 MARCH 1999, Entire Document	1-25
A,P	US 5,915,027 A [COX et al.] 22 JUNE 1999, Entire Document	1-25
A,P	US 5,940,134 A [WIRTZ] 17 AUGUST 1999, Entire Document	1-25

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 30 JUNE 2000	Date of mailing of the international search report 18 AUG 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer PAUL E. CALLAHAN Telephone No. (703) [illegible]

Form PCT/ISA/210 (second sheet) (July 1998)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/06522

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A,P	US 5,991,426 A [COX et al.] 23 NOVEMBER 1999, Entire Document	1-25
A,E	US 6,069,914 A [COX] 30 MAY 2000, Entire Document	1-25
A,P	US 5,943,422 A [VAN WIE et al.] 24 AUGUST 1999, Entire Document	1-25

Form PCT/ISA/210 (continuation of second sheet) (July 1998)*

European Patent Office

SUPPLEMENTARY EUROPEAN SEARCH REPORT

Application Number
EP 00 91 9398

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	WO 98 37513 A (TELSTRA R & D MAN PTY LTD ;BIGGAR MICHAEL (AU); JOHNSON ANDREW (AU) 27 August 1998 (1998-08-27) * page 6, line 25 - page 7, line 10 *	6	H04N7/167 H04N7/26 H04N1/32 G06F17/30
Y	US 4 969 204 A (MELNYCHUCK PAUL W ET AL) 6 November 1990 (1990-11-06) * column 2, line 9 - column 2, line 48 *	1-10	
Y	EP 0 651 554 A (EASTMAN KODAK CO) 3 May 1995 (1995-05-03) * column 6, line 43 - column 9, line 19; figure 2 *	1-10	
A	JOHNSON A ET AL: "TRANSFORM PERMUTED WATERMARKING FOR COPYRIGHT PROTECTION OF DIGITAL VIDEO" IEEE GLOBECOM 1998. GLOBECOM '98. THE BRIDGE TO GLOBAL INTEGRATION. SYDNEY, NOV. 8 - 12, 1998, IEEE GLOBAL TELECOMMUNICATIONS CONFERENCE, NEW YORK, NY: IEEE, US, vol. 2, 1998, pages 684-689, XP000825846 ISBN: 0-7803-4985-7 * page 685, left-hand column, paragraph 2 - page 685, left-hand column, paragraph 3 *	1-10	
P,X	WO 99 62044 A (HANDEL THEODORE G ;UNIV CALIFORNIA (US); SANDFORD MAXELL T II (US)) 2 December 1999 (1999-12-02) * abstract * * page 4, line 17 - page 5, line 5 *	6	

TECHNICAL FIELDS SEARCHED (Int.Cl.7)
H04N
G06F

The supplementary search report has been based on the last set of claims valid and available at the start of the search.

Place of search	Date of completion of the search	Examiner
MUNICH	27 June 2002	Schoeyer, M

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C04)

3

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 January 2002 (10.01.2002)

PCT

(10) International Publication Number
WO 02/03385 A1

(51) International Patent Classification[7]: G11B 20/00, G06F 1/00

(21) International Application Number: PCT/US00/18411

(22) International Filing Date: 5 July 2000 (05.07.2000)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant and
(72) Inventor: **MOSKOWITZ, Scott, A.** [US/US]; 16711 Collins Avenue #2505, Miami, FL 33160 (US).

(74) Agents: **CHAPMAN, Floyd, B.** et al.; Wiley Rein & Fielding, Intellectual Property Department, 1776 K Street, N.W., Washington, DC 20006 (US).

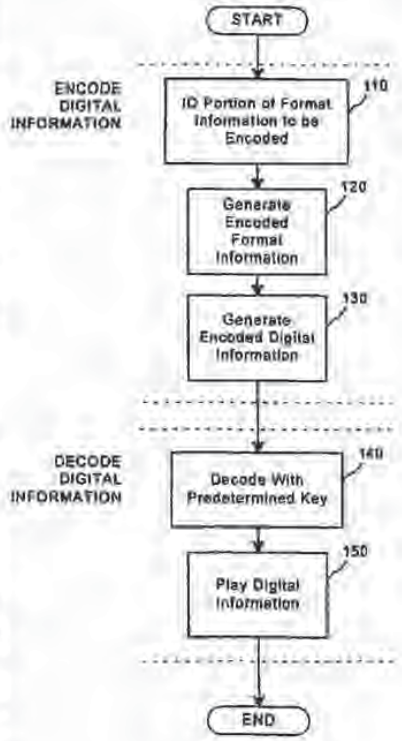
(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: COPY PROTECTION OF DIGITAL DATA COMBINING STEGANOGRAPHIC AND CRYPTOGRAPHIC TECHNIQUES



(57) Abstract: A method for combining transfer functions with predetermined key creation. In one embodiment, digital information, including a digital sample and format information, is protected by identifying and encoding a portion of the format information. Encoded digital information, including the digital sample and the encoded format information generated to protect the original digital information. In another embodiment, a digital signal, including digital samples in a file format having an inherent granularity, is protected by creating a predetermined key. The predetermined key is comprised of a transfer function-based mask set to manipulate data at the inherent granularity of the file format of the underlying digitized samples.

WO 02/03385 A1

COPY PROTECTION OF DIGITAL DATA COMBINING STEGANOGRAPHIC AND CRYPTOGRAPHIC TECH NIQUES

BACKGROUND OF THE INVENTION

Increasingly, commercially valuable information is being created and stored in "digital" form. For example, music, photographs and video can all be stored and transmitted as a series of numbers, such as 1's and 0's. Digital techniques let the original information be recreated in a very accurate manner. Unfortunately, digital techniques also let the information be easily copied without the information owner's permission.

Because unauthorized copying is clearly a disincentive to the digital distribution of valuable information, it is important to establish responsibility for copies and derivative copies of such works. For example, if each authorized digital copy of a popular song is identified with a unique number, any unauthorized copy of the song would also contain the number. This would allow the owner of the information, such as a song publisher, to investigate who made the unauthorized copy. Unfortunately, it is possible that the unique number could be erased or altered if it is simply tacked on at the beginning or end of the digital information.

As will be described, known digital "watermark" techniques give creators and publishers of digitized multimedia content localized, secured identification and authentication of that content. In considering the various forms of multimedia content, such as "master," stereo, National Television Standards Committee (NTSC) video, audio tape or compact disc, tolerance of quality will vary with individuals and affect the underlying commercial and aesthetic value of the content. For example, if a digital version of a popular song sounds distorted, it will be less valuable to users. It is therefore desirable to embed copyright, ownership or purchaser information, or some combination of these and related data, into the content in a way that will damage the content if the watermark is removed without authorization.

To achieve these goals, digital watermark systems insert ownership information in a way that causes little or no noticeable effects, or "artifacts," in the underlying content signal. For example, if a digital watermark is inserted into a

digital version of a song, it is important that a listener not be bothered by the slight changes introduced by the watermark. It is also important for the watermark technique to maximize the encoding level and "location sensitivity" in the signal to force damage to the content signal when removal is attempted. Digital watermarks address many of these concerns, and research in the field has provided extremely robust and secure implementations.

What has been overlooked in many applications described in the art, however, are systems which closely mimic distribution of content as it occurs in the real world. For instance, many watermarking systems require the original un-watermarked content signal to enable detection or decode operations. These include highly publicized efforts by NEC, Digimarc and others. Such techniques are problematic because, in the real world, original master copies reside in a rights holders vaults and are not readily available to the public.

With much activity overly focused on watermark survivability, the security of a digital watermark is suspect. Any simple linear operation for encoding information into a signal may be used to erase the embedded signal by inverting the process. This is not a difficult task, especially when detection software is a plug-in freely available to the public, such as with Digimarc. In general, these systems seek to embed cryptographic information, not cryptographically embed information into target media content.

Other methods embed ownership information that is plainly visible in the media signal, such as the method described in US Patent No. 5,530,739 to Braudaway et al. The system described in Braudaway protects a digitized image by encoding a visible watermark to deter piracy. Such an implementation creates an immediate weakness in securing the embedded information because the watermark is plainly visible. Thus, no search for the embedded signal is necessary and the watermark can be more easily removed or altered. For example, while certainly useful to some rights owners, simply placing the symbol "©" in the digital information would only provide limited protection. Removal by adjusting the brightness of the pixels forming the "©" would not be difficult with respect to the computational resources required.

Other relevant prior art includes US Patents No. 4,979,210 and 5,073,925 to Nagata et al., which encodes information by modulating an audio signal in the amplitude/time domain. The modulations introduced in the Nagata process carry a "copy/don't copy" message, which is easily found and circumvented by one skilled in the art. The granularity of encoding is fixed by the amplitude and frequency modulation limits required to maintain inaudibility. These limits are relatively low, making it impractical to encode more information using the Nagata process.

Although US Patent No. 5,664,018 to Leighton describes a means to prevent collusion attacks in digital watermarks, the disclosed method may not actually provide the security described. For-example, in cases where the watermarking technique is linear, the "insertion envelope" or "watermarking space" is well-defined and thus susceptible to attacks less sophisticated than collusion by unauthorized parties. Over-encoding at the watermarking encoding level is but one simple attack in such linear implementations. Another consideration not made by Leighton is that commercially-valuable content may already exist in a un-watermarked form somewhere, easily accessible to potential pirates, gutting the need for any type of collusive activity. Digitally signing the embedded signal with preprocessing of watermark data is more likely to prevent successful collusion. Furthermore, a "baseline" watermark as disclosed is quite subjective. It is simply described elsewhere in the art as the "perceptually significant" regions of a signal. Making a watermarking function less linear or inverting the insertion of watermarks would seem to provide the same benefit without the additional work required to create a "baseline" watermark. Indeed, watermarking algorithms should already be capable of defining a target insertion envelope or region without additional steps. What is evident is the Leighton patent does not allow for initial prevention of attacks on an embedded watermark as the content is visibly or audibly unchanged.

It is also important that any method for providing security also function with broadcasting media over networks such as the Internet, which is also referred to as "streaming." Commercial "plug-in" products such as RealAudio and RealVideo, as well as applications by vendors VDONet and Xtreme, are common in such network environments. Most digital watermark implementations focus on

common file base signals and fail to anticipate the security of streamed signals. It is desirable that any protection scheme be able to function with a plug-in player without advanced knowledge of the encoded media stream.

Other technologies focus solely on file-based security. These technologies illustrate the varying applications for security that must be evaluated for different media and distribution environments. Use of cryptolopes or cryptographic containers, as proposed by IBM in its Cryptolope product, and InterTrust, as described in U.S. Patents No. 4,827,508, 4,977,594, 5,050,213 and 5,410,598, may discourage certain forms of piracy. Cryptographic containers, however, require a user to subscribe to particular decryption software to decrypt data. IBM's InfoMarket and InterTrust's DigiBox, among other implementations, provide a generalized model and need proprietary architecture to function. Every user must have a subscription or registration with the party which encrypts the data. Again, as a form of general encryption, the data is scrambled or encrypted without regard to the media and its formatting. Finally, control over copyrights or other neighboring rights is left with the implementing party, in this case, IBM, InterTrust or a similar provider. Methods similar to these "trusted systems" exist, and Cerberus Central Limited and Liquid Audio, among a number of companies, offer systems which may functionally be thought of as subsets of IBM and InterTrust's more generalized security offerings. Both Cerberus and Liquid Audio propose proprietary player software which is registered to the user and "locked" in a manner parallel to the locking of content that is distributed via a cryptographic container. The economic trade-off in this model is that users are required to use each respective companies' proprietary player to play or otherwise manipulate content that is downloaded. If, as is the case presently, most music or other media is not available via these proprietary players and more companies propose non-compatible player formats, the proliferation of players will continue. Cerberus and Liquid Audio also by way of extension of their architectures provide for "near-CD quality" but proprietary compression. This requirement stems from the necessity not to allow content that has near-identical data make-up to an existing consumer electronic standard, in Cerberus and Liquid Audio's case the so-called Red Book audio CD standard of 16 bit 44.1 kHz, so that comparisons with the proprietary file may not

yield how the player is secured. Knowledge of the player's file format renders its security ineffective as a file may be replicated and played on any common player, not the intended proprietary player of the provider of previously secured and uniquely formatted content. This is the parallel weakness to public key crypto-systems which have gutted security if enough plain text and cipher text comparisons enable a pirate to determine the user's private key.

Many approaches to digital watermarking leave detection and decoding control with the implementing party of the digital watermark, not the creator of the work to be protected. A set of secure digital watermark implementations address this fundamental control issue forming the basis of key-based approaches. These are covered by the following patents and pending applications, the entire disclosures of which are hereby incorporated by reference: US Patent No. 5,613, 004 entitled "Steganographic Method and Device" and its derivative US patent application Serial No. 08/775,216, US patent application Serial No. 08/587,944 entitled "Human Assisted Random Key Generation and Application for Digital Watermark System," US Patent Application Serial No. 08/587,943 entitled "Method for Stega-Cipher Protection of Computer Code," US patent application Serial No. 08/677,435 entitled "Optimization Methods for the Insertion, Protection, and Detection of Digital Watermarks in Digitized Data," and US Patent Application Serial No. 08/772,222 entitled "Z-Transform Implementation of Digital Watermarks." Public key crypto-systems are described in US Patents No. 4,200,770, 4,218,582, 4,405,829 and 4,424,414, the entire disclosures of which are also hereby incorporated by reference.

In particular, an improved protection scheme is described in "Method for Stega-Cipher Protection of Computer Code," US patent application Serial No. 08/587,943. This technique uses the key-based insertion of binary executable computer code within a content signal that is subsequently, and necessarily, used to play or otherwise manipulate the signal in which it is encoded. With this system, however, certain computational requirements, such as one digital player per digital copy of content, may be necessitated. For instance, a consumer may download many copies of watermarked content. With this technique, the user would also be downloading as many copies of the digital player program. While this form of

security may be desirable for some applications, it is not appropriate in many circumstances. Finally, even when digital information is distributed in encoded form, it may be desirable to allow unauthorized users to play the information with a digital player, perhaps with a reduced level of quality. For example, a popular song may be encoded and freely distributed in encoded form to the public. The public, perhaps using commonly available plug-in digital players, could play the encoded content and hear the music in some degraded form. The music may sound choppy, or fuzzy or be degraded in some other way. This lets the public decide, based on the available lower quality version of the song, if they want to purchase a key from the publisher to decode, or "clean-up," the content. Similar approaches could be used to distribute blurry pictures or low quality video. Or even "degraded" text, in the sense that only authenticated portions of the text can be determined with the predetermined key or a validated digital signature for the intended message.

In view of the foregoing, it can be appreciated that a substantial need exists for a method allowing encoded content to be played, with degraded quality, by a plug-in digital player, and solving the other problems discussed above.

SUMMARY OF THE INVENTION

The disadvantages of the art are alleviated to a great extent by a method for combining transfer functions with predetermined key creation. In one embodiment, digital information, including a digital sample and format information, is protected by identifying and encoding a portion of the format information. Encoded digital information, including the digital sample and the encoded format information, is generated to protect the original digital information.

In another embodiment, a digital signal, including digital samples in a file format having an inherent granularity, is protected by creating a predetermined key. The predetermined key is comprised of a transfer function-based mask set to manipulate data at the inherent granularity of the file format of the underlying digitized samples.

With these and other advantages and features of the invention that will become hereinafter apparent, the nature of the invention may be more clearly understood by reference to the following detailed description of the invention, the appended claims and to the several drawings attached herein.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block flow diagram of a method for copy protection or authentication of digital information according to an embodiment of the present invention.

DETAILED DESCRIPTION

In accordance with an embodiment of the present invention, a method combines transfer functions with predetermined key creation. Increased security is achieved in the method by combining elements of "public-key steganography" with cryptographic protocols, which keep in-transit data secure by scrambling the data with "keys" in a manner that is not apparent to those with access to the content to be distributed. Because different forms of randomness are combined to offer robust, distributed security, the present invention addresses an architectural "gray space" between two important areas of security: digital watermarks, a subset of the more general art of steganography, and cryptography. One form of randomness exists in the mask sets that are randomly created to map watermark data into an otherwise unrelated digital signal. The second form of-randomness is the random permutations of data formats used with digital players to manipulate the content with the predetermined keys. These forms can be thought of as the transfer function versus the mapping function inherent to digital watermarking processes.

According to an embodiment of the present invention, a predetermined, or randomly generated, key is used to scramble digital information in a way that is unlike known "digital watermark" techniques and public key crypto-systems. As used herein, a key is also referred to as a "mask set" which includes one or more random or pseudo-random series of bits. Prior to encoding, a mask can be generated by any cryptographically secure random generation process. A block cipher, such as a Data Encryption Standard (DES) algorithm, in combination with a sufficiently random seed value, such as one created using a Message Digest 5 (MD5) algorithm, emulates a cryptographically secure random bit generator. The keys are saved in a database, along with information matching them to the digital signal, for use in descrambling and subsequent viewing or playback. Additional file format or transfer property information is prepared and made available to the encoder, in a bit addressable manner. As well, any authenticating function can be

combined, such as Digital Signature Standard (DSS) or Secure Hash Algorithm (SHA).

Using the predetermined key comprised of a transfer function-based mask set, the data representing the original content is manipulated at the inherent granularity of the file format of the underlying digitized samples. Instead of providing, or otherwise distributing, watermarked content that is not noticeably altered, a partially "scrambled" copy of the content is distributed. The key is necessary both to register the sought-after content and to descramble the content into its original form.

The present invention uses methods disclosed in "Method for Stega-Cipher Protection of Computer Code," US Patent Application Serial No. 08/587,943, with respect to transfer functions related to the common file formats, such as PICT, TIFF, AIFF, WAV, etc. Additionally, in cases where the content has not been altered beyond being encoded with such functional data, it is possible for a digital player to still play the content because the file format has not been altered. Thus, the encoded content could still be played by a plug-in digital player as discrete, digitally sampled signals, watermarked or not. That is, the structure of the file can remain basically unchanged by the watermarking process, letting common file format based players work with the "scrambled" content.

For example, the Compact Disc-Digital Audio (CD-DA) format stores audio information as a series of frames. Each frame contains a number of digital samples representing, for example, music, and a header that contains file format information. As shown in FIG. 1, according to an embodiment of the present invention some of the header information can be identified and "scrambled" using the predetermined key at steps 110 to 130. The music samples can remain unchanged. Using this technique, a traditional CD-DA player will be able to play a distorted version of the music in the sample. The amount of distortion will depend on the way, and extent, that the header, or file format, information has been scrambled. It would also be possible to instead scramble some of the digital samples while leaving the header information alone. In general, the digital signal would be protected by manipulating data at the inherent granularity, or "frames," of the CD-

DA file format. To decode the information, a predetermined key is used before playing the digital information at steps 140 and 150.

A key-based decoder can act as a "plug-in" digital player of broadcast signal streams without foreknowledge of the encoded media stream. Moreover, the data format orientation is used to partially scramble data in transit to prevent unauthorized descrambled access by decoders that lack authorized keys. A distributed key can be used to unscramble the scrambled content because a decoder would understand how to process the key. Similar to on-the-fly decryption operations, the benefits inherent in this embodiment include the fact that the combination of watermarked content security, which is key-based, and the descrambling of the data, can be performed by the same key which can be a plurality of mask sets. The mask sets may include primary, convolution and message delimiter masks with file format data included. r

The creation of an optimized "envelope" for insertion of watermarks provides the basis of much watermark security, but is also a complementary goal of the present invention. The predetermined or random key that is generated is not only an essential map to access the hidden information signal, but is also the descrambler of the previously scrambled signal's format for playback or viewing.

In a system requiring keys for watermarking content and validating the distribution of the content, different keys may be used to encode different information while secure one way hash functions or one-time pads may be incorporated to secure the embedded signal. The same keys can be used to later validate the embedded digital signature, or even fully decode the digital watermark if desired. Publishers can easily stipulate that content not only be digitally watermarked but that distributors must check the validity of the watermarks by performing digital signature-checks with keys that lack any other functionality. The system can extend to simple authentication of text in other embodiments.

Before such a market is economically feasible, there are other methods for deploying key-based watermarking coupled with transfer functions to partially scramble the content to be distributed without performing full public key encryption, i.e., a key pair is not necessarily generated, simply, a predetermined key's function is created to re-map the data of the content file in a lossless process.

Moreover, the scrambling performed by the present invention may be more dependent on the file in question. Dissimilarly, encryption is not specific to any particular media but is performed on data. The file format remains unchanged, rendering the file useable by any conventional viewer/player, but the signal quality can be intentionally degraded in the absence of the proper player and key. Public-key encryption seeks to completely obscure the sensitive "plaintext" to prevent comparisons with the "ciphertext" to determine a user's private keys. Centralized encryption only differs in the utilization of a single key for both encryption and decryption making the key even more highly vulnerable to attacks to defeat the encryption process. With the present invention, a highly sought after photograph may be hazy to the viewer using any number of commonly available, nonproprietary software or hardware, without the authorized key. Similarly, a commercially valuable song may sound poor.

The benefit of some form of cryptography is not lost in the present invention. In fact, some piracy can be deterred when the target signal may be known but is clearly being protected through scrambling. What is not anticipated by known techniques, is an ala carte method to change various aspects of file formatting to enable various "scrambled states" for content to be subsequently distributed. An image may lack all red pixels or may not have any of the most significant bits activated. An audio sample can similarly be scrambled to render it less-than-commercially viable.

The present invention also provides improvements over known network-based methods, such as those used for the streaming of media data over the Internet. By manipulating file formats, the broadcast media, which has been altered to "fit" within electronic distribution parameters, such as bandwidth availability and error correction considerations; can be more effectively utilized to restrict the subsequent use of the content while in transit as well as real-time viewing or playing.

The mask set providing the transfer function can be read on a per-use basis by issuing an authorized or authenticating "key" for descrambling the signal that is apparent to a viewer or a player or possessor of the authenticating key. The mask set can be read on a per-computer basis by issuing the authorized key that is

more generalized for the computer that receives the broadcast signals. Metering and subscription models become viable advantages over known digital watermark systems which assist in designating the ownership of a copy of digitized media content, but do not prevent or restrict the copying or manipulation of the sampled signal in question. For broadcast or streamed media, this is especially the case. Message authentication is also possible, though not guaranteeing the same security as an encrypted file as with general crypto systems.

The present invention thus benefits from the proprietary player model without relying on proprietary players. No new players will be necessary and existing multimedia file formats can be altered to exact a measure of security which is further increased when coupled with digital watermarks. As with most consumer markets for media content, predominant file formats exist, de facto, and corresponding formats for computers likewise exist. For a commercial compact disc quality audio recording, or 16 bit 44.1 kHz, corresponding file formats include: Audio Interchange File Format (AIFF), Microsoft WAV, Sound Designer II, Sun's .au, Apple's Quicktime, etc. For still image media, formats are similarly abundant: TIFF, PICT, JPEG, GIF, etc. Requiring the use of additional proprietary players, and their complementary file formats, for limited benefits in security is wasteful. Moreover, almost all computers today are multimedia-capable, and this is increasingly so with the popularity of Intel's MMX chip architecture and the PowerPC line of microchips. Because file formatting is fundamental in the playback of the underlying data, the predetermined key can act both as a map, for information to be encoded as watermark data regarding ownership, and a descrambler of the file that has been distributed. Limitations will only exist in how large the key must be retrofitted for a given application, but any manipulation of file format information is not likely to exceed the size of data required versus that for an entire proprietary player.

As with previous disclosures by the inventor on digital watermarking techniques, the present invention may be implemented with a variety of cryptographic protocols to increase both confidence and security in the underlying system. A predetermined key is described as a set of masks. These masks may include primary, convolution and message delimiter mask. In previous disclosures,

the functionality of these masks is defined solely for mapping. The present invention includes a mask set which is also controlled by the distributing party of a copy of a given media signal. This mask set is a transfer function which is limited only by the parameters of the file format in question. To increase the uniqueness or security of each key used to scramble a given media file copy, a secure one way hash function can be used subsequent to transfer properties that are initiated to prevent the forging of a particular key. Public and private keys may be used as key pairs to further increase the unlikeliness that a key may be compromised.

These same cryptographic protocols can be combined with the embodiments of the present invention in administering streamed content that requires authorized keys to correctly display or play the streamed content in an unscrambled manner. As with digital watermarking, symmetric or asymmetric public key pairs may be used in a variety of implementations. Additionally, the need for certification authorities to maintain authentic key-pairs becomes a consideration for greater security beyond symmetric key implementations. The cryptographic protocols makes possible, as well, a message of text to be authenticated by a message authenticating function in a general computing device that is able to ensure secure message exchanges between authorizing parties.

Although various embodiments are specifically illustrated and described herein, it will be appreciated that modifications and variations of the present invention are covered by the above teachings and within the purview of the appended claims without departing from the spirit and intended scope of the invention.

What is claimed is:

1. A method for copy protection of digital information, the digital information including a digital sample and format information, comprising the steps of:

identifying a portion of the format information to be encoded;

generating encoded format information from the identified portion of the format information; and

generating encoded digital information, including the digital sample and the encoded format information.

2. The method of claim 1, further comprising the step of requiring a predetermined key to decode the encoded format information.

3. The method of claim 2, wherein the digital sample and format information are configured to be used with a digital player, and wherein information output from the digital player will have a degraded quality unless the encoded format information is decoded with the predetermined key.

4. The method of claim 3, wherein the information output from the digital player represents a still image, audio or video.

5. The method of claim 3, wherein the information output represents text data to be authenticated.

6. A method for protecting a digital signal, the digital signal including digital samples in a file format having an inherent granularity, comprising the step of:

creating a predetermined key comprised of a transfer function-based mask set to manipulate data at the inherent granularity of the file format of the underlying digitized samples.

7. The method of claim 6, wherein the digital signal represents a continuous analog waveform.

8. The method of claim 6, wherein the predetermined key comprises a plurality of mask sets.

9. The method of claim 6, wherein the digital signal is a message to be authenticated.

10. The method of claim 6, wherein the mask set is ciphered by a key pair comprising a public key and a private key.

11. The method of claim 6, further comprising the step of:

using a digital watermarking technique to encode information that identifies ownership, use, or other information about the digital signal, into the digital signal.

12. The method of claim 6, wherein the digital signal represents a still image, audio or video.

13. The method of claim 6, further comprising the steps of:

selecting the mask set, including one or more masks having random or pseudo-random series of bits; and

validating the mask set at the start of the transfer function-based mask set.

14. The method of claim 13, wherein said step of validating comprises the step of:

comparing a hash value computed at the start of the transfer function-based mask set with a determined transfer function of the hash value.

15. The method of claim 6, further comprising the steps of:

selecting the mask set, including one or more masks having random or pseudo-random series of bits; and

authenticating the mask set by comparing a hash value computed at the start of the transfer function-based mask set with a determined transfer function of the hash value.

16. The method of claim 13, wherein said step of validating comprises the step of:

comparing a digital signature at the start of the transfer function-based mask set with a determined transfer function of the digital signature.

17. The method of claim 6, further comprising the steps of:

selecting the mask set, including one or more masks having random or pseudo-random series of bits; and

authenticating the mask set by comparing a digital signature at the start of the transfer function-based mask set with a determined transfer function of the digital signature.

18. The method of claim 13, further comprising the step of:

using a digital watermarking technique to embed information that identifies ownership, use, or other information about the digital signal, into the digital signal; and

wherein said step of validating is dependent on validation of the embedded information.

19. The method of claim 6, further comprising the step of:

computing a secure one way hash function of carrier signal data in the digital signal, wherein the hash function is insensitive to changes introduced into the carrier signal for the purpose of carrying the transfer function-based mask set.

20. A method for protecting a digital signal, the digital signal including digital samples in a file format having an inherent granularity, comprising the steps of:

creating a predetermined key comprised of a transfer function-based mask set that can manipulate data at the inherent granularity of the file format of the underlying digitized samples;

authenticating the predetermined key containing the correct transfer function-based mask set during playback of the data; and

metering the playback of the data to monitor content.

21. The method of claim 20, wherein the predetermined key is authenticated to authenticate message information.

22. A method to prepare for the scrambling of a sample stream of data, comprising the steps of:

generating a plurality of mask sets to be used for encoding, including a random primary mask, a random convolution mask and a random start of message delimiter;

obtaining a transfer function to be implemented;

generating a message bit stream to be encoded;

loading the message bit stream, a stega-cipher map truth table, the primary mask, the convolution mask and the start of message delimiter into memory;

initializing the state of a primary mask index, a convolution mask index, and a message bit index; and

setting a message size equal to the total number of bits in the message bit stream.

23. A method to prepare for the encoding of stega-cipher information into a sample stream of data, comprising the steps of:

generating a mask set to be used for encoding, the set including a random primary mask, a random convolution mask, and a random start of message delimiter;

obtaining a message to be encoded;

compressing and encrypting the message if desired;

generating a message bit stream to be encoded;

loading the message bit stream, a stega-cipher map truth table, the primary mask, the convolution mask and the start of message delimiter into memory;

initializing the state of a primary mask index, a convolution mask index, and a message bit index; and

setting the message size equal to the total number of bits in the message bit stream.

24. The method of claim 23 wherein the sample stream of data has a plurality of windows, further comprising the steps of:

calculating over which windows in the sample stream the message will be encoded;

computing a secure one way hash function of the information in the calculated windows, the hash function generating hash values insensitive to changes in the samples induced by a stega-cipher; and

encoding the computed hash values in an encoded stream of data.

25. The method of claim 13, wherein said step of selecting comprises the steps of:

collecting a series of random bits derived from keyboard latency intervals in random typing;

processing the initial series of random bits through an MD5 algorithm;

using the results of the MD5 processing to seed a triple-DES encryption loop;

cycling through the triple-DES encryption loop, extracting the least significant bit of each result after each cycle; and

concatenating the triple-DES output bits into the random series of bits.

26. A method for copy protection of digital information, the digital information including a digital sample and format information, comprising the steps of:

a identifying a portion of the digital sample to be encoded;

generating an encoded digital sample from the identified portion of the digital sample; and

generating encoded digital information, including the encoded digital sample and the format information.

27. The method of claim 26, further comprising the step of requiring a predetermined key to decode the encoded digital sample.

28. The method of claim 27, wherein the digital sample and format information are configured to be used with a digital player, and wherein information output from the digital player will have a degraded quality unless the encoded digital sample is decoded with the predetermined key.

29. The method of claim 27, wherein information output will have non authentic message data unless the encode digital sample is decoded with the predetermined key.

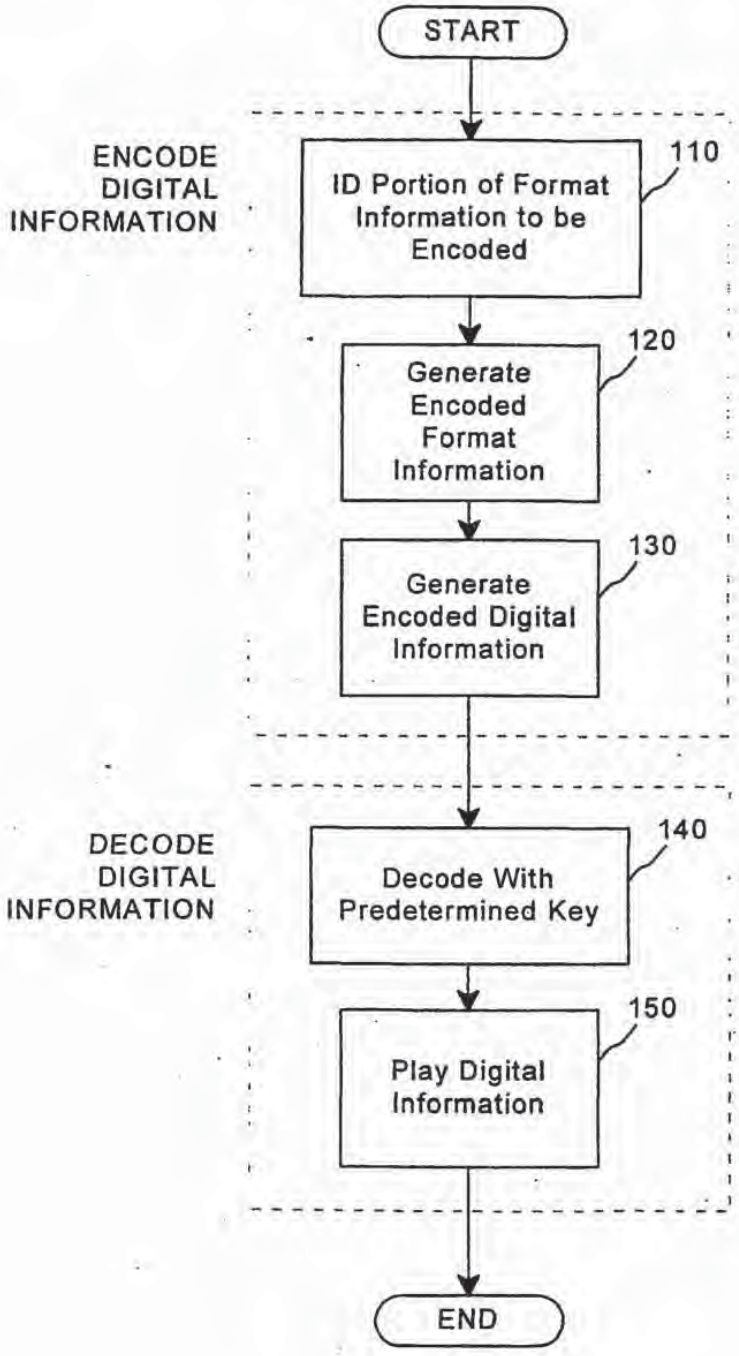

START
ENCODE DIGITAL INFORMATION
ID Portion of Format Information to be Encoded
110
Generate Encoded Format Information
120
Generate Encoded Digital Information
130
DECODE DIGITAL INFORMATION
Decode With Predetermined Key
140
Play Digital Information
150
END

FIG. 1

INTERNATIONAL SEARCH REPORT

ional Application No
PCT/US 00/18411

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G11B20/00 G06F1/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G11B G06F H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	NL 1 005 523 C (EINDHOVEN TECH HOCHSCHULE) 15 September 1998 (1998-09-15) abstract; figure 4 page 1, line 35 -page 3, line 9 page 9, line 21 -page 10, line 5	1,2, 26-29
X	WO 97 44736 A (APPLE COMPUTER) 27 November 1997 (1997-11-27) abstract; figure 4 page 2, line 35 -page 3, line 27 page 9, line 10 -page 11, line 28	1,2
Y		3,4

-/--

[X] Further documents are listed in the continuation of box C. [X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 July 2001	30/07/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Sigalo, A

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Ional Application No
PCT/US 00/18411

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 687 236 A (MOSKOWITZ SCOTT A ET AL) 11 November 1997 (1997-11-11) cited in the application column 5, line 1 -column 6, line 37 column 7, line 54 -column 10, line 11 column 11, line 31 -column 12, line 10 column 15, line 42 -column 16, line 32	6-12, 19-21
A		22,23
A	US 5 974 141 A (SAITO MAKOTO) 26 October 1999 (1999-10-26) abstract; figures 4A-4G column 8, line 24 - line 67	5,26
X	WO 99 52271 A (MOSKOWITZ SCOTT A) 14 October 1999 (1999-10-14) abstract page 11, line 15 -page 13, line 13	6,7,10
Y	EP 0 649 261 A (CANON KK) 19 April 1995 (1995-04-19) page 3, line 53 -page 4, line 5 page 7, line 18 - line 23	3,4
A	WO 99 63443 A (DATAMARK TECHNOLOGIES PTE LTD; HO ANTHONY TUNG SHUEN (SG); TAM SIU) 9 December 1999 (1999-12-09) page 2, line 10 -page 5, line 16	6-8,11, 12

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

ernational application No.
PCT/US 00/18411

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☒ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-5,26-29

   Protecting the distribution of digital data to be used with a digital player characterized by encrypting format information and allowing low quality play back in case of lack of decrypting key.

2. Claims: 6-25

   Digital signature encrypting technique combining transfer functions with predetermined key creation.

INTERNATIONAL SEARCH REPORT

Information on patent family members

ional Application No
PCT/US 00/18411

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
NL 1005523	C	15-09-1998	NONE		
WO 9744736	A	27-11-1997	AU	3206397 A	09-12-1997
US 5687236	A	11-11-1997	US	5613004 A	18-03-1997
			EP	0872073 A	21-10-1998
			WO	9642151 A	27-12-1996
US 5974141	A	26-10-1999	US	6076077 A	13-06-2000
			US	6002772 A	14-12-1999
			US	6097818 A	01-08-2000
WO 9952271	A	14-10-1999	US	6205249 B	20-03-2001
			EP	1068720 A	17-01-2001
EP 0649261	A	19-04-1995	JP	7115638 A	02-05-1995
			US	5933499 A	03-08-1999
WO 9963443	A	09-12-1999	AU	7683398 A	20-12-1999
			EP	1103026 A	30-05-2001

Form PCT/ISA/210 (patent family annex) (July 1992)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 March 2001 (15.03.2001)

PCT

(10) International Publication Number
WO 01/18628 A2

(51) International Patent Classification[7]: G06F

(21) International Application Number: PCT/US00/21189

(22) International Filing Date: 4 August 2000 (04.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/147,134 4 August 1999 (04.08.1999) US
60/213,489 23 June 2000 (23.06.2000) US

(71) Applicant *(for all designated States except US)*: **BLUE SPIKE, INC.** [US/US]; 16711 Collins Avenue #2505, Miami, FL 33160 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **MOSKOWITZ, Scott, A.** [US/US]; 16711 Collins Avenue #2505, Miami, FL 33160 (US). **BERRY, Michael** [US/US]; 12401 Princess Jeanne, Alburquerque, NM 87112 (US).

(74) Agents: **CHAPMAN, Floyd, B. et al.**; Baker Botts, LLP, The Warner, 1299 Pennsylvania Avenue, N.W., Washington, DC 20004 (US).

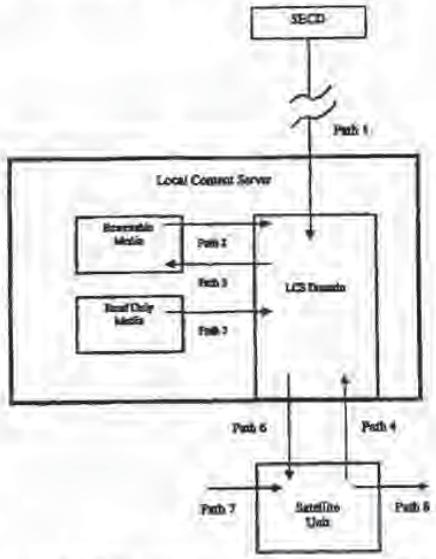
(81) Designated States *(national)*: JP, US.

(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published:
— *Without international search report and to be republished upon receipt of that report.*

[Continued on next page]

(54) Title: A SECURE PERSONAL CONTENT SERVER



(57) Abstract: A local content server system (LCS) for creating a secure environment for digital content is disclosed, which system comprises: a communications port in communication for connecting the LCS via a network to at least one Secure Electronic Content Distributor (SECD), which SECD is capable of storing a plurality of data sets, is capable of receiving a request to transfer at least one content data set, and is capable of transmitting the at least one content data set in a secured transmission; a rewritable storage medium whereby content received from outside the LCS may be stored and retrieved; a domain processor that imposes rules and procedures for content being transferred between the LCS and devices outside the LCS, and a programmable address module which can be programmed with an identification code uniquely associated with the LCS. The LCS is provided with rules and procedures for accepting and transmitting content data. Optionally, the system may further comprise: an interface to permit the LCS to communicate with one or more Satellite Units (SU) which may be connected

[Continued on next page]

WO 01/18628 A2

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

to the system through the interface, which SUs are capable of receiving and transmitting digital content; at least one SU; and/or at least one SECD. The SECD may have a storage device for storing a plurality of data sets, as well as a transaction processor for validating the request to purchase and for processing payment for a request to retrieve one of the data sets. The SECD typically includes a security module for encrypting or otherwise securitizing data which the SECD may transmit. A method for creating a secure environment for digital content for a consumer is also disclosed. As part of the method, a LCS requests and receives a digital data set that may be encrypted or scrambled. The digital data set may be embedded with at least one robust open watermark, which permits the content to be authenticated. The digital data set is preferably embedded with additional watermarks which are generated using information about the LCS requesting the copy and/or the SECD which provides the copy. Once received by the LCS, the LCS exercises control over the content and only releases the data to authorized users. Generally, the data is not released until the LCS embeds at least one additional watermark based upon protected information associated with the LCS and/or information associated with the user.

A SECURE PERSONAL CONTENT SERVER

Field of Invention

The present invention relates to the secure distribution of digitized value-added information, or media content, while preserving the ability of publishers to make available unsecured versions of the same value-added information, or media content, without adverse effect to the systems security.

Authentication, verification and authorization are all handled with a combination of cryptographic and steganographic protocols to achieve efficient, trusted, secure exchange of digital information.

Cross-Reference To Related Application

This application is based on and claims the benefit of pending U.S. Patent Application Serial No. 60/147,134, filed 08/04/99, entitled, "A Secure Personal Content Server" and pending U.S. Patent Application Serial No. 60/213,489, filed 06/23/2000, entitled "A Secure Personal Content Server."

This application also incorporates by reference the following applications: pending U.S. Patent Application Serial No. 08/999,766, filed 7/23/97, entitled "Steganographic Method and Device"; pending U.S. Patent Application Serial No. 08/772,222, filed 12/20/96, entitled "Z-Transform Implementation of Digital Watermarks"; pending U.S. Patent Application Serial No. 09/456,319, filed 12/08/99, entitled "Transform Implementation of Digital Watermarks"; pending U.S. Patent Application Serial No. 08/674,726, filed 7/2/96, entitled "Exchange Mechanisms for Digital Information Packages with Bandwidth Securitization, Multichannel Digital Watermarks, and Key Management"; pending U.S. Patent Application Serial No. 09/545,589, filed 04/07/2000, entitled "Method and System for Digital Watermarking"; pending U.S. Patent Application Serial No. 09/046,627, filed 3/24/98, entitled "Method for Combining Transfer Function with Predetermined Key Creation"; pending U.S. Patent Application Serial No. 09/053,628, filed 04/02/98, entitled "Multiple Transform Utilization and Application for Secure Digital Watermarking"; pending U.S. Patent Application Serial No. 09/281,279, filed 3/30/99, entitled "Optimization Methods for the Insertion, Protection, and Detection..."; U.S. Patent Application Serial No.09/594,719, filed June 16, 2000, entitled "Utilizing Data Reduction in Steganographic and

Cryptographic Systems" (which is a continuation-in-part of PCT application No. PCT/US00/06522, filed 14 March 2000, which PCT application claimed priority to U.S. Provisional Application No. 60/125,990, filed 24 March 1999); and pending U.S. Application No 60/169,274, filed 12/7/99, entitled "Systems, Methods And Devices For Trusted Transactions." All of the patent applications previously identified in this paragraph are hereby incorporated by reference, in their entireties.

Background of the Invention

The music industry is at a critical inflection point. Digital technology enables anyone to make perfect replica copies of musical recordings from the comfort of their home, or as in some circumstances, in an offshore factory. Internet technology enables anyone to distribute these copies to their friends, or the entire world. Indeed, virtually any popular recording is already likely available in the MP3 format, for free if you know where to look.

How the industry will respond to these challenges and protect the rights and livelihoods of copyright owners and managers and has been a matter of increasing discussion, both in private industry forums and the public media. Security disasters like the cracking of DVD-Video's CSS security system have increased doubt about the potential for effective robust security implementations. Meanwhile, the success of non-secure initiatives such as portable MP3 players lead many to believe that these decisions may have already been made.

Music consumers have grown accustomed to copying their music for their own personal use. This fact of life was written into law in the United States via the Audio Home Recording Act of 1992. Millions of consumers have CD players and purchase music in the Compact Disc format. It is expected to take years for a format transition away from Red Book CD Audio to reach significant market penetration.

Hence, a need exists for a new and improved system for protecting digital content against unauthorized copying and distribution.

Summary of the Invention

A local content server system (LCS) for creating a secure environment for digital content is disclosed, which system comprises: a communications port in communication for connecting the LCS via a network to at least one Secure Electronic Content Distributor (SECD), which SECD is capable of storing a

plurality of data sets, is capable of receiving a request to transfer at least one content data set, and is capable of transmitting the at least one content data set in a secured transmission; a rewritable storage medium whereby content received from outside the LCS may be stored and retrieved; a domain processor that imposes rules and procedures for content being transferred between the LCS and devices outside the LCS; and a programmable address module which can be programmed with an identification code uniquely associated with the LCS. The LCS is provided with rules and procedures for accepting and transmitting content data. Optionally, the system may further comprise: an interface to permit the LCS to communicate with one or more Satellite Units (SU) which may be connected to the system through the interface, which SUs are capable of receiving and transmitting digital content; at least one SU; and/or at least one SECD. The SECD may have a storage device for storing a plurality of data sets, as well as a transaction processor for validating the request to purchase and for processing payment for a request to retrieve one of the data sets. The SECD typically includes a security module for encrypting or otherwise securitizing data which the SECD may transmit.

A method for creating a secure environment for digital content for a consumer is also disclosed. As part of the method, a LCS requests and receives a digital data set that may be encrypted or scrambled. The digital data set may be embedded with at least one robust open watermark, which permits the content to be authenticated. The digital data set is preferably be embedded with additional watermarks which are generated using information about the LCS requesting the copy and/or the SECD which provides the copy. Once received by the LCS, the LCS exercises control over the content and only releases the data to authorized users. Generally, the data is not released until the LCS embeds at least one additional watermark based upon protected information associated with the LCS and/or information associated with the user.

Another embodiment of the method of the present invention comprises: connecting a Satellite Unit to an local content server (LCS), sending a message indicating that the SU is requesting a copy of a content data set that is stored on the LCS, said message including information about the identity of the SU; analyzing the message to confirm that the SU is authorized to use the LCS; retrieving a copy of the

requested content data set; assessing whether a secured connection exists between the LCS and the SU; if a secured connection exists, embedding a watermark into the copy of the requested content data set, said watermark being created based upon information transmitted by the SU and information about the LCS; and delivering the content data set to the SU for its use.

The SU may also request information that is located not on the LCS, but on an SECD, in which case, the LCS will request and obtain a copy from the SECD, provided the requesting SU is authorized to access the information.

Digital technology offers economies of scale to value-added data not possible with physical or tangible media distribution. The ability to digitize information both reduces the cost of copying and enables perfect copies. This is an advantage and a disadvantage to commercial publishers who must weigh the cost reduction against the real threat of unauthorized duplication of their value-added data content. Because cost reduction is an important business consideration, securing payment and authenticating individual copies of digital information (such as media content) presents unique opportunities to information service and media content providers. The present invention seeks to leverage the benefits of digital distribution to consumers and publishers alike, while ensuring the development and persistence of trust between all parties, as well as with any third parties involved, directly or indirectly, in a given transaction.

In another approach that is related to this goal, there are instances where transactions must be allowed to happen after perceptually-based digital information can be authenticated. (Perceptually based information is information whose value is in large part, based upon its ability to be perceived by a human, and includes for example, acoustic, psychoacoustic, visual and psychovisual information.) The process of authenticating before distributing will become increasingly important for areas where the distributed material is related to a trust-requiring transaction event. A number of examples exist. These include virtual retailers (for example, an on-line music store selling CDs and electronic versions of songs); service providers (for example, an on-line bank or broker who performs transactions on behalf of a consumer); and transaction providers (for example, wholesalers or auction houses). These parties have different authentication interests and requirements. By using the

teachings of this application, these interests and requirements may be separated and then independently quantified by market participants in shorter periods of time.

All parties in a transaction must authenticate information that is perceptually observable before trust between the parties can be established. In today's world, information (including perceptually rich information) is typically digitized, and as a result, can easily be copied and redistributed, negatively impacting buyers, sellers and other market participants. Unauthorized redistribution confuses authenticity, non-repudiation, limit of ability and other important "transaction events." In a networked environment, transactions and interactions occur over a transmission line or a network, with buyer and seller at different points on the line or network. While such electronic transactions have the potential to add value to the underlying information being bought and sold (and the potential to reduce the cost of the transaction), instantaneous piracy can significantly reduce the value of the underlying data, if not wholly destroy it. Even the threat of piracy tends to undermine the value of the data that might otherwise exist for such an electronic transaction.

Related situations range from the ability to provably establish the "existence" of a virtual financial institution to determining the reliability of an "electronic stamp." The present invention seeks to improve on the prior art by describing optimal combinations of cryptographic and steganographic protocols for "trusted" verification, confidence and non-repudiation of digitized representations of perceptually rich information of the actual seller, vendor or other associated institutions which may not be commercial in nature (confidence building with logo's such as the SEC, FDIC, Federal Reserve, FBI, etc. apply). To the extent that an entity plays a role in purchase decisions made by a consumer of goods and services relating to data, the present invention has a wide range of beneficial applications. One is enabling independent trust based on real world representations that are not physically available to a consumer or user. A second is the ability to match informational needs between buyers and sellers that may not be universally appealing or cost effective in given market situations. These include auction models based on recognition of the interests or demand of consumers and market participants—which make trading profitable by focusing specialized buyers and

sellers. Another use for the information matching is to establish limits on the liability of such institutions and profit-seeking entities, such as insurance providers or credit companies. These vendors lack appropriate tools for determining intangible asset risk or even the value of the information being exchanged. By encouraging separate and distinct "trust" arrangements over an electronic network, profitable market-based relationships can result.

The present invention can make possible efficient and openly accessible markets for tradable information. Existing transaction security (including on-line credit cards, electronic cash or its equivalents, electronic wallets, electronic tokens, etc.) which primarily use cryptographic techniques to secure a transmission channel--but are not directly associated or dependent on the information being sold--fails to meet this valuable need. The present invention proposes a departure from the prior art by separating transactions from authentication in the sale of digitized data. Such data may include videos, songs, images, electronic stamps, electronic trademarks, and electronic logos used to ensure membership in some institutional body whose purpose is to assist in a dispute, limit liability and provide indirect guidance to consumers and market participants, alike.

With an increasingly anonymous marketplace, the present invention offers invaluable embodiments to accomplish "trusted" transactions in a more flexible, transparent manner while enabling market participants to negotiate terms and conditions. Negotiation may be driven by predetermined usage rules or parameters, especially as the information economy offers potentially many competitive marketplaces in which to transact, trade or exchange among businesses and consumers. As information grows exponentially, flexibility becomes an advantage to market participants, in that they need to screen, filter and verify information before making a transaction decision. Moreover, the accuracy and speed at which decisions can be made reliably enables confidence to grow with an aggregate of "trusted transactions". "Trusted transactions" beget further "trusted transactions" through experience. The present invention also provides for improvements over the prior art in the ability to utilize different independently important "modules" to enable a "trusted transaction" using competitive cryptographic and steganographic elements, as well as being able to support a wide variety of perceptually-based

media and information formats. The envisioned system is not bound by a proprietary means of creating recognition for a good or service, such as that embodied in existing closed system. Instead, the flexibility of the present invention will enable a greater and more diverse information marketplace.

The present invention is not a "trusted system", *per se*, but "trusted transactions" are enabled, since the same value-added information that is sought may still be in the clear, not in a protected storage area or closed, rule-based "inaccessible virtual environment".

A related additional set of embodiments regards the further separation of the transaction and the consumer's identification versus the identification of the transaction only. This is accomplished through separated "trusted transactions" bound by authentication, verification and authorization in a transparent manner. With these embodiments, consumer and vendor privacy could be incorporated. More sophisticated relationships are anticipated between parties, who can mix information about their physical goods and services with a transparent means for consumers, who may not be known to the seller, who choose not to confide in an inherently closed "trusted system" or provide additional personal information or purchasing information (in the form of a credit card or other electronic payment system), in advance of an actual purchase decision or ability to observe (audibly or visibly) the content in the clear. This dynamic is inconsistent with the prior art's emphasis on access control, not transparent access to value-added information (in the form or goods or services), that can be transacted on an electronic or otherwise anonymous exchange.

These embodiments may include decisions about availability of a particular good or service through electronic means, such as the Internet, or means that can be modularized to conduct a transaction based on interconnection of various users (such as WebTV, a Nintendo or Sony game console with network abilities, cellular phone, PalmPilot, etc.). These embodiments may additionally be implemented in traditional auction types (including Dutch auctions). Consumers may view their anonymous marketplace transactions very differently because of a lack of physical human interactions, but the present invention can enable realistic transactions to occur by maintaining open access and offering strict authentication and verification of the

information being traded. This has the effect of allowing legacy relationships, legacy information, and legacy business models to be offered in a manner which more closely reflects many observable transactions in the physical world. The tremendous benefits to sellers and consumers is obvious; existing transactions need not reduce their expectations of security. As well, the ability to isolate and quantify aspects of a transaction by module potentially allows for better price determinations of intangible asset insurance, transaction costs, advertising costs, liability, etc. which have physical world precedent.

It is contemplated that the publisher and/or owner of the copyrights will want to dictate restrictions on the ability of the purchaser to use the data being sold. Such restrictions can be implemented through the present invention, which presents a significant advantage over the prior art (which attempts to effect security through access control and attempted tight reigns over distribution). See US Pat. No. 5,428,606 for a discussion on democratizing digital information exchange between publishers and subscribers of said information.

A goal for providers of value-added content is to maximize profits for the sale of their content. Marketing and promotion of the informational content cannot be eliminated, considering the ever increasing amount of information vying for consumers and other market participant's attention. Nonetheless, in a market where the goods are speculatively valued, marketing budgets are inherently constrained, as you are trying to create demand for a product with little inherent value. Where such markets have participants, both buyers and sellers and their respective agents, with access to the same information in real time, market mechanisms efficiently price the market goods or services. These markets are characterized by "price commoditization" so buyers and sellers are limited to differentiating their offerings by selection and service. If the markets are about information itself, it has proven more difficult to accurately forecast the target price where sellers can maximize their profits. Quality and quantity provide different evaluation criteria of selection and service relating to the information being traded. The present invention regards a particular set of implementations of value-added content security in markets which may include unsecured and secure versions of the same value-added data (such as

songs, video, research, pictures, electronic logos, electronic trademarks, value-added information, etc.).

Transactions for value-added information can occur without any physical location. So, there is a need for a secure personal content server for which the value added information can be offered for transactions in a manner similar to real world transactions. One feature is to offer seemingly similar value added information in differing quality settings. These settings have logical relationships with fidelity and discreteness and are determined by market participants. Another issue is that because purchasers may be anonymous to sellers, it is more important to have a particular value-added information object available so that market participants can fulfill their role are consumers.

One fundamental weakness of current information markets is the lack of mechanisms to ensure that buyers and sellers can reach pricing equilibrium. This deficit is related to the "speculative" , "fashion", and "vanity" aspects of perceptual content (such as music, video, and art or some future recognition to purchasers). For other goods and services being marketed to an anonymous marketplace, market participants may never see (and indeed, may choose to never see, an actual location where the transaction may physically occur. A physical location may simply not exist. There are a number of such virtual operations in business today, which would benefit from the improvements offered under the present system.

The present invention also seeks to provide improvements to the art in enabling a realistic model for building trust between parties (or their agents) not in a "system", per se. Because prior art systems lack any inherent ability to allow for information to flow freely to enable buyers and sellers to react to changing market conditions. The present invention can co-exist with these "trusted systems" to the extent that all market participants in a given industry have relatively similar information with which to price value-added data. The improvement over such systems, however, addresses a core features in most data-added value markets: predictions, forecasts, and speculation over the value of information is largely an unsuccessful activity for buyers and sellers alike. The additional improvement is the ability to maintain security even with unsecured or legacy versions of value-added information available to those who seek choices that fit less quantitative criteria—

"aesthetic quality" of the information versus "commercial price". Purchase or transaction decisions can be made first by authenticating an electronic version of a song, image, video, trademark, stamp, currency, etc.

Additional anticipated improvements include the ability to support varying pricing models such as auctions that are difficult or impossible to accomplish under existing prior art that leaves all access and pricing control with the seller alone, and the separation of the transaction from the exchange of the value-added information, which gives more control to buyers over their identities and purchasing habits, (both sensitive and separately distinct forms of "unrelated" value-added information). Essentially, no system known in the art allows for realistic protocols to establish trust between buyers and sellers in a manner more closely reflecting actual purchasing behavior of consumers and changing selling behavior of sellers. The goal in such transactions is the creation of trust between parties as well as "trusted relationships" with those parties. The present invention is an example of one such system for media content where the "aesthetic" or "gestalt" of the underlying content and its characteristics is a component of buying habits. Without an ability to open distribution systems to varying buyers and sellers, media content may be priced at less than maximum economic value and buyers may be deprived of a competitive, vigorous marketplace for exciting media content from many different creative participants.

To the extent that recognition plays such a key role in an information economy, value-added data should be as accessible as possible to the highest number of market participants in the interests of furthering creativity and building a competitive marketplace for related goods and services. This is to the benefit of both buyers and sellers as well as the other participants in such an economic ecosystem. The Internet and other transmission-based transactions with unknown parties presents a number of challenges to information vendors who wish to develop customer relations, trust and profitable sales. The information economy is largely an anonymous marketplace, thus, making it much more difficult to identify consumers and sellers. The present invention provides remedies to help overcome these weaknesses.

The present invention is concerned with methods and systems which enable secure, paid exchange of value-added information, while separating transaction protocols. The present invention improves on existing means for distribution control by relying on authentication, verification and authorization that may be flexibly determined by both buyers and sellers. These determinations may not need to be predetermined, although pricing matrix and variable access to the information opens additional advantages over the prior art. The present invention offers methods and protocols for ensuring value-added information distribution can be used to facilitate trust in a large or relatively anonymous marketplace (such as the Internet's World Wide Web).

We now define components of the preferred embodiments for methods, systems, and devices.

Definitions:

Local Content Server (LCS): A device or software application which can securely store a collection of value-added digital content. The LCS has a unique ID.

Secure Electronic Content Distributor (SECD): An entity, device or software application which can validate a transaction with a LCS, process a payment, and deliver digital content securely to a LCS. In cryptographic terms, the SECD acts as a "certification authority" or its equivalent. SECDs may have differing arrangements with consumers and providers of value-added information. (The term "content" is used to refer generally to digital data, and may comprise video, audio, or any other data that is stored in a digital format).

Satellite Unit (SU): A portable medium or device which can accept secure digital content from a LCS through a physical, local connection and which can either play or make playable the digital content. The SU may have other functionality as it relates to manipulating the content, such as recording. The SU has a unique ID. An SU may be a CD player, a video camera, a backup drive, or other electronic device which has a storage unit for digital data.

LCS Domain: A secure medium or area where digital content can be stored, with an accompanying rule system for transfer of digital content in and out of the LCS Domain. The domain may be a single device or multiple devices—all of which have some common ownership or control. Preferably, a LCS domain is linked to a

single purchasing account. Inside the domain, one can enjoy music or other digital data without substantial limitations—as typically a license extends to all personal use.

SecureChannel™: A secure channel to pass individualized content to differentiate authentic content from legacy or unauthorized, pirated content. For example, the Secure Channel may be used as an auxiliary channel through which members of the production and distribution chain may communicate directly with individual consumers. Preferably, the Secure Channel is never exposed and can only be accessed through legitimate methods. SecureChannel may carry a value-adding component (VAC). The ability to provide consumers with value adding features will serve to give consumers an incentive to purchase new, secure hardware and software that can provide the additional enhanced services. The SecureChannel may also include protected associated data—data which is associated with a user and/or a particular set of content.

Standard Quality: A transfer path into the LCS Domain which maintains the digital content at a predetermined reference level or degrades the content if it is at a higher quality level. In an audio implementation, this might be defined as Red Book CD Quality (44100 Hz., 16 bits, 2 channels). This transfer path can alternately be defined in terms of a subset of VAC's or a quality level associated with particular VAC's. If a VAC is not in the subset, it is not passed. If a VAC is above the defined quality level, it is degraded.

Low Quality: A transfer path into the LCS Domain which degrades the digital content to a sub-reference level. In an audio implementation, this might be defined as below CD Quality (for instance, 32000 Hz., 16 bits, 2 channels). This transfer path can alternately be defined in terms of an absence of VAC's or a degraded quality level associated with particular VAC's.

High Quality: A transfer path into the LCS Domain which allows digital content of any quality level to pass unaltered. This transfer path can alternately be defined in terms of a complete set of VAC's or the highest quality level available associated with particular VAC's.

Rewritable Media: An mass storage device which can be rewritten (e.g. hard drive, CD-RW, Zip cartridge, M-O drive, etc...).

Read-Only Media: A mass storage device which can only be written once (e.g. CD-ROM, CD-R, DVD, DVD-R, etc...). Note: pre-recorded music, video, software, or images, etc. are all "read only" media.

Unique ID: A Unique ID is created for a particular transaction and is unique to that transaction (roughly analogous to a human fingerprint). One way to generate a Unique ID is with a one-way hash function. Another way is by incorporating the hash result with a message into a signing algorithm will create a signature scheme. For example, the hash result may be concatenated to the digitized, value added information which is the subject of a transaction. Additional uniqueness may be observed in a hardware device so as to differentiate that device, which may be used in a plurality of transactions, from other similar devices.

Value-added: Value-added information is differentiated from non-commoditized information in terms of its marketability or demand, which can vary, obviously, from each market that is created for the information. By way of example, information in the abstract has no value until a market is created for the information (i.e., the information becomes a commodity). The same information can be packaged in many different forms, each of which may have different values. Because information is easily digitized, one way to package the "same" information differently is by different levels of fidelity and discreteness. Value is typically bounded by context and consideration.

Authentication: A receiver of a "message" (embedded or otherwise within the value-added information) should be able to ascertain the original of the message (or by effects, the origin of the carrier within which the message is stored). An intruder should not be able to successfully represent someone else. Additional functionality such as Message Authentication Codes (MAC) could be incorporated (a one-way hash function with a secret key) to ensure limited verification or subsequent processing of value-added data.

Verification: In cryptographic terms, "verification" serves the "integrity" function to prevent an intruder from substituting false messages for legitimate ones. In this sense, the receiver of the message (embedded or otherwise present within the value-added information) should be assured that the message was not modified or altered in transit.

One-way hash function: One-way hash functions are known in the art. A hash function is a function which converts an input into an output, which is usually a fixed-sized output. For example, a simple hash function may be a function which accepts a digital stream of bytes and returns a byte consisting of the XOR function of all of the bytes in the digital stream of input data Roughly speaking, the hash function may be used to generate a "fingerprint" for the input data. The hash function need not be chosen based on the characteristics of the input. Moreover, the output produced by the hash function (i.e., the "hash") need not be secret, because in most instances it is not computationally feasible to reconstruct the input which yielded the hash. This is especially true for a "one-way" hash function--one that can be used to generate a hash value for a given input string, but which hash cannot be used (at least, not without great effort) to create an input string that could generate the same hash value.

Authorization: A term which is used broadly to cover the acts of conveying official sanction, permitting access or granting legal power to an entity.

Encryption: For non digitally-sampled data, encryption is data scrambling using keys. For value-added or information rich data with content characteristics, encryption is typically slow or inefficient because content file sizes tend to be generally large. Encrypted data is called "ciphertext".

Scrambling: For digitally-sampled data, scrambling refers to manipulations of the value-added or information rich data at the inherent granularity of the file format. The manipulations are associated with a key, which may be made cryptographically secure or broken into key pairs. Scrambling is efficient for larger media files and can be used to provide content in less than commercially viable or referenced quality levels. Scrambling is not as secure as encryption for these applications, but provides more fitting manipulation of media rich content in the context of secured distribution. Scrambled data is also called "ciphertext" for the purposes of this invention. Encryption generally acts on the data as a whole, whereas scrambling is applied often to a particular subset of the data concerned with the granularity of the data, for instance the file formatting. The result is that a smaller amount of data is "encoded" or "processed" versus strict encryption, where all of the data is "encoded" or "processed." By way of example, a cable TV signal

can be scrambled by altering the signal which provides for horizontal and vertical tracking, which would alter only a subset of the data, but not all of the data—which is why the audio signal is often untouched. Encryption, however, would generally so alter the data that no recognizable signal would be perceptually appreciated. Further, the scrambled data can be compared with the unscrambled data to yield the scrambling key. The difference with encryption is that the ciphertext is not completely random, that is, the scrambled data is still perceptible albeit in a lessened quality. Unlike watermarking, which maps a change to the data set, scrambling is a transfer function which does not alter or modify the data set.

Detailed Discussion of Invention

The LCS Domain is a logical area inside which a set of rules governing content use can be strictly enforced. The exact rules can vary between implementations, but in general, unrestricted access to the content inside the LCS Domain is disallowed. The LCS Domain has a set of paths which allow content to enter the domain under different circumstances. The LCS Domain also has paths which allow the content to exit the domain.

A simple example provides insight into the scope of an LCS domain. If an LCS is assigned to an individual, then all music, video, and other content data which has lawfully issued to the individual may be freely used on that persons LCS domain (though perhaps "freely" is misleading, as in theory, the individual has purchased a license). A LCS Domain may comprise multiple SUs, for example, a video player, a CD player, etc. An individual may be authorized to take a copy of a song and play it in another's car stereo, but only while the individual's device or media is present. Once the device is removed, the friend's LCS will no longer have a copy of the music to play.

The act of entering the LCS Domain includes a verification of the content (an authentication check). Depending upon the source of the content, such verification may be easier or harder. Unvalidateable content will be subjected to a quality degradation. Content that can be validated but which belongs to a different LCS Domain will be excluded. The primary purpose of the validation is to prevent unauthorized, high-quality, sharing of content between domains.

When content leaves the LCS Domain, the exiting content is embedded with information to uniquely identify the exiting content as belonging to the domain from which the content is leaving. It is allowed to leave at the quality level at which the content was originally stored in the LCS Domain (i.e. the quality level determined by the validation path). For example, the exiting content may include an embedded digital watermark and an attached hash or digital signature; the exiting content may also include a time stamp—which itself may be embedded or merely attached). Once it has exited, the content cannot return to the domain unless both the watermark and hash can be verified as belonging to this domain. The presence of one or the other may be sufficient to allow re-entry, or security can be set to require the presence of more than one identification signal.

This system is designed to allow a certifiable level of security for high-quality content while allowing a device to also be usable with unsecured content at a degraded quality level. The security measures are designed such that a removal of the watermark constitutes only a partial failure of the system. The altered content (i.e., the content from which the watermark has been removed or the content in which the watermark has been degraded) will be allowed back into the LCS Domain, but only at a degraded quality level, a result of the watermark destruction and subsequent obscurity to the system, consumers will not be affected to the extent that the unauthorized content has only been degraded, but access has not been denied to the content. Only a complete forgery of a cryptographically-secure watermark will constitute a complete failure of the system. For a discussion on such implementations please see US Pat. No. 5,613,004, US Pat No. 5,687,236, US Pat. No. 5,745,569, US Pat. No. 5,822,432, US Pat. No. 5,889,868, US Pat. No. 5,905,800, included by reference in their entirety and pending U.S. patent applications with Serial No. 09/046,627 "Method for Combining Transfer Function...", Serial No. 09/053,628 "Multiple Transform Utilization and Application for Secure Digital Watermarking", Serial No. 08/775,216 "Steganographic Method and Device", Serial No. 08/772,222 "Z-Transform Implementation ...", Serial No. 60/125990 "Utilizing Data Reduction in Steganographic and Cryptographic Systems".

Provable security protocols can minimize this risk. Thus the embedding system used to place the watermark does not need to be optimized for robustness, only for imperceptibility (important to publishers and consumers alike) and security (more important to publishers than to consumers). Ideally, as previously disclosed, security should not obscure the content, or prevent market participants from accessing information, which in the long term, should help develop trust or create relationships.

The system can flexibly support one or more "robust" watermarks as a method for screening content to speed processing. Final validation, however, relies upon the fragile, secure watermark and its hash or digital signature (a secure time stamp may also be incorporated). Fragile watermarks, meaning that signal manipulations would affect the watermark, may be included as a means to affect the quality of the content or any additional attributes intended to be delivered to the consumer.

LCS Functions

The LCS provides storage for content, authentication of content, enforcement of export rules, and watermarking and hashing of exported content. Stored content may be on an accessible rewritable medium, but it must be stored as ciphertext (encrypted or scrambled), not plain text, to prevent system-level extraction of the content. This is in contrast to the prior art which affix or otherwise attach meta-data to the content for access control by the variously proposed systems.

Typically, an LCS receives secured data from one or more SECDs. The SECD transfers content only after it has been secured. For example, the SECD may use an individualized cryptographic container to protect music content while in transit. Such a container may use public/private key cryptography, ciphering and/or compression, if desired.

The LCS may be able to receive content from a SECD, and must be able to authenticate content received via any of the plurality of implemented paths. The LCS must monitor and enforce any rules that accompany received content, such as number of available copies. Finally, it is preferred for the LCS to watermark all exported material (with the exception of Path 6 - see below) and supply a hash made from the unique ID of the LCS and the content characteristics (so as to be

maintained perceptually within the information and increase the level of security of the watermark).

SU Functions

The SU enables the content to be usable away from the LCS. The SU is partially within the LCS Domain. A protocol must exist for the SU and LCS to authenticate any connection made between them. This connection can have various levels of confidence set by the level of security between the SU and LCS and determinable by a certification authority or its equivalent, an authorized site for the content, for example. The transfer of content from the SU to the LCS without watermarking is allowed. However, all content leaving the SU must be watermarked. Preferably, the SU watermark contains a hash generated from the SU's Unique ID and the content characteristics of the content being transferred. If the content came from a LCS, the SU watermark must also be generated based, in part, upon the hash received from the LCS. The LCS and SU watermarking procedures do not need to be the same. However, the LCS must be able to read the SU watermarks for all different types of SU's with which it can connect. The SU does not need to be able to read any LCS watermarks. Each LCS and SU must have separate Unique IDs.

Sample Embodiment

BRIEF DESCRIPTION OF THE DRAWINGS

For a more complete understanding of the present invention, the objects and advantages thereof, reference is now made to the following descriptions taken in connection with the accompanying drawings in which:

FIG. 1 shows in block diagram form a system for one embodiment of an LCS, showing the possible paths for content to enter and exit the system.

FIG. 2 is flow diagram illustrating the functions performed by the LCS of FIG. 1 when content enters the LCS Domain from the rewritable media.

FIG. 3 is flow diagram illustrating the functions performed by the LCS of FIG. 1 when content enters the LCS Domain from the read-only media.

FIG. 4 is flow diagram illustrating the functions performed by the LCS of FIG. 1 when content enters the LCS Domain from the satellite unit.

FIG. 5 is flow diagram illustrating the functions performed by the LCS of FIG. 1 when content leaves the LCS Domain.

FIG. 6 is flow diagram illustrating the functions performed by the LCS of FIG. 1 when content leaves the LCS Domain from the read-only media.

FIG. 7 is flow diagram illustrating the functions performed by the LCS of FIG. 1 when content leaves the SU to a receiver other than the LCS.

DETAILED DESCRIPTION OF THE INVENTION

The preferred embodiment of the present invention and its advantages are best understood by referring to FIGs. 1 through 7 of the drawings, like numerals being used for like and corresponding parts of the various drawings.

FIG. 1 is a block diagram showing the components of a sample LCS system and showing the possible paths for content to enter and leave the LCS. In the embodiment of Figure 1, the LCS is a general purpose computing device such as a PC with software loaded to emulate the functions of a LCS. The LCS of Figure 1 has a Rewritable media (such as a hard drive), a Read-Only media (such as a CD-ROM drive), and software to control access (which software, in effect, defines the "LCS Domain"). The Secure Electronic Content Distributor (SECD) is connected via a network (such as the Internet, intranet, cable, satellite link, cellular communications network, or other commonly accepted network). The Satellite Unite (SU) is a portable player which connects to the LCS and/or to other players where applicable (for example by way of a serial interface, USB, IEEE 1394, infrared, or other commonly used interface protocol). FIG. 1 also identifies seven (7) path ways.

Path 1 depicts a secure distribution of digital content from a SECD to a LCS. The content can be secured during the transmission using one or more 'security protocols' (e.g., encryption or scrambling). Moreover, a single LCS may have the capability to receive content transmissions from multiple SECDs, and each SECD may use the same security protocols or different security protocols. In the context of FIG. 1, however, only a single SECD is displayed. It is also contemplated that the same SECD may periodically or randomly use different security protocols. A typical security protocol uses an asymmetric cryptographic system, an example being a public key cryptography system where private and public key pairs allow the

LCS to authenticate and accept the received content. Another security protocol may involve the ability to authenticate the received content using a signature scheme.

In FIG. 2, content enters the LCS Domain from the rewritable media (such as a hard drive). This communication path is identified as Path 2 on FIG. 1. The LCS Domain analyzes the content to determine if a watermark is present in the content. If no watermark is present, then the quality of the content is downgraded to Low Quality before it is stored in the LCS Storage. If a watermark is present, then the watermark is extracted and compared with the watermark of the LCS in order to determine if a match exists. In the event of a match, the content is permitted to be stored on the LCS Storage at the same level of quality which the content entered the LCS Domain. Optionally, if a watermark is present, the hash may be checked as further verification; and if the hash matches, the content is allowed in at High Quality. If it does not match, the content is rejected. If the extracted watermark does not match the expected watermark, then the content is denied access to the LCS Storage (i.e., the content is rejected).

In FIG. 3, content enters the LCS Domain from the Read-Only media. This communication path is identified as Path 3 on FIG. 1. The LCS Domain analyzes the content to determine if a watermark is present in the content. If no watermark is present, then the LCS attempts to further analyze the content using other methods (i.e., other than watermarking) to try and verify the content for originality. If the content cannot be verified or is deemed to have been altered, then the content is downgraded to Standard Quality (or even Low Quality) before it is stored in the LCS Storage. If a watermark is present, then the watermark is extracted and compared with the watermark of the LCS in order to determine if a match exists. In the event of a match, or in the event that the content is verified by means other than the watermark, the content is permitted to be stored on the LCS Storage at the same level of quality which the content entered the LCS Domain (which is likely to be High Quality). For example, the Read-Only media may also contain an media-based identifier which verifies the content as an original, as opposed to a copy—and hence, a non-watermark method may be used to verify authenticity.

Optionally, even in the event of a watermark match, a hash may be checked as further verification; and if the hash matches, the content is allowed in at High

Quality, but if there is no match, the content is rejected. If the extracted watermark does not match the expected watermark, or if the LCS is unable to identify any other method for verifying the content's authenticity, then the content may be denied access to the LCS Storage (i.e., the content may be rejected), or if preferred by the user, the content may be permitted into the system at a degraded quality level. It is the user's prerogative to decide how the system will treat non-authenticated content, as well as legacy content.

In FIG. 4, content enters the LCS Domain from the satellite unit. This communication path is identified as Path 4 on FIG. 1. Content from an SU is marked with an SU watermark before exiting the SU. The LCS analyzes the content from the SU for watermarks, and in particular to determine if there is a watermark that matches that of the LCS. If the watermarks match, the content is permitted access to the LCS at the highest quality level. If there is a mismatch, then the content is denied access (i.e., the content is rejected). If the content does not contain a watermark, the quality is downgraded to Low Quality before permitting access to the LCS. Optionally, even in the event of a watermark match, a hash may be checked as further verification; and access at the highest quality level may depend upon both a match in watermarks and a match in hashes.

In FIG. 5, content is shown leaving the LCS Domain. This communication path is identified as Path 5 on FIG. 1. Content is retrieved from the LCS storage and then the content may be watermarked with a watermark that is unique to the LCS (for example, one that is based upon the LCS's Unique ID). Optionally, a hash may be attached to the watermarked content, and/or the hash may be embedded as part of the watermark. If an external hash is used, preferably, for security purposes, the external hash should be created in a different manner from the embedded, watermark hash. Optionally, other information may be included in the watermark, for example, information to specify a time stamp, the number of allowable copies, etc. After watermarking, the content may be permitted to exit the LCS Domain, and may be exported to a device outside the LCS Domain, including for example, a rewritable media, a viewer, player, or other receiver.

In FIG. 6, content is shown leaving the LCS Domain. This communication path is identified as Path 6 on FIG. 1. This path is similar to Path 5, with a few

important differences. The output receiver is an SU, and because the receiver is an SU, the content may leave the LCS without being watermarked. Path 6 requires a secure protocol to determine that the receiver is in fact an SU. Once the path is verified, the content can be exported without a watermark. The LCS may optionally transmit the content together with a hash value which will be uniquely associated with the content.

In FIG. 7, content is shown leaving the SU, to a receiver other than the LCS. This communication path is identified as Path 7 on FIG. 1. Content is retrieved from the SU storage and then the content may be watermarked with a watermark that is unique to the SU (for example, one that is based upon the SU's Unique ID). Optionally, a hash may be attached to the watermarked content, and/or the hash may be embedded as part of the watermark. If an external hash is used, preferably, for security purposes, the external hash should be created in a different manner from the embedded, watermark hash. Optionally, other information may be included in the watermark, for example, information to specify a time stamp, the number of allowable copies, etc., and may even include the hash which the LCS attached to the content After watermarking, the content may be permitted to exit the SU, and may be exported to a device other than the LCS, including for example, a rewritable media, a viewer, player, or other receiver. The quality level of the content leaving the LCS is generally the same quality level as that of the content when stored internally to the LCS.

The system of the present invention is utilized to complete digital data transactions. A typical transaction would have the following steps:

1.) Using an LCS, a user connects to a SECD.

2.) The user reviews a collection of data sets which are available for license (which for purposes of this application, may be equated with a purchase). The user then selects a data set (e.g., a song or other content), and purchases (or otherwise obtains the right to receive) a copy of the data set. (The user may transmit purchase information, for example, credit card information, using digital security that is known in the art of electronic commerce.)

3.) The SECD transmits the secured content to the LCS. Before transmitting any digital content, the SECD embeds at least one watermark and may

also transmit (perhaps through cryptography) at least one hash value along with the data being transmitted. The at least one hash value may be embedded with the at least one watermark or may be attached to the beginning or end of the data being transmitted. Alternately, the hash output may be combined in ways that are known in the art.

4.) The LCS optionally may send its public key to the SECD, in which case the SECD may use the LCS public key to apply an additional security measure to the data to be transmitted, before the data is actually transmitted to the LCS.

5.) The LCS receives the secured content transmitted by the SECD. The LCS may optionally use its private key to remove the additional layer of security which was applied with the LCS's public key.

6.) The LCS may authenticate the secure content that was received from the SECD by checking the watermark(s) and/or hash values. Optionally, the LCS may unpack the secured content from its security wrapper and/or remove any other layers of security. If the content can be authenticated, the content may be accepted into the LCS domain. Otherwise, it may be rejected.

Fragile Watermark Structure

A fragile watermark—one that is encoded in the LSB of each 16 bit sample—can actually hold all of the data that would typically comprise the information being transmitted in the SecureChannel™. At a typical sampling rate of 44.1 kHz, there is 88,200 16 bit samples for each second of data in the time domain (44,100 x 2 stereo channels). This provides 88,200 bits per second which may be used for storing a fragile watermark. A typical 3 minute stereo song could therefore accommodate 1.89 MB of data for a fragile watermark. (The watermark is called fragile, because it is easily removed without greatly sacrificing the quality of the audio data.) 1.89 MB represents an immense capacity relative to the expected size of the typical data to be transmitted in a SecureChannel (100 - 200 K).

Preferably, the fragile watermark is bound to a specific copy of a specific song, so that "information pirates" (i.e., would-be thieves) cannot detect a watermark and then copy it onto another song in an effort to feign authorization when none exists. A fragile watermark may also contain information which can be utilized by various receivers which might receive the signal being packaged. For

instance, a fragile watermark may contain information to optimize the playback of a particular song on a particular machine. A particular example could include data which differentiates an MP3 encoded version of a song and an AAC encoded version of the same song.

One way to bind a fragile watermark to a specific data set is through the use of hash functions. An example is demonstrated by the following sequence of steps:

1.) A digital data set (e.g., a song) is created by known means (e.g., sampling music at 44.1 kHz, to create a plurality of 16 bit data sets). The digital data set comprises a plurality of sample sets (e.g., a plurality of 16 bit data sets).

2) Information relative to the digital data set (e.g., information about the version of the song) is transformed into digital data (which we will call the SecureChannel data), and the SecureChannel data is then divided into a plurality of SecureChannel data blocks, each of which blocks may then be separately encoded.

3) A first block of the SecureChannel data is then is encoded into a first block of sample sets (the first block of sample sets comprising—at a minimum—a sufficient number of sample sets to accommodate the size of the first block of Secure Channel Data), for example by overwriting the LSB of each sample in the first block of sample sets.

4) A hash pool is created comprising the first block of encoded sample sets.

5) A first hash value is then created using i) the hash pool, ii) a random (or pseudorandom) number seeded using a code that serves to identify the owner of the digital data set, and iii) the SecureChannel data;

6) The first hash value is then encoded into a second block of sample sets, the second block of sample sets being sufficient in size to accommodate the size of the first hash value.

7.) The second block of sample sets is then added to the hash pool

8) A second block of the SecureChannel data is then is encoded into a third block of sample sets.

9) The third block of encoded sample sets is added to the hash pool.

10) A second hash value is then created using i) the hash pool, ii) a random (or pseudorandom) number seeded using a code that serves to identify the owner of the digital data set, and iii) the SecureChannel data;

11) The second hash value is then encoded into a fourth block of sample sets.

Steps 7-11 are then repeated for successive blocks of SecureChannel data until all of the SecureChannel data is encoded. Understand that for each block of SecureChannel data, two blocks of content data are utilized. Moreover, for efficiency, one could use a predetermined subset of the samples in the hash pool, instead of the whole block.

Each SecureChannel block may, for example, have the following structure:

```
{
    long    BlockIdentifier;        //A code for the type of block
    long    BlockLength;            //The length of the block
    ...                             //Block data of a length matching BlockLength
    char    IdentityHash[hashSize];
    char    InsertionHash[hashSize];
}
```

In theory, each SecureChannel block may be of a different type of block (i.e., may begin with a different BlockIdentifier). In operation, a software application (or even an ASIC) may read the BlockIdentifier and determine whether it is a recognized block type for the particular application. If the application does not recognize the block type, the application may use the BlockLength to skip this block of SecureChannel.

Certain block types will be required to be present if the SecureChannel is going to be accepted. These might include an identity block and a SecureChannel hash block. The SecureChannel data may or may not be encrypted, depending on whether the data is transfer-restricted (a type of value-adding component, that is, VAC) or simply informative. For instance, user-added SecureChannel data need not be encrypted. A BlockIdentifier may also be used to indicate whether a SecureChannel data block is encrypted or not.

Robust Open Watermark (ROW)

A Robust-Open Watermark may be used to divide content into three categories. (The term "open watermark" is used merely to indicate that the watermark relies on a secret which is shared by an entire class of devices, as opposed to a secure watermark—which is readable only by a single member of a class of devices.) A binary setting may be used, whereby one state (e.g., "1") may be used to identify secure protected content—such as content that is distributed in a secured manner. When the LCS detects a secured status (e.g., by determining that the ROW is "1"), the content must be accompanied by an authenticatable SecureChannel before the content is permitted to enter the LCS Domain (e.g., electronic music distribution or EMD content). The other binary state (e.g., "0") may be used to identify unsecured content, for example, non-legacy media that is distributed in a pre-packaged form (e.g. CD's). When the binary setting is "0", the content may or may not have a SecureChannel. Such "0 content" shall only be admitted from a read-only medium in its original file format (e.g., a 0 CD shall only be admitted if it is present on a Redbook CD medium). On the other hand, if the ROW is absent, then the LCS will understand that the content is "legacy". Legacy content may be admitted, or optionally, may be checked for a fragile watermark—and then admitted only if the fragile watermark is present. It would be possible to permit unfettered usage of legacy content—though again, it is the prerogative of the user who sets up the LCS.

Robust Forensic Watermark

Preferably, a robust forensic watermark is not accessible in any way to the consumer—or to "information pirates." A forensic watermark may be secured by a symmetric key held only by the seller. A transaction ID may be embedded at the time of purchase with a hash matching the symmetric key. The watermark is then embedded using a very low density insertion mask (< 10 %), making it very difficult to find without the symmetric key. Retrieval of such a watermark is not limited by real-time/low cost constraints. The recovery will typically only be attempted on known pirated material, or material which is suspected of piracy. A recovery time of 2 hours on a 400 MHz PC may, therefore, be reasonable.

Sample Embodiment - Renewability

The system of the present invention contemplates the need for updating and replacing previously-embedded watermarks (which may be thought of generally as "renewing" a watermark). If someone is able to obtain the algorithms used to embed a watermark—or is otherwise able to crack the security, it would be desirable to be able to embed a new watermark using a secure algorithm. New watermarks, however, cannot be implemented with complete success over night, and thus, there inevitably will be transition periods where older SPCS are operating without updated software. In such a transition period, the content must continue to be recognizable to both the old SPCSs and the upgraded SPCSs. A solution is to embed both the original and the upgraded watermarks into content during the transition periods. Preferably, it is the decision of the content owner to use both techniques or only the upgraded technique.

The operation of the system of the present invention is complicated, however, by the presence of "legacy" digital content which is already in the hands of consumer (that is, digital content that was commercially distributed before the advent of watermarking systems) because legacy content will continue to be present in the future. Moreover, pirates who distribute unauthorized content will also complicate matters because such unauthorized copies are likely to be distributed in the same formats as legacy content. As it is unlikely that such unwatermarked content can ever be completely removed, the present system must try to accommodate such content.

Hardware can be configured to read old ROW content and extract the old ROW and insert in the content a new ROW.

Sample Embodiment – SPCS Audio Server

Tables 1, 2 and 3 depict a sample embodiment for an SPCS Audio Server, and in particular show how secured content packages are created as downloadable units (Table 1), how the LCS works on the input side for an SPCS Audio Server (Table 2), and how the LCS works on the output side (Table 3).

While the invention has been particularly shown and described by the foregoing detailed description, it will be understood by those skilled in the art that various other changes in form and detail may be made without departing from the spirit and scope of the invention.

Table 1

SAMPLE EMBODIMENT- SPCS Audio Server Stage

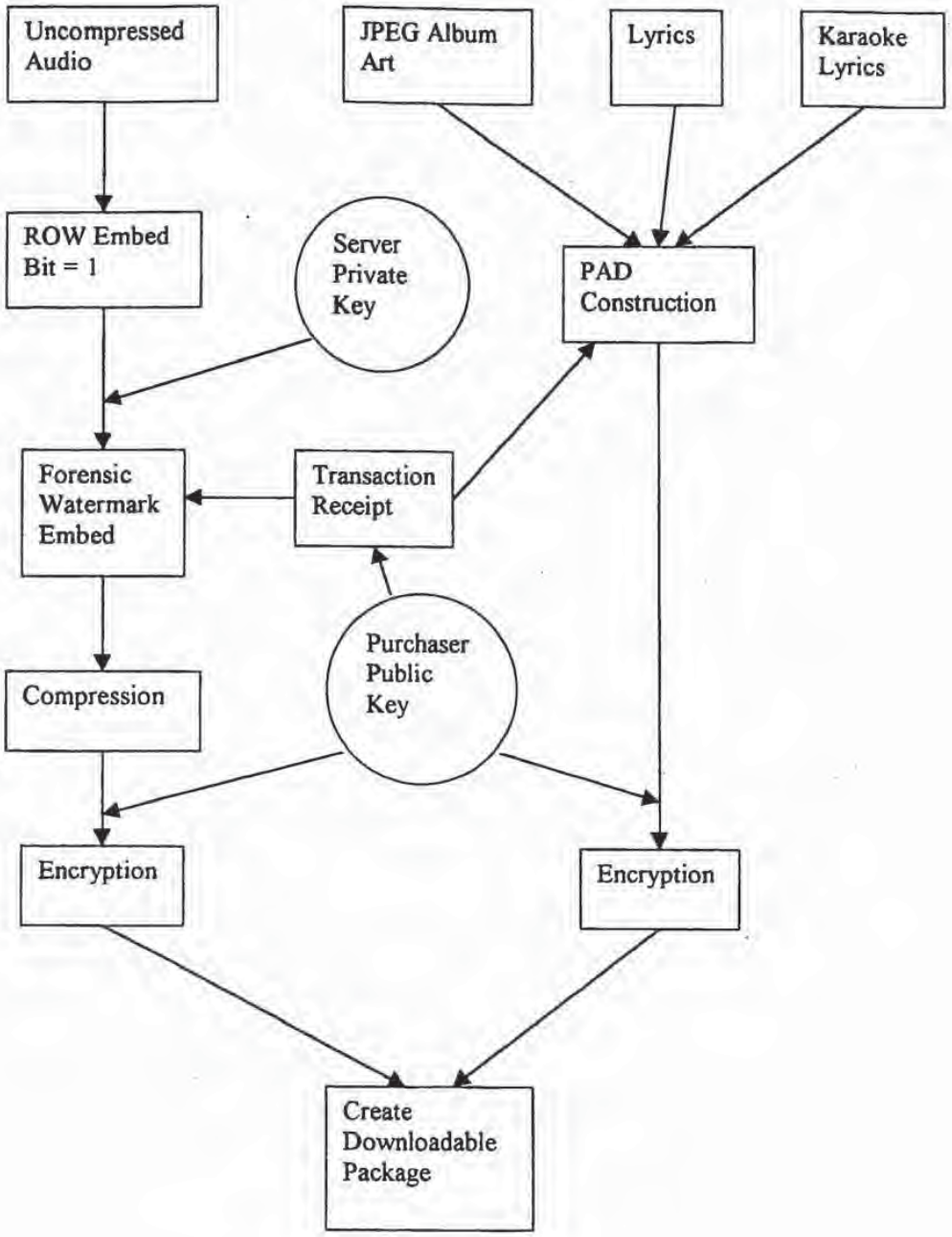


Table 2

SPCS Audio Player Input Stage

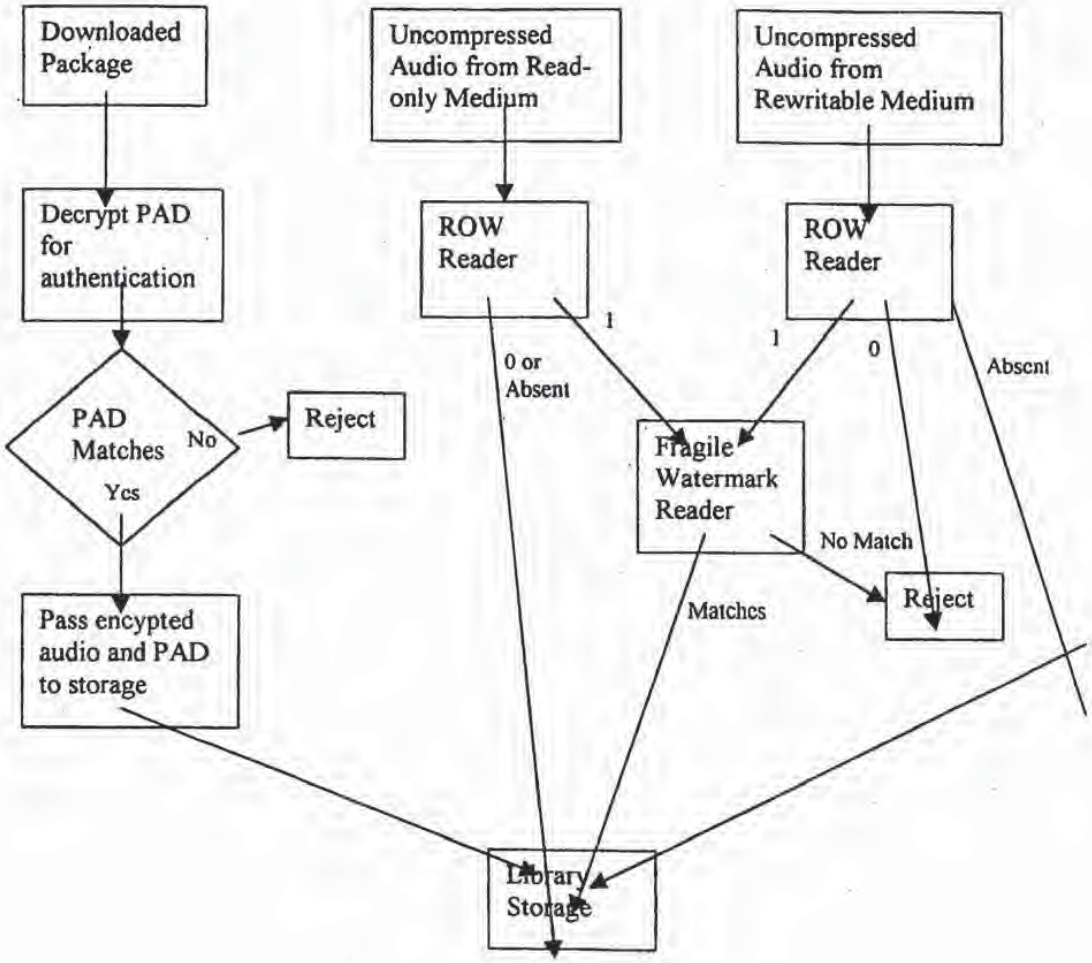
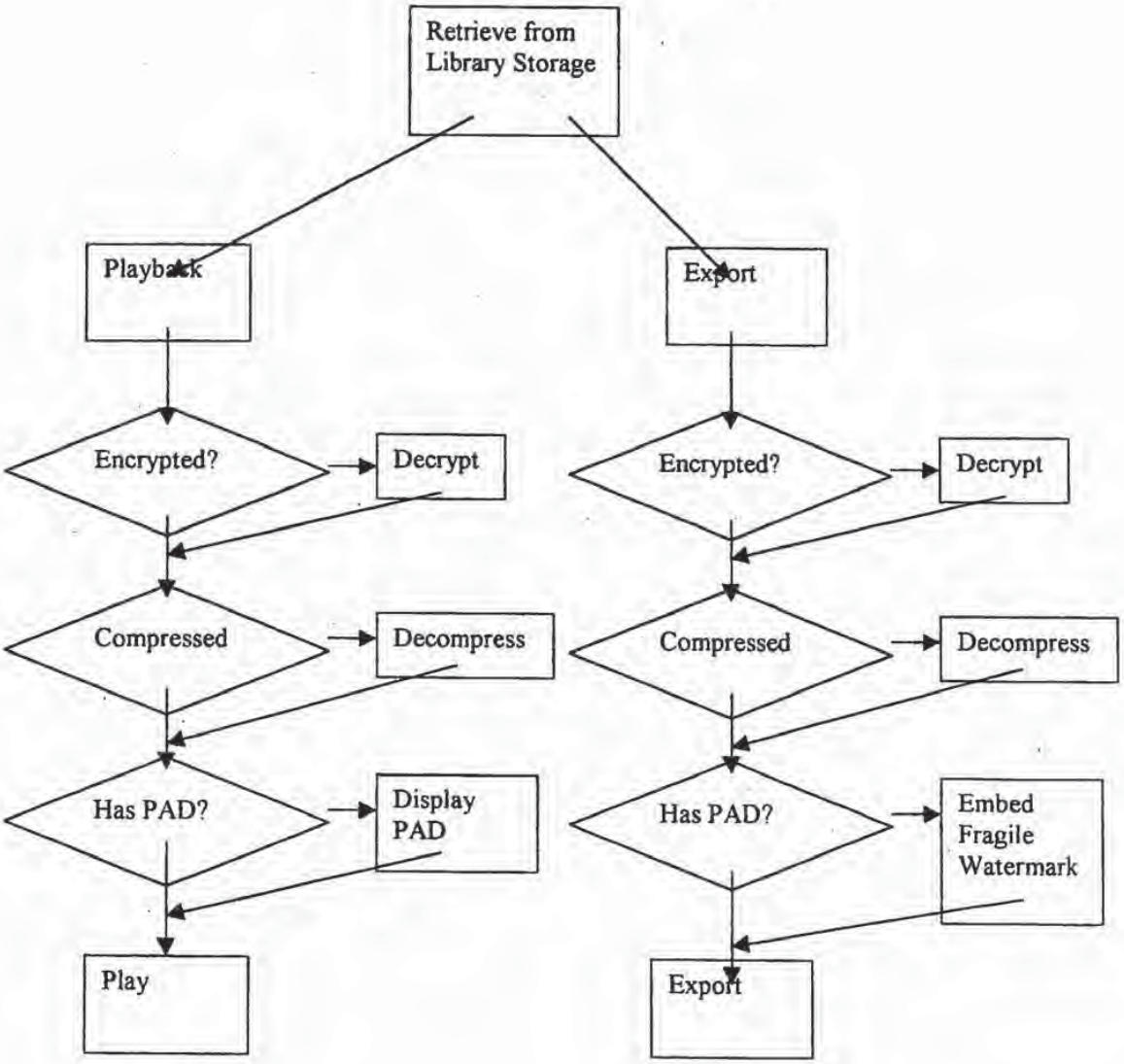


Table 3

SPCS Audio Player Output Stage



Claims:

1. A local content server system (LCS) for creating a secure environment for digital content, comprising:

a) a communications port in communication for connecting the system via a network to at least one Secure Electronic Content Distributor (SECD), said SECD capable of storing a plurality of data sets, capable of receiving a request to transfer at least one content data set, and capable of transmitting the at least one content data set in a secured transmission;

b) a rewritable storage medium whereby content received from outside the LCS may be stored and retrieved;

c) a domain processor that imposes rules and procedures for content being transferred between the LCS and devices outside the LCS; and

d) a programmable address module which can be programmed with an identification code uniquely associated with the LCS; and

said domain processor permitting the LCS to receive digital content from outside the LCS provided the LCS first determines that the digital content being delivered to the LCS is authorized for use by the LCS.

2. The LCS of claim 1 further comprising

e) an interface to permit the LCS to communicate with one or more Satellite Units (SU) which may be connected to the system through the interface, said SUs capable of receiving and transmitting digital content;

and wherein said domain processor permits the LCS to receive digital content from an SECD that is connected to the LCS's communication port, provided the LCS first determines that digital content being received is authorized for use by the LCS,

and wherein said domain processor permits the LCS to deliver digital content to an SU that may be connected to the LCS's interface, provided the LCS first determines that digital content being received is authorized for use by the SU.

3. A local content server system (LCS) for creating a secure environment for digital content, comprising:

a) a communications port in communication for connecting the system via a network to at least one Secure Electronic Content Distributor (SECD), said SECD capable of storing a plurality of data sets, capable of receiving a request to transfer at least one content data set, and capable of transmitting the at least one content data set in a secured transmission;

b) an interface to permit the LCS to communicate with one or more Satellite Units (SU) which may be connected to the system through the interface, said SUs capable of receiving and transmitting digital content; and

c) a rewritable storage medium whereby content received from an SECD and from an SU may be stored and retrieved;

d) a domain processor that imposes rules and procedures for content being transferred between the LCS and the SECD and between the LCS and the SU; and

e) a programmable address module which can be programmed with an identification code uniquely associated with the LCS;

said domain processor permitting the LCS to deliver digital content to and receive digital content from an SU that is connected to the LCS's interface, provided the LCS first determines that the digital content being delivered to the SU is authorized for use by the SU or that the digital content being received is authorized for use by the LCS,

and said domain processor permitting the LCS to receive digital content from an SECD that is connected to the LCS's communication port, provided the LCS first determines that digital content being received is authorized for use by the LCS.

4. The system of claim 3, wherein said domain processor determines whether digital content is authorized for use by extracting a watermark from the digital content being transferred.

5. The system of claim 3, wherein said domain processor comprises:

means for obtaining an identification code from an SU connected to the LCS's interface;

an analyzer to analyze the identification code from the SU to determine if the SU is an authorized device for communicating with the LCS;

means for analyzing digital content received from an SU;

said system permitting the digital content to be stored in the LCS if i) an analysis of the digital content received from the SU concludes that the content is authenticated, or ii) an analysis of the digital content received from the SU concludes that the content cannot be authenticated because no authentication data is embedded in the content, and

said system preventing the digital content from being stored on the LCS if i) an analysis of the digital content received from the SU concludes that the content is unauthenticated.

6. The system of claim 4, wherein said analyzer of the domain processor comprises means for extracting digital watermarks from the digital content received from an SU, and means for analyzing the digital watermark to determine if the digital content has been previously marked with the unique identification code of the LCS.

7. The system of claim 4, wherein said system permits the digital content to be stored in the LCS at a degraded quality level if an analysis of the digital content received from the SU concludes that the digital content received from the SU cannot be authenticated because there is no authentication data embedded in the content.

8. The system of claim 4, further comprising at least one SU, each such SU being capable of communicating with the LCS.

9. The system of claim 8, wherein the SU has means to sending a message to the LCS indicating that the SU is requesting a copy of a content data set that is stored on the LCS, said message including information about the identity of the SU, and wherein the LCS comprises:

means to analyze the message from the SU to confirm that the SU is authorized to use the LCS;

means to retrieve a copy of the requested content data set;

means to embed at least one robust open watermark into the copy of the requested content data set, said watermark indicating that the copy is authenticated;

means to embed a second watermark into the copy of the requested content data set, said second watermark being created based upon information transmitted by the SU and information about the LCS; and

means to deliver the watermarked content data set to the SU for its use.

10. The system of claim 8, further comprising a SECD, said SECD capable of receiving a request to transfer at least one data set and capable of transmitting the at least one data set in a secured transmission.

11. The system of claim 10,

wherein the SU includes means to send a message to the LCS indicating that the SU is requesting a copy of a content data set that is not stored on the LCS, but which the LCS can obtain from an SECD, said message including information about the identity of the SU;

wherein the SECD comprises:

means to retrieve a copy of the requested content data set;

means to embed at least one robust open watermark into the copy of the requested content data set, said watermark indicating that the copy is authenticated;

means to embed a second watermark into the copy of the requested content data set, said second watermark being created based upon information transmitted by the LCS; and

means to deliver the watermarked content data set to the LCS for its use; and

wherein the LCS comprises:

means to analyze the message from the SU to confirm that the SU is authorized to use the LCS;

means to receive a copy of the requested content data set as transmitted by the SECD;

means to extract at least one watermark to confirm that the content data is authorized for use by the LCS;

means to embed at least one robust open watermark into the copy of the requested content data set, said watermark indicating that the copy is authenticated;

means to embed a second watermark into the copy of the requested content data set, said second watermark being created based upon information transmitted by the SU and information about the LCS; and

means to deliver the watermarked content data set to the SU for its use.

12. The system of claim 8, wherein the SU has means to sending a message to the LCS indicating that the SU is requesting to store a copy of a content data set on a storage unit of the LCS, said message including information about the identity of the SU, and wherein the LCS comprises:

means to analyze the message from the SU to confirm that the SU is authorized to use the LCS;

means receive a copy of the content data set;

means to determine if a robust open watermark is embedded in the content data set, and to extract the robust open watermark if is it is determined that one exists;

means to analyze any extracted robust open watermarks to determine if the content data set can be authenticated;

means to permit the storage of the content data set on a storage unit of the LCS if i) the LCS authenticates the content data set, or ii) the LCS determines that no robust open watermark is embedded in the content signal.

13. The system of claim 4, further comprising at least one SU, each such SU being capable of communicating with the LCS, and being capable of using only data which has been authorized for use by the SU or which has been determined to be legacy content such the data contains no additional information to permit authentication.

14. The system of claim 5, wherein the LCS further comprises:

means to embed at least one robust open watermark into a copy of content data, said watermark indicating that the copy is authenticated;

means to embed a second watermark into the copy of content data, said second watermark being created based upon information comprising information uniquely associated with the LCS; and

means to embed a third watermark into the copy of content data, said third watermark being a fragile watermark created based upon information which can enhance the use of the content data on one or more SUs.

15. The system of claim 5, wherein the LCS further comprises:

means for encrypting or scrambling content data, such that content data may be encrypted or scrambled before it is stored in the rewritable storage medium.

16. A system for creating a secure environment for digital content, comprising:

a Secure Electronic Content Distributor (SECD);

a Local Content Server (LCS);

a communications network interconnecting the SECD to the LCS; and

a Satellite Unit (SU) capable of interfacing with the LCS;

said SECD comprising: a storage device for storing a plurality of data sets; an input for receiving a request from the LCS to purchase a selection of at least one of said plurality of data sets; a transaction processor for validating the request to purchase and for processing payment for the request; a security module for encrypting or otherwise securitizing the selected at least one data set; and an output for transmitting the selected at least one data set that has been encrypted or otherwise secured for transmission over the communications network to the LCS;

said LCS comprising: a domain processor; a first interface for connecting to a communications network; a second interface for communicating with the SU; a memory device for storing a plurality of data sets; and a programmable address module which can be programmed with an identification code uniquely associated with the LCS; and

said SU being a portable module comprising: a memory for accepting secure digital content from a LCS; an interface for communicating with the LCS; and a programmable address module which can be programmed with an identification code uniquely associated with the SU.

17. A Method for creating a secure environment for digital content for a consumer, comprising the following steps:

sending a message indicating that a user is requesting a copy of a content data set;

retrieving a copy of the requested content data set;

embedding at least one robust open watermark into the copy of the requested content data set, said watermark indicating that the copy is authenticated;

embedding a second watermark into the copy of the requested content data set, said second watermark being created based upon information transmitted by the requesting user;

transmitting the watermarked content data set to the requesting consumer via an electronic network;

receiving the transmitted watermarked content data set into a Local Content Server (LCS) of the user;

extracting at least one watermark from the transmitted watermarked content data set; and

permitting use of the content data set if the LCS determines that use is authorized.

18. The Method of claim 17, wherein the step of permitting use of the content data set if the LCS determines that use is authorized comprises:

checking to see if a watermark extracted from the content data set includes information which matches unique information which is associated with the user; and

permitting the storage of the content data set in a storage unit for the LCS.

19. The Method of claim 17, further comprising:

connecting a Satellite Unit (SU) to an LCS,

and wherein the step of permitting use of the content data set if the LCS determines that use is authorized comprises:

checking to see if a watermark extracted from the content data set includes information which matches unique information which is associated with the user; and

embedding a watermark into the content data set using information that is associated with the user and information that is associated with an SU;

delivering the content data set to the SU for its use.

20. A Method for creating a secure environment for digital content for a consumer, comprising the following steps:

connecting a Satellite Unit to an local content server (LCS),

sending a message indicating that the SU is requesting a copy of a content data set that is stored on the LCS, said message including information about the identity of the SU;

analyzing the message to confirm that the SU is authorized to use the LCS; and

retrieving a copy of the requested content data set;

assessing whether a secured connection exists between the LCS and the SU;

if a secured connection exists, embedding a watermark into the copy of the requested content data set, said watermark being created based upon information transmitted by the SU and information about the LCS; and

delivering the content data set to the SU for its use.

21. The Method of claim 20, further comprising:

embedding an open watermark into the content data to permit enhanced usage of the content data by the user.

22. The Method of claim 21, further comprising:

embedding at least one additional watermark into the content data, said at least one additional watermark being based on information about the user, the LCS and an origin of the content data, said watermark serving as a forensic watermark to permit forensic analysis to provide information on the history of the content data's use.

23. The method of claim 20, wherein the content data can be stored at a level of quality which is selected by a user.

24. A Method for creating a secure environment for digital content for a consumer, comprising the following steps:

connecting a Satellite Unit (SU) to an local content server (LCS),

sending a message indicating that the SU is requesting a copy of a content data set that is stored on the LCS, said message including information about the identity of the SU;

analyzing the message to confirm that the SU is authorized to use the LCS; and

retrieving a copy of the requested content data set;

assessing whether a secured connection exists between the LCS and the SU;

if a secured connection exists, embedding a watermark into the copy of the requested content data set, said watermark being created based upon information transmitted by the SU and information about the LCS; and

delivering the watermarked content data set to the SU for its use.

25. The method of claim 24, further comprising:

embedding at least one robust open watermark into the copy of the requested content data set before the requested content data is delivered to the SU, said watermark indicating that the copy is authenticated.

26. The method of claim 25, wherein the robust watermark is embedded using any one of a plurality of embedding algorithms.

26. The method of claim 24, further comprising:

embedding a watermark which includes a hash value from a one-way hash function generated using the content data.

27. The method of claim 25, wherein the robust watermark can be periodically replaced with a new robust watermark generated using a new algorithm with payload that is no greater than that utilized by the old robust watermark.

28. The method of claim 24, further comprising the step of:

embedding additional robust open watermarks into the copy of the requested content data set before the requested content data is delivered to the SU, using a new algorithm; and

re-saving the newly watermarked copy to the LCS.

29. The method of claim 24, further comprising the step of:

saving a copy of the requested content data with the robust watermark to the rewritable media of the LCS.

30. A Method for creating a secure environment for digital content for a consumer, comprising the following steps:

connecting a Satellite Unit (SU) to an local content server (LCS),

sending a message indicating that the SU is requesting to store a copy of a content data on the LCS, said message including information about the identity of the SU;

analyzing the message to confirm that the SU is authorized to use the LCS; and

receiving a copy of the content data set;

assessing whether the content data set is authenticated;

if the content data is unauthenticated, denying access to the LCS storage unit; and

if the content data is not capable of authentication, accepting the data at a predetermined quality level, said predetermined quality level having been set for legacy content.

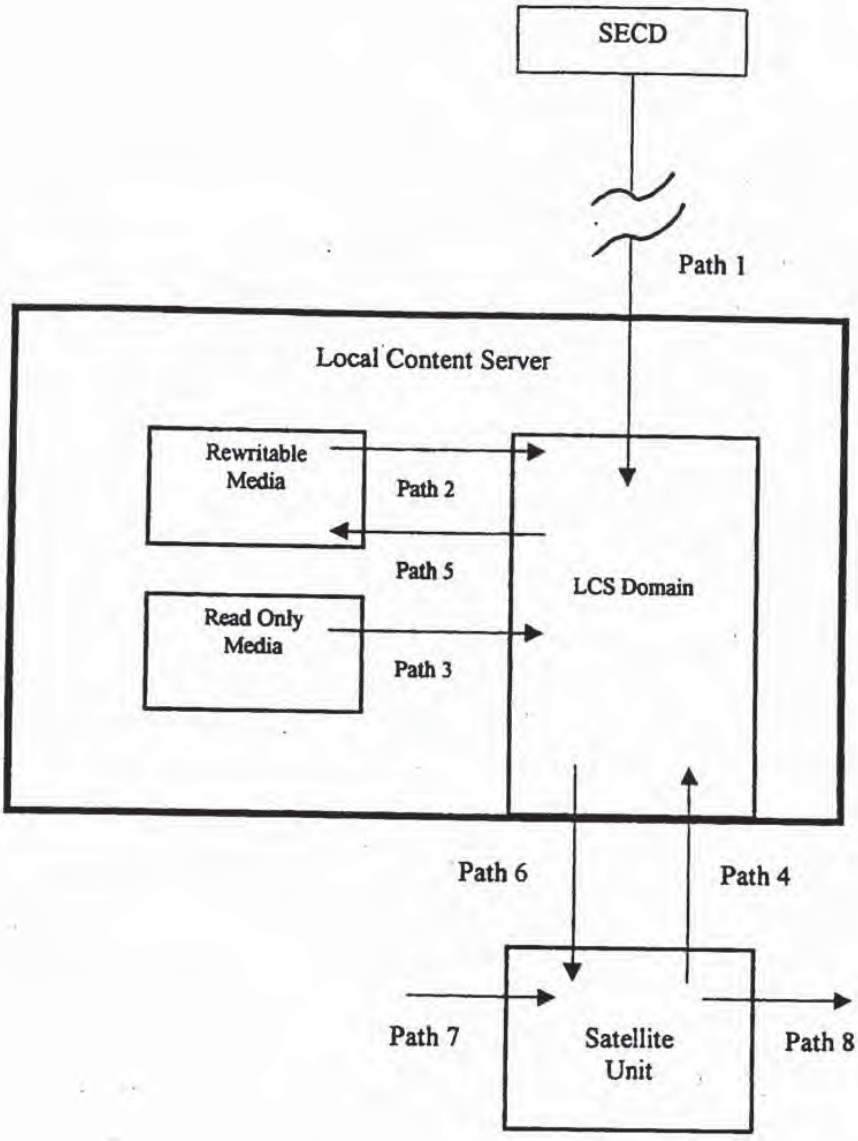

SECD
Path 1
Local Content Server
Rewritable Media
Path 2
Path 5
LCS Domain
Read Only Media
Path 3
Path 6
Path 4
Path 7
Satellite Unit
Path 8

FIG. 1

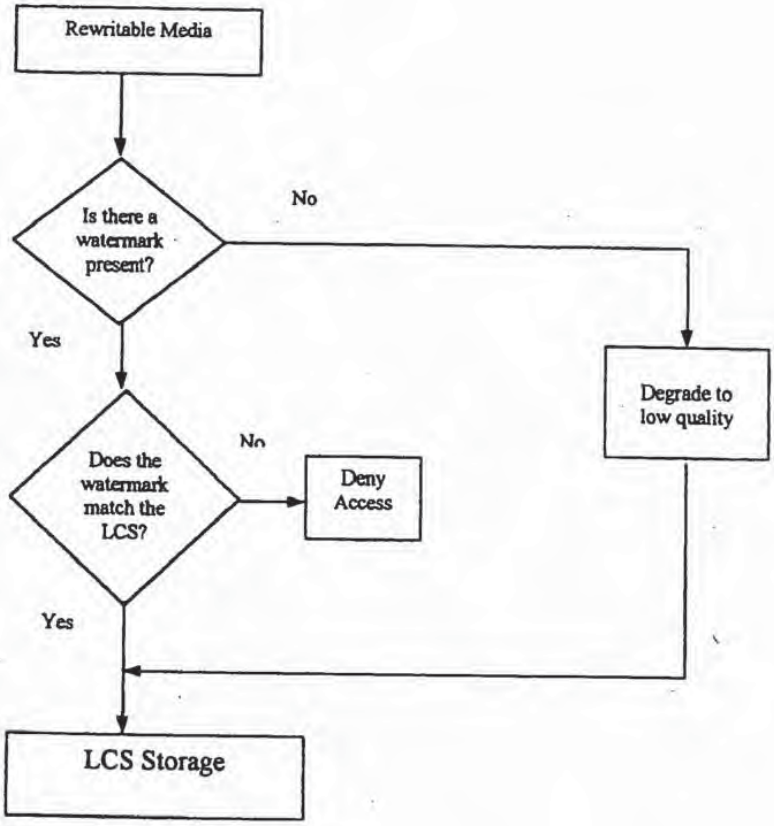

Rewritable Media
Is there a watermark present?
No
Yes
Degrade to low quality
Does the watermark match the LCS?
No
Deny Access
Yes
LCS Storage

FIG. 2

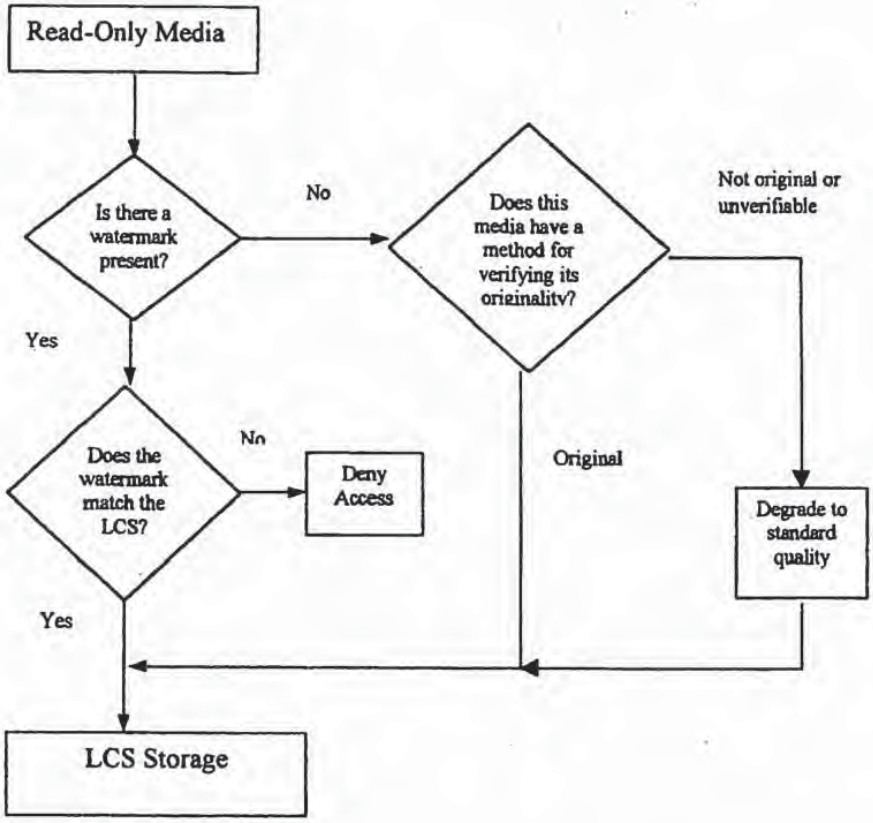

Read-Only Media
Is there a watermark present?
No
Does this media have a method for verifying its originality?
Not original or unverifiable
Yes
Does the watermark match the LCS?
No
Deny Access
Original
Degrade to standard quality
Yes
LCS Storage

FIG. 3

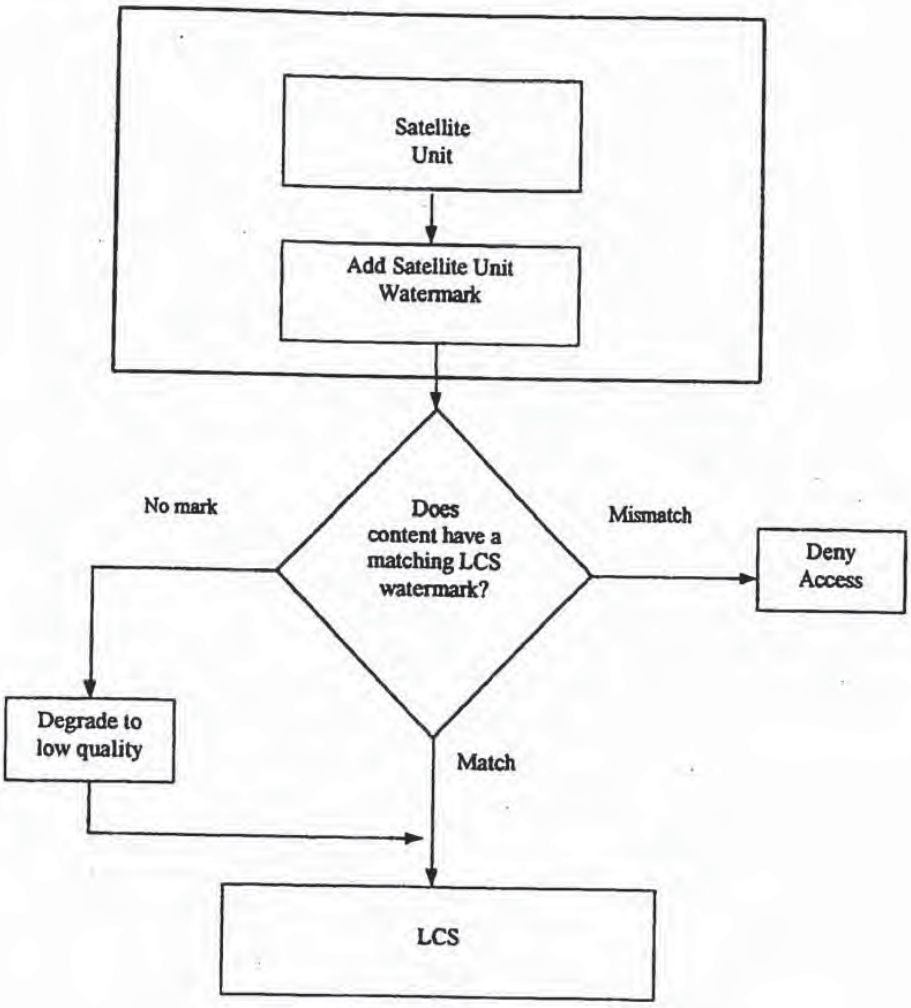

Satellite Unit
Add Satellite Unit Watermark
Does content have a matching LCS watermark?
No mark
Mismatch
Deny Access
Degrade to low quality
Match
LCS

FIG. 4

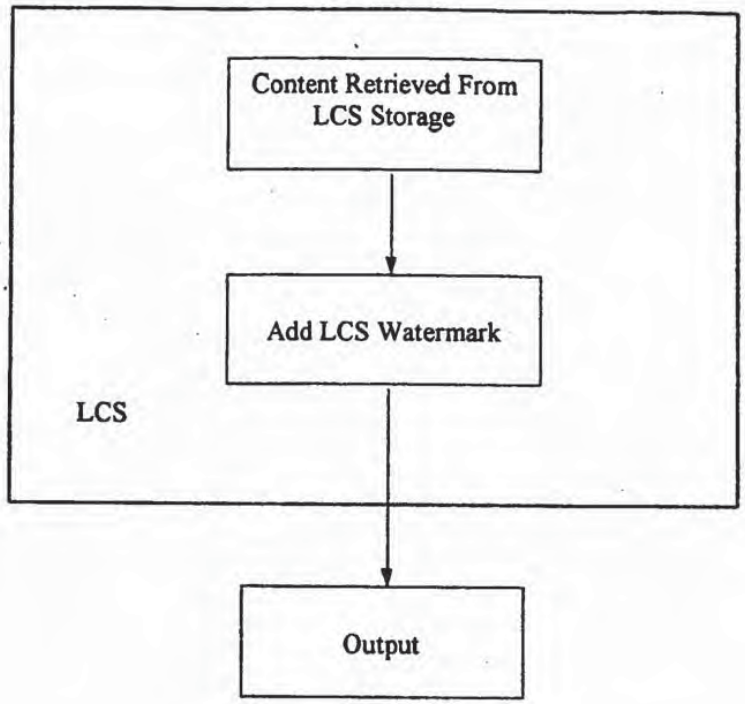

Content Retrieved From LCS Storage
Add LCS Watermark
LCS
Output

FIG. 5

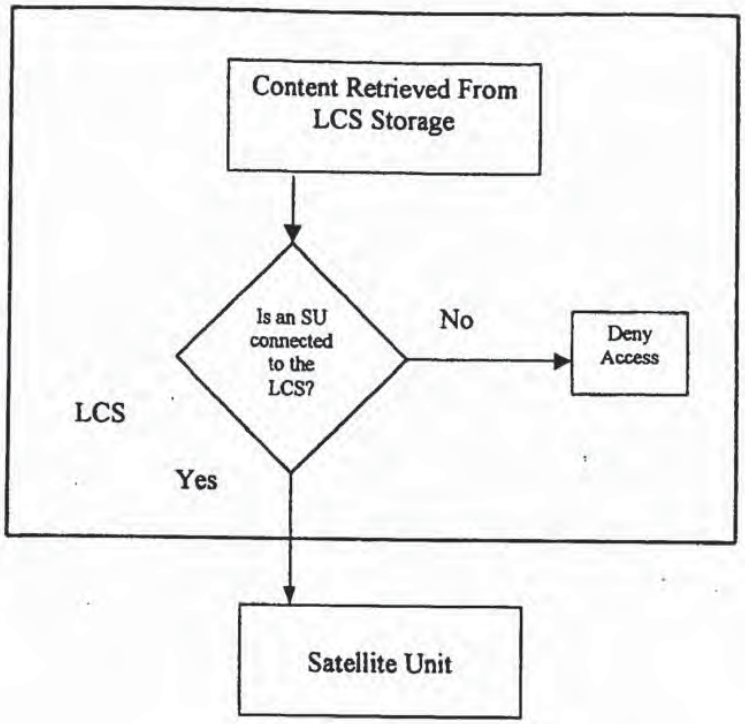

Content Retrieved From LCS Storage
Is an SU connected to the LCS?
No
Deny Access
LCS
Yes
Satellite Unit

FIG. 6

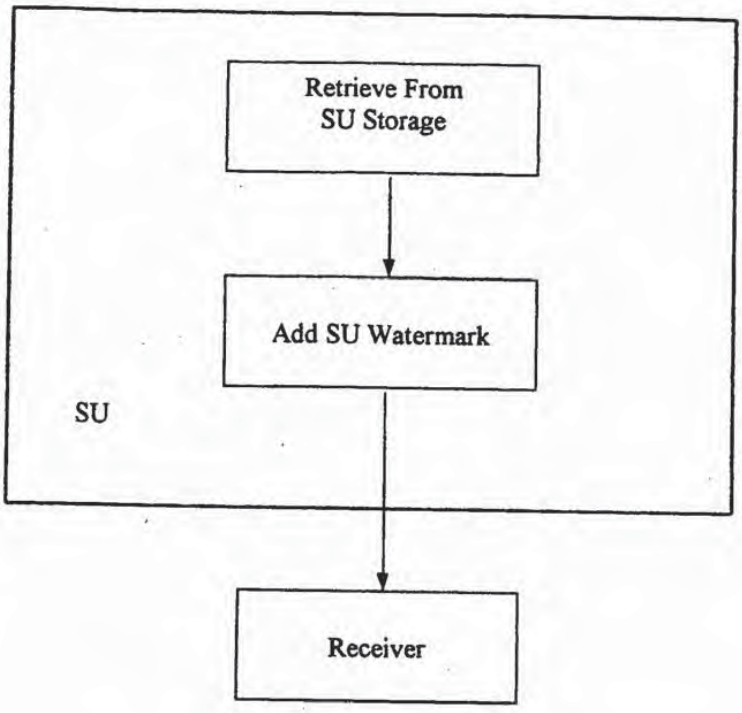


FIG. 7

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 June 2001 (14.06.2001)

PCT

(10) International Publication Number
WO 01/43026 A1

(51) International Patent Classification⁷: G06F 17/60

(21) International Application Number: PCT/US00/33126

(22) International Filing Date: 7 December 2000 (07.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/169,274	7 December 1999 (07.12.1999)	US
09/456,319	8 December 1999 (08.12.1999)	US
09/545,589	7 April 2000 (07.04.2000)	US
09/594,719	16 June 2000 (16.06.2000)	US
PCT/US00/21189	4 August 2000 (04.08.2000)	US
09/657,181	7 September 2000 (07.09.2000)	US
60/234,199	20 September 2000 (20.09.2000)	US
09/671,739	29 September 2000 (29.09.2000)	US
Not furnished	7 December 2000 (07.12.2000)	US

(71) Applicant *(for all designated States except US)*: **BLUE SPIKE, INC.** [US/US]; 16711 Collins Avenue, #2505, Miami, FL 33160 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **MOSKOWITZ, Scott, A.** [US/US]; 16711 Collins Avenue, #2505, Miami, FL 33160 (US).

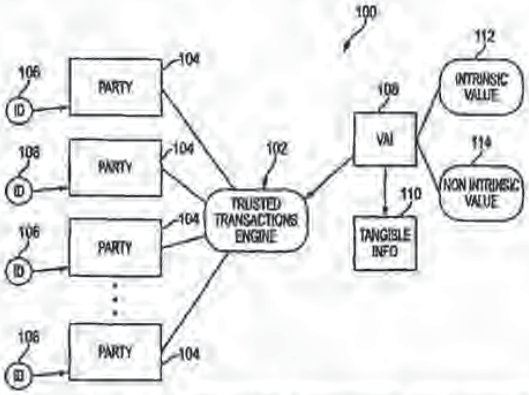
(74) Agents: **CHAPMAN, Floyd, B.** et al.; Intellectual Property Department, Brobeck, Phleger & Harrison LLP, Suite 800, 1333 H Street, N.W., Washington, DC 20005 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: SYSTEMS, METHODS AND DEVICES FOR TRUSTED TRANSACTIONS



(57) Abstract: The invention discloses a system for enhancing trust in transactions, most particularly in remote transactions between a plurality of transactional parties, for instance a seller and buyer(s) of goods and/or services over a public computer network such as the internet. Trust is disclosed to be a multivalent commodity, in that the trust that is to be enhanced relates to information about the subject matter of the transactions (e.g., the suitability of the goods and services sold), the bona fides of the supplier of the goods and services, the appropriateness of a pricing structure for a particular transaction or series of transactions, a quantum of additional transactional value that may be imparted to the transactional relationship, security of information exchange, etc. An important contributor to trust for such aspects of the transaction is disclosed to be the use of highly-secure steganographic computer processing means for data identification, authentication, and transmission, such that confidence in the transaction components is enhanced. By providing an integrated multivalent system for enhancing trust across a variety of categories (for a variety of transaction species, including those in which the need for trust is greater on the part of one party than of another, as well as those in which both require substantial trust enhancement), the invention reduces barriers to forming and optimizing transactional relationships.

WO 01/43026 A1

Published:
— *With international search report.*
— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

SYSTEMS, METHODS AND DEVICES FOR TRUSTED TRANSACTIONS

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to the transfer of information between parties; in particular, it relates to systems, methods, and devices for trusted transactions.

2. Description of the Related Art

Transactions are increasingly characterized by the amount and quality of information available to market participants. Whereas a seller seeks profit driven arrangements, which may vary over the course of a relationship with a particular buyer or consumer; buyers seek satisfaction of at least one of the following: price, selection or service. At any time the buyer or seeker of value-added information may lack recognition of the seller or provider of such information, even if coupled with a "manufactured" product or good. Sellers, or providers, similarly lack any information about individual buyers, buying groups or agents, and may only have information regarding potentially profitable transaction events defined by at least one of the following: existing market for goods or services, targeted projected market for new goods or services, or those consumers or buyers who currently engage in transactions with the provider. Transactions are the result of customer profiling, a form of recognizable pattern analysis for commerce.

Transactions conducted electronically, often in an online environment taking advantage of networks, such as the Internet and/or World Wide Web ("WWW"), form an increasingly-important subset of transactions. Most obviously, retail sales transactions in which individual customers purchase goods or services from a central web server using a WWW connection have become a prominent form of electronic transactions, though such transactions are by no means the only or even necessarily the predominant category of electronic transactions.

Electronic transactions pose special challenges for transaction parties. Some of these challenges relate to the difficulty of providing to a prospective acquirer (e.g., a purchaser) of goods or services full, accurate, and verifiable information regarding the nature, value, authenticity, and other suitability-related characteristics of the product in question. This is true in part, for instance, because the customer

cannot necessarily handle, sample, or evaluate at first hand the goods or services in question in an online transaction to the same extent to which he could evaluate them in an in-person transaction. It may also be true because of the fear of counterfeit, defective, or otherwise unsuitable products that may be viewed as more easily "passed off" (assuming a certain non-zero incidence of deceit and/or inadequate suitability verification among suppliers of products) in an electronic transaction than in an in-person transaction.

Further challenges in online transactions revolve around the serious concerns regarding security of such transactions. Such security-related concerns arise from the inherently-vulnerable nature of distributed public networks such as the internet, in which transaction parties cannot necessarily determine the path by which data travelling to and from them will take. Nor is it always possible to determine the identity of another transaction party, or to ensure that such other transaction party will take adequate precautions with sensitive data (for instance, data related to the identity or financial details (e.g., credit card number) of the first transaction party) transmitted during the course of proposing, evaluating, negotiating, executing, or fulfilling a transaction. Thus, concerns are raised about interception, inadequate safeguarding, or other unauthorized or inappropriate use of data generated or transmitted between transaction parties. Such concerns have raised the perceived need for security technologies adaptable for online transactions. Generically, these technologies have included encryption, scrambling, digital watermarking, and like methods of protecting transaction-related data.

Two conventional techniques for providing confidentiality and/or authentication currently in use involve reciprocal and non-reciprocal encrypting. Both systems use non-secret algorithms to provide encryption and decryption, and keys that are used by the algorithm.

In reciprocal algorithm systems, such as DES, the same key and algorithm is used to encrypt and decrypt a message. To assure confidentiality and authenticity, the key is preferably known only to the sending and receiving computers, and were traditionally provided to the systems by "secure" communication, such as courier.

In non-reciprocal systems, such as those described in U.S. Patent 4,218,582, a first party to a communication generates a numerical sequence and uses that -

sequence to generate non-reciprocal and different encrypting and decrypting keys. The encrypting key is then transferred to a second party in a non-secure communication. The second party uses the encrypting key (called a public key because it is no longer secure) to encrypt a message that can only be de-crypted by the decrypting key retained by the first party. The key generation algorithm is arranged such that the decrypting key cannot be derived from the public encrypting key. Similar methods are known for using non-reciprocal keys for authentication of a transmission. In the present invention, the non-secure "public" key is used to a message that has been encrypted using a secure "private" key known only to the originating party. In this method the receiving party has assurance that the origination of the message is the party who has supplied the "public" decrypting key.

SUMMARY OF THE INVENTION

Thus, a need has arisen for a system and method for enhancing trust on the part of participants in transaction. This may be with respect to all aspects of the transaction as to which trust may be an influential factor (or, viewed negatively, in which the lack of trust may be a potential bottleneck prohibiting consummation of the transaction, or of a more-optimal transaction, or of a series of transactions in a mutually-beneficial transactional relationship).

A need has also arisen for trust enhancement for transactions in connection with sophisticated security, scrambling, and encryption technology, for instance that provided by steganographic encryption, authentication, and security means.

A need has also arisen to provide these technologies in an integrated method and system, optimally requiring comparatively little processing resources so as to maximize its usefulness and minimize its cost.

The present invention represents a bridge between mathematically determinable security and analog or human measures of trust. These measures are typically perceptible or perceptual when evaluating value-added information. Additionally, a higher level of transparency between parties is assured, because information flow is recognizable and controllable by transacting parties at will.

According to one embodiment of the present invention, a method for trusted transactions is provided. The method includes the steps of (1) establishing an

agreement to exchange digitally-sampled information between a first and a second party; (2) exchanging the digitally-sampled information between the first and the second party; and (3) approving the digitally-sampled. The digitally-sampled information may be approved with an approval element, for example, a predetermined key, a predetermined message, or a predetermined cipher. The step of approving the digital information may include authorizing the digital information with the approval element, verifying the digital information with the approval element, or authenticating the digital information with the approval element. The predetermined cipher may be a steganographic cipher or a cryptographic cipher.

According to another embodiment of the present invention, a method for conducting a trusted transaction between two parties that have agreed to transact is provided. The method includes the steps of (1) establishing a secure transmission channel between the two parties; (2) verifying an identity of at least one of the parties; (3) determining an amount of value-added information to be exchanged between the parties; (4) verifying the agreement to transact; and (5) transmitting the value-added information. The value-added information may include value-adding components.

According to another embodiment of the present invention, a method for conducting at least one trusted transaction between two parties is provided. The method includes the steps of (1) authenticating the parties; (2) agreeing to a security of a transmission channel; (3) exchanging secondary value-added information; (4) determining at least one term for a primary value-added information exchange; and (5) facilitating payment for the transaction based on the terms.

According to another embodiment of the present invention, a method for conducting a trusted transaction between two parties is provided. The method includes the steps of (1) establishing a steganographic cipher; (2) exchanging secondary value-added information between the parties; (3) agreeing to terms for the exchange of primary value-added information; and (4) facilitating payment for the transaction.

According to another embodiment of the present invention, a method for conducting a trusted transaction between parties is provided. The method includes the steps of (1) identifying a unique identification for each of the parties, a unique

identification of the transaction, a unique identification of value-added information to be transacted, or a unique identification of a value-adding component; (2) applying a steganographic cipher; and (3) verifying an agreement to transact between the parties. Once the parties are identified by the unique identification, transaction identification, or the unique identification of the value-added information, secondary terms and conditions may be offered for acceptance. The transaction may take several additional steps and may include additional value-adding components to reach a legal agreement.

The agreement may cause a secondary term to be enabled for one of the parties. For example, the agreement may be related to the ability to choose ownership in the seller instead of some benefit in price, service or selection. This ownership may be priced according to traditional options pricing methodologies. Essentially the "discount" in cash value terms, may be the option price. So if there is a price, selection or service that can be equated to some cash equivalent amount, that amount can be used by the buyer as a right, but not obligation to purchase equity in the seller. Alternatively, the cash equivalent may have a direct equivalence in equity prices.

According to another embodiment of the present invention, a method for bi-directionally exchanging value-added information between parties is provided. The method includes the steps of (1) associating a plurality of unique identifiers with the value-added information, the value-added information including a digital watermark, a file header, a file attachment, and/or a file wrapper; (2) associating each of the parties with unique identifiers, the unique identifiers including a digital watermark, a file header, a file attachment, and/or a file wrapper; and (3) exchanging value-added information between the parties.

According to another embodiment of the present invention, a method for exchanging value-added information between parties is provided. The method includes the steps of (1) providing a data transmission means; (2) verifying the parties to the transaction; (3) negotiating a term. such as a price, a service, and/or a selection; and (4) binding the term to the information using a digital watermark, a file header, metadata, and/or a file wrapper. The bound transaction terms may include value-added information.

According to another embodiment of the present invention, a method for trusted transactions is provided. The method includes the steps of (1) receiving data to be processed; (2) determining a structure of the data; (3) determining if the data is authentic; and (4) determining an associated usage of the data based on the data structure and the authenticity of the data.

According to another embodiment of the present invention, a method for secure transaction is provided. The method includes the steps of (1) receiving a request to process a transaction; (2) uniquely identifying the source of the request; (3) uniquely identifying at least one term of the request; and (4) storing identification information for transaction negotiation.

According to another embodiment of the present invention, a method for the facilitation of the exchange of information data between at least a first party and a second party is provided. The method includes the steps of (1) receiving a rule governing information data from a first party; (2) receiving a request for the information data from a second party; (3) matching the predetermined rule with the request; and (4) uniquely identifying the information data and the first and second parties. The information data may include unstructured data or structured data.

According to another embodiment of the present invention, a method for the management of rights is provided. The method includes the steps of (1) receiving information; (2) determining whether the information is structured information or unstructured information; (3) identifying the information with a steganographic cipher; (4) authenticating the information with a digital signature or a digital watermark check; and (5) associating the identification and authentication results with a predetermined record, a predetermined rule, or a predetermined function.

According to another embodiment of the present invention, a method for risk management is provided. The method includes the steps of (1) receiving information; (2) determining whether the information is structured or unstructured; (3) identifying information with a predetermined ciphered key; (4) authenticating information with a digital signature, a digital watermark check, or a predetermined ciphered key; (5) associating identification and authentication results with a predetermined rule; and (6) limiting access based on a predetermined exposure of a decision maker.

According to another embodiment of the present invention, a method for securely exchanging information data between parties is provided. The method includes the steps of (1) creating a private key; (2) deriving a corresponding public key corresponding to the information data sought and at least one of (a) verifiable data associated with different versions of the information data, (b) verifiable data associated with a transmitting device, and (c) verifiable data associated with an identity of the party seeking the information data; (3) establishing a set of one time signatures relating to the information data; (4) establishing a hierarchy of access to the set of one time signatures; (5) creating a public key signature, the public key signature being verifiable with the public key, including the hierarchy of access to the set of one time signatures; (6) providing the information to a certification authority for verification; and (7) verifying the one time signature and the hierarchy of access to enable transfer of predetermined data.

According to another embodiment of the present invention, a method for authenticating an exchange of a plurality of sets of information data between parties is provided. The method includes the steps of (1) creating a plurality of hierarchical classes based on a perceptual quality of the information data; (2) assigning each set of information data to a corresponding hierarchical class; (3) defining access to each hierarchical classes and to each set of information data based on at least one recognizable feature of the information data to be exchanged; (4) predetermining access to the sets of information data by perceptually-based quality determinations; (5) establishing at least one connection between the exchanging parties; (6) perceptually recognizing at least one of the sets of information data dependent on user provided value-added information data; and (7) enabling a trusted transaction based on verification, and associated access, governing at least one of a set of information data sets.

According to another embodiment of the present invention, a method for authenticating the exchange of perceptual information data between parties over a networked system is provided. The method includes the steps of (1) creating a plurality of hierarchical classes based on a perceptual quality of the information data; (2) assigning each set of information data to a corresponding hierarchical class; (3) defining access to each hierarchical classes and to each set of information data

based on at least one recognizable feature of the information data to be exchanged; (4) perceptually recognizing at least one of the sets of information data dependent on user provided value-added information data; (5) enabling a trusted transaction of the information data based on verification of means of payment, and associated access, governing at least one copy of the information data sought; (6) associating the transaction event with the information data prior to transmission of the information data; and (7) transmitting and confirming delivery of the information data

According to another embodiment of the present invention, a device for conducting a trusted transaction between parties who have agreed to transact is provided. The device includes means for uniquely identifying unique identification information, such as a unique identification of one of the parties, a unique identification of the transaction, a unique identification of value-added information to be transacted, or a unique identification of a value-adding component; a steganographic cipher; and a means for verifying an agreement to transact between the parties.

According to another embodiment of the present invention, a device for conducting a trusted transaction between parties who have agreed to transact is provided. The device includes means for uniquely identifying unique identification information such as a unique identification of one of the parties, a unique identification of the transaction, a unique identification of value-added information to be transacted, or a unique identification of a value-adding component; and means for enabling a subsequent mutually agreed to at least one term.

According to another embodiment of the present invention, a device for conducting trusted transactions between parties us provided. The device includes a steganographic cipher; a controller for receiving input data or outputting output data; and an input/output connection. The device may have a unique identification code.

According to another embodiment of the present invention, a trusted transaction device for transmitting authentic value-added information data between parties is provided. The device includes a display; a unique identifier; means for ciphering information that is input and output; means for interacting with other similarly functional devices; and means for storing or retrieving value-added information and a value-adding component.

According to another embodiment of the present invention, a device for securely exchanging information data is provided. The device includes means for creating a private key by the party seeking information; means for deriving a corresponding public key based on the predetermined data and verifiable data associated with different versions of the information, verifiable data associated with a transmitting device, or verifiable data associated with the identity of the party seeking information; means for creating a set of one-time signatures relating to the predetermined data; means for validating a predetermined hierarchy of access of the set of one-time signatures; means for creating a public key signature, verifiable with the public key, including the access hierarchy of one time signatures; means for securely transacting predetermined data by providing information relating to a proposed transaction; and means for verifying the one time signature and the hierarchy of access to enable transfer of predetermined data.

According to one embodiment of the present invention, a system for the secure exchange of predetermined, verifiable information data between parties is provided. The system includes at least one condition for the use of the information; means for differentiating between predetermined information and other seemingly identical information based on an authentication protocol; means for associating authenticity of verifiable information data with at least one condition for use; a storage unit for storing the predetermined, verifiable information; and means for communicating with the predetermined, verifiable information storage.

According to one embodiment of the present invention, a system for the exchange of information is provided. The system includes at least one sender; at least a receiver; a verifiable message; and a verification of the message by at least one of the senders and the receivers. A verification of the message may enable a decision over receiving additional related information.

According to one embodiment of the present invention, a system for computer based decision protocol is provided. The system includes a means for identifying between structured and unstructured information; a means for authenticating structured information; and a means for enabling a decision rule based on the identity and authenticity of the information.

According to one embodiment of the present invention, a system for computer-based decision protocol is provided. The system includes means for identifying between structured and unstructured information; means for identifying structured information; and means for enabling a predetermined decision rule based on the identity of the information.

BRIEF DESCRIPTION OF THE DRAWINGS

For a more complete understanding of the present invention, the objects and advantages thereof, reference is now made to the following descriptions taken in connection with the accompanying drawings in which:

Fig. 1 is a block diagram of a system for trusted transactions according to one embodiment of the present invention;

Fig. 2 is a schematic of a local content server environment according to one embodiment of the present invention;

Fig. 3 is a flowchart depicting an example of an authentication according to one embodiment of the present invention;

Fig. 4 is a flowchart depicting an example of content flow according to one embodiment of the present invention;

Fig. 5 is a flowchart depicting an example of content flow according to one embodiment of the present invention;

Fig. 6 is a flowchart depicting an example of content flow according to one embodiment of the present invention;

Fig. 7 is a flowchart depicting an example of content flow according to one embodiment of the present invention;

Fig. 8 is a flowchart depicting an example of content flow according to one embodiment of the present invention;

Fig. 9 is a flowchart of a method for trusted transactions according to one embodiment of the present invention;

Fig. 10 depicts a device for trusted transactions according to one embodiment of the present invention.

Fig. 11 is a block diagram of a person information device according to one embodiment of the present invention;

Fig. 12 is a block diagram of an authentication device according to one embodiment of the present invention; and

Fig. 13 is a flowchart depicting an authentication process according to one embodiment of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

In order to assist in the understanding of the present invention, the following definitions are provided and are intended to supplement the ordinary and customary meaning of the terms:

Authentication: A receiver of a "message" (embedded or otherwise within the value-added information) preferably is able to ascertain the origin of the message (or by effects, the origin of the carrier within which the message is stored). An intruder preferably cannot successfully represent someone else. Additional functionality, such as message authentication codes, may be incorporated (a one-way hash function with a secret key) to ensure limited verification or subsequent processing of value-added data.

Authorization: A term which is used broadly to cover the acts of conveying official sanction, permitting access or granting legal power to an entity.

Encryption: Encryption is a method of securitizing data. For example, encryption may be data scrambling using keys. For value-added or information rich data with content characteristics, encryption is typically slow or inefficient because content file sizes tend to be generally large. Encrypted data is sometimes referred to as "ciphertext."

High Quality: A transfer path into the LCS Domain that allows digital content of any quality level to pass unaltered. "High Quality" can also mean unfettered access to all VACs.

Local Content Server (LCS): A device or software application that can securely store a collection of value-added digital information, such as entertainment media. The LCS has a unique ID.

LCS Domain: A secure medium or area where digital content can be stored, with an accompanying rule system for transfer into and out of itself.

Low Quality: A transfer path into the LCS Domain that degrades the digital content to a sub-reference level. In an audio implementation, this might be defined

as below CD Quality. Low Quality can also mean no VACs are allowed in to the system.

One way hash function: One-way hash functions are known in the art. A hash function is a function which converts an input into an output, which is usually a fixed-sized output. For example, a simple hash function may be a function which accepts a digital stream of bytes and returns a byte consisting of the XOR function of all of the bytes in the digital stream of input data Roughly speaking, the hash function may be used to generate a "fingerprint" for the input data. The hash function need not be chosen based on the characteristics of the input. Moreover, the output produced by the hash function (i.e., the "hash") need not be secret, because in most instances it is not computationally feasible to reconstruct the input which yielded the hash. This is especially true for a "one-way" hash function--one that can be used to generate a hash value for a given input string, but which hash cannot be used (at least, not without great effort) to create an input string that could generate the same hash value.

Read-Only Media: A mass storage device that can only be written once (e.g., CD-ROM, CD-R, DVD, DVD-R, etc.) Note: pre-recorded music, video, game software, or images, etc. are all "read only" media.

Re-writable Media: An mass storage device that can be rewritten (e.g., hard drive, CD-RW, Zip cartridge, M-O drive, etc.).

Satellite Unit: A portable medium or device that can accept secure digital content from a LCS through a physical, local connection and that can either play or make playable the digital content. The satellite unit may have other functionality as it relates to manipulating the content, such as recording. The satellite unit has a Unique ID.

Scrambling: For digitally-sampled data, scrambling refers to manipulations of the data. Value-added or information rich data may be manipulated at the inherent granularity of the file format, essentially through the use of a transfer function. The manipulations are associated with a key, which may be made cryptographically secure or broken into key pairs. The manipulation may be associated with a predetermined key, which may be made cryptographically secure or made into asymmetric key pairs. Scrambling is efficient for larger media files

and can be used to provide content in less than commercially viable or referenced quality levels. Scrambling is not as secure as encryption for these applications, but provides more fitting manipulation of media rich content in the context of secured distribution. Scrambled data is also called "ciphertext" for the purposes of this invention.

Encryption generally acts on the data as a whole, whereas scrambling is applied often to a particular subset of the data concerned with the granularity of the data, for instance the file formatting. The result is that a smaller amount of data is "encoded" or "processed" versus strict encryption, where all of the data is "encoded" or "processed." By way of example, a cable TV signal can be scrambled by altering the signal which provides for horizontal and vertical tracking, which would alter only a subset of the data, but not all of the data—which is why the audio signal is often untouched. Encryption, however, generally alters the data such that no recognizable signal would be perceptually appreciated. Further, the scrambled data can be compared with the unscrambled data to yield the scrambling key. The difference with encryption is that the ciphertext is not completely random, that is, the scrambled data is still perceptible albeit in a lessened quality. Unlike watermarking, which maps a change to the data set, scrambling is a transfer function which does not alter or modify the data set.

Secure Electronic Content Distributor (SECD): An entity that can validate a transaction with a LCS, process a payment, and deliver digital content securely to a LCS. This may be referred to as a "certification authority." SECDs may have differing arrangements with consumers and providers of value-added information or other parties that may conduct transactions, such as business to business relationships. The level of trust place into an SECD can be dynamically adjusted as transactions warrant or parties agree.

Standard Quality: A transfer path into the LCS Domain that maintains the digital content at a predetermined reference level or degrades the content if it is at a higher quality level. In an audio implementation, this might be defined as Red Book CD Quality. Standard Quality may also refer to a particular set of VACs that are allowed into the system.

Unique Identification, or Unique ID: A Unique ID is created for a particular transaction and is unique to that transaction (roughly analogous to a human fingerprint). One way to generate a Unique ID is with a one-way hash function. Another way is by incorporating the hash result with a message into a signing algorithm will create a signature scheme. For example, the hash result may be concatenated to the digitized, value-added information which is the subject of a transaction. Additional uniqueness may be observed in a hardware device so as to differentiate that device, which may be used in a plurality of transactions, from other similar devices.

Value-Adding Component (VAC): An attachment to the content that enhances the user's experience of the content. VACs may be metadata, headers, usage rules, etc. For music, some examples are: album art, lyrics, promotional material, specialized playback instructions. For other embodiments, the value-adding component may relate to the consumer's personal information, preferences, payment options, membership, or expectations over a transaction.

The agglomeration of value-adding components is "value-added information." In the aggregate, value creation on an informational level can be observed and measured.

Value-added Information: Value-added information is generally differentiated from non-commoditized information in terms of its marketability or demand, which can vary, obviously, from each market that is created for the information. By way of example, information in the abstract has no value until a market is created for the information (i.e., the information becomes a commodity). The same information can be packaged in many different forms, each of which may have different values. Because information is easily digitized, one way to package the "same" information differently is by different levels of fidelity and discreteness. Value is typically bounded by context and consideration.

Verification: Called "integrity," in cryptography, an intruder preferably cannot substitute false messages for legitimate ones; the receiver of the message (embedded or otherwise within the value-added information) preferably is assured that the message (or by effects, the origin of the carrier within which the message is stored) that the message was not modified or altered in transit.

Note: The above definitions may be interchanged in different embodiments of the present invention and serve as parameters in breaking down value-added information exchange and trusted transactions.

Embodiments of the present invention and their technical advantages may be better understood by referring to **Figs. 1** through **13**, like numerals referring to like and corresponding parts of the various drawings.

Increasingly, a premium is being placed on both recognition and trust. These intangible elements are both expensive to create and to maintain given the ever-decreasing amount of human contact during transactions. To the extent that many transactions are now possible without any human contact, the present invention is a unique improvement over the art in enabling bi-directional authentication of information between parties to enable "trusted transactions" between those parties

For anonymous market exchanges, transparency and data integrity, as well as confidence, serve to promote confidence and growth in product, goods and service offerings. Perception is an expensive trigger to trusted transactions reinforced by the experience of market participants.

Confidence as well as experience enable trust: in an anonymous marketplace, it is desirable for the authenticity of value-added information and value-added components to be made more transparent and independently verifiable by all concerned parties. Transparency is valued in education and experience.

A purchase decision between a buyer and a seller is equivalent to the temporal establishment of a mutually agreed "abstraction of value" in the information sought or exchanged, which may be represented in both tangible and intangible forms. Perception is the natural limit of "fair pricing," and drives value determination of a particular good or service. Perception may be structured by context, history, and/or condition. The "value" of a particular transaction has an intrinsic meaning (financial, economic, legal, political, social, statistical or actuarial meaning), temporally (at the instant of the transaction), for both the buyer and seller (reached an agreement including offer acceptance and consideration), with any inclusive terms and conditions (hereinafter, "terms") governing the transaction (price, credit terms, delivery options, and other parameters concerning the good or service with respect to which the transaction takes place). As a result of such trusted

transactions, the parties gain confidence. Even parties who may be anonymous benefit from the contemplated improvements over the art.

Referring to **Fig. 1**, a block diagram of a system for trusted transactions is provided. System 100 includes trusted transaction engine 102, which interacts with a plurality of parties 104. Each party 104 has a unique identity 106.

Value-added information 108, as defined above, includes both intrinsic value 112 and nonintrinsic value 114. A vendor (who may be a party 104) may decide what information has value (i.e., should be considered to have intrinsic value or not), and this decision may be made on a per transaction basis.

The present invention may provide advantages to all parties involved, including pricing flexibility, a reduction (or optimization) of transaction costs, a recognition of value-adding components, and the ability to provide provable security and trust among parties. Each will be discussed in greater detail, below.

1. Pricing flexibility for parties

Because buyers and sellers have complementary but competitive goals in consummating a transaction, variable pricing in the present invention is supported without any detrimental affect on the potential relationship between the buyer and the seller, or their agents. Known systems depend primarily on securing payment; payment alone, however, does not ensure the buyer and the seller of lasting protection of their respective "intangible assets," especially those that are increasingly based on value-adding information (e.g., trademarks, copyright, patents, credit history, health condition, etc.). The buyer fears identity theft ("first party," or "sentimental" piracy), while the seller fears piracy of valuable information assets ("third party," or "positional" piracy). The separation of authentication of perceptually-represented goods and services and value-adding information, from payment security, is an important novel feature of the present invention.

Known systems specify a number of methods for ensuring "security." However, the primary feature of these approaches is access control based solely on proof that a purchase has been completed. This means that if a purchase can be enabled only by determinations that a transaction was successful, the ability to entice more transactions or otherwise increase the development of maintainable trusted transactions is undermined. Simply, the fact that a purchase was completed does not

mean that a trusted transaction has, in fact, been enabled. No provision for establishing a trusted relationship between the buyer and the seller takes place absent some authenticable exchange of additional value-adding information. The present invention increases the likelihood of a successful trusted transaction and extends beyond the ability to pay (assuming no "identity theft" has occurred). The present invention provides additional means for verifiable information exchange that enhance the experience of the buyer and the seller in seeking trusted transactions.

Because many manufactured goods are likely to have similar costs from a strict manufacturing standpoint, the value-added service, or services, that are provided to the buyer are likely to encourage additional opportunities for trusted transaction. The seller can benefit by leveraging a single purchase into a profitable relationship. Even distribution costs may be commoditized for all similar tangible goods. A series of non-contiguous or non-temporal transactions alone would constitute a profitable relationship if the buyer is satisfied and the seller is profiting. That pricing, and its terms, may be varied dynamically or supported flexibly (based on information exchange at the time or leading to a transaction), is another improvement over the art. The incorporation of micropayments becomes more feasible as the cost of trust has been reduced and thus smaller discrete increments of monetary consideration are easier to support to the benefit of buyers and sellers seeking higher granularity or discreteness over the information or tangible goods they transact. Simply put, identification and authentication of specific information and value-added components is inherently important to further segmentation of units of payment (e.g., micropayments). Micropayments may be interpreted as a value-added component in facilitating transactions.

Pricing may also be bi-directional and asymmetric, and is preferably determined by the seller in order to define "profitability." Some sellers may choose to maintain fixed pricing for their goods or services, but may incorporate variable pricing in the value-added component. For instance, while the price of a given good or service may be fixed, the value-added component may be the terms of the pricing as it effects the buyer. The seller may also entice the buyer to provide demographic value-added components, or related data, which has intrinsic, sentimental value to the buyer. To the seller, the pattern, or structure, of demographic datum serves as a

valuable filter in which to position its offerings. Simply put, while barter is relatively inefficient, cash, being anonymous, may not reveal enough information to provide an incentive for the seller to vary credit terms or offer a greater variety of goods and services, even if there is a single underlying value-added information good (the seller can still offer perceptually similar but nonequivalent versions of the information without threatening secure, higher quality, limited, or more expensive versions).

The ability to offer both secure and unsecure, or legacy, versions of the same information based on a mutual disclosure and mutual understanding of both the buyer and the seller is particularly novel in the art. Moreover, privacy can be enhanced and new, unproven and yet unsecure information can be offered without jeopardizing the security of any pre-existing primary value-added information whether it be music, images, currency, electronic documents, chip designs, source code, legacy versions, prior art, etc.

The period of payment, like the discreteness of the actual payment, interest rate relating to a payment period, grace periods, early payment benefits, variable interest rate based on the seller's ability to assess the credit risk/worthiness of the buyer or its agent, etc. is an element or component (a value-added component) that may be changed to affect a transaction. Making these components more transparent to buyers improves the opportunity for enhancing and maintaining trust. It also enables buyers and sellers to make mutually beneficial decisions based on transparent, verifiable information or value-added components. Moreover, buyer-driven pricing, as with Dutch auctions, or market-based pricing, are not possible without compromising the access-based security in known systems. With the present invention, goods and services are better able to realize full market value because access to the good or service is not restricted (such as with new music or new endeavors by "unknown" or "unrecognized" artists, designers, creators or engineers). The market participants are better able to assess the good or service in question, and/or the related value-adding information/component, when experience and information sharing is encouraged. The prior art is restrictive by necessity in information sharing precisely because security cannot be maintained by prior art systems with such open access to information.

For goods or services that are difficult to value (e.g., media content, legal advice, design, non-commodity items, etc.) and decision-intensive, pricing becomes a barrier to entry in a marketplace that puts a premium on recognition. Highly recognized artists, lawyers, designers, retailers, etc. have a competitive advantage over their unrecognized competitors. One approach to gaining recognition is freely distributing or providing goods or services. Ultimately, the seller still needs to profit from this initial positioning to the extent that financing of operations is available (the seller can stay in business as long as investors or financing is available to enable such operations). The same goods or services may be offered in a "tiered" manner, which relates to the purchase price or to the quality of the underlying good or service to be exchanged. Examples of this include providing music in MP3 quality audio instead of CD quality; providing 10 hours of customer support instead of charging per hour; charging service charges instead of free checking or ATM access; charging a price per bit or bandwidth; etc.

Segmenting also plays a role in the "freshness" or "newness" of the information good or service. Live concerts or lectures may be worth more to the buyer than pre-recorded versions offered later or separately. The performer or creator of the information to be performed, or conveyed live, can only be at one place at a time, and may be a premium for that time. Live broadcasts may similarly have a higher value. Physical advice may be worth more than printed literature to the buyer as well. These dynamics create an impetus for flexible and dynamic pricing that does not undercut the security of the overall "trusted transaction" methods and systems envisioned in the present invention.

In known systems, legacy information, relationships, etc. systemically undermine the ability to ensure a "trusted system." The buyer and the seller in the art have no means for differentiating between the secure and unsecure versions of a good, service, or value-adding component. The present invention provides such protocols by incorporating additional bits of data, which do not necessarily represent added data, but imperceptibly replace data with identifying or authenticating data, enabling market participants to determine whether a value-added information "package" is secure. This also enables uniqueness of information packages to be consistently created and checked or maintained for later reference. The prior art

relies on the denial of access or access restriction, a clear disadvantage in increasing the availability of value-added information. With trusted transactions market participants are able to verify, identify, and price information and then decide which versions are appropriate for a given or existing demand.

Pricing may be better understood if the cost or time of computation is measured as a tangible asset. Similarly, the natural limit to theft of tangible assets has always been in the cost of the tangible assets. As information can increasingly be traded for value in excess of the cost of its storage or transmission, pricing becomes less tangible and more subjective. Delivery of information accurately and quickly becomes a valued service. Measuring such value is based on the same principles that allow cost estimates of the delivery of fixed weight parcel packages. The existence of hackers indicates a lowered economic barrier to entry for informational crime, including identity theft and piracy. Dissemination of binary code, which is similarly detrimental, at little or no cost to the originator of the valuable information, introduces novel concepts to the approaches of information pricing. Tangible goods become substitutes for cash payment.

An example of pricing based on effort is illustrated by a watchmaker who takes six months to finish a watch that he prices at $70,000. This includes a "reasonable" profit and the cost of materials. The buyer is a watch fanatic and earns $140,000 a year. The exchange of a tangible good that has intrinsic value, which is converted into monetary terms for negotiation, as agreed by the parties in the exchange, becomes more prominent if information concerning value is transparent or fluid for all market participants. Transparency is inherently favored by markets seeking to appropriately price goods or services based on all available information at the moment of pricing. Conversely, risk can be priced based on the financial context or structure of an organization. Those who earn $20,000 should have to have confirmation by others with additional financial or fiduciary responsibilities before validating or approving transactions that exceed an individual's earnings for the period in question. At any time responsibility can be linked to authority, as a pricing mechanism for decisions concerning similar amounts of monetary consideration. With pricing mechanisms and use rules, trusted transactions offer flexible pricing not possible with current systems.

Value-adding components, which may include pricing, is preferably viewed as a separate and distinct means for the buyer and the seller to separate information that may or may not be essential to any given transaction and may also be viewed as nonessential unless both parties can stipulate such information exchange. This is invaluable as multiple channel distribution of the "same" goods (e.g., download music over the Internet versus purchasing a CD from a store) or services (obtaining a mortgage online versus processing physical loan documents) can be offered by the seller. Determinations of which channel, or channels, are profitable requires verification of unsecure and secure versions of these "same" goods.

Value-adding components may also include an offer, an acceptance, a bid, a purchase, and a sale of a securities instrument, including an option, a warrant, or equity.

Security is inherently intended for the party seeking value or authentication over the information or transaction and conversely protecting sentimental information or identity from being stolen or defrauded. For the long term, buyers are able to differentiate that personal information value-added components are appropriate for dissemination to a seller to affect a transaction, or to get better terms. Either the buyer or the seller, or both, are better able to determine that transactions or relationships are favorable on a transaction to transaction basis, and thus "transact" accordingly.

Pricing of the value-added information may include a value-adding component relating to the present value of recognition/non-cash equivalent cost/service that is handled in a separate negotiation or transaction, or a subsequent negotiation or transaction

The present invention may include limits of liability, or may consider the time value of money when determining a limit of liability threshold. The present invention may enable rules/access/authorization based on the result of that operation. In one embodiment, an actuarial estimate of liability (future time) or cost (present time) may serve as a rule for enabling another rule.

2. Reduction or optimization of transaction costs

In instances where the buyer and the seller, or their agents, seek to transact products or services that include value-added information, the seller generally seeks

to maximize profit, but may forego profit in the short term to ensure recognition or market share in the short term. The buyer seeks "satisfaction," which is dependent on one or more of the following product/service determinants: 1) price; 2) service; and 3) selection. These determinants may be quantitatively or qualitatively assessed and may be based on available bandwidth, time of transaction, and transaction event conditions.

A priori, the buyer may not recognize the seller. In an information economy, such events are not a disincentive to pursuing a trusted transaction, but instead present market opportunities for valuing, authenticating, and verifying information (all may be value-added components) concerning potential transactions are inefficient. Conversely, the seller may not have enough information about the buyer to determine what type of potential transaction can be enabled, based on the buyer's ability to purchase now, or at any point in the future. The seller may be inclined to make a sale with the buyer (or the buyer's agents) with or without confidence that the initial transaction will lead to further transactions or trusted relationships that are profitable for the seller. The seller may use purchasing options (e.g., barter, cash or its equivalent, or credit) to enable a purchase by the buyer. According to one embodiment of the present invention, because value-adding information and its components may be bi-directional, both the buyer and the seller may chose to negotiate the transaction, including variable terms for payment, as one form of value-added component or service and support for the information to be transacted.

Transactions, as defined by a purchase event (payment can be preliminarily assured), may happen before or after the buyer and the seller have "agreed" to transact. When the seller requires value-adding components/information about the buyer before entering the transaction, the seller generally has higher risks than the buyer, which may affect its profitability. Where there is a high risk for piracy, such as the digital copy problem (that can render individual copies of value-added information worthless), the seller may not be able to establish trust with an unknown buyer. The seller is not assured of any potential profitable transactions or long-term relationship with the buyer, which poses a significant risk to the seller if the buyer pirates information goods or services. A lack of dynamic authentication, even in

real time, at least initially, and adjusted as needs arise over time, and flexibility in negotiable terms, may cause the seller's assets to be economically undervalued.

Conversely, in those events where the buyer requires value-adding components/information about the seller in advance of entering a transaction, the buyer generally has higher risks than the seller with regard to its ability to enter into transactions. "Identity theft" is an example of a risk that is higher for the buyer than the seller in these types of transactions. Additional transactions include on-line brokering, auctions, searches, bots, webcrawlers, recognition, and determination of goods or services absent proof of privacy guarantees. This applies to noncommercial information as well (e.g. the FDIC logo, currency, driver's license, etc.)

The establishment of mutual trust may be asymmetric depending on the risk profile of the buyer and the seller. Risk/reward tradeoffs are implicit to some transactions, while the time required to establish a trusted transaction or eventual profitable relationship may not be contiguous. In many on-line transactions, the per transaction risk is generally higher to the buyer, who may suffer fraud and may need to be more diligent about what value-adding information it chooses to exchange in the interests of enabling a trusted transaction. It is true, however, that in business to business transactions ("B2B"), or in financial information exchange, the relative risks to each party are relatively equivalent, and requiring a more symmetric exchange of value-adding components relating to verification and purchasing power (in the form of barter, cash, cash equivalents or financing that would also constitute value-adding components) is not as necessary. Reducing the cost of creating and maintaining trust is an advantage of the present invention over known systems.

3. <u>"Reintermediation": recognition as a Value-added Component</u>

Asymmetry exists in recognition as well. Where word-of-mouth may constitute an acceptable means for creating recognition for a particular good or service, the buyer and the seller may wish to expand their respective abilities to capture more of the increasingly available goods and services, or value-adding information (about themselves, or terms for a trusted transaction). With advertising and other forms of marketing, the push and pull of value-adding information between the buyer and the seller also contributes to potential purchase decisions by

both parties or their agents. The buyer may control certain criteria it seeks, such as price, selection, and/or service. The seller, conversely, seeks the highest profits from a given potential buyer or his agents, which may not be quantifiable from the first transaction or may not be the primary focus of the seller (such as seeking a valuable, marquis client). Both the buyer and seller may compare patterns or structure that, when recognized, help in forming opinions about the history, condition or context of the information.

In general, recognition serves to encourage more recognition. The seller will likely seek trusted transactions in the interests of profitably leveraging the time, cost and expense of generating the initial exchange of goods and services with the buyer. Over the longer term (defined as any additional transactions beyond the initial transaction), a profitable relationship is sought by the seller. The buyer and the seller may still maintain flexibility as expectations or needs concerning the relationship change. The present invention allows for such variability and flexibility by enabling real time adjustments to the terms that prevail between market participants. While terms are conditions are negotiatiable, security of the overall system is not jeopardized because secure and unsecure versions of the "same" value-added information and value-added components can be adjusted bi-directionally. In an information-based transaction, there is value in reintermediation by sellers seeking to ensure that their information is provably identifiable and verifiable.

The buyer and the seller may seek recognition or use means for increasing visibility of their respective interests. The buyer ultimately seeks to satisfy itself through a trusted transaction preserving private or financial information for select transactions requiring higher amounts of information exchange or verification (real time references, "membership reward programs" such as frequent flier airline points, or financing options that can be dynamically offered, are two incentives to the buyer and are likely to differentiate vendors, large and small, really or perceptually); the seller ultimately seeks to profit from the trusted transaction. Recognition of this potential exchange between the parties is not assumed to be high enough to enable a transaction, but high enough to create exposure for the buyer or the seller. Trust is assumed to not be pre-existing, or it may be variable between the buyer and/or the seller, requiring additional exchanges of value-adding information to enable a

trusted transaction. The seller, in the extreme, seeks the highest profit for each transaction. The buyer, in the extreme, seeks the highest satisfaction for each transaction. As discussed above, both goals are complementary and competitive, thereby increasing the need for dynamic exchange of value-adding information. Recognition can enhance the potential for a successful trusted transactions and serves as a form of abstract experience for both parties to efficiently make decisions. With experience, value assessments become possible. Abstractions of value become experience as trusted transactions beget more trusted transactions.

4. Provable security and trust

Trusted transactions are characterized primarily by bridging the gap between "provable security" and the imprecise nature of trust. Encryption, cryptographic containers, digital watermarks and other forms of electronic data security can be mathematically demonstrated -- discrete algorithms can be designed to meet certain pre-defined specifications or pre-defined expectations.

Encryption and secure digital watermarking (e.g., steganographic ciphering) offer tools for determining data integrity, authenticity and confidence. Transactions, however, still require human decision-making. Known systems describe a number of approaches for ensuring transactional security based solely on transmission security and fail to differentiate between what could be called "positional piracy" (e.g., the fraud or theft of universally recognized goods, products, and services) and "sentimental piracy" (e.g., the fraud or theft of personal, private or financial information).

For the purposes of this disclosure, the extreme case of sentimental piracy is identity theft. So long as information can be represented in binary digits (0s and 1s), and can be easily copied, stored or transferred, identity fraud becomes an increasingly insidious problem. There is a temporal limit whereby the actual person is able to "reclaim" their identity at some point in time. The extreme case of positional piracy is zero returns on an intangible asset that has been pirated. As well, the present invention offers advantages over known systems for positional piracy that enable the continuation of legacy business, customer relations and existing information formats, without sufficiently weakening any overall system security for trusted transactions. Simply, unlike known systems, access restriction is

not an adequate or appropriate means for ensuring the security of information data for a wide variety of applications.

To the extent that "security by obscurity" is typically representative of weak security to those skilled in the art of cryptography, more transparency for parties to a transaction over security protocols and information transfer are inherently necessary to ensure trusted transactions. Although information between parties may be asymmetrically exchanged (i.e., the value-added information or value-adding components is not equivalent in quality or quantity between parties, such as a difference in the amount of information exchanged, the identification of the parties, etc.), the level and degree of authenticity or verification only differs among the goods, products or services to be transacted, as well as the demands of the market participants. For the purposes of this disclosure, the value-added information is the fundamental good to be transacted between parties, while value-added components represent an atomic unit of data that is defined as the least amount of data that can either add functionality or be perceptibly recognized to a system for trusted transactions. Data may be represented in analog or binary terms in order to establish uniqueness and assist in identification and authentication. Value-added components may be added, subtracted, or changed to vary the underlying value-added information sought.

Because humans have difficulty remembering passwords, personal identification numbers (PINs), and the like, dependence on such datum is increasingly problematic as more anonymous transactions are enabled between parties over electronic networks, such as the Internet, or between businesses in private networks. While passwords, or PINs, are commonly thought to be secure, the ability to check all combinations of numbers or crack passwords becomes less computationally expensive with increases in both processing speed and availability of bandwidth. Cost is reduced to the detriment of security if any individual has the means for high order computation or network-based bandwidth in discovering or hacking any given secret. Quantum computing speeds up the ability to test and discover such data at even greater speeds, and presents unique problems to security systems described in the art. Quantum computing also enables the definition or predetermination of the physical limitations of communicating or securing

information. Where difference between binary or digital signal processing and quantum mechanical limits is higher, better security is enabled.

Biometrics have been suggested to remedy this problem, but do not offer any way to create truly cryptographic secrets to be shared between parties. Iris scans, fingerprints, and the like, are easily stolen because they are easily perceptible to those seeking to defraud. Once stored electronically, biometrics be stolen for unauthorized use. Combining a biometric with a digital signature may provide a means to ensure that a given representation of a fingerprint or iris is fixed, temporally at the time the certificate is created, but does not prevent dedicated attacks at determining the fingerprint or iris to be used at some subsequent time. Real time authentication and verification are improvements envisioned with the present invention. Assuring that a particular fingerprint, signature or iris "data set" is that of the intended user, is fundamentally important to embodiments described herein. This becomes especially invaluable with increasing number of anonymous transactions. Although uniqueness may be enhanced with digital signatures and digital iris or fingerprint records, the advantage with the present invention is that more secure forms of uniqueness based on a predetermination of the discreteness of time and a predetermination of the limits of information conversion and transfer are absent in the art.

Moreover, real time authentication is not enhanced with systems described in the art, since such biometric data is easily stored or transferred, and thus suffers the same pitfalls for any binary data that is sought by a party seeking to defraud. Biometrics may be great for forensics (e.g., to determine after the fact who is responsible for a particular act), but they do not effectively address an inherent problem in enabling trusted transactions; that is, real time verification of parties or real time association of parties with information being transacted (in an auction, for instance). They are also not representative of a cryptographic key, which, as is well-known in the art, requires secrecy, randomness, and an ability to update or destroy the cryptographic key.

Another advantage of the present invention is the ability to serialize or individualize "personal secrets" that are shared between parties to boost confidence and transparency of transactions. That control, and the inherent uniqueness of

personal entropy, constructed from such information as a hometown, favorite restaurant, or high school sweetheart, is a means for perceptible representations of "secret data" that enhances the ease-of-use and application of appropriate shared secrets to be exchanged in conducting trusted transactions. Associating such secrets with primary value-added information or value-added components being transacted is an additional novel feature of the present invention. Essentially, the present invention provides the ability to personalize or serialize, informationally, an actual "transaction event," including: the buyer; the seller; primary information; value-added components and tangible assets created, manufactured, or manipulated; and any additional reference that can be made perceptible and secure to any observer. Bridging cryptographic with real world perception is a benefit over the prior art.

Essentially, randomness alone, whether pre-determined or not, is not sufficient for the creation of a "secret" that may be used with high levels of confidence repeatedly in assuring the validity of information or verify the identity of a party. Encryption systems cipher the randomness according to available data capacity; digital watermarking ciphers the randomness according to perceptible features or characteristics of the carrier signal (a humanly-perceptible measure of data capacity, which distinguishes applications for encryption from secure watermarking). That such information can be made more computationally difficult to discover, even by brute force attacks (since such experience is only limited by the experience of individuals) is of particular benefit to the art. The computational complexity added by use of a steganographic cipher is discussed in the U.S. Patent No. 5,613,004, the disclosure of which is incorporated by reference in its entirety, and offers a means for human observers to see the actual tampering of information represented perceptibly. This proof is self-similar to that which is obvious in the real world, i.e., the ease at which one can observe that a watermark is missing from currency. Handling information as contemplated by the present invention for trusted transactions is unique in bridging computational benefits from both digital signal processing and cryptography to the benefits of all parties to a transaction. The present invention is the enhancement of transactions through bi-directional verification of parties and verification of primary or secondary information exchanged.

An additional advantage of the present invention is the ability to continue to offer legacy business relationships, legacy products, legacy services and other means that will not reduce the overall security maintained by a system for trusted transactions. Known applications lack this feature, and instead rely on denial of access or authorized access to information. Information need not be restricted, and is preferably freely exchanged to widen the opportunities for transactions with a greater potential number of parties. The present invention is an improvement, in that the elements necessary for generating trusted transactions may be made more flexible, and those elements that are "secret," those elements that will be available at predetermined times, as well as those elements that are made more obscure to unintended parties, increase the overall computational difficulties in defeating a system for trusted transactions.

An additional consequence is improvements in enterprise resource planning and data mining. To the extent that transactions are made unique and may be atomized into data, functions, value-added components and any associated information, the cost of maintaining or referencing stored data, a goal in data mining technologies, can be made more efficient and effective in assisting with an optimized appropriation of resources, individual or corporate. Without such uniqueness, serialization, authentication, verification or identification, particular transaction events cannot be analyzed, manipulated or optimally used to create additional trusted transaction opportunities. Caching technologies are similarly effected by the present invention. The choice about what information should be maintained locally based on identification or authentication of that information available on a network, such as the World Wide Web, enables higher efficiency in sorting and referencing data for repeated use without increased demands on the network.

The ability to serialize individual transactions by particularizing trusted transaction elements between parties is handled more consistently than in known systems. Access is not denied, and rules for access are not pre-determined for goods or services that require exposure, testing or additional information for consummating a transaction. Ease-of-use, maintenance of more human-like and physical world expectations of trust are made more transparent. Identity and authentication risk is

reduced, and confidence is increased. Overall expectations are handled according to the needs of individual parties to any number of transactions. What results from trusted transactions is a more vibrant and competitive marketplace for information, value-added or not. Anonymity and legacy relationships may be maintained, unlike requirements in known systems.

The application of steganographic ciphers enables an "optimized envelope" for securely inserting, detecting, and protecting informational signals, or data, or digital watermarks (predetermined messages) in a given digitized sample stream (e.g., a predetermined carrier signal, such as audio, video, image, multimedia, virtual reality, etc.). As the perceptible qualities of the content stream have a basis as analog waveforms, steganographic ciphering increases the computational difficulty of crypto-analysis and makes unauthorized removal or tampering of the watermark a costly operation. With perceptible damage to a carrier signal a result of such tampering, tampering is more easily observable by parties, including those who are involved in a particular transaction event. Moreover, such tampering enables higher transparency and verification of carrier signals of datum that are marked for secure exchange, even if over unsecure transmission channels. The prior art relies overly on secure transmission channels while ignoring the potential benefits of securing datum (with secure watermarking, scrambling, or chaffing, for instance) over any available transmission channel. Such tampering is also transparent to vendors handling or accepting the information that enables less costly validation of claims made after some event must be confirmed and verified to the satisfaction of transacting parties. These unique features are an improvement over the art.

What differentiates the "digital marketplace" from the physical marketplace is the absence of any scheme that establishes rights and responsibility, or trust, in the authenticity of digitized goods, services or value-added information. For physical products, corporations and governments watermark "goods" and monitor manufacturing capacity and sales to estimate loss from piracy. Reinforcement mechanisms, including legal, electronic, and informational campaigns also exist to better educate consumers. Evidentiary levels of confidence must exist to support claims that are typically competitive between parties to a transaction.

Currently, security parameters may be coded into the actual physical transaction system or instrument. Similar to the security inherent in the randomness of the magnetic strip on most credit cards, these security parameters are designed to be tamper-resistant. Cracking such codes would not present insurmountable barriers to a dedicated effort at cracking a PIN. Access authorization is easily compromised by fraudulent reconstruction of an instrument, such as a credit card. Although storage of the security parameters in volatile, or nonpermanent, memory appears to offer advantages, including higher security required for many transactions, absent this higher level of security, real time authentication becomes a crucial benefit to parties in ensuring the validity of many forms of transactions. Insurance, identity, and purchases of expensive items or services are not generally confidently handled. Use of trusted transactions to process value-added information is unique and beneficial.

Several components may be used for separation of "trusted elements" for a given device or method for ensuring "trust" according to one embodiment of the present invention. First, a general purpose computing device is comprised of a CPU, a memory or storage, input and output devices, and a power supply. A device or card holder decides whether and when to use the device. For additional benefits described herein, personal information or privacy data may be controlled by the user in sample embodiments envisioned, unlike other pre-determinations of data in non-trusted transaction smart cards (e.g., a credit card).

A data owner, who may or may not be the device holder, is provided. Where the device holder and data owner are the same, as contemplated by some embodiments of the present invention, such data as digital certificates, time stamps, Unique IDs of data coming into and out of the device (personal or financial information being a large class of such data), etc. can be authenticated in a humanly-perceptible manner. This may be accomplished by a transducer, or a screen, that can transfer analog-based information of device holder, or be inputted and transmitted by the device holder for secure watermarking, or hashing of data to be exchanged.

A terminal, controlling input and output to and from the device (e.g., phone cards are controlled by the phone service provider's terminals, ATMs are controlled by financial institutions, set-top boxes controlled or owned by entertainment

distribution providers, etc. that may be made physically secure by separate means) or a system that may interact with a device, such as that contemplated in embodiments herein, to enable real time authentication or verification where such checks may fail from time-to-time with existing pre-defined trust arrangements or pre-determined protocols that require inefficient updating by one or both parties. In lieu of a physical visit to a vendor, the present invention anticipates more convenient anonymous updates, in those markets where it is possible to the benefit of both buyers and seller -- both parties have a market demand or need and are able to agree to such arrangements.

Embodiments of the present invention may include a simple Internet browser plug-in, with complementary system software for the provider of "information goods or services," that would identity, verify, authenticate, enable transfer, enable copying or other manipulations of the various primary value-added information and value-added components. Some of the functionality may strictly indicate what, if any, security exists within a particular primary value-added information set. This need not be settled within a system of trust, but be inherently imperceptible to any casual observer or market participant interested in the information or the transaction events that can be observed. Essentially, encouragement of provable differentiation between different classes of primary value-added information (secure, unsecure, legacy, etc.), value-added components (not the primary information but value-adding to the transaction event, and any information concerning market participants (private, history, condition, or financial) is enabled, using simple steganographic ciphers with mapping and transfer functions without compromising the underlying security.

A device issuer controls the operation of the device according to mutually agreed to terms between parties. The device issuer may limit the use or functionality of the device.

For the device hardware manufacturer, fraud may be attempted by the various parties, subcontractors, etc, who are involved in the manufacture of the devices. The device issuer requires protocols that cannot be defeated by typical "rogue engineer" attacks, where security is dependent on an understanding of the methodologies, device, or system design. In fact, the ability to transparently and

provably manufacture secure smart devices may be accomplished with such protocols as digital time stamping (using successive temporally related hashes that seed other hashes to create a universally acceptable means for establishing the time of manufacturer, with time being the universal constant), or digital watermarking (where instead of time, other predetermined data is concatenated with data for provably establishing ownership, over the device). Tampering must be provably perceptibly evident upon tamper detection of the device (as with device used for limiting theft of clothing or physical items in retail stores). Prevention of the rogue engineer problem is not anticipated by known systems.

A software manufacturer usually requires clear specifications or transparency such as open source code, providing the underlying ciphering algorithms and other specifications for analysis. Similar trust issues as with device hardware manufacturing exist. Stega-ciphering the operating system, the simple system or engine for determining authenticity and identification of available data, to prevent memory capture, cloning, write once memory specific to the device holder provide additional benefits of security. A discussion of such is provided in U.S. Patent No. 5,745,569, the disclosure of which is incorporated by reference in its entirety. As well, using transfer functions with associated predetermined keys is also a means for accomplishing confidence and authenticity in transaction. This is described in U.S. Patent Application Serial No. 09/046,627, entitled "Method for Combining Transfer Functions with Predetermined Key Creation," the disclosure of which is incorporated by reference in its entirety.

In general, security requires: fewer splits of trust (poor tying arrangements that may encourage fraud or piracy), better transparency of data (it should be perceptibly apparent, or mathematically, or actuarially possible to observe risks and quantify them to enable security design with a clear understanding of potential threats for each system, method or device), and use of cryptographically strong protocols, where security is both provable and perceptible such that market-driven features are both fundamental at the earliest development and design of appropriate systems and devices, in order to build confidence and trust that is acceptable and transparent to all parties to a transaction.

Application of a steganographic cipher to the operating system or operation of the contemplated systems and devices ensures further security from tampering. Such methods are disclosed in U.S. Patent No. 5,745,569, and offer additional benefits when coupled with the embodiments disclosed herein. System or device operations may be controlled with minimum functionality, objects or executable code. As value-added information is checked for authenticity, decoding any embedded operation objects or code, executing the operation of the system, and deleting the object or code from memory, or randomizing it in memory to avoid capture, would greatly increase the security of both value-added information and the systems or devices intended for manipulation of the value-added information. Alternatively, certain base functions, such as play, record, copy, manipulate, and transfer data, may be problematic. These functions may be atomized into objects that must be first authenticated by the trusted transaction device before they are operable for the given format, or before they provide additional information.

Time of use has traditionally been a typical constraint for securing smart cards and similar devices, but may become ineffective and inconvenient to users. Enabling a smart card to capture or transduce information (even converting analog information or input into secure digitally-sampled representations of the analog information for analysis and authorization, as with a stega-ciphered digital watermark) about the time, location, identity or any number of specific datum greatly enhances smart card and similar device security, trust and confidence. Such benefits over known systems are valuable contemplated with the present invention.

Valuations of trust also enables the described sample embodiment of a trusted transaction system or device to compare private information with financial information, essentially bridging determinations of risk in financial transactions and insurability. Private, or sentimental, information disclosure is more highly sought in determining insurance risk. The ability to pay, and other financial information, are being commoditized. Insofar as the described method and device for such deployment of trusted transaction technology can be assessed for different products and markets, the example of an insurance device could easily be called a trusted transaction privacy/financial information device or card. Users can control what information they disclose given the risk coverage or credit they seek, and providers

being able to decide, with more current and transparent information disclosure possible, what to underwrite or what to finance.

For the authentication or identification device, there is a risk of identity theft to both buyers and sellers, or information that is limited by law. Examples include Medicare-covered drugs, local legal constraints, etc. Risk may be predetermined or limited by a government agency (FDIC, FBI, Social Security, IRS, DMV, Federal Reserve, etc.), a similarly outfitted organization (trust is held in perceived and observable representations of the organization, food stamps, stamps), or an equivalent transaction event enabler (traveler's check provider, medication, etc.). In these cases, systemic risk is limited by enforcement agencies held in trust by a government or body politic. The restrictions are predetermined and dependent on successful authentication or identification of a product, label, or other similar item. Laws may differ between localities and may be dependent on some form of identification, proof of age, or proof of residency. To properly serve local residents becomes a data security issue. This embodiment offers advantages over the art in its flexibility and real time, perceptible authentication properties.

Both the provider and the agency involved may have higher levels of risk, because the nature of the information is characterized by high value, general or universal recognizability, and a genuine threat of fraud. Most people casually accept that $10 and $20 bills are real even if they prove not to be later. Governments try to limit such liability without damaging the overall trust in the currency. As abstractions of value are exchanged, a smart identifying device, instead of value replacement device (predetermined, fixed spending or authorization in a device), is necessary to capture "personal entropy," or information about oneself that can be more closely guarded and less open to theft versus a password or pass phrase. Secrets must differ from identification. The larger body of data to search to discover these secrets act as a higher form of secrecy. These datum may be converted to readable text in some embodiments or maintained in digitally-sampled but humanly perceptible form in other embodiments (favorite restaurant is represented as an actual image of the restaurant, mother's maiden name is actually the voice of an individual's maternal grandparents, highly specialized forms of personal information

that may be dynamically changed or checked quickly and conveniently without undue risk exposure to the system).

For governments and individuals, piracy of identity is the most insidious risk exposure. Identity theft may be curtailed with devices that can transduce, in real time, an iris scan, fingerprint or other biometric and compare securely transmitted results with a secured stored record at the time of initialization. Alternatively, this may be accomplished with an unrelated Unique ID that confirms the identity of the user, and may be created and stored on the device. Because governments are arbiters of trust in markets (their actions in the collective affect trust and confidence in products and markets), these devices are able to alert consumers to potential risk for a given product or service (represented by some ruling or law that is important to convey to the consumer, such as with alcohol, medications, or tobacco). These devices could, at the discretion of the user, indicate related warnings for which the government has an interest in safety. In one embodiment, by checking an actual cigarette carton, or drug packaging, with the enabled device, counterfeit packaging may also be detected. In one embodiment of the present invention, bar code scanners may be "required" to also check for embedded or associated signals indicating authenticity. The devices may also check if supposedly "real" prescription drugs are authentic. Such a check may occur when using the device to communicate with a vendor and check to see if any complaints or problems exist in stored records; again the packaging may be checked for authenticity in cases where counterfeits are high and difficult to check without some form of secure watermarking or perception-based authentication that can be efficiently handled by an enabled device.

According to one embodiment of the present invention, digital content may be distributed through a local content sever, or LCS. In general, the LCS environment is a logical area inside which a set of rules governing content use may be strictly enforced. The exact rules may vary between implementations, but in general, unrestricted access to the content inside the LCS environment is disallowed. The LCS environment has a set of paths, or paths that allow content to enter the domain under different circumstances. The LCS environment also has paths that allow the content to exit the domain.

The act of entering the LCS environment may include a verification of the content (an authentication check). Depending upon the source of the content, such verification may be easy or hard. Invalidatable content may be subjected to a quality degradation. This degradation may be to the content itself, or it may be removal of value-added components. Content that can be validated, but that belongs to a different LCS environment may be excluded. The primary purpose of the validation is to prevent unauthorized, high-quality, sharing of content between environments.

When content leaves the LCS environment, it may be watermarked as belonging to that environment. It is allowed to leave the LCS environment at the quality level at which it was stored (i.e., the quality level determined by the path). The watermark on the exiting content may be both an embedded digital watermark and an attached hash or digital signature (it may also include a secure time stamp). Content cannot return into the environment unless both the watermark and hash can be verified as belonging to this environment. The presence of one or the other is generally sufficient to allow re-entry.

This system may allow a certifiable level of security for high-quality content, and may allow the use of unsecure content at a degraded quality level. The security measures are such that a removal of the watermark constitutes only a partial failure of the system. The "wiped" content may be allowed back into the LCS environment, but only at a degraded quality level, a result of the watermark destruction and subsequent obscurity to the system. Consumers will not be affected to the extent that the unauthorized content has only been degraded, but access has not been denied to the content. Only a complete forgery of a cryptographically-secure watermark will constitute a complete failure of the system. For a discussion on such implementations please see U.S. Patent No. 5,613,004; U.S. Patent No. 5,687,236; U.S. Patent No. 5,745,569; U.S. Patent No. 5,822,432; U.S. Patent No. 5,889,868; U.S. Patent No. 5,905,800, U.S. Patent No. 6,078,664, U.S. Patent Application No. 09/046,627 U.S. Patent Application No. 09/053,628, and U.S. Patent Application No. 09/594,719

Provable security protocols may minimize this risk. Thus, the embedding system that embeds the watermark does not need to be optimized for robustness, only for imperceptibility (important to publishers and consumers alike) and security

(more important to publishers and commercial interests in the content than to consumers). Ideally, as previously disclosed, security preferably does not obscure the content, nor prevent market participants from accessing information contained therein, and for the longer term, developing trust or creating relationships.

The system can flexibly support "robust" watermarks as a method for screening content to speed processing. Final validation, however, is relied upon the fragile, secure watermark and its hash or digital signature (a secure time stamp may also be incorporated).

The LCS provides storage for content, authentication of content, enforcement of export rules, and watermarking and hashing of exported content. Stored content may be on an accessible rewritable medium, but is preferably stored as ciphertext (encrypted or scrambled), not plain text, to prevent system-level extraction of the content. This is in contrast to known systems, which affix or otherwise attach meta-data to the content for access control by the variously proposed systems.

The LCS may be able to receive content from a secure electronic content distributor, or SECD, and may be able to authenticate content received via any of the plurality of implemented paths. The LCS may monitor and enforce any rules that accompany received content, such as number of available copies. Finally, unless being transmitted to a satellite unit, the LCS may watermark all exported material and supply a hash made from the Unique ID and the content characteristics (so as to be maintained perceptually within the information and increase the level of security of the watermark).

The satellite unit enables the content to be usable apart from the LCS. The satellite unit is partially within the LCS environment. A protocol may exist for the satellite unit and LCS to authenticate any path made between them. This path may have various levels of confidence set by the level of security between the satellite unit and LCS, and determinable by a certification authority or its equivalent, such as an authorized site for the content. The transfer of content from the satellite unit to the LCS without watermarking may be allowed. However, all content leaving the satellite unit is preferably watermarked. The satellite unit watermark may contain a hash generated from the satellite unit Unique ID and the content characteristics. If the content came from a LCS, the satellite unit may also add the hash received from

the LCS to the watermark. The LCS and satellite unit watermarking procedures do not need to be the same. However, the LCS is preferably able to read the satellite unit watermarks for all different types of satellite units with which it can connect. The satellite unit does not need to be able to read any LCS watermarks. Each LCS and satellite unit preferably has a separate Unique ID.

Referring to **Fig. 2**, a schematic of a local content server environment according to one embodiment of the present invention is provided. LCS 202 may be a software device running on a general purpose computing device, such as a personal computer (including, in general, a central processing unit, an input, an output, a memory, and a power supply). LCS 202 may include local content server domain 204, rewritable media 206 (such as a hard disk drive, a CD-R/W, etc), and read-only media 208 (such as a CD-ROM). LCS 202 may communicate with at least one satellite unit 210 via an interface.

In one embodiment, LCS 202 may have a Unique ID. Similarly, in one embodiment, satellite unit 210 may have a Unique ID.

LCS 202 may communicate with SECD 212 via a network, including a local area network, a wide area network, an intranet, and the Internet. This communication may also be established by a telephone link, a cable connection, a satellite connection, a wireless connection, etc.

In one embodiment, a single LCS 202 may interface with more than one SECD 212.

A plurality of paths 220, 222, 224, 226, 228, 230, 232, and 234 may exist among LCS 202, SECD 212, Satellite unit 210, LCS domain 204, rewritable media 206, and read-only media 208. Each will be discussed in greater detail, below.

Digital content may be securely distributed to LCS 202 from SECD via path 220. The content may be secured during the transmission using one or more security protocols (e.g., encryption or scrambling of the content). In one embodiment, if LCS 202 interfaces with multiple SECDs 212, each path may use a different security protocol.

The security protocol may use an asymmetric cryptographic system. An example of such a system includes a public key cryptography system. The private and public key pairs allow LCS 202 to authenticate and accept the received content.

Referring to **Fig. 3**, a flowchart depicting an example of an authentication by LCS 202 is provided. In step 302, the user connects to the SECD, makes a selection, and completes a sale.

In step 304, the LCS provides its public key to the SECD.

In step 306, the SECD uses the LCS public key to initiate transmission security.

In step 308, the SECD transmits the secured digital content to the LCS.

In step 310, the LCS receives the digital content, authenticates that the digital content was unchanged during transmission, and unpacks it from its security wrapper (that may include a secured transmission line, such as SSL). In one embodiment, the digital content may be authenticated by a watermark and hash check. If the content can be authenticated, the content is accepted into the LCS domain. If the content cannot be authenticated, it is rejected.

Referring again to **Fig. 2**, path 222 connects LCS domain 204 with rewritable media 206. Referring to **Fig. 4**, a flowchart depicting the process for content entering LCS domain 204 from rewritable media 206 is provided. In step 402, the content is provided. In step 404, the content is checked for the presence of a watermark, such as a watermark for the particular LCS. If there is not a watermark, in step 406, the content is degraded to Low Quality and, in step 408, the content is stored in the LCS domain.

If, in step 404, a watermark is present, in step 410, the watermark is checked to determine if it matches the LCS. This may be achieved by a hash. If the watermark is verified, in step 408, the content is stored in the LCS. If the hash does not match, the content is rejected.

Referring again to **Fig. 2**, LCS domain 204 may export content to any receiver (other than satellite unit 210) through path 224. This may include copying content to a rewritable media, creating a read-only media, rendering the content for use (e.g., playing, viewing, etc), etc.

Referring to **Fig. 5**, a flowchart depicting the process for content leaving LCS domain 204 is provided. In step 502, the content is retrieved from storage within the LCS. In step 504, the content is embedded with a watermark. In one embodiment, the watermark may be unique to the particular LCS, as determined by

the LCS Unique ID. The watermark may contain a hash that is created from the combination of the content characteristics (such as signal features, etc.) and the Unique ID. The watermark may optionally contain other data, such as a timestamp, a number of allowable copies, etc. This would be described as parameters of use, usage data, etc. which could be referenced when content is exported. If the export is to a storage medium, the LCS optionally can add a second hash to the file, external to the content, which can be used for further authentication. For security purposes, in one embodiment, the external hash may be created in a different manner from the embedded, watermark hash.

In step 506, the content is output from the LCS to the receiver.

Referring again to **Fig. 2**, path 226 connects LCS domain 204 with read-only media 208. Referring to **Fig. 6**, a flowchart depicting the process for content entering LCS domain 204 from read-only media 208 is provided. In step 602, the content is provided. In step 604, the content is checked for the presence of a watermark, such as a watermark for the particular LCS. If there is no watermark, a check is made in step 610 to see if the originality of the content can be determined. An example of such includes a media-based identifier that identifies the content as original.

If the content can be verified as an original, in step 608, it is stored as High Quality in the LCS domain. If the originality cannot be verified, in step 610, the quality is degraded to Standard Quality, and, in step 608, the content is stored in the LCS domain.

If a watermark is identified in step 604, in step 612, the hash is checked to verify that the content matches this LCS. If it matches, in step 608, the content is stored in LCS domain at High Quality. If it does not match, in step 614, the content is rejected.

Referring again to **Fig. 2**, path 228 connects LCS 202 with satellite unit 210. Referring to **Fig. 7**, a flowchart depicting the process for content entering LCS 202 from satellite unit 210 is provided. In step 702, the content may be watermarked before it is transmitted to the LCS. In step 704, the content is transmitted to the LCS.

In step 706, the content is checked by the LCS. This may include checking the LCS hash. If the hash matches, in step 708, the content is stored in the LCS domain as High Quality. If there is no hash, in step 710, the content is degraded to Low Quality, and in step 708, the content is stored in the LCS domain. If the hash does not match, in step 712, the content is rejected.

Referring again to **Fig. 2**, path 230 connects LCS 202 with satellite unit 210. Referring to **Fig. 8**, a flowchart depicting the process for exporting data from the LCS 202 to satellite unit 210 is provided. In step 802, the content is retrieved from storage within the LCS. In step 804, the security of the path between the LCS and the satellite unit is verified. Once the security is verified, in step 806, the content is exported to the satellite unit without a watermark.

If the security of the path cannot be verified, the export process mirrors that of an export to a receiver, depicted in **Fig. 5**.

Referring again to **Fig. 2**, path 232 is a path for content to be stored in satellite unit 210. In one embodiment, all content may be allowed to be imported into satellite unit 210, but may be automatically degraded to Low Quality when it is stored.

Path 234 is an export path for content rendered by satellite unit 210. In one embodiment, this content may be marked with a satellite unit watermark that contains a hash from the satellite unit Unique ID and any hash that is associated with the content from an LCS .

It should be noted that a hash function may be converted into a digital signature by performing a hash and encrypting the result of the hash. The uniqueness of the hash can vary with the hash function, while the digital signature adds a layer of confidence to the integrity of the data.

Other types of encryption, including transfer functions, may also be used.

Referring to **Fig. 9**, a flowchart of a method for trusted transactions according to one embodiment of the present invention is provided. In step 902, value-added information, or its tangible equivalent, is provided. This may be provided by a user that wishes to verify the value-added information.

In step 904, the perceptible data for verification may be maintained by a vendor or provider, and may be updated by a public-key secure digital watermark in

the observable packaging (if applicable). In those cases where security must be high, real time, or simply faster, key generation or signature generation functions may be enabled with embodiments of the present invention.

In step 906, the user provides a public key based on the identify held in the device to enable an authentication check.

In step 908, a response may be sent to the user.

Steps 906 and 908 may be repeated with further prompting for higher levels of authentication, or for additional checks. If the remote location provides the confirmation, or if a certification authority is involved, the response may be sent via secure transmission lines (e.g., encrypted transmission that can only be decrypted with the user's device and access to the user's stored private key). Alternatively, information may not need to be sent in a secure manner and may be checked upon delivery to the device to limit any remote communications breaches by unintended third parties.

Referring to **Fig. 10**, a device for trusted transactions according to one embodiment of the present invention is provided. Device 1000 may include steganographic cipher 1002. Steganographic cipher 1002 may be governed by at least the following elements: (1) a predetermined message; (2) a predetermined key/key pair; and (3) a predetermined carrier signal (image data, so images will be the primary data represented and ciphered).

Transducer 1004 may be provided. Transducer 1004 may include a charged coupled device (CCD), a personal entropy capture device (e.g., a retinal scanner, a thumbprint scanner, etc.), a touch pad (e.g., a pad for receiving a signature), an image capture device, a bar code reader, a magnetic card reader, etc. Transducer 904 receives the data in a physical format and converts it to an analog or digital format.

In one embodiment, the data from transducer 1004 may be marked with a timestamp for time-critical input.

Analog/digital converter 1006 may be provided. A/D converter 1004 may be used to convert analog information from transducer 1004 into predetermined digital format. In one embodiment, signatures may be converted in one format, images that

are captured in another format, and fingerprint/iris scans may be converted in another format.

A memory may be provided. The memory may include both volatile memory, and re-writable memory, such as DataSlim™.

A volatile device may be provided, such as a one time pad (private key of card holder/user), a one time memory or floating in the volatile memory to evade capture (stega-cipher computer code). This may be provided in a tamperproof casing.

Device 1000 may also include output 1020. Output 1020 may be any suitable output, including a connection port, a wireless port, a radio transmitter, etc. Before information is output from device 1000, it may be encrypted. In one embodiment, the information may be digitally watermarked. In another embodiment, the information may be digitally signed. In another embodiment, the information is not encrypted, and instead is transmitted over a secure transmission channel. Number generator 1008 may be provided. Number generator may be a random number generator, or it may be a pseudo-random number generator.

In addition, the device may include a controller, a power source, and an input and an output.

Information may be converted into a humanly perceptible form (chemical/electrical/magnetic such as a humanly visible chemical test result, as with a pregnancy tests, an EKC, an MRI or CatScan image, are all converted into "humanly perceptible form for "human" analysis) prior to authorization of a transaction/decision event.

EXAMPLES

In order to better understand the present invention, several examples are provided. These example do not limit the present invention in any way, and are intended to illustrate embodiments of the present invention.

1. Smart Telecommunications

At present, large volumes of commerce and commerce-related activities are performed using telephone connections. Authentication of identity is an ongoing concern in such transactions. Present technology allows the verification of the

origin of a landline phone call (POT), but offers no assurances as to the identity of the user. Furthermore, simple identification of the origin of the call is only useful insofar as that phone number can be used to index a database of callers. The present invention allows for bi-directional verification of identity during a phone call, with the option of partial or full concealment of identity.

A consumer may wish to make a purchase on the phone. Presently, the consumer's identity is established by the seller using personal information from the consumer, such as a credit card number, an address, a phone number, etc. However, all of this information may be known by an imposter. A smart phone transmits identity information (perhaps embedded as a watermark in the audio connection), in response to a query from the seller. The receiver verifies the buyer's identity with a certification authority. Furthermore, the consumer may also verify the authenticity of the seller's identity at the same time, by the same method. The consumer may choose not to respond to certain queries in real time.

The smart phone may require a level of identity disclosure before it accepts an incoming call. For instance, telemarketers may be required to reveal the name of their company before the call is accepted by the smart phone. Consumers may protect themselves from fraudulent sellers by requiring such identification. Further, legitimate sellers may be assured that their customers know that they are legitimate. The certification authority assures the consumer and seller that they are receiving authentic identifications.

2. Equity Programs As A Value-added Component

Another embodiment of the present invention relates to methods and means of payment includes a novel means for encouraging alignment of buyer and seller interests. Similar to cooperatives, membership programs (in proprietary form, co-branded with a financial institution, or implemented as a specialty device that can handle these equity transactions) may be enhanced to offer buyers the opportunity to purchase options in equity of the seller's company or related institution. Instead of being given cash or points, at some fixed point in time, consumers and sellers may be provided with the opportunity to purchase equity as available on some public or private market or exchange.

These options may be built into the functionality of the actual transaction device and may be coupled with both trusted transactions or general transaction systems. Settlement of the option may be based on any known option pricing mechanism (such as the well-known Black-Scholes model) and predetermination of terms for settlement and conversion of the option. This approach incentivizes and encourages clearer alignment of all market participants in the value and condition of the equity of the entity with which transactions are being handled or negotiated. Independent certification authorities, or infomediaries that are able to ensure or verify a transaction or related information, may be used to ensure that such equity programs can be trusted. Any relevant disclosures concerning legal or financial restrictions are simply additional value-added components for consideration.

3. More security - body movements for entropy and pharmaceutical use control

A related embodiment according to another embodiment of the present invention includes an interface for detection of body movements (eye movements, blinks, voice pass phrases, etc.). These movements may include predetermined sequences of movements that may be ciphered in a manner similar to encrypting ASCII pass phrases. This is a novel implementation of human movement in generating symmetric or asymmetric cryptographic keys. The transducer may include any number of means of capturing human-based body movements in real time for instantaneous verification of an authorized user. Moreover, unlike simple biometrics, a series of body movements (similar to the act of signing in writing, but likely to be more difficult to capture for unauthorized misuse -- a signature, like a fingerprint, is able to be observed and copied without permission or knowledge of the signature author) is difficult to copy.

The movements or similar biological entropy (transduced from biomedical, bioengineered, biochemical or biophysical information that may be made perceptible and encrypted or securely watermarked for later comparison or real time verification) may be captured by a transducer of analog signals and converted into digital binary information used for comparison with any number of stored corresponding instructions or messages to be decrypted. These signals may be multidimensional (2D, 3D, 4D- with a time component, etc.) to increase the information space and make discovery of hidden secrets more computationally

difficult. Images, medical or human-condition based, audio signals, video, virtual reality, multimedia, etc. all provide rich media information in which to enhance the security of any embodiment contemplated by the present invention. Combinations of multidimensional media for varying ciphering options as well as steganographic embedding are also contemplated as a means for furthering ensuring computational complexity to any unauthorized user. Steganographic-mapping (watermarking) or transfer functions (scrambling or "chaffing") may be combined with encryption ciphers as a means for making each unique implementation or tangible device -- serialization or personalization of a method for engaging in trusted transactions, high risk, information-intensive or sensitive decision (military use, security use, restricted government use, privacy use, or any number similar commercial or noncommercial decision or transaction events).

Additional embodiments include actual control over the use or access to pharmaceuticals based on medical risk, condition or personalized advice to the user. Tangible methods for transfer of chemical, biological or physical agents intended for medical use or individualized control based on third party conditions (legal, medical, governmental, etc.) are governed by manipulation of the apparatus, device or system used to introduce foreign agents (informational, intangible or tangible) into patients (the intended, authorized or verified user).

Highly secure and artificial environments, such as aircraft flying simulations or visual financial trading information, may be representative of more risk to owners of actual tangible planes or tangible assets related to any financial information. Recognition of a digitized iris does not enable movement based confirmation of future secrets (the movements) that may be changed, destroyed or updated to ensure consistent or higher degrees of security maintenance. For some body movements, it may be possible to maintain better security than with written information. In other words, certain body movements may be prevented, or made difficult to perform even under rigorous demand by unauthorized agents. Blinking or other facial movements may be made impossible to verify the real time identity of the user. This adds a layer of security and increases the difficulty of defeating a cipher or a series of related ciphers (encryption-based or steganographically-based, where the digitized signal has humanly-perceptible fidelity or characteristics) depending on access or

sensitivity of information. It also maybe psychologically or human-rule driven. Certain humanly observable body movements, or detectable "telemetry-type" data (brain activity, heart beat, pulse, or any other medically observable information) may be either unique to an individual or simply general to certain behavior. This data may be important to use as a means of preventing poor decision-making, or requiring higher diligence before transacting or executing a given operation. At the least, the movements are a means for predetermining and assisting the generation of a binary key or seeding the generation of a cryptographic key, message or signature.

Any particular instance may be successively stored in subsets of any primary value information or value-added components (single key or key pair associated with a single message or signature to further serialize data that may have steganographic capacity for imperceptible embedding in the carrier signal, primary or value-added components data). The operation may be highly demanding, or may require human-based or driven or initiated decisions. The instructor, or the user, may have predetermined the conditions that indicate confidence or lack thereof at the time of the verification or authentication of the user. This may be for security reasons, or simply risk management, as information is increasingly processed at higher speeds and may require greater care in ensuring information data integrity. As well, humanly-observable (and convertible into binary data for deciphering) movements enable a form of bridging analog, human trust with digital or mathematically provable, actuarially, statistically, deterministically known or predictable measures of risk and trust. This novel feature is an additional benefit over the prior art and ensures future human-like characteristics in "digital" (underlying, "measurable" or "estimable" data integrity, authentication and confidence), electronic (analog transducers and transmitters), or binary transaction systems. Further security or serialization of transaction event information (human movement or observable condition used for secret key or equivalent generation) enable additional forms of trusted transactions.

Additional security may be assured with temporal-based limits on human body movement or biologically observable human condition (by use of a medical or human directed transducer). Interlocking keys and messages with blind signatures, or onion routing transmission techniques to obscure the identity of the user, are

further enhancements that may guarantee a high level of privacy to the user of the system or device. Information formats may be encrypted or have stored primary or value-added component information that has to arrive to the user without any digitally evident tampering for the user to make the best possible decision regarding the observed information.

Unlike the prior art, embodiments of the present invention consider the perceptibility of information to bridge human trust and confidence with cryptographic or "mathematical" measures or estimates of "security," "data integrity" or "trust." This is novel to the art of data security and secured transaction or transmission technologies.

4. Algorithmic Information Theory (AIT) for additional security

By implementing predetermined indications of mathematically provable randomness, the ability to discover secrets and human choice, based on unprovability or incompleteness, as discussed and is well-known in the art as originating with Godel (incompleteness theorem) and Turing (halting problem, uncomputability). Chaitin "discovered" randomness, stating essentially that randomness can be described mathematically, and thus differentiations between discrete and infinite randomness are logically observable. Because truth is relative in a quantum mechanical sense, degrees of credibility concern the level of trust that may be offered in any trusted transaction system. While the primary value that concerns us is information, the ability to describe programming size complexity (that is optimized functional data) enables self-limiting software to be programmed. To the extent that trusted transactions can never be physically perfect operations, uniqueness of information, as both data and code, is particularly important to providing higher security when computational cost and bandwidth is extraordinarily cheap.

Essentially, choice over answers to questions that cannot be characterized as "True" or "False," such as "This statement is false," have inherent randomness and are thus ripe for paradoxical response. More intricate paradoxes, Berry's Paradox, Turing's halting problem, as well as Chaitin's definition of "randomness," are sure to enable predictable infinite and finite (discrete) randomness with which to seed and cryptographic secret or generation of a symmetric, asymmetric key or digital

signature. Human perception as a means for enabling analog trust may be made inherently more secure by choosing responses to paradoxes that have no computable value. That Chaitin can describe "randomness" with logically structured instructions for the halting problem, in LISP or C programming languages, including the computer programming language of Mathematica, enabled the development of a randomness constant.

The equations of randomness may be implemented in software and offer a unique and novel means for further securing the generation of cryptographic or steganographic seeds, secrets, keys or messages. Of course, differences between any of these information elements as to the means for securing or authenticating data would enable flexible architectures combining various ciphers and methods for arriving at a rule for validation, authenticity, data integrity, confidence or enabling any subsequent manipulation of the associated data (primary value-added or value-added components).

5. Entertainment media exchange

According to one embodiment of the present invention, the device may be used for the exchange of entertainment media. This may include audio, video, multimedia, etc. In such an exchange, the perceived risk of value-added information piracy is relatively high for the seller or provider, while the perceived risk is relatively low for the purchaser. The obvious risk is that all potential "consumers" of the media access and copy the entertainment media for free. For music or video, or similar entertainment good, according the present invention provides the following structure may be used.

a) Fragile watermark structure

The fragile watermark, according to one embodiment of the present invention, can actually hold an entire value-added component, encoded in the least significant bit (LSB) of each 16-bit sample. This gives a data rate of 88200 bits per second in a stereo CD file, or a capacity of 1.89 M in a 3 minute song. This is an immense capacity relative to the expected size of the value-added component (100 - 200 K).

The fragile watermark is preferably bound to a specific copy (Unique ID) of a specific song (Unique ID), so that it cannot be transferred to other songs. This binding can be achieved through use of a hash in the following sequence:

(1) A block of value-added component is encoded into a block of samples.

(2) A hash of the value-added component block and a random number seeded by the owner's identity (Device or system Unique ID) is generated and encoded into the subsequent block of samples.

(3) A hash of the first two blocks of samples and a random number seeded by the owner's identity is generated and encoded into a third block of samples.

(4) Repeat steps 1-3 as necessary.

Each value-added component block may have the following structure:

```
{
    long  BlockIdentifier;     //A code for the type of block
    long  BlockLength;         //The length of the block
    ....                       //Block data of a length matching
BlockLength
    char  IdentityHash[hashSize];
    char  InsertionHash[hashSize];
}
```

An application can read the block identifier and determine if it recognizes the block type. If it does not recognize the block type, it can use the BlockLength to skip this block.

Certain Block Types are required to be present if the value-added component is to be accepted. These may include and identity block and a value-added component Hash block. The Block Data may or may not be encrypted, depending on whether the data is transfer-restricted (value-adding) or simply informative. For instance, user-added value-added component data would not need to be encrypted. The BlockIdentifier would indicate whether the block data was encrypted or not.

b) Robust open watermark

This is the mark that may indicate non-legacy content. In one embodiment, there may be two possible settings. "1" indicates non-legacy content that must be accompanied by a authenticable value-added component for entry into the domain (e.g., EMD or Electronic Media Distribution media content). "0", on the other hand, indicates non-legacy media that was distributed in a pre-packaged form (e.g., CDs, DVDs, game software, etc.). "0" content may or may not have a value-added component. "0" content may only be admitted from a read-only medium in its original file format (e.g., a "0" CD may only be admitted if it is present on a Red Book CD Specification medium).

c) Robust forensic watermark

This watermark may not be accessible to the consumer in any way. It may be secured by a symmetric key held only by the seller (or an asymmetric key pair that may be desired for some embodiments). A transaction ID may be embedded at the time of purchase with a hash matching the symmetric key (or key pair). The watermark may then be embedded using a very low density insertion mask (< 10 %), making it very difficult to find without the symmetric key. Retrieval of this watermark is not limited by real-time/low cost constraints. The recovery will only be attempted on pirated material. A recovery time of 2 hours on a 400 MHz PC is reasonable.

6. Additional parameters for value-adding components

Physical shipment of packaged goods or services (value-added information) is anticipated as being a potential option to consumers or purchasers as well as sellers and providers. That the value-adding information may be packaged or represented tangibly does not obviate the need for trusted transactions to ensure payment and the appropriate division of rights and responsibilities for various goods (a DVD for music or video), services (smart credit card or insurance card) or markets (trusted telephone system, government identification schemes). This type of transaction represents additional benefits over embodiments in the existing art -- on-demand trusted transactions and physical manufacture/delivery of goods is enabled, without risk to the overall system and its value-added information security. This amounts essentially to serializing or personalizing, depending on the

perspective in the transaction, each and every transaction, while building trusted transactions for the benefit of the marketplace for goods services and information.

7. Financial Or Insurance Device

The present invention enables systems and supported devices that are useful in situations where parties need to have pre-defined limits to risk exposure, such as an insurance policy or a claim. These systems are generally characterized by an emphasis on transmission and data security, which reduces the perceived risk of the insurer (a seller of risk coverage for pre-determined events). To the extent that insurance takes into account the history and existing condition of an asset, a measure of context or structure (tangible as well as intangible) to be covered, as well as an economically-based replacement value (though to confuse matters, there are also issues concerning such items as after market versus brand new, brand versus generic, etc.), there exist differences with more transparent financial devices. Financial devices (essentially a "credit agreement" or credit facility based on an imprecise estimate of condition but also experience or trust) rely on the ability, perceived or actuarially observable, to repay credit extended on behalf of the device holder. Whereas financial or credit history is transparent in many cases, private information about an individual's history or condition are perceived to be have higher implicit value to the user. Financial devices and insurance devices converge at those points where privacy or personal information are equivalent with financial or credit information. Both types of risk have differing requirements for updating or adjustment over the course of use of a particular line of credit or insurance policy.

Cars may be embedded with telemetry sensors to determine the real time condition of various components, such as the frame, engine, brakes, or any combination of components mutually deemed to justify such monitoring. Alternatively, a smart card-like device equipped with a transducer may be used to "capture" images of items that are packed (for travel insurance purposes), insurable items in a residence (for homeowner's insurance purposes), etc. Any image captured may be securely watermarked by the device and then exported to an insurance provider via a transmission line (an ATM, a wireless connection such as a mobile phone, a PC modem connection, etc.). An insurance provider may offer such

services at auto service/repair facilities, airports, etc. with a mutual reduction in claims costs and adjustments costs.

Medical information may similarly be digitally stored, securely watermarked, and time-stamped (for any perceptible data stored, such as images or voice) for reference to an individual's health. based on varying levels of access to stored information, which may be distributed among different physicians or handled by a central medical information infomediary. The secured image may be sent to an insurance provider as a secured image (both the device and storage facility may independently verify the security or tamperproofing of the perceptibly represented information). The doctor, patient, health care provider, government agencies can all have varying degrees of access that can be made transparent to the patient. This is an inherent benefit over the prior art in that the patient can see those records that are then watermarked and securely stored.

Additionally, the present invention provides the novel feature of enabling the same information, at the request or demand of the patient, to be sent to a personal or secure storage "space," so that patients may have more accessibility and control over their own medical records and medical conditions. In one embodiment, the information may be provided as digitized bits. In another embodiment, the data may be provided in a tangible form.

The information may be stored as tangible records or intangible, bit-represented records. Doctors may use tamperproofed signals (watermarked audio, image, video, virtual reality, any humanly-perceptible signal) and records that are perceptible to lower insurance costs and potential liability. The prior art ignores the mutual benefits afforded by bi-directional information exchange (that can be tamperproofed with secure watermarking) and transparency in creating opportunities for trusted transactions.

Additional data, such as the transaction information that may be evidenced on a credit card bill or statement, may also be automatically associated with the stored image(s) for later use. In one embodiment, the user may send the same secured data to a private data storage facility, or create personalized records, which may serve as a secondary set of records against which other data sent to the insurance or financial provider may be verified or validated. According to another

embodiment of the present invention, authorized mechanics, physicians, and pharmacists, may add to, but not access or manipulate, previously stored data. These individuals may also be bound by rules for establishing the history and condition of any person or physical good that is being underwritten or financed.

The present invention provides certification authorities the ability to determine the authenticity of data. In cases where public-key steganography or cryptosystems are preferred, the embodiments extend to those implementations as well. Moreover, they enable secure transmission capabilities over unsecured data transmission lines.

Referring to **Fig. 11**, a personal information device according to one embodiment of the present invention is provided. Personal information device (PID) 1102 may be used with financial institutions, insurance companies, etc.

In one embodiment, PID 1102 may be smart card; that is, a device that resembles a credit card, but includes a processor, a power supply, a memory, and an input and output device. In another embodiment, PID 1102 may be a card including a magnetic strip.

PID 1102 preferably has a Unique ID. In one embodiment, the Unique ID of PID 1102 may be a policy number, a social security number, etc.

PID 1102 may receive information from several sources. In one embodiment, telemetry data 1104 may be input to PID 1102. Perceptible data 1106, such as images, photos, etc. may be input to PID 1102. In still another embodiment, associated data, such as purchase receipts, descriptions, serial numbers, registrations, etc., which may be value-adding components, may be input to PID 1102.

PID 1102 may provide output data 1110 to a variety of entities. In one embodiment, output data 1110 may be provided to company 1112 and to storage 1114. Company 1112 may include any organization the may receive output data 1110, including an insurance company, a financial institution, etc. Storage 1114 may include any personal use for output data 1110, including a private data storage such as a fixed storage media, paper records, etc. Company 1112 and storage 1114 may receive output data 1110 in different formats. In one embodiment, output data 1110 is provided according to predetermined parameters for the entity.

Output data 1110 may be watermarked, or it may be time stamped, or it may include both. Other types of encryption are provided.

In general, output data 1110 is preferably provided to the entity via a secure communication link. Transmission of output data 1110 may be controlled by the entity (e.g., company 1112 or storage 1114) or by the user.

8. Authentication Device

According to another embodiment of the present invention, an authentication device may be provided. Referring to **Fig. 12**, authentication device 1202 may be a credit-card sized "smart card," including a processor, a power supply, a memory, and an input and output device. In another embodiment, authentication device 1202 may be a palm sized computing device.

A variety of input devices may be provided. In one embodiment, a bar code scanner may be used. In another embodiment, a keypad may be used. Other input devices may be used as necessary.

In one embodiment, authentication device 1202 may include a display, such as a LCD screen. Other display technologies are within the contemplation of the present invention.

In one embodiment, authentication device 1202 may be a government-issued device.

Anonymous authentication 1204 may be provided. Anonymous authentication 1204 may be used to authenticate a product, a medicine, a label, etc. Anonymous authentication 1204 communicates with authentication device 1202 to authenticate the item in question. In one embodiment, authentication device 1202 may display relevant information, such as known warnings, recommended dosages, etc. regarding the item in question.

In another embodiment, image capture device 1206 may be provided. Image capture device 1206 may include a digital camera, a scanner, etc. In one embodiment, image capture device 1206 may time stamp the image as it is captured.

Identity exchange 1208 may be provided. Identity exchange 1208 includes a Unique ID that may be authenticated or modified by the user. In one embodiment, in order to verify the identity of an individual, additional independent identify

verification may be required in addition to identity exchange 1208. This is because authentication device 1202 may be stolen, borrowed, etc.

Certification authority 1210 may be provided. Certification authority may be bound by federal, state, and local laws. In addition, private restrictions may apply to certification authority 1210.

In one embodiment, certification authority may be further bound by geographical (e.g., location) or age basis (e.g., date of birth, age, etc.) to verify.

Referring to **Fig. 13**, a method of use for an authentication device is provided. In step 1302, a user locates information to be authenticated. This may include a variety of information. The information is then entered into the authentication device.

In step 1304, perceptible data is marked with a public key secure watermark. In one embodiment, this may be done in real time.

In step 1306, the user provides a public key to initiate the authentication.

In step 1308, a response is sent from the certification authority, or additional prompts for higher access levels are provided.

In one embodiment, transmissions between any elements may be over a secure communication link, including SSL or similar transmission exchange.

In another embodiment of the present invention, an authentication device may comprise a Internet web browser. For example, the authentication device may be a "plug in" for a web browser. Such a authentication device may be used to verify, or authenticate, items on web pages. For instance, according to one embodiment of the present invention, the authentication device may be used to verify that an Internet bank that displays the FDIC logo is authorized to display this logo. In one embodiment, real time verification will allow a user to verify such, and govern transactions accordingly.

It will be evident to those of ordinary skill in the art that the above-described modes and embodiments of the present invention, while they disclose useful aspects of the present invention and its advantages, are illustrative and exemplary only, and do not describe or delimit the spirit and scope of the present invention, which are limited only by the claims that follow below.

I CLAIM:

1. A method for trusted transactions, comprising:

establishing an agreement to exchange digitally-sampled information between a first and a second party;

exchanging the digitally-sampled information between the first and the second party; and

approving the digitally-sampled information using an approval element selected from the group consisting of a predetermined key, a predetermined message, and a predetermined cipher, the step of approving the digitally-sampled information using an approval element consisting of a step selected from the group consisting of verifying the digitally-sampled information with the approval element, authenticating the digitally-sampled information with the approval element, and authorizing the digitally-sampled information with the approval element.

2. The method of claim 1, wherein the step of approving the digitally-sampled information precedes the step of exchanging digitally-sampled information.

3. The method of claim 1, wherein the step of approving the digitally-sampled information comprises:

transmitting a first party approval element from the first party to the second party; and

transmitting a second party approval element from the second party to the first party.

4. The method of claim 3, wherein the steps of transmitting the first party approval element and transmitting the second party approval element occur substantially simultaneously.

5. The method of claim 3, wherein the first party approval element and the second party approval element are symmetric.

6. The method of claim 3, wherein the first party approval element and the second party approval element are asymmetric.

7. The method of claim 1, wherein the approving step is accomplished using predetermined key pairs.

8. The method of claim 7, wherein the predetermined key pairs are created by a cipher selected from the group consisting of steganographic and cryptographic ciphers.

9. The method of claim 1, wherein the predetermined cipher is selected from the group consisting of a steganographic cipher and a cryptographic cipher.

10. The method of claim 1, wherein the predetermined message is selected from the group consisting of a unique identification, a unique time, data associated with a predetermined information function, and combinations thereof.

11. The method of claim 1, wherein the predetermined message has value independent from at least one primary value-adding component.

12. The method of claim 1, wherein the predetermined message contains at least one value-adding component.

13. The method of claim 1, wherein the step of approving the digitally-sampled information comprises:

verifying the digitally-sampled information with the approval element.

14. The method of claim 1, wherein the step of approving the digitally-sampled information comprises:

authenticating the digitally-sampled information with the approval element.

15. The method of claim 1, wherein the step of approving the digitally-sampled information comprises:

authorizing the digitally-sampled information with the approval element.

16. The method of claim 1, further comprising:

entering into a security arrangement based on the exchange.

17. The method of claim 16, wherein the security arrangement is a non-cash right.

18. The method of claim 16, wherein the security arrangement is an option for a non-cash right.

19. The method of claim 16, wherein the security arrangement is an equity purchase right.

20. A method for conducting a trusted transaction between two of a plurality of parties who have reached an agreement to transact, comprising:

establishing a secure transmission channel between the two parties;

approving an identity of at least one of the two parties;

determining an amount of value-added information to be exchanged between the parties, the value-added information comprising a plurality of value-adding components;

verifying the agreement to transact; and

transmitting the value-added information.

21. The method of claim 20, wherein the step of approving an identity of at least one of the two parties comprises:

at least one of the parties verifying at least one value-adding component.

22. The method of claim 20, wherein the step of approving an identity of at least one of the two parties comprises:

at least one of the parties authorizing at least one value-adding component.

23. The method of claim 20, wherein the step of approving an identity of at least one of the two parties comprises:

at least one of the parties authenticating at least one value-adding component.

24. The method of claim 20, wherein the step of establishing a secure transmission channel between two of a plurality of parties comprises:

exchanging data between the two parties;

selecting a pre-determined key to exchange over the secure transmission channel; and

securing the transmission channel by at least one of a password, a pass phrase entry, a query to a user, and real-time biometric data transfer.

25. The method of claim 20, wherein the step of approving an identity of at least one of the two parties comprises:

exchanging a value-adding component for each party to the other party.

26. The method of claim 20, wherein the step of approving an identity of at least one of the two parties comprises:

at least one of the parties independently verifying a value-adding component of the other party.

27. The method of claim 20, wherein a bandwidth of the primary value-added information comprises a description including at least one of a bandwidth requirement for transmission, a bandwidth requirement for storage, and a bandwidth requirement for playback.

28. The method of claim 20, wherein at least one term for the exchange of primary value-added information is negotiated between parties, the terms selected from the group consisting of an offer, an acceptance, and consideration.

29. The method of claim 28, wherein the at least one term changes in real time.

30. The method of claim 28, wherein access to the at least one term is restricted by at least one of a pass phrase, a password, a correct answer to a query, a real time authentication with a biometric, a real time authentication with personal entropy information, real time telemetry data, and access to additional transaction records.

31. The method of claim 28, wherein the at least one term is referenced by a subsequent transaction.

32. The method of claim 28, wherein the at least one term is access restricted by a provider of at least one value-adding component.

33. The method of claim 28, wherein the at least one term is traced by a provider of at least one value-adding component.

34. The method of claim 28, wherein the at least one term is authenticated by a provider of at least one value-adding component.

35. The method of claim 28, wherein the at least one term is accessed for at least one of verification, authentication, and authorization.

36. The method of claim 28, wherein the at least one term comprises at least one of readable text, visible color, voice command, and visual instructions.

37. The method of claim 28, wherein the at least one term comprises humanly perceptible information.

38. The method of claim 20, wherein the value-added information is convertible into a tangible good.

39. The method of claim 20, further comprising verifying the value-added information.

40. The method of claim 20, further comprising authenticating the value-added information.

41. The method of claim 20, wherein the value-adding components comprise at least one of an equity purchase right, an option, a warrant, and a security instrument.

42. The method of claim 20, wherein the value-adding components comprise a non-cash service.

43. A method for conducting at least one trusted transaction between at least two parties, comprising:

authenticating the at least two parties;

agreeing to a security of a transmission channel;

exchanging secondary value-added information;

determining at least one term for a primary value-added information exchange; and

facilitating payment for the transaction based on the terms.

44. The method of claim 43, wherein the step of facilitating payment for the transaction is accomplished in real-time.

45. The method of claim 44, wherein the at least one term includes micropayment systems.

46. The method of claim 43, wherein the transaction is governed by at least one of legal restrictions that apply to at least one of the parties, a timing of the transaction, a geographic location of the transaction, and value-added information.

47. The method of claim 43, wherein the value-added information is represented physically.

48. The method of claim 43, wherein the secondary value-added information comprises at least one of an equity option and at least one term from a previous trusted transaction.

49. The method of claim 43, wherein the secondary value-added information derives benefit from a previous trusted transaction.

50. The method of claim 49, wherein the at least two trusted transactions are substantially contiguous.

51. The method of claim 49, wherein the at least two trusted transactions have at least one of a time or an event limitation.

52. The method of claim 43, further comprising the step of:

agreeing to at least one term for a different transaction.

53. The method of claim 43, wherein the first trusted transaction enables manipulation of information in a subsequent transaction.

54. A method for conducting a trusted transaction between at least two parties, comprising:

establishing a steganographic cipher;

exchanging secondary value-added information between the parties;

agreeing to at least one term for the exchange of primary value-added information; and

facilitating payment for the transaction.

55. The method of claim 54, wherein the step of facilitating payment for the transaction is accomplished in real-time.

56. The method of claim 54, wherein the step of facilitating payment for the transaction is based on the at least one term for the primary value-added information exchange.

57. The method of claim 54, wherein the transaction is governed by at least an age and a geographical limitation.

58. The method of claim 54, wherein the transaction is governed by at least one of legal restrictions that apply to at least one of the parties, a timing of the transaction, a geographic location of the transaction, and value-added information.

59. The method of claim 54, wherein at least one of the primary and secondary value-added information is represented physically.

60. A method for conducting a trusted transaction between at least two parties, comprising:

identifying at least one of a unique identification for each of the at least two parties, a unique identification of the transaction, a unique identification of value-

added information to be transacted, and a unique identification of a value-adding component;

applying a steganographic cipher; and

verifying an agreement to transact between the parties.

61. The method of claim 60, wherein the trusted transaction is governed by at least one of a transaction age and a geographical location of the transaction.

62. The method of claim 60, wherein the trusted transaction is governed by legal restrictions that apply to at least one of the parties, a timing of the transaction, and value-added information.

63. The method of claim 60, wherein the value-added information is represented physically.

64. The method of claim 60, further comprising the step of:

transmitting the value-added information.

65. The method of claim 60, wherein the agreement causes at least one secondary term to be enabled for at least one of the parties.

66. The method of claim 60, wherein the agreement creates at least one term for a second trusted transaction.

67. The method of claim 60, further comprising the step of:

agreeing to at least one term for a second trusted transaction.

68. A method for bi-directionally exchanging value-added information between at least two parties, comprising:

associating a plurality of unique identifiers with the value-added information, the value-added information including at least one of a digital watermark, a file header, a file attachment, and a file wrapper;

associating each of the at least two parties with unique identifiers, the unique identifiers including at least one of a digital watermark, a file header, a file attachment, and a file wrapper; and

exchanging value-added information between the at least two parties.

69. The method of claim 68, wherein the transaction and the unique identifiers are stored for subsequent reference.

70. The method of claim 68, wherein unique identifiers are access restricted by at least one pre-determined rule.

71. The method of claim 68, wherein the unique identifiers are asymmetrically access restricted.

72. The method of claim 70, wherein the access restriction is dependent on verification of a querying party.

73. The method of claim 70, wherein the access restriction allows value-added information to be transmitted in an altered format.

74. The method of claim 68, further comprising the step of:

associating the bi-directional exchange of value-added information with a subsequent exchange of additional value-added information.

75. The method of claim 74, wherein the additional value-added information is governed by at least one separate term.

76. The method of claim 74, wherein the additional value-added information comprises a right to purchase equity in at least one of the parties to the transaction.

77. The method of claim 68, further comprising the step of agreeing to at least one term for a subsequent transaction.

78. A method for exchanging value-added information between at least two parties, comprising:

providing a data transmission means;

verifying the parties to the transaction;

negotiating at least one term selected from the group consisting of a price, a service, a selection, and combinations thereof; and

binding the at least one term to the information using at least one of a digital watermark, a file header, metadata, and a file wrapper;

wherein the at least one bound transaction term comprises value-added information.

79. The method of claim 78, wherein the at least one bound term cannot be removed without altering the value-added information.

80. The method of claim 78, wherein an authentication of the value-added information requires successful verification of the at least one bound term.

81. A method for trusted transactions, comprising the steps of:

receiving data to be processed;

determining a structure of the data;

determining if the data is authentic; and

determining an associated usage of the data based on the data structure and the authenticity of the data.

82. The method of claim 81, wherein the data is comprises at least one of aesthetic data and functional data.

83. The method of claim 81, wherein the structure of the data is determined based on at least one of a digital signature, a digital watermark, and a digital notary.

84. The method of claim 81, wherein the authenticity of the data is determined based on at least one of a digital signature, a digital watermark and a digital notary.

85. The method of claim 83, further comprising the step of verifying at least one of the digital signature, the digital watermark, and the digital notary by at least one of a trusted third party and a certification authority

86. The method of claim 83, wherein a bit from at least one of the digital signature, the digital watermark and the digital notary can be verified by at least one of a trusted third party and a certification authority.

87. A method for secure transaction, comprising:

receiving a request to process a transaction;

uniquely identifying a source of the request;

uniquely identifying at least one term of the request; and

storing identification information for transaction negotiation.

88. The method of claim 87, wherein the at least one term of the request includes at least one of a condition and a timing of the request.

89. The method of claim 87, wherein the request may be received over at least one of a secure and an unsecure transmission line.

90. The method of claim 87, wherein the source of the request is identified by at least one of a determinable origin of the source and a predetermined routing of the request by the seller.

91. The method of claim 87, wherein the at least one term of the request comprises a value-adding component.

92. The method of claim 87, wherein the transaction is noncontiguous with the request.

93. The method of claim 87, wherein the transaction and the request are processed in real time.

94. A method for the facilitation of the exchange of information data between at least a first party and a second party, comprising:

receiving a rule governing information data from a first party;

receiving a request for the information data from a second party;

matching the rule with the request; and

uniquely identifying the information data and the first and second parties;

wherein the information data is selected from the group consisting of unstructured data and structured data.

95. The method of claim 94, wherein the rule governs a use of the information data.

96. The method of claim 95, wherein the use comprises manipulating the information data.

97. The method of claim 95, wherein the use comprises transferring the information data.

98. The method of claim 95, wherein the use comprises subsequently changing to the information data.

99. The method of claim 95, wherein the use comprises playing the information data.

100. The method of claim 95, wherein the use comprises recording the information data.

101. The method of claim 95, wherein the use comprises converting the information data from at least one of analog to digital format and digital to analog format.

102. The method of claim 94, wherein the structured data comprises at least one of source code and executable code.

103. The method of claim 94, wherein the request may be filtered according to at least one of a characteristic, a function, an aesthetic, a condition, a history, a context, a consideration, a cost, a time, a bandwidth requirement, a storage

requirement, an available format, an owner identification, a creator identification, a seller identification, an infomediary identification, a distributor identification, a distribution parameter, an age in unit of time, and a upcoming information data.

104. The method of claim 94, wherein the unique identification is cryptographically secure.

105. The method of claim 104, wherein the unique identification may be cryptographically secured by using at least one of a cryptographic cipher, a stegnographic cipher for digital signatures, a special one-way hash, a digital watermark, and a time stamp, and combinations thereof.

106. The method of claim 94, further comprising the step of verifying the unique identification by an independent third party

107. The method of claim 106, wherein the independent third party comprises at least one of a certification authority, a creator of the information, an owner of the information, and a mutually agreed to third party.

108. The method of claim 94, wherein the exchange is in real time.

100. The method of claim 94, wherein the exchange is substantially noncontiguous.

110. A method for rights management, comprising:

receiving information;

determining whether the information is structured information or unstructured information;

identifying the information with a steganographic cipher;

authenticating the information with at least one of a digital signature and digital watermark check; and

associating the identification and authentication results with at least one of a predetermined record, a predetermined rule, and a predetermined function.

111. The method of claim 110, further comprising the step of:

limiting an access to the information based on a predetermined exposure of a decision maker.

112. The method of claim 110, further comprising the step of:

limiting a financial exposure based on a predetermined exposure of a decision maker.

113. A method for rights management, comprising:

exchanging information between at least two parties;

verifying the information, the verification performed by at least one of the parties; and

activating at least one of a predetermined act and a rule based on the result of the verification of information.

114. The method of claim 113, wherein information is exchanged in a format selected from the group consisting of an analog waveform and binary data.

115. The method of claim 113, further comprising the step of authenticating the verification by a trusted third party.

116. The method of claim 113, wherein an anonymity of each party is maintained during the step of verifying the information.

117. The method of claim 113, further comprising the step of making the verification publicly available for additional verification.

118. The method of claim 113, wherein the predetermined rule is activated noncontiguously with verification.

119. The method of claim 113, further comprising the step of making the accessible for further authentication and identification.

120. A method for risk management, comprising:

receiving information;

determining whether the information is structured or unstructured;

identifying information with a predetermined ciphered key;

authenticating information with at least one of a digital signature, a digital watermark check, and a predetermined ciphered key;

associating identification and authentication results with a predetermined rule; and

limiting access based on a predetermined exposure of a decision maker.

121. A method for securely exchanging information data between at least two parties, comprising:

creating a private key;

deriving a corresponding public key corresponding to the information data sought and at least one of (a) verifiable data associated with different versions of the

information data, (b) verifiable data associated with a transmitting device, and (c) verifiable data associated with an identity of the party seeking the information data;

establishing a set of one time signatures relating to the information data;

establishing a hierarchy of access to the set of one time signatures;

creating a public key signature that is verifiable with the public key, including the hierarchy of access to the set of one time signatures;

providing the information to a certification authority for verification; and

verifying the one time signature and the hierarchy of access to enable transfer of predetermined data.

122. A method for authenticating an exchange of a plurality of sets of information data between at least two parties, comprising:

creating a plurality of hierarchical classes based on a perceptual quality of the information data;

assigning each set of information data to a corresponding hierarchical class;

defining access to each hierarchical classes and to each set of information data based on at least one recognizable feature of the information data to be exchanged;

predetermining access to the sets of information data by perceptually-based quality determinations;

establishing at least one connection between the exchanging parties;

perceptually recognizing at least one of the sets of information data dependent on user provided value-added information data; and

enabling a trusted transaction based on verification, and associated access, governing at least one of a set of information data sets.

123. The method of claim 122, further comprising the step of grouping each hierarchical class by at least one of a quality, a price, and a service.

124. The method of claim 123, wherein the grouping is determined by at least one of a buyer and a seller.

125. The method of claim 123, wherein the grouping enables greater exchange of information.

126. A method for authenticating the exchange of perceptual information data between at least two parties over a networked system, comprising:

creating a plurality of hierarchical classes based on a perceptual quality of the information data;

assigning each set of information data to a corresponding hierarchical class;

defining access to each hierarchical classes and to each set of information data based on at least one recognizable feature of the information data to be exchanged;

perceptually recognizing at least one of the sets of information data dependent on user provided value-added information data;

enabling a trusted transaction of the information data based on verification of means of payment, and associated access, governing at least one copy of the information data sought;

associating the transaction event with the information data prior to transmission of the information data; and

transmitting and confirming delivery of the information data

127. The method of claim 126, further comprising the step of grouping the class of data by at least one of quality, price, and service.

128. The method of claim 127, wherein the grouping is determined by at least one of a buyer and a seller.

129. The method of claim 127, wherein the grouping enables greater exchange of information.

130. The method of claim 126, further comprising the step of:

confirming both a digital and an analog copy of the transmission.

131. The method of claim 127, further comprising the step of:

associating the transaction event with the buyer or seller to develop trust with other party

132. The method of claim 126, further comprising the step of:

charging at least one party based on a transaction bandwidth requirement.

133. A device for conducting a trusted transaction between at least two parties who have agreed to transact, comprising:

means for uniquely identifying unique identification information selected from the group consisting of a unique identification of one of the parties, a unique

identification of the transaction, a unique identification of value-added information to be transacted, and a unique identification of a value-adding component;

a steganographic cipher; and

means for verifying an agreement to transact between the parties.

134. The device of claim 133, wherein the unique identification information seeds the steganographic cipher.

135. The device of claim 133, wherein the unique identification information is verifiable.

136. The device of claim 133, further comprising:

means for transmitting value-added information.

137. The device of claim 136, wherein the means for transmitting value-added information transmits the value-added information by a method selected from the group consisting of electrical and physical.

138. The device of claim 136, wherein the wherein the means for transmitting value-added information transmits the value-added information in a medium selected from the group consisting of a pre-determined file format and a predetermined carrier medium.

139. A device for conducting a trusted transaction between at least two parties who have agreed to transact, comprising:

means for uniquely identifying unique identification information selected from the group consisting of a unique identification of one of the parties, a unique identification of the transaction, a unique identification of value-added information to be transacted, and a unique identification of a value-adding component; and

means for enabling a subsequent mutually agreed to at least one term.

140. The method of claim 139, wherein the at least one subsequent term concerns at least one of equity, service, and recognition.

141. A device for conducting trusted transactions between at least two parties, comprising:

a steganographic cipher;

a controller for receiving input data or outputting output data; and

at least one input/output connection,

wherein the device has a unique identification code.

142. The device of claim 141, wherein the unique identification code is predetermined.

143. The device of claim 141, wherein the unique identification code is upgradeable.

144. The device of claim 141, wherein the steganographic cipher comprises:

a number generator selected from the group consisting of a pseudo-random number generator and a random number generator;

a predetermined key generation algorithm selected from the group consisting of a hash function and a special one-way function;

a predetermined message information selected from the group consisting of a digital signature, a time stamp, a digital watermark, and function-dependent data;

a predetermination of the information carrier signals characteristics selected from the group consisting of a perceptual characteristic and a signal feature.

145. The device of claim 141, wherein the steganographic cipher manipulates the input data.

146. The device of claim 141, wherein the steganographic cipher manipulates the output data

147. The device of claim 141, wherein the input of input data is controlled by predetermined information selected from the group consisting of a pass phrase, a password, biometric data, and a personal entropy query.

148. The device of claim 144, wherein an identification of a device holder requires at least one additional iteration of verification by at least one of a pass phrase, a password, biometric data, and a personal entropy query.

149. The device of claim 141, wherein the device converts at least one value-added information metrics selected from the group consisting of a price, a selection, and a service into humanly perceptible information.

150. The device of claim 149, wherein the humanly perceptible information relates to at least one of a present value cost to the party, at least one term for use, a level of confidence over the transaction, a level of confidence over transmission security, and a data integrity metric of the value-added information.

151. The device of claim 141, wherein the device is manufactured as a device selected from the group consisting of a smart card, a microchip, and a software application.

152. The device of claim 151, wherein the manufactured device is tamper-resistant.

153. The device of claim 151, wherein the manufactured device ceases to function if at least one function of the manufactured device is altered by an unauthorized party.

154. The device of claim 151, wherein the software application is subject to a steganographic cipher for serialization or creating unique instances of individual copies of the application.

155. The device of claim 141, further comprising an analog to digital converter.

156. The device of claim 141, wherein the device is securely linked to at least one of a means for payment and a transmission channel for private key exchange and approval.

157. The device of claim 156, wherein the key approval is selected from the group consisting of identification, authentication, and authorization.

158. The device of claim 141, wherein the device transacts according to at least one predetermination of at least an identity of the vendor, a plurality of conditions of the information transfer, a payment, and an identity of a separate but similar device.

159. The device of claim 141, wherein the device further comprises:

an internal memory.

160. A trusted transaction device for transmitting authentic value-added information data between at least two parties, comprising:

a display;

a unique identifier;

means for ciphering information input and output;

means for interacting with other similarly functional devices; and

means for storing or retrieving value-added information and a value-adding component.

161. The device of claim 160, wherein the display transceives cryptographically verifiable information.

162. The device of claim 161, wherein the cryptographically verifiable information is observed by a user.

163. The device of claim 160, wherein the unique identifier is upgradeable.

164. The device of claim 160, wherein the unique identifier is serialized.

165. The device of claim 160, wherein the unique identifier comprises at least one of a means for facilitating transaction authorization, a means for facilitating bandwidth requirements, and a means for associating the unique identifier with information.

166. The device of claim 160, wherein the means for ciphering information comprises at least one of a means for facilitating transaction authorization, a means for facilitating bandwidth requirements, and a means for associating the unique identifier with information.

167. The device of claim 160, further comprising:

a means for establishing communications/connecting with other similarly outfitted devices;

a means for storing or retrieving trusted transaction value-adding component data; and

a means for attaching storage or transducers to the device.

168. The device of claim 167, further comprising:

means for anonymous tracing of the transaction.

169. The device of claim 167, wherein information is processed in real time.

170. A device for securely exchanging information data, comprising:

means for creating a private key by the party seeking predetermined data;

means for deriving a corresponding public key based on the predetermined data and at least one of verifiable data associated with different versions of the information, verifiable data associated with a transmitting device, and verifiable data associated with the identity of the party seeking information;

means for creating a set of one-time signatures relating to the predetermined data;

means for validating a predetermined hierarchy of access of the set of one-time signatures;

means for creating a public key signature, verifiable with the public key, including the access hierarchy of one time signatures;

means for securely transacting predetermined data by providing information relating to a proposed transaction; and

means for verifying the one time signature and the hierarchy of access to enable transfer of predetermined data.

171. The device of claim 170, further comprising

a means for interacting with other equipped devices.

172. The device of claim 171, further comprising:

means for establishing a secure transmission.

173. A system for the secure exchange of predetermined, verifiable information data between at least two parties, comprising:

at least one condition for the use of the information;

means for differentiating between predetermined information and other seemingly identical information based on an authentication protocol;

means for associating authenticity of verifiable information data with at least one condition for use;

a storage unit for storing the predetermined, verifiable information; and

means for communicating with the predetermined, verifiable information storage.

174. The system of claim 173, wherein the means for differentiating between predetermined information and the seemingly identical information based on an authentication protocol comprises at least one of a hash, a signature, and a secure watermark.

175. The system of claim 173, further comprising:

means for authenticating verifiable information flow between transacting parties.

176. The system of claim 173, wherein the system securely exchanges predetermined, verifiable information data prior to consummating verifiable financial transaction between the parties.

177. A system for the exchange of information, comprising:

at least one sender;

at least a receiver;

a verifiable message; and

a verification of the message by at least one of the senders and the receivers; wherein a verification of the message enables a decision over receiving additional related information.

178. A system for computer based decision protocol comprising:

a means for identifying between structured and unstructured information;

a means for authenticating structured information; and

a means for enabling a decision rule based on the identity and authenticity of the information.

179. The system of claim 178, further comprising:

a means for comparing decision results with at least one predetermined rule.

180. A system for computer-based decision protocol, comprising:

means for identifying between structured and unstructured information;

means for identifying structured information; and

means for enabling a predetermined decision rule based on the identity of the information.

181. The system of claim 180, wherein the structured information is defined by at least one of a digital signal processor and a general purpose computing device.

182. The system of claim 180, wherein the structured information comprises binary data.

183. The system of claim 180, wherein the structured information is humanly perceptible.

184. The system of claim 180, wherein the structured information is defined in a bit addressable manner.

185. The system of claim 180, wherein the structured information has at least one mathematically definable characteristic.

186. The system of claim 180, wherein the structured information is selected from the group consisting of pseudo-random and random.

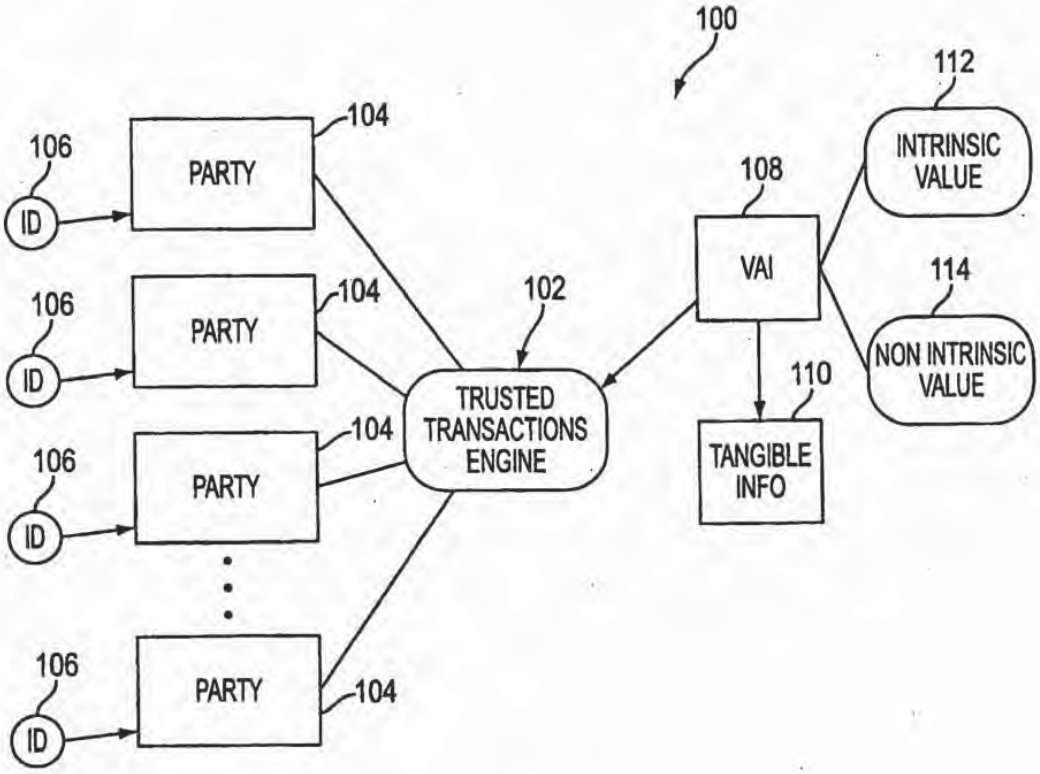

100
112
INTRINSIC VALUE
108
VAI
114
NON INTRINSIC VALUE
110
TANGIBLE INFO
102
TRUSTED TRANSACTIONS ENGINE
104
PARTY
106
ID
104
PARTY
106
ID
104
PARTY
106
ID
104
PARTY
106
ID

FIG. 1

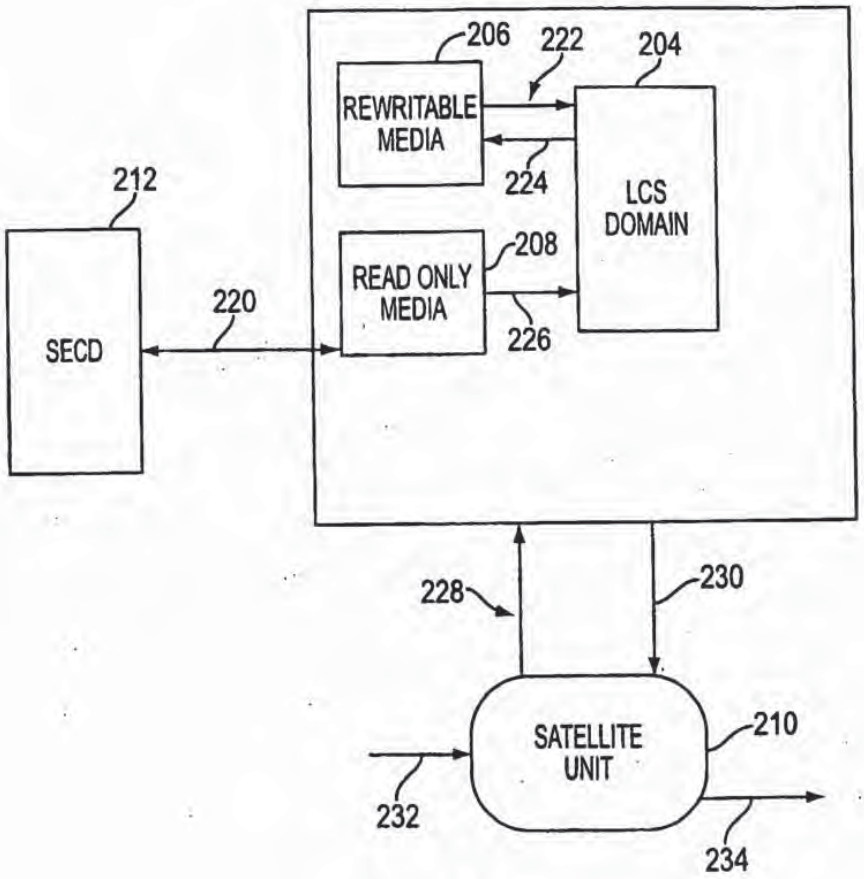


FIG. 2

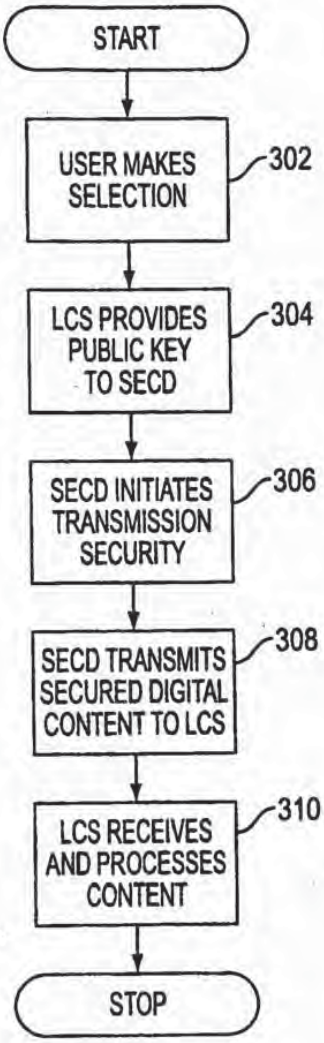

START
USER MAKES SELECTION
302
LCS PROVIDES PUBLIC KEY TO SECD
304
SECD INITIATES TRANSMISSION SECURITY
306
SECD TRANSMITS SECURED DIGITAL CONTENT TO LCS
308
LCS RECEIVES AND PROCESSES CONTENT
310
STOP

FIG. 3

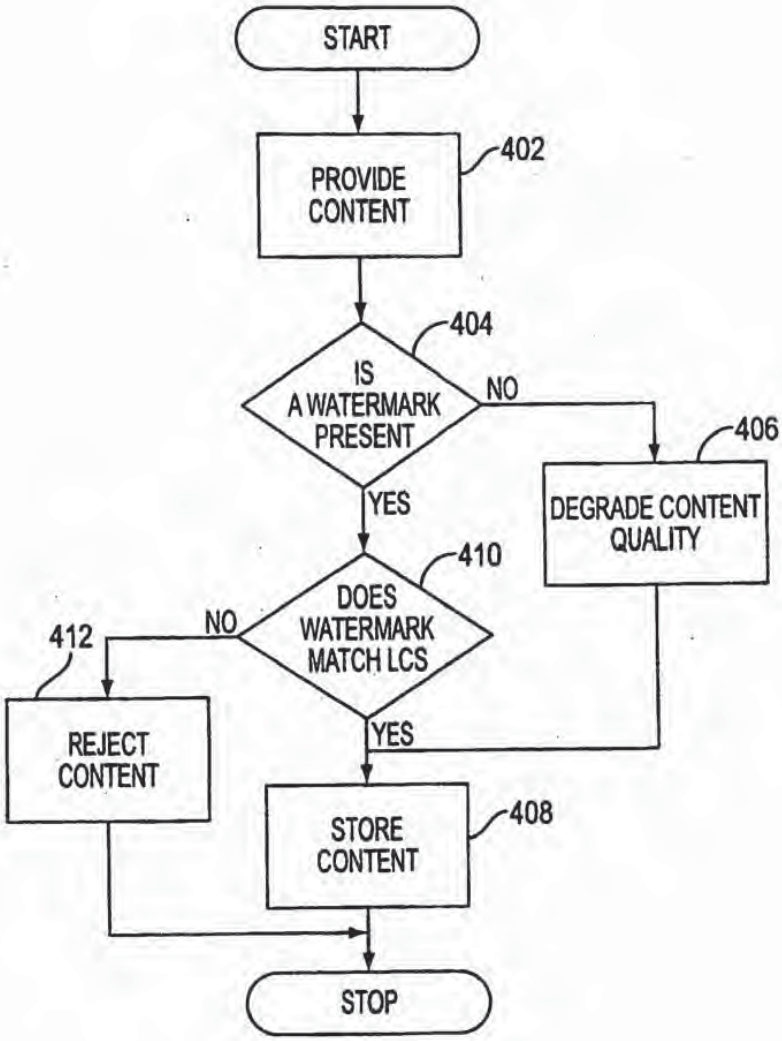


FIG. 4

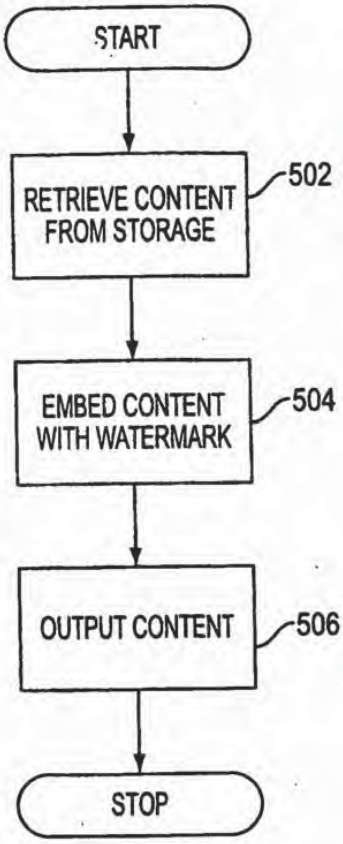

START
RETRIEVE CONTENT FROM STORAGE
502
EMBED CONTENT WITH WATERMARK
504
OUTPUT CONTENT
506
STOP

FIG. 5

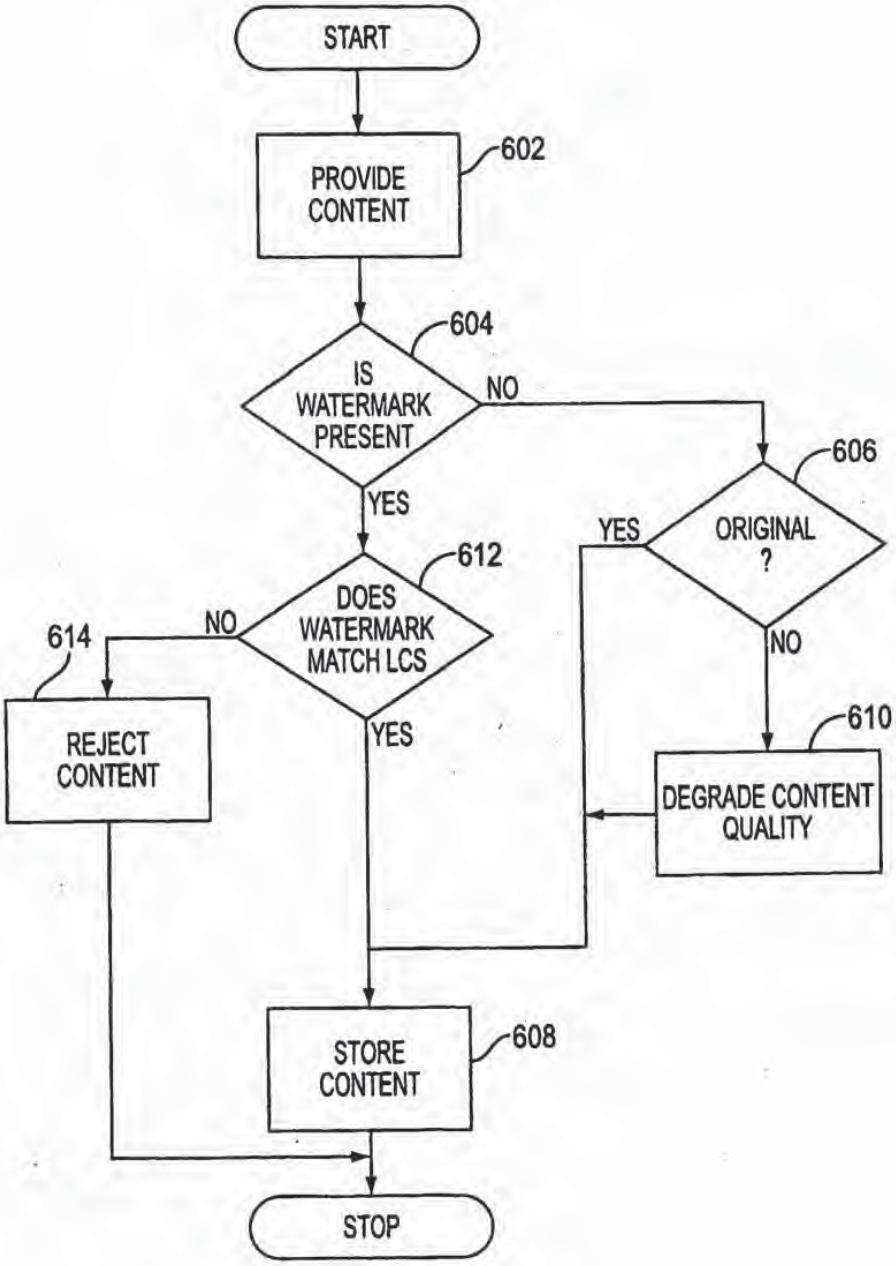

START
PROVIDE CONTENT
602
IS WATERMARK PRESENT
604
NO
YES
ORIGINAL ?
606
YES
NO
DOES WATERMARK MATCH LCS
612
NO
YES
614
REJECT CONTENT
DEGRADE CONTENT QUALITY
610
STORE CONTENT
608
STOP

FIG. 6

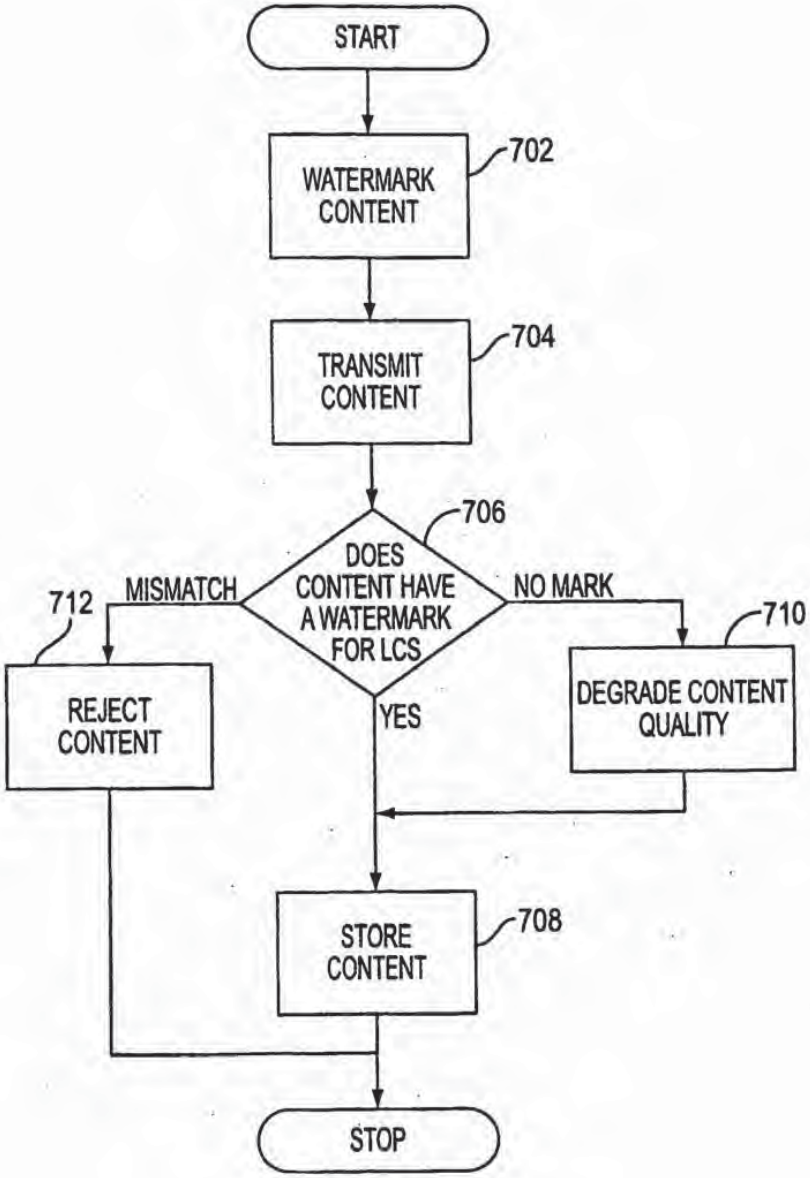

START
WATERMARK CONTENT
702
TRANSMIT CONTENT
704
706
DOES CONTENT HAVE A WATERMARK FOR LCS
MISMATCH
NO MARK
YES
712
REJECT CONTENT
710
DEGRADE CONTENT QUALITY
STORE CONTENT
708
STOP

FIG. 7

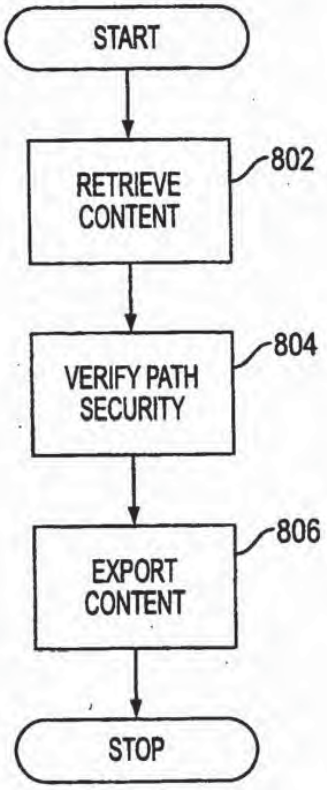

START
RETRIEVE CONTENT
802
VERIFY PATH SECURITY
804
EXPORT CONTENT
806
STOP

FIG. 8

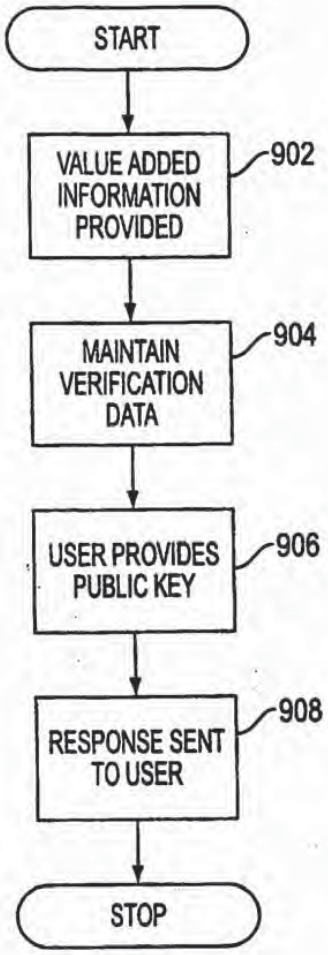

START
VALUE ADDED INFORMATION PROVIDED
902
MAINTAIN VERIFICATION DATA
904
USER PROVIDES PUBLIC KEY
906
RESPONSE SENT TO USER
908
STOP

FIG. 9

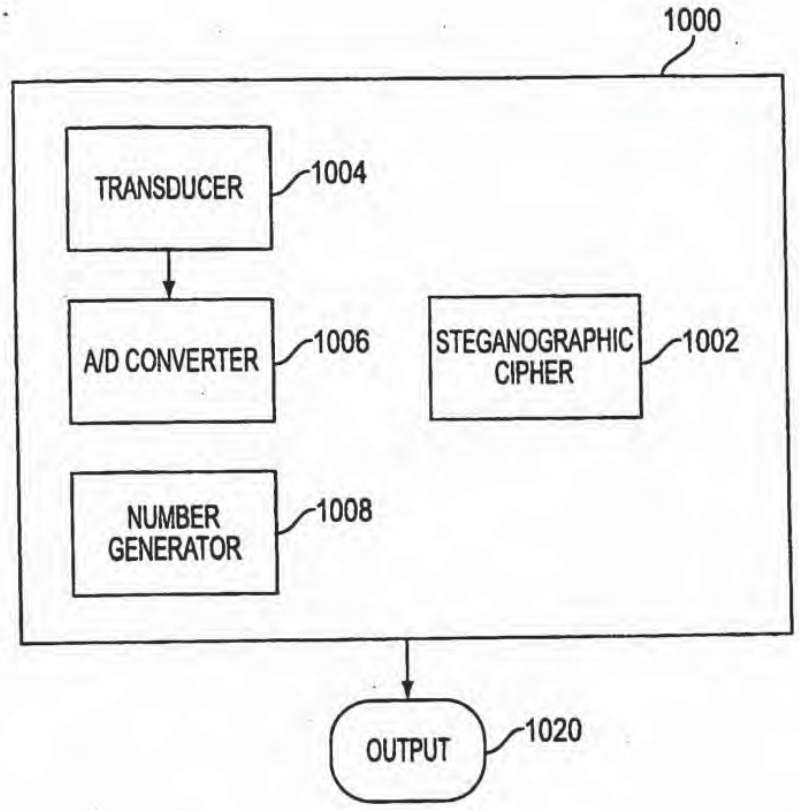

1000
TRANSDUCER
1004
A/D CONVERTER
1006
STEGANOGRAPHIC CIPHER
1002
NUMBER GENERATOR
1008
OUTPUT
1020

FIG. 10

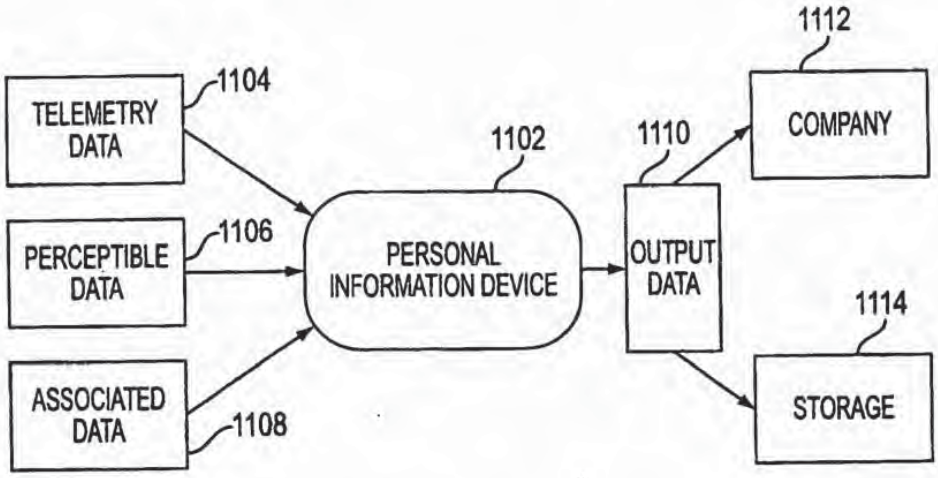

1112
1104
TELEMETRY DATA
1102
1110
COMPANY
1106
PERCEPTIBLE DATA
PERSONAL INFORMATION DEVICE
OUTPUT DATA
1114
ASSOCIATED DATA
1108
STORAGE

FIG. 11

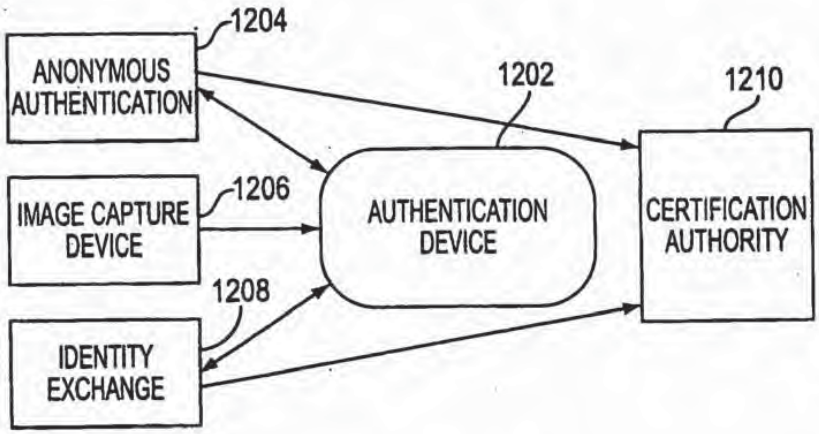

1204
ANONYMOUS AUTHENTICATION
1202
1210
IMAGE CAPTURE DEVICE
1206
AUTHENTICATION DEVICE
CERTIFICATION AUTHORITY
1208
IDENTITY EXCHANGE

FIG. 12

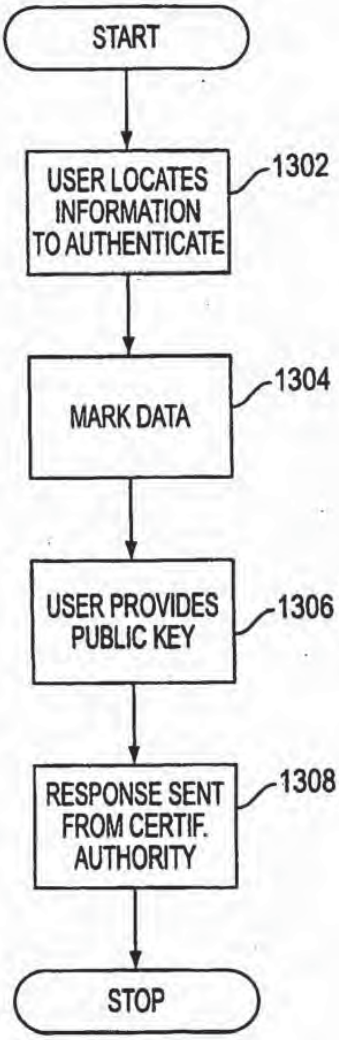

START
USER LOCATES INFORMATION TO AUTHENTICATE
1302
MARK DATA
1304
USER PROVIDES PUBLIC KEY
1306
RESPONSE SENT FROM CERTIF. AUTHORITY
1308
STOP

FIG. 13

INTERNATIONAL SEARCH REPORT

Int. anal Application No
PCT/US 00/33126

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06F17/60

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 903 721 A (SIXTUS TIMOTHY) 11 May 1999 (1999-05-11) abstract column 3, line 26 -column 5, line 31	1-19
X	US 5 790 677 A (SPELMAN JEFFREY F ET AL) 4 August 1998 (1998-08-04) abstract column 2, line 6 -column 4, line 39	1-19
X	WO 96 29795 A (MICALI SILVIO) 26 September 1996 (1996-09-26) abstract page 5, line 27 -page 8, line 6	1-19

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 March 2001	04.04.01

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Corcoran, P

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/33126

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 24833 A (MICALI SILVIO) 10 July 1997 (1997-07-10) abstract page 2, line 12 -page 5, line 8	1-19
A	US 5 539 735 A (MOSKOWITZ SCOTT A) 23 July 1996 (1996-07-23) abstract column 1, line 60 -column 4, line 29	1-19
A	SIRBU M ET AL: "NETBILL: AN INTERNET COMMERCE SYSTEM OPTIMIZED FOR NETWORK DELIVERED SERVICES" DIGEST OF PAPERS OF THE COMPUTER SOCIETY COMPUTER CONFERENCE (SPRING) COMPCON,US,LOS ALAMITOS, IEEE COMP. SOC. PRESS, vol. CONF. 40, 5 March 1995 (1995-03-05), pages 20-25, XP000577034 ISBN: 0-7803-2657-1 The whole document	1-19
A	SCHUNTER M ET AL: "A status report on the SEMPER framework for secure electronic commerce" COMPUTER NETWORKS AND ISDN SYSTEMS,NL,NORTH HOLLAND PUBLISHING. AMSTERDAM, vol. 30, no. 16-18, 30 September 1998 (1998-09-30), pages 1501-1510, XP004138681 ISSN: 0169-7552 2. Model for electronic commerce 3. The SEMPER framework	1-19
A	KONRAD K ET AL: "Trust and electronic commerce-more than a technical problem" PROCEEDINGS OF THE 18TH IEEE SYMPOSIUM ON RELIABLE DISTRIBUTED SYSTEMS, PROCEEDINGS 18TH IEEE SYMPOSIUM ON RELIABLE DISTRIBUTED SYSTEMS, LAUSANNE, SWITZERLAND, 19-22 OCT. 1999, pages 360-365, XP002162270 1999, Los Alamitos, CA, USA, IEEE Comput. Soc, USA ISBN: 0-7695-0290-3 3. Trust, Security and Electronic Commerce 4. Technology and Institutions	1-19

-/--

INTERNATIONAL SEARCH REPORT

Inte onal Application No
PCT/US 00/33126

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KINI A ET AL: "Trust in electronic commerce: definition and theoretical considerations" PROCEEDINGS OF THE THIRTY-FIRST HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES (CAT. NO.98TB100216), PROCEEDINGS OF THE THIRTY-FIRST HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES, KOHALA COAST, HI, USA, 6-9 JAN. 1998, pages 51-61, XP002162271 1998, Los Alamitos, CA, USA, IEEE Comput. Soc, USA ISBN: 0-8186-8255-8 1.3 The Significance of Trust in Electronic Commerce, ---	1-19
A	STEINAUER D D ET AL: "Trust and traceability in electronic commerce" STANDARD VIEW, SEPT. 1997, ACM, USA, vol. 5, no. 3, pages 118-124, XP002162272 ISSN: 1067-9936 The whole document ---	1-19
A	US 5 687 236 A (MOSKOWITZ SCOTT A ET AL) 11 November 1997 (1997-11-11) abstract ---	8,9
A	US 5 745 569 A (MOSKOWITZ SCOTT A ET AL) 28 April 1998 (1998-04-28) abstract -----	8,9

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte onal Application No
PCT/US 00/33126

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5903721 A	11-05-1999	AU 6549498 A	29-09-1998
		DE 1008022 T	25-01-2001
		EP 1008022 A	14-06-2000
		ES 2150892 T	16-12-2000
		NO 994428 A	09-11-1999
		WO 9840809 A	17-09-1998
US 5790677 A	04-08-1998	NONE	
WO 9629795 A	26-09-1996	WO 9806198 A	12-02-1998
		CA 2215908 A	26-09-1996
		EP 0815671 A	07-01-1998
		US 5553145 A	03-09-1996
		US 5629982 A	13-05-1997
		US 5666420 A	09-09-1997
		US 6137884 A	24-10-2000
		US 6141750 A	31-10-2000
		EP 0917781 A	26-05-1999
		JP 2000515649 T	21-11-2000
WO 9724833 A	10-07-1997	US 5615269 A	25-03-1997
		AU 1951497 A	28-07-1997
US 5539735 A	23-07-1996	US 5428606 A	27-06-1995
		WO 9701892 A	16-01-1997
US 5687236 A	11-11-1997	US 5613004 A	18-03-1997
		EP 0872073 A	21-10-1998
		WO 9642151 A	27-12-1996
US 5745569 A	28-04-1998	AU 1829497 A	11-08-1997
		WO 9726732 A	24-07-1997

Form PCT/ISA/210 (patent family annex) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/33126

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 20-186
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 20-186

In view of the large number and also the wording of the claims presently on file, which render it difficult, if not impossible, to determine the matter for which protection is sought, the present application fails to comply with the clarity and conciseness requirements of Article 6 PCT (see also Rule 6.1(a) PCT) to such an extent that a meaningful search is impossible.

Moreover, the proliferation of independent claims and the broad manner in which these have been worded make it impossible to determine which parts of the claims may be said to define subject-matter for which protection might legitimately be sought (Article 6 PCT). For these reasons, a meaningful search over the whole breadth of the claim(s) is impossible.

Consequently, the search has been restricted to the subject matter recited in claims 1-19.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

European Patent Office
Postbus 5818
2280 HV RIJSWIJK
NETHERLANDS
Tel.: +31 70 340 2040
Fax: +31 70 340 3016

Europäisches Patentamt | European Patent Office | Office européen des brevets

Moskowitz, Scott A.

Townhouse 4, 20191 East Country Club Drive
North Miami Beach, FL 33180
ETATS-UNIS D'AMERIQUE

EPO Customer Services
Tel.: +31 (0)70 340 45 00

Date
01.10.07

Reference	Application No./Patent No.
	07112420.0-1228
Applicant/Proprietor	
Wistaria Trading, Inc.	

Designation as inventor - communication under Rule 17(3) EPC

You have been designated as inventor in the above-mentioned European patent application. Below you will find the data contained in the Designation of Inventor and further data mentioned in Art. 128(5) EPC:

DATE OF FILING : 07.06.96

PRIORITY : US/07.06.95/ USA 489172

TITLE : Steganographic method and device

DESIGNATED STATES : AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

INVENTOR (PUBLISHED = 1, NOT PUBLISHED = 2):

1/Cooperman, Marc S./ 20 Wildwood/Short Hills, NJ 07078/US
1/Moskowitz, Scott A./ Townhouse 4, 20191 East Country Club Drive/North Miami Beach, FL 33180/US

DECLARATION UNDER ARTICLE 81 EPC:
The applicant(s) has (have) acquired the right to the European patent as employer(s).

RECEIVING SECTION

Europäisches Patentamt • European Patent Office • Office européen des brevets

PayWord and MicroMint: Two simple micropayment schemes

Ronald L. Rivest* and Adi Shamir**

April 27, 2001

*MIT Laboratory for Computer Science
545 Technology Square, Cambridge, Mass. 02139

**Weizmann Institute of Science
Applied Mathematics Department
Rehovot, Israel

{rivest,shamir}@theory.lcs.mit.edu

1 Introduction

We present two simple micropayment schemes, "PayWord" and "MicroMint," for making small purchases over the Internet. We were inspired to work on this problem by DEC's "Millicent" scheme[10]. Surveys of some electronic payment schemes can be found in Hallam-Baker [6], Schneier[16], and Wayner[18].

Our main goal is to minimize the number of public-key operations required per payment, using hash operations instead whenever possible. As a rough guide, hash functions are about 100 times faster than RSA signature verification, and about 10,000 times faster than RSA signature generation: on a typical workstation, one can sign two messages per second, verify 200 signatures per second, and compute 20,000 hash function values per second.

To support micropayments, exceptional efficiency is required, otherwise the cost of the mechanism will exceed the value of the payments. As a consequence, our micropayment schemes are light-weight compared to full macropayment schemes. We "don't sweat the small stuff": a user who loses a micropayment is similar to someone who loses a nickel in a candy machine. Similarly, candy machines aren't built with expensive mechanisms for detecting forged coins, and yet they work well in practice, and the overall level of abuse is low. Large-scale and/or persistent fraud must be detected and eliminated, but if the scheme delivers a volume of payments to the right parties that is roughly correct, we're happy.

In our schemes the players are brokers, users, and vendors. Brokers authorize users to make micropayments to vendors, and redeem the payments collected by the vendors. While user-vendor relationships are transient, broker-user and broker-vendor relationships are long-term. In a typical transaction a vendor sells access to a World-Wide Web page for one cent. Since a user may access only a few pages before moving on, standard credit-card arrangements incur unacceptably high overheads.

The first scheme, "PayWord," is a credit-based scheme, based on chains of "paywords" (hash values). Similar chains have been previously proposed for different purposes: by Lamport [9] and Haller (in S/Key) for access control [7], and by Winternitz [11] as a one-time signature scheme. The application of this idea for micropayments has also been independently discovered by Anderson et al. [2] and by Pederson [14], as we learned after distributing the initial draft of this paper. We discuss these related proposals further in Section 5. The user authenticates a complete chain to the vendor with a single public-key signature, and then successively reveals each payword in the chain to the vendor to make micropayments. The incremental cost of a payment is thus one hash function computation per party. PayWord is optimized for sequences of micropayments, but is secure and flexible enough to support larger variable-value payments as well.

The second scheme, "MicroMint," was designed to eliminate public-key operations altogether. It has lower security but higher speed. It introduces a new paradigm of representing coins by k-way hash-function collisions. Just as for a real mint, a broker's "economy of scale" allows him to produce large quantities of such coins at very low cost per coin, while small-scale forgery attempts can only produce coins at a cost exceeding their value.

2 Generalities and Notation

We use public-key cryptography (e.g. RSA with a short public exponent). The public keys of the broker B, user U, and vendor V are denoted PK_B, PK_U, and PK_V, respectively; their secret keys are denoted SK_B, SK_U, and SK_V. A message M with its digital signature produced by secret key SK is denoted $\{M\}_{SK}$. This signature can be verified using the corresponding public key PK.

We let h denote a cryptographically strong hash function, such as MD5[15] or SHA[13]. The output (nominally 128 or 160 bits) may be truncated to shorter lengths as described later. The important property of h is its one-wayness and collision-resistance; a very large search should be required to find a single input producing a given output, or to find two inputs producing the same output. The input length may, in some cases, be equal to the output length.

3 PayWord

PayWord is credit-based. The user establishes an account with a broker, who issues her a digitally-signed PayWord Certificate containing the broker's name, the user's name and IP-address, the user's public key, the expiration date, and other information. The certificate has to be renewed by the broker (e.g. monthly), who will do so if the user's account is in good standing. This certificate authorizes the user to make Payword chains, and assures vendors that the user's paywords are redeemable by the broker. We assume in this paper that each payword is worth exactly one cent (this could be varied).

In our typical application, when U clicks on a link to a vendor V's non-free web page, his browser determines whether this is the first request to V that day. For a first request, U computes and signs a "commitment" to a new user-specific and vendor-specific chain of paywords $w_1, w_2, \ldots, w_n$. The user creates the payword chain in reverse order by picking the last payword w_n at random, and then computing

$$w_i = h(w_{i+1})$$

for $i = n-1, n-2, \ldots, 0$. Here w_0 is the root of the payword chain, and is not a payword itself. The commitment contains the root w_0, but not any payword w_i for $i > 0$. Then U provides this commitment and her certificate to V, who verifies their signatures.

The i-th payment (for $i = 1, 2, \ldots$) from U to V consists of the pair (w_i, i), which the vendor can verify using w_{i-1}. Each such payment requires no calculations by U, and only a single hash operation by V.

At the end of each day, V reports to B the last (highest-indexed) payment (w_l, l) received from each user that day, together with each corresponding commitment. B charges U's account l cents and pays l cents into V's account. (The broker might also charge subscription and/or transaction fees, which we ignore here.)

A fundamental design goal of PayWord is to minimize communication (particularly on-line communication) with the broker. We imagine that there will be only a few nationwide

brokers; to prevent them from becoming a bottleneck, it is important that their computational burden be both reasonable and "off-line." PayWord is an "off-line" scheme: V does not need to interact with B when U first contacts V, nor does V need to interact with B as each payment is made. Note that B does not even receive every payword spent, but only the *last* payword spent by each user each day at each vendor.

PayWord is thus extremely efficient when a user makes repeated requests from the same vendor, but is quite effective in any case. The public-key operations required by V are only signature verifications, which are relatively efficient. We note that Shamir's probabilistic signature screening techniques[17] can be used here to reduce the computational load on the vendor even further. Another application where PayWord is well-suited is the purchase of pay-per-view movies; the user can pay a few cents for each minute of viewing time.

This completes our overview; we now give some technical details.

3.1 User-Broker relationship and certificates

User U begins a relationship with broker B by requesting an account and a PayWord Certificate. She gives B over a secure authenticated channel: her credit-card number, her public key PK_U, and her "delivery address" A_U. Her aggregated PayWord charges will be charged to her credit-card account. Her delivery address is her Internet/email or her U.S. mail address; her certificate will only authorize payments by U for purchases to be delivered to A_U.

The user's certificate has an expiration date E. Certificates might expire monthly, for example. Users who don't pay their bills won't be issued new certificates.

The broker may also give other (possibly user-specific) information I_U in the certificate, such as: a certificate serial number, credit limits to be applied per vendor, information on how to contact the broker, broker/vendor terms and conditions, etc.

The user's certificate C_U thus has the form:

$$C_U = \{B, U, A_U, PK_U, E, I_U\}_{SK_B} .$$

The PayWord certificate is a statement by B to any vendor that B will redeem authentic paywords produced by U turned in before the given expiration date (plus a day's grace).

PayWord is not intended to provide user anonymity. Although certificates could contain user account numbers instead of user names, the inclusion of A_U effectively destroys U's anonymity. However, some privacy is provided, since there is no record kept as to which documents were purchased.

If U loses her secret key she should report it at once to B. Her liability should be limited in such cases, as it is for credit-card loss. However, if she does so repeatedly the broker may refuse her further service. The broker may also keep a "hot list" of certificates whose users have reported lost keys, or which are otherwise problematic.

As an alternative to hot-lists, one can use hash-chains in a different manner as proposed by Micali [12] to provide daily authentication of the user's certificate. The user's certificate would additionally contain the root w'_0 of a hash chain of length 31. On day $j-1$ of the month, the broker will send the user (e.g. via email) the value w'_j if and only if the user's

account is still in good standing. Vendors will then demand of each user the appropriate w' value before accepting payment.

3.2 User-Vendor relationships and payments

User-vendor relationships are transient. A user may visit a web site, purchase ten pages, and then move on elsewhere.

Commitments

When U is about to contact a new vendor V, she computes a fresh payword chain w_1, ..., w_n with root w_0. Here n is chosen at the user's convenience; it could be ten or ten thousand. She then computes her commitment for that chain:

$$M = \{V, C_U, w_0, D, I_M\}_{SK_U} .$$

Here V identifies the vendor, C_U is U's certificate, w_0 is the root of the payword chain, D is the current date, and I_M is any additional information that may be desired (such as the length n of the payword chain). M is signed by U and given to V. (Since this signature is necessarily "on-line," as it contains the vendor's name, the user might consider using an "on-line/off-line" signature scheme[5].)

This commitment authorizes B to pay V for any of the paywords w_1, ..., w_n that V redeems with B before date D (plus a day's grace). Note that paywords are *vendor-specific* and *user-specific*; they are of no value to another vendor.

Note that U must sign a commitment for each vendor she pays. If she rapidly switches between vendors, the cost of doing so may become noticeable. However, this is PayWord's only significant computational requirement, and the security it provides makes PayWord usable even for larger "macropayments" (e.g. software selling at $19.99).

The vendor verifies U's signature on M and the broker's signature on C_U (contained within M), and checks expiration dates.

The vendor V should cache verified commitments until they expire at the end of the day. Otherwise, if he redeemed (and forgot) paywords received before the expiration date of the commitment, U could cheat V by replaying earlier commitments and paywords. (Actually, to defeat this attack, V need store only a short hash of each commitment he has reported to B already today.)

The user should preferably also cache her commitment until she believes that she is finished ordering information from V, or until the commitment expires. She can always generate a fresh commitment if she re-visits a vendor whose commitment she has deleted.

Payments

The user and vendor need to agree on the amount to be paid. In our exemplary application, the price of a web page is typically one cent, but could be some other amount. A web page should presumably be free if the user has already purchased it that day, and is just requesting it again because it was flushed from his cache of pages.

A payment P from U to V consists of a payword and its index:

$$P = (w_i, i) .$$

The payment is short: only twenty or thirty bytes long. (The first payment to V that day would normally accompany U's corresponding commitment; later payments are just the payword and its index, unless the previous chain is exhausted and a new chain must be committed to.) The payment is not signed by U, since it is self-authenticating (using the commitment).

The user spends her paywords in order: w_1 first, then w_2, and so on. If each payword is worth one cent, and each web page costs one cent, then she discloses w_i to V when she orders her i-th web page from V that day.

This leads to the PayWord payment policy: *for each commitment a vendor V is paid l cents, where (w_l, l) is the corresponding payment received with the largest index.* This means that V needs to store only one payment from each user: the one with the highest index. Once a user spends w_i, she can not spend w_j for $j < i$. The broker can confirm the value to be paid for w_l by determining how many applications of h are required to map w_l into w_0.

PayWord supports variable-size payments in a simple and natural manner. If U skips paywords, and gives w_7 after giving w_2, she is giving V a nickel instead of a penny. When U skips paywords, during verification V need only apply h a number of times proportional to the value of the payment made.

A payment does not specify what item it is payment for. The vendor may cheat U by sending him nothing, or the wrong item, in return. The user bears the risk of losing the payment, just as if he had put a penny in the mail. Vendors who so cheat their customers will be shunned. This risk can be moved to V, if V specifies payment *after* the document has been delivered. If U doesn't pay, V can notify B and/or refuse U further service. For micropayments, users and vendors might find either approach workable.

3.3 Vendor-Broker relationships and redemption

A vendor V needn't have a prior relationship with B, but does need to obtain PK_B in an authenticated manner, so he can authenticate certificates signed by B. He also needs to establish a way for B to pay V for paywords redeemed. (Brokers pay vendors by means outside the PayWord system.)

At the end of each day (or other suitable period), V sends B a redemption message giving, for each of B's users who have paid V that day (1) the commitment C_U received from U, (2) the last payment $P = (w_l, l)$ received from U.

The broker then needs to (1) verify each commitment received (he only needs to verify user signatures, since he can recognize his own certificates), including checking of dates, etc., and (2) verify each payment (w_l, l) (this requires l hash function applications). We assume that B normally honors all valid redemption requests.

Since hash function computations are cheap, and signature verifications are only moderately expensive, B's computational burden should be reasonable, particularly since it is more-or-less proportional to the payment volume he is supporting; B can charge transaction or subscription fees adequate to cover his computation costs. We also note that B never needs to respond in real-time; he can batch up his computations and perform them off-line overnight.

3.4 Efficiency

We summarize PayWord's computational and storage requirements:

- The broker needs to sign each user certificate, verify each user commitment, and perform one hash function application per payment. (All these computations are off-line.) The broker stores copies of user certificates and maintains accounts for users and vendors.
- The user needs to verify his certificates, sign each of his commitments, and perform one hash function application per payword committed to. (Only signing commitments is an on-line computation.) He needs to store his secret key SK_U, his active commitments, the corresponding payword chains, and his current position in each chain.
- The vendor verifies all certificates and commitments received, and performs one hash function application per payword received or skipped over. (All his computations are on-line.) The vendor needs to store all commitments and the last payment received per commitment each day.

3.5 Variations and Extensions

In one variation, $h(\cdot)$ is replaced by $h_s(\cdot) = h(s, \cdot)$, where s is a "salt" (random value) specified in the commitment. Salting may enable the use of faster hash functions or hash functions with a shorter output length (perhaps as short as 64–80 bits).

The value of each payword might be fixed at one cent, or might be specified in C_U or M. In a variation, M might authenticate several chains, whose paywords have different values (for penny paywords, nickel paywords, etc.).

The user name may also need to be specified in a payment if it is not clear from context. If U has more than one payword chain authorized for V, then the payment should specify which is relevant.

Paywords could be sold on a debit basis, rather than a credit basis, but only if the user interacts with the broker to produce each commitment: the certificate could require that the broker, rather than the user, sign each commitment. The broker can automatically refund the user for unused paywords, once the vendor has redeemed the paywords given to him.

In some cases, for macropayments, it might be useful to have the "commitment" act like an electronic credit card order or check without paywords being used at all. The commitment would specify the vendor and the amount to be paid.

The broker may specify in user certificates other terms and conditions to limit his risk. For example, B may limit the amount that U can spend per day at any vendor. Or, B may refuse payment if U's name is on B's "hot list" at the beginning of the day. (Vendors can down-load B's hot-list each morning.) Or, B may refuse to pay if U's total expenditures over all vendors exceeds a specified limit per day. This protects B from extensive liability if SK_U is stolen and abused. (Although again, since C_U only authorizes delivery to A_U, risk is reduced.) In these cases vendors share the risk with B.

Instead of using payword chains, another method we considered for improving efficiency was to have V *probabilistically* select payments for redemption. We couldn't make this idea work out, and leave this approach as an open problem.

4 MicroMint

MicroMint is designed to provide reasonable security at very low cost, and is optimized for unrelated low-value payments. MicroMint uses *no* public-key operations at all.

MicroMint "coins" are produced by a broker, who sells them to users. Users give these coins to vendors as payments. Vendors return coins to the broker in return for payment by other means.

A coin is a bit-string whose validity can be easily checked by anyone, but which is hard to produce. This is similar to the requirements for a public-key signature, whose complexity makes it an overkill for a transaction whose value is one cent. (PayWord uses signatures, but not on every transaction.)

MicroMint has the property that generating many coins is very much cheaper, per coin generated, than generating few coins. A large initial investment is required to generate the first coin, but then generating additional coins can be made progressively cheaper. This is similar to the economics for a regular mint, which invests in a lot of expensive machinery to make coins economically. (It makes no sense for a forger to produce coins in a way that costs more per coin produced than its value.)

The broker will typically issue new coins at the beginning of each month; the validity of these coins will expire at the end of the month. Unused coins are returned to the broker at the end of each month, and new coins can be purchased at the beginning of each month. Vendors can return the coins they collect to the broker at their convenience (e.g. at the end of each day).

We now describe the "basic" variant of MicroMint. Many extensions and variations are possible on this theme; we describe some of them in section 4.2.

Hash Function Collisions

MicroMint coins are represented by *hash function collisions*, for some specified one-way hash function h mapping m-bit strings x to n-bit strings y. We say that x is a pre-image of y if $h(x) = y$. A pair of distinct m-bit strings (x_1, x_2) is called a *(2-way) collision* if $h(x_1) = h(x_2) = y$, for some n-bit string y.

If h acts "randomly," the only way to produce even one acceptable 2-way collision is to hash about $\sqrt{2^n} = 2^{n/2}$ x-values and search for repeated outputs. This is essentially the "birthday paradox." (We ignore small constants in our analyses.)

Hashing c times as many x-values as are needed to produce the first collision results in approximately c^2 as many collisions, for $1 \leq c \leq 2^{n/2}$, so producing collisions can be done increasingly efficiently, per coin generated, once the threshold for finding collisions has been passed.

Coins as k-way collisions

A problem with 2-way collisions is that choosing a value of n small enough to make the

broker's work feasible results in a situation where coins can be forged a bit too easily by an adversary. To raise the threshold further against would-be forgers, we propose using k-way collisions instead of 2-way collisions.

A k-way collision is a set of k distinct x-values $x_1, x_2, \ldots, x_k$ that have the same hash value y. The number of x-values that must be examined before one expects to see the first k-way collision is then approximately $2^{n(k-1)/k}$. If one examines c times this many x-values, for $1 \leq c \leq 2^{n/k}$, one expects to see about c^k k-way collisions. Choosing $k > 2$ has the dual effect of delaying the threshold where the first collision is seen, and also accelerating the rate of collision generation, once the threshold is passed.

We thus let a k-way collision $(x_1, \ldots, x_k)$ represent a coin. The validity of this coin can be easily verified by anyone by checking that the x_i's are distinct and that

$$h(x_1) = h(x_2) = \cdots = h(x_k) = y$$

for some n-string y.

Minting coins

The process of computing $h(x) = y$ is analogous to tossing a ball (x) at random into one of 2^n bins; the bin that ball x ends up in is the one with index y. A coin is thus a set of k balls that have been tossed into the same bin. Getting k balls into the same bin requires tossing a substantial number of balls altogether, since balls can not be "aimed" at a particular bin. To mint coins, the broker will create 2^n bins, toss approximately $k2^n$ balls, and create one coin from each bin that now contains at least k balls. With this choice of parameters each ball has a chance of roughly $1/2$ of being part of a coin.

Whenever one of the 2^n bins has k or more balls in it, k of those balls can be extracted to form a coin. Note that if a bin has more than k balls in it, the broker can in principle extract k-subsets in multiple ways to produce several coins. However, an adversary who obtains two different coins from the same bin could combine them to produce multiple new coins. Therefore, we recommend that *a MicroMint broker should produce at most one coin from each bin.* Following this rule also simplifies the Broker's task of detecting multiply-spent coins, since he needs to allocate a table of only 2^n bits to indicate whether a coin with a particular n-bit hash value has already been redeemed.

A small problem in this basic picture, however, is that computation is much cheaper than storage. The number of balls that can be tossed into bins in a month-long computation far exceeds both the number of balls that can be memorized on a reasonable number of hard disks and the number of coins that the broker might realistically need to mint. One could attempt to balance the computation and memory requirements by utilizing a very slow hash algorithm, such as DES iterated many times. Unfortunately, this approach also slows down the verification process.

A better approach, which we adopt, is to make most balls unusable for the purpose of minting coins. To do so, we say that a ball is "good" if the high-order bits of the hash value y have a value z specified by the broker. More precisely, let $n = t + u$ for some specified nonnegative integers t and u. If the high-order t bits of y are equal to the specified value z then the value y is called "good," and the low-order u bits of y determine the index of the bin into which the (good) ball x is tossed. (General x values are referred to merely as

"balls," and those that are not good can be thought of as having been conceptually tossed into nonexistent virtual bins that are "out of range.")

A proper choice of t enables us to balance the computational and storage requirements of the broker, without slowing down the verification process. It slows down the generation process by a factor of 2^t, while limiting the storage requirements of the broker to a small multiple of the number of coins to be generated. The broker thus tosses approximately $k2^n$ balls, memorizes about $k2^u$ good balls that he tosses into the 2^u bins, and generates from them approximately $(1/2) \cdot 2^u$ valid coins.

Remark: We note that with standard hash functions, such as MD5 and DES, the number of ouput bits produced may exceed the number n of bits specified in the broker's parameters. A suitable hash function for the broker can be obtained by discarding all but the low-order n bits of the standard hash function output. This discarding of bits other than the low-order n bits is a different process than that of specifying a particular value for the high-order t bits out of the n that was described above.

A detailed scenario

Here is a detailed sketch of how a typical broker might proceed to choose parameters for his minting operating for a given month. The calculations are approximate (values are typically rounded to the nearest power of two), but instructive; they can be easily modified for other assumptions.

The broker will invest in substantial hardware that gives him a computational advantage over would-be forgers, and run this hardware continuously for a month to compute coins valid for the next month. This hardware is likely to include many special-purpose chips for computing h efficiently.

We suppose that the broker wishes to have a net profit of $1 million per month (approximately 2^{27} cents/month). He charges a brokerage fee of 10%. That is, for every coin worth one cent that he sells, he only gives the vendor 0.9 cents when it is redeemed. Thus, the broker needs to sell one billion coins per month (approximately 2^{30} coins/month) to collect his $1M fee. If an average user buys 2500 ($25.00) coins per month, he will need to have a customer base of 500,000 customers.

The broker chooses $k = 4$; a coin will be a good 4-way collision.

To create 2^{30} coins, the broker chooses $u = 31$, so that he creates an array of 2^{31} (approximately two billion) bins, each of which can hold up to 4 x-values that hash to an n-bit value that is the concantenation of a fixed t-bit pattern z and the u-bit index of the bin.

The broker will toss an average of 4 balls into each bin. That is, the broker will generate $4 \cdot 2^{31} = 2^{33}$ (approximately eight billion) x-values that produce good y-values. When he does so, the probability that a bin then contains 4 or more x-values (and thus can yield a coin) is about 1/2. (Using a Poisson approximation, it can be calculated that the correct value is approximately 0.56.) Since each of the 2^{31} bins produces a coin with probability 1/2, the number of coins produced is 2^{30}, as desired.

In order to maximize his advantage over an adversary who wishes to forge coins, the broker invests in special-purpose hardware that allows him to compute hash values very quickly. This will allow him to choose a relatively large value of t, so that good hash values are relatively rare. This increases the work factor for an adversary (and for the broker) by a

factor of 2^t. The broker chooses his hash function h as the low-order n bits of the encryption of some fixed value v_0 with key x under the Data Encryption Standard (DES):

$$h(x) = [DES_x(v_0)]_{1\ldots n} .$$

The broker purchases a number of field-programmable gate array (FPGA) chips, each of which is capable of hashing approximately 2^{25} (approximately 30 million) x-values per second. (See [3].) Each such chip costs about $200; we estimate that the broker's actual cost per chip might be closer to $400 per chip when engineering, support, and associated hardware are also considered. The broker purchases 2^8 (= 256) of these chips, which costs him about $100,000. These chips can collectively hash 2^{33} (approximately 8.6 billion) values per second. Since there are roughly 2^{21} (two million) seconds in a month, they can hash about 2^{54} (approximately 18 million billion) values per month.

Based on these estimates the broker chooses $n = 52$ and $t = 21$ and runs his minting operation for one month. Of the $k2^n = 2^{54}$ hash values computed, only one in 2^{21} will be good, so that approximately 2^{33} good x-values are found, as necessary to produce 2^{30} coins.

Storing a good $(x, h(x))$ pair takes less than 16 bytes. The total storage required for all good pairs is less than 2^{37} bytes (128 Gigabytes). Using standard magnetic hard disk technology costing approximately $300 per Gigabyte, the total cost for storage is less than $40,000. The total cost for the broker's hardware is thus less than $150,000, which is less than 15% of the first month's profit.

Rather than actually writing each pair into a randomly-accessible bin, the broker can write the 2^{33} good pairs sequentially to the disk array, and then sort them into increasing order by y value, to determine which are in the same bin. With a reasonable sorting algorithm, the sorting time should be under one day.

Selling coins

Towards the end of each month, the broker begins selling coins to users for the next month. At the beginning of each month, B reveals the new validity criterion for coins to be used that month. Such sales can either be on a debit basis or a credit basis, since B will be able to recognize coins when they are returned to him for redemption. In a typical purchase, a user might buy $25.00 worth of coins (2500 coins), and charge the purchase to his credit card. The broker keeps a record of which coins each user bought. Unused coins are returned to the broker at the end of each month.

Making payments

Each time a user purchases a web page, he gives the vendor a previously unspent coin $(x_1, x_2, \ldots, x_k)$. (This might be handled automatically by the user's web browser when the user clicks on a link that has a declared fee.) The vendor verifies that it is indeed a good k-way collision by computing $h(x_i)$ for $1 \le i \le k$, and checking that the values are equal and good. Note that while the broker's minting process was intentionally slowed down by a factor of 2^t, the vendor's task of verifying a coin remains extremely efficient, requiring only k hash computations and a few comparisons (in our proposed scenario, $k = 4$).

Redemptions

The vendor returns the coins he has collected to the broker at the end of each day. The broker checks each coin to see if it has been previously returned, and if not, pays the vendor

one cent (minus his brokerage fee) for each coin. We propose that if the broker receives a specific coin more than once, he does not pay more than once. Which vendor gets paid can be decided arbitrarily or randomly by the broker. This may penalize vendors, but eliminates any financial motivation a vendor might have had to cheat by redistributing coins he has collected to other vendors.

4.1 Security Properties

We distinguish between small-scale attacks and large-scale attacks. We believe that users and vendors will have little motivation to cheat in order to gain only a few cents; even if they do, the consequences are of no great concern. This is similar to the way ordinary change is handled: many people don't even bother to count their change following a purchase. Our security mechanisms are thus primarily designed to discourage large-scale attacks, such as massive forgery or persistent double-spending.

Forgery

Small-scale forgery is too expensive to be of interest to an adversary: with the recommended choice of $k = 4$, $n = 54$, and $u = 31$, the generation of the first forged coin requires about 2^{45} hash operations. Since a standard work-station can perform only 2^{14} hash operations per second, a typical user will need 2^{31} seconds (about 80 years) to generate just one forged coin on his workstation.

Large-scale forgery can be detected and countered as follows:

- All forged coins automatically become invalid at the end of the month.
- Forged coins can not be generated until after the broker announces the new monthly coin validity criterion at the beginning of the month.
- The use of hidden predicates (described below) gives a finer time resolution for rejecting forged coins without affecting the validity of legal coins already in circulation.
- The broker can detect the presence of a forger by noting when he receives coins correspondings to bins that he did not produce coins from. This works well in our scenario since only about half of the bins produce coins. To implement this the broker need only work with a bit-array having one bit per bin.
- The broker can at any time declare the current period to be over, recall all coins for the current period, and issue new coins using a new validation procedure.
- The broker can simultaneously generate coins for several future months in a longer computation, as described below; this makes it harder for a forger to catch up with the broker.

Theft of coins

If theft of coins is judged to be a problem during initial distribution to users or during redemption by vendors, it is easy to transmit coins in encrypted form during these operations.

User/broker and vendor/broker relationships are relatively stable, and long-term encryption keys can be arranged between them.

To protect coins as they are being transferred over the Internet from user to vendor, one can of course use public-key techniques to provide secure communication. However, in keeping with our desire to minimize or eliminate public-key operations, we propose below another mechanism, which makes coins user-specific. This does not require public-key cryptography, and makes it harder to re-use stolen coins.

Another concern is that two vendors may collude so that both attempt to redeem the same coins. The recommended solution is that a broker redeem a coin at most once, as discussed earlier. Since this may penalize honest vendors who receive stolen coins, we can make coins vendor-specific as well as user-specific, as described below.

Double-spending

Since the MicroMint scheme is not anonymous, the broker can detect a doubly-spent coin, and can identify which vendors he received the two instances from. He also knows which user the coin was issued to. With the vendors' honest cooperation, he can also identify which users spent each instance of that coin. Based on all this information, the broker can keep track of how many doubly-spent coins are asssociated with each user and vendor. A large-scale cheater (either user or vendor) can be identified by the large number of duplicate coins associated with his purchases or redemptions; the broker can then drop a large-scale cheater from the system. A small-scale cheater may be hard to identify, but, due to the low value of individual coins, it is not so important if he escapes identification.

MicroMint does not provide any mechanism for preventing purely malicious framing (with no financial benefit to the framer). We believe that the known mechanisms for protecting against such behavior are too cumbersome for a light-weight micropayment scheme. Since MicroMint does not use real digital signatures, it may be hard to legally prove who is guilty of duplicating coins. Thus, a broker will not be able to pursue a cheater in court, but can always drop a suspected cheater from the system.

4.2 Variations

User-specific coins

We describe two proposals for making coins that are user-specific in a way that can be easily checked by vendors. Such coins, if stolen, are of no value to most other users. This greatly reduces the motivation for theft of coins.

In the first proposal, the broker splits the users into "groups," and gives each user coins whose validity depends on the identity of the group. For example, the broker can give user U coins that satisfy the additional condition $h'(x_1, x_2, \ldots, x_k) = h'(U)$, where hash function h' produces short (e.g. 16-bit) output values that indicate U's group. A vendor can easily check this condition, and reject a coin that is not tendered by a member of the correct group.

The problem with this approach is that if the groups are too large, then a thief can easily find users of the appropriate group who might be willing to buy stolen coins. On the other hand, if the groups are too small (e.g. by placing each user is in his own group), the broker may be forced to precompute a large excess of coins, just to ensure that he has a large enough

supply to satisfy each user's unpredictable needs.

In the second proposal, we generalize the notion of a "collision" to more complicated combinatorial structures. Formally, a coin $(x_1, \ldots, x_k)$ will be valid for a user U if the images $y_1 = h(x_1)$, $y_2 = h(x_2)$, $\ldots$, $y_k = h(x_k)$ satisfy the condition

$$y_{i+1} - y_i = d_i \pmod{2^n}$$

for $i = 1, 2, \ldots, k-1$, where

$$(d_1, d_2, \ldots, d_{k-1}) = h'(U)$$

for a suitable auxiliary hash function h'. (The original proposal for representing coins as collisions can be viewed as the special case where all the distances d_i's between the k bins are zero.)

To mint coins of this form, the broker fills up most of his bins by randomly tossing balls into them, except that now it is not necessary to have more than one ball per bin. We emphasize that this pre-computation is not user-specific, and the broker does not need to have any prior knowledge of the number of coins that will be requested by each user, since each good ball can be used in a coin for *any* user. After this lengthy pre-computation, the broker can quickly create a coin for any user U by

- Computing $(d_1, \ldots, d_{k-1}) = h'(U)$.
- Picking a random bin index y_1. (This bin should have been previously unused as a y_1 for another coin, so that y_1 can be used as the "identity" of the coin when the broker uses a bit-array to determine which coins have already been redeemed.)
- Computing $y_{i+1} = y_i + d_i \pmod{2^n}$ for $i = 1, 2, \ldots, k-1$.
- Taking a ball x_1 out of bin y_1, and taking a copy of one ball out of each bin $y_2, \ldots, y_k$. (If any bin y_i is empty, start over with a new y_1.) Note that balls may be re-used in this scheme.
- Producing the ordered k-tuple $(x_1, \ldots, x_k)$ as the output coin.

A convenient feature of this scheme is that it is easy to produce a large number of coins for a given user even when the broker's storage device is a magnetic disk with a relatively slow seek time. The idea is based on the observation that if the y_1 values for successive coins are consecutive, then so also will be the y_i values for each i, $1 < i \leq k$. Therefore, a request for 2500 new coins with $k = 4$ requires only four disk seeks, rather than 10,000 seeks: at 10 milliseconds per seek, this reduces the total seek time from 100 seconds to only 40 milliseconds.

Note that in principle coins produced for different users could re-use the same ball x_i. Conceivably, someone could forge a new coin by combining pieces of other coins he has seen. However, he is unlikely to achieve much success by this route unless he sees balls from a significant fraction of all the bins. For example, suppose that there are 2^{31} bins, of which the forger has seen a fraction 2^{-10} (i.e., he has collected 2^{21} balls from coins spent by other users). Then the expected number of coins he can piece together from these balls that satisfy

the condition of being a good coin for himself is only $2^{31}(2^{-10})^3 = 2$. (Even if he had 1000 customers for these coins, he would expect to make only 2000 coins total, or two coins per customer on the average.) Thus, we are not too concerned about this sort of "cut-and-paste" forgery.

Vendor-specific coins

To further reduce the likelihood that coins will be stolen, the user can give coins to vendors in such a way that each coin can be redeemed only by a small fraction of the vendors. This technique makes a stolen coin less desirable, since it is unlikely to be accepted by a vendor other than the one where it was originally spent. The additional check of validity can be carried out both by the vendor and by the broker. (Having vendor-specific coins is also a major feature of the Millicent [10] scheme.)

The obvious difficulty is that neither the broker nor the user can predict ahead of time which vendors the user will patronize, and it is unreasonable to force the user to purchase in advance coins specific for each possible vendor. Millicent adopts the alternative strategy whereby the user must contact the broker in real-time whenever the user needs coins for a new vendor. (He also needs to contact the broker to return excess unused coins that are specific to that vendor.) We can overcome these problems with an extension of the user-specific scheme described above, in which the user purchases a block of "successive" MicroMint coins.

Intuitively, the idea is the following. Choose a value v (e.g. 1024) less than u. Let a u-bit bin-index y be divided into a $u-v$-bit upper part y' and a v-bit lower part y''. We consider that y' specifies a "superbin" index and that y'' specifies a bin within that superbin. A user now purchases balls in bulk and makes his own coins. He purchases balls by the superbin, obtaining 2^v balls per superbin with one ball in each bin of the superbin. He buy k superbins of balls for 2^v cents. A coin from user U is valid for redemption by vendor V if:

$$y'_{i+1} = y'_i + d'_i \pmod{2^{u-v}} \text{ for } i = 1, \ldots, k-1,$$

and

$$y''_{i+1} = y''_i + d''_i \pmod{2^v} \text{ for } i = 1, \ldots, k-1,$$

where

$$h'(U) = (d'_1, \ldots, d'_{k-1})$$

and

$$h''(V) = (d''_1, \ldots, d''_{k-1}) .$$

The broker chooses the next available superbin as the first superbin to give the user; the other superbins are then uniquely determined by the differences $\{d'_i\}$ defined by the user's identity and the choice of the first superbin. Analogously, to make a coin for a particular vendor the user chooses a ball from the next bin from his first superbin, and must use balls from bins in the other superbins that are then uniquely determined by the differences $\{d''_i\}$ defined by the vendor's identity and the choice of the first bin. Note that balls from the first superbin are used only once, to permit detection of double-spending, whereas balls from the other superbins may appear more than once (in coins paid to different vendors), or not at all. It may be difficult for a broker to create superbins that are perfectly full even if he

throws more balls. He might sell superbins that are almost full, but then a user may have difficulty producing some coins for some vendors. To compensate, the broker can reduce the price by one cent for each empty bin sold.

Simultaneously generating balls for multiple months

Our major line of defense against large-scale forgery is the fact that the broker can compute coins in advance, whereas a forgery attempt can only be started once the new validity condition for the current month is announced. We now describe a technique whereby computing the balls for a single month's coins takes eight months, but the broker doesn't fall behind because he can generate balls for eight future months concurrently. The forger will thus have the dual problems of starting late and being too slow, even if he uses the same computational resources as the real broker.

In this method, the broker changes the monthly validity criterion, not by changing the hash function h, but by announcing each month a new value z such that ball x is good when the high-order t bits of $h(x)$ are equal to z. The broker randomly and secretly chooses in advance the values z that will be used for each of the next eight months. Tossing a ball still means performing one hash function computation, but the tossed ball is potentially "good" for any of the next eight months, and it is trivial for the broker to determine if this is the case. In contrast, the forger only knows the current value of z, and can not afford to memorize all the balls he tosses, since memory is relatively expensive and only a tiny fraction (e.g., 2^{-24} in our running example) of the balls are considered "good" at any given month.

We now describe a convenient way of carrying out this calculation. Assume that at the beginning of the month j, the broker has all of the balls needed for month j, 7/8 of the balls needed for month $j+1$, 6/8 of the balls needed for month $j+2$, ..., and 1/8 of the balls needed in for month $j+7$. During month j, the broker tosses balls by randomly picking x values, calculating $y = h(x)$, and checking whether the top-most t bits of y are equal to any of the z values to be used in months $j+1, \ldots, j+8$. To slow the rate at which he generates good balls for each upcoming month, he increases n and t each by three. After the month-long computation, we expect him to have all the coins he needs for month $j+1$, 7/8 of the coins he needs for month $j+2$, and so on; this is the desired "steady-state" situation. The broker needs four times as much storage to hold the balls generated for future months, but balls for future months can be temporarily stored on inexpensive magnetic tapes because he doesn't need to respond quickly to user requests for those coins yet.

Hidden Predicates

The "hidden predicate" technique for defeating forgers works as follows. We choose $m > n$, and require each m-bit pre-image to satisfy a number of hidden predicates. The hidden predicates should be such that generating pre-images satisfying the predicates is easy (if you know the predicate). To generate an x_i, one can pick its last n bits randomly, and define the j-th bit of x_i, for $j = m-n, \ldots, 1$, to be the j-th hidden predicate applied to bits $j+1, \ldots, m$ of x_i. The hidden predicates must be balanced and difficult to learn from random examples. Suggestions of hard-to-learn predicates exist in the learning-theory literature. For example the parity/majority functions of Blum et al.[4] (which are the exclusive-or of some of the input bits together with the majority function on a disjoint set of input bits) are interesting, although slightly more complicated functions may be appropriate in this application when word lengths are short. With $m-n = 32$, the broker can have one hidden

predicate for each day of the month. He could reveal a new predicate each day, and ask vendors to check that the coins they receive satisfy these predicates (otherwise the coins will not be accepted by the broker). This would not affect the validity of legitimate coins already in circulation, but makes forgery extremely difficult, since the would-be forger would have to discard much of his precomputation work as each new predicate is revealed. We feel that such techniques are strongly advisable in MicroMint.

Other Extensions

Peter Wayner (private communication) has suggested a variation on MicroMint in which coins of different values are distinguished by publicly-known predicates on the x-values.

5 Relationship to Other Micropayment Schemes

In this section we compare our proposals to the Millicent[10], NetBill [1], NetCard [2], and Pederson [14] micropayment schemes.

NetBill offers a number of advanced features (such as electronic purchase orders and encryption of purchased information), but it is relative expensive: digital signatures are heavily used and the NetBill server is involved in each payment.

Millicent uses hash functions extensively, but the broker must be on-line whenever the user wishes to interact with a new vendor. The user buys vendor-specific scrip from the broker. For applications such as web browsing, where new user-vendor relationships are continually being created, Millicent can place a heavy real-time burden on the broker. Compared to Millicent, both PayWord and MicroMint enable the user to generate vendor-specific "scrip" without any interaction with the broker, and without the overhead required in returning unused vendor-specific scrip. Also, PayWord is a credit rather than debit scheme.

Anderson, Manifavas, and Sutherland [2] have developed a micropayment system, "NetCard," which is very similar to PayWord in that it uses chains of hash values with a digitally signed root. (The way hash chains are created differs in a minor way.) However, in their proposal, it is the bank rather than the user who prepares the chain and signs the root, which adds to the overall burden of the bank. This approach prevents the user from creating new chains, although a NetCard user could spend a single chain many times. Compared to PayWord, NetCard is debit-based, rather than credit-based. We have heard that a patent has been applied for on the NetCard system.

Torben Pedersen outlines a micropayment proposal[14] that is also based on hash chains. His motivating application was for incremental payment of telephone charges. His paper does not provide much detail on many points (e.g. whether the system is credit or debit-based, how to handle exceptions, whether chains are vendor-specific, and other auxiliary security-related matters). The CAFE project has filed for a patent on what we believe is an elaboration of Pedersen's idea. (The details off the CAFE scheme are not available to us.)

Similarly following Pedersen's exposition, the iKP developers Hauser, Steiner, and Waidner have independently adopted a similar approach [8].

6 Conclusions and Discussion

We have presented two new micropayment schemes which are exceptionally economical in terms of the number of public-key operations employed. Furthermore, both schemes are *off-line* from the broker's point of view.

References

[1] The NetBill Electronic Commerce Project, 1995. `http://www.ini.cmu/NETBILL/home.html`.

[2] Ross Anderson, Harry Manifavas, and Chris Sutherland. A practical electronic cash system, 1995. Available from author: `Ross.Anderson@cl.cam.ac.uk`.

[3] Matt Blaze, Whitfield Diffie, Ronald L. Rivest, Bruce Schneier, Tsutomu Shimomura, Eric Thompson, and Michael Wiener. Minimal key lengths for symmetric ciphers to provide adequate commercial security: A report by an ad hoc group of cryptographers and computer scientists, January 1996. Available at `http://www.bsa.org`.

[4] Avrim Blum, Merrick Furst, Michael Kearns, and Richard J. Lipton. Cryptographic primitives based on hard learning problems. In Douglas R. Stinson, editor, *Proc. CRYPTO 93*, pages 278–291. Springer, 1994. Lecture Notes in Computer Science No. 773.

[5] Shimon Even, Oded Goldreich, and Silvio Micali. On-line/off-line digital signatures. In G. Brassard, editor, *Proc. CRYPTO 89*, pages 263–277. Springer-Verlag, 1990. Lecture Notes in Computer Science No. 435.

[6] Phillip Hallam-Baker. W3C payments resources, 1995. `http://www.w3.org/hypertext/WWW/Payments/overview.html`.

[7] Neil M. Haller. The S/KEY one-time password system. In *ISOC*, 1994.

[8] Ralf Hauser, Michael Steiner, and Michael Waidner. Micro-Payments based on iKP, December 17, 1995. Available from authors. `sti@zurich.ibm.com`.

[9] Leslie Lamport. Password authentication with insecure communication. *Communications of the ACM*, 24(11):770–771, November 1981.

[10] Mark S. Manasse. Millicent (electronic microcommerce), 1995. `http://www.research.digital.com/SRC/personal/Mark_Manasse/uncommon/ucom.html`.

[11] Ralph C. Merkle. A certified digital signature. In G. Brassard, editor, *Proc. CRYPTO 89*, pages 218–238. Springer-Verlag, 1990. Lecture Notes in Computer Science No. 435.

[12] Silvio Micali. Efficient certificate revocation. Technical Report TM-542b, MIT Laboratory for Computer Science, March 22, 1996.

[13] National Institute of Standards and Technology (NIST). *FIPS Publication 180: Secure Hash Standard (SHS)*, May 11, 1993.

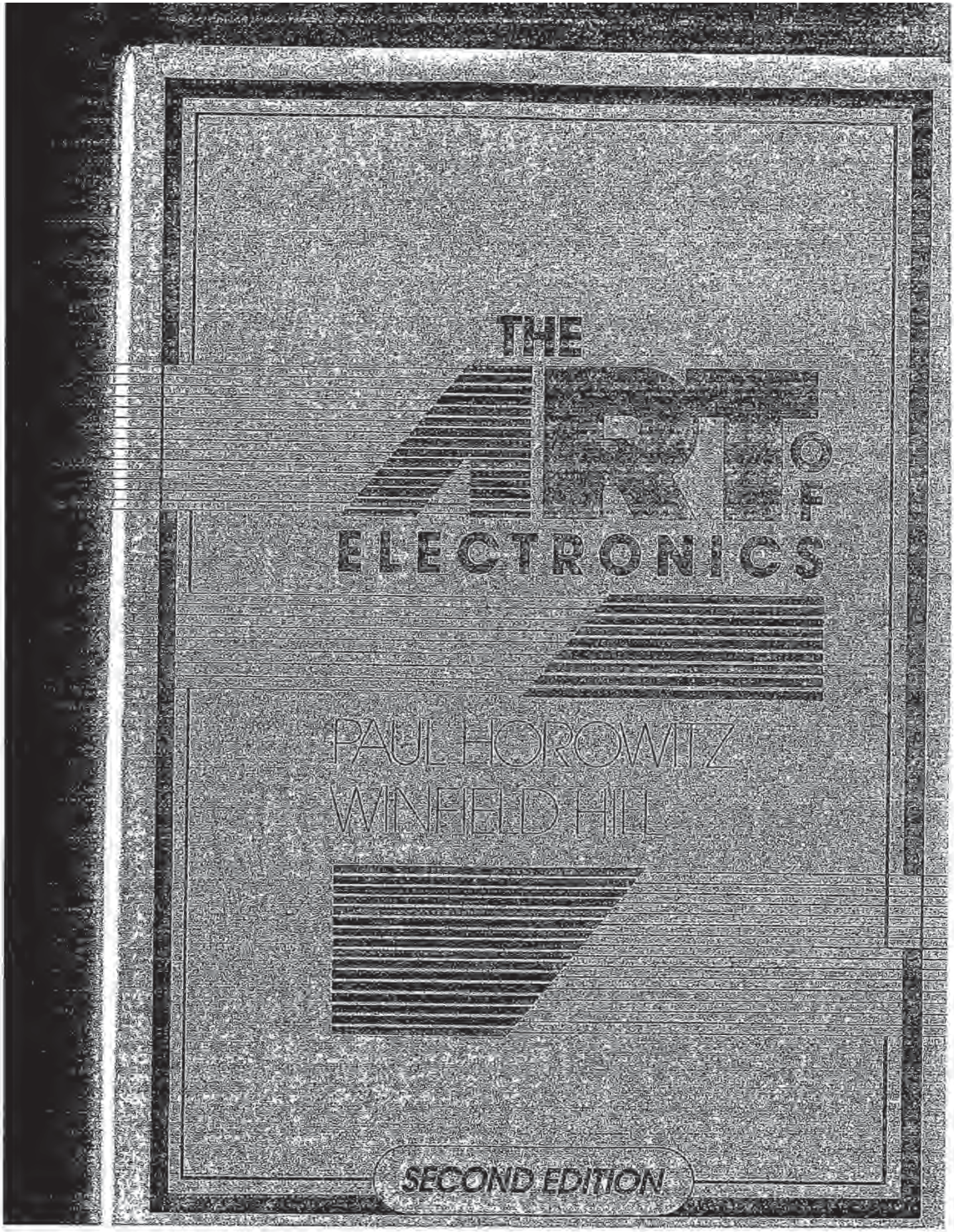
[14] Torben P. Pedersen. Electronic payments of small amounts. Technical Report DAIMI PB-495, Aarhus University, Computer Science Department, Århus, Denmark, August 1995.

[15] Ronald L. Rivest. The MD5 message-digest algorithm. Internet Request for Comments, April 1992. RFC 1321.

[16] Bruce Schneier. *Applied Cryptography (Second Edition)*. John Wiley & Sons, 1996.

[17] Adi Shamir. Fast signature screening. CRYPTO '95 rump session talk; to appear in RSA Laboratories' *CryptoBytes*.

[18] Peter Wayner. *Digital Cash: Commerce on the Net*. Academic Press, 1996.


THE
ART OF
ELECTRONICS
PAUL HOROWITZ
WINFIELD HILL
SECOND EDITION

Some more home-grown philosophy: There is a tendency among beginners to want to compute resistor values and other circuit component values to many significant places, and the availability of inexpensive calculators has only made matters worse. There are two reasons you should try to avoid falling into this habit: (a) the components themselves are of finite precision (typical resistors are ±5%; the parameters that characterize transistors, say, frequently are known only to a factor of two); (b) one mark of a good circuit design is insensitivity of the finished circuit to precise values of the components (there are exceptions, of course). You'll also learn circuit intuition more quickly if you get into the habit of doing approximate calculations in your head, rather than watching meaningless numbers pop up on a calculator display.

In trying to develop intuition about resistance, some people find it helpful to think about *conductance*, $G = 1/R$. The current through a device of conductance G bridging a voltage V is then given by $I = GV$ (Ohm's law). A small resistance is a large conductance, with correspondingly large current under the influence of an applied voltage.

Viewed in this light, the formula for parallel resistors is obvious: When several resistors or conducting paths are connected across the same voltage, the total current is the sum of the individual currents. Therefore the net conductance is simply the sum of the individual conductances, $G = G_1 + G_2 + G_3 + \cdots$, which is the same as the formula for parallel resistors derived earlier.

Engineers are fond of defining reciprocal units, and they have designated the unit of conductance the siemens ($S = 1/\Omega$), also known as the mho (that's ohm spelled backward, given the symbol $\mho$). Although the concept of conductance is helpful in developing intuition, it is not used widely; most people prefer to talk about resistance instead.

Power in resistors

The power dissipated by a resistor (or any other device) is $P = IV$. Using Ohm's law, you can get the equivalent forms $P = I^2R$ and $P = V^2/R$.

EXERCISE 1.5

Show that it is not possible to exceed the power rating of a 1/4 watt resistor of resistance greater than 1k, no matter how you connect it, in a circuit operating from a 15 volt battery.

EXERCISE 1.6

Optional exercise: New York City requires about 10^{10} watts of electrical power, at 110 volts (this is plausible: 10 million people averaging 1 kilowatt each). A heavy power cable might be an inch in diameter. Let's calculate what will happen if we try to supply the power through a cable 1 foot in diameter made of pure copper. Its resistance is $0.05\mu\Omega$ (5×10^{-8} ohms) per foot. Calculate (a) the power lost per foot from "I^2R losses," (b) the length of cable over which you will lose all 10^{10} watts, and (c) how hot the cable will get, if you know the physics involved ($\sigma = 6 \times 10^{-12}\text{W/°K}^4\text{cm}^2$).

If you have done your computations correctly, the result should seem preposterous. What is the solution to this puzzle?

Input and output

Nearly all electronic circuits accept some sort of applied *input* (usually a voltage) and produce some sort of corresponding *output* (which again is often a voltage). For example, an audio amplifier might produce a (varying) output voltage that is 100 times as large as a (similarly varying) input voltage. When describing such an amplifier, we imagine measuring the output voltage for a given applied input voltage. Engineers speak of the *transfer function* H, the ratio of (measured) output divided by (applied) input; for the audio amplifier above, H is simply a constant (H = 100). We'll get to amplifiers soon enough, in the next chapter. However, with just resistors we can already look at a very important circuit fragment, the *voltage divider* (which you might call a "de-amplifier").

Digital watermarking

J.-F. Delaigle, C. De Vleeschouwer, B. Macq

Laboratoire de Télécommunications et Télédétection
Université catholique de Louvain
Bâtiment Stévin - 2, place du Levant
B-1348 Louvain-la-Neuve
Tel.: +32 10 47.80.72 - Fax: +32 10 47.20.89
E-mail: delaigle@tele.ucl.ac.be

ABSTRACT

This paper presents a process able to mark digital pictures with an invisible and undetectable secrete information, called the watermark. This process can be the basis of a complete copyright protection system.
The process first step consists in producing a secrete image . The first part of the secret resides in a basic information that forms a binary image. That picture is then frequency modulated. The second part of the secret is precisely the frequencies of the carriers. Both secrets depends on the identity of the copyright owner and on the original picture contents. The obtained picture is called the stamp.
The second step consists in modulating the ampitude of the stamp according to a masking criterion stemming from a model of human perception. That too theoretical criterion is corrected by means of morphological tools helping to locate in the picture the places where the criterion is supposed not to match.
This is followed by the adaptation of the level of the stamp at that places. The so formed watermark is then added to the original to ensure its protection.
That watermarking method allows the detection of watermarked pictures in a stream of digital images, only with the knowledge of the picture owner's secrets.

Keywords: copyright protection, watermark, secrete key, masking, human vision model, perceptive components, morphology, robustness, detection, correlation.

1 GENERAL INTRODUCTION

With the increasing availability of digitally stored information and the development of new multimedia services, security questions are becoming even more urgent. The acceptance of new services depends on whether suitable techniques for the protection of the work providers' interests are available.[1]

Moreover the nature of digital media threatens its own viability:

- First the replication of digital works is very easy and, what is more dangerous, really perfect. The copy is identical to the original.

0-8194-2033-6/96/$6.00

- The ease of transmission and multiple uses is very worrying, too. Once a single pirate copy has been made, it is instantaneously accessible to anyone who wants it, without any control of the original picture owner.
- Eventually the plasticity of digital media is a great menace. Any malevolent user *(a pirate)* can modify an image at will. Such maniplations are really easy for a pirate and put many copyright protection methods at risk.

According to these considerations the conception of a copyright protection system is really vital and it constitutes a great challenge, because it should cope with all these threats. Without watermarking, most authors will not dare to broadcast their work.

This paper presents an additive watermarking technique. It consits in producing a synthetic picture (also called the stamp) which holds informations about the ownership of the original image and depends on the picture contents. That stamp is added to the original in a way that resulting picture is perceptually identical to the original one and so that the stamp is indetectable by a pirate computer. The aim of that technique is not the authentication of the picture content nor the identication of the owner. It is to allow a controller (i.e. the owner's computer or a Trusted Third Part) to find out watermarked pictures in a stream of images with the knowledge of the owner's secret key in order to detect broadcast of illegal copies.

The most interesting part of that method is the embedding process i.e. the weighting of each pixels of the stamp before adding it to the original. This is based on the masking concept coming from a model of human vision (the perceptive model). From this concept was deduced a method which reveals itself actually efficient. Another interesting part is the presentation of two methods used for the detection of watermarked pictures without the original. This last point is fundemental for the management of the copyright protection.
Eventually this paper ends with the analyse of the results and the system robustness.

2 THE MASKING

2.1 Introduction

The aim of a watermarking technique is to provide an invisible embedding of a secrete information, the watermark. This watermark must be masked (hidden) by the picture it is inlayed in. Precisely a master thesis has lead to a masking criterion deduced from physiological and psychophysic studies.[2] Nevertheless, this theoretical criterion having been formulated for monochromatic signals, it had to be adaptated to suit real images.

2.2 The perceptive model: approximation of the eye functionment

It is now admitted that the retina of the eye splits an image in several components. These components circulate from the eye to the cortex by different tuned channels, one channel being tuned to one component.

The characteristics of one component are:

- the location in the visual field (in the image).
- the spatial frequency (in the Fourier domain: the amplitude in polar coordinates).
- the orientation (in the Fourier domain: the phase in polar coordinates)

So, one perceptive channel can only be excited by one component of a signal whose characteristics are tuned to its. Components that have different characteristics are independent.

2.3 The masking concept

According to perceptive model of human vision,[3] signals that have same (near) components take the same channels from the eye to the cortex. It appears that such signals interact and are submitted to non-linear effects. The masking is one of those effects.

Definition: *the detection threshold* is the minimum level below which a signal can not be seen.

Definition: *the masking* occurs when the detection threshold is increased because of the presence of another signal.

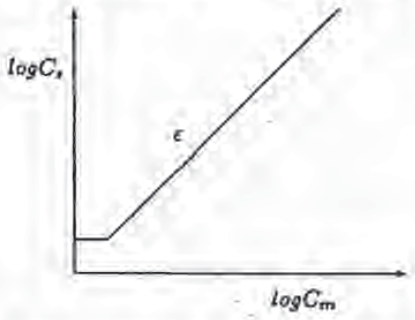
In other words, there is masking when a signal can not be seen because of another with near characteristics and at a higher level.

2.4 The masking model

With the object of modalizing the masking phenomenon, tests have been made on monochromatic signals, also called *gratings*. It appears that the eye is sensitive to the contrast of those gratings. This contrast is defined by:

$$C = \frac{2(Lmax - Lmin)}{Lmax + Lmin} \tag{1}$$

where L is the luminance.
It is possible to determine experimentally the detection threshold of one signal of contrast C_s with respect to the contrast C_m of the masking signal. That threshold can be modalized as follows:



Such bilogarithmic curves are traced for signals of one single frequency and one orientation (f_0, θ_0).
The expression of the detection threshold is thus:

$$C_s = max[C_0, C_0(\frac{C_m}{C_0})^\epsilon] \tag{2}$$

where ϵ (the slope) depends on (f_0, θ_0), typically, $0.6 \leq \epsilon \leq 1.1$.

It is possible to extend that expression to introduce frequency dependence. The general expression of the detection threshold is becomes:

$$C_s(C_m, f, \theta) = C_0 + k_{(f_0,\theta_0)}(f,\theta)[C_{s(f_0,\theta_0)}(C_m) - C_0] \quad (3)$$

where:

$$k_{(f_0,\theta_0)}(f,\theta) = exp[-(\frac{log^2(\frac{f}{f_0})}{F^2(f_0)} + \frac{(\theta-\theta_0)^2}{\Theta^2(f_0)}] \quad (4)$$

In that expression, f_0 and θ_0 are relevant to the masking signal, f and θ are relevant to the masked signal, $F(f_0)$ and $\Theta(f_0)$ are parameters that represent the spreading of the Gaussian function, C_0 is often negligable. The spread of the gaussian function depends upon the frequency f_0: For frequency, typical bandwith at half response are 2,5 octaves at 1 c/d and 1,5 octaves at 16 c/d with a linear decrease between both frequencies.[4] For orientation, half bandwith at half response depends on f_0 and it takes typical values like 30 degrees at 1 c/d and 15 degrees at 16 c/d.[5]

After this expression, the frequency dependence of the detection threshold has a Gaussian form. Only near frequency signals can interact. When the frequency of the masking signal (the mask) is far from this of the signal to mask, the detection threshold is almost equal to C_0.

2.5 The masking criterion

It is important to notice that those results concern only gratings signals. To deduce a masking criterion that will apply to signals like real images, the preceding masking condition has to be adaptated. So, it is necessary to define a new concept able to take the place of the contrast, because the contrast is not define for real images. That new concept,[2] is the *local energy*.

The local energy is defined on narrowband signals centered around one frequency and one orientation. A picture which is a broadband signal is first filtered by Gabor narrowband filters, whose characteristics are near to human perception. The local energy around one frequency and one orientation is calculated following the scheme presented in this figure:

The masking criterion: If the local energy of one picture is less than the local energy of the mask, around all the frequencies (f_0, θ_0) and for each pixel (x, y), then one can say that the picture is masked by the mask. Strictly, a picture is masked by a mask if $\forall(x, y)$ and $\forall(f_0, \theta_0)$, $E_{mask,(f_0,\theta_0)}(x, y) \geq E_{picture,(f_0,\theta_0)}(x, y)$. For real images, a good approximation of this criterion can be obtained by using a bank of filters whose central frequencies correspond to independent components and which are spread on all the Fourier space. It is admitted that 4 or 5 frequencies and 4 to 9 orientations are sufficient. The standard choice is twenty filters (5 frequencies and 4 orientations).

Figure 1: Example of basic information used

2.6 Conclusion

This section has lead to the expression of an easily implementable masking criterion appliable to any image. But this criterion is only an extension of a theoretic criterion appliable to monochromatic signals. Thus cases where that criterion does not match are possible.

3 PRINCIPLE OF THE SYSTEM

3.1 Basic information of the watermark

This information is a binary picture looking like a modified checkerboard (figure 1). As explained later, the pixels value of the square forming that picture can correspond to a binary sequence deduced from the copyright owner's (CO) *secrete key*.

3.2 The stamp

In order to take advantage of the eye behaviour, the basic information is modulated at different frequencies and orientations corresponding to rather independent components. Moreover, we take care to filter the initial checkerboard with a low pass filter (LPF) (i.e. a Butterworth LPF) so that the resulting signal is bandlimited. This point is very important because it permits to limit the verification of the masking criterion in the coresponding channel.
The position of the modulating carriers is *secret*. It can be deduced from CO's secret key. In practice, the frequency plan is divided into sectors. Each sector is relevant to one perceptive component and defined a group of couples (f, θ) where basic information can be modulated. Only one couple is chosen for each sector (because couples of a same sector don't stimulate independent components). The picture obtained from the sum of each modulated grid is called *the stamp* $S(x, y)$.

$$S(x, y) = \sum_{j \in K} G(x, y).cos(f_{x_j}.x + f_{y_j}.y) \tag{5}$$

K represents the set of sectors and (f_{x_j}, f_{y_j}) correspond to the couple chosen in sector j (this couple is designed by the CO's secrete key).

3.3 The position of the process in a global copyright scheme

The process should be placed in a copyright protection scheme like drawn at figure 2.
The skeletization function consists in an image processing program extracting essential characteristics from an image. The result is a bitsteam. This must be followed by a *hash-function*[6] whose result is a succession of blocks of bits. Every block has the same length. The skeletization function gives the same result for two near images (i.e. original image and watermarked image). But the H-function always gives different results from different bitstreams as inputs. So, the inscription keys will be different for perceptually distinct pictures. After the H-function, the ciphering function is a trapdoor function.[6] Thanks to this function the inscription keys used to deduce the basic grid and the position of the carriers depends on the CO's secret key.
The aim of the use of a trapdoor function is to prevent someone from reproducing the same inscription keys with the knowledge of the H-function result. But it is possible for anyone to inverse that trapdoor function and to find the H-function result from the inscription keys. It can be interesting in a proof procedure.

4 IMPLEMENTATION

4.1 Inscription

The purpose of the inscription is to adapt the level of each part of the stamp (for all frequencies) to make it invisible once added to the picture. As mentioned above, each part of the stamp is narrow band. Inscriptions at different frequencies are thus independent and one can treat the different components of the stamp one at a time. For each frequency designed by the inscription keys, the procedure is divided in three steps : the modulation, the regulation of the level and the correction.

- Modulation

  The first step consists in the modulation of the particular carrier by the lowpass grid $G(x,y)$. The result is $G(x,y).cos(f_{x_j}.x + f_{y_j}.y)$, where f_{x_j} and f_{y_j} are the carrier position.

- Regulation of the level

  According to the perceptual model, in order to guarantee the invisibility of the watermark its local energy has to be inferior to the picture local energy for each pixel around the inscription frequency. A way to reach this objective is to multiplicate the modulated grid by a weighting mask $Weight_j(x,y)$ reducing the amplitude of the stamp where energy in the correponding component of the original picture is weak. Nevertheless, one must take care to keep the narrow band characteristic of the resulting signal $S_j(x,y)$ ($= Weight_j(x,y).G(x,y).cos(f_{x_j}.x + f_{y_j}.y)$) in order to avoid non linear interactions between different parts of the stamp. In conclusion, $\forall j$, we have to find a signal $Weight_j(x,y)$ so that:

  - $\forall (x,y)\ E_{S_j}(x,y) < E_{I.(f_{x_j},f_{y_j})}(x,y)$
  - S_j is narrow band

  For simplification, lets consider $Weight_j(x,y)$ be composed of two factors:

  - α_j, a constant factor (fixing the global level of the stamp).
  - $M_j(x,y)$, a mask whose values $\in [0,1]$.

  When α_j is chosen, the way to find $M_j(x,y)$ so that $Weight_j(x,y)$ satisfy the conditions defined above is the following:

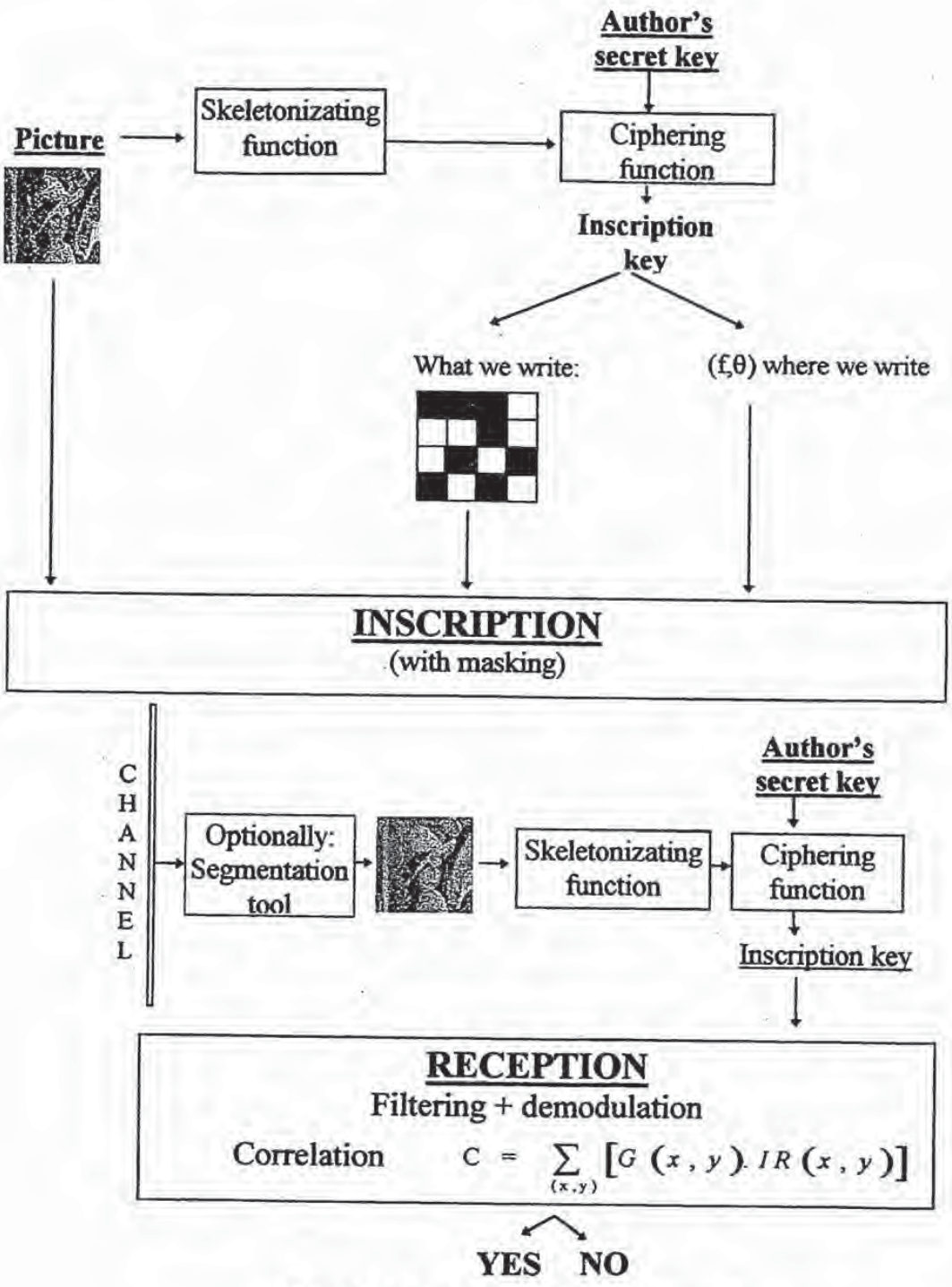
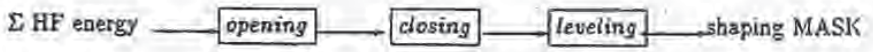


Figure 2: Global scheme for copyright protection

- Firstly, $M_j(x,y)$ is a binary mask. $M_j(x,y) = 1$ when the local energy of the stamp pemits the masking and $M_j(x,y) = 0$ when the local energy of the stamp is too important. It is obvious that the initial choice of α_j has a direct influence on $M_j(x,y)$. Indeed, a great α_j value will lead to put most of the $M_j(x,y)$ values to zero, while a small α_j value will lead to keep most of $M_j(x,y)$ values at one.
- Secondly, $Weight_j(x,y)$ is filtered so that the stamp remains narrow band.
- After this second step, one has found a signal $\alpha_j.M_j(x,y).G(x,y)$ which is better masked than $\alpha_j.G(x,y)$. In order to really satisfy the masking criterion $\forall(x,y)$, this procedure must be repeated iteratively, taking $M_j(x,y).G(x,y)$ as new $G(x,y)$. Experiments have shown that only two iterations are sufficient to have a result satisfying the masking criterion everywhere.

One important question remains: how to choose α_j?
It has already been said that the more α_j increases, the more $M_j(x,y)$ has points equal to zero. A trade off has to be found by means of a defined criterion. Maximizing the correlation at the detection (by maximizing $\sum \alpha_j.M_j(x,y).G(x,y)$) could have been a good criterion, but such a criterion often tends to impose an optimum with a lot of points equal to zero and a small number of points with a great value. The addition of the so obtained watermark generally entails a degradation of the picture quality. This emphasizes the lack of the masking criterion used.
As mentioned in section 2.6, the invisibility criterion used here is an extension for real images. It appears that this extension entails some imperfections. This criterion being insufficient, some improvements have been brought thanks to experimental results.

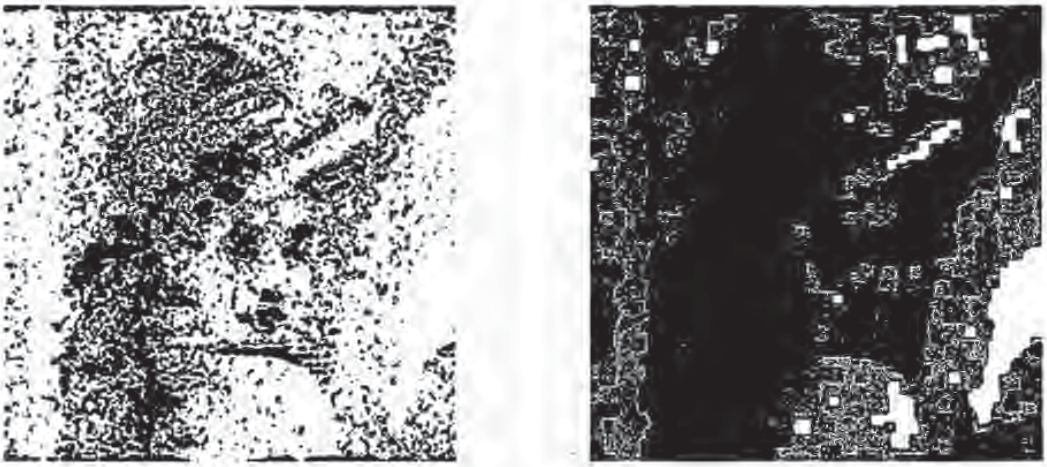
The conclusion of these observations is that the invisibility is only strictly observed in high activity regions, where the local energy of high fequencies is important. These regions have to be favoured during the inscription in the sense that the level of the watermark will be increased in those regions while it has to be decreased in other regions.
The correction process first isolates the high activity regions (figure 3.a). Then, an homogeneization of this picture is performed by use of morphological tools, e.g. one opening and one closing (figure 3.b). After a leveling (in fact, a division by the mean or mean square value of the homogenized mask), we obtain a new mask used to multiply the picture local energy and so, giving an advantage to regions of highfrequency energy in comparison with other areas. After that correction, the process is identical to the one described previously. Moreover, the complexity is not increased. Indeed, we first work on the inscription at high frequencies (where there is no quality problems). The value of high frequency local energy is then used for the calculation of the correcting mask used for inscription at lower frequencies. The correction scheme is drawn in the following schema.



4.2 Detection

The aim is to detect if a watermark has been embedded. This can be done with the use of a correlation, but first it is necessary to isolate the watermark and then to demodulate it in order to reconstruct something that is highly correlated with the basic information (the grid).
The formulation of the watermark is:

$$W(x,y) = \sum_{j \in K} A_j.cos(f_{x_j}.x + f_{y_j}.y) \qquad (6)$$



(a) (b)

Figure 3: Correcting mask for Lena: (a) Areas of high frequencies, (b) Morphological homogeneization of the mask.

$$where A_j = \alpha_j . G(x,y) . M(x,y) \tag{7}$$

In this expression, $M(x,y)$ adjusts the level of the grid in order it becomes invisible, it is called *a mask*, and its maximal value is one.
α_j is a constant that used to normalize the mask, it must be as high as possible.

The detection is divided in three steps : teh demodulation, the correlation and the decision.

- Demodulation

$$I_W(x,y) = \sum_{j \in K} A_j . cos(f_{x_j} . x + f_{y_j} . y) + I_O + N(x,y) \tag{8}$$

where $I_W(x,y)$ is the watermarked picture, $I_O(x,y)$ is the original picture and $N(x,y)$ is an additive noise from the channel.

The demodulation consists in multiplying I_W by $cos(f_{x_j} . x + f_{y_j} . y), \forall j \in K$ and then to filter with a LP filter.
The result will be :

$$D_j(x,y) = \frac{1}{2} . A_j(x,y) + N^*(x,y) \tag{9}$$

$N^*(x,y)$ depends on the image and on the additive noise. The other parts of the stamp will be eliminated by the LP filter.

- Correlation It consists in mutiplying the demodulated information $D(x,y) = \sum_{j \in K} D_j(x,y)$ with the basic grid $G(x,y)$. If the picture has not been too deteriorated, $D(x,y)$ and $G(x,y)$ should be similar.

$$C = \sum_{j \in K} \sum_{x,y} D_j(x,y) . G(x,y) \tag{10}$$

$$= \sum_{j \in K} \alpha_j \sum_{x,y} [G^2(x,y).M_j(x,y) + G(x,y).N^*(x,y)] \quad (11)$$

In 11, the first term is even greater than the second, because G and N^* have null average values.
So C exclusively depends on the watermark value.
in the case the grid is not the good one, the correlation gives:

$$C^* = \sum_{j \in K} \alpha_j \sum_{x,y} G(x,y).G^*(x,y)..M_j(x,y) \quad (12)$$

$C^* \ll C$ if the choice of the basic information has been appropriate.

- decision

The detection algorithm performs demodulations and correlations at diverse frequencies and with diverse grids.The decision is made after the comparison of these correlations.

5 RESULTS

The first and probably mosty important result is the invisibility of the stam in all images that were tested. Figure 4.a and b compares the original and stamped picture for Lena. In figure 4.e, omne observes the watermark that was added to the original picture.

Two methods were used to determine whether an image is watermarked or not. The first one consists in comparing the result of C the correlation made with the right grid $G(x,y)$ from the right key with C^* the correlation made with $G^*(x,y)$, the grid obtained by random keys see 12. If the picture is watermarked, the correlation with the right key is even greater than the random correlations. The results below (Figure 5) show the pertinence of this method.

The second method uses a grid $G(x,y)$ formed from a MLS sequence, having good correlation properties. Correlations are made with shifted versions of the basic grid. Due to these good correlation properties, the correlation with the the right grid gives a result even greater than the correlations with shifted grids. Results are presented below (figure 4.c and d), if a picture is watermarked, a pick appears in the center.

6 SYSTEM ROBUSTNESS

Many tests have been performed concerning usual pictures deteriorations in image processing like blurring and compression. The inspection of these results are quite satisfying, but expected due to the frequency approach.
For all classical pirate attacks like zoom, cropping, overwatermarking it is not as simple. The overwatermarking makes no problem, the presence of the watermark is still detected. But for zoom and cropping, the remaing point is to find a few tools permitting to complete the process. The concept of these tools is already defined but yet no implementation has been acheived.[7]

7 CONCLUSION

The process developed here allows the watermarking of the ownership of any picture. The perceptual approach used here is probably the best one, that is why the results obtained are so satifying compared with other methods and this method is so performant. Nevertheless studies are still running to acheive a new goal, consisting in

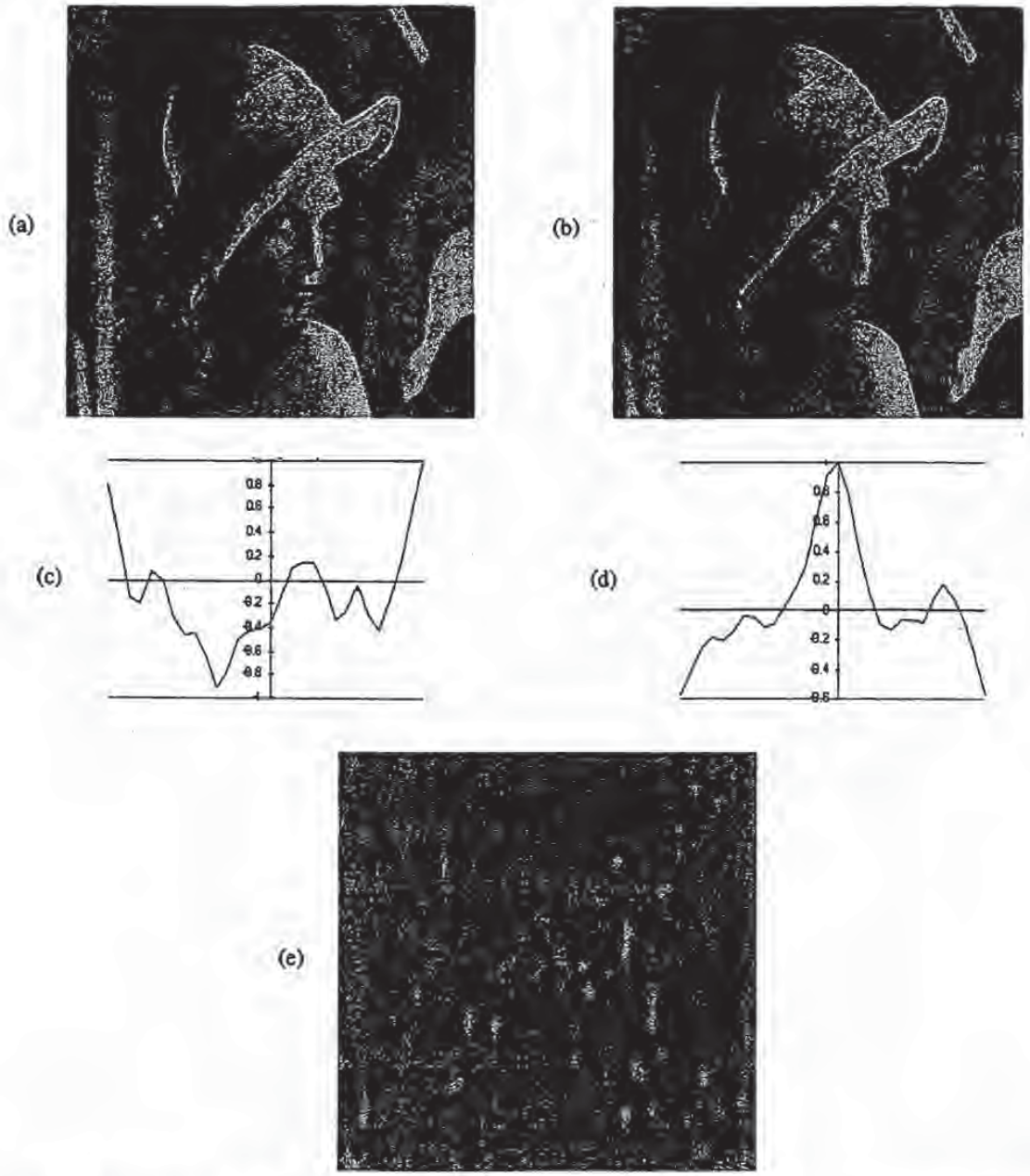


Figure 4: Results for Lena: (a) Original, (b) Watermarked one, (c) Correlation grafic for original, (d) Correlation grafic for watermarked, (e) Watermark.

Image Name	Optimal correlation	Random correlation 1	Random correlation 2	Random correlation3	Random correlation 4	Conclusion
Lena watermarked	584609	92605	133920	80534	143633	*watermarked*
Lena original	94538	98099	135492	76739	137120	*Non watermarked*

Figure 5: Results of correlation for Lena and decision.

making more information (e.g. ownership, date of marking) readable by the key owner from the watermark. This could be useful for real copyright protection protocols[8,9]

8 REFERENCES

[1] Kahin B. The strategic environment for protecting multimedia. volume 1, pages 1–8. IMA Intellectual Property Project Proceedings, January 1994.

[2] Comes S. *Les traitements perceptifs d'images numérisées.* PhD thesis, Université Catholique de Louvain, June 1995.

[3] Okak L.A. and Thomas J.P. Handbook of perception and human performance vol.1: Seeing spatial patterns. chapter 7.

[4] G.C. Phillips H.R. Wilson, D.K. McFarlane. Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research*, 23(9):873–847, 1983.

[5] G.C. Phillips H.R. Wilson. Orientation bandwiths of spatial mechanisms measured by masking. *J. Opt. Soc. Am. A*, 1(2):226–232, February 1984.

[6] Edited by Gustavus J. Simmons. Section 1: Chapter 4: 'public key cryptography' and section 2: Chapter 6: 'authentication: Digital signature' from 'contemporary cryptology: the science of information integrity' ieee press, 1992.

[7] J.F. Delaigle and C. De Vleeschouwer. Etiquetage d'images numériques en vue de la protection des droits d'auteur, Juin 1995.

[8] J.F. Delaigle C. Simon and B. Macq. Talisman (ac019): Technical state of the art. January 1996.

[9] O. Bruyndonckx J.M. Boucqueau and B. Macq. Watermarking: workpackage 5 of accopi. June 1995.

A ROBUST CONTENT BASED DIGITAL SIGNATURE FOR IMAGE AUTHENTICATION

Marc Schneider and Shih-Fu Chang

Columbia University
Image and Advanced Television Laboratory
Room 801 Schapiro Research Building
530 West 120th Street
New York, NY 10027-6699
USA
E-mail: {mars, sfchang}@ctr.columbia.edu

Abstract

A methodology for designing content based digital signatures which can be used to authenticate images is presented. A continuous measure of authenticity is presented which forms the basis of this methodology. Using this methodology signature systems can be designed which allow certain types of image modification (e.g. lossy compression) but which prevent other types of manipulation. Some experience with content based signatures is also presented.

The idea of signature based authentication is extended to video, and a system to generate signatures for video sequences is presented. This signature also allows smaller segments of the secured video to be verified as unmanipulated.

1.0 Motivation

Powerful, and easy to use image manipulation software has made it possible to alter digital images. It has been suggested that the authenticity of digital images can be preserved by having a camera "sign" the image using a digital signature. [1] However, applying a signature scheme directly to the image has some drawbacks. For many applications, image compression is desired to reduce transmission bandwidth, storage space, etc. Authenticity, the ability to detect image manipulation, is also desired. These two function are at odds with each other since lossy compression is a form of manipulation. Our goal is to develop a way to be able to prove some form of authenticity, while still allowing desired forms of manipulation, such as lossy compression. Ideally, a robust signature scheme should not declare an image modified under these circumstances.

2.0 Previous Work

Previous work on image authentication falls into two groups, digital signatures[1] and digital watermarks[3]. A digital signature is based upon the idea of public key encryption. A private key is used to encrypt a hashed version of the image. This encrypted file then forms a unique "signature" for the image since only the entity signing the image has knowledge of the private key used. An associated public key can be used to decrypt the signature. The image under question can be hashed using the same hashing function as used originally. If these hashes match then the image is authenticated.

Digital signatures can be used for more than just image authentication. In particular when combined with secure timestamp, a digital signature can be used as a proof of first authorship.

A watermark, on the other hand, is a code secretly embedded into the image. The watermark allows for verification of the origin of an image. However, a watermark alone is not enough to prove first authorship, since an image could be marked with multiple watermarks. It has also been pointed out [6] that digital watermarks are not well suited to protecting the authenticity of an image.

3.0 Content Based Signatures

The key to developing a robust digital signature for images is to examine what the digital signature should protect. Ideally the signature should protect the message conveyed by the content of the image, and not the particular representation of that content. Thus the robust signature can be used to verify the authenticity of an image which has been modified by processing that does not affect the content of the image. Examples of this type of processing are removal of noise or lossy compression. However, manipulation of the image which changes the content, such as removal of a person from a scene, can still be detected by the use of this signature.

Additionally, the use of a content based signature fits well with other content based image processing, such as content based coding and queries. By using the same content for both the signature and the compression algorithm, the signature will be able to authenticate images highly compressed using content based coding. With content based queries a signature can be the basis of a query.

4.0 Authenticity and Feature Selection

Often people think of authenticity as a binary quantity, either an image is authentic or it is not authentic. However, this not always what people want when they are concerned with detecting image manipulation. We propose a continuous interpretation of authentic. An image which is bit by bit identical to the original image is considered completely authentic (authenticity measure of 1.0). An image which has nothing in common with the original image would be considered unauthentic (authenticity measure of 0.0). All other images would be partially authentic. Partially authentic is a loosely defined concept and measurement of the authenticity is subjective, and changes from application domain to application domain. One

0-7803-3258-X/96/$5.00 © 1996 IEEE

way of thinking of this authenticity measure is as an authenticity vs. modification curve (see Figure 1). For example a curve could be drawn relating authenticity to the bit rate of a compressed image. Thus for each different type of modification there would be a corresponding curve. The old concept of authenticity would be represent as a Dirac delta function or a unit step function for all of the possible types of modification.

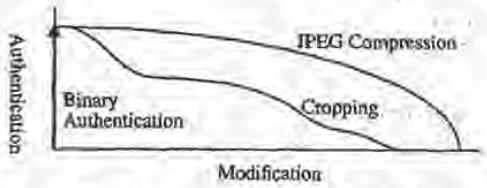


Figure 1. Authenticity vs. Modification Curve

Since authenticity is a subjective quantity, it is difficult to use directly as the basis of an authenticity verification system. We therefore need an approximation to authenticity which is analytical and can be computed from an image. The approach we are taking is to define a concept call feature authenticity A_f, which is one minus the normalized distance between a feature vector computed for the original image, I_o and the same feature vector computed from the image whose authenticity is to be measured, I_m. The key

$$A_f = 1 - \|feature(I_o) - feature(I_m)\|_{normalized}$$

is to find a set of features such that the feature authenticity closely approximates the image authenticity curves for the allowable forms of modification (e.g. lossy compression). Additionally, the feature authenticity curves for undesired forms of modification should be significantly below the curves for allowable forms of manipulation. Using the continuous measure of feature authenticity, a minimum acceptable authenticity can be defined. This can be defined directly, or defined in terms of some acceptable amount of manipulation. For example, the minimum acceptable authenticity can be defined in terms of maximum compression ratio. This minimum authenticity becomes a constraint on the optimization of the feature set. Once acceptable forms of manipulation are specified (e.g. compression, noise reduction, etc.) and the unacceptable forms are specified (e.g. cropping, cut and paste, etc.) the optimal set of features can be found. The goal is to have the authenticity vs. modification curve have a gentle slope for desired type of manipulation, and to have a very steep slope for the undesired forms of manipulation.

5.0 Generating and Verifying a Content Based Signature

The general procedure for generating a content based signature is diagramed in figure 2. First, the content of interested C_o is extracted from the image I_o to be signed, using an extraction function f_c. The content is then possibly hashed, using a hash function f_h, to reduce the amount of data. This may be necessary since the size of the signature is dependent upon the amount of data

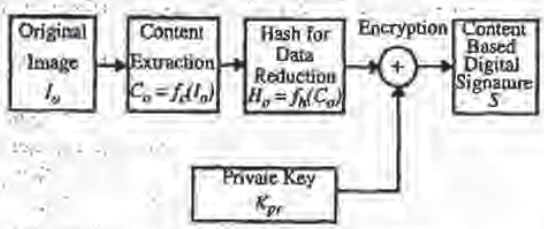


Figure 2. Generating a Content Based Signature

encrypted. The hash H_o is then encrypted using the private key K_{pr} of the signing entity to produce the final signature S. To verify

$$C_o = f_c(I_o)$$
$$H_o = f_h(C_o)$$
$$S = H_o \oplus K_{pr}$$

the authenticity of an questionable image I_t, the signature is decrypted using the public key K_{pu} and is compared to the hashed content extracted from the questionable image. If the distance between the feature vectors is less then a threshold value *tau*, then the questionable image is declared unmanipulated. This procedure is shown in figure 3.

$$C_t = f_c(I_t)$$
$$H_t = f_h(C_t)$$
$$H_o = S \oplus K_{pu}$$
$$\|H_o - H_t\| \le \tau \Rightarrow authentic$$

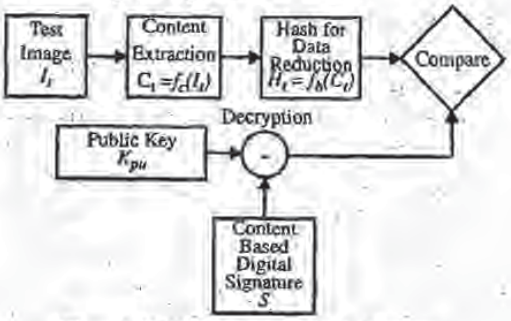


Figure 3. Verifying a Content Based Signature

The threshold value can be set by examining the amount of error introduce into an image by lossy compression. The difference between the image content hash at a target compression rate and the original image content hash can be used to set the threshold value. Note: If the threshold value is non zero, a cryptographically based hashing function can not be used, since there is no significance to closeness once a cryptographic hash has been applied to the content. A non cryptographic hashing function can be used to reduce the size of the signature, however this will typically weaken the signature.

It should be noted that a cryptographic hash can be used as the content extraction function. This content is of course not related to the way visual information is processed. Thus operations such as lossy compression as well as most other forms of image modification will have very steep authentication vs. modification curves. This is of course the signature scheme proposed in [1].

The problem is now one of finding a set of features which adequately describe the content of an image. Several different features can be used such as edge information, DCT coefficients, and color or intensity histograms.

We examined using the intensity histogram to sign the image. However, the histogram of the entire image itself is not very useful, since it contains no spatial information about the image intensities. Thus the images were divided into blocks and the intensity histogram for each block was computed separately. This allows some spatial information to be incorporated into the signature since the location of these blocks are fixed. Further spatial information can be incorporated by using a variable size block. Starting with small blocks, the histograms can be combined to form the histogram of a larger block. This can be used to produce blocks of different sizes for different parts of the image, allowing fine details to be protected by a small block size, and larger regions to be protected by a larger block size.

The distance function for detecting content changes is a subtle and challenging issue. The euclidean distance between intensity histograms was used as a measure of the content of the image. This performed well in detecting modification of the image. The amount of lossy compression which could be applied to the image and not produce a false positive was limited to at most 4:1. If we used a reduced distance function, then the maximum permissible compression ratio is increased. However, this increased robustness is achieved at the cost of sensitivity to subtle image manipulation. For example, it was found that the mean of the intensity histograms was a useful measure for detecting image content manipulation. Several different images were signed using the block average intensity technique. These images were then altered, typical altered test images are shown in figures 5 and 6, and the signatures were used to successfully detect the image manipulation. Note that the white boxes were added to highlight the modified regions and did not appear in the original test images. The original image was also compressed using JPEG compression. The signature system was not triggered even at high compression ratios of 14:1. As opposed to the euclidean distance using the full histogram, the maximum compression ratio is increased, but we can clearly see the trade-off.

Content based signatures for images can also be used to have an author sign an image. A typical use of this would be proof of first authorship. For this application the signature would have to be processed by a secure timestamp server. Additionally, it would be desirable to have the signature embedded into the image for this application. The signature should travel with the image so that authorship can readily be confirmed. It should be embedded so that even if the image is converted from one format to another the signature will remain with the image.

Embedding the signature into the image brings forth an additional issue, development of an embedding process which does not effect the signature verification process, since the embedding processes manipulates the image data. Information embedding in images is a

Figure 4. Original Image

Figure 5. JPEG Compressed 14:1

Figure 6. Manipulated Image Strip on Fireman's Jacket Removed

Figure 7 Manipulated Image Fire Hose Removed

generalization of the image watermark problem. One possible approach to embed a watermark into an image is to code the signature such that it resembles quantization noise and embed this in the image.[7] Another technique embeds a message into an image in the frequency domain. The image is broken into 8x8 blocks and a Discrete Cosine Transform is performed on each of the blocks. The signature is embedded by modifying the middle frequency coefficients and transforming the block back to the spatial domain.[8] This technique could easily be used with a signature based on the DC component, since this is unaffected by the embedding process. A third method also embeds a message in the frequency domain, however, here the image is treated as a noisy communication channel. The watermark is transmitted in this channel using a spread spectrum technique. [9]

6.0 Authentication of Video

It is also possible to extend authentication systems to video. There are two additional problems that need to be addressed when dealing with video sequences. The first is that of maintaining frame order integrity. The still image authentication techniques can be applied to each frame of the video sequence, but an addition signature must be provided to insure that the frames aren't reordered. The second problem is that we would like an unmodified clip from the larger sequence to be detected as unmodified. Here we present an extension to the original trustworthy camera[1]. A cryptographic hash is applied to each frame of the video. The hashes are ordered according to the corresponding frame order. This ordered sequence of hashes is then itself hashed (See Figure 8). All of these hashes are then encrypted using a public key cryptographic system to prevent forging of the hashes. To verify the authenticity of the entire video sequence only the second level hash is needed. The first level hashes are generated from the video sequence to be checked. These are then used to compute the second level hash. This second level hash is then compared to the original second level hash. To verify a subsection of the video sequence, first level hashes are generated from the subsection. Missing hashes are supplied from the signature. The second level hash can then be generated and checked. This system can be used to protect any group of pictures, not just video. For example, it could be used to protect the authenticity and order of multiple slices of MRI data.

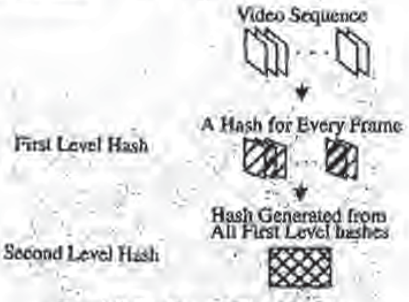


Figure 8. Two Level Video Hashing

This idea can be extended to provide a more flexible method of verifying the authenticity of still pictures. In order to do this we can break the still picture into blocks. A hash is generated for each of the blocks. The blocks are then ordered, for instance by scanning across the rows. The sequence of hashes is then hashed to generate a signature which can protect the order. This image signature can be used to verify the authenticity of sections of cropped images. The blocks not effected by the cropping can be verified using the hashes for those blocks.

7.0 Contribution and Conclusion

The contributions of this work are the idea of using the image content to form a signature which can be used to protect the authenticity of images and survive acceptable compression. We also proposed a method for the extension of the authentication system to video. We have also presented a methodology for determining the set of features which can be used to approximate the image authenticity. We have also presented a method to extend digital signatures to video sequences, which can also be used to enhance the robustness of signatures for still images.

8.0 References

[1] Friedman, Gary L., "The Trustworthy Digital Camera: Restoring Credibility to the Photographic Image", IEEE Transactions on Consumer Electronics, vol. 39 no. 4, November 1993, pp. 905-910.

[2] Netravali, Arun N. and Haskell, Barry G."Digital Pictures: Representation and Compression", New York, NY, Plenum Press, 1988.

[3] Walton, Steve, "Image Authentication for a Slippery New Age", Dr. Dobb's Journal, April 1995, pp. 18-26

[4] Stinson, Douglas R., "Cryptography: Theory and Practice", Boca Raton, FL, CRC Press, 1995.

[5] Wallace, G.K., "The JPEG still picture compression standard.", Communications of the ACM, vol. 34, no. 4, April 1991, pp. 30-40.

[6] Macq, B.M. and Quisquater, J.J., "Cryptology for Digital TV Broadcasting", Proceedings of the IEEE, vol. 83, no. 6, June 1995.

[7] Matsui, Kineo and Tanaka, Kiyoshi, "Video-Steganography: How to Secretly Embed a Signature in a Picture", IMA Intellectual Property Project Proceedings, vol. 1, pp. 187-206, 1994.

[8] Zhao, Jian and Koch, Eckhard, "Embedding Robust Label into Images for Copyright Protection", Proc. of the International Congress on Intellectual Property Rights for Specialized Information, Knowledge and New Technologies., Vienna, Austria, August 21-25, 1995.

[9] Cox, I.J., Kilian, J., Leighton, T. and Shamoon, T., "Secure Spread Spectrum Watermarking for Multimedia", NEC Research Institute Technical Report 95-10, 1995.

Secure Spread Spectrum Watermarking for Multimedia

Ingemar J. Cox, *Senior Member, IEEE*, Joe Kilian, F. Thomson Leighton, and Talal Shamoon, *Member, IEEE*

***Abstract*— This paper presents a secure (tamper-resistant) algorithm for watermarking images, and a methodology for digital watermarking that may be generalized to audio, video, and multimedia data. We advocate that a watermark should be constructed as an independent and identically distributed (i.i.d.) Gaussian random vector that is imperceptibly inserted in a spread-spectrum-like fashion into the perceptually *most* significant spectral components of the data. We argue that insertion of a watermark under this regime makes the watermark robust to signal processing operations (such as lossy compression, filtering, digital-analog and analog-digital conversion, requantization, etc.), and common geometric transformations (such as cropping, scaling, translation, and rotation) provided that the original image is available and that it can be succesfully registered against the transformed watermarked image. In these cases, the watermark detector unambiguously identifies the owner. Further, the use of Gaussian noise, ensures strong resilience to multiple-document, or collusional, attacks. Experimental results are provided to support these claims, along with an exposition of pending open problems.**

***Index Terms*— Intellectual property, fingerprinting, multimedia, security, steganography, watermarking.**

I. Introduction

THE PROLIFERATION of digitized media (audio, image, and video) is creating a pressing need for copyright enforcement schemes that protect copyright ownership. Conventional cryptographic systems permit only valid keyholders access to encrypted data, but once such data is decrypted there is no way to track its reproduction or retransmission. Therefore, conventional cryptography provides little protection against data piracy, in which a publisher is confronted with unauthorized reproduction of information. A digital watermark is intended to complement cryptographic processes. It is a visible, or preferably invisible, identification code that is permanently embedded in the data and remains present within the data after any decryption process. In the context of this work, data refers to audio (speech and music), images (photographs and graphics), and video (movies). It does not include ASCII representations of text, but does include text represented as an image. Many of the properties of the scheme presented in this work may be adapted to accommodate audio and video implementations, but the algorithms here specifically apply to images.

A simple example of a digital watermark would be a visible "seal" placed over an image to identify the copyright owner (e.g., [2]). A visible watermark is limited in many ways. It marrs the image fidelity and is susceptible to attack through direct image processing. A watermark may contain additional information, including the identity of the purchaser of a particular copy of the material. In order to be effective, a watermark should have the characteristics outlined below.

Unobtrusiveness: The watermark should be perceptually invisible, or its presence should not interfere with the work being protected.

Robustness: The watermark must be difficult (hopefully impossible) to remove. If only partial knowledge is available (for example, the exact location of the watermark in an image is unknown), then attempts to remove or destroy a watermark should result in severe degradation in fidelity before the watermark is lost. In particular, the watermark should be robust in the following areas.

- *Common signal processing:* The watermark should still be retrievable even if common signal processing operations are applied to the data. These include, digital-to-analog and analog-to-digital conversion, resampling, requantization (including dithering and recompression), and common signal enhancements to image contrast and color, or audio bass and treble, for example.
- *Common geometric distortions (image and video data):* Watermarks in image and video data should also be immune from geometric image operations such as rotation, translation, cropping and scaling.
- *Subterfuge attacks (collusion and forgery):* In addition, the watermark should be robust to collusion by multiple individuals who each possess a watermarked copy of the data. That is, the watermark should be robust to combining copies of the same data set to destroy the watermarks. Further, if a digital watermark is to be used in litigation, it must be impossible for colluders to combine their images to generate a different valid watermark with the intention of framing a third party.

Manuscript received January 14, 1996; revised January 24, 1997. Portions of this work were reprinted, with permission, from the Proceedings of the IEEE Conference on Image Processing, 1996, and from the Proceedings of the First International Conference on Data Hiding (Springer-Verlag, 1996). The associate editor coordinating the reivew of this manuscript and approving it for publication was Prof. Sarah Rajala.

I. J. Cox and J. Kilian are with NEC Research Institute, Princeton, NJ 08540 USA (e-mail: ingemar@research.nj.nec.com; joe@research.nj.nec.com).

F. T. Leighton is with the Mathematics Department and Laboratory for Computer Science, The Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: ftl@math.mit.edu).

T. Shamoon is with InterTrust STAR Laboratory, Sunnyvale, CA 94086 USA (e-mail: talal@intertrust.com).

Publisher Item Identifier S 1057-7149(97)08460-1

Universality: The same digital watermarking algorithm should apply to all three media under consideration. This is potentially helpful in the watermarking of multimedia products. Also, this feature is conducive to implementation of audio and image/video watermarking algorithms on common hardware.

Unambiguousness: Retrieval of the watermark should unambiguously identify the owner. Furthermore, the accuracy of owner identification should degrade gracefully in the face of attack.

There are two parts to building a strong watermark: the *watermark structure* and the *insertion strategy*. In order for a watermark to be robust and secure, these two components must be designed correctly. We provide two key insights that make our watermark both robust and secure: We argue that the watermark be placed explicitly in the perceptually most significant components of the data, and that the watermark be composed of random numbers drawn from a Gaussian $(N(0,1))$ distribution.

The stipulation that the watermark be placed in the perceptually significant components means that an attacker must target the fundamental structural components of the data, thereby heightening the chances of fidelity degradation. While this strategy may seem counterintuitive from the point of view of steganography (how can these components hide any signal?), we discovered that the significant components have a *perceptual capacity* that allows watermark insertion without perceptual degradation. Further, most processing techniques applied to media data tend to leave the perceptually significant components intact. While one may choose from a variety of such components, in this paper, we focus on the perceptually significant *spectral* components of the data. This simultaneously yields high perceptual capacity and achieves a uniform spread of watermark energy in the pixel domain.

The principle underlying our watermark structuring strategy is that the mark be constructed from independent, identically distributed (i.i.d.) samples drawn from a Gaussian distribution. Once the significant components are located, Gaussian noise is injected therein. The choice of this distribution gives resilient performance against collusion attacks. The Gaussian watermark also gives our scheme strong performance in the face of quantization, and may be structured to provide low false positive and false negative detection. This is discussed below, and elaborated on in [13].

Finally, note that the techniques presented herein do not provide proof of content ownership on their own. The focus of this paper are algorithms that insert messages into content in an extremely secure and robust fashion. Nothing prevents someone from inserting another message and claiming ownership. However, it is possible to couple our methods with strong authentication and other cryptographic techniques in order to provide complete, secure and robust owner identification and authentication.

Section III begins with a discussion of how common signal transformations, such as compression, quantization, and manipulation, affect the frequency spectrum of a signal. This discussion motivates our belief that a watermark should be embedded in the data's perceptually significant frequency components. Of course, the major problem then becomes how to imperceptibly insert a watermark into perceptually significant components of the frequency spectrum. Section III-A proposes a solution based on ideas from spread spectrum communications. In particular, we present a watermarking algorithm that relies on the use of the original image to extract the watermark. Section IV provides an analysis based on possible collusion attacks that indicates that a binary watermark is not as robust as a continuous one. Furthermore, we show that a watermark structure based on sampling drawn from multiple i.i.d Gaussian random variables offers good protection against collusion. Ultimately, no watermarking system can be made perfect. For example, a watermark placed in a textual image may be eliminated by using optical character recognition technology. However, for common signal and geometric distortions, the experimental results of Section V suggest that our system satisfies most of the properties discussed in the introduction, and displays strong immunity to a variety of attacks in a collusion resistant manner. Finally, Section VI discusses possible weaknesses and potential enhancements to the system and describes open problems and subsequent work.

II. Previous Work

Several previous digital watermarking methods have been proposed. Turner [25] proposed a method for inserting an identification string into a digital audio signal by substituting the "insignificant" bits of randomly selected audio samples with the bits of an identification code. Bits are deemed "insignificant" if their alteration is inaudible. Such a system is also appropriate for two-dimensional (2-D) data such as images, as discussed in [26]. Unfortunately, Turner's method may easily be circumvented. For example, if it is known that the algorithm only affects the least significant two bits of a word, then it is possible to randomly flip *all* such bits, thereby destroying any existing identification code.

Caronni [6] suggests adding *tags*—small geometric patterns—to digitized images at brightness levels that are imperceptible. While the idea of hiding a spatial watermark in an image is fundamentally sound, this scheme may be susceptible to attack by filtering and redigitization. The fainter such watermarks are, the more susceptible they are such attacks and geometric shapes provide only a limited alphabet with which to encode information. Moreover, the scheme is not applicable to audio data and may not be robust to common geometric distortions, especially cropping.

Brassil *et al.* [4] propose three methods appropriate for document images in which text is common. Digital watermarks are coded by 1) vertically shifting text lines, 2) horizontally shifting words, or 3) altering text features such as the vertical endlines of individual characters. Unfortunately, all three proposals are easily defeated, as discussed by the authors. Moreover, these techniques are restricted exclusively to images containing text.

Tanaka *et al.* [19], [24] describe several watermarking schemes that rely on embedding watermarks that resemble quantization noise. Their ideas hinge on the notion that quantization noise is typically imperceptible to viewers. Their

first scheme injects a watermark into an image by using a predetermined data stream to guide level selection in a predictive quantizer. The data stream is chosen so that the resulting image looks like quantization noise. A variation on this scheme is also presented, where a watermark in the form of a dithering matrix is used to dither an image in a certain way. There are several drawbacks to these schemes. The most important is that they are susceptible to signal processing, especially requantization, and geometric attacks such as cropping. Furthermore, they degrade an image in the same way that predictive coding and dithering can.

In [24], the authors also propose a scheme for watermarking facsimile data. This scheme shortens or lengthens certain runs of data in the run length code used to generate the coded fax image. This proposal is susceptible to digital-to-analog and analog-to-digital attacks. In particular, randomizing the least significant bit (LSB) of each pixel's intensity will completely alter the resulting run length encoding. Tanaka *et al.* also propose a watermarking method for "color-scaled picture and video sequences". This method applies the same signal transform as the Joint Photographers Expert Group (JPEG) (discrete cosine transform of 8×8 subblocks of an image) and embeds a watermark in the coefficient quantization module. While being compatible with existing transform coders, this scheme may be susceptible to requantization and filtering and is equivalent to coding the watermark in the LSB's of the transform coefficients.

In a recent paper, Macq and Quisquater [18] briefly discuss the issue of watermarking digital images as part of a general survey on cryptography and digital television. The authors provide a description of a procedure to insert a watermark into the least significant bits of pixels located in the vicinity of image contours. Since it relies on modifications of the least significant bits, the watermark is easily destroyed. Further, their method is restricted to images, in that it seeks to insert the watermark into image regions that lie on the edge of contours. Bender *et al.* [3] describe two watermarking schemes. The first is a statistical method called *patchwork*. Patchwork randomly chooses n pairs of image points, (a_i, b_i), and increases the brightness at a_i by one unit while correspondingly decreasing the brightness of b_i. The expected value of the sum of the differences of the n pairs of points is then $2n$, provided certain statistical properties of the image are true.

The second method is called "texture block coding," wherein a region of random texture pattern found in the image is copied to an area of the image with similar texture. Autocorrelation is then used to recover each texture region. The most significant problem with this technique is that it is only appropriate for images that possess large areas of random texture. The technique could not be used on images of text, for example, nor is there a direct analog for audio.

Rhoads [21] describes a method that adds or subtracts small random quantities from each pixel. Addition or subtraction is determined by comparing a binary mask of L bits with the LSB of each pixel. If the LSB is equal to the corresponding mask bit, then the random quantity is added, otherwise it is subtracted. The watermark is subtracted by first computing the difference between the original and watermarked images and then by examining the sign of the difference, pixel by pixel, to determine if it corresponds to the original sequence of additions and subtractions. This method does not make use of perceptual relevance, but it is proposed that the high frequency noise be prefiltered to provide some robustness to lowpass filtering. This scheme does not consider the problem of collusion attacks.

Koch, Rindfrey, and Zhao [14] propose two general methods for watermarking images. The first method, attributed to Scott Burgett, breaks up an image into 8×8 blocks and computes the discrete cosine transform (DCT) of each of these blocks. A pseudorandom subset of the blocks is chosen, then, in each such block, a triple of frequencies is selected from one of 18 predetermined triples and modified so that their relative strengths encode a one or zero value. The 18 possible triples are composed by selection of three out of eight predetermined frequencies within the 8×8 DCT block. The choice of the eight frequencies to be altered within the DCT block is based on a belief that the "middle frequencies...have moderate variance," i.e. they have similar magnitude. This property is needed in order to allow the relative strength of the frequency triples to be altered without requiring a modification that would be perceptually noticeable. Superficially, this scheme is similar to our own proposal, also drawing an analogy to spread spectrum communications. However, the structure of their watermark is different from ours, and the set of frequencies is not chosen based on any direct perceptual significance, or relative energy considerations. Further, because the variance between the eight frequency coefficients is small, one would expect that their technique may be sensitive to noise or distortions. This is supported by the experimental results that report that the "embedded labels are robust against JPEG compression for a quality factor as low as about 50%." By comparison, we demonstrate that our method performs well with compression quality factors as low as 5%. An earlier proposal by Koch and Zhao [15] used not triples of frequencies but pairs of frequencies, and was again designed specifically for robustness to JPEG compression. Nevertheless, they state that "a lower quality factor will increase the likelihood that the changes necessary to superimpose the embedded code on the signal will be noticeably visible." In a second method, designed for black and white images, no frequency transform is employed. Instead, the selected blocks are modified so that the relative frequency of white and black pixels encodes the final value. Both watermarking procedures are particularly vulnerable to multiple document attacks. To protect against this, Zhao and Koch propose a *distributed* 8×8 block created by randomly sampling 64 pixels from the image. However, the resulting DCT has no relationship to that of the true image and consequently may be likely to cause noticeable artifacts in the image and be sensitive to noise.

In addition to direct work on watermarking images, there are several works of interest in related areas. Adelson [1] describes a technique for embedding digital information in an analog signal for the purpose of inserting digital data into an analog TV signal. The analog signal is quantized into one of two disjoint ranges ($\{0, 2, 4 \cdots\}$, $\{1, 3, 5 \cdots\}$, for example) that are selected based on the binary digit to be transmitted. Thus,

Adelson's method is equivalent to watermark schemes that encode information into the LSB's of the data or its transform coefficients. Adelson recognizes that the method is susceptible to noise and therefore proposes an alternative scheme wherein a 2×1 Hadamard transform of the digitized analog signal is taken. The differential coefficient of the Hadamard transform is offset by zero or one unit prior to computing the inverse transform. This corresponds to encoding the watermark into the least significant bit of the differential coefficient of the Hadamard transform. It is not clear that this approach would demonstrate enhanced resilience to noise. Furthermore, like all such LSB schemes, an attacker can eliminate the watermark by randomization.

Schreiber *et al.* [22] describe a method to interleave a standard NTSC signal within an enhanced definition television (EDTV) signal. This is accomplished by analyzing the frequency spectrum of the EDTV signal (larger than that of the NTSC signal) and decomposing it into three subbands (L, M, H for low-, medium- and high-frequency, respectively). In contrast, the NTSC signal is decomposed into two subbands, L and M. The coefficients, M_k, within the M band are quantized into m levels and the high frequency coefficients, H_k, of the EDTV signal are scaled such that the addition of the H_k signal plus any noise present in the system is less than the minimum separation between quantization levels. Once more, the method relies on modifying least significant bits. Presumably, the midrange rather than low frequencies were chosen because these are less perceptually significant. In contrast, the method proposed here modifies the *most* perceptually significant components of the signal.

Finally, it should be noted that existing techniques are generally not resistant to collusion attacks by multiple documents.

III. Watermarking in the Frequency Domain

In order to understand the advantages of a frequency-based method, it is instructive to examine the processing stages that an image (or sound) may undergo in the process of copying, and to study the effect that these stages could have on the data, as illustrated in Fig. 1. In the figure, "transmission" refers to the application of any source or channel code, and/or standard encryption technique to the data. While most of these steps are information lossless, many compression schemes (JPEG, MPEG, etc.) are lossy, and can potentially degrade the data's quality, through *irretrievable* loss of information. In general, a watermarking scheme should be resilient to the distortions introduced by such algorithms.

Lossy compression is an operation that usually eliminates perceptually nonsalient components of an image or sound. Most processing of this sort takes place in the frequency domain. In fact, data loss usually occurs among the high-frequency components.

After receipt, an image may endure many common transformations that are broadly categorized as geometric distortions or signal distortions. Geometric distortions are specific to images and video, and include such operations as rotation, translation, scaling and cropping. By manually determining a minimum of four or nine corresponding points between the original and the distorted watermark, it is possible to remove any two or three-dimensional (3-D) affine transformation [8]. However, an affine scaling (shrinking) of the image leads to a loss of data in the high-frequency spectral regions of the image. Cropping, or the cutting out and removal of portions of an image, leads to irretrievable loss of image data, which may seriously degrade any spatially based watermark such as [6]. However, a frequency-based scheme spreads the watermark over the whole spatial extent of the image, and is therefore less likely to be affected by cropping, as demonstrated in Section V-E.

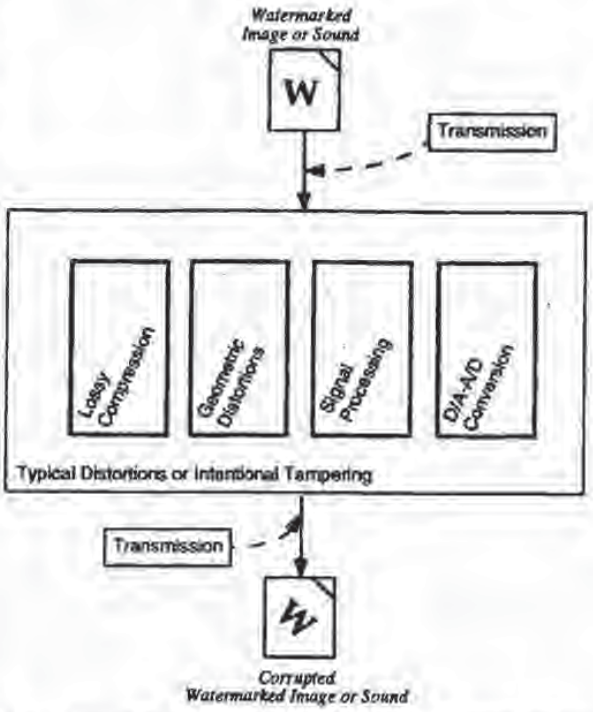


Fig. 1. Common processing operations that a media document could undergo.

Common signal distortions include digital-to-analog and analog-to-digital conversion, resampling, requantization, including dithering and recompression, and common signal enhancements to image contrast and/or color, and audio frequency equalization. Many of these distortions are nonlinear, and it is difficult to analyze their effect in either a spatial- or frequency-based method. However, the fact that the original image is known allows many signal transformations to be undone, at least approximately. For example, histogram equalization, a common nonlinear contrast enhancement method, may be removed substantially by histogram specification [10] or dynamic histogram warping [7] techniques.

Finally, the copied image may not remain in digital form. Instead, it is likely to be printed, or an analog recording made (onto analog audio or video tape). These reproductions introduce additional degradation into the image that a watermarking scheme must be robust to.

The watermark must not only be resistant to the inadvertent application of the aforementioned distortions. It must also

be immune to intentional manipulation by malicious parties. These manipulations can include combinations of the above distortions, and can also include collusion and forgery attacks, which are discussed in Section IV-E.

A. Spread Spectrum Coding of a Watermark

The above discussion illustrates that the watermark should *not* be placed in perceptually insignificant regions of the image (or its spectrum), since many common signal and geometric processes affect these components. For example, a watermark placed in the high-frequency spectrum of an image can be easily eliminated with little degradation to the image by any process that directly or indirectly performs lowpass filtering. The problem then becomes how to insert a watermark into the most perceptually significant regions of the spectrum in a fidelity preserving fashion. Clearly, any spectral coefficient may be altered, provided such modification is small. However, very small changes are very susceptible to noise.

To solve this problem, the frequency domain of the image or sound at hand is viewed as a *communication channel*, and correspondingly, the watermark is viewed as a signal that is transmitted through it. Attacks and unintentional signal distortions are thus treated as noise that the immersed signal must be immune to. While we use this methodology to hide watermarks in data, the same rationale can be applied to sending any type of message through media data.

We originally conceived our approach by analogy to spread spectrum communications [20]. In spread spectrum communications, one transmits a narrowband signal over a much larger bandwidth such that the signal energy present in any single frequency is undetectable. Similarly, the watermark is spread over very many frequency bins so that the energy in any one bin is very small and certainly undetectable. Nevertheless, because the watermark verification process knows the location and content of the watermark, it is possible to concentrate these many weak signals into a single output with high signal-to-noise ratio (SNR). However, to destroy such a watermark would require noise of high amplitude to be added to *all* frequency bins.

Spreading the watermark throughout the spectrum of an image ensures a large measure of security against unintentional or intentional attack: First, the location of the watermark is not obvious. Furthermore, frequency regions should be selected in a fashion that ensures severe degradation of the original data following any attack on the watermark.

A watermark that is well placed in the frequency domain of an image or a sound track will be practically impossible to see or hear. This will always be the case if the energy in the watermark is sufficiently small in any single frequency coefficient. Moreover, it is possible to increase the energy present in particular frequencies by exploiting knowledge of masking phenomena in the human auditory and visual systems. Perceptual masking refers to any situation where information in certain regions of an image or a sound is occluded by perceptually more prominent information in another part of the scene. In digital waveform coding, this frequency domain (and, in some cases, time/pixel domain) masking is exploited extensively to achieve low bit rate encoding of data [9], [12]. It is known that both the auditory and visual systems attach more resolution to the high-energy, low-frequency, spectral regions of an auditory or visual scene [12]. Further, spectrum analysis of images and sounds reveals that most of the information in such data is located in the low-frequency regions.

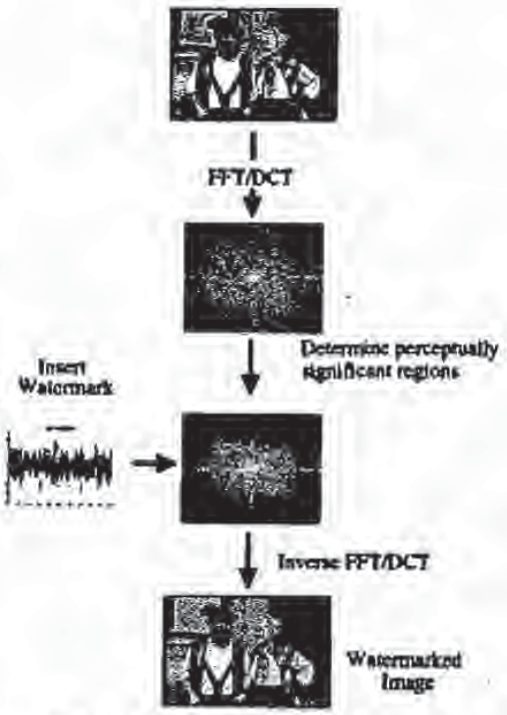


Fig. 2. Stages of watermark insertion process.

Fig. 2 illustrates the general procedure for frequency domain watermarking. Upon applying a frequency transformation to the data, a *perceptual mask* is computed that highlights perceptually significant regions in the spectrum that can support the watermark without affecting perceptual fidelity. The watermark signal is then inserted into these regions in a manner described in Section IV-B. The precise magnitude of each modification is only known to the owner. By contrast, an attacker may only have knowledge of the possible range of modification. To be confident of eliminating a watermark, an attacker must assume that each modification was at the limit of this range, despite the fact that few such modifications are typically this large. As a result, an attack creates visible (or audible) defects in the data. Similarly, unintentional signal distortions due to compression or image manipulation, must leave the perceptually significant spectral components intact, otherwise the resulting image will be severely degraded. This is why the watermark is robust.

In principle, any frequency domain transform can be used. However, in the experimental results of Section VI we use a Fourier domain method based on the DCT [16], although we are currently exploring the use of wavelet-based schemes as a variation. In our view, each coefficient in the frequency domain has a *perceptual capacity*, that is, a quantity of additional

information can be added without any (or with minimal) impact to the perceptual fidelity of the data. To determine the perceptual capacity of each frequency, one can use models for the appropriate perceptual system or simple experimentation.

In practice, in order to place a length n watermark into an $N \times N$ image, we computed the $N \times N$ DCT of the image and placed the watermark into the n highest magnitude coefficients of the transform matrix, excluding the DC component.[1] For most images, these coefficients will be the ones corresponding to the low frequencies.

In the next section, we provide a high level discussion of the watermarking procedure, describing the structure of the watermark and its characteristics.

IV. Structure of the Watermark

We now give a high-level overview of our a basic watermarking scheme; many variations are possible. In its most basic implementation, a watermark consists of a sequence of real numbers $X = x_1, \cdots, x_n$. In practice, we create a watermark where each value x_i is chosen independently according to $N(0,1)$ (where $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2). We assume that numbers are represented by a reasonable but finite precision and ignore these insignificant roundoff errors. Section IV-A introduces notation to describe the insertion and extraction of a watermark and Section IV-D describes how two watermarks (the original one and the recovered, possibly corrupted one) can be compared. This procedure exploits the fact that each component of the watermark is chosen from a normal distribution. Alternative distributions are possible, including choosing x_i uniformly from $\{1, -1\}$, $\{0, 1\}$ or $[0, 1]$. However, as we discuss in IV-D, using such distributions leaves one particularly vulnerable to attacks using multiple watermarked documents.

A. Description of the Watermarking Procedure

We extract from each document D a sequence of values $V = v_1, \cdots, v_n$, into which we insert a watermark $X = x_1, \cdots, x_n$ to obtain an adjusted sequence of values $V' = v'_1, \cdots, v'_n$. V' is then inserted back into the document in place of V to obtain a watermarked document D'. One or more attackers may then alter D', producing a new document D^*. Given D and D^*, a possibly corrupted watermark X^* is extracted and is compared to X for statistical significance. We extract X^* by first extracting a set of values $V^* = v^*_1, \ldots, v^*_n$ from D^* (using information about D) and then generating X^* from V^* and V.

Frequency-domain based methods for extracting V and V^* and inserting V' are given in Section III. For the rest of this section, we ignore the manipulations of the underlying documents.

[1] More generally, n randomly chosen coefficients could be chosen from the M, $M \geq n$ most perceptually significant coefficients of the transform. The choice of appropriate components remains a subject of research.

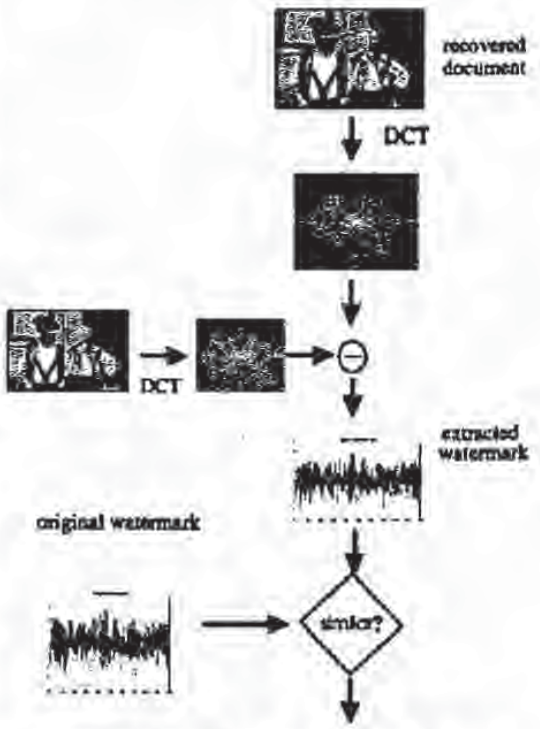


Fig. 3. Encoding and decoding of the watermark string.

B. Inserting and Extracting the Watermark

When we insert X into V to obtain V' we specify a scaling parameter α, which determines the extent to which X alters V. Three natural formulae for computing V' are

$$v'_i = v_i + \alpha x_i \quad (1)$$

$$v'_i = v_i(1 + \alpha x_i) \quad (2)$$

$$v'_i = v_i(e^{\alpha x_i}). \quad (3)$$

Equation (1) is always invertible, and (2) and (3) are invertible if $v_i \neq 0$, which holds in all of our experiments. Given V^*, we can therefore compute the inverse function to derive X^* from V^* and V.

Equation (1) may not be appropriate when the v_i values vary widely. If $v_i = 10^6$ then adding 100 may be insufficient for establishing a mark, but if $v_i = 10$ adding 100 will distort this value unacceptably. Insertion based on (2) or (3) are more robust against such differences in scale. We note that (2) and (3) give similar results when αx_i is small. Also, when v_i is positive, then (3) is equivalent to $\lg(v'_i) = \lg(v_i) + \alpha x_i$, and may be viewed as an application of (1) to the case where the logarithms of the original values are used.

1) Determining Multiple Scaling Parameters: A single scaling parameter α may not be applicable for perturbing all of the values v_i, since different spectral components may exhibit more or less tolerance to modification. More generally one can have multiple scaling parameters $\alpha_1, \cdots, \alpha_n$ and use update rules such as $v'_i = v_i(1 + \alpha_i x_i)$. We can view α_i as a relative measure of how much one must alter v_i to alter the perceptual quality of the document. A large α_i means that one

can perceptually "get away" with altering v_i by a large factor without degrading the document.

There remains the problem of selecting the multiple scaling values. In some cases, the choice of α_i may be based on some general assumption. For example, (2) is a special case of the generalized (1) ($v_i' = v_i + \alpha_i x_i$), for $\alpha_i = \alpha v_i$. Essentially, (2) makes the reasonable assumption that a large value is less sensitive to additive alterations than a small value.

In general, one may have little idea of how sensitive the image is to various values. One way of empirically estimating these sensitivities is to determine the distortion caused by a number of attacks on the original image. For example, one might compute a degraded image D^* from D, extract the corresponding values $v_1^*, \cdots, v_n^*$ and choose α_i to be proportional to the deviation $|v_i^* - v_i|$. For greater robustness, one should try many forms of distortion and make α_i proportional to the average value of $|v_i^* - v_i|$. As alternatives to taking the average deviation one might also take the median or maximum deviation.

One may combine this empirical approach with general global assumptions about the sensitivity of the values. For example, one might require that $\alpha_i \geq \alpha_j$ whenever $v_i \geq v_j$. One way to combine this constraint with the empirical approach would be to set α_i according to

$$\alpha_i \sim \max_{j | v_j \leq v_i} |v_j^* - v_j|.$$

A still more sophisticated approach would be to weaken the monotonicity constraint to be robust against occasional outliers.

In all our experiments we simply use (2) with a single parameter $\alpha = 0.1$. When we computed JPEG-based distortions of the original image, we observed that the higher energy frequency components were not altered proportional to their magnitude [the implicit assumption of (2)]. We suspect that we could make a less obtrusive mark of equal strength by attenuating our alterations of the high-energy components and amplifying our alterations of the lower energy components. However, we have not yet performed this experiment.

C. Choosing the Length, n, of the Watermark

The choice of n dictates the degree to which the watermark is spread out among the relevant components of the image. In general, as the number of altered components are increased the extent to which they must be altered decreases. For a more quantitative assessment of this tradeoff, we consider watermarks of the form $v_i' = v_i + \alpha x_i$ and model a white noise attack by $v_i^* = v_i' + r_i$ where r_i are chosen according to independent normal distributions with standard deviation σ. For the watermarking procedure we described below, one can recover the watermark when α is proportional to $\sigma/\sqrt{n}$. That is, by quadrupling the number of components used, one can halve the magnitude of the watermark placed into each component. Note that the sum of squares of the deviations will be essentially unchanged.

Note that the number of bits of information associated with the watermark can be arbitrary—the watermark is simply used as an index to a database entry associated with the watermark.

D. Evaluating the Similarity of Watermarks

It is highly unlikely that the extracted mark X^* will be identical to the original watermark X. Even the act of requantizing the watermarked document for delivery will cause X^* to deviate from X. We measure the similarity of X and X^* by

$$\mathrm{sim}(X, X^*) = \frac{X^* \cdot X}{\sqrt{X^* \cdot X^*}}. \tag{4}$$

Many other measures are possible, including the standard correlation coefficient. Further variations on this basic metric are discussed in IV-D2. To decide whether X and X^* match, one determines whether $\mathrm{sim}(X, X^*) > T$, where T is some threshold. Setting the detection threshold is a classical decision estimation problem in which we wish to minimize both the rate of false negatives (missed detections) and false positives (false alarms) [23]. We have chosen our measure so that it is particularly easy to determine the probability of false positives.

1) Computing the Probability of False Positives: There is always the possibility that X and X^* will be very similar purely by random chance; hence, any similarity metric will give "significant" values that are spurious. We analyze the probability of such false positives as follows. Suppose that the creators of document D^* had no access to X (either through the seller or through a watermarked document). Then, even conditioned on any fixed value for X^*, each x_i will be independently distributed according to $N(0,1)$. That is, X is independent of X^*.

The distribution on $X^* \cdot X$ may be computed by first writing it as $\sum_{i=1}^{n} x_i^* x_i$, where x_i^* is a constant. Using the well-known formula for the distribution of a linear combination of variables that are independent and normally distributed, $X^* \cdot X$ will be distributed according to

$$N\left(0, \sum_{i=1}^{n} x_i^{*2}\right) = N(0, X^* \cdot X^*),$$

Thus, $\mathrm{sim}(X, X^*)$ is distributed according to $N(0,1)$. We can then apply the standard significance tests for the normal distribution. For example, if X^* is created independently from X then the probability that $\mathrm{sim}(X, X^*) > 6$ is the probability of a normally distributed random variable exceeding its mean by more than six standard deviations.

Hence, for a small number of documents, setting the threshold at T equal to six will cause spurious matchings to be extremely rare. Of course, the number of tests to be performed must be considered in determining what false positive probability is acceptable. For example, if one tests an extracted watermark X^* against 10^6 watermarks, then the probability of a false positive is increased by a multiplicative factor of 10^6 as well.

We note that our similarity measure and the false-positive probability analysis does not depend on n, the size of the watermark. However, n implicitly appears, since for example, $\mathrm{sim}(X, X)$ is likely to be around $\sqrt{n}$ when X is generated in the prescribed manner. As a rule of thumb, larger values of n tend to cause larger similarity values when X and X^* are genuinely related (e.g., X^* is a distorted version of X),

Fig. 4. Bavarian couple image courtesy of Corel Stock Photo Library.

Fig. 5. Watermarked version of Bavarian couple.

without causing larger similarity values when X and X^* are independent. This benefit must be balanced against the tendency for the document to be more distorted when n is larger.

a) A remark on quantization: In the above analysis, we treated all of the vectors as consisting of ideal real numbers. In practice, the actual values inserted will be quantized to some extent. Nevertheless, it is simpler to view the watermarks as real numbers and the quantization process as yet another form of distortion. Our analysis of false positives does not depend on the distribution or even the domain of possible X^*, and hence holds regardless of quantization effects.

There is an additional, extremely low-order quantization effect that occurs because X is generated with only finite precisions. However, this effect is caused only by the arithmetic precision, and not on the constraints imposed by the document. If each $x_i \in X$ is stored as a double-precision real number, the difference between the calculated value of $\mathrm{sim}(X, X^*)$ and its "ideal" value will be quite small for any reasonable n and any reasonable bound on the dynamic range of X^*.

2) Robust Statistics The above analysis required only the independence of X from X^*, and did not rely on any specific properties of X^* itself. This fact gives us further flexibility when it comes to preprocessing X^*. We can process X^* in a number of ways to potentially enhance our ability to extract a watermark. For example, in our experiments on images we encountered instances where the average value of x_i^*, denoted $E_i(X^*)$, differed substantially from zero, due to the effects of a dithering procedure. While this artifact could be easily eliminated as part of the extraction process, it provides a motivation for postprocessing extracted watermarks. We found that the simple transformation $x_i^* \leftarrow x_i^* - E_i(X^*)$ yielded superior values of $\mathrm{sim}(X, X^*)$. The improved performance resulted from the decreased value of $X^* \cdot X^*$; the value of $X^* \cdot X$ was only slightly affected.

In our experiments, we frequently observed that x_i^* could be greatly distorted for some values of i. One postprocessing

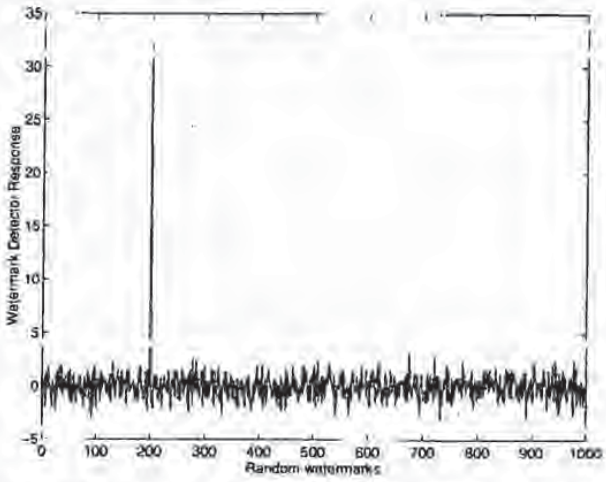


Fig. 6. Watermark detector response to 1000 randomly generated watermarks. Only one watermark (the one to which the detector was set to respond) matches that present in Fig. 5.

option is to simply ignore such values, setting them to zero. That is

$$x_i^* \leftarrow \begin{cases} x_i^*, & \text{if } |x_i^*| \le \text{tolerance} \\ 0, & \text{otherwise.} \end{cases}$$

Again, the goal of such a transformation is to lower $X^* \cdot X^*$. A less abrupt version of this approach is to normalize the X^* values to be either $-1, 0$ or 1, by

$$x_i^* \leftarrow \mathrm{sign}(x_i^* - E_i(X^*)).$$

This transformation can have a dramatic effect on the statistical significance of the result. Other robust statistical techniques could also be used to suppress outlier effects [11].

A natural question is whether such postprocessing steps run the risk of generating false positives. Indeed, the same potential risk occurs whenever there is any latitude in the

(a) (b)

Fig. 7. (a) Lowpass filtered, 0.5 scaled image of Bavarian couple. (b) Rescaled image showing noticeable loss of fine detail.

procedure for extracting X^* from D^*. However, as long as the method for generating a set of values for X^* depends solely on D and D^*, our statistical significance calculation is unaffected. The only caveat to be considered is that the bound on the probability that one of $X_1^*, \cdots X_k^*$ generates a false positive is the sum of the individual bounds. Hence, to convince someone that a watermark is valid, it is necessary to have a published and rigid extraction and processing policy that is guaranteed to only generate a small number of candidate X^*.

E. Resilience to Multiple-Document (Collusion) Attacks

The most general attack consists of using t multiple watermarked copies $D_1', \cdots, D_t'$ of document D to produce an unwatermarked document D^*. We note that most schemes proposed seem quite vulnerable to such attacks. As a theoretical exception, Boneh and Shaw [5] propose a coding scheme for use in situations in which one can insert many relatively weak $0/1$ watermarks into a document. They assume that if the ith watermark is the same for all t copies of the document then it cannot be detected, changed or removed. Using their coding scheme, the number of weak watermarks to be inserted scales according to t^4, which may limit its usefulness in practice.

To illustrate the power of multiple-document attacks, consider watermarking schemes in which v_i' is generated by either adding 1 or -1 at random to v_i. Then as soon as one finds two documents with unequal values for v_i', one can determine v_i and, hence, completely eliminate this component of the watermark. With t documents one can, on average, eliminate all but a 2^{1-t} fraction of the components of the watermark. Note that this attack does not assume anything about the distribution on v_i. While a more intelligent allocation of ± 1 values to the watermarks (following [5] and [17]) will better resist this simple attack, the discrete nature of the watermark components makes them much easier to completely eliminate. Our use of continuous valued watermarks appears to give greater resilience to such attacks. Interestingly, we have experimentally determined that if one chooses the x_i uniformly over some range, then one can remove the watermark using only five documents.

Use of the normal distribution seems to give better performance than the distributions considered above. We note that the crucial performance measure to consider is the value of $\max_i(X^* \cdot X_i)$, where X^* is the watermark extracted from an document D^* generated by attacking documents $D_1, \cdots, D_t$, with respective watermarks $X_1, \cdots, X_t$. The denominator $\sqrt{X^* \cdot X^*}$ of our similarity measure can always be made larger by, for example, adding noise. This causes the similarity measure to shrink, at the expense of distorting the image. Hence, we can view $\max_i(X^* \cdot X_i)$ as determining a fidelity/undetectability tradeoff curve and the value of $\sqrt{X^* \cdot X^*}$ as picking a point on this curve.

When X_i is inserted into D by a linear update rule, then an averaging attack, which sets

$$D^* = \frac{D_1 + \cdots + D_t}{t}$$

will result in

$$X^* = \frac{X_1 + \cdots + X_t}{t}.$$

In this case,

$$\max_i(X^* \cdot X_i) \approx \frac{1}{t} \max_i(X_i \cdot X_i) \text{ (assuming } X_i X_j \approx 0).$$

That is, there is a $1/t$ behavior in the detector output.

Note that with a naive averaging attack, the denominator, $\sqrt{X^* \cdot X^*}$, will be a (roughly) $1/\sqrt{t}$ factor smaller, so $\max_i \text{sim}(X_i, X^*)$ will be roughly $\sqrt{n}/\sqrt{t}$. However, as mentioned before, additional noise can be added so that the extracted watermark, X^*, has the same power as any of the original watermarks X_i. Then $\max_i \text{sim}(X_i, X^*)$ will be

Fig. 8. JPEG encoded version of Bavarian couple with 10% quality and 0% smoothing.

Fig. 9. JPEG encoded version of Bavarian couple with 5% quality and 0% smoothing.

roughly $\sqrt{n}/t$. Thus, the similarity measure can be shrunk by a factor of t.

We do not know of any more effective multidocument attack on normally distributed watermarks. In a forthcoming paper (see http://www.neci.nj.nec.com/tr/index.html), a more theoretical justification is given for why it is hard to achieve more than an $O(t)$ reduction in the similarity measure.

V. Experimental Results

In order to evaluate the proposed watermarking scheme, we took the Bavarian couple[2] image of Fig. 4 and produced the watermarked version of Fig. 5. We then subjected the watermarked image to a series of image processing and collusion style attacks. These experiments are preliminary, but show resilience to certain types of common processing. Of note is our method's resistance to compression such as JPEG, and data conversion (printing, xeroxing and scanning). Note that in the case of affine transforms, registration to the original image is crucial to successful extraction.

In all experiments, a watermark length of 1000 was used. We added the watermark to the image by modifying 1000 of the more perceptually significant components of the image spectrum using (2). More specifically, the 1000 largest coefficients of the DCT (excluding the DC term) were used. A fixed scale factor of 0.1 was used throughout.

A. Experiment 1: Uniqueness of Watermark

Fig. 6 shows the response of the watermark detector to 1000 randomly generated watermarks of which only one matches the watermark present in Fig. 5. The positive response due to the correct watermark is very much stronger that the response to incorrect watermarks, suggesting that the algorithm has very low false positive response rates.

[2]The common test image Lenna was originally used in our experiments, and similar results were obtained. However, Playboy Inc. refused to grant copyright permission for electronic distribution.

Fig. 10. Dithered version of the Bavarian couple image.

B. Experiment 2: Image Scaling

We scaled the watermarked image to half of its original size, as shown in Fig. 7(a). In order to recover the watermark, the quarter-sized image was rescaled to its original dimensions, as shown in Fig. 7(b), in which it is clear that considerable fine detail has been lost in the scaling process. This is to be expected since subsampling of the image requires a lowpass spatial filtering operation. The response of the watermark detector to the original watermarked image of Fig. 5 was 32.0, which compares to a response of 13.4 for the rescaled version of Fig. 7(b). While the detector response is down by over 50%, the response is still well above random chance

(a) (b)

Fig. 11. (a) Clipped version of watermarked Bavarian couple. (b) Restored version of Bavarian couple in which missing portions have been replaced with imagery from the original unwatermarked image of Fig. 4.

levels suggesting that the watermark is robust to geometric distortions. Moreover, it should be noted that 75% of the original data is missing from the scaled down image of Fig. 7.[3]

C. Experiment 3: JPEG Coding Distortion

Fig. 8 shows a JPEG encoded version of the Bavarian couple image with parameters of 10% quality and 0% smoothing, which results in clearly visible distortions of the image. The response of the watermark detector is 22.8, again suggesting that the algorithm is robust to common encoding distortions. Fig. 9 shows a JPEG encoded version of Bavarian couple with parameters of 5% quality and 0% smoothing, which results is very significant distortions of the image. The response of the watermark detector in this case is 13.9, which is still well above random.

D. Experiment 4: Dithering Distortion

Fig. 10 shows a dithered version of Bavarian couple. The response of the watermark detector is 5.2, again suggesting that the algorithm is robust to common encoding distortions. In fact, more reliable detection can be achieved simply by removing any nonzero mean from the extracted watermark, as discussed in Section IV-D2. In this case the detection value is 10.5.

E. Experiment 5: Cropping

Fig. 11(a) shows a cropped version of the watermarked image of Fig. 5 in which only the central quarter of the image remains. In order to extract the watermark from this image, the missing portions of the image were replaced with portions from the original *unwatermarked* image of Fig. 4, as shown in Fig. 11(b). In this case, the response of the watermark is 14.6. Once again, this is well above random even though 75% of the data has been removed.

Fig. 12(a) shows a clipped version of the JPEG encoded image of Fig. 8 in which only the central quarter of the image remains. As before, the missing portions of the image were replaced with portions from the original *unwatermarked* image of Fig. 4, as shown in Fig. 12(b). In this case, the response of the watermark is 10.6. Once more, this is well above random even though 75% of the data has been removed and distortion is present in the clipped portion of the image.

F. Experiment 6: Print, Xerox, and Scan

Fig. 13 shows an image of the Bavarian Couple after 1) printing, 2) xeroxing, then 3) scanning at 300 dpi using a UMAX PS-2400X scanner, and finally 4) rescaling to a size of 256 × 256. Clearly, this image suffers from several levels of distortion that accompany each of the four stages. High-frequency pattern noise is especially noticeable. The detector response to the watermark is 4.0. However, if the nonzero mean is removed and only the sign of the elements of the watermark are used, then the detector response is 7.0, which is well above random.

G. Experiment 7: Attack by Watermarking Watermarked Images

Fig. 14 shows an image of Bavarian Couple after five successive watermarking operations, i.e., the original image is watermarked, the watermarked image is watermarked, etc. This may be considered another form of attack in which it is clear that significant image degradation eventually occurs as the process is repeated. This attack is equivalent to adding noise to the frequency bins containing the watermark. Interestingly, Fig. 15 shows the response of the detector to

[3] However, subsequent experiments have revealed that if small changes of scale are not corrected, then the response of the watermark detector is severely degraded.

(a)

(b)

Fig. 12. (a) Clipped version of JPEG encoded (10% quality, 0% smoothing) Bavarian couple. (b) Restored version of Bavarian couple in which missing portions have been replaced with imagery from the original unwatermarked image of Fig. 4.

Fig. 13. Printed, xeroxed, scanned, and rescaled image of Bavarian couple.

Fig. 14. Image of Bavarian couple after five successive watermarks have been added.

1000 randomly generated watermarks, which include the five watermarks present in the image. Five spikes clearly indicate the presence of the five watermarks and demonstrate that successive watermarking does not unduly interfere with the process.

H. Experiment 8: Attack by Collusion

In a similar experiment, we took five separately watermarked images and averaged them to form Fig. 16 in order to simulate a simple collusion attack. As before, Fig. 17 shows the response of the detector to 1000 randomly generated watermarks, which include the five watermarks present in the image. Once again, five spikes clearly indicate the presence of the five watermarks and demonstrate that simple collusion based on averaging a few images is an ineffective attack.

VI. Conclusion

A need for electronic watermarking is developing as electronic distribution of copyright material becomes more prevalent. Above, we outlined the necessary characteristics of such a watermark. These are: fidelity preservation, robustness to common signal and geometric processing operations, robustness to attack, and applicability to audio, image and video data.

To meet these requirements, we propose a watermark whose structure consists of k i.i.d. random numbers drawn from a $N(0,1)$ distribution. We rejected a binary watermark because it is far less robust to attacks based on collusion of several independently watermarked copies of an image. The length of the watermark is variable and can be adjusted to suit the characteristics of the data. For example, longer watermarks

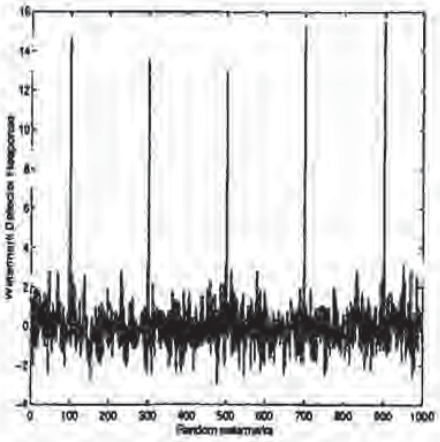


Fig. 15. Watermark detector response to 1000 randomly generated watermarks (including the five specific watermarks) for the watermarked image of Fig. 14. Each of the five watermarks is clearly indicated.

Fig. 16. Image of Bavarian couple after averaging together five independently watermarks versions of the Bavarian couple image.

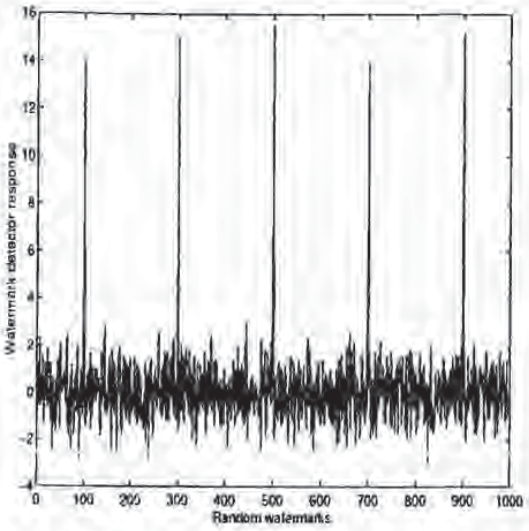


Fig. 17. Watermark detector response to 1000 randomly generated watermarks (including the five specific watermarks) for the watermarked image of Fig. 16. Each of the five watermarks is clearly detected, indicating that collusion by averaging is ineffective.

may be used for an image that is especially sensitive to large modifications of its spectral coefficients, thus requiring weaker scaling factors for individual components.

We recommend that the watermark be placed in the perceptually *most* significant components of the image spectrum. This maximizes the chances of detecting the watermark even after common signal and geometric distortions. Further, modification of these spectral components results in severe image degradation long before the watermark itself is destroyed. Of course, to insert the watermark, it is necessary to alter these very same coefficients. However, each modification can be extremely small and, in a manner similar to spread spectrum communication, a strong narrowband watermark may be distributed over a much broader image (channel) spectrum. We have not performed an objective evaluation of the image quality, in part because the image quality can be adjusted to any desired quality by altering the relative power of the watermark using the scale factor term. Of course, as the watermark strength is reduced to improve the image quality, the robustness of the method is also reduced. It will ultimately be up to content owners to decide what image degradation and what level of robustness is acceptable. This will vary considerably from application to application.

Detection of the watermark then proceeds by adding all of these very small signals, and concentrating them once more into a signal with high SNR. Because the magnitude of the watermark at each location is only known to the copyright holder, an attacker would have to add much more noise energy to each spectral coefficient in order to be sufficiently confident of removing the watermark. However, this process would destroy the image fidelity.

In our experiments, we added the watermark to the image by modifying the 1000 largest coefficients of the DCT (excluding the DC term). These components are heuristically perceptually more significant than others. An important open problem is the construction of a method that would identify perceptually significant components from an analysis of the image and the human perceptual system. Such a method may include additional considerations regarding the relative predictability of a frequency based on its neighbors. The latter property is important in combating attacks that may use statistical analyzes of frequency spectra to replace components with their maximum likelihood estimate. For example, the choice of the DCT is not critical to the algorithm and other spectral transforms, including wavelet type decompositions, are also possible.

We showed, using the Bavarian couple image, that our algorithm can extract a reliable copy of the watermark from imagery that we degraded with several common geometric and signal processing procedures. An important caveat here is that any affine geometric transformation must first be inverted. These procedures include translation, rotation, scale

change, and cropping. The algorithm displays strong resilience to lossy operations such as aggressive scale changes, JPEG compression, dithering and data conversion. The experiments presented are preliminary, and should be expanded in order to validate the results. We are conducting ongoing work in this area. Further, the degree of precision of the registration procedures used in undoing affine transforms must be characterized precisely across a large test set of images.

Application of the method to color images is straightforward. The most common transformation of a color image is to convert it to black and white. Color images are therefore converted into a YIQ representation and the brightness component Y is then watermarked. The color image can then be converted to other formats, but must be converted back to YIQ prior to extraction of the watermark. We therefore expect color images to be robust to the signal transformations we applied to gray-level images. However, robustness to certain color image processing procedures should be investigated. Similarly, the system should work well on text images, however, the binary nature of the image together with its much more structured spectral distribution need more work. We expect that our watermarking methodology should extend straightforwardly to audio and video data. However, special attention must be paid to the time-varying nature of these data.

Broader systems issues must be also addressed in order for this system to be used in practice. For example, it would be useful to be able to prove in court that a watermark is present without publicly revealing the original, unmarked document. This is not hard to accomplish using secure trusted hardware; an efficient purely cryptographic solution seems much more difficult. It should also be noted that the current proposal only allows the watermark to be extracted by the owner, since the original unwatermarked image is needed as part of the extraction process. This prohibits potential users from querying the image for ownership and copyright information. This capability may be desirable but appears difficult to achieve with the same level of tamper resistance. However, it is straightforward to provide if a much weaker level of protection is acceptable and might therefore be added as a secondary watermarking procedure. Finally, we note that while the proposed methodology is used to hide watermarks in data, the same process can be applied to sending other forms of message through media data.

Acknowledgment

I. Cox and T. Shamoon thank L. O'Gorman of AT&T Bell Laboratories for bringing this problem to their attention, and S. Roy for testing the robustness of the algorithm. I. Cox thanks H. Stone for advice on image transforms.

References

[1] E. H. Adelson, "Digital signal encoding and decoding apparatus," U.S. Patent 4939515, 1990.
[2] G. W. Braudaway, K. A. Magerlein, and F. C. Mintzer, "Color correct digital watermarking of images," U.S. Patent 5530759, 1996.
[3] W. Bender, D. Gruhl, and N. Morimoto, "Techniques for data hiding," in *Proc. SPIE*, vol. 2420, p. 40, Feb. 1995.
[4] J. Brassil, S. Low, N. Maxemchuk, and L. O'Gorman, "Electronic marking and identification techniques to discourage document copying," in *Proc. Infocom'94*, pp. 1278–1287.
[5] D. Bonch and J. Shaw, "Collusion-secure fingerprinting for digital data," in *Advances in Cryptology: Proc. CRYPTO'95*. New York: Springer-Verlag, 1995.
[6] G. Caronni, "Assuring ownership rights for digital images," in *Proc. Reliable IT Systems, VIS'95*.
[7] I. J. Cox, S. Roy, and S. L. Hingorani, "Dynamic histogram warping of images pairs for constant image brightness," in *IEEE Int. Conf. Image Processing*, 1995.
[8] O. Faugeras, *Three Dimensional Computer Vision: A Geometric Viewpoint*. Cambridge, MA: MIT Press, 1993.
[9] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Boston, MA: Kluwer, 1992.
[10] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. New York: Addison-Wesley, 1993.
[11] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.
[12] N. Jayant, J. Johnston, and R. Safranek, "Signal compression based on models of human perception," in *Proc. IEEE*, vol. 81, no. 10, 1993.
[13] J. Kilian *et al.*, "Resistance of watermarked documents to collusional attacks," in preparation.
[14] E. Koch, J. Rindfrey, and J. Zhao, "Copyright protection for multimedia data," in *Proc. Int. Conf. Digital Media and Electronic Publishing*, 1994.
[15] E. Koch and Z. Zhao, "Toward robust and hidden image copyright labeling," in *Proc. 1995 IEEE Workshop on Nonlinear Signal and Image Processing*, June 1995.
[16] J. S Lim, *Two-Dimensional Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1990.
[17] F. T. Leighton and S. Micali, "Secret-key agreement without public-key cryptography," in *Proc. Cryptology*, 1993.
[18] B. M. Macq and J.-J. Quisquater, "Cryptology for digital TV broadcasting," in *Proc. IEEE*, vol. 83, pp. 944–957, 1995.
[19] K. Matsui and K. Tanaka, "Video-steganography," in *Proc. IMA Intellectual Property Project*, 1994, vol. 1, pp. 187–206.
[20] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein, "Theory of spread spectrum communications—A tutorial," *IEEE Trans. Commun.*, vol. COMM-30, pp. 855–884, 1982.
[21] G. B. Rhoads, "Indentification/authentication coding method and apparatus," Rep. WIPO WO 95/14289, World Intellect. Property Org., 1995.
[22] W. F. Schreiber, A. E. Lippman, E. H. Adelson, and A. N. Netravali, "Receiver-compatible enhanced definition television system," U.S. Patent 5010405, 1991.
[23] C. W. Therrien, *Decision Estimation and Classification: An Introduction to Pattern Recognition and Related Topics*. New York: Wiley, 1989.
[24] K. Tanaka, Y. Nakamura, and K. Matsui, "Embedding secret information into a dithered multi-level image," in *Proc. 1990 IEEE Military Communications Conf.*, 1990, pp. 216–220.
[25] L. F. Turner, "Digital data security system," Patent IPN WO 89/08915, 1989.
[26] R. G. van Schyndel, A. Z. Tirkel, and C. F. Osborne, "A digital watermark," in *Int. Conf. Image Processing*, 1994, vol. 2, pp. 86–90.

Ingemar J. Cox (S'79–M'83–SM'95) received the Ph.D. degree from Oxford University, Oxford, U.K., in 1983.

From 1984 to 1989, he was a principal investigator in the Robotics Principles Department, AT&T Bell Laboratories, Murray Hill, NJ, where his research interests focused on issues of autonomous mobile robots. He joined NEC Research Institute, Princeton, NJ, as a senior research scientist in 1989. His principal research interests are broadly in computer vision, specifically tracking, stereo and 3-D estimation, and multimedia, especially image database retrieval and electronic watermarking for copyright protection.

Joe Kilian received the B.S. degree in computer science and in mathematics in 1985, and the Ph.D. in mathematics in 1989, both from the Massachusetts Institute of Technology, Cambridge.

He is a Research Scientist with NEC Research Institute, Princeton, NJ. His research interests are in complexity theory and cryptography.

F. Thomson Leighton received the B.S.E. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, in 1978, and the Ph.D. degree in applied mathematics from the Massachusetts Institute of Technology (MIT), Cambridge, in 1981.

He is a Professor of applied mathematics and a member of the Laboratory for Computer Science (LCS) at MIT. He was a Bantrell Postdoctoral Research Fellow at LCS from 1981 to 1983, and he joined the MIT faculty as an Assistant Professor of applied mathematics in 1982. He is a leader in the development of networks and algorithms for message routing in parallel machines, particularly in the use of randomness in wiring to overcome problems associated with congestion, blocking, and faults in networks. He has published over 100 research papers on parallel and distributed computing and related areas. He is the author of two books, including a leading text on parallel algorithms and architectures.

Talal Shamoon (S'84–M'95) received the Ph.D. degree in electrical engineering from Cornell University, Ithaca, NY, in January 1995.

He joined the NEC Research Institute (NECI), Princeton, NJ, in December of 1994, where he held the title of Scientist. He joined the InterTrust STAR Laboratory, Sunnyvale, CA, in 1997, where he is currently a Member of the Research Staff working on problems related to trusted rights management of multimedia content. His research interests include algorithms for audio, image and video coding and processing, multimedia security, data compression, and acoustic transducer design. He has worked on high-fidelity audio coding and fast search algorithms for large image data bases. Since joining NECI, he has been actively involved in research on watermarking for multimedia systems.

A Public Key Watermark for Image Verification and Authentication

Ping Wah Wong
Hewlett Packard Company
11000 Wolfe Road
Cupertino, CA 95014

Abstract

We propose in this paper a public key watermarking algorithm for image integrity verification. This watermark is capable of detecting any change made to an image, including changes in pixel values and image size. This watermark is important for several imaging applications, including trusted camera, legal usage of images, medical archiving of images, news reporting, commercial image transaction, and others. In each of these applications, it is important to verify that the image has not been manipulated and that the image was originated by either a specific camera or a specific user. The verification (the watermark extraction) procedure uses a public key as in public key cryptography, and hence it can be performed by any person without the secure exchange of a secret key. This is very important in many applications (e.g., trusted camera, new reporting) where the exchange of a secret key is either not possible or undesirable.

1 Introduction

Digital watermarking is a technique to insert a digital signature into an image so that the signature can be extracted for the purposes of ownership verification and/or authentication. This type of technology is becoming increasingly important due to the popularity of the usage of digital images on the World Wide Web and in electronic commerce.

There are different types of watermarking schemes that are designed for different applications [1, 2]. One type of watermark is designed to ensure the integrity of images, i.e., it can detect any change to an image as well as localizing the areas that have been changed. Since digital images can be altered or manipulated with ease, the ability to detect changes to digital images is very important for many applications such as news reporting, medical archiving, or legal usages. Another need for image authentication arises in, for example, electronic commerce where a buyer purchases a digital image from a seller, and then the seller transmits the digital image to the buyer over the network. In this case the buyer wants to ensure that the received image is indeed the genuine image sent by the seller. Here we not only want to verify the integrity of an image, we also want to check the original ownership.

Previously, the idea of a trusted camera [3] was proposed. This scheme computes for each captured image a standard digital signature. The digital signature is stored and transmitted along with the image. The integrity of the output digital image can be checked using standard digital signature techniques. Recently Yeung and Mintzer [4] propose a verification watermarking method based on indexing to a random sequence. This method detects changes to the pixel values, but it does not detect changes in image size due to scaling or cropping. Wong [5] proposes a secret key watermarking method where a user can detect any change to the pixel values and to the size of the image. The security of this method resides in a secret user key used in conjunction with a cryptographic hash function. Since this is a secret key scheme, only the user who has possession of the secret key can carry out the verification procedure. There is also the undesirable requirement whereby the secret key must be communicated through a separate secure channel.

In this paper, we extend the secret key verification watermark into a public key scheme so that the integrity and ownership of the image can be verified using a public key. In such a system, the owner of the image inserts a watermark using a private key K'. In the watermark extraction procedure, any person can use the public key K (corresponding to the private key K') to extract a watermark. Any change made to the watermarked image can be detected by a visual inspection of the extracted watermark. As in public key cryptographic systems [6, 7], the public key in this watermarking scheme can be published without compromising the security of the system.

0-8186-8821-1/98 $10.00 © 1998 IEEE

2 Watermark Insertion and Extraction

We describe in this section our public key authentication watermarking algorithm for grayscale images. For color images, the same technique can be applied independently to the color planes of the image, either in the RGB color space or in any other color space such as YUV.

Consider a grayscale image $x_{m,n}$ of size M by N pixels. We want to insert a binary watermark image $b_{m,n}$ to $x_{m,n}$ to obtain the watermarked image $y_{m,n}$. To this end, we partition the image into blocks of size I by J pixels, and insert a block of the watermark into each block of image data.

Let $a_{m,n}$ be a bi-level image that we will use as our watermark, to be embedded in $x_{m,n}$. Note that $a_{m,n}$ need not be of the same size as $x_{m,n}$. From $a_{m,n}$, we form another bi-level image $b_{m,n}$ of size M by N (same size as $x_{m,n}$). In our example, we form $b_{m,n}$ by tiling $a_{m,n}$, i.e., periodically replicating $a_{m,n}$ to the desired size. We then partition $b_{m,n}$ into blocks of I by J pixels. Each block of $b_{m,n}$ is then inserted into the corresponding block $x_{m,n}$ to give a watermarked block of $y_{m,n}$.

The watermark insertion and extraction procedures for each block are shown in Figs. 1 and 2, respectively.

2.1 Watermark Insertion

Let X_r denote the r^{th} block of data within the image $x_{m,n}$. We form the corresponding block $\widetilde{X}_r$ where each element in $\widetilde{X}_r$ equals the corresponding element in X_r except that the least significant bit is set to zero. Let $H(\cdot)$ be a cryptographic hash function such as the MD5 [8]. We compute the hash

$$H(M, N, \widetilde{X}_r) = (p_1^r, p_2^r, \ldots, p_s^r) \qquad (1)$$

where p_i^r denotes the output bits from the hash function, and s is size of the output bits that is dependent on the specific hash function used. For example, $s = 128$ for MD5. In our algorithm we need to make sure to select a block size such that $IJ \leq s$. Let P_r be the first IJ bits from the bit stream, i.e.,

$$P_r \triangleq (p_1^r, p_2^r, \ldots, p_{IJ}^r).$$

We combine P_r with a corresponding block B_r in $b_{m,n}$ using an exclusive or function. That is, we compute $W_r = P_r \oplus B_r$ where $\oplus$ denotes the element-wise exclusive OR operation between the two blocks. Finally we encrypt W_r with a public key cryptographic system [7] to give

$$C_r = E_{K'}(W_r)$$

where $E(\cdot)$ is the encryption function of the public key system, and K' is the private key. The binary block of data C_r is then embedded into the least significant bit of $\widetilde{X}_r$ to form a block Y_r of the watermarked image.

2.2 Watermark Extraction

In the extraction procedure, we split the input image block Z_r into two pieces; the first piece G_r contains the least significant bits, and the other piece $\widetilde{Z}_r$ contains the pixel values except that the least significant bits have been zeroed out. We then calculate the hash of M, N and $\widetilde{Z}_r$, and denote the first 64 bits of the output by Q_r. We use a public key decryption algorithm [7] to decrypt G_r with the *public key* K that corresponds to the private key K' used in the watermark insertion procedure. That is, we calculate

$$U_r = D_K(Z_r).$$

Finally, we compute the output block $O_r = Q_r \oplus U_r$ using an element-wise exclusive or procedure.

3 Experimental Results

We implemented both the public key watermark insertion and extraction procedures as described in the previous section. For the experiments, we used the MD5 [8] as our hash function, and the RSA public key encryption algorithm [7] for encryption and decryption. A vase image shown in Fig. 3 is used for testing the validity and properties of the algorithm. The binary watermark image is the logo image shown in the upper right hand corner of Fig. 3.

Note from the algorithm that if both the watermarked image block and the image size had not been changed since the insertion of a watermark, i.e., if $Z_r = Y_r$, then $\widetilde{Z}_r = \widetilde{X}_r$ and $G_r = C_r$. This implies $P_r = Q_r$ and $U_r = W_r$. Hence the output binary image O_r is identical to the block B_r. If the watermarked image was changed, the output block O_r will appear similar to random noise due to the nature of the hash function. As a result, this algorithm can detect any change to the pixel values to the block level. Since the block sizes are relatively small (we used 8 by 8 in our experiments), we consider the detection to be sufficiently localized.

Since the image size parameters M and N are used in the watermark insertion and extraction procedures of every block, any change in image size will result in the detection of changes in every block of the image. Hence the entire extracted watermark appears like random noise as shown in Fig. 3. In summary, this public key algorithm allows an authentication of image integrity as it can detect any change to an image including changes in pixel values and image size.

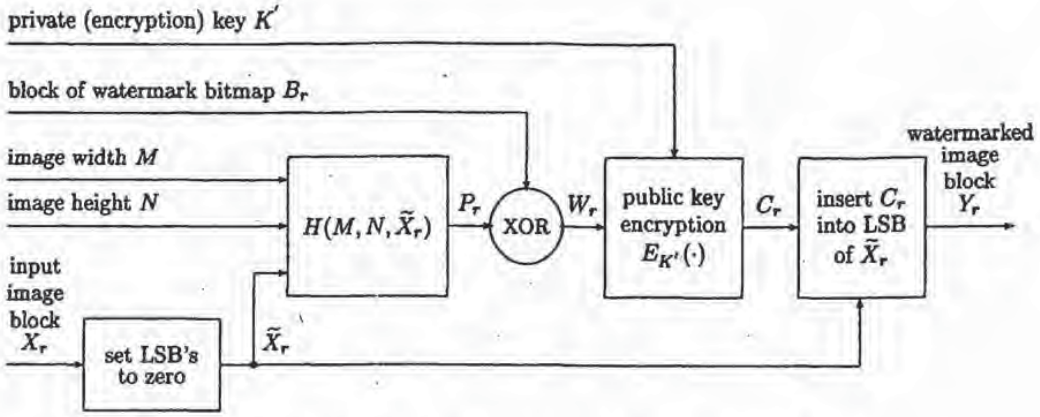


Figure 1: Public key verification watermark insertion procedure.

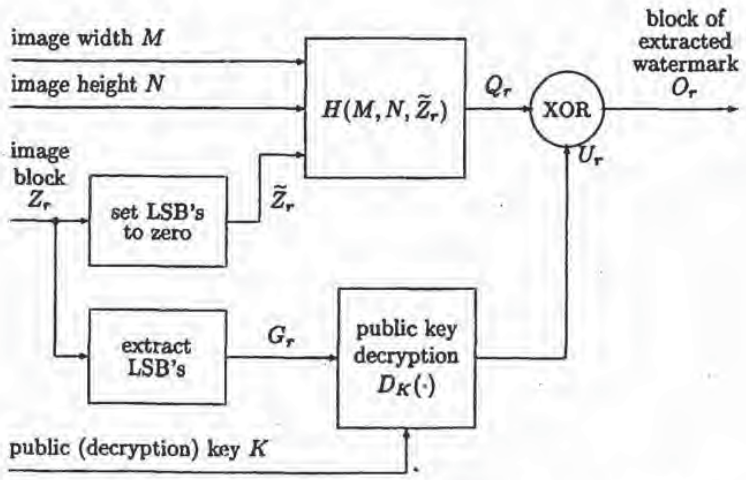


Figure 2: Public key verification watermark extraction procedure.

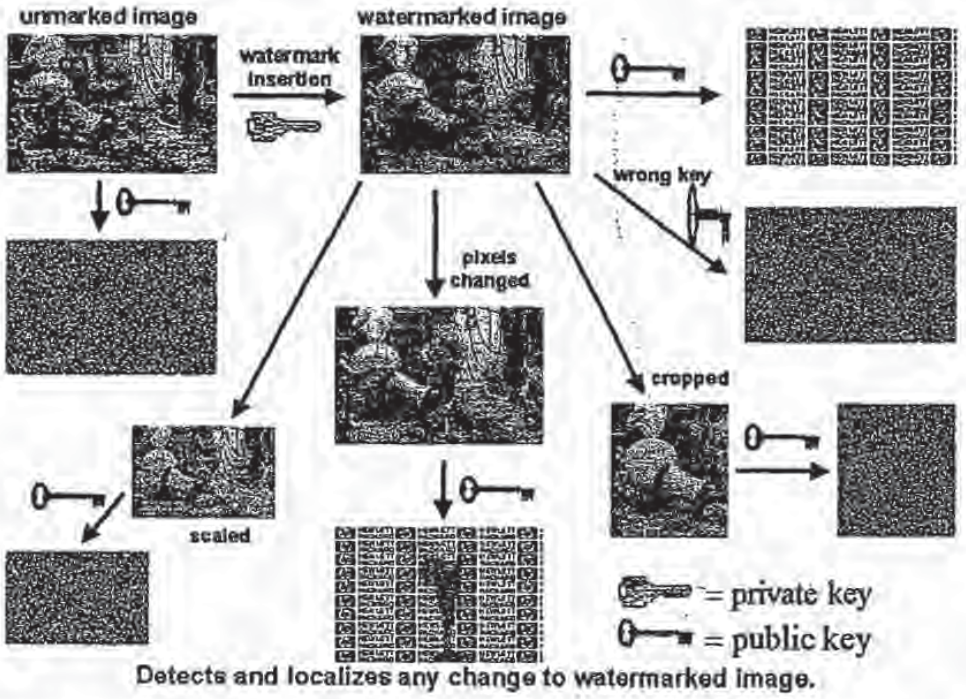


Figure 3: Experimental results summarizing the properties of the public key verification watermark.

4 Properties

As described in the previous section, this public key authentication watermark can detect any change to an image. The verification is performed using the public key of the owner, which also implies the original ownership of the image. Experimental results indicating the properties of this public key verification watermark is summarized in Fig. 3. Here we summarize the properties of the public key authentication watermark.

- The watermark is invisible.
- If one uses the correct public key K in the watermark extraction procedure, one obtains an appropriate watermark.
- If an image is unmarked, i.e., if it does not contain a watermark, the watermark extraction procedure returns an output that resembles random noise as shown in Fig. 3.
- If one applies an incorrect key (for example, if one uses the public key of different owner), then the watermark extraction procedure returns an output that resembles random noise.
- If a watermarked image is cropped or scaled, then the watermark extraction procedure returns an output that resembles random noise.
- If one changes certain pixels in the watermarked image, then the specific locations of the changes are reflected at the output of the watermark extraction procedure. This is shown in the middle and bottom part of Fig. 3 where a glass is pasted onto a watermarked image and the extracted watermark indicates the location of the glass.
- Despite embedding the watermark in the least significant bit of the image, the watermark is still secure. Recall that this watermark is designed for authentication purposes, i.e., to detect any change to the image. If someone attempts to remove the watermark by changing some bit planes of the image, the watermark extraction procedure will detect the changes.

5 Conclusion

We described a public key watermarking algorithm in this paper for image verification and authentication purposes. This is an extension of our previous work [5] on a secret key watermarking algorithm for image verification. The importance of the public key extension is that while a private key (secret) is used in watermark insertion, the watermark can be checked using a public key. As a result, any person can perform the integrity check using a public key without the secure exchange of a secret key.

References

[1] F. Mintzer, G. Braudaway, and M. M. Yeung, "Effective and ineffective digital watermarks," in *Proceedings of ICIP*, (Santa Barbara, CA), October 1997.

[2] N. Memon and P. W. Wong, "Protecting digital media content: Watermarks for copyrighting and authentication," *Communications of ACM*, July 1998.

[3] G. L. Friedman, "The trustworthy digital camera: restoring credibility to the photographic image," *IEEE Transactions on Consumer Electronics*, vol. 39, pp. 905–910, November 1993.

[4] M. M. Yeung and F. Mintzer, "An invisible watermarking technique for image verification," in *Proceedings of ICIP*, (Santa Barbara, CA), October 1997.

[5] P. W. Wong, "A watermark for image integrity and ownership verification," in *Proceedings of IS&T PIC Conference*, (Portland, OR), May 1998.

[6] W. Diffie and M. E. Hellman, "New directions in cryptography," *IEEE Transactions on Information Theory*, vol. 67, pp. 644–654, November 1976.

[7] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, pp. 120–126, February 1978.

[8] R. L. Rivest, "The MD5 message digest algorithm." Internet RFC 1321, April 1992.

David Aucsmith, Ed., *Second workshop on information hiding*, in vol. 1525 of *Lecture Notes in Computer Science*, Portland, Oregon, USA, 14–17 April, 1998, pp. 218–238.
ISBN 3-540-65386-4

Attacks on Copyright Marking Systems

Fabien A.P. Petitcolas*, Ross J. Anderson, and Markus G. Kuhn**

University of Cambridge, Computer Laboratory
Pembroke Street, Cambridge CB2 3QG, UK
{fapp2,rja14,mgk25}@cl.cam.ac.uk
<http://www.cl.cam.ac.uk/Research/Security/>

Abstract. In the last few years, a large number of schemes have been proposed for hiding copyright marks and other information in digital pictures, video, audio and other multimedia objects. We describe some contenders that have appeared in the research literature and in the field; we then present a number of attacks that enable the information hidden by them to be removed or otherwise rendered unusable.

1 Information Hiding Applications

The last few years have seen rapidly growing interest in ways to hide information in other information. A number of factors contributed to this. Fears that copyright would be eroded by the ease with which digital media could be copied led people to study ways of embedding hidden copyright marks and serial numbers in audio and video; concern that privacy would be eroded led to work on electronic cash, anonymous remailers, digital elections and techniques for making mobile computer users harder for third parties to trace; and there remain the traditional 'military' concerns about hiding one's own traffic while making it hard for the opponent to do likewise.

The first international workshop on information hiding [3] brought these communities together and a number of hiding schemes were presented there; more have been presented elsewhere. We formed the view that useful progress in steganography and copyright marking might come from trying to attack all these first generation schemes. In the related field of cryptology, progress was iterative: cryptographic algorithms were proposed, attacks on them were found, more algorithms were proposed, and so on. Eventually, theory emerged: fast correlation attacks on stream ciphers and differential and linear attacks on block ciphers, now help us understand the strength of cryptographic algorithms in much more detail than before. Similarly, many cryptographic protocols were proposed and almost all the early candidates were broken, leading to concepts of protocol robustness and techniques for formal verification [7].

* The first author is grateful to Intel Corporation for financial support under the grant 'Robustness of Information Hiding Systems'
** The third author is supported by a European Commission Marie-Curie grant

So in this paper, we first describe the copyright protection context in which most recent schemes have been developed; we then describe a selection of these schemes and present a number of attacks, which break most of them. We finally make some remarks on the meaning of robustness in the context of steganography in general and copyright marking in particular.

1.1 Copyright Protection Issues

Digital recording media offer many new possibilities but their uptake has been hindered by widespread fears among intellectual property owners such as Hollywood and the rock music industry that their livelihoods would be threatened if users could make unlimited perfect copies of videos, music and multimedia works.

One of the first copy protection mechanisms for digital media was the serial copy management system (SCMS) introduced by Sony and Phillips for digital audio tapes in the eighties [34]. The idea was to allow consumers to make a digital audio tape of a CD they owned in order to use it (say) in their car, but not to make a tape of somebody else's tape; thus copies would be limited to first generation only. The implementation was to include a Boolean marker in the header of each audio object. Unfortunately this failed because the hardware produced by some manufacturers did not enforce it.

More recently the Digital Video Disk, also known as Digital Versatile Disk (DVD) consortium called for proposals for a copyright marking scheme to enforce serial copy management. The idea is that the DVD players sold to consumers will allow unlimited copying of home videos and time-shifted viewing of TV programmes, but cannot easily be abused for commercial piracy [21, 46]. The proposed implementation is that videos will be unmarked, or marked 'never copy', or 'copy once only'; compliant players would not record a video marked 'never copy' and when recording one marked 'copy once only' would change its mark to 'never copy'. Commercially sold videos would be marked 'never copy', while TV broadcasts and similar material would be marked 'copy once only' and home videos would be unmarked.

Electronic copyright management schemes have also been proposed by European projects such as Imprimatur and CITED [47, 68, 69], and American projects such as the proposed by the Working Group on Intellectual Property Rights [71].

1.2 Problems

Although these schemes might become predominant in areas where they can be imposed from the beginning (such as DVD and video-on-demand), they suffer from a number of drawbacks. Firstly, they rely on the tamper-resistance of consumer electronics – a notoriously unsolved problem [5]. The tamper-resistance mechanisms being built into DVD players are fairly rudimentary and the history of satellite TV piracy leads us to expect the appearance of 'rogue' players which will copy everything. Electronic copyright management schemes also conflict with applications such as digital libraries, where 'fair use' provisions are

strongly entrenched. According to Samuelson, '*Tolerating some leakage may be in the long run of interest to publishers* [...] *For educational and research works, pay-per-use schemes may deter learning and deep scholarship*' [58]. A European legal expert put it even more strongly: that copyright laws are only tolerated because they are not enforced against the large numbers of petty offenders [35].

Similar issues are debated within the software industry; some people argue, for example, that a modest level of amateur software piracy actually enhances revenue because people may 'try out' software they have 'borrowed' from a friend and then go on to buy it (or the next update).

For all these reasons, we may expect leaks in the primary copyright protection mechanisms and wish to provide independent secondary mechanisms that can be used to trace and prove ownership of digital objects. It is here that marking techniques are expected to be most important.

2 Copyright Marks

There are two basic kinds of mark: fingerprints and watermarks. One may think of a fingerprint as an embedded serial number while a watermark is an embedded copyright message. The first enables us to trace offenders, while the second can provide some of the evidence needed to prosecute them. It may also, as in the DVD proposal, form part of the primary copy management system; but it will more often provide an independent back-up to a copy management system that uses overt mechanisms such as digital signatures.

In [8], we discussed the various applications of fingerprinting and watermarking, their interaction, and some related technologies. Here, we are concerned with the robustness of the underlying mechanisms. What sort of attacks are possible on marking schemes? What sort of resources are required to remove marks completely, or to alter them so that they are read incorrectly? What sort of effect do various possible removal techniques have on the perceptual quality of the resulting audio or video?

We will use the terminology agreed at the first international workshop on Information Hiding [54]. The information to be hidden (watermark, fingerprint, or in the general case of steganography, a secret message) is *embedded* in a *cover* object (a cover CD, a cover video, a cover text, etc.) giving a *stego* object, which in the context of copyright marking we may also call a *marked* object (CD, video, etc). The embedding is performed with the help of a *key*, a secret variable that is in general known to the object's owner. Recovery of the embedded mark may or may not require a key; if it does the key may be equal to, or derived from, the key used in the embedding process.

In the rest of this section, we will first discuss simple hiding methods and the obvious attacks on them. We will then present, as an example of the 'state of the art', robustness requirements that appeared in a recent music industry request for proposals [1]. We will then present the main contending techniques used in currently published and fielded systems. Attacks on these systems will then be presented.

2.1 Simple Hiding Methods

The simplest schemes replace all the bits in one or more of the less significant bit planes of an image or audio sample with the 'hidden' information [12, 26, 39, 67]. This is particularly easy with pictures: even when the four least significant bits of the cover image are replaced with the four most significant bits of the embedded image, the eye cannot usually tell the difference [39]. Audio is slightly harder, as the randomisation of even the least significant bit of 8-bit audio adds noise that is audible during quiet passages of music or pauses in speech. Nonetheless, several systems have been proposed: they include embedding, in the regular channels of an audio CD, another sound channel [27, 70] and a steganographic system in which secret messages are hidden in the digitised speech of an ISDN telephone conversation [26].

However, bit-plane replacement signals are not only easy to detect. They violate Kerckhoffs' principle that the security of a protection system should not rely on its method of operation being unknown to the opponent, but rather on the choice of a secret key [36]. Better approaches use a key to select some subset of pixels or sound samples which then carry the mark.

An example of this approach is Chameleon [6], a system which enables a broadcaster to send a single ciphertext to a large population of users, each of which is supplied with a slightly different decryption key; the effect of this is to introduce a controlled number of least-significant-bit errors into the plaintext that each user decrypts. With uncompressed digital audio, the resulting noise is at an acceptably low level and then Chameleon has the advantage that the decrypted audio is fingerprinted automatically during decryption without any requirement that the consumer electronic device be tamper-resistant.

In general, schemes which use a key to choose some subset of least significant bits to tweak may provide acceptable levels of security in applications where the decrypted objects are unlikely to be tampered with. However, in many applications, a copyright pirate may be able and willing to perform significant filtering operations and these will destroy any watermark, fingerprint or other message hidden by simple bit tweaking. So we shall now consider what it *means* for a marking scheme to be robust.

2.2 Robustness Requirements

The basic problem is to embed a mark in the digital representation of an analogue object (such as a film or sound recording) in such a way that it will not reduce the perceived value of the object while being difficult for an unauthorised person to remove. A first pass at defining robustness in this context may be found in a recent request for proposals for audio marking technology from the International Federation for the Phonographic Industry, IFPI [1]. The goal of this exercise was to find a marking scheme that would generate evidence for anti-piracy operations, track the use of recordings by broadcasters and others and control copying. The IFPI robustness requirements are as follows:

- the marking mechanism should not affect the sonic quality of the sound recording;
- the marking information should be recoverable after a wide range of filtering and processing operations, including two successive D/A and A/D conversions, steady-state compression or expansion of 10%, compression techniques such as MPEG and multi-band nonlinear amplitude compression, adding additive or multiplicative noise, adding a second embedded signal using the same system, frequency response distortion of up to 15 dB as applied by bass, mid and treble controls, group delay distortions and notch filters;
- there should be no other way to remove or alter the embedded information without sufficient degradation of the sound quality as to render it unusable;
- given a signal-to-noise level of 20 dB or more, the embedded data channel should have a bandwidth of 20 bits per second, independent of the signal level and type (classical, pop, speech).

Similar requirements could be drawn up for marking still pictures, videos and multimedia objects in general. However, before rushing to do this, we will consider some systems recently proposed and show attacks on them that will significantly extend the range of distortions against which designers will have to provide defences, or greatly reduce the available bandwidth, or both.

2.3 General Techniques

We mentioned schemes that modify the least significant bits of digital media; by repeating such marks, or employing more robust encoding methods, we can counter some filtering attacks. We can also combine coding with various transform techniques (DCT, wavelet and so on).

The *Patchwork* algorithm [11], for instance, successively selects random pairs of pixels; it makes the brighter pixel brighter and the duller pixel duller and the contrast change in this pixel subset encodes one bit. To maintain reasonable robustness against filtering attacks, the bandwidth of such systems has to be limited to at most a few hundred bits per image [40, 41]. In a similar way, marks can be embedded in audio by increasing the amplitude contrast of many pairs of randomly chosen sound samples and using a suitable filter to minimise the introduction of high-frequency noise.

More sophisticated variants on this theme involve spread-spectrum techniques. Although these have been used since the mid-fifties in the military domain because of their anti-jamming and low-probability-of-intercept properties [61], their applicability to image watermarking has only been noticed recently by Tirkel *et al.* [66]. Since then a number of systems based on this technique have been proposed [67, 72, 73]: typically a maximal length sequence is added to the signal in the spatial domain and the watermark is detected by using the spatial cross-correlation of the sequence and the watermarked image.

Another kind of marking technique embeds the mark in a transform domain, typically one that is widely used by compression algorithms. Thus when marking sound one could add a pseudorandom sequence to the excitation signal in

an LPC or CELP coded audio signal [45] and when marking an image one could use the DCT domain. Langelaar *et al.* remove certain high frequency DCT coefficients [41]; Cox *et al.* modulate the 1000 largest DCT coefficients of an image with a random vector [19]; Koch *et al.* change the quantisation of the DCT coefficients and modify some of them in such a way that a certain property (order in size) is verified [37]; while Ó Ruanaidh *et al.* modulate the DCT coefficient with a bi-directional coding [49].

Techniques of this kind are fairly robust against various kinds of signal processing and may be combined with exploitation of the perceptual masking properties of the human auditory system in [16, 17] and of the human vision system in [28, 65, 64]. The basic idea here is to amplify the mark wherever the changes will be less noticeable and also to embed it in the *perceptually significant* components of the signal [20]. Masking may also be used to avoid placing marks in places such as the large expanses of pure colour found in cartoons; the colour histogram of such images has sharp peaks, which are split into twin peaks by some naïve marking methods as the colour value c is replaced by $c-\delta$ and $c+\delta$, thus allowing the mark to be identified and removed [44].

3 Attacks

This leads us to the topic of attacks and here we present some quite general kinds of attack that destroy, or at least reveal significant limitations of, several marking schemes: PictureMarc 1.51 [24, 56], SysCoP [37, 74, 75], JK_PGS (EPFL algorithm, part of the European TALISMAN project), SureSign [63], EIKONAmark [25, 55], Echo Hiding, and the NEC method [19]. We suspect that systems that use similar techniques are also vulnerable to our attacks.

3.1 The Jitter Attack

Our starting point in developing a systematic attack on marking technology was to consider audio marking schemes that tweak low order bits whose location is specified by a key. A simple and devastating attack on these schemes is to add jitter to the signal. In our first implementation, we split the signal into chunks of 500 samples, either duplicated or deleted a sample at random in each chunk (resulting in chunks of 499 or 501 samples long) and stuck the chunks back together. This turned out to be almost imperceptible after filtering, even in classical music; but the jitter prevents the marked bits from being located.

In a more sophisticated implementation, we resample these chunks at a lower or higher frequency. This relies on the properties of the ear's pitch resolution:

> In pitch perception experiments in the mid-audio frequency range, subjects are able to perceive changes in frequency of pure tones of approximately 0.1%. [...] At frequencies above 4 kHz pitch discrimination reduces substantially. [...] In the case of complex signals, such as speech, it is very much less clear what the capabilities and processes of the auditory system are. [...] There is evidence that peaks in the spectrum of

the audio signal are detected more easily than features between spectral peaks. *J.N. Holmes* [33]

If n_i is the number of samples in the ith chunk, n_i' the number of samples after resampling and α the maximum relative change of frequency allowed then, in the mid-audio range, we are roughly limited, for pure tones, by $|\Delta n_i| \leq \alpha n_i$ (because α is small), where $\Delta n_i := n_{i+1}' - n_i'$. This can be simplified as $0 < k \leq \frac{\alpha n}{2}$ when the n_i are equal and when the number k of removed or added samples is constant for each chunk. This is the approach we chose; it allowed us to introduce a long jitter. Then the strategy for choosing k and n depends on the input signal. With this technique we were able to tweak up to one sample in 50 of a 44 kHz sampled voice recording without any perceptible effect.

We also applied a similar attack to SysCoP Demo 1.0. In that case we simply deleted columns of pixels and duplicated others in order to preserve the image size. Fig. 1 gives an example of this attack.

Of course, there are much more subtle distortions that can be applied. For instance, in [30], Hamdy *et al.* present a way to increase or decrease the length of a music performance without changing the pitch; this was developed to enable radio broadcasters to slightly increase or decrease the playing time of a musical track. As such tools become widely available, attacks involving sound manipulation will become easy. Most simple spread-spectrum based techniques are subject to this kind of attacks. Indeed, although spread-spectrum signal are very robust to distortion of their amplitude and to noise addition, they do not survive timing errors: synchronisation of the chip signal is very important and simple systems fail to recover this synchronisation properly.

3.2 StirMark

Following this attack and after evaluating some watermarking software, it became clear that although many of the seriously proposed schemes could survive basic manipulations – that is, manipulations that can be done easily with standard tools, such as rotation, shearing, resampling, resizing and lossy compression – they would not cope with combinations of them. This motivated us to implement StirMark.

StirMark is a generic tool developed for simple robustness testing of image marking algorithms and other steganographic techniques. In its simplest version, StirMark simulates a resampling process, i.e. it introduces the same kind of errors into an image as printing it on a high quality printer and then scanning it again with a high quality scanner. It applies a minor geometric distortion: the image is slightly stretched, sheared, shifted and/or rotated by an unnoticeable random amount[1] (Fig. 2 – middle drawing) and then resampled using either bi-linear or

[1] If A, B, C and D are the corners of the image, a point M of the said image can be expressed as $M = \alpha[\beta A + (1-\beta)D] + (1-\alpha)[\beta B + (1-\beta)C]$ where $0 \leq \alpha, \beta \leq 1$ are the coordinates of M relatively to the corners. The distortion is done by moving the corners by a small random amount in both directions. The new coordinates of M are given by the previous formula, keeping (α, β) constant.

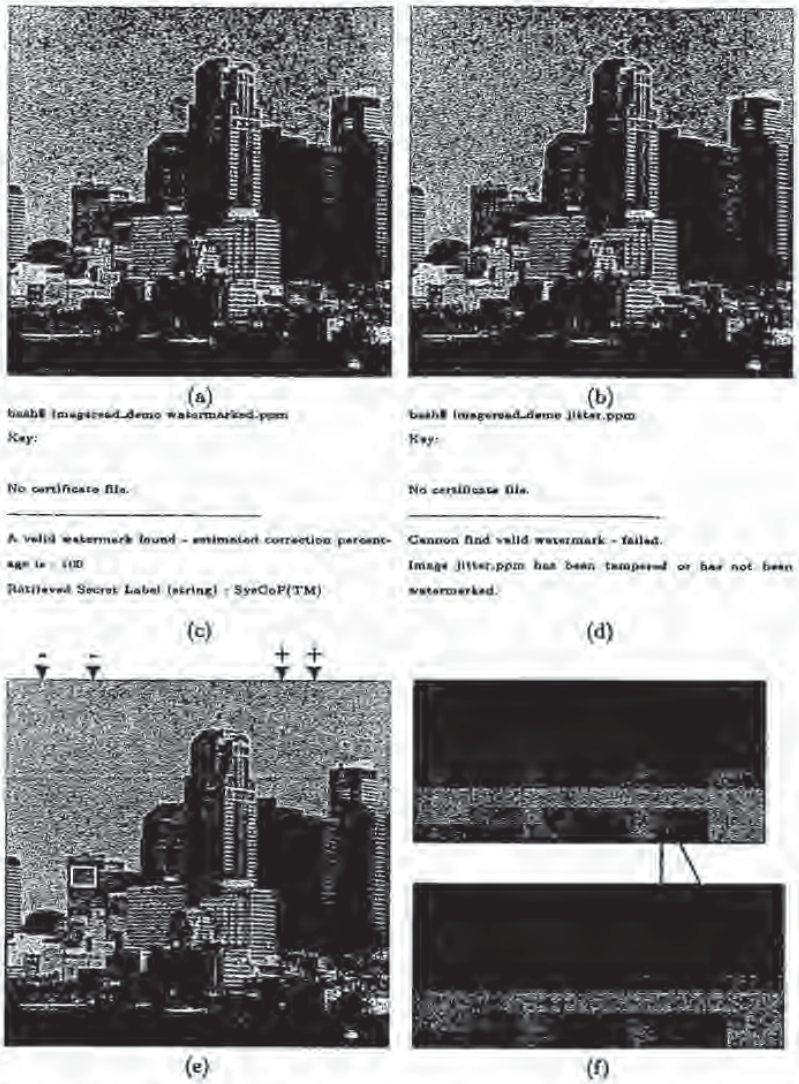


Fig. 1. A successful jitter attack on SysCoP. We used the demo software release 1.0 available on SysCoP's Web site [76]. (a) shows an image watermarked with SysCoP and (b) the same image but after the attack. In the first case the software detects the watermark correctly (c) but the check fails on the modified image (d). Here, the attack simply consists in deleting and duplicating some columns of pixels such that the original size of the picture is conserved. (e) shows the columns which have been deleted (-) and duplicated (+). Finally, (f) is a magnified view of the white rectangle in (e); the bottom part corresponds to the original image.

Nyquist interpolation. In addition, a transfer function that introduces a small and smoothly distributed error into all sample values is applied. This emulates the small non-linear analog/digital converter imperfection typically found in scanners and display devices. StirMark introduces a practically unnoticeable quality loss in the image if it is applied only once. However after a few iterated applications, the image degradation becomes noticeable.

With those simple geometrical distortions we could confuse most marking systems available on the market. More distortions – still unnoticeable – can be applied to a picture. We applied a global 'bending' to the image: in addition to the general bi-linear property explained previously a slight deviation is applied to each pixel, which is greatest at the center of the picture and almost null at the borders. On top of this a higher frequency displacement of the form $\lambda \sin(\omega_x x) \sin(\omega_y y) + n(x, y)$ – where n is a random number – is added. In order for these distortions to be most effective, a medium JPEG compression is applied at the end.

Fig. 2. We exaggerate here the distortion applied by StirMark to still pictures. The first drawing corresponds to the original picture; the others show the picture after StirMark has been applied – without and with bending and randomisation.

For those unfamiliar with digital image signal processing we shall now summarise briefly the main computation steps. Apart from a few simple operations such as rotations by 90 or 180 degrees, reflection and mirroring, image manipulation usually requires resampling when destination pixels do not line up with source pixels. In theory, one first generates a continuous image from the digital one, then modifies the continuous image, finally samples this to create a new digital image. In practice, however, we compute the inverse transform of a new pixel and evaluate the reconstruction function at that point.

There are numerous reconstruction filters. In a first version of the software we simply used a linear interpolation but, as foreseen, this tended to blur the image too much, making the validity of the watermark removal arguable. Then we implemented the sinc function as a reconstruction filter, which gives theoretically perfect reconstruction for photo images and can be described as follows. If (x, y) are the coordinates of the inverse transform – which, in our case is a distortion of the picture – of a point in the new image and f the function to be reconstructed,

then, an estimate of f at (x, y) is given by $\tilde{f}(x, y) = \sum_{i=-n}^{n} \sum_{j=-n}^{n} \mathrm{sinc}(x - i)\mathrm{sinc}(y - j) f_{i,j}$. This gives very much better results than the simple filter; an example of the removal of an NEC watermark is given in Fig. 3.

We suggest that image watermarking tools which do not survive StirMark – with default parameters – should be considered unacceptably easy to break. This immediately rules out the majority of commercial marking schemes.

(a) (b)

Fig. 3. Kings' College Chapel, courtesy of John Thompson, JetPhotographic, Cambridge. For this example we watermarked a picture with NEC's algorithm [19]. We used the default parameters suggested by their paper ($N = 1000$ and $\alpha = 0.1$). (a) is the watermarked image. We then applied StirMark (b) and tested the presence of the watermark. The similarity between the original watermark and the extracted watermark was 3.74 instead of 21.08. This is well below the decision threshold.

One might try to increase the robustness of a watermarking system by trying to foresee the possible transforms used by pirates; one might then use techniques such as embedding multiple versions of the mark under suitable inverse transforms; for instance Ó Ruanaidh and Pereira suggest to use the Fourier-Mellin transform[2] to cope with rotation and scaling [50]. However, the general theme of the attacks we have developed and described above is that given a target marking scheme, we invent a distortion (or a combination of distortions) that will remove it or at least make it unreadable, while leaving the perceptual value of the previously marked object undiminished. We are not limited in this process to the distortions produced by common analogue equipment, or considered in the IFPI request for proposals cited above.

[2] The Fourier-Mellin transform is equivalent to the Fourier transform on a log-polar map: $(x, y) \to (\mu, \theta)$ with $x = e^{\mu} \cos\theta$ and $y = e^{\mu} \sin\theta$.

As an analogy, one might consider the 'chosen protocol attack' on authentication schemes [60]. It is an open question whether there is any marking scheme for which a chosen distortion attack cannot be found.

3.3 The Mosaic Attack

This point is emphasised by a 'presentation' attack, which is of quite general applicability and which possesses the initially remarkable property that a marked image can be unmarked and yet still rendered pixel for pixel in exactly the same way as the marked image by a standard browser.

The attack was motivated by a fielded automatic system for copyright piracy detection, consisting of a watermarking scheme plus a web crawler that downloads pictures from the net and checks whether they contain a watermark.

It consists of chopping an image up into a number of smaller subimages, which are embedded in a suitable sequence in a web page. Common web browsers render juxtaposed subimages stuck together, so they appear identical to the original image (Fig. 4). This attack appears to be quite general; all marking schemes require the marked image to have some minimal size (one cannot hide a meaningful mark in just one pixel). Thus by splitting an image into sufficiently small pieces, the mark detector will be confused [53]. The best that one can hope for is that the minimal size could be quite small and the method might therefore not be very practical.

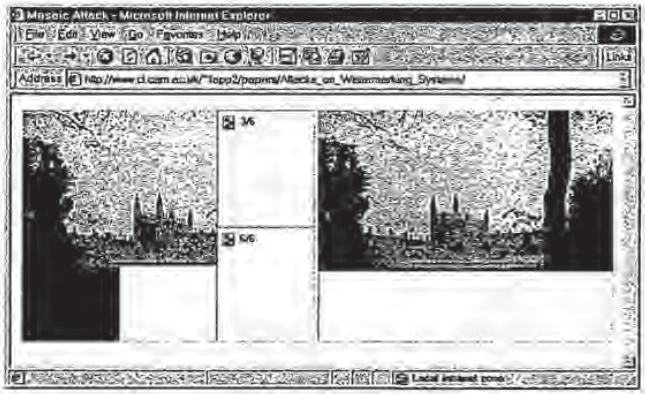


Fig. 4. Screen-shot of a web browser while downloading an image after the *mosaic attack*. This attack chops a watermarked image into smaller images which are stuck back together when the browser renders the page. We implemented software that reads a JPEG picture and produces a corresponding mosaic of small JPEG images as well as the necessary HTML code automatically [53]. In some cases downloading the mosaic is even faster than downloading the full image! In this example we used a 350×280-pixel image watermarked using PictureMarc 1.51.

There are other problems with such 'crawlers'. Java applets, ActiveX controls, etc. can be embedded to display a picture inside the browser; the applet could even de-scramble the picture in real time. Defeating such techniques would entail rendering the web page, detecting pictures and checking whether they contain a mark. An even more serious problem is that much current piracy is of pictures sold via many small services, from which the crawler would have to purchase them using a credit card before it could examine them. A crawler that provided such 'guaranteed sales' would obviously become a target.

3.4 Attack on *Echo Hiding*

One of the few marking schemes to be robust against the jitter attack is echo hiding, which hides information in sound by introducing echoes with very short delays. *Echo hiding* [29] relies on the fact that we cannot perceive short echoes (say 1 ms) and embeds data into a cover audio signal by introducing an echo characterised by its delay τ and its relative amplitude α. By using two types of echo it is possible to encode ones and zeros. For this purpose the original signal is divided into chunks separated by spaces of pseudo-random length; each of these chunks will contain one bit of information.

The echo delays are chosen between 0.5 and 2 milliseconds and the best relative amplitude of the echo is around 0.8. According to its creators, decoding involves detecting the initial delay and the auto-correlation of the cepstrum of the encoded signal is used for this purpose.

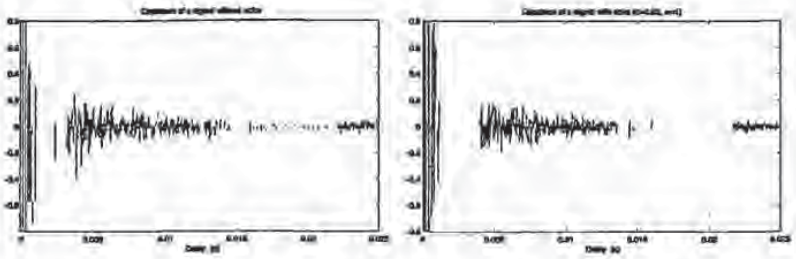
The 'obvious' attack on this scheme is to detect the echo and then remove it by simply inverting the convolution formula; the problem is to detect the echo without knowledge of either the original object or the echo parameters. This is known as 'blind echo cancellation' in the signal processing literature and is known to be a hard problem in general.

We tried several methods to remove the echo. Frequency invariant filtering [51, 59] was not very successful. Instead we used a combination of cepstrum analysis and 'brute force' search.

The underlying idea of cepstrum analysis is presented in [15]. Suppose that we are given a signal $y(t)$ which contains a simple single echo, i.e. $y(t) = x(t) + \alpha x(t-\tau)$. If we note Φ_{xx} the power spectrum of x then $\Phi_{yy}(f) = \Phi_{xx}(f)[1 + 2\alpha\cos(2\pi f\tau) + \alpha^2]$ whose logarithm is approximately $\log\Phi_{yy}(f) \approx \log\Phi_{xx}(f) + 2\alpha\cos(2\pi f\tau)$. This is a function of the frequency f and taking its power spectrum raises its 'quefrency' τ, that is the frequency of $\cos(2\pi\tau f)$. The auto-covariance of this later function emphasises the peak that appears at 'quefrency' τ (Fig. 5).

To remove the echos, we need a method to detect the echo delay τ. For this, we used a slightly modified version of the cepstrum: $\mathrm{C}\circ\Phi\circ\ln\circ\Phi$ where C is the auto-covariance function[3], Φ the power spectrum density function and $\circ$ the composition operator. Experiments on random signals as well as on music show that this method returns quite accurate estimators of the delay (Fig. 6) when an artificial echo has been added to the signal. In the detection function we only

[3] $\mathrm{C}(x) = \mathrm{E}[(x-\overline{x})(x-\overline{x})^*]$.



(a) (b)

Fig. 5. Graph (a) represents the cepstrum of a signal without echo. Graph (b) is the cepstrum of the same signal with a 20 ms echo which is emphasised by the very clear peak at 0.02 s.

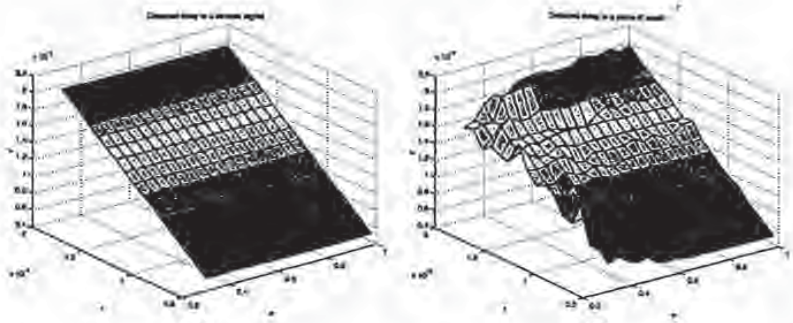
consider echo delays between 0.5 and 3 milliseconds. Below 0.5 ms the function does not work properly and above 3 ms the echo becomes too audible.

Our first attack was to remove an echo with random relative amplitude, expecting that this would introduce enough modification in the signal to prevent watermark recovery. Since echo hiding gives best results for α greater than 0.7 we could use $\tilde{\alpha}$ – an estimation of α – drawn from, say a normal distribution centred on 0.8. It was not really successful so our next attack was to iterate: we re-apply the detection function and vary $\tilde{\alpha}$ to minimise the residual echo. We could obtain successively better estimators of the echo parameters and then remove this echo. When the detection function cannot detect any more echo, we have got the correct value of $\tilde{\alpha}$ (as this gives the lowest output value of the detection function). Results obtained using this algorithm are presented in Fig. 6.

3.5 Protocol Considerations

The main threat addressed in the literature is an attack by a pirate who tries to remove the watermark directly. As a consequence, the definition commonly used for robustness includes only resistance to signal manipulation (cropping, scaling, resampling, etc.). Craver *et al.* show that this is not enough by exhibiting a 'protocol' level attack [22].

The basic idea is that many schemes provide no intrinsic way of detecting which of two watermarks was added first: the process of marking is often additive, or at least commutative. So if the owner of the document d encodes a watermark w and publishes the marked version $d + w$ and has no other proof of ownership, a pirate who has registered his watermark as w' can claim that the document is his and that the original unmarked version of it was $d + w - w'$. Their paper ([23]) extends this idea to defeat a scheme which is non-invertible (an inverse needs only be approximated).



(a) (b)

Fig. 6. Performances of the echo detector. We added different echoes characterised by their relative amplitude α and their delay τ to a signal and each time we used our echo detector to find an estimation $\hat{\tau}$ of τ. These graphs show the detected echo delay as a function of α and τ for random signals (a) and for a piece of music (b).

Craver *et al.* argue for the use of information-losing marking schemes whose inverses cannot be approximated closely enough. However, our alternative interpretation of their attack is that watermarking and fingerprinting methods must be used in the context of a larger system that may use mechanisms such as timestamping and notarisation to prevent attacks of this kind.

Registration mechanisms have not received very much attention in the copyright marking literature to date. The existing references such as [18, 32, 31, 52] mainly focus on protecting the copyright holder and do not fully address the rights of the consumers who might be fooled by a crooked reseller.

3.6 Implementation Considerations

The robustness of embedding and retrieving techniques is not the only issue. Most attacks on fielded cryptographic systems have come from the opportunistic exploitation of loopholes that were found by accident; cryptanalysis was rarely used, even against systems that were vulnerable to it [2].

We cannot expect copyright marking systems to be any different and the pattern was followed in the first attack to be made available on the Internet against the most widely used picture marking scheme, PictureMarc, which is bundled with Adobe Photoshop and Corel Draw. This attack [13] exploited weaknesses in the implementation rather than the underlying marking algorithms, even although these are weak (the marks can be removed using StirMark).

Each user has an ID and a two-digit password, which are issued when she registers with Digimarc and pays for a subscription. The correspondence between IDs and passwords is checked using obscure software in the implementation and although the passwords are short enough to be found by trial and error, the

attack first uses a debugger to break into the software and disable the password checking mechanism.

We note in passing that IDs are public, so either password search or disassembly can enable any user to be impersonated.

A deeper examination of the program also allows a villain to change the ID, thus the copyright, of an already marked image as well as the type of use (such as adult versus general public content). Before embedding a mark, the program checks whether there is already a mark in the picture, but this check can be bypassed fairly easily using the debugger with the result that it is possible to overwrite any existing mark and replace it with another one.

Exhaustive search for the personal code can be prevented by making it longer, but there is no obvious solution to the disassembly attack. If tamper resistant software [9] cannot give enough protection, then one can always have an online system in which each user shares a secret embedding key with a trusted party and uses this key to embed some kind of digital signature. Observe that there are two separate keyed operations here; the authentication (which can be done with a signature) and the embedding or hiding operation.

Although we can do public-key steganography – hiding information so that only someone with a certain private key can detect its existence [4] – we still do not know how to do the hiding equivalent of a digital signature; that is, to enable someone with a private key to embed marks in such a way that anyone with the corresponding public key can read them but not remove them. One problem is that a public decoder can be used by the attacker; he can remove a mark by applying small changes to the image until the decoder cannot find it anymore. This was first suggested by Perrig in [52]. In [42] a more theoretical analysis of this attack is presented as well as a possible countermeasure: randomisating the detection process. One could also make the decoding process computationally expensive. However neither approach is really satisfactory in the absence of tamper-resistant hardware.

Unless a breakthrough is made, applications that require the public verifiability of a mark (such as DVD) appear doomed to operate within the constraints of the available tamper resistance technology, or to use a central 'mark reading' service. This is evocative of cryptographic key management prior to the invention of public key techniques.

4 Conclusion

We have demonstrated that the majority of copyright marking schemes in the literature are vulnerable to attacks involving the introduction of sub-perceptual levels of distortion. In particular, many of the marking schemes in the marketplace provide only a limited measure of protection against attacks. Most of them are defeated by StirMark, a simple piece of software that we have placed in the public domain [38]. We have also shown a specific attack on the one serious exception to this rule (echo hiding).

This experience confirms our hypothesis that steganography would go through the same process of evolutionary development as cryptography, with an iterative process in which attacks lead to more robust systems.

Our experience in attacking the existing marking schemes has convinced us that any system which attempted to meet all the accepted requirements for marking (such as those set out by IFPI) would fail: if it met the robustness requirements then its bandwidth would be quite insufficient. This is hardly surprising when one considers that the information content of many music recording is only a few bits per second, so to expect to embed 20 bits per second against an opponent who can introduce arbitrary distortions is very ambitious.

Our more general conclusion from this work is that the 'marking problem' has been over-abstracted; there is not one 'marking problem' but a whole constellation of them. We do not believe that any general solution will be found. The trade-offs and in particular the critical one between bandwidth and robustness, will be critical to designing a specific system.

We already remarked in [8] on the importance of whether the warden was active or passive – that is, whether the mark needed to be robust against distortion. In general, we observe that most real applications do not require all of the properties in the IFPI list. For example, when auditing radio transmissions, we only require enough resistance to distortion to deal with naturally occurring effects such as multipath. Many applications will also require supporting protocol features, such as the timestamping service that we mentioned in the context of reversible marks.

So we do not believe that the intractability of the 'marking problem' is a reason to abandon this field of research. On the contrary; practical schemes for most realistic application requirements are probably feasible and the continuing process of inventing schemes and breaking them will enable us to advance the state of the art rapidly.

Finally, we suggest that the real problem is not so much inserting the marks as recognising them afterwards. Thus progress may come not just from devising new marking schemes, but in developing ways to recognise marks that have been embedded using the obvious combinations of statistical and transform techniques and thereafter subjected to distortion. The considerable literature on signal recognition may provide useful starting points.

Acknowledgements

Some of the ideas presented here were clarified by discussion with Roger Needham, David Wheeler, John Daugman, Peter Rayner, David Aucsmith, Stewart Lee, Scott Craver, Brian Moore, Mike Roe, Peter Wayner, Jon Honeyball, Scott Moskowitz and Matt Blaze.

References

1. Request for proposals – Embedded signalling systems issue 1.0. International Federation of the Phonographic Industry, 54 Regent Street, London W1R 5PJ, June 1997.
2. Ross J. Anderson. Why cryptosystems fail. *Communications of the ACM*, 37(11):32–40, November 1994.
3. Ross J. Anderson, editor. *Information hiding: first international workshop*, volume 1174 of *Lecture Notes in Computer Science*, Isaac Newton Institute, Cambridge, England, May 1996. Springer-Verlag, Berlin, Germany.
4. Ross J. Anderson. Stretching the limits of steganography. In IH96 [3], pages 39–48.
5. Ross J. Anderson and Markus G. Kuhn. Tamper resistance – A cautionary note. In *Second USENIX Workshop on Electronic Commerce*, pages 1–11, Oakland, CA, USA, November 1996.
6. Ross J. Anderson and Charalampos Manifavas. Chameleon – a new kind of stream cipher. In Biham [14], pages 107–113.
7. Ross J. Anderson and Roger M. Needham. Programming satan's computer. In J.van Leeuwen, editor, *Computer Science Today – Commemorative Issue*, volume 1000 of *Lecture Notes in Computer Science*, pages 426–441. Springer-Verlag, Berlin, Germany, 1995.
8. Ross J. Anderson and Fabien A. P. Petitcolas. On the limits of steganography. *IEEE Journal of Selected Areas in Communications*, 16(4):474–481, May 1998. Special Issue on Copyright & Privacy Protection.
9. David Aucsmith. Tamper resistant software: An implementation. In Anderson [3], pages 317–333.
10. David Aucsmith, editor. *Information Hiding: Second International Workshop*, volume 1525 of *Lecture Notes in Computer Science*, Portland, Oregon, USA, 1998. Springer-Verlag, Berlin, Germany.
11. Walter Bender, Daniel Gruhl, and Norishige Morimoto. Techniques for data hiding. In Niblack and Jain [48], pages 164–173.
12. Walter Bender, Daniel Gruhl, Norishige Morimoto, and Anthony Lu. Techniques for data hiding. *IBM Systems Journal*, 35(3 & 4):313–336, 1996.
13. Anonymous (<zguan.bbs@bbs.ntu.edu.tw>). Learn cracking IV – another weakness of PictureMarc. <news:tw.bbs.comp.hacker> mirrored on <http://www.cl.cam.ac.uk/~fapp2/watermarking/image_watermarking/digimarc_crack.html>, August 1997. Includes instructions to override any Digimarc watermark using PictureMarc.
14. Eli Biham, editor. *Fast Software Encryption – 4th International Workshop, FSE'97*, volume 1267 of *Lecture Notes in Computer Science*, Haifa, Israel, January 1997. Springer-Verlag, Germany.
15. Bruce P. Bogert, M.J.R. Healy, and John W. TEnglandey. The quefrency alanysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-ceptstrum and saphe cracking. In M. Rosenblatt, editor, *Symposium on Time Series Analysis*, pages 209–243, New York, NY, USA, 1963. John Wiley & Sons, Inc.
16. Laurence Boney, Ahmed H. Tewfik, and Khaled N. Hamdy. Digital watermarks for audio signals. In *European Signal Processing Conference, EUSIPCO '96*, Trieste, Italy, September 1996.
17. Laurence Boney, Ahmed H. Tewfik, and Khaled N. Hamdy. Digital watermarks for audio signals. In *International Conference on Multimedia Computing and Systems*, pages 473–480, Hiroshima, Japan, 17–23 June 1996. IEEE.

18. Marc Cooperman and Scott A. Moskowitz. Steganographic method and device. US Patent 5,613,004, March 1995.
19. Ingemar J. Cox, Joe Kilian, Tom Leighton, and Talal Shamoon. A secure, robust watermark for multimedia. In Anderson [3], pages 183–206.
20. Ingemar J. Cox and Matt L. Miller. A review of watermarking and the importance of perceptual modeling. In Rogowitz and Pappas [57].
21. Ingemar J. Cox and Kazuyoshi Tanaka. NEC data hiding proposal. Technical report, NEC Copy Protection Technical Working Group, July 1997. Response to call for proposal issued by the Data Hiding SubGroup.
22. Scott Craver, Nasir Memon, Boon-Lock Yeo, and Minerva M. Yeung. Can invisible watermark resolve rightful ownerships? In Sethin and Jain [62], pages 310–321.
23. Scott Craver, Nasir Memon, Boon-Lock Yeo, and Minerva M. Yeung. Resolving rightful ownerships with invisible watermarking techniques: Limitations, attacks, and implications. *IEEE Journal of Selected Areas in Communications*, 16(4):573–586, May 1998. Special Issue on Copyright & Privacy Protection. ISSN 0733-8716.
24. Digimarc home page. <http://www.digimarc.com/>, April 1997.
25. Eikonamark. Alpha Tec Ltd., <http://www.generation.net/~pitas/sign.html>, October 1997.
26. Elke Franz, Anja Jerichow, Steffen Möller, Andreas Pfitzmann, and Ingo Stierand. Computer based steganography: how it works and why therefore any restriction on cryptography are nonsense, at best. In Anderson [3], pages 7–21.
27. Michael A. Gerzon and Peter G. Graven. A high-rate buried-data channel for audio CD. *Journal of the Audio Engineering Society*, 43(1/2):3–22, January–February 1995.
28. François Goffin, Jean-François Delaigle, Christophe De Vleeschouwer, Benoît Macq, and Jean-Jacques Quisquater. A low cost perceptive digital picture watermarking method. In Sethin and Jain [62], pages 264–277.
29. Daniel Gruhl, Walter Bender, and Anthony Lu. Echo hiding. In Anderson [3], pages 295–315.
30. Khaled N. Hamdy, Ahmed H. Tewfik, Ting Chen, and Satoshi Takagi. Time-scale modification of audio signals with combined harmonic and wavelet representations. In *International Conference on Acoustics, Speech and Signal Processing – ICASSP '97*, volume 1, pages 439–442, Munich, Germany, April 1997. IEEE, IEEE Press. Session on Hearing Aids and Computer Music.
31. Alexander Herrigel, Joseph J. K. Ó Ruanaidh, Holger Petersen, Shelby Pereira, and Thierry Pun. Secure copyright protection techniques for digital images. In Aucsmith [10], pages 169–190.
32. Alexander Herrigel, Adrian Perrig, and Joseph J. K. Ó Ruanaidh. A copyright protection environment for digital images. In *Verläßliche IT-Systeme '97*, Albert-Ludwigs Universität, Freiburg, Germany, October 1997.
33. J.N. Holmes. *Speech Synthesis and Recognition*, chapter 3.6 Analysis of simple and complex signals, pages 47–48. Aspects of Information Technology. Chapman & Hall, London, England, 1988.
34. International Electrotechnical Commission, Geneva, Switzerland. *Digital audio interface, IEC 60958*, February 1989.
35. Alastair Kelman. Electronic copyright management – the way ahead. Security Seminars, University of Cambridge, February 1997.
36. A. Kerckhoffs. La Cryptographie Militaire. *Journal des Sciences Militaires*, 9:5–38, January 1883.

37. E. Koch and J. Zhao. Towards robust and hidden image copyright labeling. In *Workshop on Nonlinear Signal and Image Processing*, pages 452–455, Neos Marmaras, Greece, June 1995. IEEE.
38. Markus G. Kuhn and Fabien A. P. Petitcolas. StirMark. <http://www.cl.cam.ac.uk/~fapp2/watermarking/stirmark/>, November 1997.
39. Charles Kurak and John McHugh. A cautionary note on image downgrading. In *Computer Security Applications Conference*, pages 153–159, San Antonio, TX, USA, December 1992.
40. Gerrit C. Langelaar, Jan C.A van der Lubbe, and J. Biemond. Copy protection for multimedia data based on labeling techniques. In *17th Symposium on Information Theory in the Benelux*, Enschede, The Netherlands, May 1996.
41. Gerrit C. Langelaar, Jan C.A. van der Lubbe, and Reginald L. Lagendijk. Robust labeling methods for copy protection of images. In Sethin and Jain [62], pages 298–309.
42. Jean-Paul M.G. Linnartz and Marten van Dijk. Analysis of the sensitivity attack against electronic watermarks in images. In Aucsmith [10], pages 258–272.
43. Mark Lomas, Bruno Crispo, Bruce Christianson, and Mike Roe, editors. *Security Protocols: Proceeding of the 5th International Workshop*, volume 1361 of *Lecture Notes in Computer Science*, École Normale Supérieure, Paris, France, April 1997. University of Cambridge, Isaac Newton Institute, Springer-Verlag, Berlin, Germany.
44. Maurice Maes. Twin peaks: The histogram attack on fixed depth image watermarks. In Aucsmith [10], pages 290–305.
45. Kineo Matsui and Kiyoshi Tanaka. Video-steganography: How to secretly embed a signature in a picture. *Journal of the Interactive Multimedia Association Intellectual Property Project*, 1(1):187–205, January 1994.
46. Norishige Morimoto and Daniel Sullivan. IBM DataHiding proposal. Technical report, IBM Corporation, September 1997. Response to call for proposal issued by the Data Hiding SubGroup.
47. Peter Nancarrow. Digital technology – Bane or boon for copyright? Computer Laboratory Seminars, University of Cambridge, November 1997.
48. Wayne Niblack and Ramesh C. Jain, editors. *Storage and Retrieval for Image and Video Database III*, volume 2420, San Jose, California, USA, February 1995. IS&T, The Society for Imaging Science and Technology and SPIE, The International Society for Optical Engineering, SPIE.
49. Joseph J. K. Ó Ruanaidh, W. J. Dowling, and F. M. Boland. Watermarking digital images for copyright protection. *IEE Proceedings on Vision, Signal and Image Processing*, 143(4):250–256, August 1996.
50. Joseph J. K. Ó Ruanaidh and Shelby Pereira. A secure robust digital image watermark. In *International Symposium on Advanced Imaging and Network Technologies – Conference on Electronic Imaging: Processing, Printing and Publishing in Colour*, Europto, Zürich, Switzerland, May 1998. International Society for Optical Engineering, European Optical Society, Commission of the European Union, Directorate General XII.
51. Alan V. Oppenheim and Ronald W. Schafer. *Discrete-Time Signal Processing*, chapter 12, pages 768–834. Prentice-Hall International, Inc., Englewood Cliffs, NJ, USA, international edition, 1989.
52. Adrian Perrig. A copyright protection environment for digital images. Diploma dissertation, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, February 1997.

53. Fabien A. P. Petitcolas. Weakness of existing watermarking schemes. <http://www.cl.cam.ac.uk/~fapp2/watermarking/image_watermarking/>, October 1997.
54. Birgit Pfitzmann. Information hiding terminology. In Anderson [3], pages 347–350. Results of an informal plenary meeting and additional proposals.
55. I. Pitas. A method for signature casting on digital images. In *International Conference on Image Processing*, volume 3, pages 215–218, September 1996.
56. Geoffrey B. Rhoads. Steganography methods employing embedded calibration data. US Patent 5,636,292, June 1997.
57. Bernice E. Rogowitz and Thrasyvoulos N. Pappas, editors. *Human Vision and Electronic Imaging II*, volume 3016, San Jose, CA, USA, February 1997. IS&T, The Society for Imaging Science and Technology and SPIE, The International Society for Optical Engineering, SPIE.
58. Pamela Samuelson. Copyright and digital libraries. *Communications of the ACM*, 38(4):15–21, 110, April 1995.
59. Ronald W. Schafer. Echo removal by discrete generalized linear filtering. Technical Report 466, Massachusetts Institute of Technology, February 1969.
60. Bruce Schneier. Protocol interactions and the chosen protocol attack. In Lomas et al. [43], pages 91–104.
61. Robert A. Scholtz. The origins of spread-spectrum communications. *IEEE Transactions on Communications*, 30(5):822–853, May 1982.
62. Ishwar K. Sethin and Ramesh C. Jain, editors. *Storage and Retrieval for Image and Video Database V*, volume 3022, San Jose, CA, USA, February 1997. IS&T, The Society for Imaging Science and Technology and SPIE, The International Society for Optical Engineering, SPIE.
63. Signum Technologies – SureSign digital fingerprinting. <http://www.signumtech.com/>, October 1997.
64. Mitchell D. Swanson, Bin Zhu, and Ahmed H. Tewfik. Transparent robust image watermarking. In *International Conference on Image Processing*, volume III, pages 211–214. IEEE, 1996.
65. Mitchell D. Swanson, Bin Zu, and Ahmed H. Tewfik. Robust data hiding for images. In *7th Digital Signal Processing Workshop (DSP 96)*, pages 37–40, Loen, Norway, September 1996. IEEE.
66. A.Z. Tirkel, G.A. Rankin, R.M. van Schyndel, W.J. Ho, N.R.A. Mee, and C.F. Osborne. Electronic watermark. In *Digital Image Computing, Technology and Applications – DICTA '93*, pages 666–673, Macquarie University, Sidney, 1993.
67. R.G. van Schyndel, A.Z. Tirkel, and C.F. Osborne. A digital watermark. In *International Conference on Image Processing*, volume 2, pages 86–90, Austin, Texas, USA, 1994. IEEE.
68. Georges Van Slype. Natural language version of the generic CITED model – ECMS (Electronic Copyright Management System) design for computer based applications. Report 2, European Commission, ESPRIT II Project, Bureau Vam Dijk, Brussel, Belgium, May 1995.
69. Georges Van Slype. Natural language version of the generic CITED model – Presentation of the generic model. Report 1, European Commission, ESPRIT II Project, Bureau Vam Dijk, Brussel, Belgium, May 1995.
70. A. Werner, J. Oomen, Marc E. Groenewegen, Robbert G. van der Waal, and Raymond N.J. Veldhuis. A variable-bit-rate buried-data channel for compact disc. *Journal of the Audio Engineering Society*, 43(1/2):23–28, January–February 1995.
71. The Working Group on Intellectual Property Rights is part of the US Information Infrastructure Task Force, formed in February 1993.

72. Raymond B. Wolfgang and Edward J. Delp. A watermark for digital images. In *International Conference on Images Processing*, pages 219–222, Lausanne, Switzerland, September 1996. IEEE.
73. Raymond B. Wolfgang and Edward J. Delp. A watermarking technique for digital imagery: further studies. In *International Conference on Imaging, Systems, and Technology*, pages 279–287, Las Vegas, NV, USA, 30 June–3 July 1997. IEEE.
74. J. Zhao and E. Koch. Embedding robust labels into images for copyright protection. In *International Congress on Intellectual Property Rights for Specialised Information, Knowledge and New Technologies*, Vienna, Austria, August 1995.
75. Jian Zhao. A WWW service to embed and prove digital copyright watermarks. In *European Conference on Multimedia Applications, Services and Techniques*, pages 695–710, Louvain-la-Neuve, Belgium, May 1996.
76. Jian Zhao. The syscop home page. `<http://syscop.igd.fhg.de/>` or `<http://www.crcg.edu/syscop/>`, February 1997.

Lecture Notes in Computer Science 1174

Series Editors

Gerhard Goos, Karlsruhe University, Germany

Juris Hartmanis, Cornell University, NY, USA

Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editor

Ross Anderson
Cambridge University, Computer Laboratory
Pembroke Street, Cambridge CB2 3QG, UK
E-mail: rja14@cl.cam.ac.uk

Cataloging-in-Publication data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Information hiding : first international workshop, Cambridge, UK, May 30 - June 1, 1996 ; proceedings / Ross Anderson (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Budapest ; Hong Kong ; London ; Milan ; Paris ; Santa Clara ; Singapore ; Tokyo : Springer, 1996
(Lecture notes in computer science ; Vol. 1174)
ISBN 3-540-61996-8
NE: Anderson, Ross [Hrsg.]; GT

CR Subject Classification (1991): E.3, K.6.5, D.4.6, E.4, C.2, J.1, K.4.1, K.5.1, H.4.3

ISSN 0302-9743
ISBN 3-540-61996-8 Springer-Verlag Berlin Heidelberg New York

Printed on acid-free paper

Stretching the Limits of Steganography

Ross Anderson

Cambridge University Computer Laboratory
Pembroke Street, Cambridge CB2 3QG, UK
Email rja14@cl.cam.ac.uk

Abstract. We present a number of insights into information hiding. It was widely believed that public key steganography was impossible; we show how to do it. We then look at a number of possible approaches to the theoretical security of hidden communications. This turns out to hinge on the inefficiency of practical compression algorithms, and one of the most important parameters is whether the opponent is active or passive (i.e., whether the censor can add noise, or will merely allow or disallow a whole messages). However, there are covertexts whose compression characteristics are such that even an active opponent cannot always eliminate hidden channels completely.

1 Introduction

Steganography is about concealing the existence of messages, and it goes back to ancient times. Kahn tells of a classical Chinese practice of embedding a code ideogram at a prearranged place in a dispatch; of the warning the Greeks received of Xerxes' intentions from a message underneath the wax of a writing tablet; and a trick of dotting successive letters in a covertext with secret ink, due to Aeneas the Tactician [8].

The opponent may be passive, and merely observe the covertext, but he may also be active. In the US post office during the second world war, postal censors deleted lovers' X's, shifted watch hands, and replaced items such as loose stamps and blank paper. They also rephrased telegrams; in one case, a censor changed 'father is dead' to 'father is deceased', which elicited the reply 'is father dead or deceased?'

The study of this subject in the open scientific literature may be traced to Simmons, who in 1983 formulated it as the prisoners' problem [16]: Alice and Bob are in jail, and wish to hatch an escape plan. All their communications pass through the warden, Willy. If Willy sees any encrypted messages, he will frustrate their plan by putting them into solitary confinement. So they must find some way of hiding their ciphertext in an innocuous looking covertext. As in the related field of cryptography, we assume that the mechanism in use is known to the warden, and so the security must rely solely on a secret key.

There are many real life applications of steganography. Apparently, during the 1980's, British Prime Minister Margaret Thatcher became so irritated at

press leaks of cabinet documents that she had the word processors programmed to encode their identity in the word spacing of documents, so that disloyal ministers could be traced. Similar techniques are now undergoing trials in an electronic publishing project, with a view to hiding copyright messages and serial numbers in documents [10].

Simmons' real application was more exotic — the verification of nuclear arms control treaties. The US and the USSR wanted to place sensors in each others' nuclear facilities that would transmit certain information (such as the number of missiles) but not reveal other kinds of information (such as their location). This forced a careful study of the ways in which one country's equipment might smuggle out the forbidden information past the other country's monitoring facilities [17, 19].

Steganography must not be confused with cryptography, where we transform the message so as to make its meaning obscure to a person who intercepts it. Such protection is often not enough: the detection of enciphered message traffic between a soldier and a hostile government, or between a known drug-smuggler and someone not yet under suspicion, has obvious implications.

However, we still have no comprehensive theory of steganography, in the way that Shannon gave us a theory of encryption [15] and Simmons of authentication [18]. In this article, we will try to move a few small steps towards such a theory.

2 The State of the Art

A number of computer programs are available that will embed a ciphertext file in an image. The better systems assume that both sender and receiver share a key and use a conventional cryptographic keystream generator [13] to expand this into a long pseudo-random keystream. The keystream is then used to select pixels in which the bits of the ciphertext are embedded.

Of course, not every pixel may be suitable for encoding ciphertext: changes to pixels in large fields of monochrome colour, or that lie on sharply defined boundaries, might be visible. So some systems have an algorithm that determines whether a candidate pixel can be used by checking that the variance in luminosity of the eight surrounding pixels is neither very high (as on a boundary) nor very low (as in a monochrome field). A bit can be embedded in a pixel that passes this test by some rule such as setting its low order bit to the parity of the surrounding pixels (though in practice one might use something slightly more complicated to avoid leaving telltale statistics).

Of course, the more bits per pixel, the less correlated the low order bits will be with neighbouring bits and with higher order bits in the same pixel. Some quantitative measurements of the correlations between pixels on different bit planes in digital video may be found in [20]. In effect, the bits that Alice can use to embed covert data are redundant in that Willy will be unaware that they have been altered. It follows that they might be removed by an efficient compression scheme, if one exists for the image or other covertext in use.

So when the image is to be subjected to compression (whether before or after the insertion of covert material), things become more complicated, and we have to tailor the embedding method. For example, with `.gif` files one can swap colours for similar colours that are adjacent in the current palette [7], while if we want to embed a message in a file that may be subjected to JPEG compression and filtering, we can embed it in multiple locations [9] or in the frequency domain by altering components of the image's discrete cosine transform [3] [23]. Further papers on the topic may be found in this volume.

So the general model is that Alice embeds information by tweaking some bits of some transform of the covertext. The transform enables her to get at one or more bits which are redundant in the sense that tweaking them cannot be detected easily or at all. To a first approximation, we will expect that such transforms will be similar to those used for compression, and that there are many low-bandwidth stego channels arising from redundancy whose elimination, by compression or otherwise, is uneconomic for normal users of the cover system. We will not expect to find many high bandwidth channels, as these would normally correspond to redundancy that could economically be removed.

3 Public Key Steganography

So far, we have merely stated the general intuition of people who have thought about these topics. They generally assume that steganography, in the presence of a capable motivated opponent who is aware of the general methods that might be used, requires the pre-existence of a shared secret so that the two communicating parties can decide on which bits to tweak. So there has been a general assumption that public-key steganography is impossible.

However, this is not the case. We will now show how a hidden message can be sent to a recipient with whom the sender has no shared secret, but for whom an authentic public key is available.

Given a covertext in which any ciphertext at all can be embedded, then there will usually be a certain rate at which its bits can be tweaked without the warden noticing (we will discuss this more fully below). So suppose that Alice can modify at least one out of every hundred bits of the covertext. This means that Willy cannot distinguish the parity of each successive block of a hundred bits from random noise, and it follows that she can encode an arbitrary pseudorandom string in these parities.

This pseudorandom material will lie in plain sight; anyone will be able to read it. So Willy cannot simply check a covertext by seeing whether a pseudorandom string can be found in it. Indeed, a suitable parity check function will extract pseudorandom-looking data from any message in which covert information can be inserted at all.

Now suppose that Alice and Bob did not have the opportunity to agree a secret key before they were imprisoned, but that Bob has a public key that is known to Alice. She can take her covert message, encrypt it under his public key,

and embed it as the parity of successive blocks. Each possible recipient will then simply try to decrypt every message he sees, and Bob alone will be successful. In practice, the value encrypted under a public key could be a control block consisting of a session key plus some padding, and the session key would drive a conventional steganographic scheme as described elsewhere in this volume.

Normal public key cryptography means that users can communicate confidentially in the absence of previously shared secrets; our construction of public key steganography shows that they can also communicate covertly (if this is at all possible for people with previously shared secrets). Public key stego scales less well than public key crypto, as every recipient has to try to decrypt every message. However, this appears to be an intrinsic property of anonymous communications.

4 Theoretical Limits

Can we get a scheme that gives unconditional covertness, in the sense that the one-time pad provides unconditional secrecy?

Suppose that Alice uses an uncompressed digital video signal as the covertext, and encodes ciphertext at a very low rate. For example, the kth bit of ciphertext might become the least significant bit of one of the pixels of the kth frame of video, with the choice of pixel being specified by the kth word of a shared one time pad. Then we intuitively expect that attacks will be impossible: the ciphertext will be completely swamped in the covertext's intrinsic noise. Is there any way this intuitively obvious fact could be rigorously proved?

This leads us to ask what a proof of perfect covertness would look like. A working definition of a secure stegosystem might be one for which Willy cannot differentiate between raw covertext and the stegotext containing embedded information, unless he has knowledge of the key. As in the case of cryptography, we might take Willy to be a probabilistic polynomial Turing machine in the case where we require computational security, and assume that he can examine all possible keys in the case where we require unconditional security.

In the latter case, he will see the actual message, so the system must generate enough plausible messages from any given stegotext, and the number of such messages must not vary in any usable way between the stegotext and a wholly innocent covertext.

This much is straightforward, but what makes the case of steganography more difficult than secrecy or authenticity is that we are dependent on the model of the source. There are a number of ways in which we can tackle this dependence, and we will present three of them. It is an open question whether any of them will yield useful results in any given application.

4.1 Selection channel

Our first idea is inspired by the correction channel that Shannon uses to prove his second coding theorem. This is the channel which someone who can see both the transmitted and received signals uses to tell the receiver which bits to tweak, and produces various noise and error correction bounds [14].

In a similar way, when Alice and Bob use a shared one-time pad to decide which covertext bit will contain the next ciphertext bit, we can think of the pad as a selection channel. If Willie is computationally unbounded, he can try all possible pads (including the right one), so the number of them which yield a plausible ciphertext must be large enough that he cannot reasonably accuse Alice of sending stegotext rather than an innocent message.

It may be useful at this point to recall the book cipher. The sender and receiver share a book and encipher a message as a series of pointers to words. So the cipher group '78216' might mean page 78, paragraph 2 and the 16th word. Book codes can be secure provided that the attacker does not know which book is in use, and care is taken not to reuse a word (or a word close enough to it) [8]. The book cipher is just a selection channel. The model of computation may be different, in that with a book cipher we start off with the book and then generate the ciphertext, whereas in a stegosystem, we start off with the text to be embedded and then create the stegotext; but they are clearly related.

A repetitive book will have a lower capacity, as we will be able to use a smaller percentage of its words before inference attacks from the context become possible. Similarly, if the covertext to be used in a stegosystem has unusual statistics (such as an unequal number of zeros and ones) then its stego capacity will be lower, as only a small proportion of candidate ciphertexts would look random enough.

We mentioned systems that generate a number of candidate locations for a ciphertext bit and then filter out the locations where actually embedding a bit would have a significant effect on the statistics thought to be relevant (in the case of hiding in an image, this could mean avoiding places where the local variance in luminosity is either very low or very high).

Our information theoretic approach suggests a better way. We will use our keystream generator to select not one pixel but a set of them, and embed the ciphertext bit as their parity. This way, the information can be hidden by changing whichever of the pixels can be changed least obtrusively.

From the information theoretic point of view, if the covertext is '1' with probability 0.6, and we encode in bit pairs, then the probability that a bit pair will have parity 1 is 0.52; if we move to triples, the parity is 1 with probability 0.504, and so on. As the improvement is geometric, we can with little difficulty get the discrepancy as low as we like and thus — provided that our selection channel is broad enough — we can hide information quite efficiently.

Indeed, there is a tradeoff: the more bits in the selection channel, the more bits we can hide in the covertext. In practice our selection channel will be a

cryptographic pseudorandom number generator, and we can draw from it as many bits as we like.

But clearly there is an ultimate limit to the amount of information that can be hidden.

For example, suppose that there is an allowed set of cover texts M (we might be using the cover of a news agency; we have to report a reasonably truthful version of events, and transmit photographs — perhaps slightly doctored — of events that actually took place). Suppose also that there is an allowed set of encodings E. Then the covert capacity will be at most $H(E) - H(M)$. Can we tighten this upper bound or — better still — get any kind of lower bound?

4.2 Equivalence classes

We can consider equivalence classes of messages. Suppose Alice uses a keyed cryptographic hash function to derive one bit from each sentence of a letter. As she prepares her stegotext letter to Bob, she has a routine which checks this bit and beeps if it is wrong. This will go off about every other sentence, which she can then rewrite.

Of course, if she uses standard changes such as [is able $\leftrightarrow$ can], then clearly she must not alter their statistics to the point that Willy can detect the change. It is even an open question whether a computer can alter a natural language text in a way that is undetectable to a human [21] — that is, embed a ciphertext using the technique described above — and the problem is commended to the research community as a weak form of the Turing test.

The converse is that writing a program to scan for human inserted steganography might be rather hard. This is an area in which competitive research might prove interesting!

Equivalence classes can also arise naturally. For example, when making a map from a larger scale map, many arbitrary decisions have to be taken about which features to incorporate, especially with features such as coastlines that are to some extent fractal [12]. Also, when software is written, it contains 'birthmarks' such as the order in which registers are pushed and popped, and these were used by IBM in litigation against software pirates who had copied their PC-AT ROM [6].

Equivalence classes of messages are tied up with compression. If covertext C_1 means the same as covertext C_2, then a compression algorithm need only select one representative from this equivalence class. However, if $C_1 \neq C_2$, then this choice throws away information, and so the compression is lossy.

Information theorists assume that any signal can in theory be completely compressed. But if this could ever be done in practice, then the steganography problem would become trivial: Alice can just 'uncompress' her ciphertext getting a comprehensible message, and Willy would have to pass the result. So Willy can only do his job at all because perfect compression is unfeasible. So regardless of whether we talk in terms of compression or of equivalence classes, Willy's

performance of his job depends on his having a model of the source, and the danger to Alice and Bob is that his model might be better than theirs.

4.3 Conditional entropy

If Alice does not trust her ability to spot which bits of the covertext (or some transform of it) will appear noise-like to Willy, she might try to use conditional entropy. If the ciphertext is random, its mutual information with the covertext will be zero, and entropy will be additive — the entropy of the loaded covertext will be exactly equal to the entropy of the pure covertext plus that of the ciphertext.

Thus the maximum amount of ciphertext that she can hide from Willy depends on the uncertainty of the covertext entropy. Confidence levels can be computed in the obvious way: if the ciphertext rate is R_1, then the warden must not be able to detect with probability P a change of R_1 in the covertext rate R_2.

Thanks to the Central Limit Theorem, the more covertext we give the warden, the better he will be able to estimate its statistics, and so the smaller the rate at which Alice will be able to tweak bits safely. The rate might even tend to zero, as was noted in the context of covert channels in operating systems [11]. However, as a matter of empirical fact, there do exist channels in which ciphertext can be inserted at a positive rate [4], so measuring entropy may be useful in a number of applications.

However, it still does not give us a way to prove the unconditional covertness of a system. The reason for this is that once Alice assumes that Willy is smarter than she is, she has no way of estimating the variance in his estimates of the entropy of her covertext. A purist might conclude that the only circumstance in which she can be certain that Willy cannot detect her messages is when she uses a subliminal channel in the sense of Simmons; that is, a channel in which she chooses some random bits (as in an ElGamal digital signature) and these bits can be recovered by the message recipient [1].

5 Active and Passive Wardens

The applications discussed above include both passive wardens, who monitor traffic and signal to some process outside the system if unauthorised message traffic is detected, and active wardens who try to remove all possible covert messages from covertexts that pass through their hands. A good example of the latter was the world war two postal censor described in the introduction, and a highly topical example is given by software piracy.

Software birthmarks, as mentioned above, have been used to prove the authorship of code so that pirates could be prosecuted. They were serviceable with hand assembled system software, but might be harder to find now that

most code is produced by a compiler. A possible remedy is to embed copyright information by mangling the object code in some way. The automatic, random replacement of code fragments with equivalent ones is used by Intel to customise security code [2]. This may be adequate in that application, where the goal is to prevent a single patch defeating all instances of a protective mechanism; but copyright marking is harder. One could imagine a contest between software authors and pirates to see who can mangle code most thoroughly without affecting its performance too much. If the author has the better mangler, then some of the information he adds will be left untouched by the pirate.

In fact, the World Intellectual Property Organisation has proposed a system of numbering for all digital works, including books, sound and video recordings, and computer programs; it claims that the boundaries between these are breaking down. Software publishers are sceptical; they claim to have had no difficulty yet in establishing ownership [5]. But whatever the legal value of copyright marking, the software pirate is a good example of an active warden.

In such a case, the simple public key scheme described in section two above will not work. Even in the shared-key model, there are cases where an active warden can completely block the stego channel. For example, if (a) his model of the communication at least as good as the prisoners' (b) the covertext information separates cleanly from the covert information, then he can replace the latter with noise. This is the case of a software pirate who has a better code mangler than the software author.

6 Limits on Active Wardens

However, there are many other cases where the stego channel is highly bound up with the covertext. For example, Jagpal [7] measured the noise that can be added to a `.gif` file before the image quality is degraded, while Möller and others have done the same for digitised speech [4].

The point here is that if Alice can add an extra X% of noise without affecting the picture, then so can Willy; but she can stop him finding out which X% carries the covert message by using a keystream to select which bits of covertext to tweak. In this case, all Willy will be able to do is to cut the bandwidth of the channel — a scenario that Trostle and others have explored in the context of covert channels in operating systems [22].

This bandwidth limitation will also be effective against systems that embed each ciphertext bit as a parity check of a number of covertext bits. When the warden is active, the more covertext bits we use in each parity check, the more easily he will be able to inject noise into our covertext.

It is an open question whether public key steganography can be made to work against an active warden who can add only a limited amount of noise. It may also be of interest to consider whether one can implement other cryptographic primitives, such as the wiretap channel and bit commitment [13]. If it turns out that the kind of public key steganography that we have described here cannot be

made to work, then key exchange well might be possible by combining techniques like these.

7 Conclusions

We have stretched the limits of steganography somewhat. Firstly, we have shown how to do public key steganography. Secondly, we have discussed a number of possible approaches to a theory of the subject, which suggest various practical techniques for improving the covertness of existing steganographic schemes. Thirdly, we have highlighted one of the most important topics, namely whether the warden is active or passive, and shown how this interacts with both the public key and theoretical approaches to the subject.

Acknowledgements: Some of the ideas presented here were clarified by discussion with David Wheeler, John Daugman, Roger Needham, Gus Simmons, Markus Kuhn, John Kelsey, Ian Jackson, Mike Roe, Mark Lomas, Stewart Lee, Peter Wayner and Matt Blaze. I am also grateful to the Isaac Newton Institute for hospitality while this paper was being written.

References

1. "The Newton Channel", RJ Anderson, S Vaudenay, B Preneel, K Nyberg, *this volume*
2. "Tamper Resistant Software: An Implementation", D Aucsmith, *this volume*
3. "Watermarking Digital Images for Copyright Protection", FM Boland, JJK Ó Ruanaidh, C Dautzenberg, *Proceedings, IEE International Conference on Image Processing and its Applications, Edinburgh 1995*
4. "Computer Based Steganography", E Franz, A Jerichow, S Moeller, A Pfitzmann, I Stierand, *this volume*
5. "A voluntary international numbering system — the latest WIPO proposals", R Hart, *Computer Law and Security Report* v 11 no 3 (May-June 95) pp 127–129
6. Talk on software birthmarks, counsel for IBM Corporation, BCS Technology of Software Protection Special Interest Group, London 1985
7. *'Steganography in Digital Images'*, G Jagpal, Thesis, Cambridge University Computer Laboratory, May 1995
8. *'The Codebreakers'*, D Kahn, Macmillan 1967
9. "Towards Robust and Hidden Image Copyright Labeling", E Koch, J Zhao, *Proceedings of 1995 IEEE Workshop on Nonlinear Signal and Image Processing* (Neos Marmaras, Halkidiki, Greece, June 20–22, 1995)
10. "Electronic Document Distribution", NF Maxemchuk, *AT & T Technical Journal* v 73 no 5 (Sep/Oct 94) pp 73–80
11. "Covert Channels — Here to Stay?", IS Moskowitz, MH Kang, *Compass 94* pp 235–243
12. RM Needham, *private conversation*, December 1995
13. *'Applied Cryptography — Protocols, Algorithms and Source Code in C'* B Schneier (second edition), Wiley 1995

14. "A Mathematical Theory of Communication", CE Shannon, in *Bell Systems Technical Journal* v 27 (1948) pp 379–423, 623–656
15. "Communication theory of secrecy systems", CE Shannon, in *Bell Systems Technical Journal* v 28 (1949) pp 656–715
16. "The Prisoners' Problem and the Subliminal Channel", GJ Simmons, in *Proceedings of CRYPTO '83*, Plenum Press (1984) pp 51–67
17. "How to Insure that Data Acquired to Verify Treaty Compliance are Trustworthy", GJ Simmons, *Proceedings of the IEEE* v 76 (1984) p 5
18. "A survey of information authentication", GJ Simmons, in *Contemporary Cryptology — the Science of information Integrity*, IEEE Press 1992, pp 379–419
19. "The History of Subliminal Channels", GJ Simmons, *this volume*
20. *'High Quality De-interlacing of Television Images'*, N van Someren, PhD Thesis, University of Cambridge, September 1994
21. K Spärck Jones, *private communication*, August 1995
22. "Modelling a Fuzzy Time System", JT Trostle, *Proc. IEEE Symposium in Security and Privacy 93* pp 82 - 89
23. "Embedding Robust Labels Into Images For Copyright Protection", J Zhao, E Koch, *Proc. Int. Congr. on IPR for Specialized Information, Knowledge and New Technologies* (Vienna, Austria, August 21-25, 1995)

Rotation, Scale and Translation Invariant Digital Image Watermarking

Joseph J.K. Ó Ruanaidh Thierry Pun

Groupe de Vision par Ordinateur,
Centre Universitaire d'Informatique,
Université de Genève,
CH-1211 Genève 4, Switzerland

Abstract

A digital watermark is an invisible mark embedded in a digital image which may be used for Copyright Protection. This paper proposes that Fourier-Mellin transform-based invariants can be used for digital image watermarking. The embedded marks may be designed to be unaffected by any combination of rotation, scale and translation transformations. The original image is not required for extracting the embedded mark.

1 Introduction

Computers, printers and high rate digital transmission facilities are becoming less expensive and more widespread. Digital networks provide an efficient cost-effective means of distributing digital media. Unfortunately however, digital networks and multimedia also afford virtually unprecedented opportunities to pirate copyrighted material. The idea of using a robust digital watermark to detect and trace copyright violations has therefore stimulated significant interest among artists and publishers. As a result, digital image watermarking has recently become a very active area of research. Techniques for hiding watermarks have grown steadily more sophisticated and increasingly robust to lossy image compression and standard image processing operations, as well as to cryptographic attack.

Many of the current techniques for embedding marks in digital images have been inspired by methods of image coding and compression. Information has been embedded using the Discrete Cosine Transform (DCT) [6, 2] Discrete Fourier Transform magnitude and phase [5], Wavelets [6], Linear Predictive Coding and Fractals. The key to making watermarks robust has been the recognition that in order for a watermark to be robust it must be embedded in the *perceptually significant* components of the image [6, 2]. The term "perceptually significant" is somewhat subjective but it suggests that a good watermark is one which takes account of the behaviour of human visual system. Objective criteria for measuring the degree to which an image component is significant in watermarking have gradually evolved from being based purely on energy content [6, 2] to statistical [7] and psychovisual [3] criteria.

The ability of humans to perceive the salient features of an image regardless of changes in the environment is something which humans take for granted [10]. We can recognize objects and patterns independently of changes in image contrast, shifts in the object or changes in orientation and scale. It seems clear that an embedded watermark should have the same invariance properties as the image it is intended to protect.

Digital watermarking is also fundamentally a problem in digital communications [6, 9, 2]. In parallel with the increasing sophistication in modelling and exploiting the properties of the human visual system, there has been a corresponding development in communication techniques. Tirkel and Osborne [11] were the first to note the applicability of spread spectrum techniques to digital image watermarking. Since then there has been an increasing use of spread spectrum communications in digital watermarking. It has several advantageous features such as cryptographic security [11, 2], and is capable of achieving error free transmission of the watermark near or at the limits set by Shannon's noisy channel coding theorem [6, 9]. Note that the shorter is the core information or "payload" contained in a watermark then the greater is the chances of the watermark being communicated reliably. Spread spectrum is also an example of a symmetric key [8] cryptosystem where system security is based on proprietary knowledge of the keys (or the

seeds (or pseudorandom generators) required to embed, extract or remove an image watermark.

Synchronization of the watermark signal is of the utmost importance during watermark extraction. If watermark extraction is carried out in the presence of the original image then synchronization is relatively trivial. The problem of synchronizing the watermark signal is much more difficult to solve in the case where there is no original image. If the watermarked image is translated, rotated and scaled then synchronization necessitates a search over a four dimensional parameter space (X-offset, Y-offset, angle of rotation and scaling factor). The search space grows even larger if one takes into account the possibility of shear and a change of aspect ratio. In this paper, the aim is to investigate the possibility of using invariant representations of a digital watermark to help avoid the need to search for synchronization during the watermark extraction process. A digital watermark that is invariant to these transformations requires no such search. The tradeoff here is between using a fully invariant representation which may be numerically unstable and expensive to compute with the expense of carrying out a search.

2 Integral Transform Invariants

There are many different kinds of image invariant such as moment, algebraic and projective invariants. In this section we will briefly outline the development of several integral transform based invariants [1].

The invariants described below depend on the properties of the Fourier transform. There are a number of advantages in using a transform based representation. First, using integral transform-based invariants is a a relatively simple generalization of transform domain watermarking. Second, the number of robust invariant components is relatively large which makes it suitable for spread spectrum techniques. Third, as we shall see, mapping to and from the invariant domain to the spatial domain is well-defined and it is, in general, not computationally expensive.

2.1 The Fourier Transform

Let the image be a real valued continuous function $f(x_1, x_2)$ defined on an integer-valued Cartesian grid $0 \leq x_1 < N_1, 0 \leq x_2 < N_2$. Let the two dimensional Discrete Fourier Transform (DFT) $F(k_1, k_2)$ where $0 \leq k_1 < N_1, 0 \leq k_2 < N_2$ be defined in the usual way [4].

2.1.1 The Translation Property

Shifts in the spatial domain cause a linear shift in the phase component:

$$F(k_1, k_2) \exp\left[-j(ak_1 + bk_2)\right] \leftrightarrow f(x_1 + a, x_2 + b) \quad (1)$$

Note that both $F(k_1, k_2)$ and its dual $f(x_1, x_2)$ are periodic functions so it is implicitly assumed that translations cause the image to be "wrapped around". We shall refer to this as a *circular translation*.

2.1.2 Reciprocal Scaling

Scaling the axes in the spatial domain causes an inverse scaling in the frequency domain:

$$\frac{1}{\rho} F\left(\frac{k_1}{\rho}, \frac{k_2}{\rho}\right) \leftrightarrow f(\rho x_1, \rho x_2) \quad (2)$$

2.1.3 The Rotation Property

Rotating the image through an angle θ in the spatial domain causes the Fourier representation to be rotated through the same angle:

$$\begin{aligned} &F(k_1 \cos\theta - k_2 \sin\theta, k_1 \sin\theta + k_2 \cos\theta) \\ &\quad \leftrightarrow f(x_1 \cos\theta - x_2 \sin\theta, x_1 \sin\theta + x_2 \cos\theta) \end{aligned} \quad (3)$$

2.2 Translation Invariance

From the translation property of the Fourier transform it is clear that spatial shifts affect only the phase representation of an image. This leads to the well known result that the DFT magnitude is a circular translation invariant. An ordinary translation can be represented as a cropped circular translation.

2.3 Rotation and Scale Invariance

The basic translation invariants described in section 2.2 may be converted to rotation and scale invariants by means of a *log-polar mapping*. Consider a point $(x, y) \in \Re^2$ and define:

$$\begin{aligned} x &= e^{\mu} \cos\theta \\ y &= e^{\mu} \sin\theta \end{aligned} \quad (4)$$

where $\mu \in \Re$ and $0 \leq \theta < 2\pi$. One can readily see that for every point (x, y) there is a point (μ, θ) that uniquely corresponds to it. Note that in the new coordinate system *scaling* and *rotation* are converted to a translation of the μ and θ coordinates respectively. At this stage one can implement a rotation and scale invariant by applying a translation invariant in the log-polar coordinate system. Taking the Fourier transform of a log-polar map (LPM) is equivalent to computing the Fourier-Mellin transform [1].

2.4 Rotation, Scale and Translation Invariance

Consider two invariant operators: $\mathcal{F}$ which extracts the modulus of the Fourier transform and $\mathcal{F}_{\mathcal{M}}$ which extracts the modulus of the Fourier-Mellin transform.

Applying the hybrid operator $\mathcal{F}_{\mathcal{M}} \circ \mathcal{F}$ to an image $f(x, y)$ we obtain:

$$I_1 = [\mathcal{F}_{\mathcal{M}} \circ \mathcal{F}]\, f(x, y) \tag{5}$$

Let us also apply this operator to an image that has been translated, rotated and scaled:

$$\begin{aligned} I_2 &= [\mathcal{F}_{\mathcal{M}} \circ \mathcal{F} \circ \mathcal{R}(\theta) \circ \mathcal{S}(\rho) \circ \mathcal{T}(\alpha, \beta)]\, f(x, y) \\ &= [\mathcal{F}_{\mathcal{M}} \circ \mathcal{R}(\theta) \circ \mathcal{F} \circ \mathcal{S}(\rho) \circ \mathcal{T}(\alpha, \beta)]\, f(x, y) \\ &= \left[\mathcal{F}_{\mathcal{M}} \circ \mathcal{R}(\theta) \circ \mathcal{S}(\tfrac{1}{\rho}) \circ \mathcal{F} \circ \mathcal{T}(\alpha, \beta)\right] f(x, y) \\ &= [\mathcal{F}_{\mathcal{M}} \circ \mathcal{F}]\, f(x, y) \\ &= I_1 \end{aligned} \tag{6}$$

Hence $I_1 = I_2$ and the represention is rotation, scale and translation invariant. The rotation, scale and translation ($\mathcal{RST}$) invariant just described is sufficient to deal with any combination of rotation, scale and translation transformations in any order [1].

3 Watermarking Implementation

Figure 1 illustrates the process of obtaining the $\mathcal{RST}$ transformation invariant from a digital image. Figure 1 is for illustrative purposes only since the process used in practice is more complicated; the main difficulty being that the time and frequency domain are both discretely sampled spaces. The watermark takes the form of a two dimensional spread spectrum signal in the $\mathcal{RST}$ transformation invariant domain. Note that the size of the $\mathcal{RST}$ invariant representation depends on the resolution of the log-polar map which can be kept the same for all images. This is a convenient feature of this approach which helps to standardise the embedding and detection algorithms.

4 Examples

Figure 2 is a standard image which contains a 104 bit rotational and scale invariant watermark. The watermark is encoded as a spread spectrum signal which was embedded in the RS invariant domain. Figure 2 was rotated by 143^O and scaled by a factor of 75% along each axis. The embedded mark which read "The watermark" in ASCII code was recovered from this watermarked image. It was also found that the watermark survived lossy image compression using JPEG at normal settings (75% quality factor). Other methods exist that tolerate JPEG compression down to 5% quality factor [2, 6]; work is underway to combine these with this approach. In addition, the mark is also reasonably resistant to cropping and could be recovered from a segment approximately 50% of the size of the original image.

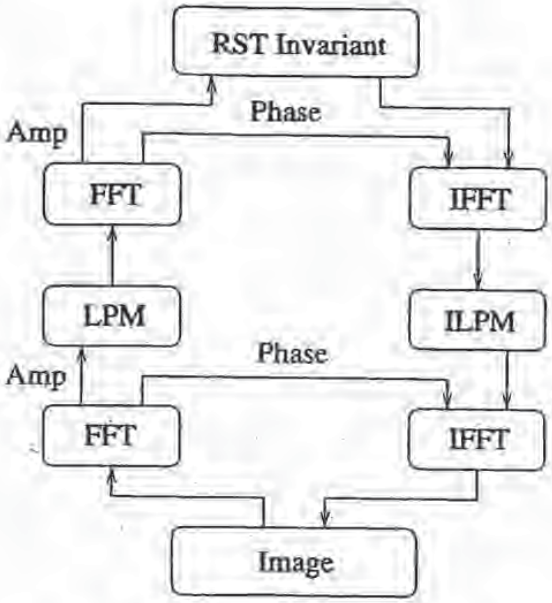


Figure 1: *A diagram of a prototype RST invariant watermarking scheme.*

5 Conclusion

This paper has outlined the theory of integral transform invariants and proposed that this can be used to produce watermarks that are resistant to translation, rotation and scaling. The importance of invertibility of the invariant representation was emphasised. One of the significant points is the application of the Fourier-Mellin transform[1] to digital image watermarking.

An example of a rotation and scale invariant watermark was presented. As one might expect, this proved to be robust to changes in scale and rotation. It was also found to be weakly resistant to lossy image compression and cropping. The robustness of the embedded mark to these attacks will be greatly improved with future work.

On its own, the invariant watermark discussed in this paper cannot resist changes in aspect ratio or shear transformations. There is no obvious means

[1] Digimarc Corporation have independently produced their PictureMarc software which uses the Fourier-Mellin transformation to achieve invariance to rotation and scale transformations. The technical details, which are included in a patent application, are not available to the authors at the time that this paper is being written.

Figure 2: *A watermarked image of a mandrill that has been rotated by 143 degrees and scaled by 75%. The embedded mark was recovered from this image.*

of constructing an integral transform-based operator that is invariant to these transformations. However, work is currently in progress to find a means of searching for the most likely values of aspect ratio and shear factor, and then to apply the necessary corrections during watermark extraction.

Acknowledgments

This work is supported by the Swiss National Science Foundation (grant no. 5003-45334). We wish to thank Dr David McG. Squire, Sergei Starchik and Dr Feng-Lin for their extremely helpful advice on the theory of invariants and Dr A. Z. Tirkel for many stimulating conversations and for exchanging many ideas. We are also grateful to Dr Alexander Herrigel and Adrian Perrig for their useful comments. Thanks also to Geoff Rhoads for answering our queries about PictureMarc.

References

[1] R. D. Brandt and F. Lin. Representations that uniquely characterize images modulo translation, rotation and scaling. *Pattern Recognition Letters*, 17:1001–1015, August 1996.

[2] I. Cox, J. Killian, T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for images, audio and video. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 243–246, Lausanne, Switzerland, September 16-19 1996.

[3] J.F. Delaigle, C. De Vleeschouwer, and B. Macq. Digital Watermarking. In *Conference 2659 - Optical Security and Counterfeit Deterrence Techniques*, San José, February 1996. SPIE Electronic Imaging : Science and Technology. pp. 99-110.

[4] Jae S. Lim. *Two-Dimensional Signal and Image Processing*. Prentice-Hall International, 1990.

[5] J. J. K. Ó Ruanaidh, W. J. Dowling, and F. M. Boland. Phase watermarking of digital images. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 239–242, Lausanne, Switzerland, September 16-19 1996.

[6] J. J. K. Ó Ruanaidh, W. J. Dowling, and F. M. Boland. Watermarking digital images for copyright protection. *IEE Proceedings on Vision, Image and Signal Processing*, 143(4):250–256, August 1996. Invited paper, based on the paper of the same title at the IEE Conference on Image Processing and Its Applications, Edinburgh, July 1995.

[7] I Pitas. A method for signature casting on digital images. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 215–218, Lausanne, Switzerland, September 16-19 1996.

[8] B. Schneier. *Applied Cryptography*. Wiley, 2nd edition, 1995.

[9] J. Smith and B. Comiskey. Modulation and information hiding in images. In Ross Anderson, editor, *Proceedings of the First International Workshop in Information Hiding*, Lecture Notes in Computer Science, pages 207–226, Cambridge, UK, May/June 1996. Springer Verlag.

[10] D. McG. Squire. *Model-based Neural Networks for Invariant Pattern Recognition*. PhD thesis, Curtin University of Technology, Perth, Western Australia, October 1996.

[11] A. Z. Tirkel, G. A. Rankin, R. G. van Schyndel, W. J. Ho, N. R. A. Mee, and C. F. Osborne. Electronic watermark. In *Dicta-93*, pages 666–672, Macquarie University, Sydney, December 1993.

Rotation, Scale and Translation Invariant Digital Image Watermarking

Joseph J.K. Ó Ruanaidh and Thierry Pun

Centre Universitaire d'Informatique, Université de Genève, 24 rue Général Dufour, CH-1211 Genève 4, Switzerland

A digital watermark is an invisible mark embedded in a digital image which may be used for Copyright Protection. This paper describes how Fourier-Mellin transform-based invariants can be used for digital image watermarking. The embedded marks are designed to be unaffected by any combination of rotation, scale and translation transformations. The original image is not required for extracting the embedded mark.

1 Introduction

Computers, printers and high rate digital transmission facilities are becoming less expensive and more widespread. Digital networks provide an efficient cost-effective means of distributing digital media. The popularity of the World Wide Web has clearly demonstrated the commercial potential of the digital multimedia market. Unfortunately however, digital networks and multimedia also afford virtually unprecendented opportunities to pirate copyrighted material. The idea of using a robust digital watermark to detect and trace copyright violations has therefore stimulated significant interest among artists and publishers. As a result, digital image watermarking has recently become a very active area of research. Techniques for hiding watermarks have grown steadily more sophisticated and increasingly robust to lossy image compression and standard image processing operations, as well as to cryptographic attack.

Many of the current techniques for embedding marks in digital images have been inspired by methods of image coding and compression. Information has been embedded using the Discrete Cosine Transform (DCT) [16,34,5,6] Discrete Fourier Transform magnitude and phase [15], Wavelets [16], Linear Predictive Coding [13] and Fractals [9,22]. The key to making watermarks robust has been the recognition that in order for a watermark to be robust it must be embedded in the *perceptually significant* components of the image [16,5,6]. The term "perceptually significant" is somewhat subjective but it suggests that a

* This work is supported by the Swiss National Science Foundation (grant no. 5003-45334)

good watermark is one which takes account of the behaviour of human visual system. Objective criteria for measuring the degree to which an image component is significant in watermarking have gradually evolved from being based purely on energy content [16,5,6] to statistical [20] and psychovisual [27,10] criteria.

Digital watermarking is also fundamentally a problem in digital communications [16,25,5,6]. In parallel with the increasing sophistication in modelling and exploiting the properties of the human visual system, there has been a corresponding development in communication techniques. Early methods of encoding watermarks were primitive and consisted of no more than incrementing an image component to encode a binary '1' and decrementing to encode a '0' [3,16]. Tirkel and Osborne [29] were the first to note the applicability of spread spectrum techniques to digital image watermarking. Since then there has been an increasing use of spread spectrum communications in digital watermarking. It has several advantageous features such as cryptographic security [29,30,6], and is capable of achieving error free transmission of the watermark near or at the limits set by Shannon's noisy channel coding theorem [16,25].

Spread spectrum is an example of a symmetric key [24] cryptosystem. System security is based on proprietary knowledge of the keys (or the seeds for pseudorandom generators) which are required to embed, extract or remove an image watermark. One proviso in the use of a spread spectrum system is that it is important that the watermarking process incorporate some non-invertible step which may depend on a private key or a hash function of the original image. Only in this way can true ownership of the copyright material be resolved [8].

The ability of humans to perceive the salient features of an image regardless of changes in the environment is something which humans take for granted [26,14]. We can recognize objects and patterns independently of changes in image contrast, shifts in the object or changes in orientation and scale. Gibson [12] makes the hypothesis that the human visual system is strongly tied to the ability to recognize invariants. It seems clear that an embedded watermark should have the same invariance properties as the image it is intended to protect. In this paper, we propose that an image watermark should be, so far as possible, encoded to be *invariant* to image transformations. We shall also demonstrate how image invariants can be used to construct watermarks that are unaltered by some of the most basic operations encountered in image processing; namely rotation, translation and changes of scale.

1.1 Nomenclature

This paper will make use of terms agreed during the 1996 Workshop on Information Hiding [18]. The term "cover image" will be used to describe the unmarked original image and "stegoimage" for an image with one or more hidden embedded marks. One significant deviation from the recommended steganographic nomenclature is the frequent use of the term "watermark" to describe the embedded mark. The authors believe this usage is perfectly acceptable because it has become the norm.

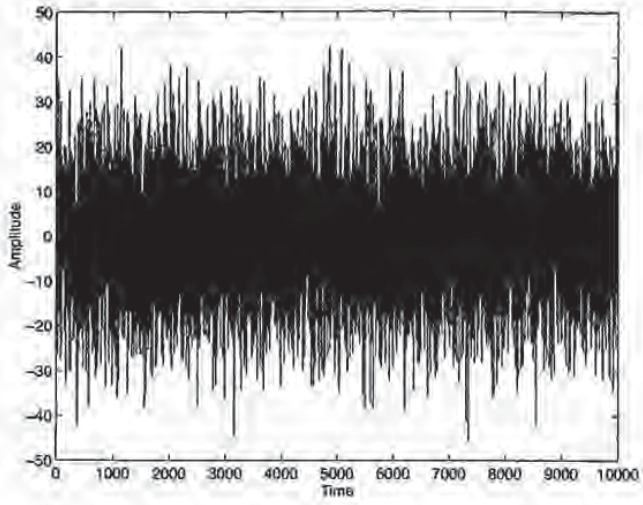


Fig. 1. *An example of a spread spectrum signal used as a digital watermark.*

2 Spread Spectrum

Pickholtz et al. [19] define spread spectrum communications as follows:

> Spread spectrum is a means of transmission in which the signal occupies a bandwidth in excess of the minimum necessary to send the information; the band spread is accomplished by a code which is independent of the data, and a synchronized reception with the code at the receiver is used for despreading and subsequent data recovery.

Spread spectrum systems are also capable of approaching the Shannon limit for reliable communication. The fundamental information theoretic limits to reliable communication and its implications to digital watermarking have been discussed by some authors [16,25]. Note that the smaller is the number of bits of core information or "payload" contained in a watermark, the greater the chance of it being communicated without error.

Cox et al [7,6] recover a watermark by explicitly computing the correlation between the (noise corrupted) watermark recovered from the image with the perfect watermarks stored in a database. This is a very robust technique for watermark recovery but it is not very useful in practice because of the need for access to the database of marks and the large amount of computation required. In this paper the approach is similar to other spread spectrum approaches in that the watermark is embedded in the form of a pseudorandom sequence. However the approach is different to that of Cox in that it does not require access to a database of watermarks and is not particularly expensive computationally. In common with other spread spectrum techniques, in order

to embed a mark or to extract it, it is important to have access to the key which is simply the seed used to generate pseudo-random sequences. In the case of a public watermarking scheme the key is generally available and may even be contained in publically available software. In a private watermarking scheme the key is proprietary. A mark may be embedded or extracted by the key owner which in our model is the Copyright Holder. In this form spread spectrum is a symmetric key cryptosystem. The infrastructure required to generate, issue and store the keys is not described here.

From the point of view of embedding watermarks in documents given the keys or seeds the sequences themselves can be generated with ease. A good spread spectrum sequence is one which combines desirable statistical propeties such as uniformly low cross correlation with cryptographic security. Examples of sequences used in spread spectrum systems used in digital watermarking include m-sequences, Gold codes, Kasami codes and Legendre sequences.

2.1 CDMA coding of digital watermarks

A method for encoding binary messages which can later be recovered given knowledge of the key used is described here. Suppose we are given a message which, without loss of generality, is in binary form $b_1, b_2 \cdots b_L$ where b_i are the bits. This can be written in the form of a set of symbols $s_1, s_2 \cdots s_M$, most generally by a change in a number base from 2 to B with $L \leq M \log_2 B$. The conversion from base 2 to a base which is a power of two is trivial. The next stage is to encode each symbol s_i in the form of a pseudorandom vector of length N. To encode the first symbol a pseudorandom sequence $\vec{v}$ of length $N + B - 1$ is generated. To encode a symbol of values where $0 \leq s < B$ the elements $v_s, v_{s+1} \cdots v_{s+N}$ are extracted as a vector $\vec{r}_1$ of length N. For the next symbol another independent pseudorandom sequence is generated and the symbol encoded as a random vector $\vec{r}_2$. Each successive symbol is encoded in the same way. Note that even if the same symbol occurs in different positions in the sequence that no collision is possible because the random sequences used to encode them are different - in fact they are statistically independent. Finally the entire sequence of symbols is encoded as the summation :

$$\vec{m}(t_i) = \sum_{j=1}^{L} \vec{r}_j(t_i) \tag{1}$$

The pseudo-random vector $\vec{m}$ is decoded by generating all of the random vectors $\vec{r}_i$ in turn and recovering the symbols which the largest value of cross correlation. In this example the pseudo-random generator (PRG) is an m-sequence generator but this is not material to the issue since any "good" generator will do. In addition, one may use two dimensional or higher dimensional arrays in place of the pseudorandom vectors described in the communications system above. One interesting point is that for M sufficiently large the statistical distribution of the message m should approach a Gaussian. This follows from the Central Limit Theorem. A Gaussian distributed watermark has the advantage that it is more difficult to detect. The variance increases with order M - in other words, the expected peak excursion of the sequence is only order

M . One can expect that a message with M = 100 symbols will only have ten times the ampliude of a message with M = 1 symbols. This is very good from the point of view of minimising the visibility of the watermark

Figure 1 shows a spread spectrum signal $s(t)$ composed of a linear combination of L random vectors $r_i(t)$ as given by equation 1. Each random vector is specifically chosen to represent a particular symbol occupying a position in the message. A symbol may be composed of any number of bits. In our case each symbol is eight bits long and the number of random vectors L is nineteen. This is a form of Direct Sequence Code Division Multiple Access (DS-CDMA) spread spectrum communications. The encoded message in Figure 1 reads "This is a watermark".

This form of spread spectrum is resistant to cropping (providing it is resynchronised), non-linear distortions of amplitude and additive noise. Also, if it has good statistical properties it should be mistaken for noise and go undetected by an eavesdropper. The specific choice of method for generating the pseudorandom sequence has direct implications for reliability and cryptographic security of the embedded mark. Pseudorandom number generators described in watermarking literature include Gold Codes, Kasami codes, m-sequences [32,29,33,30] and perfect maps [31].

There are however some drawbacks to using direct sequence spread spectrum. Although a spread spectrum signal as described above is extremely resistant to non-linear distortion of its amplitude and additive noise it is also intolerant of timing errors. Synchronization is of the utmost importance during watermark extraction. If watermark extraction is carried out in the presence of the cover image then synchronization is relatively trivial. The problem of synchronizing the watermark signal is much more difficult to solve in the case where there is no cover image. If the stegoimage is translated, rotated and scaled then synchronization necessitates a search over a four dimensional parameter space (X-offset, Y-offset, angle of rotation and scaling factor). The search space grows even larger if one takes into account the possibility of shear and a change of aspect ratio.

In this paper, the aim is to investigate the possibility of using invariant representations of a digital watermark to help avoid the need to search for synchronization during the watermark extraction process.

2.2 Error control codes

It is desirable to incorporate some form of error control coding into the above scheme. The method is symbol based rather than binary bit based as in normal error codes. Because in this implementation each symbol may be correctly received or not, one finds that errors in the bit stream after despreading will occur in bursts, where each burst is due to an incorrectly decoded symbol. Reed Solomon (RS) codes [4,28,1] are powerful codes which are particularly suited to this application. RS codes can correct both errors (the locations of which are unknown) and erasures (the locations of which are exactly known). The probability of a false detection is extremely low. Reed Solomon codes are

particularly suited to this application for the following reasons : RS codes correct symbol errors rather than bit errors. RS codes can correct erasures as well as errors. Erasures can be factored out of the key equation which means that "erased symbols can be ignored. They do not play any role in the error control mechanism - an erasure is useless redundancy. We recognise that this property of being able to discard erased symbols has two advantages : If the posterior probability of a received symbol is low then it may be ignored. RS codes only come in standard sizes. For example a 255x8 bit code is common. Most commonly used RS error control codes appear to be too large to be used in watermarking. However, it is possible to make almost any RS code fit a watermarking application by judiciously selecting symbols as being erased (because they were never embedded in the document in the first place). For a symbol length of eight bits the corresponding RS code (based on a Galois extension field GF(2^8)) will be 255 symbols long. This is considerably longer than a watermark (typically approximately 100 bits only). However, this is not a problem since the unneeded symbols can be flagged as erasures and they play no part in the decoding process.

3 Integral Transform Invariants

There are many different kinds of image invariant such as moment, algebraic and projective invariants [23,26]. In this section we will briefly outline the development of several integral transform based invariants [26].

The invariants described below depend on the properties of the Fourier transform. There are a number of reasons for this. First, using integral transform-based invariants is a a relatively simple generalization of transform domain watermarking. Second, the number of robust invariant components is relatively large which makes it suitable for spread spectrum techniques. Third, as we shall see, mapping to and from the invariant domain to the spatial domain is well-defined and it is in general not computationally expensive.

3.1 The Fourier Transform

Let the image be a real valued continuous function $f(x_1, x_2)$ defined on an integer-valued Cartesian grid $0 \leq x_1 < N_1, 0 \leq x_2 < N_2$.

The Discrete Fourier Transform (DFT) is defined as follows:

$$F(k_1, k_2) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} f(x_1, x_2) e^{-j2\pi x_1 k_1/N_1 - j2\pi x_2 k_2/N_2} \tag{2}$$

The inverse transform is

$$f(x_1, x_2) = \frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} F(k_1, k_2) e^{j2\pi k_1 x_1/N_1 + j2\pi k_2 x_2/N_2} \tag{3}$$

The DFT of a real image is generally complex valued.

This leads to a

magnitude and phase representation for the image:

$$A(k_1, k_2) = |F(k_1, k_2)| \tag{4}$$

$$\Phi(k_1, k_2) = \angle F(k_1, k_2) \tag{5}$$

We now discuss the properties of the Fourier representation that are crucial to the construction of translation, rotation and scaling invariants.

3.1.1 The Translation Property

Shifts in the spatial domain cause a linear shift in the phase component.

$$F(k_1, k_2) \exp[-j(ak_1 + bk_2)] \leftrightarrow f(x_1 + a, x_2 + b) \tag{6}$$

Note that both $F(k_1, k_2)$ and its dual $f(x_1, x_2)$ are periodic functions so it is implicitly assumed that translations cause the image to be "wrapped around". We shall refer to this as a *circular translation.*

3.1.2 Reciprocal Scaling

Scaling the axes in the spatial domain causes an inverse scaling in the frequency domain.

$$\frac{1}{\rho} F\left(\frac{k_1}{\rho}, \frac{k_2}{\rho}\right) \leftrightarrow f(\rho x_1, \rho x_2) \tag{7}$$

An important example of this property is the Fourier transform of a delta function (which is infinitely narrow) which has a uniformly flat amplitude spectrum (and is infinitely wide).

3.1.3 The Rotation Property

Rotating the image through an angle θ in the spatial domain causes the Fourier representation to be rotated through the same angle.

$$\begin{aligned} &F(k_1 \cos\theta - k_2 \sin\theta, k_1 \sin\theta + k_2 \cos\theta) \\ &\quad \leftrightarrow f(x_1 \cos\theta - x_2 \sin\theta, x_1 \sin\theta + x_2 \cos\theta) \end{aligned} \tag{8}$$

Note that the grid is rotated so the new grid points may not be defined. The value of the image at the nearest valid grid point may be estimated by interpolation.

3.2 Translation Invariance

From property 6 of the Fourier transform it is clear that spatial shifts affect only the phase representation of an image. This leads to the well known result that the DFT magnitude is a circular translation invariant. An ordinary translation can be represented as a cropped circular translation.

It is less well known that it is possible to derive invariants based on the phase representation. To do this involves eliminating the translation dependent linear term from the phase representation. Brandt and Lin [2] present two such translation invariants, namely the *Taylor invariant* which removes the linear phase term in the Taylor expansion of the phase and the *Hessian invariant* which removes this linear phase term by double differentiation.

We shall see in section 3.3 that properties 7 and 8 allow one to extend the basic translation invariants to cover changes of rotation and scale.

3.3 Rotation and Scale Invariance

The basic translation invariants described in section 3.2 may be converted to rotation and scale invariants by means of a *log-polar mapping*.

Consider a point $(x, y) \in \Re^2$ and define:

$$\begin{aligned} x &= e^{\mu} \cos\theta \\ y &= e^{\mu} \sin\theta \end{aligned} \tag{9}$$

where $\mu \in \Re$ and $0 \leq \theta < 2\pi$. One can readily see that for every point (x, y) there is a point (μ, θ) that uniquely corresponds to it.

The new coordinate system has the following properties:

Scaling is converted to a translation.

$$(\rho x, \rho y) \mapsto (\mu + \log\rho, \theta) \tag{10}$$

Rotation is converted to a translation.

$$\begin{aligned} &(x\cos(\theta+\delta) - y\sin(\theta+\delta), x\sin(\theta+\delta) + y\cos(\theta+\delta)) \\ &\quad \mapsto (\mu, \theta+\delta) \end{aligned} \tag{11}$$

At this stage one can implement a rotation and scale invariant by applying a translation invariant in the log-polar coordinate system. Taking the Fourier transform of a log-polar map is equivalent to computing the Fourier-Mellin transform:

$$F_M(k_1,k_2) = \int_{-\infty}^{\infty}\int_{0}^{2\pi} f(e^{\mu}\cos\theta, e^{\mu}\sin\theta)\exp\left[i(k_1\mu + k_2\theta)\right]d\mu\, d\theta \qquad (12)$$

The modulus of the Fourier-Mellin transform is rotation and scale invariant.

Many useful invariants are derived by finding an alternative coordinate system in which the effect of the transformation is replaced by a translation and applying a translation invariant operator in the new coordinate syatem. Squire [26] demonstrates how such invariants can be derived formally using the methods of Lie Group algebra.

3.3.1 The Commutative Property

It is interesting to show that the single parameter group of rotation transformations $\mathcal{R}(\theta)$ and the single parameter group of scale tranformations $\mathcal{S}(\rho)$ commute.

$$\begin{aligned}\mathcal{R}(\theta)\circ\mathcal{S}(\rho)\,f(x,y) &= \mathcal{R}(\theta)\,f(\rho x,\rho y)\\ &= f(\rho x\cos\theta - \rho y\sin\theta, \rho x\sin\theta + \rho y\cos\theta)\\ &= \mathcal{S}(\rho)\,f(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)\\ &= \mathcal{S}(\rho)\circ\mathcal{R}(\theta)\,f(x,y)\end{aligned} \qquad (13)$$

Similarly one can show [2] that the two parameter group of translation transformations $\mathcal{T}(\alpha,\beta)$ commutes neither with $\mathcal{R}(\theta)$, nor with $\mathcal{S}(\rho)$ nor with the joint transformation $\mathcal{RS}(\theta,\rho)$.

3.4 Rotation, Scale and Translation Invariance

Consider two invariant operators: $\mathcal{F}$ which extracts the modulus of the Fourier transform and $\mathcal{F}_{\mathcal{M}}$ which extracts the modulus of the Fourier-Mellin transform. Applying the hybrid operator $\mathcal{F}_{\mathcal{M}}\circ\mathcal{F}$ to an image $f(x,y)$ we obtain:

$$I_1 = [\mathcal{F}_{\mathcal{M}}\circ\mathcal{F}]\; f(x,y) \qquad (14)$$

Let us also apply this operator to an image that has been translated, rotated and scaled:

$$I_2 = [\mathcal{F}_{\mathcal{M}}\circ\mathcal{F}\circ\mathcal{R}(\theta)\circ\mathcal{S}(\rho)\circ\mathcal{T}(\alpha,\beta)]\; f(x,y)$$

$$= [\mathcal{F}_{\mathcal{M}} \circ \mathcal{R}(\theta) \circ \mathcal{F} \circ \mathcal{S}(\rho) \circ \mathcal{T}(\alpha, \beta)]\ f(x, y) \tag{15}$$

$$= \left[\mathcal{F}_{\mathcal{M}} \circ \mathcal{R}(\theta) \circ \mathcal{S}(\frac{1}{\rho}) \circ \mathcal{F} \circ \mathcal{T}(\alpha, \beta)\right]\ f(x, y) \tag{16}$$

$$= [\mathcal{F}_{\mathcal{M}} \circ \mathcal{F}]\ f(x, y) \tag{17}$$

$$= I_1 \tag{18}$$

Hence $I_1 = I_2$ and the represention is rotation, scale and translation invariant. Steps 15 and 16 follow from properties 8 and 7 of the Fourier transform respectively. The contraction in equation 17 is due to the invariance properties of $\mathcal{F}$ and $\mathcal{F}_{\mathcal{M}}$.

The rotation, scale and translation ($\mathcal{RST}$) invariant just described is sufficient to deal with any combination or permutation of rotation, scale and translation in any order [2].

To give a concrete example of its application, consider a copy of a stegoimage placed on a scanner from which we wish to extract an embedded mark. The image may be reduced or increased in size and will be, more often than not, at an angle of $\pm\epsilon$, $\pm 90 \pm \epsilon$ or even $180 \pm \epsilon$ degrees where $\pm\epsilon$ is some small random angle. The image is also likely to be translated. Using the invariants derived above it should be possible to to extract an embedded mark regardless of orientation, scale or position.

3.5 Complete and Strong invariants

Brandt and Lin [2] define the important concept of *completeness*. A complete invariant represents "all the information contained in the image modulo the given transformation". In this sense a complete invariant is *almost* invertible. If two images have the same complete translation invariant then, by the definition of completeness, one must be a shifted version of the other. Such an invariant cannot be inverted uniquely because the mapping to the invariant domain is not a bijective function. Brandt and Lin [2] present an example where a complete Hessian invariant is inverted to yield the original image, albeit with the origin shifted and image wrapped around at the edges.

Ferraro and Caelli [11] in an earlier paper defined the related concept of *strong* invariance. "An integral transform is defined to be invariant in the strong sense if ..." the amplitude representation is constant for all states of the transformation and different states are uniquely encoded in the phase component. The phase component may therefore be used to invert the invariant representation.

For convenience, the invariants used in this paper are strongly invariant. In image watermarking it is more convenient to use strong invariants because the last stage of the process of *embedding* a mark involves inverting the invariant representation to obtain the (marked) stegoimage. Invertibility is of no concern whatsoever during the extraction process.

4 Watermark Implementations

In this section we describe two different prototype schemes for embedding watermarks in digital images using $\mathcal{RST}$ invariants. Typically, the watermark is embedded in a gray scale image or the luminance component of a colour image.

4.1 General scheme

Figure 2 illustrates the process of obtaining the $\mathcal{RST}$ transformation invariant from a digital image. The watermark takes the form of a two dimensional spread spectrum signal in the $\mathcal{RST}$ transformation invariant domain. Note that the size of the $\mathcal{RST}$ invariant representation depends on the resolution of the log-polar map which can be kept the same for all images. This is a convenient feature of this approach which helps to standardise the embedding and detection algorithms.

4.1.1 Embedding a watermark

A Fourier transform (FFT) is first applied which is then followed by a Fourier-Mellin transform (A log-polar mapping (LPM) followed by a Fourier transform (FFT)). The invariant coefficients preselected for their robustness to image processing are marked using a spread spectrum signal. The inverse mapping is computed as an inverse FFT (IFFT) followed by an inverse Fourier-Mellin transform (An inverse log-polar mapping (ILPM) followed by an inverse FFT) Note that the inverse transformation from RST invariant domain to the image domain uses the phase computed during the forward transformations from image domain to the RST invariant domain.

4.1.2 Extracting a watermark

A watermark may be extracted without or without a cover image. In the case where there is no cover image the image is transformed to the $\mathcal{RST}$ invariant domain and the watermark is decoded. This is similar in principle to the scheme described by Smith and Comiskey [25] whose approach is to "treat the image as noise" and overcome the interference from the stegoimage using spread spectrum communication. When a cover image is available it should be subtracted from the stegoimage and the difference transformed to the $\mathcal{RST}$ invariant domain (since the operations in Figure 2 are linear with respect to image amplitude). Subtracting the cover image improves the performance of the detector because, as Smith and Comiskey point out, it eliminates the noise interference due to the stegoimage [25]. In many cases, image contrast may be distorted, for example by a scanner, in which case the effects of change of contrast must be compensated for in some way. Cox et al. [5,6] describe a method known as dynamic histogram warping [7] to carry this out.

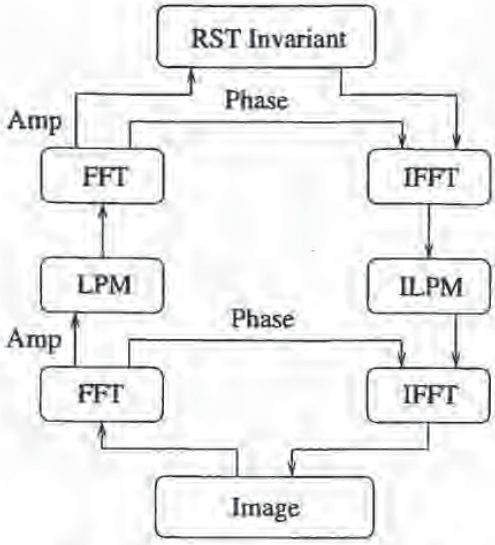


Fig. 2. *A diagram of a prototype RST invariant watermarking scheme.*

4.1.3 Practical considerations

There are a number of complications in implementing the processing steps depicted in Figure 2. The stegoimage must be real which in turn means that its amplitude spectrum ($A(k,l)$ where $0 \leq k < M$ and $0 \leq l < N$) as well as being positive ($A(k,l) \geq 0$) must also be positively symmetric:

$$A(k,l) = A(M-k, N-l) \tag{19}$$

The log-polar map of a positively symmetric matrix consists of two identical halves. This follows from the fact that the positive symmetry condition in equation 19 is written in polar coordinates as:

$$A(r,\theta) = A(r,\pi+\theta) \tag{20}$$

where $(M/2, N/2)$ is the centre of rotation. Since both halves of the log-polar map are identical then only one half need be used in the upper FFT of Figure 2. The spread spectrum signal is determined from the amplitude spectrum of this FFT. Applying the above in reverse gives an embedding algorithm which yields a real valued watermark.

The scheme described in Figure 2 works in principle but has some serious deficiencies in practice. The first difficulty is that both the log-polar mapping and the inverse log-polar mapping can cause a loss of image quality. The change of coordinate system means that some form of interpolation should be used.

Two simple forms of interpolation, nearest neighbour and bilinear interpolation [21], are in common use. Non-stationary low pass filtering can improve the performance by eliminating frequency aliasing. In practice the resolution of the log-polar map must be at least 512×512 for even a quite poor quality image. The second difficulty is numerical. Interpolation only performs well if neighbouring samples are of the same scale. This makes the computation of the Fourier-Mellin transform of the modulus of a Fourier transform somewhat problematic. A typical Fourier transform representation of an image is quite badly behaved in this respect since there are generally a few components of relatively large magnitude. This difficulty is resolved in the next section.

4.2 Cover Image Independent Scheme

The problems in embedding watermarks using the previous implementation described in Figure 2 can be circumvented by using the method illustrated in Figure 3. In this case the mark must be embedded in the $\mathcal{RST}$ invariant domain independently of the original image. The advantage of using this approach is that the distortions caused by the inverse log-polar map are suffered only by the embedded mark itself and do not affect the stegoimage. Figure 4 shows the corresponding detection process which is relatively straightforward.

Note that when embedding the mark there is no phase component available for the first inverse Fourier transform. The first FFT operates on a random phase signal to keep the amplitude distribution of the inverse FFT reasonably flat. This is beneficial to the inverse log-polar map which performs best when the input is a smooth image. The second FFT uses the phase component directly from the cover image. The advantage in doing this is that matching the phase component of the embedded mark to that of the cover image helps to hide it because the embedded mark resembles the cover image. This follows from the research of Oppenheim and Lim [17] which demonstrates that image phase is far more important to image structure than image amplitude.

5 Examples

Figure 5 depicts a standard image of a mandrill. Figure 6 is the log-polar map of Figure 5. This image was computed using 600 grid points along the θ (angle) axis, 600 grid points along the μ (log-radial) axis and bilinear interpolation. Figure 7 is the inverted log-polar map computed using just 100 angular and 100 log-radial grid points and nearest neighbour interpolation. Note that the restoration grows progressively poorer away from the centre.

Figure 5 is in fact a stegoimage which contains a 104 bit rotational and scale invariant watermark. The watermark is encoded as a spread spectrum signal which was embedded in the $\mathcal{RS}$ invariant domain. Figure 5 was rotated by 143^{o} and scaled by a factor of 75% along each axis to give the image shown in Figure 8. The embedded mark which read "The watermark" in ASCII code was recovered from this stegoimage. It was also found that the watermark survived

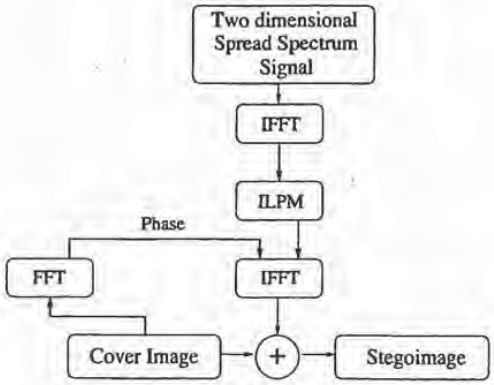


Fig. 3. *A method of embedding a mark in an image which avoids mapping the cover image into the RST invariant domain.*

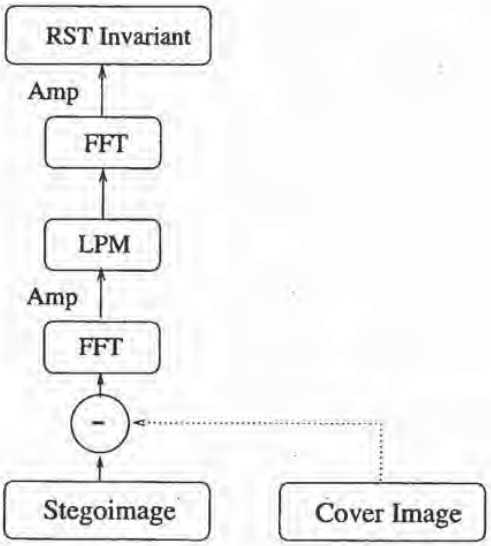


Fig. 4. *A scheme to extract a mark from an image.*

lossy image compression using JPEG at normal settings (75% quality factor). Other methods exist that tolerate JPEG compression down to 5% quality factor [7,6,16,15]; work is underway to combine these with this approach. In addition, the mark is also reasonably resistant to cropping and could be recovered from a segment approximately 50% of the size of the original image.

Fig. 5. *A standard* 500 × 480 *image of a mandrill.*

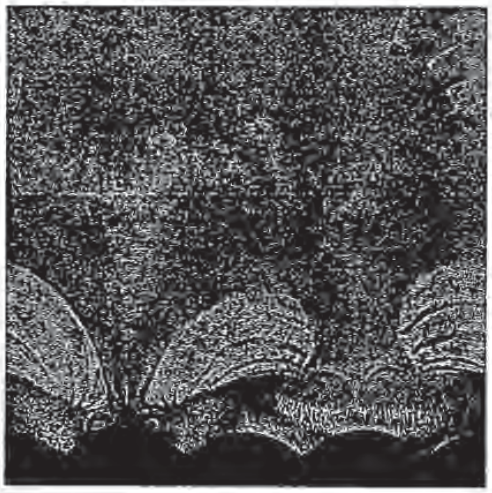


Fig. 6. *A log polar map of the image of a mandrill. The log-polar map employs bilinear interpolation and the log-polar grid is 600 × 600 samples.*

6 Conclusion

This paper has outlined the theory of integral transform invariants and showed that this can be used to produce watermarks that are resistant to translation,

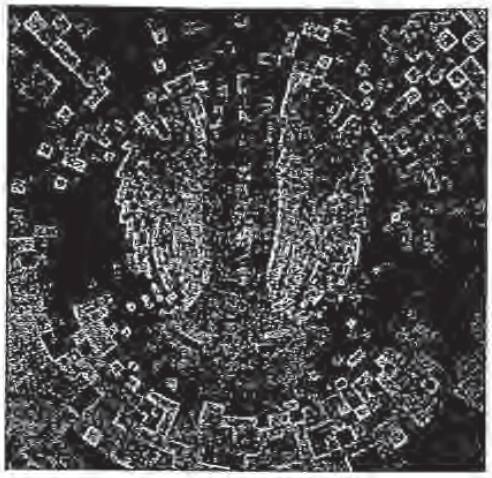


Fig. 7. *The image of a mandrill reconstructed from a log polar map of size 100 × 100 samples. This reconstruction uses nearest neighbour interpolation.*

Fig. 8. *A watermarked image of a mandrill that has been rotated by 143 degrees and scaled by 75%. The embedded mark was recovered from this image.*

rotation and scaling. The importance of invertibility of the invariant representation was emphasised. One of the significant points is the novel application of the Fourier-Mellin transform to digital image watermarking.

There are several advantages in using integral transform domain marks. The main advantage is that the transforms can be computed very quickly (although in practice it has been found that the inverse log-polar mapping is a computational bottleneck). In addition, transform space contains a large number of samples which can be used to hide a spread spectrum signal.

An example of a rotation and scale invariant watermark was presented. As one might expect, this proved to be robust to changes in scale and rotation. It was also found to be weakly resistant to lossy image compression and cropping. The robustness of the embedded mark to these attacks will be greatly improved with future work.

On its own, the invariant watermark discussed in this paper cannot resist changes in aspect ratio or shear transformations. There is no obvious means of constructing an integral transform-based operator that is invariant to these transformations. However, work is currently in progress to find a means of searching for the most likely values of aspect ratio and shear factor, and then to apply the necessary corrections during watermark extraction.

In addition to the above, we intend to investigate the possible use of phase-based complete invariants. This would have some advantage over only marking strong invariants, since a complete invariant presents a maximal number of potential communications channels through which watermark information may be transmitted.

Acknowledgement

We wish to thank Dr David McG. Squire, Sergei Starchik and Dr Feng-Lin for their extremely helpful advice on the theory of invariants and Dr A. Z. Tirkel for many stimulating conversations and for exchanging many ideas. We are also grateful to Dr Alexander Herrigel and Adrian Perrig for their useful comments.

References

[1] R. E. Blahut. *The theory and practice of error control codes.* Addison-Wesley, 1983.

[2] R. D. Brandt and F. Lin. Representations that uniquely characterize images modulo translation, rotation and scaling. *Pattern Recognition Letters*, 17:1001–1015, August 1996.

[3] G. Caronni. Assuring Ownership Rights for Digital Images. In H. H. Brueggemann and W. Gerhardt-Haeckl, editors, *Reliable IT Systems VIS '95.* Vieweg Publishing Company, Germany, 1995.

[4] W. G. Chambers. *Basics of Communications and Coding.* Oxford Science Publications. Clarendon Press Oxford, 1985.

[5] I. Cox, J. Killian, T. Leighton, and T. Shamoon. Secure spread spectrum communication for multimedia. Technical report, N.E.C. Research Institute, 1995. ftp://ftp.nj.nec.com/pub/ingemar/papers/watermark.ps.Z.

[6] I. Cox, J. Killian, T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for images, audio and video. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 243–246, Lausanne, Switzerland, September 16-19 1996.

[7] I. Cox, S. Roy, and S. L. Hingorani. Dynamic histogram warping of image pairs for constant image brightness. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-95*, Austin, Texas, 1995.

[8] S. Craver, N. Memon, B. Yeo, and M. Yeung. Can invisible marks resolve rightful ownerships? In *IS&T/SPIE Electronic Imaging '97: Storage and Retrieval of Image and Video Databases*, 1997.

[9] P. Davern and M. Scott. Fractal based image steganography. In Ross Anderson, editor, *Proceedings of the First International Workshop in Information Hiding*, Lecture Notes in Computer Science, pages 279–294, Cambridge, UK, May/June 1996. Springer Verlag.

[10] J.F. Delaigle, C. De Vleeschouwer, and B Macq. A psychovisual approach for digital picture watermarking, submitted to the Journal of Electronic Imaging, 1996.

[11] Mario Ferraro and Terry M. Caelli. Lie transform groups, integral transforms, and invariant pattern recognition. *Spatial Vision*, 8(1):33–44, 1994.

[12] James Gibson. *The Senses Considered as Perceptual Systems*. Houghton-Mifflin, Boston, Massachusetts, 1966.

[13] K. Matsui and K. Tanaka. Video-Steganography: How to secretly embed a signature in a picture. In *IMA Intellectual Property Project Proceedings*, pages 187–206, January 1994.

[14] R. Milanese, S. Gil, and T. Pun. Attentive mechanisms for dynamic and static scene analysis. *Optical Engineering*, 34(8):2428–2434, August 1995.

[15] J. J. K. Ó Ruanaidh, W. J. Dowling, and F. M. Boland. Phase watermarking of digital images. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 239–242, Lausanne, Switzerland, September 16-19 1996.

[16] J. J. K. Ó Ruanaidh, W. J. Dowling, and F. M. Boland. Watermarking digital images for copyright protection. *IEE Proceedings on Vision, Image and Signal Processing*, 143(4):250–256, August 1996. Invited paper, based on the paper of the same title at the IEE Conference on Image Processing and Its Applications, Edinburgh, July 1995.

[17] A. V. Oppenheim and J. S. Lim. The importance of phase in signals. *Proceedings of the IEEE*, 69(5):529–541, May 1981.

[18] B Pfitzmann. Information hiding terminology. In Ross Anderson, editor, *Proceedings of the First International Workshop in Information Hiding*, Lecture Notes in Computer Science, pages 347–350, Cambridge, UK, May/June 1996. Springer Verlag.

[19] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein. Theory of spread spectrum communications – a tutorial. *IEEE Transactions on Communications*, COM-30(5):855–884, May 1982.

[20] I Pitas. A method for signature casting on digital images. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 215–218, Lausanne, Switzerland, September 16-19 1996.

[21] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C*. Cambridge University Press, second edition, 1992.

[22] J. Puate and F. Jordan. Using fractal compression scheme to embed a digital signature into an image. http://ltswww.epfl.ch/kutter/publications/fvt.html, November 1996.

[23] T. H. Reiss. *Recognizing planar Objects Using Invariant Image Features*. Lecture Notes in Computer Science. Springer-Verlag, 1993.

[24] B. Schneier. *Applied Cryptography*. Wiley, 2nd edition, 1995.

[25] J. Smith and B. Comiskey. Modulation and information hiding in images. In Ross Anderson, editor, *Proceedings of the First International Workshop in Information Hiding*, Lecture Notes in Computer Science, pages 207–226, Cambridge, UK, May/June 1996. Springer Verlag.

[26] D. McG. Squire. *Model-based Neural Networks for Invariant Pattern Recognition*. PhD thesis, Curtin University of Technology, Perth, Western Australia, October 1996.

[27] M. D. Swanson, B. Zhu, and A. Tewfik. Transparent robust image watermarking. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 211–214, Lausanne, Switzerland, September 16-19 1996.

[28] P. Sweeney. *Error Control Coding: An Introduction*. Prentice-Hall, 1991.

[29] A. Z. Tirkel, G. A. Rankin, R. G. van Schyndel, W. J. Ho, N. R. A. Mee, and C. F. Osborne. Electronic watermark. In *Dicta-93*, pages 666–672, Macquarie University, Sydney, December 1993.

[30] A. Z. Tirkel, R. G. van Schyndel, and C. F. Osborne. A two-dimensional digital watermark. In *ACCV'95*, pages 378–383, University of Queensland, Brisbane, December 6-8 1995.

[31] A.Z. Tirkel. Image and watermark registration. Submitted to Signal processing, January 1997.

[32] R. G. van Schyndel, A. Z. Tirkel, and C. F. Osborne. A digital watermark. In *IEEE Int. Conf. on Image Processing ICIP-95*, pages 86–90, Austin, Texas, 1994.

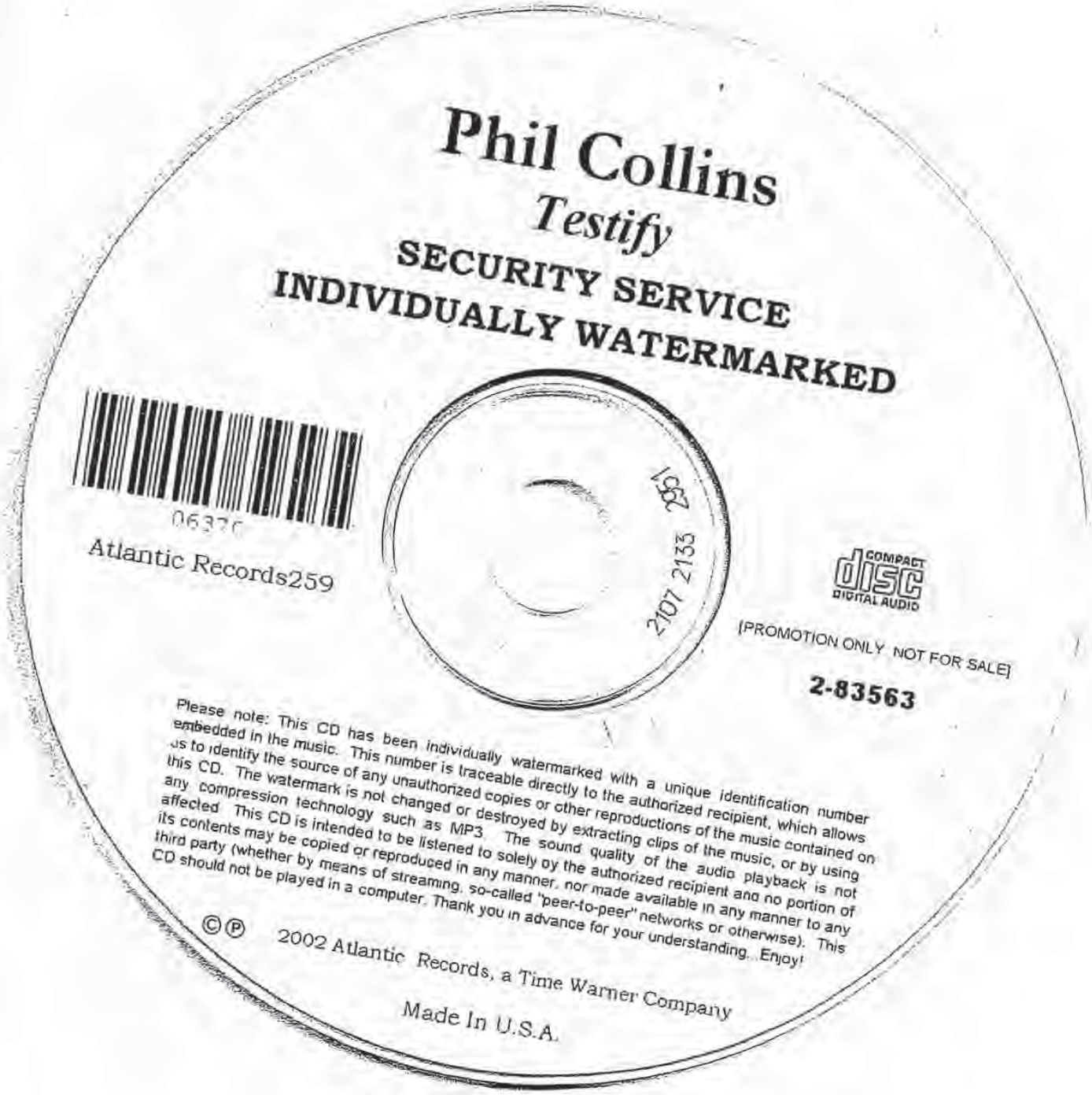
[33] R. G. van Schyndel, A. Z. Tirkel, and C. F. Osborne. Towards a robust digital watermark. In *Dicta-95*, pages 504–508, Nanyang Technological University, Singapore, December 5-8 1995.

[34] J. Zhao and E. Koch. Embedding robust labels into images for copyright protection. Technical report, Fraunhofer Institute for Computer Graphics, Darmstadt, Germany, 1994.

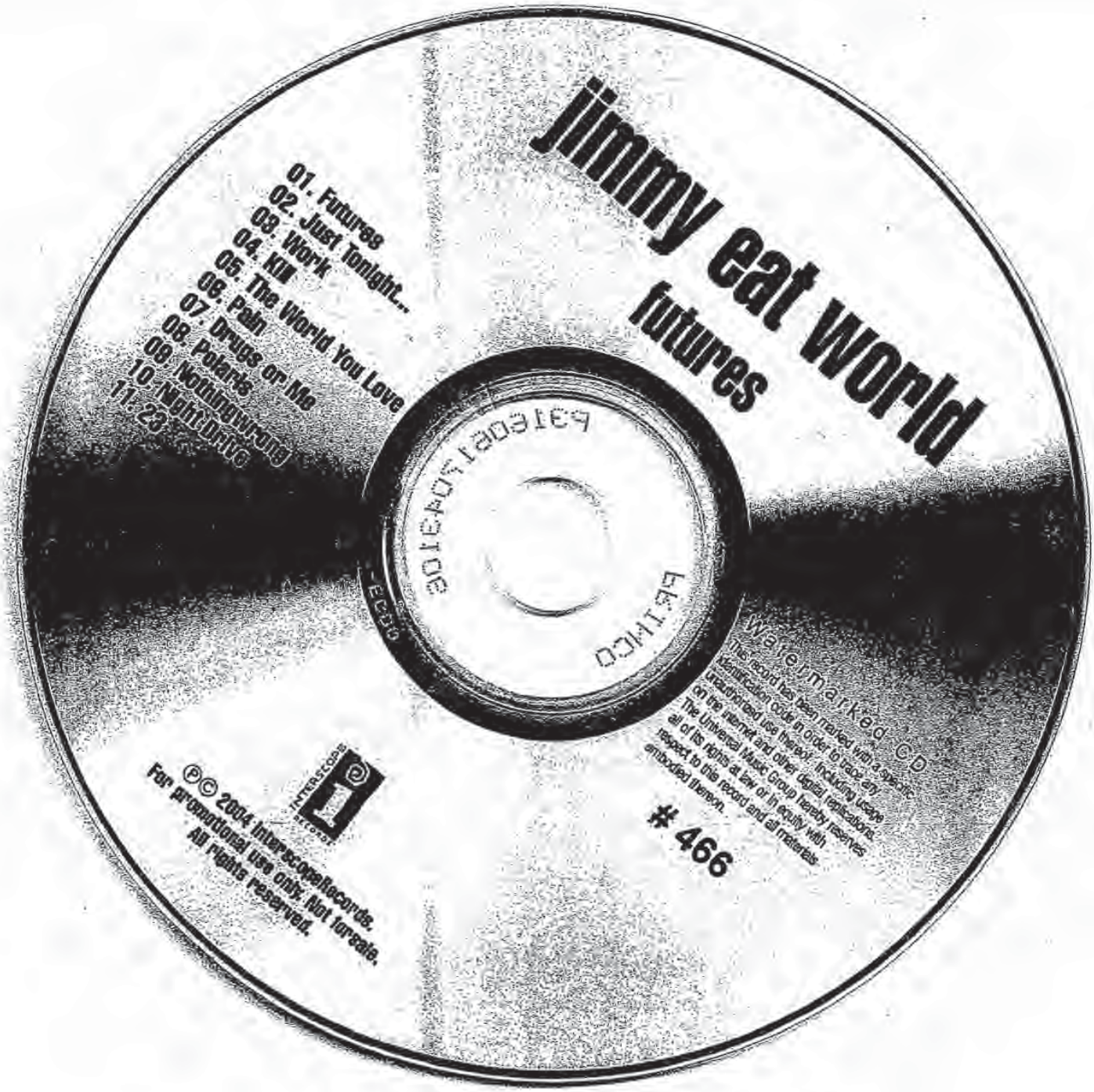
Staind
The Singles 1996-2006
Atlantic
THIS CD BELONGS TO:
Ashley White 142
WATERMARKED DISC!
RESTRICTED RELEASE
USE, REPRODUCTION, TRANSMISSION AND DISTRIBUTION
CD AND THE MUSIC ON IT IS SUBJECT TO AN AGREEMENT SET FORTH ON THE ORIGINAL PACKAGE IN WHICH THIS CD WAS PROVIDED. THIS CD AND THE MUSIC ON IT
DISTRIBUTED EXCEPT AS PROVIDED IN THAT AGREEMENT.


Arctic Monkeys
46. Chris Price
Domino
P 2005 Domino Recording Co
C 2005 Domino Recording Co Ltd


RADIOHEAD
CDP 7243 5 84543 2 1
Capitol
215
VERBATIM CD-R
DataLifePlus CD-R 80
This record has been marked with a specific identification code in order to trace any unauthorized use thereof, including usage on the Internet and other digital replications. The EMI Music Group hereby reserves all of its rights at law or in equity with respect to this record and all materials embodied thereon.
"Hail To The Thief"
ADVANCE COPY - INTERNAL USE ONLY - DO NOT DUPLICATE


Phil Collins
Testify
SECURITY SERVICE
INDIVIDUALLY WATERMARKED
06376
Atlantic Records259
2107 2133
COMPACT disc DIGITAL AUDIO
[PROMOTION ONLY NOT FOR SALE]
2-83563
Please note: This CD has been individually watermarked with a unique identification number embedded in the music. This number is traceable directly to the authorized recipient, which allows us to identify the source of any unauthorized copies or other reproductions of the music contained on this CD. The watermark is not changed or destroyed by extracting clips of the music, or by using any compression technology such as MP3. The sound quality of the audio playback is not affected. This CD is intended to be listened to solely by the authorized recipient and no portion of its contents may be copied or reproduced in any manner, nor made available in any manner to any third party (whether by means of streaming, so-called "peer-to-peer" networks or otherwise). This CD should not be played in a computer. Thank you in advance for your understanding. Enjoy!
©℗ 2002 Atlantic Records, a Time Warner Company
Made In U.S.A.


AEROSMITH
"Just Push Play"
(Final)
(This Product Has Been Watermarked)
Source Date:
February 1, 2001
February 2, 2001
MITSUI COLOR THERMAL
1. Beyond Beautiful (4:45) 2. Just Push Play (3:51)
3. Jaded (3:34) 4. Fly Away From Here (5:01)
5. Trip Hoppin' (4:27) 6. Sunshine (3:37)
7. Under My Skin (3:45) 8. Luv Lies (4:26)
9. Outta Your Head (3:22)
10. Drop Dead Gorgeous (3:42)
11. Light Inside (3:34)
12. Avant Garden (4:52)


jimmy eat world
futures
01. Futures
02. Just Tonight..
03. Work
04. Kill
05. The World You Love
06. Pain
07. Drugs or Me
08. Polaris
09. Nothingwrong
10. Night Drive
11. 23
Watermarked CD
This record has been marked with a specific identification code in order to trace any unauthorized use thereof, including usage on the internet and other digital replications. The Universal Music Group hereby reserves all of its rights at law or in equity with respect to this record and all materials embodied therein.
466
Ⓟ© 2004 InterscopeRecords.
For promotional use only. Not for sale.
All rights reserved.

PATENT
ATTORNEY DOCKET NO. 066112.0132

METHOD AND DEVICE FOR MONITORING AND ANALYZING SIGNALS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of pending U.S. Patent Application Serial No. 08/999,766, filed July 23, 1997, entitled "Steganographic Method and Device"; pending U.S. Patent Application Serial No. 08/772,222, filed December 20, 1996, entitled "Z-Transform Implementation of Digital Watermarks"; pending U.S. Patent Application Serial No. 09/456,319, filed December 8, 1999, entitled "Transform Implementation of Digital Watermarks"; pending U.S. Patent Application Serial No. 08/674,726, filed July 2, 1996, entitled "Exchange Mechanisms for Digital Information Packages with Bandwidth Securitization, Multichannel Digital Watermarks, and Key Management"; pending U.S. Patent Application Serial No. 09/545,589, filed April 7, 2000, entitled "Method and System for Digital Watermarking"; pending U.S. Patent Application Serial No. 09/046,627, filed March 24, 1998, entitled "Method for Combining Transfer Function with Predetermined Key Creation"; pending U.S. Patent Application Serial No. 09/053,628, filed April 2, 1998, entitled "Multiple Transform Utilization and Application for Secure Digital Watermarking"; pending U.S. Patent Application Serial No. 09/281,279, filed March 30, 1999, entitled "Optimization Methods for the Insertion, Protection, and Detection..."; U.S. Patent Application Serial No.09,594,719, filed June 16, 2000, entitled "Utilizing Data Reduction in Steganographic and Cryptographic Systems" (which is a continuation-in-part of PCT application No. PCT/US00/06522, filed March 14, 2000, which PCT application claimed priority to U.S. Provisional Application No. 60/125,990, filed March 24, 1999); pending U.S. Application No 60/169,274, filed December 7, 1999, entitled "Systems, Methods And Devices For Trusted Transactions"; and PCT Application No. PCT/US00/21189, filed August 4, 2000 (which claims priority to U.S. Patent Application Serial No. 60/147,134,

filed August 4, 1999, and to US Patent Application No. 60/213,489, filed June 23, 2000, both of which are entitled, "A Secure Personal Content Server"). The previously identified patents and/or patent applications are hereby incorporated by reference, in their entireties.

In addition, this application hereby incorporates by reference, as if fully stated herein, the total disclosures of US Patent 5,613,004 "Steganographic Method and Device"; U.S. Patent 5,745,569 "Method for Stega-Cipher Protection of Computer Code"; and U.S. Patent 5,889,868 "Optimization Methods for the Insertion, Protection, and Detection of Digital Watermarks in Digitized Data."

BACKGROUND OF THE INVENTION

1. Field of the Invention

The invention relates to the monitoring and analysis of digital information. A method and device are described which relate to signal recognition to enhance identification and monitoring activities.

2. Description of the Related Art

Many methods and protocols are known for transmitting data in digital form for multimedia applications (including computer applications delivered over public networks such as the internet or World Wide Web ("WWW"). These methods may include protocols for the compression of data, such that it may more readily and quickly be delivered over limited bandwidth data lines. Among standard protocols for data compression of digital files may be mentioned the MPEG compression standards for audio and video digital compression, promulgated by the Moving Picture Experts Group. Numerous standard reference works and patents discuss such compression and transmission standards for digitized information.

Digital watermarks help to authenticate the content of digitized multimedia information, and can also discourage piracy. Because piracy is clearly a disincentive to the digital distribution

of copyrighted content, establishment of responsibility for copies and derivative copies of such works is invaluable. In considering the various forms of multimedia content, whether "master," stereo, NTSC video, audio tape or compact disc, tolerance of quality will vary with individuals and affect the underlying commercial and aesthetic value of the content. It is desirable to tie copyrights, ownership rights, purchaser information or some combination of these and related data into the content in such a manner that the content must undergo damage, and therefore reduction of its value, with subsequent, unauthorized distribution, commercial or otherwise. Digital watermarks address many of these concerns. A general discussion of digital watermarking as it has been applied in the art may be found in U.S. Patent No. 5,687,236 (whose specification is incorporated in whole herein by reference).

Further applications of basic digital watermarking functionality have also been developed. Examples of such applications are shown in U.S. Patent No. 5,889,868 (whose specification is incorporated in whole herein by reference). Such applications have been drawn, for instance, to implementations of digital watermarks that were deemed most suited to particular transmissions, or particular distribution and storage mediums, given the nature of digitally sampled audio, video, and other multimedia works. There have also been developed techniques for adapting watermark application parameters to the individual characteristics of a given digital sample stream, and for implementation of digital watermarks that are feature-based - i.e., a system in which watermark information is not carried in individual samples, but is carried in the relationships between multiple samples, such as in a waveform shape. For instance, natural extensions may be added to digital watermarks that may also separate frequencies (color or audio), channels in 3D while utilizing discreteness in feature-based encoding only known to those with pseudo-random keys (i.e., cryptographic keys) or possibly tools to access such information, which may one day exist on a quantum level.

A matter of general weakness in digital watermark technology relates directly to the manner of implementation of the watermark. Many approaches to digital watermarking leave

detection and decode control with the implementing party of the digital watermark, not the creator of the work to be protected. This weakness removes proper economic incentives for improvement of the technology. One specific form of exploitation mostly regards efforts to obscure subsequent watermark detection. Others regard successful over encoding using the same watermarking process at a subsequent time. Yet another way to perform secure digital watermark implementation is through "key-based" approaches.

SUMMARY OF THE INVENTION

A method for monitoring and analyzing at least one signal is disclosed, which method comprises the steps of: receiving at least one reference signal to be monitored; creating an abstract of the at least one reference signal; storing the abstract of the at least one reference signal in a reference database; receiving at least one query signal to be analyzed; creating an abstract of the at least one query signal; and comparing the abstract of the at least one query signal to the abstract of the at least one reference signal to determine if the abstract of the at least one query signal matches the abstract of the at least one reference signal.

A method for monitoring a plurality of reference signals is also disclosed, which method comprises the steps of: creating an abstract for each one of a plurality of reference signals; storing each of the abstracts in a reference database; receiving at least one query signal to be analyzed; creating an abstract of each at least one query signal; locating an abstract in the reference database that matches the abstract of each at least one query signal; and recording the identify of the reference signal whose abstract matched the abstract of each at least one query signal.

A computerized system for monitoring and analyzing at least one signal is also disclosed, which system comprises: a processor for creating an abstract of a signal using selectable criteria; a first input for receiving at least one reference signal to be monitored, the first input being coupled to the processor such that the processor may generate an abstract for each reference

signal input to the processor; a reference database, coupled to the processor, for storing abstracts of each at least one reference signal; a second input for receiving at least one query signal to be analyzed, the second input being coupled to the processor such that the processor may generate an abstract for each query signal; and a comparing device, coupled to the reference database and to the second input, for comparing an abstract of the at least one query signal to the abstracts stored in the reference database to determine if the abstract of the at least one query signal matches any of the stored abstracts.

Further, an electronic system for monitoring and analyzing at least one signal is disclosed, which system comprises: a first input for receiving at least one reference signal to be monitored, a first processor for creating an abstract of each reference signal input to the first processor through the first input; a second input for receiving at least one query signal to be analyzed, a second processor for creating an abstract of each query signal; a reference database for storing abstracts of each at least one reference signal; and a comparing device for comparing an abstract of the at least one query signal to the abstracts stored in the reference database to determine if the abstract of the at least one query signal matches any of the stored abstracts.

DETAILED DESCRIPTION OF THE INVENTION

While there are many approaches to data reduction that can be utilized, a primary concern is the ability to reduce the digital signal in such a manner as to retain a "perceptual relationship" between the original signal and its data reduced version. This relationship may either be mathematically discernible or a result of market-dictated needs. The purpose is to afford a more consistent means for classifying signals than proprietary, related text-based approaches. A simple analogy is the way in which a forensic investigator uses a sketch artist to assist in determining the identity of a human.

In one embodiment of the invention, the abstract of a signal may be generated by the following steps: 1) analyze the characteristics of each signal in a group of audible/perceptible

variations for the same signal (e.g., analyze each of five versions of the same song—which versions may have the same lyrics and music but which are sung by different artists); and 2) select those characteristics which achieve remain relatively constant (or in other words, which have minimum variation) for each of the signals in the group. Optionally, the null case may be defined using those characteristics which are common to each member of the group of versions.

Lossless and lossy compression schemes are appropriate candidates for data reduction technologies, as are those subset of approaches that are based on perceptual models, such as AAC, MP3, TwinVQ, JPEG, GIF, MPEG, etc. Where spectral transforms fail to assist in greater data reduction of the signal, other signal characteristics can be identified as candidates for further data reduction. Linear predictive coding (LPC), z-transform analysis, root mean square (rms), signal to peak, may be appropriate tools to measure signal characteristics, but other approaches or combinations of signal characteristic analysis are contemplated. While such signal characteristics may assist in determining particular applications of the present invention, a generalized approach to signal recognition is necessary to optimize the deployment and use of the present invention.

Increasingly, valuable information is being created and stored in digital form. For example, music, photographs and motion pictures can all be stored and transmitted as a series of binary digits -- 1's and 0's. Digital techniques permit the original information to be duplicated repeatedly with perfect or near perfect accuracy, and each copy is perceived by viewers or listeners as indistinguishable from the original signal. Unfortunately, digital techniques also permit the information to be easily copied without the owner's permission. While digital representations of analog waveforms may be analyzed by perceptually-based or perceptually-limited analysis it is usually costly and time-consuming to model the processes of the highly effective ability of humans to identify and recognize a signal. In those applications where analog signals require analysis, the cost of digitizing the analog signal is minimal when compared to the benefits of increased accuracy and speed of signal analysis and monitoring when the processes

contemplated by this invention are utilized.

The present invention relates to identification of digitally-sampled information, such as images, audio and video. Traditional methods of identification and monitoring of those signals do not rely on "perceptual quality," but rather upon a separate and additional signal. Within this application, such signals will be called "additive signals" as they provide information about the original images, audio or video, but such information is in addition to the original signal. One traditional, text-based additive signal is title and author information. The title and author, for example, is information about a book, but it is in addition to the text of the book. If a book is being duplicated digitally, the title and author could provide one means of monitoring the number of times the text is being duplicated, for example, through an Internet download. The present invention, however, is directed to the identification of a digital signal—whether text, audio, or video—using only the digital signal itself and then monitoring the number of times the signal is duplicated. Reliance on an additive signal has many shortcomings. For example, first, someone must incorporate the additive signal within the digital data being transmitted, for example, by concatenation or through an embedding process. Such an additive signal, however, can be easily identified and removed by one who wants to utilize the original signal without paying for its usage. If the original signal itself is used to identify the content, an unauthorized user could not avoid payment of a royalty simply by removing the additive signal—because there is no additive signal to remove. Hence, the present invention avoids a major disadvantage of the prior art.

One such additive signal that may be utilized is a digital watermark—which ideally cannot be removed without perceptually altering the original signal. A watermark may also be used as a monitoring signal (for example, by encoding an identifier that uniquely identifies the original digital signal into which the identifier is being embedded). A digital watermark used for monitoring is also an additive signal, and such a signal may make it difficult for the user who wants to duplicate a signal without paying a royalty—mainly by degrading the perceptual quality of the original signal if the watermark (and hence the additive monitoring signal) is removed.

This is, however, is a different solution to the problem.

The present invention eliminates the need of any additive monitoring signal because the present invention utilizes the underlying content signal as the identifier itself. Nevertheless, the watermark may increase the value of monitoring techniques by increasing the integrity of the embedded data and by indicating tampering of either the original content signal or the monitoring signal. Moreover, the design of a watermarking embedding algorithm is closely related to the perceptibility of noise in any given signal and can represent an ideal subset of the original signal: the watermark bits are an inverse of the signal to the extent that lossy compression schemes, which can be used, for instance, to optimize a watermarking embedding scheme, can yield information about the extent to which a data signal can be compressed while holding steadfast to the design requirement that the compressed signal maintain its perceptual relationship with the original, uncompressed signal. By describing those bits that are candidates for imperceptible embedding of watermark bits, further data reduction may be applied on the candidate watermarks as an example of retaining a logical and perceptible relationship with the original uncompressed signal.

Of course, the present invention may be used in conjunction with watermarking technology (including the use of keys to accomplish secure digital watermarking), but watermarking is not necessary to practice the present invention. Keys for watermarking may have many forms, including: descriptions of the original carrier file formatting, mapping of embedded data (actually imperceptible changes made to the carrier signal and referenced to the predetermined key or key pairs), assisting in establishing the watermark message data integrity (by incorporation of special one way functions in the watermark message data or key), etc. Discussions of these systems in the patents and pending patent applications are incorporated by reference above. The "recognition" of a particular signal or an instance of its transmission, and its monitoring are operations that may be optimized through the use of digital watermark analysis.

A practical difference between the two approaches of using a separate, additive monitoring signal and using the original signal itself as the monitoring signal is control. If a separate signal is used for monitoring, then the originator of the text, audio or video signal being transmitted and the entity doing the monitoring have to agree as to the nature of the separate signal to be used for monitoring—otherwise, the entity doing the monitoring would not know where to look, for what to look, or how to interpret the monitoring signal once it was identified and detected. On the other hand, if the original signal is used itself as a monitoring signal, then no such agreement is necessary. Moreover, a more logical and self-sufficient relationship between the original and its data-reduced abstract enhances the transparency of any resulting monitoring efforts. The entity doing the monitoring is not looking for a separate, additive monitoring system, and further, need not have to interpret the content of the monitoring signal.

Monitoring implementations can be handled by robust watermark techniques (those techniques that are able to survive many signal manipulations but are not inherently "secure" for verification of a carrier signal absent a logically-related watermarking key) and forensic watermark techniques (which enable embedding of watermarks that are not able to survive perceptible alteration of the carrier signal and thus enable detection of tampering with the originally watermarked carrier signal). The techniques have obvious trade-offs between speed, performance and security of the embedded watermark data.

In other disclosures, we suggest improvements and implementations that relate to digital watermarks in particular and embedded signaling in general. A digital watermark may be used to "tag" content in a manner that is not humanly-perceptible, in order to ensure that the human perception of the signal quality is maintained. Watermarking, however, must inherently alter at least one data bit of the original signal to represent a minimal change from the original signal's "unwatermarked state." The changes may affect only a bit, at the very least, or be dependent on information hiding relating to signal characteristics, such as phase information, differences between digitized samples, root mean square (RMS) calculations, z-transform analysis, or similar

signal characteristic category.

There are weaknesses in using digital watermark technology for monitoring purposes. One weakness relates directly to the way in which watermarks are implemented. Often, the persons responsible for encoding and decoding the digital watermark are not the creator of the valuable work to be protected. As such, the creator has no input on the placement of the monitoring signal within the valuable work being protected. Hence, if a user wishing to avoid payment of the royalty can find a way to decode or remove the watermark, or at least the monitoring signal embedded in the watermark, then the unauthorized user may successfully duplicate the signal with impunity. This could occur, for example, if either of the persons responsible for encoding or decoding were to have their security compromised such that the encoding or decoding algorithms were discovered by the unauthorized user.

With the present invention, no such disadvantages exist because the creator need not rely on anyone to insert a monitoring signal—as no such signal is necessary. Instead, the creator's work itself is used as the monitoring signal. Accordingly, the value in the signal will have a strong relationship with its recognizability.

By way of improving methods for efficient monitoring as well as effective confirmation of the identity of a digitally-sampled signal, the present invention describes useful methods for using digital signal processing for benchmarking a novel basis for differencing signals with binary data comparisons. These techniques may be complemented with perceptual techniques, but are intended to leverage the generally decreasing cost of bandwidth and signal processing power in an age of increasing availability and exchange of digitized binary data.

So long as there exist computationally inexpensive ways of identifying an entire signal with some fractional representation or relationship with the original signal, or its perceptually observable representation, we envision methods for faster and more accurate auditing of signals as they are played, distributed or otherwise shared amongst providers (transmitters) and consumers (receivers). The ability to massively compress a signal to its essence—which is not

strictly equivalent to "lossy" or "lossless" compression schemes or perceptual coding techniques, but designed to preserve some underlying "aesthetic quality" of the signal—represents a useful means for signal analysis in a wide variety of applications. The signal analysis, however, must maintain the ability to distinguish the perceptual quality of the signals being compared. For example, a method which analyzed a portion of a song by compressing it to a single line of lyrics fails to maintain the ability to distinguish the perceptual quality of the songs being compared. Specifically, for example, if the song "New York State of Mind" were compressed to the lyrics "I'm in a New York State of Mind," such a compression fails to maintain the ability to distinguish between the various recorded versions of the song, say, for example between Billy Joel's recording and Barbara Streisand's recording. Such a method is, therefore, incapable of providing accurate monitoring of the artist's recordings because it could not determine which of the two artists is deserving of a royalty—unless of course, there is a separate monitoring signal to provide the name of the artist or other information sufficient to distinguish the two versions. The present invention, however, aims to maintain some level of perceptual quality of the signals being compared and would deem such a compression to be excessive.

This analogy can be made clearer if it is understood that there are a large number of approaches to compressing a signal to, say, $1/10{,}000^{th}$ of its original size, not for maintaining its signal quality to ensure computational ease for commercial quality distribution, but to assist in identification, analysis or monitoring of the signal. Most compression is either lossy or lossless and is designed with psychoacoustic or psychovisual parameters. That is to say, the signal is compressed to retain what is "humanly-perceptible." As long as the compression successfully mimics human perception, data space may be saved when the compressed file is compared to the uncompressed or original file. While psychoacoustic and psychovisual compression has some relevance to the present invention, additional data reduction or massive compression is anticipated by the present invention. It is anticipated that the original signal may be compressed to create a realistic or self-similar representation of the original signal, so that the compressed

signal can be referenced at a subsequent time as unique binary data that has computational relevance to the original signal. Depending on the application, general data reduction of the original signal can be as simple as massive compression or may relate to the watermark encoding envelope parameter (those bits which a watermarking encoding algorithm deem as candidate bits for mapping independent data or those bits deemed imperceptible to human senses but detectable to a watermark detection algorithm). In this manner, certain media which are commonly known by signal characteristics, a painting, a song, a TV commercial, a dialect, etc., may be analyzed more accurately, and perhaps, more efficiently than a text-based descriptor of the signal. So long as the sender and receiver agree that the data representation is accurate, even insofar as the data-reduction technique has logical relationships with the perceptibility of the original signal, as they must with commonly agreed to text descriptors, no independent cataloging is necessary.

The present invention generally contemplates a signal recognition system that has at least five elements. The actual number of elements may vary depending on the number of domains in which a signal resides (for example, audio is at least one domain while visual carriers are at least two dimensional). The present invention contemplates that the number of elements will be sufficient to effectively and efficiently meet the demands of various classes of signal recognition. The design of the signal recognition that may be used with data reduction is better understood in the context of the general requirements of a pattern or signal recognition system.

The first element is the reference database, which contains information about a plurality of potential signals that will be monitored. In one form, the reference database would contain digital copies of original works of art as they are recorded by the various artists, for example, contain digital copies of all songs that will be played by a particular radio station. In another form, the reference database would contain not perfect digital copies of original works of art, but digital copies of abstracted works of art, for example, contain digital copies of all songs that have been preprocessed such that the copies represent the perceptual characteristics of the original songs. In another form, the reference database would contain digital copies of processed data

files, which files represent works of art that have been preprocessed in such a fashion as to identify those perceptual differences that can differentiate one version of a work of art from another version of the same work of art, such as two or more versions of the same song, but by different artists. These examples have obvious application to visually communicated works such as images, trademarks or photographs, and video as well.

The second element is the object locator, which is able to segment a portion of a signal being monitored for analysis (i.e., the "monitored signal"). The segmented portion is also referred to as an "object." As such, the signal being monitored may be thought of comprising a set of objects. A song recording, for example, can be thought of as having a multitude of objects. The objects need not be of uniform length, size, or content, but merely be a sample of the signal being monitored. Visually communicated informational signals have related objects; color and size are examples.

The third element is the feature selector, which is able to analyze a selected object and identify perceptual features of the object that can be used to uniquely describe the selected object. Ideally, the feature selector can identify all, or nearly all, of the perceptual qualities of the object that differentiate it from a similarly selected object of other signals. Simply, a feature selector has a direct relationship with the perceptibility of features commonly observed. Counterfeiting is an activity which specifically seeks out features to misrepresent the authenticity of any given object. Highly granular, and arguably successful, counterfeiting is typically sought for objects that are easily recognizable and valuable, for example, currency, stamps, and trademarked or copyrighted works and objects that have value to a body politic.

The fourth element is the comparing device which is able to compare the selected object using the features selected by the feature selector to the plurality of signals in the reference database to identify which of the signals matches the monitored signal. Depending upon how the information of the plurality of signals is stored in the reference database and depending upon the available computational capacity (e.g., speed and efficiency), the exact nature of the comparison

will vary. For example, the comparing device may compare the selected object directly to the signal information stored in the database. Alternatively, the comparing device may need to process the signal information stored in the database using input from the feature selector and then compare the selected object to the processed signal information. Alternatively, the comparing device may need to process the selected object using input from the feature selector and then compare the processed selected object to the signal information. Alternatively, the comparing device may need to process the signal information stored in the database using input from the feature selector, process the selected object using input from the feature selector, and then compare the processed selected object to the processed signal information.

The fifth element is the recorder which records information about the number of times a given signal is analyzed and detected. The recorder may comprise a database which keeps track of the number of times a song, image, or a movie has been played, or may generate a serial output which can be subsequently processed to determine the total number of times various signals have been detected.

Other elements may be added to the system or incorporated into the five elements identified above. For example, an error handler may be incorporated into the comparing device. If the comparing device identifies multiple signals which appear to contain the object being sought for analysis or monitoring, the error handler may offer further processing in order to identify additional qualities or features in the selected object such that only one of the set of captured signals is found to contain the further analyzed selected object that actually conforms with the object thought to have been transmitted or distributed.

Moreover, one or more of the five identified elements may be implemented with software that runs on the same processor, or which uses multiple processors. In addition, the elements may incorporate dynamic approaches that utilize stochastic, heuristic, or experience-based adjustments to refine the signal analysis being conducted within the system, including, for example, the signal analyses being performed within the feature selector and the comparing

device. This additional analyses may be viewed as filters that are designed to meet the expectations of accuracy or speed for any intended application.

Since maintenance of original signal quality is not required by the present invention, increased efficiencies in processing and identification of signals can be achieved. The present invention concerns itself with perceptible relationships only to the extent that efficiencies can be achieved both in accuracy and speed with enabling logical relationships between an original signal and its abstract.

The challenge is to maximize the ability to sufficiently compress a signal to both retain its relationship with the original signal while reducing the data overhead to enable more efficient analysis, archiving and monitoring of these signals. In some cases, data reduction alone will not suffice: the sender and receiver must agree to the accuracy of the recognition. In other cases, agreement will actually depend on a third party who authored or created the signal in question. A digitized signal may have parameters to assist in establishing more accurate identification, for example, a "signal abstract" which naturally, or by agreement with the creator, the copyright owner or other interested parties, can be used to describe the original signal. By utilizing less than the original signal, a computationally inexpensive means of identification can be used. As long as a realistic set of conditions can be arrived at governing the relationship between a signal and its data reduced abstract, increases in effective monitoring and transparency of information data flow across communications channels is likely to result. This feature is significant in that it represents an improvement over how a digitally-sampled signal can be cataloged and identified, though the use of a means that is specifically selected based upon the strengths of a general computing device and the economic needs of a particular market for the digitized information data being monitored. The additional benefit is a more open means to uniformly catalog, analyze, and monitor signals. As well, such benefits can exist for third parties, who have a significant interest in the signal but are not the sender or receiver of said information.

As a general improvement over the art, the present invention incorporates what could best

be described as "computer-acoustic" and "computer-visual" modeling, where the signal abstracts are created using data reduction techniques to determine the smallest amount of data, at least a single bit, which can represent and differentiate two digitized signal representations for a given predefined signal set. Each of such representations must have at least a one bit difference with all other members of the database to differentiate each such representation from the others in the database. The predefined signal set is the object being analyzed. The signal identifier/detector should receive its parameters from a database engine. The engine will identify those characteristics (for example, the differences) that can be used to distinguish one digital signal from all other digital signals that are stored in its collection. For those digital signals or objects which are seemingly identical, excepting that the signal may have different performance or utilization in the newly created object, benefits over additive or text-based identifiers are achieved. Additionally, decisions regarding the success or failure of an accurate detection of any given object may be flexibly implemented or changed to reflect market-based demands of the engine. Appropriate examples are songs or works or art which have been sampled or re-produced by others who are not the original creator.

In some cases, the engine will also consider the NULL case for a generalized item not in its database, or perhaps in situations where data objects may have collisions. For some applications, the NULL case is not necessary, thus making the whole system faster. For instance, databases which have fewer repetitions of objects or those systems which are intended to recognize signals with time constraints or capture all data objects. Greater efficiency in processing a relational database can be obtained because the rules for comparison are selected for the maximum efficiency of the processing hardware and/or software, whether or not the processing is based on psychoacoustic or psychovisual models. The benefits of massive data reduction, flexibility in constructing appropriate signal recognition protocols and incorporation of cryptographic techniques to further add accuracy and confidence in the system are clearly improvements over the art. For example, where the data reduced abstract needs to have further

uniqueness, a hash or signature may be required. And for objects which have further uniqueness requirements, two identical instances of the object could be made unique with cryptographic techniques.

Accuracy in processing and identification may be increased by using one or more of the following fidelity evaluation functions:

1) RMS (root mean square). For example, a RMS function may be used to assist in determining the distance between data based on mathematically determinable Euclidean distance between the beginning and end data points (bits) of a particular signal carrier.

2) Frequency weighted RMS. For example, different weights may be applied to different frequency components of the carrier signal before using RMS. This selective weighting can assist in further distinguishing the distance between beginning and end points of the signal carrier (at a given point in time, described as bandwidth, or the number of total bits that can be transmitted per second) and may be considered to be the mathematical equivalent of passing a carrier signal difference through a data filter and figuring the average power in the output carrier.

3) Absolute error criteria, including particularly the NULL set (described above) The NULL may be utilized in two significant cases: First, in instances where the recognized signal appears to be an identified object which is inaccurately attributed or identified to an object not handled by the database of objects; and second, where a collision of data occurs. For instance, if an artist releases a second performance of a previously recorded song, and the two performances are so similar that their differences are almost imperceptible, then the previously selected criteria may not be able to differentiate the two recordings. Hence, the database must be "recalibrated" to be able to differentiate these two versions. Similarly, if the system identifies not one, but two or more, matches for a particular search, then the database may need

"recalibration" to further differentiate the two objects stored in the database.

4) Cognitive Identification. For example, the present invention may use an experience-based analysis within a recognition engine. Once such analysis may involve mathematically determining a spectral transform or its equivalent of the carrier signal. A spectral transform enables signal processing and should maintain, for certain applications, some cognitive or perceptual relationship with the original analog waveform. As a novel feature to the present invention, additional classes may be subject to humanly-perceptible observation. For instance, an experience-based criteria which relates particularly to the envisioned or perceived accuracy of the data information object as it is used or applied in a particular market, product, or implementation. This may include a short 3 second segment of a commercially available and recognizable song which is used for commercials to enable recognition of the good or service being marketed. The complete song is marketed as a separately valued object from the use of a discrete segment of the song (that may be used for promotion or marketing—for the complete song or for an entirely different good or service). To the extent that an owner of the song in question is able to further enable value through the licensing or agreement for use of a segment of the original signal, cognitive identification is a form of filtering to enable differentiations between different and intended uses of the same or subset of the same signal (object). The implementation relating specifically, as disclosed herein, to the predetermined identification or recognition means and/or any specified relationship with subsequent use of the identification means can be used to create a history as to how often a particular signal is misidentified, which history can then be used to optimize identification of that signal in the future. The difference between use of an excerpt of the song to promote a separate and distinct good or service and use of the excerpt to promote recognition of the song itself (for example, by the artist to sell copies of the

song) relates informationally to a decision based on recognized and approved use of the song. Both the song and applications of the song in its entirety or as a subset are typically based on agreement by the creator and the sender who seeks to utilize the work. Trust in the means for identification, which can be weighted in the present invention (for example, by adjusting bit-addressable information), is an important factor in adjusting the monitoring or recognition features of the object or carrier signal, and by using any misidentification information, (including any experience-based or heuristic information), additional features of the monitored signal can be used to improve the performance of the monitoring system envisioned herein. The issue of central concern with cognitive identification is a greater understanding of the parameters by which any given object is to be analyzed. To the extent that a creator chooses varying and separate application of his object, those applications having a cognitive difference in a signal recognition sense (e.g., the whole or an excerpt), the system contemplated herein includes rules for governing the application of bit-addressable information to increase the accuracy of the database.

5) Finally, the predetermined parameters that are associated with a discrete case for any given object will have a significant impact upon the ability to accurately process and identify the signals. For example, if a song is transmitted over a FM carrier, then one skilled in the art will appreciate that the FM signal has a predetermined bandwidth which is different from the bandwidth of the original recording, and different even from song when played on an AM carrier, and different yet from a song played using an 8-bit Internet broadcast. Recognition of these differences, however, will permit the selection of an identification means which can be optimized for monitoring a FM broadcasted signal. In other words, the discreteness intended by the sender is limited and directed by the fidelity of the transmission means. Objects may be cataloged and assessing with the understanding

that all monitoring will occur using a specific transmission fidelity. For example, a database may be optimized with the understanding that only AM broadcast signals will be monitored. For maximum efficiency, different data bases may be created for different transmission channels, e.g., AM broadcasts, FM broadcasts, Internet broadcasts, etc.

For more information on increasing efficiencies for information systems, see The Mathematical Theory of Communication (1948), by Shannon.

Because bandwidth (which in the digital domain is equated to the total number of bits that can be transmitted in a fixed period of time) is a limited resource which places limitations upon transmission capacity and information coding schemes, the importance of monitoring for information objects transmitted over any given channel must take into consideration the nature and utilization of a given channel. The supply and demand of bandwidth will have a dramatic impact on the transmission, and ultimately, upon the decision to monitor and recognize signals. A discussion of this is found in a co-pending application by the inventor under U.S. Patent Application No. 08/674,726 "Exchange Mechanisms for Digital Information Packages with Bandwidth Securitization, Multichannel Digital Watermarks, and Key Management" (which application is incorporated herein by reference as if fully setforth herein).

If a filter is to be used in connection with the recognition or monitoring engine, it may be desirable for the filter to anticipate and take into consideration the following factors, which affect the economics of the transmission as they relate to triggers for payment and/or relate to events requiring audits of the objects which are being transmitted: 1) time of transmission (i.e., the point in time when the transmission occurred), including whether the transmission is of a live performance); 2) location of transmission (e.g., what channel was used for transmission, which usually determines the associated cost for usage of the transmission channel); 3) the point of origination of the transmission (which may be the same for a signal carrier over many distinct channels); and 4) pre-existence of the information carrier signal (pre-recorded or newly created

information carrier signal, which may require differentiation in certain markets or instances).

In the case of predetermined carrier signals (those which have been recorded and stored for subsequent use), "positional information carrier signals" are contemplated by this invention, namely, perceptual differences between the seemingly "same" information carrier that can be recognized as consumers of information seek different versions or quality levels of the same carrier signal. Perceptual differences exist between a song and its reproduction from a CD, an AM radio, and an Internet broadcast. To the extent that the creator or consumer of the signal can define a difference in any of the four criteria above, means can be derived (and programmed for selectability) to recognize and distinguish these differences. It is, however, quite possible that the ability to monitor carrier signal transmission with these factors will increase the variety and richness of available carrier signals to existing communications channels. The differentiation between an absolute case for transmission of an object, which is a time dependent event, for instance a live or real time broadcast, versus the relative case, which is prerecorded or stored for transmission at a later point in time, creates recognizable differences for signal monitoring.

The monitoring and analysis contemplated by this invention may have a variety of purposes, including, for example, the following: to determine the number of times a song is broadcast on a particular radio broadcast or Internet site; to control security though a voice-activated security system; and to identify associations between a beginner's drawing and those of great artists (for example to draw comparisons between technique, compositions, or color schemes). None of these examples could be achieved with any significant degree of accuracy using a text-based analysis. Additionally, strictly text-based systems fail to fully capture the inherent value of the data recognition or monitoring information itself.

SAMPLE EMBODIMENTS

In order to better appreciate and understand the present invention, the following sample embodiments are provided. These sample embodiments are provided for exemplary purposes

only, and in no way limit the present invention.

SAMPLE EMBODIMENT 1

A database of audio signals (e.g., songs) is stored or maintained by a radio station or Internet streaming company, who may select a subset of the songs are stored so that the subset may be later broadcast to listeners. The subset, for example, may comprise a sufficient number of songs to fill 24 hours of music programming (between 300 or 500 songs). Traditionally, monitoring is accomplished by embedding some identifier into the signal, or affixing the identifier to the signal, for later analysis and determination of royalty payments. Most of the traditional analysis is performed by actual persons who use play lists and other statistical approximations of audio play, including for example, data obtained through the manual (i.e., by persons) monitoring of a statistically significant sample of stations and transmission times so that an extrapolation may be made to a larger number of comparable markets.

The present invention creates a second database from the first database, wherein each of the stored audio signals in the first database is data reduced in a manner that is not likely to reflect the human perceptual quality of the signal, meaning that a significantly data-reduced signal is not likely to be played back and recognized as the original signal. As a result of the data reduction, the size of the second database (as measured in digital terms) is much smaller than the size of the first database, and is determined by the rate of compression. If, for example, if 24 hours worth of audio signals are compressed at a 10,000:1 compression rate, the reduced data could occupy a little more than 1 megabyte of data. With such a large compression rate, the data to be compared and/or analyzed may become computationally small such that computational speed and efficiency are significantly improved.

With greater compression rates, it is anticipated that similarity may exist between the data compressed abstractions of different analog signals (e.g., recordings by two different artists of the same song). The present invention contemplates the use of bit-addressable differences to

distinguish between such cases. In applications where the data to be analyzed has higher value in some predetermined sense, cryptographic protocols, such as a hash or digital signature, can be used to distinguish such close cases.

In a preferred embodiment, the present invention may utilize a centralized database where copies of new recordings may be deposited to ensure that copyright owners, who authorize transmission or use of their recordings by others, can independently verify that the object is correctly monitored. The rules for the creator himself to enter his work would differ from a universally recognized number assigned by an independent authority (say, ISRC, ISBN for recordings and books respectively). Those skilled in the art of algorithmic information theory (AIT) can recognize that it is now possible to describe optimized use of binary data for content and functionality. The differences between objects must relate to decisions made by the user of the data, introducing subjective or cognitive decisions to the design of the contemplated invention as described above. To the extent that objects can have an optimized data size when compared with other objects for any given set of objects, the algorithms for data reduction would have predetermined flexibility directly related to computational efficiency and the set of objects to be monitored. The flexibility in having transparent determination of unique signal abstracts, as opposed to independent third party assignment, is likely to increase confidence in the monitoring effort by the owners of the original signals themselves. The prior art allows for no such transparency to the copyright creators.

SAMPLE EMBODIMENT 2

Another embodiment of the invention relates to visual images, which of course, involve at least two dimensions.

Similar to the goals of a psychoacoustic model, a psychovisual model attempts to represent a visual image with less data, and yet preserve those perceptual qualities that permit a human to recognize the original visual image. Using the very same techniques described above

in connection with an audio signal, signal monitoring of visual images may be implemented.

One such application for monitoring and analyzing visual images involves a desire to find works of other artists that relate to a particular theme. For example, finding paintings of sunsets or sunrises. A traditional approach might involve a textual search involving a database wherein the works of other artists have been described in writing. The present invention, however, involves the scanning of an image involving a sun, compressing the data to its essential characteristics (i.e., those perceptual characteristics related to the sun) and then finding matches in a database of other visual images (stored as compressed or even uncompressed data). By studying the work of other artists using such techniques, a novice, for example, could learn much by comparing the presentations of a common theme by different artists.

Another useful application involving this type of monitoring and analyzing is the identification of photographs of potential suspects whose identity matches the sketch of a police artist.

Note that combinations of the monitoring techniques discussed above can be used for audio-visual monitoring, such as video-transmission by a television station or cable station. The techniques would have to compensate, for example, for a cable station that is broadcasting a audio channel unaccompanied by video.

Other embodiments and uses of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. The specification and examples should be considered exemplary only with the true scope and spirit of the invention indicated by the following claims. As will be easily understood by those of ordinary skill in the art, variations and modifications of each of the disclosed embodiments can be easily made within the scope of this invention as defined by the following claims.

WHAT IS CLAIMED IS:

1. A method for monitoring and analyzing at least one signal comprising:

receiving at least one reference signal to be monitored;

creating an abstract of said at least one reference signal;

storing the abstract of said at least one reference signal in a reference database;

receiving at least one query signal to be analyzed;

creating an abstract of said at least one query signal;

comparing the abstract of said at least one query signal to the abstract of said at least one reference signal to determine if the abstract of said at least one query signal matches the abstract of said at least one reference signal.

2. The method of claim 1 wherein

the step of creating an abstract of said at least one reference signal comprises:

inputting the reference signal to a processor;

creating an abstract of the reference signal using perceptual qualities of the reference signal such that the abstract retains a perceptual relationship to the reference signal from which it is derived; and

the step of creating an abstract of said at least one query signal comprises:

inputting the at least one query signal to the processor;

creating an abstract of the at least one query signal using perceptual qualities of the at least one query signal such that the abstract retains a perceptual relationship to the at least one query signal from which it is derived.

3. The method of claim 1 further comprising:

creating at least one counter corresponding to one of said at least one reference signals, said at least one counter being representative of the number of times a match is found between

the abstract of said at least one query signal and the abstract of said at least one reference signal; and

incrementing the counter corresponding to a particular reference signal when a match is found between an abstract of said at least one query signal and the abstract of the particular reference signal.

4. The method of claim 1 further comprising:

recording an occurrence of a match between the abstract of said at least one query signal and the abstract of said at least one reference signal; and

generating a report that identifies the reference signal whose abstract matched the abstract of said at least one query signal.

5. The method of claim 4, further comprising:

recording an occurrence of a match between the abstract of said at least one query signal and the abstract of said at least one reference signal.

6. The method of claim 1, further comprising permitting access to a secured area when the abstract of said at least one query signal matches the abstract of said at least one reference signal.

7. The method of claim 1, wherein

the step of creating an abstract of said at least one reference signal comprises:

using a portion of said at least one reference signal to create an abstract of said at least one reference signal; and

the step of creating an abstract of said at least one query signal comprises:

using a portion of said at least one query signal to create an abstract of said at

least one query signal.

8. A method for monitoring a plurality of reference signals, comprising:

creating an abstract for each of the plurality of reference signals;

storing each of said abstracts in a reference database;

receiving at least one query signal to be analyzed;

creating an abstract of each of the at least one query signals;

locating an abstract in the reference database that matches the abstract of each at least one query signal; and

recording the identify of the reference signal whose abstract matched the abstract of each at least one query signal.

9. The method of claim 8, wherein

the step of creating an abstract for each of a plurality of reference signals comprises:

inputting each of the plurality of reference signals to a processor;

creating an abstract of each one of the plurality of reference signals using perceptual qualities of each one of a plurality of reference signals such that the abstract retains a perceptual relationship to the reference signal from which it is derived; and

the step of creating an abstract of each of the at least one query signals comprises:

inputting each of the at least one query signals to a processor;

creating an abstract of each one of a plurality of reference signals using perceptual qualities of each one of a plurality of reference signals such that the abstract retains a perceptual relationship to the reference signal from which it is derived

10. The method of claim 8, wherein

the step of creating an abstract of said at least one reference signal comprises:

using a portion of said at least one reference signal to create an abstract of said at least one reference signal;

and the step of creating an abstract of said at least one query signal comprises:

using a portion of said at least one query signal to create an abstract of said at least one query signal.

11. The method of claim 8, further comprising:

creating at least one counter corresponding to one of said plurality of reference signals, said at least one counter being representative of the number of times a match is found between the abstract of said at least one query signal and an abstract of one of said plurality of reference signals; and

incrementing the counter corresponding to a particular reference signal when a match is found between an abstract of said at least one query signal and the abstract of the particular reference signal.

12. The method of claim 8, further comprising permitting access to a secured area when the abstract of said at least one query signal matches an abstract of one of said plurality of reference signals.

13. A computerized system for monitoring and analyzing at least one signal:

a processor that creates an abstract of a signal using selectable criteria;

a first input that receives at least one reference signal to be monitored, said first input being coupled to said processor such that said processor may generate an abstract for each reference signal input to said processor;

a reference database, coupled to said processor, that stores abstracts of each at least one reference signal;

a second input that receives at least one query signal to be analyzed, said second input being coupled to said processor such that said processor may generate an abstract for each query signal;

a comparing device, coupled to said reference database and to said second input, that compares an abstract of said at least one query signal to the abstracts stored in the reference database to determine if the abstract of said at least one query signal matches any of the stored abstracts.

14. The system of claim 13, further comprising:

a storage medium coupled to said first input, that stores each of said at least one reference signals to be monitored; and

a controller coupled to the first input, the processor, the comparing device, the reference database and the storage medium, said controller causing an abstract for each reference signal being input for the first time to be compared to all previously stored abstracts in the reference database, such that in the event that the comparing device determines that it cannot distinguish between the abstract of a reference signal being input for the first time from a previously stored abstract in the reference database, the controller adjusts the criteria being used by the processor and re-generates the reference database, by re-processing each reference signal stored on the storage medium to create new abstracts and storing said new abstracts in the reference database.

15. The system of claim 14, wherein the controller includes a means to adjust compression rates at which the processor processes a signal to create an abstract.

16. The system of claim 13, wherein the comparing device identifies at least two abstracts in the reference database that match the abstract of said at least one query signal and an index

of relatedness to said at least one query signal for each of said at least two matching abstracts.

17. The system of claim 13, further comprising:

a security controller that controls access to a secured area, such that access is granted only if the comparing device confirms that an abstract of said at least one query signal matches an abstract of said at least one reference signal.

18. The system of claim 13, wherein said first input and said second input are the same.

19. The system of claim 13, wherein said second input is remotely coupled to the processor.

20. The system of claim 13, further comprising:

a recorder that records the identify of the reference signal whose abstract matched the abstract of said at least one query signal; and

a report generator that generates a report that identifies the reference signals whose abstracts matched the abstract of said at least one query signal.

21. A electronic system for monitoring and analyzing at least one signal, comprising:

a first input that receives at least one reference signal to be monitored,

a first processor that creates an abstract of each reference signal input to said first processor through said first input;

a second input that receives at least one query signal to be analyzed,

a second processor that creates an abstract of each query signal;

a reference database that stores abstracts of each at least one reference signal;

a comparing device that compares an abstract of said at least one query signal to the

abstracts stored in the reference database to determine if the abstract of said at least one query signal matches any of the stored abstracts.

22. The system of claim 21, wherein said second input is remotely coupled to the system.

23. The system of claim 21, wherein said second processor is remotely coupled to the system.

24. The system of claim 21, wherein the system transmits the criteria that are being used by the first processor to the second processor.

25. The system of claim 21, further comprising:

a storage medium coupled to said first input, that stores each of said at least one reference signals to be monitored; and

a controller that compares an abstract for each reference signal being input for the first time to be compared to all previously stored abstracts in the reference database, such that in the event that the comparing device determines that it cannot distinguish between the abstract of a reference signal being input for the first time from a previously stored abstract in the reference database, the controller adjusts the criteria being used by the processor and re-generates the reference database, by re-processing each reference signal stored on the storage medium to create new abstracts and storing said new abstracts in the reference database.

METHOD AND DEVICE FOR MONITORING AND ANALYZING SIGNALS

ABSTRACT OF THE DISCLOSURE

A method and system for monitoring and analyzing at least one signal are disclosed. An abstract of at least one reference signal is generated and stored in a reference database. An abstract of a query signal to be analyzed is then generated so that the abstract of the query signal can be compared to the abstracts stored in the reference database for a match. The method and system may optionally be used to record information about the query signals, the number of matches recorded, and other useful information about the query signals. Moreover, the method by which abstracts are generated can be programmable based upon selectable criteria. The system can also be programmed with error control software so as to avoid the re-occurrence of a query signal that matches more than one signal stored in the reference database.

EXCHANGE MECHANISMS FOR DIGITAL INFORMATION PACKAGES WITH BANDWIDTH SECURITIZATION, MULTICHANNEL DIGITAL WATERMARKS, AND KEY MANAGEMENT

RELATED APPLICATIONS

This application is related to patent applications entitled "Steganographic Method and Device", Serial No. 08/489,172 filed on June 7, 1995; "Method for Human-Assisted Random Key Generation and Application for Digital Watermark System", Serial No. 08/587,944 filed on January 17, 1996; "Method for Stega-Cipher Protection of Computer Code", Serial No. 08/587,943 filed on January 17, 1996; "Digital Information Commodities Exchange", Serial No. 08/365,454 filed on December 28, 1994, which is a continuation of Serial No. 08/083,593 filed on June 30, 1993; and "Optimization Methods For The Insertion, Protection, and Detection of Digital Watermarks In Digital Data", Serial No. _____, filed on _____.

These related applications are all incorporated herein by reference.

This application is also related to U.S. Patent No. 5,428,606,

"Digital Information Commodities Exchange", issued on June 27, 1995, which is incorporated herein by reference.

BACKGROUND OF THE INVENTION

The present invention relates to digital watermarks.

Digital watermarks exist at a convergence point where creators and publishers of digitized multimedia content demand localized, secured identification and authentication of that content. Because piracy is clearly a disincentive to the digital distribution of copyrighted content, establishment of responsibility for copies and derivative copies of such works is invaluable. It is desirable to tie copyrights, ownership rights, purchaser information or some combination of these and related data into the content in such a manner that the content must undergo damage, and therefore a reduction of its value, in order to remove such data for the purpose of subsequent, unauthorized distribution, commercial or otherwise. Legal precedent or attitudinal shifts recognizing the importance of digital watermarks as a necessary component of commercially-distributed content (audio, video, game, etc.) will further the development of acceptable parameters for the exchange of such content by the various parties engaged in such activities. These may include artists, engineers, studios, INTERNET access providers, publishers, agents, on-line service providers, aggregators of content for some form of electronic delivery, on-line retailers, individuals and other related parties that participate in the

transfer of funds or arbitrate the actual delivery of content to intended recipients.

There are a number of hardware and software approaches that attempt to provide protection of multimedia content, including encryption, cryptographic containers, cryptographic envelopes or "cryptolopes", and trusted systems in general. None of these systems places control of copyrights in the hands of the content creator as content is created. Further, none of these systems provide an economically feasible model for the content to be exchanged with its identification embedded within the signals that comprise the content. Given the existence of over 100 million personal computers and many more noncopyright-protected consumer electronic goods (such as audio clips, still pictures and videos), copyrights are most suitably placed within the digitized signals. Playing content is necessary to determine or "establish" its commercial value. Likewise, advertising and broadcast of samples or complete works reinforces demand for the content by making its existence known to market participants (via radio, television, print media or even the INTERNET).

Generally, encryption and cryptographic containers serve copyright holders as a means to protect data in transit between a publisher or distributor and the purchaser of the data. That is, a method of securing the delivery of copyrighted material from one location to another is performed by using variations of public key cryptography or other cryptosystems. Cryptolopes are

suited specifically for copyrighted text that is time sensitive, such as newspapers, where intellectual property rights and origin are made a permanent part of the file.

The basis for public key cryptography is provided, for example, in a number of patented inventions. Information on public-key cryptosystems can be obtained from U.S. Patent No. 4,200,770 to Hellman et al., U.S. Patent No. 4,218,582 to Hellman et al., U.S. Patent No. 4,405,829 to Riverst et al., and U.S. Patent No. 4,424,414 to Hellman et al. Digitally-sampled copyrighted material is a special case because of its long term value coupled with the ease and perfection in creating copies and transmitting by general purpose computing and telecommunications devices. In this special case of digitally-sampled material, there is no loss of quality in derivative works and no identifiable differences between one copy and any other subsequent copy.

For creators of content, distribution costs may be minimized with electronic transmission of copyrighted works. Unfortunately, seeking some form of informational or commercial return via electronic exchange is ill-advised, absent the establishment of responsibility of specific copies or instances of copies or some form of trusted system in general.

SUMMARY OF THE INVENTION

The present invention allows the establishing of responsibility of specific copies or instances of copies using digital watermarks.

The present invention relates to methods for the management and distribution of digital watermark keys (e.g., private, semiprivate and public) and the extension of information associated with such keys in order to create a mechanism for the securitization of multimedia titles to which the keys apply.

The present invention additionally relates to "distributed" keys to better define rights that are traded between transacting parties in exchanging information or content.

The present invention additionally provides improvements in using digital watermark information. For example, the speed of performing a key search for watermarks within content is increased. Additionally, more than one party can cooperate in adding distinguished watermarks at various stages of distribution without destroying watermarks previously placed in the content.

Digital watermarks make possible more objective commercial exchanges of content. Trusted systems are more costly but achieve the same goal by establishing the identity of all electronic exchange participants. Digital watermark per copy systems, however, are not on a simple level of establishing responsibility of a master work and its derivative copy only. Multichannel watermarks with private, semiprivate and public keys used as different levels of neighboring rights assist in the creation of a self-contained model for the exchange of copyrighted works. Private key watermarks can be inserted into content to establish ownership rights (copyright, master right, etc.) with the content creator or an agent of the content creator maintaining control

over the key. Semiprivate watermark keys can exist in a separate channel of the information signals that make up the work to be exchanged for subsequently delegating responsibility to distributors or sales entities to restrict resale rights in the same manner that physical goods have an exchange of title corresponding to their sale. And finally, public watermark keys exist as an independent component of the identification, authentication or advertising of a given work to be widely distributed over networks for initiating the purchase of a sought-after work. The market will still rely upon trusted parties who report any distribution or exchange of derivative watermarked copies of these "protected" works. Recognition of copyrights as well as the desire to prevent piracy is a fundamental motive of enforcement which uses the mechanism of digital watermarks to alleviate fears of copyright holders and transacting parties that responsibility and payment for copyrights cannot be established and accomplished.

A necessity has arisen for a system that better defines methods for recognizing these rights and, with the further creation of bandwidth rights, as in the present invention, makes possible a distributed model for digital distribution of content which combines the security of a digital watermark system with efficient barter mechanisms for handling the actual delivery of digital goods.

The present invention relates to methods for the management and distribution of digital watermark keys (e.g., private, semiprivate and public) and the extension of information associated with such keys in order to create a

mechanism for the securitization of multimedia titles to which the keys apply. To differentiate the present invention from public key cryptography, use of "private ", "semiprivate", and "public" keys herein refers to the use of such "information" with the stated purpose of distributing goods and watermarking content, not encryption or cryptography in the general sense.

The present invention additionally relates to "distributed" keys to better define rights that are traded between transacting parties in exchanging information or content. Such keys can carry additional pricing and timing information, and represent coupons, warrants or similar financial instruments for purchase of copies of the corresponding title at particular prices within a specified period of time. These instruments, as extended keys, can be collected on servers, distributed to individuals and redeemed as part of a transaction to purchase the content. The basis for this type of content trading system is described in U.S. Patent No. 5,428,606 entitled "Digital Information Commodities Exchange" (hereinafter, also referred to as "the DICE patent"). The present invention improves on the invention described in the DICE patent by integrating into the DICE exchange (i.e., The Digital Information Commodities Exchange) the copyright protection mechanism of digital watermarks. Digital watermarks are described in the following patent applications assigned to The DICE Company: "Steganographic Method and Device", Serial No. 08/489,172; "Method for Stega-Cipher Protection of Computer Code", Serial No. 08/587,943; "Method for Human Assisted

Random Key Generation and Application for Digital Watermark System", Serial No. 08/587,944; and "Optimization Methods for the Insertion, Protection, and Detection of Digital Watermarks in Digitized Data", Serial No. _____.

In addition, the present invention improves upon the techniques of digital watermark systems, described in the patent applications listed above, by adding methods for the use of this information which allow for improvements in the speed of performing a key search for watermarks within content, and by allowing for more than one party to cooperate in adding distinguished watermarks at various stages of distribution without destroying watermarks previously placed in the content. At the same time, these methods minimize the amount of information which any one party must divulge to another party, and prevent "downstream" parties from compromising or otherwise gaining control of watermarks embedded by "upstream" parties.

Further improvements of the present invention include the incorporation of retail models using well-known commodities exchanges to accomplish more efficient means of advertising, negotiating, and delivering digital goods in an anonymous marketplace as commonly characterized by such systems as the INTERNET. Video-on-demand models, quality of service reservations considered in subscriber models, and related models that have been referred to as "time shares" for parceling up processing time in a general computing network will also be differentiated.

DETAILED DESCRIPTION

There are several issues preventing greater volumes of electronic distribution of multimedia content. While such distribution is in fact technically feasible at the present time, attempts at commercially-viable systems are still plagued by these problems, and render digital multimedia exchanges unsatisfactory on a scale comparable to mass retailing in consumer goods markets, such as that of digital audio recordings on compact discs (CDs). While it is possible to transmit a single copy of a digital recording, as 16-bit 44.1 kHz stereo (CD-quality), to an individual from an archive, making such copies available to a large number of paying consumers on demand is still not yet being implemented. The problems fall into several classes, including distribution bandwidth, copyright protection, technological complexities, and "efficient shopping."

In a similar vein to distribution of physical goods in the real world, bandwidth and developments that effectively increase bandwidth are creating profound new business models in how content creators and publishers can distribute their works. From the simplest compression schemes, to actual use of "wired" technology including ISDN, cable modems, ATM and fiber optic lines, the trend is moving toward greater amounts of bandwidth available to on-line users. It is a conundrum of the digital age that the object of bandwidth use will most likely require downloads of copyrighted works, or transaction-based models, to justify such increases in bandwidth availability.

The actual works sought exist as a predefined set of protocols or standards that, when adhered to by hardware or software, can be played back flawlessly many times over. Such works include 74 minute CDs and 300 MB CD-ROMs, among the many physical transport media that now exist. However, the actual digital signals that make up the audio or video clip are not dependent on new playback standards or PC playback software. Simply put, "clips" do not need additional steps to be played back. The signals that a CD carries are not dependent on the CD for its commercial value and could just as easily be carried on a DAT, Minidisc, DVD or any other physical medium that can carry to a consumer audio signals (for example) in a format of 44.1 kHz and 16 bits ("CD quality"). The most apparent drawback is that CDs are not recordable mediums, like cassettes or the above mentioned mediums, so that they are not as economical when coupled with prevalent recording devices such as DAT recorders, PC hard drives, DVD recorders, etc., or when coupled with the advent of electronic lines or "pipes" to the home.

Compression can be both lossless and lossy and has an effect on how a given piece of content can be commercially-valued in the marketplace. Physical goods pricing can be thought of similarly with cassette tapes and CDs which trade at divergent values because of audio quality and degradation, or lack thereof, of such quality over time. Although manufacturing costs of CDs are lower than cassettes, CDs are actually more expensive than cassettes in the marketplace. Presumably a premium is placed on the quality of the stored

content, music or otherwise, and the durability of the medium itself, which can be played without loss of quality far more times than any analog tape. However, the CD is a storage media that must be manufactured, put into inventory, sent by carrier to physical locations, etc., and has an inherent tendency to standardization (the CD is actually a specification determined by manufacturers of both the hardware and software).

Hard costs for marketing and promotion may be better spent across a larger geographical segment, easily accomplished by such electronic networks as the INTERNET but harder to assess in terms of actual sales. Determining market reception is also difficult when buyers are relatively unknown and not available for localized comment or analysis in typical, physical retail store sites (such as Tower Records, Sam Goody's, Blockbuster, etc.).

What equalizes physical mediums such as DAT, CD and DVD, are the lines running between geographic locations, including POTs (i.e., Plain Old Telephone), cable, fiber optic, electric power lines and wireless access points including radio, satellite, cellular phones, and the like. The digitization of these access points and the networks that make them possible ultimately dictate what devices will be appropriate to consumers of the present day and the future. That is, matters of cost and even reputation will increasingly dictate the economics of the distribution of digital content, much the way matters of costs and reputation dictate sales in other consumer goods markets. No longer will it necessarily be important to manufacture X number of copies of a given work

for distribution at N number of sites to capture the optimal market of consumers. The present invention is predicated on not only the existence of a plurality of access points, as discussed in the DICE patent (U.S. Patent No. 5,428,606), but also on a domain where digital content can pass freely between networks much as the INTERNET works with a common protocol (TCP/IP) to facilitate the exchange of data files. However, the ability and desire to orient delivery of digitized content around the specs that describe the content, rather than protocols necessary to redefine the content for exchange over a specific protocol (such as TCP/IP), can better define more convenient delivery of the content between publishers and subscribers given the heterogeneous nature of transmission media (POTs, cable, etc.), the unchanging behavior of "consumer electronically-described" media content (FM-quality, CD-quality, etc.), and the varying configurations of pipes utilized by both publishers and subscribers more concerned with the distribution and exchange of digital goods, not configurations of the immediate input and output devices that are linked by a multitude of electronic exchanges (cable, POTs, wireless, electric power, etc.). Indeed, shifting only the recordable media cost to consumers that, for the most part, already own one or more such devices and may have exposure to a number of broadcast and advertising media (INTERNET, on-line services, radio, cable, print, etc.) may afford both buyers and sellers the cheapest means of profitably exchanging digital goods.

At present, over 15% of the U.S. population has more than one phone

line, 60 million households have cable television, and 15 million consumers are on-line subscribers. ISDN is also experiencing growing demand in the U.S. to give consumers higher bandwidth in the interim. Projected increases of bandwidth portend future supply and demand of larger data files of copyrighted passive works (e.g., music, pictures, video, etc.) and interactive works (e.g., games, software, etc.), thus putting pressure on the need for increases of bandwidth. Never before has increased available bandwidth suffered from a lack of demand by users. In other words, new bandwidth seems to create its own demand. Much of the presumption in increased investments in creating the bandwidth has been to enable the transfer of audio, video, and multimedia files that typically occupy more than 5 MB of space per file. The misanalyzed aspect of these investment plans is a method for addressing digital piracy of copyrighted works and efficient, market-based allocation of the subsequent bandwidth by users. The present invention better defines maximized operations dependent more on the specs that describe playback of content than redefining additional protocols which add additional and unnecessary levels to the playback of the content. With such advances, exchanging media content can potentially be made as easy as exchanging physical content.

The present invention additionally reduces costs in the distribution process, provides the monitoring of, and thus ability to protect, copyrights within the media, and allows the implementation of better payment systems suited to the distribution of digital goods. What is clear is that bandwidth may

never be unlimited, but with consideration made to real world economics, efficient and realistic methods for considering "fill rate" (the actual titles "delivered" to a purchaser versus the titles "ordered"), speed (actual time it takes for a consumer to receive desired content), and cost (expense given trade-offs of immediate availability at a given price point to the consumer, e.g., immediate fulfilment equates to higher pricing, versus delayed delivery of the same content at a lower price) all represent input variables in a real world "retail experience" that may be replicated in the digital domain. The present invention takes into consideration the behavior of parties engaged in selling content that may not be initially valued at the same price by all market participants and is subject to the same promotion hype as goods in the real world. In the digital domain, sampling, trailers, and pre-release hype can be replicated to foster demand for a given title of a digital good with many of the same results that are experienced in the real world.

Evidence of supposedly more efficient schemes for retail include U.S. Patent No. 4,528,643 to Freeny, which shifts much of the manufacturing costs to physical retail sites, thus increasing the cost of doing business on the retail side with possible increases of convenience to the consumer. In the Freeny patent, retailers are envisioned to have localized reproduction of given digitized products (music, video, etc.) and a means to use "owner authorization codes" to verify the electronic transmission of a given work from some "master file unit" to recordable media (VCR, recordable CD, etc.). Freeny refers to mail order

clubs and other direct marketing efforts as being inefficient versus the localized manufacturing structure. These predictions have since been proven false. It is because of the nebulous concept of intellectual property coupled with the extreme expense on retailers for the in-store manufacturing units that makes clear the benefit of leveraging available bandwidth to content creators, publishers, consumers and "pipe owners." The efficiency of such operations as Federal Express in delivering even small packages in under 24 hours and the ability of "fulfilment houses" to effectively carry all but the most obscure titles (music, books, videos, etc.) has made actual "manufacturing" of a given physical media object (CD, VHS tape, etc.) or what Freeny describes as a "material object" simply uneconomical and increasingly irrelevant in an age when bandwidth and digital recording devices such as PCs, Minidiscs, digital video disks (DVD), etc. make physical retail-based, or in-store, copying more of an inconvenience.

The paradox of digital copies is the ease and relatively inexpensive operation of making perfect copies from a single instance of a work, thus providing the potential of unauthorized copies or piracy. The binary data that comprises a digitized work is an approximation of an analog signal. As is well known binary ones and zeros can be manipulated to form words, audio, pictures, video, etc. Manners in which individual copies can be marked so that responsibility can be assigned to individual copies that are derivatives of the master copy is documented in the patent applications by The DICE Company

referenced above (i.e., U.S. Patent No. 5,428,606, and the "Steganographicc Method and Device", "Method for Human-Assisted Random Key Generation and Application for Digital Watermark System", "Method for Stega-Cipher Protection of Computer Code", "Digital Information Commodities Exchange" and "Optimization Methods for the Insertion, Protection, and Detection of Digital Watermarks In Digital Data" applications), and in alternative proposals by Digimarc Corporation (a form of pseudo-randomly encoding digital signatures into images), Bolt Beranek & Newman (Preuss et al. patent, U.S. Patent No. 5,319,735) (embedded signaling) and others. Additional proposals for cryptolopes and cryptographic containers by IBM and Electronic Publishing Resources (EPR) place control of copyrights and other "rights" in the control of IBM and EPR, not the individual content creator or publisher. IBM and EPR are creating a form of "trusted systems." What is clear is that trusted systems, where all parties are known in some way to establish responsibility for instances of copied files, are not realistically possible with the number and ease of manufacture of digitization systems such as general purpose computing devices. At present, over 100 million such devices are in existence, and it is not possible to guarantee that all of these systems will be made to adhere to the defined parameters of a trusted machine for verification and the establishment of responsibility for individual copies made of digital works. Profit motives continue to exist for individuals to make perfect copies and distribute these copies without paying the parties responsible for creating and

distributing the content. Moreover, beyond considerations of digital exchanges that do establish responsibility for the goods being sought, the digital bits that comprise the commercially-valuable works suffer both from lack of use by parties seeking more secured means of distributing and marking content, and legal tanglings by parties that own the copyrights and seek any entity deemed to copy works illicitly for settlement of disputes. That is, with the great number of untrusted systems in existence, many copyright holders have resorted to legal challenges of on-line services and individuals found to be in possession of unauthorized copies of copyrighted works. The resultant digital marketplace tends to favor larger companies who can afford to seek legal settlements without delivering any substantial benefit over smaller companies that for many reasons would otherwise favor digital distribution of content to minimize overall costs. The remedy for such problems is addressed in the previously discussed related U.S. patent and patent applications by The DICE Company and other parties mentioned above (e.g., NEC, Digimarc, EPR, IBM, etc.)

The present invention relates to methods for parceling rights to benefit buyers and sellers of digital works in ways that even the playing field of the marketplace given the resource of electronic marketplaces that can work with such networks as the INTERNET. Too often physical world solutions are offered where digital domain considerations are completely ignored.

Another issue relating to the present invention involves haphazard grafting of physical world pricing and automated payment systems onto digital

systems. Issues of inventory, physical movement, and manufacture of goods are completely muted in digital exchanges, but are replaced by bandwidth utilization and efficiency, one-to-one connections, and one-to-many connections, i.e., seeking and reaching customers in an anonymous marketplace. It is these issues that will better determine the price of a given digital good. Timing of the good (that is, live versus broadcast rerelease of the same digital good) and the necessity of filters or brokers which guide individuals to acceptable goods are variables that will play roles in determining the ultimate efficiency of exchanging digital goods.

Among some of the proposed systems are a proposal by Wave Systems, which necessitates the use of proprietary boxes using encryption to tie the user's "exchange device " to some party that can determine the validity of the box, a trusted system. Unfortunately, adoption of such a solution would necessitate the purchase of separate boxes for separate vendors of particular works or the routing of all digital goods through a proprietary system that then resembles closed cable, video-on-demand, and private networks. Similar approaches are used by merchants using credit card processors and the use of credit card authorization devices and paying incremental costs for the use and security delivered by the credit card processor. Further systems include log-in procedures to validate the accessing party's identification. The premium paid for such systems is arguably excessive when compared to content creator-controlled implementation of digital watermarks and an exchange by

which all distribution parties are engaged in the marketplace to pay for bandwidth rights to market-test given digital goods. The only alternative available to smaller content creators and artists is to sell content at no charge, thus jeopardizing potential future returns, or purchasing outright the hardware to plug-in to existing networks, an excessive cost if such "bandwidth" could be more fairly-priced in a need-based system such as that discussed in this disclosure.

As an improvement to the system discussed in U.S. Patent No. 5,428,606, the present invention ties so-called "header" files into the actual content. U.S. Patent No. 5,428,606 addresses the separation of content from its references ("header") to facilitate more efficient access and exchange of digital content. The "headers" described in this patent might be construed in the real world as options or futures, and is discussed below. The present invention concerns itself with creating a method for introducing a layer of price and distribution determination given the necessity of payment in delivering digital content between points in the digital domain which may not suffer from any physical limitations but are limited by bandwidth considerations.

Some attempts at the exchange of content are being tried with existing networks such as the INTERNET. The complexities extant are apparent in the requirements of the operating protocols and the dependence of TCP/IP for orienting content and subsequently playing it back through "players" that are TCP/IP compliant, if the INTERNET is solely considered. More issues

regarding the INTERNET are further discussed below.

Conceptually, "agents" partially meet some of the expectations of a content-based system, except agents are also dependent on participation by sites willing to allow for pure price comparisons and later reporting to the purchasing party. At present, many sites lock out such agents as they seek to profit by value-added services which are not considered by an agent when "shopping prices." Video-on-demand systems also propose a more closed system that is reliant on a proprietary network to deliver a video (or audio for that matter) to a consumer with the least amount of time delay while satisfying the demands for the video by many other consumers seeking the same video at the same time. The difference between such a system and that disclosed in the present invention is that such video-on-demand networks propose "subscriber" models where all consumers are deemed to have the same right to a given, demanded, piece of content at any time. That is, all participants are "subscribers" who prepay a fee structure that cannot necessarily be justified given bandwidth and processing limitations for delivering digital goods "on demand." In such a system, infrastructure cost can run as high as 5,000 dollars per subscriber, as with Time Warner's system in Orlando, Florida.

In the present invention, time is not an absolute standard to measure satisfaction. In the same manner that retail stores cannot always have a given audio or video work "on demand," other factors may play into the competitiveness of that entity to contribute to the satisfaction of a given

consumer. These issues include a depth (number of copies or copyrights of a given title) or breadth (number) of titles offered, a variety of delivery mediums to satisfy customers with varying access infrastructure (cable, telephone, fiber optic, electric power, etc.), pricing, and, finally, service as it can be applied in an anonymous marketplace. Services may include the know-how of buyers employed by a given digital broker in offering samples of new releases or unknown artists, as well as special price offers given the amount and types of digital goods being purchased. What is certain is that a "subscriber" model is subject to the same deficiencies of a cable model or proprietary on-line service that may not be able to balance financial considerations with the variety and cost of titles sought by individuals at any given time. On the seller side, maximizing profit per title cannot always be satisfied if distribution control or proprietary rights are granted to any single entity which, by the present nature of the INTERNET and future interpretations of on-line commerce, cannot be guaranteed. Indeed, the above-mentioned U.S. Patent No. 5,428,606 discusses a situation where all subscribers can be publishers. For smaller parties, naturally lacking sufficient resources to initially and adequately market and promote titles, a more open system for negotiating distribution rights must be sought by commoditizing the good that most effects exchange of their goods in the digital domain (i.e., bandwidth).

Moreover, in an anonymous marketplace, even small aggregators of content may be able to adequately promote the digital properties of other small

content creators with value-added services. These services, such as samples of content, used to entice buyers, just as trailers create demand for upcoming movies, could be delivered to a differing type of subscriber, much as the music aficionados who subscribe to College Music Journal (CMJ) and other resources to sample new, relatively uncommercial music. Samples of 10-30 seconds could be sent directly to consumer e-mail addresses replicating the prevalent listening bars set up by physical music retailers seeking to introduce new titles to eager listeners. Other services might be more representative of "music chat rooms" or special title web-sites, to more fully entice potential buyers with a greater amount of purchase information. Much of the premise of such services and fulfilling demand for content, however, will require a more efficient means to allocate bandwidth according to an embodiment of the present invention. Without such bandwidth allocation, even small digital goods vendors will need to purchase substantial hardware, from T1 lines to high-powered UNIX machines, meaning high entry or fixed costs, to effectively market what may only be a single title in a year.

The present invention deals with commoditization of the digital distribution of multimedia content. It is important to note that in creating such a market, one must consider two commodities. One is the title, or data itself, of which there is a theoretical unlimited supply over time (limited only by how many copies of a given title that can be made). The second commodity is bandwidth. This is a commodity which must be treated more like traditional

commodities, since its supply is physically limited over discrete periods of time "Fatter" pipes and compression can only increase upper limits given the observed tendency for larger data files to accompany bandwidth increases in the short term. In practice, bandwidth limits act as a parameter on the capacity of a distribution channel at any given moment in time, since there is a fixed amount of bandwidth. In dealing with commercial markets, where, for example, 80% of the consumers want 20% of the products, (and for digital marketplaces, generally all at the same instant), some premium can be observed as with "first come first serve" principles in physical sales channels. The difference is that an additional copy of a digital work can be made almost instantaneously, although additional bandwidth cannot be replicated. Even in instances with theoretically infinite time to fill all orders, most buyers will have given up and "left" the exchange after waiting a short period, during which time they get no satisfaction, measured explicitly by an access or download of a specifically desired title. On-line services today are typically plagued by this shortfall, leading most users to complaints of access and speed. Market-based principles could alleviate some of this problem on both the buyer and seller side if bandwidth is treated as the commodity it is. "Quality-of-service" proposals partially address this issue, though costs are stacked on the seller side because such systems are almost always proprietary given the requirement of high infrastructure expenses to enable timely delivery to all subscribers to the "private" network.

The present invention combines "efficient shopping" principles with the commoditization of bandwidth and titles to create an exchange, under principles as described in the DICE patent, where in place of a security, one can buy titles where a component of the title price is actually a bandwidth option, or bandwidth right. The purchaser buys a right on the underlying title to take delivery of the title via a particular transport medium which uses a particular allocation of transmission bandwidth at a particular time. According to an additional embodiment of the present invention, distributor or content aggregator-only purchases of bandwidth are stipulated as options for digital distribution increase, in terms of available channels (such as cable, satellite, etc.). In this case, the end user never deals with the bandwidth right, although the costs of such rights may by passed on in the retail price of the title which is purchased and downloaded. In other words, the distributor must purchase rights in advance to support a projected volume level of distribution. These pre-purchased rights are then attached to individual downloads. These instruments can vary in price, much like stock options, based on time. Only, in this case, it is the amount of time required to receive the underlying security, which implicitly indicates how much bandwidth will be used by the buyer. The bandwidth actually implies time. The spectrum could range from lowest bandwidth, such as an e-mail delivery by POTs lines, which uses bandwidth when it is otherwise not in use and is at the convenience of the seller (sender), and not the buyer (receiver), to highest bandwidth that may be parallel or direct

access fiber optic line which may be necessary for users acting as wholesalers between electronically -linked parties who seek content for negotiated delivery.

U.S. Patent No. 5,428,606 uses the concept of a "DIP" ("digital information packet") header to create an advertising, distribution, and pricing device which allows for the dissemination of references to and description of particular titles available electronically. The DICE Company's related digital watermark patent and patent applications as discussed previously disclose an exchange model for digitally-watermarked content and digital watermark keys whereby keys which allow a party to scan or imprint watermarks are distributed, possibly electronically, at the discretion of the controlling party. Both these methods have in common the fact that they allow tor the distribution of some information related to an underlying work, without distributing the work itself. It is in the interest of simplicity, therefore, to allow for the combination or conjunction of these information items in addition to associating them with a bandwidth right or option for the downloading of the copyrighted work.

Essentially, some of this negotiation of bandwidth takes place between the "Baby Bells" and AT&T or other long distance providers when settling rights-of-way between points of a telephone conversation. At present, a key difference is that the utility value of a phone call sets the value of the "phone time" being sold. Bandwidth rights as envisioned in an embodiment of the present invention price the commodity of bandwidth given the luxury item being sought (i.e., data or content). The present invention seeks to value the

immediacy as well as convenience (of which price may play a role) in receiving a given packet of data (media content, software, etc.) from one or many locations where it may be available to other locations. The lines may be heterogeneous between points, thus offering a more open bidding system between line owners, content creators and publishers, and end users or consumers. At present, no such "negotiation" can be handled by network operators running lines to the same home or office. Indeed, lines are usually charged at a fixed fee, not by what amount they are used. In some cases, lines are billed by a raw measure of the data transferred, but not in relation to the actual value of such data nor with respect to the value of other transfers which might occur simultaneously via the same line. This sort of billing-by-byte tends to discourage use, but it is a very coarse tool with which to manage utilization. To fill the middle market for demand of these lines for telecommunications lines in particular, long distance carriers such as AT&T, MCI and Sprint sell excess capacity to "wholesalers," while the larger companies generally have price constraints.

The potential demand for bandwidth is clearly evident with such widespread use of networks, epitomized by the INTERNET. But, as previously discussed, smaller, specialist "retailers" and "wholesalers" of services or content that could be marketed over these lines are not efficient. The potential for efficient pricing exists as demonstrated by "call-back" services, which route calls from one location through a third party location, benefitting from that

location's line pricing, though the overall market for such services is still only about $300 million annually. What restricts more open allocation of bandwidth is political in nature. At the same time, cross subsidization of local phone access from more expensive long distance and international service is open for rationalization envisioned by the present invention. Even if more network services could offer greater returns for line use, and thus bandwidth use, public telephony accounts for over 85% of the market. A particular model being evaluated is called "sender takes all" where the access point, or the party that provides access to an end user, would take all the access charges. This is similar to the INTERNET, but is still stacked against smaller players, of which content providers are the least favored if they seek "distribution channels" over networks that still lack proper market incentives for use of bandwidth. Some other models being considered include a single access charge, which is an improvement over current international accounting standards being negotiated between countries. Still, this model does not take into consideration the available bandwidth controlled by non-telecommunications parties, such as cable companies, though ultimately the commodity being brokered is actually common bandwidth. The uneasy balance in negotiating access is being tempered by the steady increase by telecommunications companies to upgrade their lines to offer comparable bandwidth access as that presently available through cable companies. A final issue for consideration is the mobile market of cellular phones and other similar technologies though there

are far more restrictions on the amount of available bandwidth for content distribution, the move to free up more radio spectrum for digital signals may lead to increases as high as a hundredfold in the capacity of the network which would make the electronic delivery of a single audio track realistic. Still, the present invention seeks the imposition of market-based pricing of available bandwidth to end users and content providers given the absence of any such system currently.

With the recent removal of barriers which previously prevented competition between cable companies, telecommunications companies, and regional Bell operating companies (RBOCs) the matter of cost of services or content being delivered over common pipes and the concept of a single entity dominating the "network" will almost surely come to an end as many companies are strongly positioned in their local markets. At present, "local loop" access to end users still presents formidable barriers to competition-- 40-45% of the cost of a long distance call is paid to the RBOC whose lines run into the home or business making the call. In total, the cost to a network for local distribution is approximately 80%. Proposals for separating a network into its infrastructure and service components would likely benefit from the invention being outlined. In such a scenario, the owner of the network would offer access to providers on the same terms, while managing the operation of the infrastructure. Simple models, such as flat rate INTERNET access, are problematic in the overall model for market-based pricing of bandwidth in that

capital costs are completely ignored though such costs are the parameter by which any business model must be judged. Though the cost of an extra phone call over a given network may be negligible, the cost of pumping large multimedia files, which have far different utility value to users of the network versus a "telephone conversation, " is relatively high in the aggregate and can be witnessed with the progressively slow performance of many on-line providers and the INTERNET. The goal for network providers will be to offer value-added services to users as well as value-added access to content that is controlled by copyright holders seeking maximum distribution (given speed and quality) to content seekers. These parties may only need the network at certain times or for certain releases of content. Meanwhile, periphery services such as music sampling, game testing, beta software distribution, will most likely comprise value-added services beyond the present scope of strict telephony. The pressure, generated from capital cost concerns, to provide a system that prices speed and line capacity is aptly answered with the creation of bandwidth rights and incorporation of such rights into the electronic distribution of content. In this way, specialist companies will strive through buying bandwidth of transmission capacity and adding value by attracting customers seeking said companies' accessible content.

Bandwidth rights are necessary as an improvement over the art. The INTERNET currently dominates any discussion of digital distribution. The INTERNET is built over lines or pipes. It is an important observation that a)

these pipes cost money to build, deploy and maintain, and b) the owners of the pipes must pay for their investment and earn some return, which is their motivation for building the infrastructure. The means by which files are transferred over the World Wide Web, the most mainstream segment of the INTERNET, is the use and interpretation of Hypertext Mark-up Language (HTML) and embedded URLs (Uniform Resource Locators) which is designed to "alias" and designate a single path between the party that is viewing a reference of a file and the underlying file. The user is unnecessarily "connected" to the actual file, which is called "aliasing," and has effectively created more network traffic and thus wasted bandwidth. This shortfall in HTML is affecting the INTERNET through inefficiencies resultant from the underlying connection-based TCP/IP protocol. In short, a lot of needless, bandwidth-wasting connections are continuously being created and destroyed. The current mechanics of the INTERNET will not be conducive to electronic commerce, and must necessarily change. This fundamental aspect of splitting content from references to that content is amply addressed in U.S. Patent No. 5,428,606.

The biggest problem can be summed up by observing that users of the INTERNET generally live under the misconception that data or content is, or should be, free. Although one can find specific instances of goods and services sold over the INTERNET, even downloadable software, the basic mechanism that underlies the sale is subject to this "fallacy of the free." There

are actually many hidden costs, some of which were discussed above. As for the content creator or publisher of said works, monitoring of sites and legal enforcement of copyrights is still significantly difficult without better education of consumers and site administrators, as well as a means for detecting unauthorized copies on an archive as disclosed in the digital watermark filings. Recent legal actions against parties that distribute copyrighted music titles and game software has resulted in setting a "for price " trend that can be made more efficient by the present invention.

The present invention deals with creating a coherent pricing model for on-line distribution, which accounts for bandwidth utilization, maximizes pricing options and efficiency for sellers and buyers, and, additionally, as a result of the process of trading and pricing of the bandwidth options, ensures that usage of the limited bandwidth is orderly. All orders result from requests filled and thus are generally a function of the price of the so -called option on bandwidth. The present invention also presents improvements over exchanges that exist for the purpose of trading commodities such as stocks, bonds and other such securities. The distinctive feature of the preferred embodiment described below is the nature of the commodities being traded, bandwidth, and the unbounded potential of derivative copies of copyrighted works.

In current trading mechanisms NASDAQ (National Association of Securities Dealers Automated Quote system) is a well-known model. Looking at details of the NASDAQ market will illuminate exchange operations and the

present invention's improvements over the present art for both market exchange mechanisms and implementations of a content-based system that monitors copyrights and optimizes the distribution of the underlying content.

The NASDAQ Market

NASDAQ is an exchange that trades in a finite number of "titles" or stock certificates, whereas the present invention is concerned with the potential of an infinite number of "titles" made up of digital bits-- each derivative copy having the same potential commercial value as the original master copy that was intended for trade. The limited or finite commodity in question on a DICE exchange is available bandwidth for the actual transmission and thus delivery of a demanded, digitized "piece" of content (audio clip, picture, video, virtual reality, software, etc.). Bandwidth is characterized by the pipes that connect buyers and sellers of digital information and include POTs, cable, fiber optic, ISDN, satellite, electric power lines, etc. On the other hand, NASDAQ deals with basic stock securities, publicly-traded shares in companies. There are a small number of derivative securities traded, notably warrants, but the mechanisms for supporting a particular security are fairly uniform. NASDAQ is primarily an electronic bulletin board where market makers advertise at what prices they are willing to buy and sell a particular security. These market makers maintain an inventory of tradeable securities for sale to other parties, whether agency or principal-based transactions. A market maker does not

necessarily equal a broker, although a market maker can also be a broker. Both market makers and brokers can participate in the system, but market makers are the heart of it. A market maker is a paying member of the NASD (National Association of Securities Dealers). In effect, they own a stake in the market governing body, and agree to be obligated to buy or sell a certain minimal amount of shares, in order to provide liquidity in the market "Confidence" in the market mechanism, that is NASDAQ itself, is in the best interests of the participants or the ultimate buyers of securities will not be willing to bid on securities at uncompetitive prices. Similarly, an artist wishing to sell their commercially-valuable copyrighted content, must be relatively confident that each derivative, a perfect digital copy, has some mechanism for identifying the initial purchaser and give all subsequent market participants a way of ensuring the copy of the content they possess is not an illicit or unauthorized copy. Previously discussed disclosures on digital watermarks cover these issues as a means to bring more artists and publishers into the digital marketplace to increase activity and liquidity.

Like the "specialists" on the NYSE (New York Stock Exchange), NASDAQ market makers earn a profit on the spread between the BUY and SELL price of a stock, assuming they can buy low and sell high (or short high and buy low). Market makers risk their own capital, trading a group of stocks, and can generally make profits trading shares for incremental profits. Such an instance would be selling at 10 and buying at 9 7/8. Many market makers

trade the same stocks competitively, and in general, the more firms that make a market in a given stock, the more liquid the trading of that stock is, simply because there are more ready buyers and sellers. Again as a means to describe the present invention some understanding of these market participants may be required in implementing the proposed system.

Although NASDAQ can be thought of as an "electronic" market, it is electronic, for the most part, only in the sense that instead of shouting across a floor at each other, traders generally advertise their price levels on a BBS (Bulletin Board System), which legally binds them to honor the price. They then field phone calls from traders at other member firms, who have seen the advertisements on the BBS, and agree to trades over the phone. Then, each side enters their transaction (if one side is a BUY, the other is a SELL) into on -site computers, which all feed into central mainframes and link up with each other. Many errors are introduced by this process, and an error report is produced at the end of the day, to be settled among the parties involved through after-hours reporting. So, there is really still a large low-tech component to NASDAQ which leads to discrepancies and inefficiencies.

The general public interacts with the market through brokers, who might also happen to work for a member firm. The chain of contact is individual to broker to trader, with traders interacting among each other, and filling orders for brokers. This also touches the issues of primary and secondary markets. When a stock goes public, called an IPO (Initial Public Offering), shares are

bought up by a syndicate of market makers. This is the primary market. The proceeds of the IPO go to the issuing company, minus the underwriting fees, which are divided among the syndicate. The syndicate then sells shares to the public through brokers, and any other traders who want to trade them. The syndicate may profit again by selling the shares at higher prices than the original purchase price. This trading continues indefinitely or until bankruptcy. This is the secondary market. Prices in the secondary market can vary continuously and widely from the price set in the primary market.

Having summarized the system, we can discuss some of the inefficiencies and idiosyncrasies of NASDAQ to establish the parameters of the present invention in the preferred embodiment

One major problem is the uniform distribution of information. Theoretically, all traders should get the same information at the same time. However, NASDAQ does not accomplish this well. Since there are intermediate "concentrators" between the terminals and the hub, and specific terminals tend to watch specific groups of stocks, some of which may be significantly more active than others, generating a larger volume of information per second, which can cause back-ups, in general, the system is plagued by delays of an intermittent and non-uniformly distributed nature. There is no mechanism for detecting these problems, which may cause the display of old or incorrect prices for some stocks, and delay the dissemination of electronic orders on an unequal basis. Traders generally have several sources of

information, and need to be "on their feet", so the burden of detection is, in effect, placed on humans. NASDAQ terminals do maintain a "heartbeat." If the terminal cannot get a response from the hub for a prescribed period of time, a problem is signaled by turning the screen a uniform yellow on black. However, most significant information delays do not trip this mechanism. Market makers have cooperated to run independent tests, and are well aware that one trader may see information up to several minutes before another. There is no aging of information. The present invention partially concerns itself with information aging as content can be time-sensitive, and up-to-date bandwidth rights pricing is important. Such instances include news reports, live broadcasts, initial "be first" demand for a particular piece of media content, and the like.

A NASDAQ hub may send out information to all routes simultaneously, but there can be large delays before it arrives at the destination. An example of a timing performance protocol, which can be employed to counter such problems, is NTP (Network Time Protocol) on UNIX networks. NTP does advanced diagnosis of point-to-point network performance to forecast timing delays between pairs of machines. It is used with time critical applications, but not widely so, as it is still considered quite esoteric. NASDAQ makes no use of such protocols. For more trustworthy information about bandwidth rights and the aging of a media content good, the present invention takes into account forecasted timing delays for pricing the subsequent bandwidth right as an

overall component of the pricing of the media content being demanded, and delays in actually distributing this information. This is an improvement over the art as it is a more appropriate aspect of pricing media versus disseminating stock price information.

Before considering the present invention's clearing operations, which are vital to simplifying the otherwise tremendous task of figuring out who owes what to whom at the end of the day, a description of the art, a la NASDAQ, is required. Basically, clearing is the matching up of trades. If one side reports a SELL, and the other a BUY, these two sides must be put together to form a trade which results in the transfer of money to the seller, and the transfer of the security to a buyer. Any halves of trades that do not match are kicked back to the member firm who entered them, for resolution. Provided the trade is resolved, both sides again enter their sides, only late/ The securities can be held in street name, meaning the brokerage house can hold the physical shares for the buyer. However, the task of transferring stock certificates and cash among brokerage houses is onerous. Instead, a special holding organization was created. This organization is independent of the stock exchanges, but works with their clearing computers. The holding organization maintains vaults filled with stock certificates, held for the brokerage, which in turn hold the stock in the names of their clients. Everyone maintains records of who owns what relative to their own organization. Should an owner actually request their certificates, they can be removed from the vault and delivered by

way of the brokerage firm. At the end of a day's trading, the hub computers at each exchange (whether NASDAQ or NYSE) net out the differences among the member firms, in cash and stock, over many trades, and produce a report of who owes what to who, in net terms, relative to each stock. The firms have a certain number of days to settle the trades (which allows for correction of errors, and transfer of funds). This allows a single day to result in one transaction for each trading firm for each stock it trades. This sort of clearing is key to the efficiency of any trading system. With the exception of a certificate delivery request, no security certificates need be moved, and cash can be transferred by wire.

Defining the Value of Bandwidth Rights

It is an object of this invention to create a trading instrument which will break bandwidth resources into discrete, usable component pieces, and allow an electronic market system to set a price for this scarce commodity which sets an equilibrium level of supply and demand. The net effect of this instrument, and its trading system, will be to efficiently apportion bandwidth to users who wish to download or upload valuable information, in whatever form it takes. Bandwidth affects the speed of information transfer. If more bandwidth is used, speed increases, and the transfer is accomplished in less time. If an individual instance of this instrument is a bandwidth right, it can be observed that several factors will affect its value:

· Intrinsic Value

This value is measured versus a minimal standard telecommunications cost. If there is a single underlying telecommunications cost to the owner of the right of X dollars per minute, let min 0 represent the number of minutes it takes to download the information using the minimal bandwidth, and min 1 represent the number of minutes a to transfer the information at the bandwidth represented by this right. Note that min 0 >= min1.

Then the intrinsic value VI = X x (min0 - min1), or the amount of money saved in telecom costs at the higher bandwidth. The intrinsic value can be negative, which would imply a compensating premium placed on the time saved by using the more expensive transport.

· Percentage Chance of Failure

This probability recognizes the generally unreliable nature of the current telecommunications and transmission mediums as well as underlying computer systems. Rather than be burdened with the task of solving all of the "bugs" in a given piece of commercial software, it would be better to account for failure in the valuation. This value could be adjusted over time, as the failure probability of a system becomes more apparent, or changes. In short, this represents the percentage chance a user cannot exercise their right. It affects the expected value of the right. In this baseline approach, if the probability of failure is Pf,

where 0 <= Pf <= 1, and the value of the right is V0, in the absence of failure, then Vf = (I-Pf)V0.

· Convenience Premium

This represents some premium, VC that a person is willing to pay to transfer their information within a specified period of time (i.e. "now" or "in the next 10 minutes"). This premium is likely to come out as the market sets the price for a right. If there is a formula for what the price should be, then the premium is simply the difference between the result of that formula, and the actual market price. This really measures the balance between supply and demand. The more demand in excess of supply, the higher C will rise. VC is then a function of supply and demand.

Vreal = Vtheoretical + VC

·Time Value

This is a function of the exercise period of the bandwidth right. It is proportional to Pf, since more time allows for recovery from an individual failure to transfer. There are two components of time, over what period a transfer can be initiated and for how long the transfer can last once it is initiated. Note that this is made more complex by congestion factors. For instance, if a user has a right for 10,000 kbps for 10 seconds, and the user wants to transfer 100,000 kb, it is not likely that the transfer can be done in exactly 10 seconds. Protocol

overhead and congestion will add some increment of time. It is advisable to leave room in the exercise period for these factors, rather than trying to value the time value in some manner which accounts for these transient conditions. Thus:

$$V = (I\text{-}Pf)(VI + VT + VC)$$

or $V = (1 - Pf)\,((X(min0\text{-}min1) + VT) + VC$

The convenience premium, VC, should be independent of all other values (except V).

The equation behaves as such:

With increased failure probability decreasing rights value, independent of other variables, while increased demand relative to supply would drive up VC. We might try to compute VC by accounting for known demand and supply values, and in fact, it is of vital importance to know the supply, and to allocate it so that any right issued can be exercised within its exercise period.

Additionally, it is observed that a method is needed to allocate supply based on demand which accounts for unused rights. In other words, the system needs to over allocate supply to some degree, knowing that some rights may go unexercised, so that demand is filled as much as possible. This is similar to airlines' practice of overbooking flights.

Some mechanism must be in place to prevent attacks on the system, by a party, who, in effect, tries to corner the market in bandwidth, with no intention

of using it, so that it goes unused. Naively, one would think that since one has to pay for the bandwidth, why would someone want to corner the market? Although bandwidth is not free, it should only comprise a small fraction of the value of the information to be transferred, and so this is not an unthinkable situation. The likeliest preventive measure is the existence of competition in transmission.

Another option is the potential need to necessitate a secondary market for the trading of bandwidth, which could be divided up by a trading syndicate, and traded on a secondary basis to users. In a manner of operations, telecommunications companies perform this role between national telecommunications systems to facilitate international phone usage. But the difference with the system envisioned in the present system is that "any" user could buy bandwidth rights at times of low demand, and hope to sell them at a profit in times of higher demand. This would seem to imply the exchange itself should do some proprietary trading in this manner, both to profit, and to ensure some bandwidth is available for sale to users when they need it. This will have a purpose to serve in making the market efficient in the future.

Bandwidth rights instruments are likely to be highly localized to specific subnets. Especially since certain types of connections may be available only from certain exchanges, and since failure probabilities are likely to vary with specific hardware, operating systems, and service providers. Additionally, the basic valuation equations above do not address telecommunications costs

across various types of lines. This problem at least, might be solved by active maintenance of cost tables, designation codes for types of lines, and the designation of a low cost standard. The problem of moving rights between exchanges is made more difficult since supply/demand planning for one exchange will not translate to another, unless some means for interconnecting exchanges is developed, and exchange bandwidth planning is global. The race by many parties to link users to the INTERNET via varying access links (modem) including ISDN, POTs, cable, may further the need for common bandwidth pricing. What is clear is that the basic structure of the present invention would facilitate such planning to the benefit of all market participants: telecoms providers, INTERNET access companies, users and publishers as well as more general aggregators of content and bandwidth such as, phone companies, cable companies and satellite companies intending on providing services across multifarious line types.

Bandwidth Rights Accounting and Clearing

If a bandwidth right is securitized, the creation and supply of certificates, made unique by cryptographic methods to manage them, will also be necessary. Transferring certificates between individuals is complicated and unnecessary. Following the general principles of the securities clearing model described above seems to be in order. In this case, the exchange needs to create and manage an account for each party that can own or trade bandwidth

rights. Additionally, a method for authenticating the party is required. With these two elements, a trading market can be implemented by the following methods:

The exchange creates and manages a supply of uniquely distinguished bandwidth rights certificates. These certificates are good for a specific period only. They may traded over the course of time, anywhere from the moment they are created to the expiration time. It is questionable whether a right should be exercisable once it is clear that even if a transfer is initiated, it cannot be completed given that right only. However, consider that the right is usable, but its value decreases rapidly as it approaches expiration (i.e. value is based on time left, not total transfer time). Once a certificate is expired it is deleted. Hash values incorporating a time-stamp could be used to serialize certificates. Such a cryptographic method is well noted in the art. US Pat No 5,136,646 and 5,136,647 ("Digital Document Time-Stamping With Catenate Certificate" and "Method For Secure Time-Stamping Of Digital Documents" respectively) describe methods for cryptographic time-stamping.

The exchange creates a central hub for planning bandwidth supply, accounting, and disseminating pricing information. Client-side software will value the rights relative to a particular user's needs, and used by any party trading rights. A seller creates a SELL advertisement, which is entered into the "exchange". The exchange verifies that the seller actually holds the right in their account. A buyer then enters a BUY offer against the sell advertisement.

The exchange validates the buyers, and then clears the transaction, transferring money from the buyer's payment method (credit card, etc.) to the seller's account, and the right to the buyer's account. The unbundled right may be so infinitesimal that the actual cost of the right must be bundled with the underlying content or information being sought. The rights could also be bound to underlying titles. This may be similar to attaching sales taxes, handling charges, and credit card use charges that are typically bundled with the cost of a given physical goods purchase.

Multichannel Watermarking Mechanisms and Techniques

One problem with previous digital watermark systems is the need for a mechanism by which multiple parties may add watermarks to a given piece of content at different stages of distribution, without requiring any one party to compromise the security of its watermarks to any other party. Although an "exchange" system allows for two-way communication, a particular "distribution path" may be taken to be the path by which a package of data travels from a source party to a destination party. So, a distribution may be a single side of an "exchange". In this context, it is useful to speak of parties to the distribution as "upstream" or "downstream" in relation to each other. The initial source would be farthest upstream, while the ultimate destination party would be farthest downstream, with any number of parties along points in the middle. If the data in a distribution flows from party A,

through party B, to party C, then:

party A is upstream from parties B and C;

party B is downstream from party A, but upstream from party C;

and party C is downstream from parties A and B.

The above example should make clear the relationships between upstream and downstream parties.

It is a useful goal, and an accomplishment of embodiments of the present invention, to provide a mechanism and technique for the purpose of allowing any party to the distribution to add at least one channel of watermark information, which exists separately and is secured by means of a separate key, to the data of the distribution in such a manner as to ensure that one or more watermarks of the other parties to the distribution remain present in the data when it reaches its final destination.

A significant improvement over traditional metering systems is that exchange mechanisms are beneficially tied into content for more realistic metering of playing or recording content. With multichannel digital watermarks, a more robust means for metering content is made possible by parties not willing to create expensive proprietary distribution channels, but who do wish to capitalize on selling content in the economic method of metering. There are two immediately apparent schemes which might accomplish this. The first is described as a "passive" scheme and the second is described as an "active" scheme.

In a passive scheme, several assumptions must be decided and jointly agreed upon beforehand by all parties who wish to add watermarks. Based upon the total number of watermark channels to be used, where each party that wants to add a watermark is assumed to use at least one watermark channel, and the amount of data, and the desired minimal level of watermark security, a watermark system could encode watermarks at an appropriate sparsity such that random chance will cause some watermarks added by downstream parties to obliterate watermarks added by upstream parties. But by the same token, random chance will allow some of the watermarks of upstream parties to survive the encoding of watermarks by downstream parties by virtue of the fact that such watermarks do not occupy enough of the same data space to cause one to significantly interfere with the reading of another. The end result is that at least one watermark added by each party will be readable at the final destination. While such a passive scheme is appealing because of its relative simplicity, in which each party can add watermarks without considering the impact of any other party, once some initial parameters are set, this type of scheme requires a lot of testing to determine optimal settings given various initial conditions, and does not guarantee any particular level of watermark redundancy. It is quite haphazard, although technically feasible.

According to an advantageous embodiment of the present invention, an

active scheme is implemented which is described as follows. The farthest party upstream, who presumably controls the ultimate copyrights and distribution rights of the data generates two keys. The first key is a regular watermark key, as described in previous related patent application disclosures by The DICE Company, particularly, including the "Method for Stega-Cipher Protection of Computer Code" application. This key is used for actual encoding and decoding of information from the watermark channel "owned" by this party. The second key is a new type of watermark key, called a master framework key, which dictates

how the entire data stream in general is to be packetized;

how the data stream packets are to be allocated among a predetermined number of reserved watermark channels; and

how the channels are to be assigned to downstream parties.

This information is the minimal amount of information which must be shared with downstream parties to enable them to add watermarks using their own regular watermark keys to their assigned channels. Notice that within a given channel, another key is still needed to extract a watermark. Therefore, while some information is potentially leaked, the watermarks are still secure. The master framework key, in effect, creates several virtual data streams within the real data stream, each of which can be accessed separately by the watermark system. The master framework key can then be shared on a limited or protected basis with only those downstream parties who the upstream party

chooses to participate in the distribution. Such master keys could be distributed using well-known cryptographic art for key transmission. Each downstream party is responsible for generating their own regular watermark key, and watermarking their assigned channel with appropriately generated information using the combination of the master framework key and the regular watermark key, as the data is received and forwarded. This active scheme is much better than the passive scheme, since it ensures that watermarks added by downstream parties do not interfere in any way with those added by upstream parties, thus guaranteeing a maximal level of watermark redundancy, which is desirable, while minimizing the disclosure of watermark information necessary to downstream parties, which is undesirable. It is envisioned that systems that use a hybrid approach, incorporating some mechanisms and methods of the active scheme, but also relying on some methods of the passive scheme may be developed.

Keysearch Optimization Mechanisms and Techniques

Another issue of digital watermark system which must be adequately addressed is key search. When a suspect copy of content is obtained, the amount of work done to extract watermark information from the copy is bounded by the set of watermark keys which are potential candidates which may have been used to encode the hypothetical watermark(s) in the suspect data. It is an object of the invention described herein to minimize the amount of

work and hence time required to search this set of keys, or keyspace, while ensuring confidence that all potential candidate keys have been searched, or at least those candidates with a significant probability of constituting the actual target of the search.

The watermark decode operation proceeds generally as follows: First a candidate key search group is generated, then a decode process is run using each candidate key until either all keys are exhausted and no watermark is extracted, or a watermark is extracted using a candidate key. Depending on the nature of the information in the extracted watermark, the search might continue with remaining keys, or terminate. One obvious method for improvement is to perform parallel searches trying multiple keys at the same time. Using powerful parallel hardware, real gains may be obtained using this method simply.

On slower, serial CPU-based hardware, real parallel gains are more difficult to make. However, using dynamic programming techniques and intelligent search scoring and management, one could configure the search engine to start with several or all keys, checking each packet of data against each key before proceeding. As each iteration is completed, factoring in the next data packet, cumulative "scores" for the results of each key may be computed and compared. Keys which appear to have more potential to ultimately yield a match and extract a watermark continue to be used in the

process, while those with lower potential, as measured by score, are dropped from the process. This process has an attractive characteristic that it gets faster as more keys are progressively eliminated from the search space, and can consider a large number of keys. Its drawback, in the absence of other techniques, is that the initial key space may be very large, and it may take considerable time to narrow the search keys to the point where the search proceeds at a reasonably fast pace. It is also possible that the process of finding a match does not score in a monotonically increasing manner, resulting in the early elimination of the correct key. In other words, scores may get worse before they get better.

Without considering any information about the source copy used to generate the suspect copy, one could limit the search work done by imposing a limit on how much time a decoder can spend checking data versus a particular key, or a maximal percentage, or number of packets of the copy to process before giving up on a given key. One could do well with a heuristic rule that says, "if I have checked 50% of the recording without finding a watermark, then in all likelihood I will not find a watermark in the other 50% of the recording with this particular key," for instance. However, the best gains can be made by eliminating as many keys as possible from the initial search pool. In order to do this the keys are expanded to included several items of information regarding the source copy or master that was watermarked using the key in question. This information includes any of the following items:

Title, Artist, Date, size of recording, format of the recording, quality of the recording;

and may also include mathematically calculated properties of the recording which can identify the recording to some significant degree of probability while using only a small amount of data (i.e. localized hash values, etc.). When a suspect copy is obtained, this same set of information describing the suspect copy is generated by the decoder system, which can then select a set of candidate keys which match to a desired degree, any or all the criteria stored with the keys.

Finally, the best potential results may be obtained by taking advantage of the multiple access levels made possible by the watermark system described in previous filings. A watermark embedded in a higher privacy channel corresponds with a particular key. Every key has a unique identification which allows the key custodian to find the key in a database, but provides no information on the key itself. This identification may have no meaning outside the custodial system. If the higher privacy key identification is included in a lower privacy watermark such as a protected or public watermark, then the party searching for the higher privacy watermark makes use of an intentionally limited set of lower privacy keys to first extract the key identification of the higher privacy key. At this point, no additional key search is necessary, thus allowing significant time savings. This assumes the lower

privacy watermark has not somehow been removed from the digital sample stream.

An embodiment of the decoder key search system encodes private key identifiers in lower privacy watermarks and uses descriptive information in the keys to compare versus the suspect copy to narrow the key search space. This embodiment makes use of parallel hardware to facilitate as much gain as possible from parallel search techniques described above, including progressive elimination of keys which appear to diverge from a match as the comparison progresses.

In an exchange mechanism according to an embodiment of the present invention, the exchange is not the source of any of the sought-after works or digital information packages (DIPs). The exchange is ultimately measured by available transmission resources. Whereas DIPs are measured in a digitization system, the size of the underlying data file, its file structure, which dictates any potential compression and buffering, and data overhead for error correction, will provide exchange participants with an estimate for the resources, including time required to distribute said DIP. Given the heterogeneous nature of existing and proposed line infrastructure, any DIP can potentially be exchanged over vastly different lines between points. These may include copper, coaxial, fiber optic, etc. Distribution of a given DIP may occur on different lines for the same work (say for instances of a work available over POTs and satellite, etc.) or over a number of different media in the distribution

of a work as it is transmitted over a network with a plurality of transmission media (say, the backbone of the network may be fiber but the end loop is coax, etc.). Given the existence of other traffic over these lines, including telephony, the pricing of a given DIP should necessarily include the price of the bandwidth resources necessary to transfer the DIP between at least two parties. As previously discussed, the difference in this embodiment and systems such as video-on-demand or proprietary cable and satellite systems is the necessity to value bandwidth between points in a network to facilitate the exchange of a demanded work at a given instant in time not continuously as with traditional "subscriber models." Similarly, "time-share" systems are oriented around selling a parcel of time to users seeking "processor" access to perform some activity, while, bandwidth is not the commodity being bid, time shares are reservation systems not capable of bidirectional or end-to-end "negotiation" of resources to facilitate the exchange of a DIP in real or next-to-real time. Further, the preferred embodiment differs in that all participants may have significantly different access infrastructure (differing modems, cable, electric powerline, satellite, etc.) and pricing preferences given demand for a particular DIP.

The price of the bandwidth resources is, thus, proportional to the percentage of bandwidth allocated to the transfer of the DIP and inversely proportional to the duration of the transfer. With these factors, the aggregate of available bandwidth must change with time and can appropriately be priced

given the demand of certain DIPs or publishers seeking to effectively distribute DIPs. Bandwidth allocation can then be securitized to reflect the varying needs of market participants to exchange DIPs. How this security is priced relates to the nature of the underlying DIP which is most likely a luxury item such as a musical recording or video game. The securities must then trade independently of the DIPs and are based in part on a convenience premium, given demand for bandwidth allocation at any given time. Additionally, network resources as measured by present digital packet switches provide the variable of "supply of bandwidth resources" and estimated demand for said resources at a given time. For networks that are more centralized, such as cable or satellite, estimating bandwidth resources may actually be far easier as traffic is generally downstream to customers not bidirectional like telephone networks. Further means for computing bandwidth securitization instruments take into consideration probability of failure to exercise an instrument, the time period for which said instrument is valid, intrinsic value relative to minimum standard bandwidth utilization for the line in question. These factors, when coupled with a convenience premium, are improvements over the prior art as described in the U.S. Patent No. 5,428,606. Bidirectional exchange of content by parties who can be both subscribers or publishers or both, are possible when the party wishing to sell content or DIPs can set distribution, pricing, and other informational fields at its discretion. These issues are well documented in U.S. Patent No. 5,428,606 and are increasingly important in the growing popularity

of the World Wide Web (WWW) portion of the INTERNET. But, given that the marketplace in which digital goods can be traded digitally is itself digital, the evident or potential scarcity of bandwidth or the ability to value existing bandwidth given a commercial market for digital goods exchange is invaluable.

Further, security of the content and records of said content can be further described as an improvement over methods to undeniably identify content through the use of digital watermarks and other similar technologies. It is desirable to take appropriate measures to protect as many parties as possible in the transaction of a copyrighted work. These parties may include the copyright holder, publisher, distributor, retailer, and consumer. As with the physical monitoring of media products such as CDs, where physical checks are conducted by the label, manufacturer, distributor, retailer and even outside parties such as SoundScan, Billboard, etc. the digital domain contains far less means for "hands-on" metering without including watermarks as "secured identification" for parties involved in the distribution chain. As a preferred embodiment of the present invention, a record of a given DIP should include at least two of any of the following three elements: a digital watermark key, a DIP header, and a bandwidth securitization instrument (bandwidth right). The DIP header describes the content, its address, pricing, and distribution. The bandwidth right is unique in its instance but also varies according to network bandwidth availability for a given period of time and the duration of the actual use of bandwidth on said network.

Optimizing key searches and increased use of multichannel digital watermarks are delineated in the discussions that follow this preferred embodiment as they are additional improvements over the art. The embodiment thus far discussed makes possible a more "democratically" or "economically" feasible market for the exchange of digital goods. With bandwidth rights, multichannel watermarking, optimized key searches, content-base metering, it will be possible to more fully replicate retail and wholesale environments as they exist in the physical world. Decisions about depth and breadth of services and goods that can be offered by on-line market participants will differ only in the ability to offer access to archives (POTs, cable, satellite, wireless, etc.) which will be determined by pricing and speed of transmission as well as by content providers interested in tapping into the potential distribution market that the pipe owner's network includes. Market participants will also be able to appeal to the anonymous parties that seek content through attractiveness of a "site, " amount of processing speed available for distributing digital goods, staff responsible for purchasing or creating available content for downloads, the number of available repurchase rights of copyrighted works: "electronic window-shopping" can be realized given heterogeneous networks, many digital goods, and the creation of bandwidth rights to complement digital watermarking systems. Simply, content can better be valued given the infrastructure of the digital domain while recognizing/he importance of tracking and monitoring the exchange of digital

goods.

WHAT IS CLAIMED IS:

1. A method of pricing on-line distribution of digital information packages comprising determining an on-line distribution net price based on a price of bandwidth resources necessary to transfer the digital information package between at least two parties and based on an underlying price of the digital information package itself.

2. The method according to claim 1, wherein the price of bandwidth resources is proportional to a percentage of bandwith allocated to transfer of the digital information, and is indirectly inversely proportional to a duration of the transfer.

3. A method of creating a bandwidth securitization instrument comprising valuing bandwidth allocation as a scarce commodity.

4. A method of valuing a price and a convenience premium of bandwidth securitization instruments by facilitating an electronic market for free trading of said bandwidth securitization instruments independently of any particular digital information packages ultimately transferred using said bandwidth.

5. A method of computing a convenience premium, comprising steps of:

determining a supply of bandwidth resources;

determining a plurality of bandwidth securitization instruments which allocate the bandwidth resources; and

determining an estimated demand at a given moment in time for the bandwidth resources.

6. A method of computing a price for a bandwidth securitization security instrument as a function of its intrinsic value relative to a minimum standard bandwidth utilization, comprising steps of:

a) obtaining a minimum standard price;

b) determining an estimated convenience premium of the bandwidth securitization security instrument with respect to said minimum standard price;

c) determining a probability of failure to effect an exercise of the security;

d) determining an exercise period of the security instrument corresponding to a time during which it may be executed or redeemed; and

e) determining a price for the bandwidth securitization security instrument based on said steps a), b), c), and d).

7. A method of combining into one record, at least two of:

a digital watermark key,

a digital information packet (DIP) header, and

a bandwidth securitization instrument (Bandwidth Right);

wherein the DIP header contains information including content description, content addressing and content pricing;

wherein a bandwidth securitization instrument may be incorporated by including a serialization identification code which is unique to an individual bandwidth right, where record of said right may exist separately from the record containing the serialization identification code;

wherein the bandwidth securitization instrument is a unique security which values the right to use a specific allocation of telecommunications bandwidth for a specific duration, where such right exists for a specified period of time, and where the duration begins at or after the temporal issuance of the security, and the exercise period ends contemporaneously with the termination of the duration period.

8. The method according to claim 7, wherein the bandwidth securitization instrument provides a right to use a given bandwidth allocation for a net duration over the exercise period where the net duration may be comprised of smaller sub-durations which are not necessarily temporally contiguous.

9. A method for optimizing key search operations comprising steps of:

associating content descriptive information with a key used to watermark content for candidate keys;

comparing the content descriptive information from each candidate key in a key;

searching against a suspect copy of a title, and using said comparison to eliminate keys which are evaluated as unlikely based on the matching criteria of the content descriptive information;

wherein criteria includes at least one of:

media format;

content length;

content title;

content author; and

content signal metrics which provide heuristic characterizations of the recorded signal.

10. A method for performing multi-party, multi-channel encoding of watermarks comprising generating a master framework key, wherein the master framework key describes packetization and channel allocation of a complete signal.

11. The method according claim 10, further comprising a step of:

distributing the master key and a channel assignment to each party who needs to watermark a channel described in the master key.

12. The method according to claim 11, further comprising a step of limiting distribution of the master key only to parties who need to add watermarks to the signal.

13. The method according to claim 12, further comprising a step performed at least one stage thereafter of:

generating a general watermark key, for use with the master key which dictates watermarking of packets assigned to a single channel of the master key watermarking said packets with said key.

14. A method of including a key identifier for a distinct watermark channel in the watermark contained in an additional separate and distinct watermark channel in the same digital sample stream, which is encoded and decoded with its own distinct key.

15. The method according to claim 14 further comprising a step of:

including the key identifier of a higher privacy watermark channel in the watermark contained in a lower privacy watermark channel for a purpose of expediting watermark search operations.

ABSTRACT OF THE DISCLOSURE

Responsibility can be established for specific copies or instances of copies of digitized multimedia content using digital watermarks. Management and distribution of digital watermark keys (e.g., private, semiprivate and public) and the extension of information associated with such keys is implemented to create a mechanism for the securitization of multimedia titles to which the keys apply. Bandwidth rights can be created to provide for a distributed model for digital distribution of content which combines the security of a digital watermark with efficient barter mechanisms for handling the actual delivery of digital goods. Distributed keys better define rights that are traded between transacting parties in exchanging information or content. More than one party can cooperate in adding distinguished watermarks at various stages of distribution without destroying watermarks previously placed in the content. Additionally, the amount of information which any one party must divulge to another party can be minimized, and "downstream" parties can be prevented from compromising or otherwise gaining control of watermarks embedded by "upstream" parties.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/08159

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :H04B 13/00; H04J 3/26; H04L 12/40
US CL :370/60, 85.11, 85.11; 375/260

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 370/32, 53, 54, 58.1, 58.2, 60, 60.1, 61, 62, 85.1, 85.11, 94.1; 375/257, 260, 267; 348/6, 7, 8, 10, 12, 16; 379/110, 219, 220

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 4,491,983, (PINNOW et al) 01 January 1985, col. 3, lines 22-45, col. 4, lines 16-33, col. 4, line 44 to col. 5, line 20.	1-7, 18-20, 26-27 and 30
Y	US, A, 4,958,341 (HEMMADY et al) 18 September 1990, col. 6, lines 4-59 and figure 2.	1-7, 18-20, 26-27 and 30

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be part of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 SEPTEMBER 1995	17 NOV 1995
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer HUY D. VU B. Harder for Telephone No. (703) 308-6602

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/08159

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Telephone Practice

I. Claims 1-7, 18-20, 26-27 and 30, drawn to an apparatus for exchanging information packets between plurality of modular expandable units over two transmission media. (375/260)

II. Claims 8-17, drawn to a method for publishing directory entries and publisher address. (375/260)

III. Claims 21-25, 28-29 and 31, drawn to a bus transmission system having a data bus and a separate control bus. (370/85.11)

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet(1))(July 1992)*

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04L 9/00 — A3

(11) International Publication Number: WO 96/42151

(43) International Publication Date: 27 December 1996 (27.12.96)

(21) International Application Number: PCT/US96/10257

(22) International Filing Date: 7 June 1996 (07.06.96)

(30) Priority Data:
08/489,172 9 June 1995 (09.06.95) US

(71) Applicant: THE DICE COMPANY [US/US]; P.O. Box 60471, Palo Alto, CA 94306-0471 (US).

(72) Inventors: COOPERMAN, Marc, S.; 2929 Ramona, Palo Alto, CA 94306 (US). MOSKOWITZ, Scott, A.; Townhouse 4, 20191 East Country Club Drive, North Miami Beach, FL 33180 (US).

(74) Agents: ALTMILLER, John, C. et al.; Kenyon & Kenyon, 1025 Connecticut Avenue, N.W., Washington, DC 20036 (US).

(81) Designated States: CA, CN, FI, JP, KR, SG, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
13 February 1997 (13.02.97)

(54) Title: STEGANOGRAPHIC METHOD AND DEVICE

(57) Abstract

An apparatus and method for encoding and decoding additional information into a stream of digitized samples in an integral manner. The information is encoded using special keys. The information is contained in the samples, not prepended or appended to the sample stream. The method makes it extremely difficult to find the information in the samples if the proper keys are not possessed by the decoder. The method does not cause a significant degradation to the sample stream. The method is used to establish ownership of copyrighted digital multimedia content and provide a disincentive to piracy of such material.