ISSN 0165-1684
SPRODR 66(3) 281-408 (1998)

SIGNAL PROCESSING

CONTENTS

A publication of the European Association for Signal Processing (EURASIP)

SIGNAL PROCESSING

Volume 66, No. 3, May 1998

CONTENTS

Special Issue on Watermarking
Guest Editors: Benoit Macq and Ioannis Pitas

Advance tables of contents can be retrieved at
http://www.elsevier.nl/locate/sigpro.
For submission of papers for the Fast Communications section please also check the journal's homepage:
http://www.elsevier.nl/locate/sigpro.

ELSEVIER

Signal Processing 66 (1998) 337–355

SIGNAL PROCESSING

Robust audio watermarking using perceptual masking[1]

Mitchell D. Swanson[a,*], Bin Zhu[a], Ahmed H. Tewfik[a], Laurence Boney[b]

[a] *Department of Electrical Engineering, University of Minnesota, Minneapolis, MN 55455, USA*
[b] *Ecole Nationale Supérieure des Télécommunications/SIG, 46 rue Barrault, 75634 Paris Cedex 13, France*

Received 10 February 1997; received in revised form 11 November 1997

Abstract

We present a watermarking procedure to embed copyright protection into digital audio by directly modifying the audio samples. Our audio-dependent watermarking procedure directly exploits temporal and frequency perceptual masking to guarantee that the embedded watermark is inaudible and robust. The watermark is constructed by breaking each audio clip into smaller segments and adding a perceptually shaped pseudo-random sequence. The noise-like watermark is statistically undetectable to prevent unauthorized removal. Furthermore, the author representation we introduce resolves the deadlock problem. We also introduce the notion of a dual watermark: one which uses the original signal during detection and one which does not. We show that the dual watermarking approach together with the procedure that we use to derive the watermarks effectively solves the deadlock problem. We also demonstrate the robustness of that watermarking procedure to audio degradations and distortions, e.g., those that result from colored noise, MPEG coding, multiple watermarks, and temporal resampling.

Zusammenfassung

Wir stellen ein Wasserzeichen-Verfahren zur Einbettung des Urheberrechtsschutzes in digitale Audiodaten vor, wobei die Audiosignalwerte direkt modifiziert werden. Unser audioabhängiges Wasserzeichen-Verfahren nützt unmittelbar die Wahrnehmungsverdeckung in Zeit-und Frequenzbereich aus, um sicherzustellen, däs das eingebettete Wasserzeichen unhörbar und robust ist. Das Wasserzeichen wird konstruiert, indem jeder Audioabschnitt in kleinere Segmente zerteilt wird und eine wahrnehmungsgerecht geformte Pseudozufallsfolge hinzuaddiert wird. Das geräuschartige Wasserzeichen ist statistisch nicht erkennbar, um unautorisiertes Entfernen zu verhindern. Weiters löst die von uns eingeführte Autorendarstellung das Pattstellungsproblem. Wir führen auch den Begriff dualer Wasserzeichen ein: eines, das das Originalsignal während der Erkennung benutzt, und eines, das es nicht benutzt. Wir zeigen, däs der Ansatz mit dualen Wasserzeichen in Verbindung mit dem Verfahren, das wir zur Herleitung der Wasserzeichen einsetzen, das Pattstellungsproblem wirksam löst. Wir zeigen auch die Robustheit des Wasserzeichen-Verfahrens gegenüber Audiostörungen und -verzerrrungen, z.B. jenen, die von farbigem Rauschen, MPEG-Codierung, mehrfachen Wasserzeichen, und Abtastratenwandlung herrühren.

Résumé

Nous présentons dans cet article une procédure de watermarking permettant d'intégrer une protection de droits d'auteur dans des données audio numériques par modification directe des échantillons audio. Cette procédure exploite

* Corresponding author.

[1] This work was supported by AFOSR under grant AF/F49620-94-1-0461. Patent pending, Media Science, Inc., 1996.

0165-1684/98/$19.00
PII S0165-1684(98)00014-0

directement les masquages perceptuels temporel et fréquentiel pour garantir que le filigrane numérique (watermark) est inaudible et robuste. Le watermark est construit en fragmentant chaque morceau audio en segments plus petits et en ajoutant une séquence pseudo-aléatoire modelée perceptuellement. Le watermark semblable à du bruit est indétectable statistiquement afin d'empêcher une suppression non autorisée de celui-ci. De plus, la représentation de l'auteur que nous introduisons résoud le problème de l'impasse. Nous introduisons également la notion de watermark dual: l'un qui utilise le signal original lors de la détection et l'autre non. Nous montrons que l'approche de watermarking dual combinée avec la procédure que nous utilisons pour dériver les watermarks résoud effectivement le problème de l'impasse. Nous mettons également en évidence la robustesse de cette procédure de watermarking vis-à-vis des dégradations et distorsions audio, telles que celles qui résultent d'un bruit coloré, d'un codage MPEG, de watermarks multiples, et de ré-échantillonnage temporel.

Keywords: Copyright protection; Masking; Digital watermarking

1. Introduction

Efficient distribution, reproduction, and manipulation have led to wide proliferation of digital media, e.g., audio, video, and images. However, these efficiencies also increase the problems associated with copyright enforcement. For this reason, creators and distributors of digital data are hesitant to provide access to their intellectual property. They are actively seeking reliable solutions to the problems associated with copyright protection of multimedia data.

Digital watermarking has been proposed as a means to identify the owner or distributor of digital data. Watermarking is the process of encoding hidden copyright information in digital data by making small modifications to the data samples. Unlike encryption, watermarking does not restrict access to the data. Once encrypted data is decrypted, the media is no longer protected. A watermark is designed to *permanently* reside in the host data. When the ownership of a digital work is in question, the information can be extracted to completely characterize the owner.

To function as a useful and reliable intellectual property protection mechanism, the watermark must be:

- *embedded* within the host media;
- *perceptually inaudible* within the host media;
- *statistically undetectable* to ensure security and thwart unauthorized removal;
- *robust* to manipulation and signal processing operations on the host signal, e.g., noise, compression, cropping, resizing, D/A conversions, etc.; and
- *readily extracted* to completely characterize the copyright owner.

In particular, the watermark may not be stored in a file header, a separate bit stream, or a separate file. Such copyright mechanisms are easily removed. The watermark must be inaudible within the host audio data to maintain audio quality. The watermark must be statistically undetectable to thwart unauthorized removal by a 'pirate'. A watermark which may be localized through averaging, correlation, spectral analysis, Kalman filtering, etc., may be readily removed or altered, thereby destroying the copyright information.

The watermark must be robust to signal distortions, incidental and intentional, applied to the host data. For example, in most applications involving storage and transmission of audio, a lossy coding operation is performed on the audio to reduce bit-rates and increase efficiency. Operations which damage the host audio also damage the embedded watermark. The watermark is required to survive such distortions to identify the owner of the data. Furthermore, a resourceful pirate may use a variety of signal processing operations to attack a digital watermarking. A pirate may attempt to defeat a watermarking procedure in two ways: (1) damage the host audio to make the watermark undetectable, or (2) establish that the watermarking scheme is unreliable, i.e., it detects a watermark when none is present. The watermark should be impossible to defeat without destroying the host audio.

Finally, the watermark should be readily extracted given the watermarking procedure and the proper author signature. Without the correct signature, the watermark cannot be removed. The extracted watermark must correctly identify the owner and solve the deadlock issue (cf. Section 2) when multiple parties claim ownership.

Watermarking digital media has received a great deal of attention recently in the literature and the research community. Most watermarking schemes focus on image and video copyright protection, e.g., [1–3,7,10,14,15,18,19,22,24]. A few audio watermarking techniques have been reported. Several techniques have been proposed in [1]. Using a phase coding approach, data is embedded by modifying the phase values of Fourier transform coefficients of audio segments. Embedding data as spread spectrum noise have also been proposed. A third technique, echo coding, employs multiple decaying echoes to place a peak in the cepstrum at a known location. Another audio watermarking technique is proposed in [21], where Fourier transform coefficients over the middle frequency bands are replaced with spectral components from a signature. some commercial products are also available. The ICE system from Central Research Laboratories inserts a pair of very short tone sequences into an audio track. An audio watermarking product MusiCode is available from ARIS technologies.

Most schemes utilize the fact that digital media contain perceptually insignificant components which may be replaced or modified to embed copyright protection. However, the techniques do not *directly* exploit spatial/temporal and frequency masking. Thus, the watermark is not guaranteed inaudible. Furthermore, robustness is not maximized. The amount of modification made to each coefficient to embed the watermark are estimated and not necessarily the maximum amount possible. In this paper, we introduce a novel watermarking scheme for audio which exploits the human auditory system (HAS) to guarantee that the embedded watermark is imperceptible. As the perceptual characteristics of individual audio signals vary, the watermark adapts to and is highly dependent on the audio being watermarked. Our watermark is generated by filtering a pseudo-random sequence (author id) with a filter that approximates the frequency masking characteristics of the HAS. The resulting sequence is further shaped by the temporal masking properties of the audio. Based on pseudo-random sequences, the noise-like watermark is statistically undetectable. Furthermore, we will show in the sequel that the watermark is extremely robust to a large number of signal processing operations and is easily extracted to prove ownership.

The work presented in this paper offers several major contributions to the field, including

A perception-based watermarking procedure: The embedded watermark *adapts* to each individual host signal. In particular, the temporal and frequency distribution of the watermark are dictated by the temporal and frequency masking characteristics of the host audio signal. As a result, the amplitude (strength) of the watermark increases and decreases with host, e.g., lower amplitude in 'quiet' regions of the audio. This guarantees that the embedded watermark is inaudible while having the maximum possible energy. Maximizing the energy of the watermark adds robustness to attacks.

An author representation which solves the deadlock problem: An author is represented with a pseudo-random sequence created by a pseudo-random generator [13] and *two* keys. One key is *author* dependent, while the second key is *signal* dependent. The representation is able to resolve rightful ownership in the face of multiple ownership claims.

A dual watermark. The watermarking scheme uses the original audio signal to detect the presence of a watermark. The procedure can handle virtually *all* types of distortions, including cropping, temporal rescaling, etc., using a generalized likelihood ratio test. As a result, the watermarking procedure is a powerful digital copyright protection tool. We integrate this procedure with a second watermark which does *not* require the original signal. The dual watermarks also address the deadlock problem.

In the next section, we introduce our noise-like author representation and the dual watermarking scheme. Our frequency and temporal masking models are reviewed in Section 3. Our watermarking design and detection algorithms are introduced in Sections 4 and 5. Finally, experimental results are presented in Section 6. Watermark statistics and fidelity results for four test audio signals are

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.