Digital Watermarking – Principles and Applications

Monica Borda
Technical University of Cluj Napoca
15 Daicoviciu Street, 400020 Cluj Napoca, Romania
E-mail: Monica.Borda@com.utcluj.ro

Ioan Nafoşniţă,
Electronics and Telecommunications Faculty, “Politehnica” University, Timișoara, ioan
ioan_naf@etc.utt.ro

Abstract
This paper presents a brief overview of digital watermarking research area with great implication in
copyright protection for digital multimedia technologies. After an introduction and a short history of the domain,
the watermarking requirements and principles are illustrated together with the main applications. Final remarks
and bibliography are ending the paper.

Keywords: digital watermarking, copyright protection, watermark applications, watermarking requirements

I. Introduction

In the last decade we were the witnesses of an outbreak of the digital multimedia
technologies. The digital audio/ video information has several advantages over its analogical
counterpart:

- superior quality in transmission, processing and storage
- simpler editing facilities, the desired fragments of the initial data can be located
  with precision and modified
- lossless simpler copying: the copy of a digital document is identical with the
  original.

For the producers and distributors of multimedia products, several of the above mentioned
advantages are handicaps, leading to important financial losses. Unauthorized copying of
audiocassettes / videocassettes is currently a major problem. Also the information contained
in WebPages, books, and the broadcasted information, are frequently copied and used without
any permission from the “editor”.

The copyright in this domain is a problem of maximum urgency. Several attempts in this
sense exist, but we cannot speak of a corresponding legislation. In 28 Oct. 1998, the president
of the United States signed an act [65] (Digital Millennium Copyright Act) that contains
recommendations to be followed in order to protect the intellectual property and also the
customer’s rights. At its turn the European Community is preparing several protection
measures for digital multimedia products such as CD and DVD (Digital Versatile Disk).

The most important technologies used in copyright protection for authors or distributors
are: encryption and watermarking.

Encryption is used for protecting data in transmission and storage. Once the information
was decrypted, it is no longer protected and can be copied without any restriction.

Watermarking is an operation, which consists in embedding an imperceptible signal called
watermark (WM) into the host information. The host information can be text, audio signal,
static images or video.

The name watermark comes from the words “water” and “mark” and designates a
transparent, invisible mark like the water transparency.
In general, the watermark contains information about the origin and destination of the host information. Event though it is not directly used in intellectual property protection, it helps identifying the host and the receiver, being useful in disputes over authors/distributors rights. From a theoretical point of view the watermark has to permanently protect the information, so it has to be robust, in such a way that any unauthorized removal will automatically lead to quality degradation. The watermark resembles to a signature, at the beginning it was called signature, but in order to eliminate the confusions with the digital signatures from cryptography the original name was dropped. Taking into account the fact that it has to be transparent, imperceptible for hearing or seeing, the resemblance with the "invisible tattoo", made by A. Tewfik [63], is suggestive.

In order to assure copyright protection, the watermarking technologies need two operations (Fig. 1 and Fig. 2):

- watermark insertion in host data, before transmission or storage;
- watermark extraction from the received data and comparison between the original watermark and the extracted one, in case of dispute.

II. Short history

Today's digital watermarking is a modern version of steganography (form the Greek words "steganos" which means covered and "graphos" meaning writing)-signifying covered writing. Steganography is a technique used for secret message hiding into other messages in such a way that the existence of the secret messages is hidden. The sender writes a secret message and hides it into an insensitive one.

The beginnings of steganography date probably since the knowledge of writing. The writing was the privilege of some classes that had the power (economical, military, and religious). Information always meant power; so it had to be protected against unauthorized persons. [37],[55],[57],[33].

Among the techniques used during the history of steganography we remind:

- use of invisible inks
- thin holes for some characters, fine modifications of the space between words
- the use of eumenograms (from Greek words "sema" meaning sign and "gramma" meaning writing, drawing)

These techniques were recently resumed and put into digital context for text watermarking [14].

Audio watermarking (audioteganography) and still / dynamic image watermarking (videosteganography) are using the same ideas as steganography.

As an example for audioteganography, we can note Bach. He used invisible watermark copyright protection, writing his name in several works using invisible watermarking; for example he counted the number of appearances of a musical note (one appearance for A, two for B, three for C and eight for H).

As for steganography, for graphical images for instance, using the least significant bit, several secret messages can be hidden. The image rests almost the same and the secret message can be extracted at the receiver. Proceeding like that for a 1024x1024 black and white image one can insert 64 KB of secret messages (several modern services are using this capacity).

For digital imaging, the firsts invisible marks appeared in 1990 in Japan [61] and independently, in 1993 in Europe [17], [64]. At the beginning the terminology used for such invisible marks was "label" or "signature"; around 1993 the words water mark were used, signifying a transparent, invisible mark. The combination of the two words, gave the word "watermark", which will be used henceforward.

Applications of digital watermarking for audio domain are known since 1996 [6]. In 1995 the first applications for uncompressed and compressed still images are done [19]. 1996 [27], 1997 [39] are marking the beginning for uncompressed, respectively compressed video signals.

After several breakthroughs between 1995-1998 it seems that the last two years can be viewed as a plateau in watermarking research. Simultaneously the industry had an increasing role in standard and recommendation elaboration. This phenomenon resembles to the development of modern cryptography and the elaboration of standards for civil applications.

III. Watermarking requirements

Each watermarking application has specific demands. However there are some general, intuitive requirements.

- **Perceptual transparency.** It is related to the fact that the watermark insertion must not affect the quality of the host data. The mark is invisible if one cannot distinguish between the original signal and the marked one, e.g. if the changes in the data are below the thresholds of human senses (hearing, seeing). Perceptual transparency test are made without knowing the input data. Original or marked data are presented independently to the subjects. If the selection percentage is equal for the two cases, this means that perceptual transparency is achieved. In real perceptual transparency applications, the subjects do not know the original data, having therefore correct testing conditions.

- **Robustness** is the watermark’s property to resist to unintentional changes, due to the inherent processing related to the transmission / storage (unintentional attacks) or to intentional changes (intentional attacks) aiming to remove the watermark.

There are some situations when one does not need this requirement. For data authentication for example, the fragile watermark needs not to be robust, an impossible watermark detection proving the fact that the data is altered, being no longer authentic.

However, for most applications the watermark has to be robust, its extraction from the host data leading to a significant quality loss, making the host data unusable.

- **Watermark payload**: the watermark payload is also known as watermark information. The watermark payload is defined as the information quantity included in the watermark. It is application dependent [38] and some usual values are:
  - 1 bit for copy protection
  - 20 bits for audio signals
  - 60 + 70 bits for video signals

Another important parameter related to the payload is the watermark granularity. This parameter shows the needed data quantity for the insertion of a single watermark information unit. In the above-mentioned example a watermark information unit has 20 bits for audio signals and, 60 + 70 bits for video signals. These bits are inserted in 1 or 5 seconds for audio segments. For video signals the watermark information unit is inserted in a single frame or is spread over multiple frames.

- **Detection with and without original signal.** Depending on the presence of the original signal there are two methods for watermark detection [38]:

42

43
• with the presence of the original signal: nonoblivious (informed) watermarking
• without original signal: oblivious (public, blind) watermarking.

The first type of detection that needs the original signal or a copy of it is used in copyright protection applications restraining the inversion attack [70], [20].

The second detection modality, not needing the original, is used in application where the presence of the original at detection is impossible, for example in copy protection.

• **Security in watermarking** can be seen as in cryptography: contained in the encryption key. Consequently the watermarking is robust if some unauthorized person cannot eliminate the watermark even if this person knows the insertion and detection algorithm. Subsequently the watermark insertion process uses one or several cryptographic robust keys. The keys are used also in the watermark detection process. There are applications, like covered communications, where encryption is necessary before marking [57].

• **Ownership deadlock**.

The ownership deadlock is known as the inversion attack, or IBM attack, [5]. Such an attack appears whenever in the same data there are several watermarks claiming the same copyright. Someone can easily insert his own watermark in the data already marked.

Watermarking schemes capable of solving this problem (who is the “right” owner or who was the first that made the mark), without using at detection the original or a copy of it, are not known until now.

Such a situation can be solved if the watermark is author and host dependent. In such a case the author will use at insertion and detection two keys: \( k_1 \) - author dependent and \( k_2 \) - signal/host dependent. Using the keys he will generate a pseudo-random sequence \( k \). The key \( k_2 \), signal dependent, can be generated using one-way hash (OWH) functions. Such generators are including: RSA, MD4, SHA, Rabin, Blum/Blum/Shub [57]. The watermark extraction at the receiver is impossible without knowing the keys \( k_1 \) and \( k_2 \). The \( k_2 \) key, being host dependent, the counterfeiting is extremely difficult. In copyright protection, the pirate will be able to give to a judge only his secret key \( k_1 \) and not \( k_2 \). The last key is computed automatically using the original signal by the insertion algorithm. The hash function being noninvertible the pirate will not be able to produce a counterfeit identical with the original.

**IV. Basic principles of watermarking**

As shown in the Introduction, watermarking has two basic processing: one at the sender and the other at the receiver:

• **Watermark insertion** in the host data. The insertion is done respecting the perceptual transparency and robustness requirements.

• **Watermark extraction** (detection) from the marked received signals (possibly altered) and the comparison between the extracted watermark and the original one, in case of deadlock.

For the robustness demand the watermark will be inserted using one or several robust cryptographic keys (secret or public). The keys will be further used at watermark detection.

The perceptual transparency is done according to a perceptible criterion, the last one being implicit or explicit. Therefore the individual samples of the host signal (audio signals, pixels or transform coefficients) used for the insertion of the watermark information will be changed only between some limits situated below the perceptiveness thresholds of the human senses (seeing, hearing).

Transparent insertion of the watermark in the digital host signal is possible only because the final user is a human being. His senses (hearing, seeing ) are imperfect detectors characterized by certain minimal perceptiveness thresholds and by the masquerade phenomenon. By masquerade, a component of a given signal may become imperceptible in the presence of another signal called masquerading signal. Most of the coding techniques for audio and video signals are using directly or indirectly the characteristics of the HAS - human audio system or HVS - human visual system [59].

The watermarking techniques cannot, therefore, be used for data representing software or numbers, perceived by a computer (machine not human).

According to the robustness demand the watermarking signal (despite the small amplitude required by the perceptual transparency demand) is spread over several samples according to the granularity demands. This makes possible the detection of the watermark signal even if the data is noise affected.

**Fig. 1** and **Fig. 2** are showing the bloc schemes of watermark insertion and detection.

**Fig. 1** Bloc scheme for watermark insertion

**Fig. 2** Bloc scheme for watermark extraction and comparison

Watermarking (Fig. 1) consists in:

• **Watermark information (I) generation (payload)**
• **Watermark generation** (distinct from I – watermark information): \( W \), that will be inserted into the host signal \( X \); usually \( W \) depends on the watermark information (I) and on the key \( K \):
  \[
  W = E_I(I, K),
  \]

  \( E_I \) is a function (in most cases modulation and spreading).
There are applications where, in order to limit the IBM attack, the watermark signal can be host signal X dependent:

\[ W = E_2(I, X, K) \]  

(2)

- **Key generation:** the key can be public or secret, leading to a possible classification of the watermarking techniques in public keys systems and private keys systems. [20].
- **Watermark signal \( W \) insertion** in the host signal \( X \). The insertion is made with respect to the robustness and perceptual transparency demands, giving the watermarked signal \( Y \):

\[ Y = E_2(X, W) \]  

(3)

where \( E_2 \) is a function (which usually makes a modulo 2 summation between \( W \) and \( X \)).

As a conclusion, in order to fulfill the perceptual transparency demands, the two models HAS or HVS, are taken into account directly or indirectly for watermarking. For the robustness requirements, the watermark information \( I \) spread over the host data (see the granularity concept in III.).

Watermarking can be done in the transform domain or in the spatial domain. It follows that, before watermark insertion or extraction, the host data needs to be converted in the domain where the processing will take place: spatial, wavelet, DCT, DFT, fractals. Each domain has specific properties that can be used in the watermarking process. [38].

Watermarking can also be done for compressed or uncompressed host data; most applications are, however, for uncompressed data. [25].

**Remark:**

Due to the perceptual transparency demands, the changes in the host data are relatively small, so the watermark signal \( W \) will be error vulnerable. In order to overcome this drawback, in transmission or storage, several protection measures can be taken using error correcting codes before watermark insertion [59], [25], [1].

**Watermark extraction** (fig. 2)

The watermark detector’s input signal is \( Y \) and it can be the result of a watermarked signal with errors or not. In order to extract the watermark information \( I \), the original signal \( X \) is necessary or not depending on the detection type: (fig. 2):  

\[ I = D(X, Y', K) \quad \text{nonoblivious detection} \]  

(4)

\[ I = D(Y', K) \quad \text{oblivious detection} \]  

(5)

In copyright applications, the detected watermark information \( I \) is compared with the ownership’s original \( I \) (fig. 2):

\[ C(I, \hat{I}) = \begin{cases} 
1, & \text{if } c \geq y \\
0, & \text{if } c < y 
\end{cases} \]  

(6)

V. Specific attacks

The causes leading to errors in the watermark extraction process are called attacks. According to the way they were produced, the attacks can be classified in two major categories:

- **Unintentional attacks:** due to the usual signal processing in transmission or storage: linear (nonlinear) filtering, JPEG compression, MPEG-2 compression [4], pixel quantisation, analog to digital conversions, digital to analog conversions for recording processes, γ correction. A detailed description of these attacks is done in [20].
- **Intentional attacks** intentionally made in order to eliminate the watermark or to insert false watermark, keeping also the perceptual fidelity. There is other attacks classifications among them we refer to: [67], [26].

A. **Simple attacks**, the watermarked signal sustains some distortions, however the intention being not to eliminate the watermark. The majority of these attacks are unintentional attacks described above.

B. **Detection disabling attacks**, including the synchronization attack. These attacks are oriented towards the watermark extraction devices; the purpose of such an attack is to avoid watermark detection. A common characteristic for such attacks is the signal decorrelation, making the correlation based watermark extraction impossible. In this case the most important distortions are geometric distortions: zooming, frame rotation, subsampling, the insertion or extraction of a pixel or a group of pixels, pixel interchanges, spatial or temporal shifts.

In the case of the Stir Mark [75], the jitter attack consists in the elimination of some columns and the multiplication of others, keeping unchanged the image dimensions. On the same category, frame modifications are included: frame removal, frame insertion or swapping.

C. **Ambiguity attacks**, also known as confusion, deadlock/ inversion-IBM/ fake watermark/ fake original attacks. These attacks are trying to create some confusion by producing a fake original.

D. **Removal attacks** are trying to decompose the watermarked image \( Y \) in a watermark \( W \) and an original \( X \), in order to eliminate the watermark. In this category we mention the **collusion attack**, noise extraction and nonlinear filtering.

In multimedia MPEG compression based applications the attacks can be done in the compressed domain (frequency- DCT), or in the spatial domain. The most important attacks are done in the spatial domain, for uncompressed signals.

There are computer programs for several kinds of attacks, among them we mention:

- **Stir - Mark**, from the Cambridge University,
- **Attack**, from the University of Essex,
- still images oriented useful also for dynamic images too.

In the following paragraphs we will expose the principle for two of the most powerful attacks: the inversion (IBM) attack, and the collusion attack.

VI. Applications

The main applications of the watermarking were mentioned in introduction. As shown in 3, each application has specific demands concerning the watermarking requirements. However, three important features define watermarking from other solutions:

- watermarks are perceptually transparent (imperceptible),
- watermarks are inseparable from the host date in which they are inserted, meaning that
- support the same processing as the host data.

The quality of the watermarking system is mainly evaluated by it’s:

- robustness, which indicates the capability of the watermark to “survive” to unintentional or intentional attacks,
- fidelity [18], describing how transparent (imperceptible) a watermark is.

The value of these quality parameters is highly dependent on the application. We will proceed presenting the most important applications, their requirements, the limitations of alternate techniques, emphasizing the necessity and advantage of the watermarking solution.

1. Broadcast monitoring

**Who is interested in it?**

- advertisers, which are paying to broadcasters for the airtime; in 1997 a huge scandal, broke out in the Japanese television advertising (some stations usually had been overbooking air time),
Broadcast monitoring ways:
- using human observers to watch the broadcast and record what they see and hear; these techniques are expansive, low tech and predisposed to errors.
- automated monitoring:
  - passive monitoring: the humans are replaced by a computer that monitors broadcasts and compares the received data with those found in a data base with known works; in case of matching the work is identified.
  
  The advantages of this system are: do not require cooperation between advertisers and broadcasters and is the least intrusive. Unfortunately it has major disadvantages related to the implementation (need of huge data base, impractical to search); if signatures of the works are used, the data base diminishes a lot, but the broadcast altering the work, will make the signature of the received work different as these corresponding to the original (an exact matching is impossible). This system is applied for marketing, but not for verification services, because its error rate.
  - active monitoring: it removes the main disadvantages of the passive monitoring. There are two ways to accomplish this aim: using computer recognizable identifiers (e.g. vertical blanking interval for analog TV or file headers for digital format) along with the work or the watermarking. The first solution has as main disadvantage the risk of spoiling in transmission and the lack of resistance to format changes. The watermarking solution, even more complicated to be embedded (compared to file headers or VBI), has not any more the risk to be lost as identifiers are.

2. Owner identification. Proof of ownership

Textual Copyright notices, placed in visible places of the work, are the most used solutions for owner identification. An exemplification for visual works is:

"Copyright date owner"
"© date owner" (© 2002 by Academic Press) the textual copyright notice for
[18]
"Copr. date owner"

Unfortunately this widespread technique has some important limitations: the copyright notice is easily omitted when copied the document (e.g. photocop of a book without the page containing the copyright notice), or cropped (Lena is a cropped version of November 1972 Playboy cover) and, if not removed, it is unaesthetic. The watermarking is a superior solution, eliminating the mentioned limitations due its properties of inseparability from the work (can’t be easily removed) and imperceptibility (it is aesthetic).

A practical implementation of a watermarking system for this application is Digimarc, distributed with Adobe Photoshop image processing software; when Digimarc’s detector recognizes a watermark, it contacts a central data base over the Internet and uses this information to find the image’s owner. This available practical system is useful only for honest people.

When malicious attacks are done, as collusion for removal and IBM for creating a fake owner, the proof of ownership is a major task. In a copyrighted work a forgery is very easy to be done in both cases presented before. As example:

"© 2000 Alice" can easily transformed by an adversary (pirate, traitor) Bob as: "© 2000 Bob" by a simple replacement of owner’s. Using the Digimarc system, the replacement of Alice’s watermark with Bob’s watermark is not difficult (if a watermark can be detected, probably can be removed too). For this situation there are some solutions as using an arbitrated protocol (which is costly) or using work dependent watermark, as presented in chapter 5.

3 Fingerprinting (transaction tracking)

The digital technology allows to make quickly cheap and identical (same quality) copies of the original and to distribute them easily and widely. Distribution of illegal copies is a major problem of our days, implying always money (huge amount) loses of copyright owners.

Fingerprinting is a new application, aiming to catch the illegal distributors or at least making it’s probability greater. A legal distributor distributes to a number of users some copies, each one being uniquely marked with a mark named fingerprint (like a serial number). A number of legal users, called pirates, cooperate in creating illegal copies for distribution. If an illegal copy is finding, the fingerprint could be extracted and traced back to the legal owner of that copy.

The idea of transaction tracking is old, being used to unveil spies. To suspects were given different information, some of them false, and by the action of enemy those who revealed the information were caught.

A practical implementation of a fingerprinting system was done by DIVX Corporation, now defunct [18]. Each DIVX player placed a unique watermark (fingerprint) into every played video. If such a video was recorded and than copies sold on the black market, the DIVX Corporation could identify the pirate by decoding the fingerprint. Unfortunately no transaction tracking was done during the life of DIVX Corporation.

Transaction tracking in the distribution of movie dailies is another application. These dailies are very confidential, being not allowed to the press. Studios prefer to use fingerprinting, and not visible texts for marking, because the last one are very easily removed. The quality (fidelity) is not necessary to be very high in this application.

4 Content authentication (Fragile watermarking)

The aim of this application is to detect data tampering and to localize the modification.

Why is it so important?

Because in digital, tampering is very easy and very hard to detect (see image modification with Adobe Photoshop). Consequences of such tampering could, be dramatically in a police investigation.

Message authentication [57] is a basically problem of cryptography, a classical solution being the digital signature, which is an encrypted summary of the message.

This solution was used in "trustworthy digital camera" [18] by computing a signature inside the camera. These signatures act as a header (identifier) necessary to be transmitted along with the work. Unfortunately the integrity of this header is not guaranteed if format changes occur. As in broadcast monitoring, instead of a header, which could be lost, an incorporated signature in the work is preferred, realized using watermarking. These embedded signatures are authentication marks. These marks become invalid at the smallest modification of the work and for this reason they are named fragile watermarks.

In this application, robustness is not anymore a requirement; the watermark designed for authentication should be fragile, as nominated. A detailed presentation of this application is done in [18], chapter 10.
5 Copy control

If the former described applications had effects after an intentional forgery, the copy control is aimed to prevent people to make illegal copies of copyrighted works. The ways for this type of protection mentioned in the Introduction are encryption and watermarking. The both, without an appropriate legislation will not solve the problem.

The first attempt of standardization for the DVD's copy protection is the Millennium watermarking system introduced by Philips, Macrovision and Digitmare in USA; it was submitted to the approval of the USA Congress, and the result was the "Digital Millennium Copyright Act" signed by the president of the USA in 28. 10.1998.

The main cause was the market explosion of digital products like DVD's, digital broadcasting of multimedia products and the producers' exposure to potential huge financial losses, in case of the non-authorized copying.

The standardization of the video DVD's provoked unusual debates in copy protection (like the 1970-1975 years for the DES standardization) influencing the whole multimedia world.

On a DVD the information is encrypted (with secret algorithms) but in order to assure the copy protection the encryption is not enough. Using encryption on a storage device: CD, DVD, or in transmission on communication channels or open interfaces copy protection can be realized using an authentication system and a session key generating mechanism for all interfaces (end to end encryption). Encryption used on DVD's supposes that the players or recorders have incorporated compliant devices. When the content is displayed in clear on a monitor or played on speaker (to the human consumer) the encryption-based protection disappears. It is now when the need for watermarking becomes clear; the watermark assures that copying is allowed for a restricted number of copies (one copy) or prohibited (never copy).

The basic demands for the DVD copy protection watermark system are:

- invisible and hard to remove
- fast watermark extraction (maximum 10 s), therefore real time processing
- cheap watermark detector, with minimum additional hardware required for players and recorders
- robust detection, the watermark has to resist when usual processing of the signal are performed: compression, noise adding, shifts, format conversions etc.
- the watermark's payload has to be at least 8 bits/ detection interval

6 Others

Beside these protection goals, watermarking can be used also for:

- Characteristic enrichment for the host signal, e.g. several language subtitling; there are several services that use this property.
- Medicine applications. Using watermarking techniques, patient data are inserted into the medical images.
- Secret message transmission. There are countries where cryptographically services are restricted; it follows that secret (private) messages can be inserted through watermarking.

VII. Some remarks

Digital watermarking was presented as a solution for copyright protection and especially for multimedia product unauthorized copying. In fact, even though several solutions were proposed, actually the domain rests untested, not experimented.

Among the great lacks shown, we can remind in the first place:
- the lack of standardization in algorithm testing and evaluation (lacks of benchmarking) [54]
- the lack of a suitable legislation.

The copyright protection problem [59] shows that watermarking it is not by no means an unfailing method. Any research teams (even the whole technical community) will not solve copyright protection, because it is related to several legal aspects including a concise definition for similarity and subsequent work. Now we are in a period of great interdisciplinary efforts for national and international recommendations and standard elaboration for ownership protection in the digital era, in which both the multimedia products manufacturers and the legislation (the political factors) have to arrive to an agreement.

VIII. References

[34] S Katzenbeisser, F. Petitoelas (2000), Information hiding techniques for steganography and digital watermarking, Artech House

[75] StirMark, http://www.cel.cam.ac.uk/