Watermarking Techniques for Multimedia Authentication

Paulo Vinicius Koerich Borges, Robinson Pizzio, Joceli Mayer

[1]LPDS: Digital Signal Processing Research Laboratory
Departamento de Engenharia Elétrica – Universidade Federal de Santa Catarina (UFSC)
Caixa Postal 476 – 88.040-900 – Florianópolis – SC – Brazil

{paulo, rpizzio, mayer}@eel.ufsc.br

Abstract. *Recently we have witnessed great interest in watermarking technologies. Watermarking can be seen as a solution to overcome many security and authentication problems. Several watermarking methods have been proposed in the literature. In this paper, we review some of these techniques and their applications, properties, advantages and shortcomings. Moreover, we indicate that watermarking technology needs to be joined to classical security methods in order to solve challenging problems in multimedia authentication and other applications.*

1. Introduction

Undoubtly digital data has many advantages over similar analog data. Unfortunately, with the growth of the Internet and the advances of copying hardware such as CD and DVD writers, it has become extremely easy to duplicate and illegally disseminate digital contents. To provide copy and copyright protection, two complementary techniques have been developed: watermarking and cryptography.

Briefly, a watermarking system aims to insert hidden information in a digital content such as a music, speech, still images, text, or video. For simplicity, the digital content that receives the watermark will be named *host* data or signal.

This paper deals with watermarking in an authentication perspective, providing a state-of-the-art review. In Section 2. we present a deeper discussion on possible applications and desirable properties of a watermarking system. Section 3. is dedicated to digital watermarking techniques that are related to data authentication, focusing on spread spectrum watermarking and authentication of documents such as text. Finally, in Section 4. we conclude the paper presenting some of the new trends in the digital watermarking area.

2. Applications and Properties

2.1. Applications

Watermarking techniques are applied to many different applications such as copyright protection, fingerprinting, copy protection, and broadcast monitoring.

To provide copyright protection, a robust watermark can be inserted into a host signal. In case the owner finds that his work is being used without authorization, content ownership can be proved in court.

To trace the distribution source of illegal copies, the owner of the content can use a fingerprinting technique. With these techniques the owner may insert different watermarks into each copy of the content.

Copy protection can be achieved by the use of watermarking techniques, where the watermark carries a copy-forbidden bit information. Watermark detectors, implemented in the recorder device, can determine if the data is allowed to be copied [Cox et al. 2002].

Advertisements of commercial products are usually costly and the task of verifying whether a given advertisement has been properly broadcasted (Broadcast monitoring) is not simple. With the help of watermarking techniques one can embed a watermark on the advertising signal (audio or video) and detect that watermark to verify if the advertisements are broadcasted as contracted.

With the current available programs that easily manipulate digital content (image, video, audio, etc), it is becoming very difficult to identify modifications in a digital content. To verify whether the content has been tampered with, one can make use of *fragile watermarking* [Wolfgang and Delp 1999]. This technique is used to check if the content has suffered any modification and also indicate where the content was altered.

2.2. Properties

The desired properties in a watermarking system depend on the application. Fidelity, also referred as perceptual transparency, is an important property in a watermark system. It refers to the similarity between the original and the watermarked version of a host data. The process of embedding a watermark is considered transparent when it is difficult for a human to distinguish between the original and watermarked versions. Understanding the behavior of the human visual system (HVS) and the human auditory system (HAS) [Cox et al. 2002] has greatly improved the performance of high fidelity watermarking, allowing modifications of the host data in specific locations and forms that result in lesser perceptual impact.

Another important property is capacity, which refers to the maximum number of bits that can be inserted in a host signal by the watermarking process.

Watermark robustness refers to the resilience of a watermark to common signal processing operations and also malicious attacks. By signal processing operations we include JPEG and other lossy compression techniques and also cropping, rescaling, filtering, D/A - A/D conversion, and other operations. By malicious attacks we mean modifications with the specific intent of removing or altering the inserted watermark. It is important to note that the watermark must only survive modifications up to the point where the host data is still perceptually acceptable.

3. Common Watermarking Techniques

3.1. Basic Correlation Method

The most common watermarking technique adds a known noise pattern to the host signal. Let $\mathbf{x}$ be a host signal of length N (an image with N pixels, for example) and let $\mathbf{s}$ be a watermarked version of it, also of length N. We define the watermarking embedding process as $\mathbf{s} = \mathbf{x} + K\mathbf{w}$ where $\mathbf{w}$ is a pseudorandom sequence (PRS) with elements w_i usually normally distributed or uniformly distributed, $w_i \in \{\pm 1\}, i = 1 \ldots N$. The key κ

used to generate $\mathbf{w}$ should be known by both the transmitter and receiver. A gain factor K is used to modulate $\mathbf{w}$, controlling the fidelity and robustness tradeoff of the watermark. A higher K provides more robustness, but also decreases fidelity.

Considering a general attack, equation $\mathbf{s} = \mathbf{x} + K\mathbf{w}$ can be extended to $\mathbf{y} = \mathbf{s} + \mathbf{n}$. In this equation $\mathbf{n}$ represents any kind of channel noise between embedding and detection. It can be additive white Gaussian noise (AWGN), compression distortions, brightness, contrast or volume adjustments, or any other malicious or non malicious attack already discussed.

To detect a watermark in a possibly watermarked image $\mathbf{s}'$ (or $\mathbf{y}'$ if we consider attacks), we perform the correlation between $\mathbf{s}'$ and $\mathbf{w}$, with $\mathbf{w}$ generated with the same key κ of the embedding process. Sequences generated with different keys have low correlation with each other. Therefore, in the detection process, the correlation value tends to be very high for a PRS generated with the correct key and tends to be very low for unwatermarked images or for images watermarked with a sequence generated with a different key.

Let us define the linear correlation $\text{Cor}(\cdot,\cdot)$ and the detection statistic d as $d \triangleq \text{Cor}(\mathbf{x},\mathbf{w}) \triangleq \frac{1}{N}\sum_{i=1}^{N} x_i w_i$. To decide whether the watermark is present or not, it is common to set a threshold T. If the correlation d exceeds a certain threshold T the watermark detector determines that $\mathbf{s}'$ contains the watermark $\mathbf{w}$.

With the basic method just described, we are inserting only one bit, given by the presence or not of the watermark. Alternatives to insert more than one bit are space or time division multiplexing [Cox et al. 2002], direct message coding [Cox et al. 2002], where a different pseudorandom sequence is associated with each message, and code division multiplexing (CDMA), also known as spread spectrum (SS) watermarking. In the following we give special attention to SS watermarking due to its great impact in literature and relative efficiency.

3.2. Spread Spectrum Watermarking

Consider we need to embed into $\mathbf{x}$ not only a single bit, but a multi-bit antipodal message of length L represented by $\mathbf{b}$, $\mathbf{b} = [b_1, b_2, \ldots, b_L], b_l \in \{-1, +1\}$. In traditional SS watermarking [Cox et al. 2002] we generate a watermark vector $\mathbf{w_m}$ of length N to carry the message by combining L reference marks from a set $\mathcal{W}$. Let us assume that each reference mark $\mathbf{w}_l$ is a normally distributed sequence of length N, $\mathcal{W} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_L\}$, $w_{li} \sim \mathcal{N}(0, \sigma_w^2)$.

The message $\mathbf{b}$ is spread into an N-dimensional sequence $\mathbf{w_m} = \frac{1}{\sqrt{L}}\sum_{l=1}^{L} b_l \mathbf{w}_l$. We scale the summation by $1/\sqrt{L}$ to enforce that $\sigma_{w_m}^2 = \sigma_w^2$. The watermarked vector $\mathbf{s}$ is obtained by additive embedding: $\mathbf{s} = \mathbf{x} + \mathbf{w_m}$. We use a detection statistic d_l to estimate the l-th bit value given by $d_l \triangleq \text{Cor}(\mathbf{s}, \mathbf{w}_l) = \frac{1}{N}\sum_{i=1}^{N} s_i w_{l_i}$. Where l represents the reference mark index and i represents the vector element index.

3.3. Alternative Methods

Many different strategies exist to insert an watermark, whether they are based on SS or not. In the following we discuss some of them, pointing out their main characteristics.

Some systems presented in literature are designed to survive specific attacks. The work developed in [Mayer and Silva 2003], for example, survives JPEG compression up to a given quality factor Q. There, each element w_i of $\mathbf{w}$ in equation $\mathbf{s} = \mathbf{x} + K\mathbf{w}$ is modulated with a distinct gain factor (instead of a constant K), just strong enough to ensure a correct detection up to a given Q. Any extra modification can destroy the watermark.

In [Mayer and Silva 2004] the authors use a high frequency wavelet mask $\mathbf{m}$ to modulate the pattern $\mathbf{w}$, such that $\mathbf{s} = \mathbf{x} + \mathbf{m} * \mathbf{w}$, where $*$ is defined as the element-by-element matrix or vector multiplication. In this way, they concentrate most of the energy of $\mathbf{w}$ over the edges and high frequency regions of $\mathbf{x}$ thus making the insertion of $\mathbf{w}$ less visible due to the characteristics of the HVS.

Different techniques are also presented in literature. A simple yet effective way to provide content authentication is to modify the least significant bits (LSB) of an image, replacing them with a binary logo, for example. In the system proposed in [Borges and Mayer 2005] several bits of information are encoded by the *position* of pseudo-random blocks. Depending on the position of the blocks, a different bit string $\mathbf{b}$ is transmitted.

4. Conclusions

In this paper we have discussed several watermarking techniques, with applications to different media. We have focused on techniques that can be used to content authentication, avoiding tampering with signals. Most of the work developed in literature has dealt with spread spectrum watermarking and it has already been widely used in commercial applications. Promising research will probably try to solve the problem of geometrical attacks in watermarking. Another challenging area in authentication is to design watermarks able to survive the print and scan process, not only for simple text, but for magazine articles, identification cards, etc. Watermarking is not yet a mature research field, and it shall gain even more attention in the years to come due to the growing concerns on security issues related to the Internet and entertainment industry.

References

Borges, P. and Mayer, J. (2005). Informed positional embedding for multi-bit watermarking. In *30th IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP 2005*.

Cox, I., Miller, M. L., and Bloom, J. A. (2002). *Digital Watermarking*. Morgan Kaufmann.

Mayer, J. and Silva, R. (2003). Informed embedding for multi-bit watermark. In *IEEE Proceedings of the XVI Brazilian Symposium on Computer Graphics and Image Processing, 2003*.

Mayer, J. and Silva, R. (2004). Efficient informed embedding of multi-bit watermark. In *29th IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP 2004*.

Wolfgang, R. B. and Delp, E. J. (1999). Fragile watermarking using the vw2d watermark. In *Electronic Imaging*, volume 3657, pages 25–27. The Internatinal Society for Optical Engineering.