Multimedia Watermarking Techniques

FRANK HARTUNG, STUDENT MEMBER, IEEE, AND MARTIN KUTTER

Invited Paper

Multimedia watermarking technology has evolved very quickly during the last few years. A digital watermark is information that is imperceptibly and robustly embedded in the host data such that it cannot be removed. A watermark typically contains information about the origin, status, or recipient of the host data. In this tutorial paper, the requirements and applications for watermarking are reviewed. Applications include copyright protection, data monitoring, and data tracking. The basic concepts of watermarking systems are outlined and illustrated with proposed watermarking methods for images, video, audio, text documents, and other media. Robustness and security aspects are discussed in detail. Finally, a few remarks are made about the state of the art and possible future developments in watermarking technology.

Keywords— *Audio, image, multimedia, review, video, watermarking.*

I. INTRODUCTION

Multimedia production and distribution, as we see it today, is all digital, from the authoring tools of content providers to the receivers. The advantages of digital processing and distribution, like noise-free transmission, software instead of hardware processing, and improved reconfigurability of systems, are all well known and obvious. Not so obvious are the disadvantages of digital media distribution. For example, from the viewpoint of media producers and content providers, the possibility for unlimited copying of digital data without loss of fidelity is undesirable because it may cause considerable financial loss. Digital copy protection or copy prevention mechanisms are only of limited value because access to cleartext versions of protected data must at least be granted to paying recipients which can then produce and distribute illegal copies. Technical attempts to prevent copying have in reality always been circumvented.

One remaining method for the protection of intellectual property rights (IPR) is the embedding of digital watermarks into multimedia data. The watermark is a digital code unremovably, robustly, and imperceptibly embedded in the host data and typically contains information about origin, status, and/or destination of the data. Although not directly used for copy protection, it can at least help identifying source and destination of multimedia data and, as a "last line of defense," enable appropriate follow-up actions in case of suspected copyright violations.

While copyright protection is the most prominent application of watermarking techniques, others exist, including data authentication by means of fragile watermarks which are impaired or destroyed by manipulations, embedded transmission of value added services within multimedia data, and embedded data labeling for other purposes than copyright protection, such as data monitoring and tracking. An example for a data-monitoring system is the automatic registration and monitoring of broadcasted radio programs such that royalties are automatically paid to the IPR owners of the broadcast data.

The development of watermarking methods involves several design tradeoffs. Watermarks should be robust against standard data manipulations, including digital-to-analog conversion and digital format conversion. Security is a special concern, and watermarks should resist even attempted attacks by knowledgeable individuals. On the other hand, watermarks should be imperceptible and convey as much information as possible. In general, watermark embedding and retrieval should have low complexity because for various applications, real-time watermarking is desirable. All of these (partly contradicting) requirements and the resulting design constraints will be discussed in more detail throughout the paper.

The paper is organized as follows. Section II gives an introductory explanation of the terms used, as well as a few remarks about the historical aspects of watermarking. In Section III, common design requirements and principles are explained that apply to all watermarking techniques, independent of the actual application. Sections IV–VII review various watermarking techniques that have been proposed for formatted text data, images, video, and audio, respectively. Watermarking of other media, including three dimensional (3-D) data and 3-D animation parameters, is discussed in Section VIII. Section IX gives detailed insight

Manuscript received October 20, 1997; revised March 26, 1998.

F. Hartung was with the Telecommunications Laboratory, University of Erlangen–Nuremberg, 91058 Erlangen, Germany. He is now with Ericsson Eurolab, Research Department, 52134 Herzogenrath, Germany.

M. Kutter is with Signal Processing Laboratory, Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland.

Publisher Item Identifier S 0018-9219(99)05174-9.

into security issues, namely attacks against watermarks, and shows the relations between watermarking and cryptology. In Section X, we extrapolate the recent development of watermarking technology and watermarking applications and try to forecast future trends. Section XI summarizes and concludes this paper on multimedia watermarking techniques.

II. Steganography and Watermarking—History and Terminology

A. History

The idea to communicate secretly is as old as communication itself. First stories, which can be interpreted as early records of covert communication, appear in the old Greek literature, for example, in Homer's *Iliad*, or in tales by Herodotus. The word "steganography," which is still in use today, derives from the Greek language and means covert communication. Kobayashi [67] and Petitcolas *et al*. [99] have investigated the history of covert communication in great detail, including the broad use of techniques for secret and covert communication before and during the two World Wars, and steganographic methods for analog signals. Although the historical background is very interesting, we do not cover it here in detail. Please refer to [67] and [99] for an in-depth investigation of historic aspects.

Paper watermarks appeared in the art of handmade papermaking nearly 700 years ago. The oldest watermarked paper found in archives dates back to 1292 and has its origin in Fabriano, Italy, which is considered the birthplace of watermarks. At the end of the thirteenth century, about 40 paper mills were sharing the paper marked in Fabriano and producing paper with different format, quality, and price. They produced raw, coarse paper which was smoothed and postprocessed by artisans and sold by merchants. Competition not only among the paper mills but also among the artisans and merchants was very high, and it was difficult to keep track of paper provenance and thus format and quality identification. The introduction of watermarks helped avoiding any possibility of confusion. After their invention, watermarks quickly spread over Italy and then over Europe, and although originally used to indicate the paper brand or paper mill, they later served as indication for paper format, quality, and strength and were also used to date and authenticate paper. A nice example illustrating the legal power of watermarks is a case in 1887 in France called "Des Decorations" [41]. The watermarks of two letters, presented as pieces of evidence, proved that the letters had been predated and resulted in considerable sensation and, in the end, in the resignation of President Grévy. For more information on paper watermarks, watermark history, and related legal issues, please refer to [144], an extensive listing of over 500 references.

The analogy between paper watermarks, steganography, and digital watermarking is obvious, and in fact, paper watermarks in money bills or stamps [135] actually inspired the first use of the term watermarking in the context of

The idea of digital image watermarking arose independently in 1990 [131], [132] and around 1993 [20], [136]. Tirkel *et al.* [136] coined the word "water mark" which became "watermark" later on. It took a few more years until 1995/1996 before watermarking received remarkable attention. Since then, digital watermarking has gained a lot of attention and has evolved very quickly, and while there are a lot of topics open for further research, practical working methods and systems have been developed. In this paper, we introduce the concepts and illustrate them with some of the work that has been published. While attempting to be as complete as possible, we can still only give a rough overview.

B. Terminology

Today, we are of course concerned with digital communication. As in classical analog communication, also in digital communication there is interest for methods that allow the transmission of information hidden or embedded in other data. While such techniques often share similar principles and basic ideas, there are also important distinguishing features, mainly in terms of robustness against attacks. Several names have been coined for such techniques. However, the terms are often confused, and therefore it is necessary to clarify the differences.

Steganography stands for techniques in general that allow secret communication, usually by embedding or hiding the secret information in other, unsuspected data. Steganographic methods generally do rely on the assumption that the existence of the covert communication is unknown to third parties and are mainly used in secret point-to-point communication between trusting parties. As a result, steganographic methods are in general not robust, i.e., the hidden information cannot be recovered after data manipulation.

Watermarking, as opposed to steganography, has the additional notion of robustness against attacks. Even if the existence of the hidden information is known it is difficult—ideally impossible—for an attacker to destroy the embedded watermark, even if the algorithmic principle of the watermarking method is public. In cryptography, this is known as *Kerkhoffs law:* a cryptosystem should be secure, even if an attacker knows the cryptographic principles and methods used but does not have the appropriate key [117]. A practical implication of the robustness requirement is that watermarking methods can typically embed much less information into host data than steganographic methods. Steganography and watermarking are thus more complementary than competitive approaches. In the remainder of this paper, we focus on watermarking methods and not on steganographic methods in general. For an overview of steganographic methods the reader is referred to [67], [99], and [124].

Data hiding and *data embedding* are used in varying contexts, but they do typically denote either steganography or applications "between" steganography and watermarking, which means applications where the existence of the

to protect it. This is typically the case for the embedded transmission of auxiliary information or services [125] that are publicly available and do not relate to copyright protection or conditional access functionalities.

Fingerprinting and *labeling* are terms that denote special applications of watermarking. They relate to copyright protection applications where information about originator and recipient of digital data is embedded as watermarks. The individual watermarks, which are unique codes out of a series of codes, are called "fingerprints" or "labels."

Bit-stream watermarking is sometimes used for data hiding or watermarking of compressed data, for example, compressed video.

The term *embedded signatures* has been used instead of "watermarking" in early publications. Because it potentially leads to confusion with cryptographic digital signatures [117], it is usually not used anymore. Cryptographic signatures serve for authentication purposes. They are used to detect alterations of the signed data and to authenticate the sender. Watermarks, however, are only in special applications used for authentication and are usually designed to *resist* alterations and modifications.

Visible watermarks, as the name says, are visual patterns, like logos, which are inserted into or overlaid on images (or video), very similar to visible paper watermarks. However, the name is confusing since visible watermarks are not watermarks in the sense of this paper. Visible watermarks are mainly applied to images, for example, to visibly mark preview images available in image databases or on the World Wide Web in order to prevent people from commercial use of such images. A visible watermarking method devised by Braudaway *et al.* [16] combines the watermark image with the original image by modifying the brightness of the original image as a function of the watermark and a secret key. The secret key determines pseudorandom scaling values used for the brightness modification in order to make it difficult for attackers to remove the visible mark.

III. Digital Watermarking

A. Requirements

The basic requirements in watermarking apply to all media and are very intuitive.

1) A watermark shall convey as much information as possible, which means the watermark data rate should be high.
2) A watermark should in general be secret and should only be accessible by authorized parties. This requirement is referred to as security of the watermark and is usually achieved by the use of cryptographic keys.
3) A watermark should stay in the host data regardless of whatever happens to the host data, including all possible signal processing that may occur, and including all hostile attacks that unauthorized parties may attempt. This requirement is referred to as robustness of the watermark. It is a key requirement for copyright protection or conditional access applications, but are not required to be cryptographically secure, for example, for applications where watermarks convey public information.
4) A watermark should, though being unremovable, be imperceptible.

Depending on the media to be watermarked and the application, this basic set of requirements may be supplemented by additional requirements.

1) Watermark recovery may or may not be allowed to use the original, unwatermarked host data.
2) Depending on the application, watermark embedding may be required in real time, e.g., for video fingerprinting. Real-time embedding again may, for complexity reasons, require compressed-domain embedding methods.
3) Depending on the application, the watermark may be required to be able to convey arbitrary information. For other applications, only a few predefined watermarks may have to be embedded, and for the decoder it may be sufficient to check for the presence of one of the predefined watermarks (hypothesis testing).

In the following, a few of the mentioned requirements and the resulting design issues are highlighted in more detail.

1) Watermark Security and Keys: If security, i.e., secrecy of the embedded information, is required, one or several secret and cryptographically secure keys have to be used for the embedding and extraction process. For example, in many schemes, pseudorandom signals are embedded as watermarks. In this case, the description and the seed of the pseudorandom number generator may be used as key. There are two levels of secrecy. In the first level, an unauthorized user can neither read or decode an embedded watermark nor can he detect if a given set of data contains a watermark. The second level permits unauthorized users to detect if data are watermarked, however, the embedded information cannot be read without having the secret key. Such schemes can, for example, embed two watermarks, one with a public key and the other with a secret key. Alternatively, a scheme has been proposed which combines one or several public keys with a private key and embeds one combined public/private watermark, rather than several watermarks [48]. When designing an overall copyright protection system, issues like secret key generation, distribution, and management (possibly by trusted third parties), as well as other system integration aspects have to be considered.

2) Robustness: In the design of any watermarking scheme, watermark robustness is typically one of the main issues, since robustness against data distortions introduced through standard data processing and attacks is a major requirement. Standard data processing includes all data manipulation and modification that the data might undergo in the usual distribution chain, such as data editing, printing, enhancement, and format conversion. "Attack" denotes data manipulation with the purpose of impairing, destroying, or removing the embedded watermarks. Section IX-B below revisits attacks and gives remedies that help to make

Although it is possible to design robust watermarking techniques, it should be noted that a watermark is only robust as long as it is not public, which means as long as it cannot be read by everyone. If watermark detector principle and key are public, and even if only a "black-box" watermark detector is public, the watermark is vulnerable to attacks [28], [64]. Hence, public watermarks, as sometimes proposed in the literature, are not robust unless every receiver uses a different key. This however is difficult in practice and gives rise to collusion attacks.

3) Imperceptibility: One of the main requirements for watermarking is the perceptual transparency. The data embedding process should not introduce any perceptible artifacts into the host data. On the other hand, for high robustness, it is desirable that the watermark amplitude is as high as possible. Thus, the design of a watermarking method always involves a tradeoff between imperceptibility and robustness. It would be optimal to embed a watermark just below the threshold of perception. However, this threshold is difficult to determine for real-world image, video and audio signals. Several measures to determine objectively perceived distortion and the threshold of perception have been proposed for the mentioned media [75]. However, most of them are still not perfect enough to replace human viewers or listeners who judge the visual or audio fidelity through blind tests. Thus, in the design of watermarking systems, it is usually necessary to do some testing with volunteers. The second problem occurs in combination with post watermarking processing, which might result in an amplification of the embedded watermark and make it perceptible. An example is zooming of watermarked images, which often makes the embedded watermarks visible, or contrast enhancement, which may amplify highly frequent watermark patterns that are otherwise invisible.

4) Watermark Recovery With or Without the Original Data: Watermark recovery is usually more robust if the original, unwatermarked data are available. Further, availability of the original data set in the recovery process allows the detection and inversion of distortions which change the data geometry. This helps, for example, if a watermarked image has been rotated by an attacker. However, access to the original data is not possible in all cases, for example, in applications such as data monitoring or tracking. For other applications, like video watermarking, it may be impractical to use the original data because of the large data volume, even if it is available. It is, however, possible to design watermarking techniques that do not need the original for watermark extraction. Most watermarking techniques perform some kind of modulation in which the original data set is considered a distortion. If this distortion is known or can be modeled in the recovery process, explicitly designed techniques allow its suppression without knowledge of the original. In fact, most recent methods do not require the original for watermark recovery. In some publications, such techniques are called "blind" watermarking techniques [2], [1].

5) Watermark Extraction or Verification of Presence for a Given Watermark: In the literature, two different types of a specific information or pattern and check the existence of the (known) information later on in the watermark recovery—usually using some sort of hypothesis testing—and systems that embed arbitrary information into the host data.

The first type, verification of the presence of a known watermark, is sufficient for most copyright-protection applications.

The second type, embedding of arbitrary information, is, for example, useful for image tracking on the Internet with intelligent agents where it might not only be of interest to discover images, but also to classify them. In such cases, the embedded watermark can serve as an image identification number. Another example where arbitrary information has to be embedded are applications for video distribution where, e.g., the serial number of the receiver has to be embedded.

Although most presented methods or systems are designed for either watermark extraction or verification of presence for a given watermark, it should be noted that in fact both approaches are inherently equivalent. A scheme that allows watermark verification can be considered as a 1-bit watermark recovery scheme, which can easily be extended to any number of bits by embedding several consecutive "1-bit watermarks." The inverse is also true: a watermark recovery scheme can be considered as a watermark verification scheme assuming the embedded information is known.

B. Basic Watermarking Principles

The basic idea in watermarking is to add a watermark signal to the host data to be watermarked such that the watermark signal is unobtrusive and secure in the signal mixture but can partly or fully be recovered from the signal mixture later on if the correct cryptographically secure key needed for recovery is used.

To ensure imperceptibility of the modification caused by watermark embedding, a perceptibility criterion of some sort is used. This can be implicit or explicit, host data adaptive or fixed, but it is necessary. As a consequence of the required imperceptibility, the individual samples (e.g., pixels or transform coefficients) that are used for watermark embedding can only be modified by an amount relatively small to their average amplitude.

To ensure robustness despite the small allowed changes, the watermark information is usually redundantly distributed over many samples (e.g., pixels) of the host data, thus providing a "holographic" robustness, which means that the watermark can usually be recovered from a small fraction of the watermarked data, but the recovery is more robust if more of the watermarked data are available for recovery.

As said before, watermark systems do in general use one or more cryptographically secure keys to ensure security against manipulation and erasure of the watermark.

There are three main issues in the design of a water-

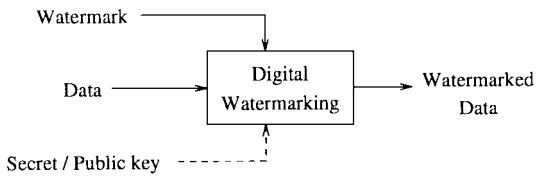

Fig. 1. Generic digital watermarking scheme.

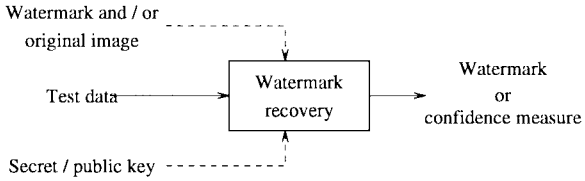

Fig. 2. Generic watermark recovery scheme.

1) Design of the watermark signal $\boldsymbol{W}$ to be added to the host signal. Typically, the watermark signal depends on a key $\mathcal{K}$ and watermark information $\boldsymbol{I}$

$$\boldsymbol{W} = f_0(\boldsymbol{I}, \mathcal{K}). \tag{1}$$

Possibly, it may also depend on the host data $\boldsymbol{X}$ into which it is embedded

$$\boldsymbol{W} = f_0(\boldsymbol{I}, \mathcal{K}, \boldsymbol{X}). \tag{2}$$

2) Design of the embedding method itself that incorporates the watermark signal $\boldsymbol{W}$ into the host data $\boldsymbol{X}$ yielding watermarked data $\boldsymbol{Y}$

$$\boldsymbol{Y} = f_1(\boldsymbol{X}, \boldsymbol{W}). \tag{3}$$

3) Design of the corresponding extraction method that recovers the watermark information from the signal mixture using the key and with help of the original

$$\hat{\boldsymbol{I}} = g(\boldsymbol{X}, \boldsymbol{Y}, \mathcal{K}) \tag{4}$$

or without the original

$$\hat{\boldsymbol{I}} = g(\boldsymbol{Y}, \mathcal{K}). \tag{5}$$

The first two issues, watermark signal design and watermark signal embedding, are often regarded as one, specifically for methods were the embedded watermark is host signal adaptive.

Figs. 1 and 2 illustrate the concept. Fig. 1 shows the generic watermarking scheme for the embedding process. The input to the scheme is the watermark, the host data, and an optional public or secret key. The host data may, depending on the application, be uncompressed or compressed, however, most proposed methods work on uncompressed data. The watermark can be of any nature, such as a number, text, or an image. The secret or public key is used to enforce security. If the watermark is not to be read by unauthorized parties, a key can be used to protect the watermark. In combination with a secret or a public key, the watermarking techniques are usually referred to as secret and public watermarking techniques, respectively. The output of the data. The generic watermark recovery process is depicted in Fig. 2. Inputs to the scheme are the watermarked data, the secret or public key, and, depending on the method, the original data and the original watermark. The output of the watermark recovery process is either the recovered watermark or some kind of confidence measure indicating how likely it is for the given watermark at the input to be present in the data under inspection.

Many proposed watermarking schemes use ideas borrowed from spread-spectrum radio communications [25], [43], [101]. They embed a watermark by adding a pseudonoise (PN) signal with low amplitude to the host data. This specific PN signal can later on be detected using a correlation receiver or matched filter. If the parameters like amplitude and the number of samples of the added PN signal are chosen appropriately, the probabilities of false-positive or false-negative detections are very low. The PN signal has the function of a secret key. The scheme can be extended if the PN signal is either added or subtracted from the host signal. In this case, the correlation receiver will calculate either a high-positive or high-negative correlation in the detection. Thus, 1 bit of information can be conveyed. If several such watermarks are embedded consecutively, arbitrary information can be conveyed.

IV. Text Document Watermarking

Methods for embedding information into text documents have been used for a long time by secret services.

For text watermarking, we have to distinguish between methods that hide information in the semantics, which means in the meaning and ordering of the words, and methods that hide information in the format, which means in the layout and the appearance.

The first class designs a text around the message to be hidden. In that sense, the information is not really embedded in existing information, but rather covered by misleading information. This class of techniques is outside the scope of this paper and will not be considered here. In the following, we concentrate on the latter type of information-embedding methods which use an existing text document into which data are embedded.

Formatted text is probably the medium where watermarking methods can be defeated most easily. If the watermark is in the format, then it can obviously be removed by "retyping" the whole text using a new character font and a new format where "retyping" can be either manual or automated using optical character recognition (OCR). OCR systems are still not perfect for many applications today and often need human supervision. Thus, removal of watermarks either yields bad results (single characters are wrong, due to OCR) or is expensive. The goal is to make watermark removal more expensive than obtaining the right to copy from the copyright owner. If this goal is achieved, text watermarking makes sense, though it can be defeated [14].

Text watermarking has applications wherever copyrighted electronic documents are distributed. Important examples

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.