University of Alberta

Digital Watermarking: Status, Limitations and Prospects

by

Stanley R. M. Oliveira
Mario A. Nascimento
Osmar R. Zaïane

Technical Report TR 02-01
January 2002

DEPARTMENT OF COMPUTING SCIENCE
University of Alberta
Edmonton, Alberta, Canada

Digital Watermarking: Its Status, Limitations and Prospects

Stanley R. M. Oliveira[1,2] oliveira@cs.ualberta.ca
Mario A. Nascimento[2] mn@cs.ualberta.ca
Osmar R. Zaïane[2] zaiane@cs.ualberta.ca

[1]Embrapa Information Technology, Campinas, São Paulo, Brazil
[2]Department of Computing Science, University of Alberta, Canada

Abstract

Research in information hiding has grown explosively. A large number of techniques have been proposed to discourage copyright infringement, tampering and unauthorized distribution of digital media (e.g. video, audio, and images). In this paper, we provide an overview of information hiding, outlining its main disciplines (covert channels, steganography, digital watermarking, and anonymity), and some applications current driving interest. We focus on the current status of and prospects for digital watermarking, devoting special attention to a taxonomy based on insertion domain, applicability, and types of existing algorithms. Finally, we offer an overview of technical progress and analyze some watermarking limitations, culminating in a discussing of some interesting topics for future research.

1 Introduction

The increase in the availability of digital data (e.g. video, audio, and images) on the Web has led to large-scale unauthorized copying and increased the opportunity for violation of copyright and tampering with (or the modification of) content [7]. The reason is simple – digital representation of media facilitates access and potentially improves the portability, efficiency, and accuracy of the information presented. As a result, there is a pressing need to manage and protect visual material against manipulation and illegal duplication.

One approach to address this problem involves embedding an invisible structure into a host multimedia data to mark ownership of them. To accomplish this, a large number of information hiding techniques have been proposed in the literature. The results achieved in the last 6 years, in a number of application areas involving audio, video, and digital images, have pointed to information hiding as one important topic related to the area of information security.

The information hiding area brings together researchers with very different backgrounds: electrical engineering, signal and image processing, computer science, and cryptography. The main disciplines studied so far have focused on covert channels, steganography, anonymity, and watermarking, as can be seen in Figure 1. This classification of information hiding techniques was first proposed by Bauer [6]. However, other researchers consider steganography, digital watermarking, and fingerprinting at the same level [77, 79]. Recent research has pointed to steganography and digital watermarking as two areas which are generally referred to as information hiding [46, 42].

In this paper we outline the major disciplines related to information hiding and the applications driving interest in them, focusing on digital watermarking in particular. First, we present the watermarking

principles and then a watermarking taxonomy based on the domain of insertion, applications area, and types of existing algorithms. We also analyze the current status of watermarking techniques, focusing on technological progress, limitations, and some challenges for future research. We do not investigate state-of-art of steganography because recent work has already been done in this area [77, 43, 3, 1].

This paper is organized as follows. A classification of information hiding techniques, a brief history, and some application areas are presented in Section 2. In Section 3, we give an overview of digital watermarking principles, focusing on terminology, framework, and properties. In Section 4, we present a watermarking taxonomy based on applicability, embedding techniques, and types of existing algorithms. In Section 5 we discuss the recent advances in watermarking techniques, some limitations, and some interesting topics for research. Related work is discussed in Section 6. Finally, in Section 7, we present our conclusions regarding this work, summarizing the main results. At the end of this paper, we also present a glossary of some key terms used in the description of digital copyright protection schemes.

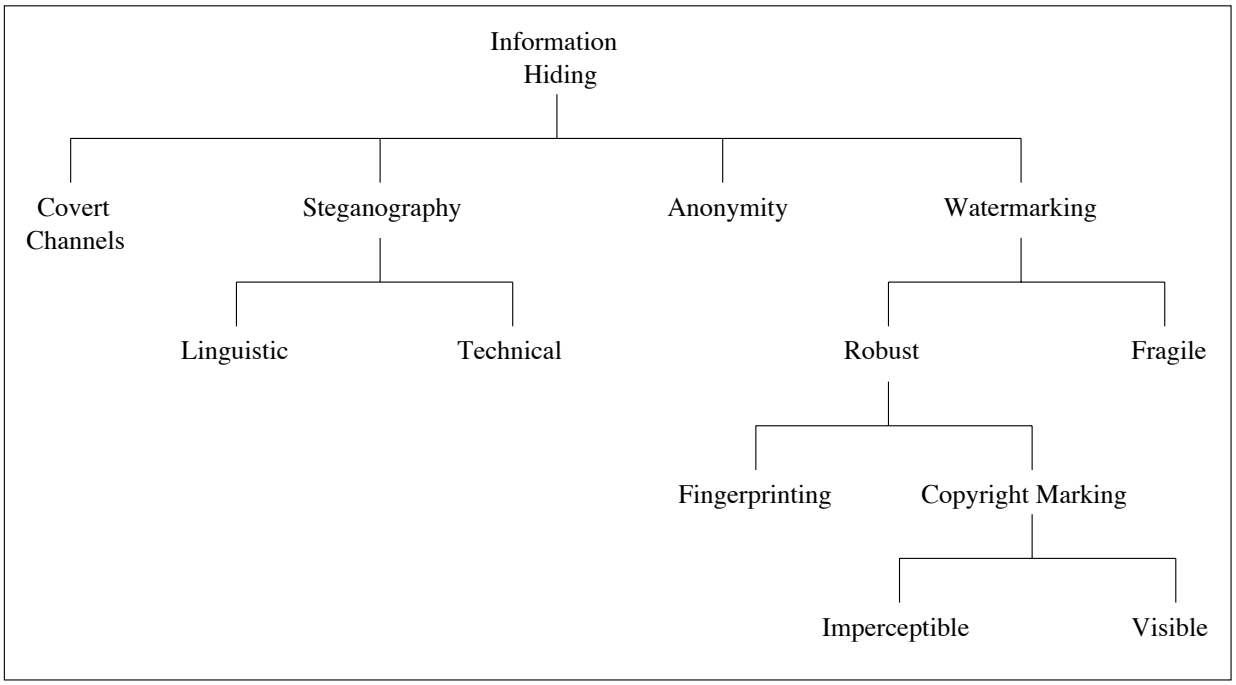

Figure 1: A classification of information hiding techniques

2 Information Hiding at a Glance

In this section, we show a taxonomy of information hiding techniques, provide a brief history of information hiding, and identify some areas to which information hiding can be applied. More details about information hiding can be found in [44, 77, 48, 2, 78, 58].

2.1 Historical Roots of Information Hiding

Although information hiding seems to some as though it is a new science, some of the first documented examples can be found in the Histories of Herodotus, in which the father of history relates several stories from the times of ancient Greece [38, 48]. One classical example refers to Histiaeus, who wished to inform his allies when to revolt against the enemy (Persians). To do so, Histiaeus shaved the head of

a trusted servant and then tattooed a message on his scalp [44, 63]. When the servant's hair grew back, he was sent through Persian territory to the allies. The slave appeared to be a harmless traveler passing through the area, but upon his arrival, the servant reported to the leader of the allies and indicated that his head should be shaved to reveal the message. This method was still used by some German spies in the beginning of the 20th century [78].

Another example related to information hiding's historical roots also goes back to antiquity. One type of writing medium was a wooden tablet covered with wax [77, 71]. A person etched letters in the wax and when he/she desired to remove the writing, the wax was melted to a smooth surface and the tablets reused. While exiled in Persia, Demeratus observed that Greece was about to be invaded and he wanted to send a warning message. As his risk of exposure was high, Demeratus decided to conceal his message by writing it directly onto the wood and then covering it with wax [71]. So the blank tablets were transported to Sparta, where the message was literally uncovered and his allies warned in advance.

A large number of techniques were invented or reported by Aineias the Tactician [94], including letters hidden in messengers' soles or women's earrings, text written on wood tablets and then white-washed, and notes carried by pigeons [46]. He also proposed hiding text by changing the heights of letterstrokes or by making very small holes above or bellow letters in a cover-text [26]. This technique was still in use during the 17th century. Later, Wilkins improved it by using invisible ink to print very small dots instead of making holes [71, 106].

Recent research has yielded more advanced techniques in information hiding. Such techniques have been used for centuries to prove the authenticity of physical materials, and recently they have been used to protect digital property as well.

2.2 A Taxonomy of Information Hiding Techniques

As discussed in Section 1, a classification of information hiding techniques was proposed by Bauer [6], as seen in Figure 1. The main sub-disciplines of information hiding are *covert channels*, *steganography*, *anonymity*, and *watermarking*. Next, we give a brief definition of each of these sub-disciplines.

Covert Channels: Lampson introduced and illustrated many covert channels in his Confinement Problem paper [55]. This technique was defined in the context of multilevel secure systems (e.g. military computer systems), in which communication paths which were not designed to transfer information at all. Instead, these channels were used by untrustworthy programs to leak information to their owner while performing a service for another program. For instance, let us suppose that a program A (service) processes some information for a program B (customer). The customer program will want to ensure that the service program cannot access (read or modify) any of its data except those items to which it explicitly grants access. In addition, the service program must be protected from intrusion by the customer program, since the service program may have its own private data. Such communications channels have been studied to find ways to confine these programs, i.e., some secure systems need to safeguard data from unauthorized access or modification or programs from unauthorized execution [33].

Steganography: Steganography (from the Greek *steganos*, or "covered," and *graphie*, or "writing") is the hiding of a secret message within an ordinary message and its extraction at a destination. Steganography takes cryptography one step further by hiding an encrypted message so that no one suspects its existence. Ideally, anyone scanning the data would fail to know that it contains encrypted data [70]. While cryptography is about protecting the content of messages, steganography is about preventing the recognition of their very existence [77]. Steganographic techniques may be roughly classified into linguistic steganography and technical steganography. The former

consists of linguistic or language forms of hidden writing. A widely used method of linguistic steganography is the acrostic. One of the most famous examples is Giovanni Boccaccio's *Amorosa visione*, which is called "the world's hugest acrostic." [78]. The latter method, such as invisible ink, tries to hide messages physically. The most famous examples go back to antiquity, as discussed in Section 2.1. However, in recent years, with the move towards digitalization, messages can be embedded into digital media using steganographic techniques, and transmitted through the Internet rapidly. There are many papers in the literature that cover the main steganographic techniques and the basic concepts as well [46, 42, 77, 2, 58, 43].

Anonymity: The main purpose of this technique is to find ways to hide the meta-content of messages, i.e., the sender and the recipients of a message [14, 34]. The basic idea is that one can obscure the trail of a message by using a set of remailers or routers as long as the intermediaries do not collude. As a result, trust remains the cornerstone of these tools [15, 103]. There are different varieties, depending on who is *anonymized*: the sender, the receiver, or both. Web applications have focused on receiver anonymity, whereas e-mail users are concerned with sender anonymity [46]. Many references on anonymity and traffic analysis can be found in [2].

Watermarking: Digital watermarking, as opposed to steganography, has the additional requirement of robustness against possible attacks [77]. The purpose of digital watermarking is to provide evidence that can be used within the legal system to prove that some copyright violation has occurred. On the other hand, digital watermarking alone is not sufficient protection against any kind of attack, and should not be considered a panacea for protecting copyrights on digital content [67]. In this sense, the meaning of *robustness* is still not very clear, as it depends on the application. Robustness has strong implications in the overall design of a watermarking system. Braudaway et al. [11] pointed out that watermarks do not need to be hidden, as some systems use visible digital watermarking; however, most of the literature has focused on imperceptible (invisible, transparent, or inaudible) digital watermarks, which have wider applications. There are three fundamental differences between steganography and watermarking [46]: (a) the information hidden by a watermarking system is always associated with the digital object to be protected or its owner, whereas steganographic systems simply hide any information; (b) the *robustness* criteria are also different, since steganography is mainly concerned with the detection of the hidden message, while watermarking is concerned with removal by a pirate; and (c) steganographic communication is usually point-to-point (between the sender and the receiver), whereas watermarking techniques are usually one-to-many.

2.3 The Application Domain of Information Hiding

Placing data in image, video and audio files is useful in a variety of applications. In this subsection, we highlight some applications driving interest in the subject of information hiding.

Unobtrusive Communication: the process of transmitting a cryptographic message may attract unwanted attention, and the use of cryptographic technology may be restricted or forbidden by law [58].The use of steganography does not advertise the cover communication and therefore avoids scrutiny concerning the sender, message, and recipient. In military and intelligence agencies, even if the content of a message is encrypted, the detection of a signal in a modern battlefield may lead to a sudden attack on the signaler. Military communications require techniques such as read spectrum modulation or meteor scatter transmission to make signals hard for the enemy to detect or jam [77, 24].

Copyright Protection: a secret watermark can be embedded inside multimedia data to identify it as intellectual property [90, 108]. When such data are sold or distributed, an identification of the recipient and time stamp can be embedded to make piracy difficult [58]. Moreover, a watermark or signature can be used as means of tracing the distribution of digital objects for an on-line news service and for photographers who sell their work for digital publication [7]. A requirement for a digital watermarking is that it must be difficult to remove, since information about legal ownership is to be included in such data.

Feature Tagging: descriptive elements, such as annotations, captions, and time stamps can be embedded inside digital objects. For instance, in an image database, keywords can be embedded to facilitate search engines. Another example is that if an image is a frame of video sequence, timing markers can be embedded into the image for synchronization with audio [58]. In general, tagging is embedded in multimedia data for later retrieval. The retrieval process starts with a search inside the data, trying to find watermarked objects in the search for a typical watermarking pattern. All objects will then be checked with an equivalent retrieval process, and the embedded information is retrieved. This application does not have the same requirements for robustness as the digital watermark. It can be assumed that since feature location is providing a service, it is unlikely that someone will maliciously try to remove the encoded information [7].

Tamper Proofing: one purpose of the information hidden in a digital object is to prevent or detect unauthorized modifications [30, 56]. Moving from the issue of robustness to that of destruction, tamper-proofing refers to the difficulty an attacker will face in altering or forging a message once it has been embedded into a digital object. These techniques must be resilient to small modifications (e.g. cropping, tone-scale, or balance or equalization for sounds), but not to large modifications such as removal from or insertion into an image, or even taking words out of context in an audio recording [7, 104].

Broadcast Monitoring: a system for automatic audit radio transmission can use a computer to listen to a radio station and look for marks, which indicate that a particular piece of a song in an advertisement has been broadcast [8]. In general, systems for broadcast monitoring are developed to prove the feasibility of professional broadcast surveillance. Other applications include verification of commercial transmissions, assessment of sponsorship effectiveness, protection against illegal transmission, statistical data collection, and analysis of broadcast content [95].

Data Augmentation: sometimes information is added for the benefit of the public. For instance, details about the work, annotations, other types of channels, or even purchasing information (e.g. the nearest shop, the price, the producer, etc) [32]. On one hand, data augmentation may be used for advertising and marketing purposes. For example, when listening to the radio, someone can simply press a button to order a compact disk. On the other hand, this technique can be applied to hide information used to index pictures or music tracks with the purpose of providing more efficient retrieval from databases [41, 78].

Other Applications: law enforcement and counter intelligence agencies are interested in information hiding technologies in order to detect and trace hidden information [77]. Another niche in which information hiding can be applied is related to schemes for digital elections and digital cash by making use of anonymous communications techniques. The marketing industry is also taking advantage of information hiding techniques by using forgery techniques to send out enormous numbers of unsolicited messages while avoiding any kind of reply from users [43, 41].

So far we have highlighted some potential areas in which information hiding may emerge as a useful solution. More details about these techniques and their applications can be found in [46, 42, 27, 7]. In the remainder of this paper, we shall focus on the specific topic of digital watermarking in the context of information hiding.

3 Digital Watermarking

Watermarking is the process of electronically attaching the identity (secret, analogue or digital) of the owner of a copyright in a way that is difficult to erase. This process is comparable to placing an electronic stamp on the document. In some cases, the word "watermark" is used exclusively for a hidden identity code [88].

This new research area is a multidisciplinary field that combines image and signal processing with cryptography, communication theory, coding theory, signal compression, and the theory of visual perception. Interest in this field has recently increased because of the wide spectrum of applications it addresses.

In this section, we review digital watermarking principles. First, we present a terminology for digital watermarking and then we describe a basic framework. We then outline several properties (or characteristics) of watermarks which might be desirable for various applications.

3.1 Terminology

Although digital watermarking has gained a lot of attention and has evolved very quickly, there is no consensus on the terminology used in this field. As a result, it is necessary to clarify some definitions in order to avoid any ambiguity.

Unlike steganography, watermarking has the additional characteristic of robustness against attacks. While steganographic data can usually be removed, an additional requirement for watermarking is that this is not possible, even if the algorithmic principles are known. Even if the existence of the hidden information is known, it is hard for an attacker to destroy the embedded watermark without destroying the data itself or without knowledge of a key [46, 22]. In cryptography, this is known as Kerkhoff's law: a cryptosystem should be secure, even if an attacker knows the cryptographic principles and methods used, but does not have the appropriate key [65]. Therefore, steganography and watermarking are complementary sub-disciplines rather than competitive approaches.

Watermarking differs from authentication or digital signatures, which prove to a receiver that the message could only have come from one particular transmitter. Usually, authentication messages are in the form of conventional hash functions that can easily be deleted by a pirate who wishes to use copyrighted material for illegal purposes. The goal is to give the copyright owner of a digital image (or other piece of information) the possibility to prove technically the origin of the image. Watermarking does not address authentication explicitly [66].

Fingerprinting refers to a particular characteristic of an object which tends to distinguish it from another similar one. Although fingerprinting has various applications (e.g. information retrieval), in digital watermarking this technique can be used for copyright protection of data. Fingerprinting does not rely on tamper-resistance, and as a result, it does not prevent users from making copies of data; however, fingerprinting enables the owner to trace authorized users distributing copies illegally [88, 66]. Fingerprinting is a special case of watermarking applications.

Bitstream watermarking is also a term used for watermarking applied to compressed data, such as compressed video [82]. For video broadcast applications, watermarking schemes operating on compressed video are desirable. For instance, a robust watermark can be embedded into an MPEG bitstream

without increasing the bit-rate and can be retrieved even from the decoded video without knowledge of the original, unwatermarked video [36].

"Embedded signatures" was a widely used term in early publications instead of watermarking, but it is usually not used anymore because it sometimes leads to confusion with cryptographic signatures. Cryptographic signatures are used for authentication purposes in the sense that they detect any alteration of the signed data and authenticate the sender, while watermarks are only used for authentication in special applications and are usually designed to resist alterations and modifications [46].

3.2 Basic Framework

The process of digital watermarking involves modifying the original data to embed a *watermark* containing key information, such as authentication or copyright codes. The embedding method must keep the original information data perceptually unchanged, and the watermark data should be detected by an extraction algorithm [45, 51].

Watermarking methods share the same generic scenarios used for hiding messages: a watermark embedding system and a watermark recovery system [39]. The latter is also called watermark extraction or watermark decoding.

Figure 2 shows the stages of a watermark embedding scheme, based on the scheme proposed by [102]. The input to this scheme is the original data (also called cover-data) and an optional public or secret key. This key may be used to enforce security measures, such as prevention of unauthorized access to the watermark through recovery and manipulation. The original data can be of any nature, including audio, video, images, formatted text, or 3D models [46]. In general, practical systems employ at least one key, or even a combination of several keys. In combination with a secret or a public key, watermarking techniques are referred to as secret and public watermarking techniques [72].

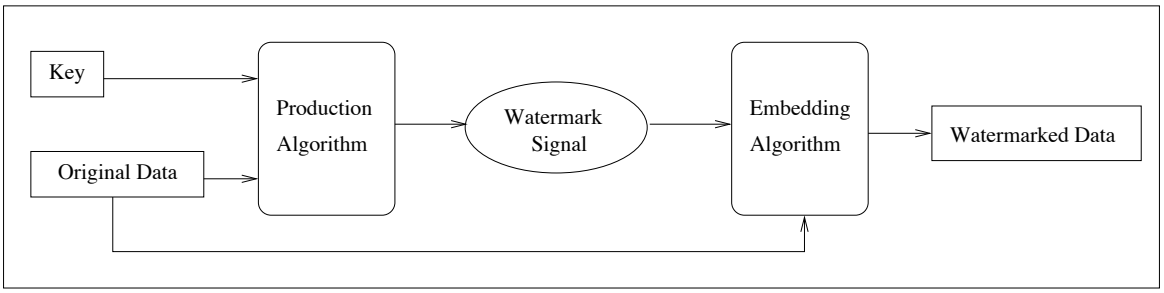

Figure 2: Stages of a Watermark Embedding Scheme

In the first stage of the watermark embedding scheme, the watermark production algorithm is applied. Its input parameters are the key and the original data. This algorithm transforms each pair (*key*, *original data*) into a watermark digital signal. By doing so, the algorithm achieves the statistical invisibility of the watermark. Generally, this production algorithm is based on pseudo-random number generators or chaotic systems [102, 100]. For this reason, the determination of a key which produces a predefined watermark signal is impossible, as this is a non-invertible process developed to prevent unauthorized parties from defining counterfeit watermarks.

Regarding the second stage, a watermark embedding algorithm is used. This algorithm requires as parameters both the original data and the produced watermark. The watermark embedding scheme in a digital product consists of producing alterations in the luminance of the pixels (e.g. video and images), and alterations in the sonority (e.g. audio) [66]. These alterations are designed to take into account the main characteristics of the human visual and auditory system [101]. Watermarking embedding

algorithms can be grouped into two classes: transform domain methods [17, 73], which embed the data by modulating the transform domain coefficients and spatial domain techniques [10, 98], which embed the data by directly modifying the pixel values of the original image. Finally, the output of the watermark embedding scheme is the watermarked data.

A generic watermark detection scheme is depicted in Figure 3. Detection is the trickiest part in the watermarking framework. The inputs of this scheme are watermarked data, the secret or public key, and, depending on the method, the original data [52].

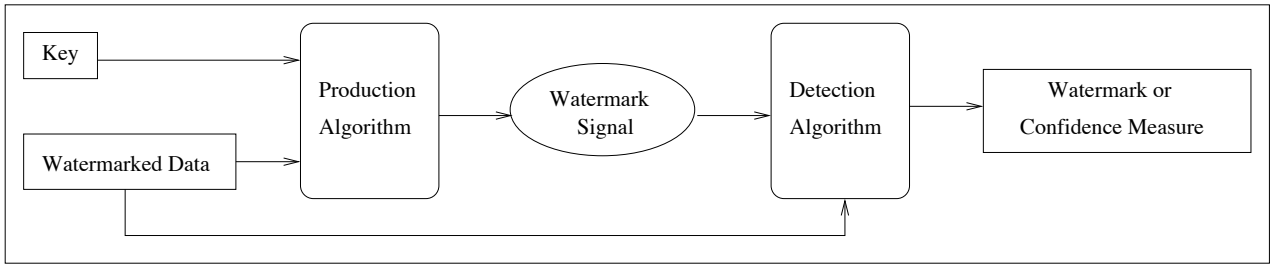

Figure 3: Stages of a Watermark Detection Scheme

Like watermark embedding scheme, a watermark detection scheme uses a production algorithm which generates a watermark signal in the first stage. The purpose of this stage is to feed the detection procedure, which is the most important stage of the detection scheme.

The next stage, the detection process, can be achieved by using a watermark detection algorithm. This algorithm should be trustworthy, producing an insignificant number of false alarm errors. Consequently, the detector should present results concerning the ownership of a digital product. This kind of algorithm relies on statistical hypothesis testing [102, 101]. For this reason, the output of this scheme is either the recovered watermark or some kind of confidence measure which provides a certain indication for a given watermark at the input under inspection.

According to Kutter and Hartung [52], two types of watermarking systems can be identified based on the nature and combination of inputs and outputs: private watermarking (also called non-blind watermarking), and public watermarking (also called blind or oblivious watermarking). The former systems extract the watermark from the watermarked data using the original data as a hint to find where the watermark is. The latter systems remain the most challenging because they require neither the secret original data nor the embedded watermark. We discuss the basic idea behind these watermarking systems in Section 4.3.

Depending on the application, the input to embedding and detection schemes can be either in the form of uncompressed or compressed data. For instance, in a watermarking scheme for video, the best input option is in an uncompressed form since this process reduces computational costs with decoding and re-encoding procedures [52].

3.3 Desirable Properties

There are some important properties desirable in a watermark design. These properties are related to difficulty of detection, robustness to common distortions, resilience to malicious attacks, support of a sufficient data rate commensurate with the application, flexibility to allow multiple watermarks to be embedded, and others [18, 66]. The relative importance of these properties depend on the application. Some of these properties are discussed in the following section.

Imperceptibility: This property, also called fidelity or invisibility, means that the watermark should not be noticeable to the viewer nor to the human auditory system. In addition, the watermark

should not degrade the quality of the content [66]. The embedded information is transparent and it is independent of the application and purpose of a watermarking system. Moreover, a transparent watermark does not create any artifacts or quality loss.

Robustness: The embedded information is said to be robust if it is still available intact after the image has been modified but not destroyed beyond recognition [88]. In other words, the watermark must be difficult to remove. Watermarks in image, video and audio files should resist any kind of distortion introduced by standard or malicious data processing. For instance, in image and video, it is desirable that the watermark survives geometric distortions such as translations, scaling, and cropping. The term "robustness" is usually used in the sense of resistance to non-targeted modifications or common image operations [29].

Security or Key Restrictions: In applications such as copyright protection, the secrecy of the embedded information has to be assured. If secrecy is a requirement, a secret key has to be used for the embedding and extraction process. Security is a property that describes whether the embedded watermarking information cannot be removed beyond reliable detection by targeted attacks. Such attacks are based on a full knowledge of the embedding algorithm and the detector, except the key, and the knowledge of at least one watermarked data [27]. Two levels of secrecy can be identified. In the first level, an unauthorized user can neither read nor decode an embedded watermark. The second level allows any user to detect if data are watermarked, but the embedded data cannot be read without the secret key [46].

Computational cost: In commercial applications, the computational costs of encoding and decoding are decisive. Depending on the application, the insertion is only done once and can be performed off-line. As a result, the cost of encoding plays a less important role than the cost of decoding. For instance, the cost of decoding may have to occur at real-time video or audio rates. So the speed requirements are highly dependent on the application. For this reason, computational requirements compel a watermark to be designed in a simple form. On the other hand, this simplicity may reduce the resistance to tampering significantly. Further, scalability is another relevant issue related to computational cost. Therefore, it is desirable to design a watermark whose decoder and/or inserter is scalable to each generation of computers [18, 66].

Modification and Multiple Watermarks: In some applications, it is necessary to alter the watermarks after the insertion process. A typical example is the case of a DVD that may be watermarked to allow for only a single copy to be made of it. Once this copy has been made, it is then necessary to alter the watermark on the original disc for the purpose of discouraging further copies. To do so, one of the following alternatives can be applied. The first watermark can be removed and a new one inserted. This alternative does not allow a watermark to be tamper-resistant, as it is easily removable. On the other hand, multiple watermarks facilitate the tracking of content from manufacturing to distribution to eventual sale. A second watermark can be inserted so that both are readable, but one must override the other [18].

False Alarm Probability: In some applications, it is necessary to distinguish between data that contain watermarks and data that do not. Even in the absence of attacks or signal distortions, the probability of failing watermark detection (false-negative error probability) and detecting a watermark when, in fact, one does not exist (false-positive error probability), must be very small to allow the use of watermarking systems as unambiguous evidence of ownership on a legal basis. To accomplish this, statistically-based algorithms can be used [60].

Data Payload: This characteristic is related to the amount of information that a watermark contains. In general, methods of storing data express data payload as a number of bits that indicate the number of distinct watermarks that might be inserted into a signal. For instance, in a watermark that carries N bits, there are 2^N different possible watermarks. However, there are $2^N + 1$ possible values returned by a watermark detector because there is the possibility that no watermark is present. [66]. Depending on the application, the watermark algorithm should allow a predefined number of bits to be hidden. This number of bits is not unlimited, and very often is fairly small because the higher the number of bits embedded, the more perceivable the watermark will be [18, 62].

4 A Watermarking Taxonomy

A wide range of watermarking techniques have been developed for images, audio and video sources, and various security services. Most of these techniques have been designed to address copyright protection [23]. Our purpose is not to present an exhaustive review of all the techniques available in the literature. Instead, we describe some categories of digital watermarking in order to show how broad this topic is. For more details about watermarking techniques, the reader can refer to [46, 42, 66, 27, 7].

In this section our objective is to provide a watermarking taxonomy based on several recent publications. The reason for this arrangement is to provide a general view of several key areas of watermarking, such as applications area, domain of insertion, and types of existing algorithms, as can be seen in Figure 4.

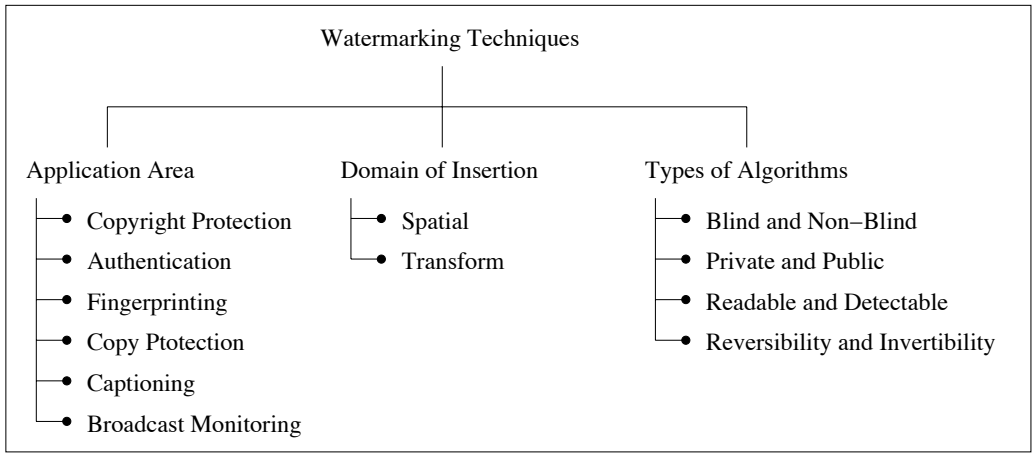

Figure 4: A watermarking taxonomy

4.1 Application Area

Watermarking techniques initially used for digital imagery and now also used for audio and 3D-models are relatively new and are growing at an exponential rate [54]. Such techniques can be categorized in a number of classes on the basis of their distinct applications, as follows. We do not provide many details about these watermarking applications since we have discussed most of them in Section 2.3.

Copyright Protection: This is the most prominent application of watermarking today. Watermarks are used to prove rightful ownership, so this application requires a very high level of robustness.

The idea is that the embedded data can be used as a proof of digital ownership in case of a copyright dispute or even to trace the image recipient that produced unauthorized copies [72].

Authentication (or Tamper-proofing): In authentication applications, the goal is to detect modifications of the data. In other words, in this kind of application, the watermark encodes the information required to determine whether the content is authentic. In contrast to copyright protection applications, data inserted for authentication purposes should be fragile in the sense that they should be modified when the data are manipulated. This can be achieved with *fragile watermarks* that have a low robustness to certain modifications (e.g. compression) and are impaired by other modifications [87, 112].

Fingerprinting for Traitor Tracking: Watermark techniques can also be used to identify the content byers. This kind of application has the purpose of conveying information about the legal recipient rather than the source of digital data in order to identify single distributed copies of data. For instance, such monitoring is useful to trace back illegally produced copies of data that may circulate and is very similar to the use of serial numbers for software products. This kind of application is called fingerprinting and involves the embedding of a different watermark into each distributed copy [52, 66].

Copy Protection: This type of application requires a watermark that contains information about the rules of usage and copying that the content owner wishes to enforce. A desirable feature in multimedia distribution systems is the existence of a copy protection mechanism which disallows unauthorized copying of the media. Copy protection is not easy to achieve in open systems, while in closed or proprietary systems it is feasible [52, 66]. DVD systems are an example in which the data contain copy information embedded as a watermark. In this case, the content may not be copied, or may be copied only once [59, 9].

Captioning (or Annotation): An invisible watermark may contain meta-information about the content in which it is embedded. In other words, captioning embeds descriptive information within images for applications, such as the labeling and annotation of medical data and video indexing. For example, a song on the radio could contain embedded information about the artist, album, recording date, etc. The amount of embedded information in this case is moderately large [26, 72].

Broadcast and Publication Monitoring: As with signatures, in this application, a watermark identifies the owner of the content, but it is different in the sense that the detection is performed by automated systems which are able to monitor distribution channels to track when and where the content appears [8]. Two important cases in which broadcast and publication monitoring can be applied are: a) for owners who wish to ensure that their multimedia data are not being illegally distributed; b) advertisers who wish to ensure that their commercials are being broadcast at the times and locations they have purchased [66].

4.2 Domain of Insertion

Although existing literature on watermarking embedding techniques deals with image, video and audio files, most of publications refer to images. Based on recent publications, we roughly divided watermarking techniques into two classes taking, into account the domain in which the watermark signal is embedded.

The first class refers to spatial domain techniques [10, 98] which embed the data by directly modifying the pixel values of the original image. The other class lays on transform domain methods [17, 73] which embed the data by modulating the transform domain coefficients.

At the end of this section, we show the extensions and adaptations of methods from still images to video and audio.

Spatial Domain Watermarking: Spatial watermarks are constructed in the image spatial domain and embedded directly into an image's pixel data [17, 73, 109]. Considering Kerckhoffs' principle [23], a watermarking algorithm should be public and the watermark should not be accessible in a straightforward manner in order to prevent casual removal. This is accomplished by the choice of cover-location where the watermark information will be embedded. For this reason, in this section, we cover the choice of host locations in the original (cover) image considering cryptographic and psychovisual aspects.

1. *Patchwork's Algorithm*: Proposed by Bender et al. [7], this algorithm does not as such allow a message to be hidden in a cover, but a secret key is used to initialize a pseudorandom number generator which outputs the locations of the cover which will host the watermark. In this insertion process, the owner selects n-pixel pairs pseudorandomly according to a secret key, K_s. The luminance values (a_i, b_i) of n pairs of pixels are modified by using the following formula:

$$\overline{a}_i = a_i + 1$$

$$\overline{b}_i = b_i - 1$$

   Thus, the owner simply adds 1 to all values a_i and subtracts 1 from every b_i. Regarding the extraction process, the n-pixel pairs which were used in the encoding step to host the watermark are retrieved, again using the secret key, K_s. So the sum

$$S = \sum_{i=1}^{n} \overline{a}_i - \overline{b}_i$$

   is computed. If the cover actually contained a watermark, we can expect the sum to be $2n$; otherwise, it should be approximately zero. This extraction is based on the statistical assumption

$$E[S] = \sum_{i=1}^{n} E[a_i] - E[b_i] = 0$$

   if someone randomly chooses several pairs of pixels in an image, assuming that they are independent and identically distributed. Consequently, only the owner who knows the modified locations can obtain a score close to $S \approx 2n$.

2. *Public key cryptography and public watermark recovery*: Watermarking algorithms based on a secret key have a drawback – they do not allow public recovery of the watermark. One alternative to such a limitation is to use public key watermarking algorithms. These algorithms consist of two keys: a public and a private one. One digital document can be watermarked using the private key, whereas the public key is used to verify the mark.

   Hartung and Girod [37] proposed the idea of deriving a public key algorithm from spread spectrum methods, as we will briefly discuss at the end of this section. The basic idea is that the direct sequence technique requires the spread sequence, S, for both spreading and unspreding processes. However, due to the robustness of the encoding, it is possible to reconstruct the original signal without knowledge

of the whole spread sequence. Thus, the secret key known only by the owner of the digital document allows the whole S to be computed, while the public key allows only a part of S (S^{pub}) to be computed so that the watermark is kept robust. The public spread S^{pub} sequence presents one bit per N equal to the original S^{orig} sequence, whereas all other bits are chosen in a random manner:

$$S^{pub} = \begin{cases} S_i^{orig} & with\ probability\ \frac{1}{N} \\ rand\{-1, 1\} & otherwise \end{cases}$$

3. *Predictive coding for psychovisual watermark management*: Predictive models are widely used in source coding to predict the new value of a signal from its former values [23]. This technique is also relevant for watermark systems because it does not require the original image. The assumption is that samples or pixels in a neighborhood are highly correlated. This case can be applied, especially in images in which neighboring pixels have related values. Thus, predictive coding consists of computing errors between predicted and original values before encoding these errors in an efficient way. Actually, it is expected that the distribution of errors will be be close to zero with a small variance, in which case an efficient binary representation may be obtained using Huffman codes, for instance. From the watermarking point of view, predictive coding is useful for psychovisual reasons. It is well known that the human visual system is less accurate in textured and edge regions, which makes the zones very suitable for watermark locations [64]. On the other hand, human eyes are very sensitive to smooth regions with uniform values, which turn out to be poor candidates for watermark locations. In predictive coding, the error distribution perfectly matches these properties, and the error signal can be used as a modulation carrier for the watermark signal.

Transform Domain Watermarking: Transform domain watermarking techniques apply some invertible transforms to the host image before embedding the watermark. The transform domain coefficients are then modified to embed the watermark and, finally, the inverse transform form is applied to obtain the marked digital object. Transform domain techniques are much more robust against compression and geometrical transformations than spatial domain techniques [51, 50]. Basically, the process of transform domain watermarking involves the following techniques.

1. *Discrete Fourier Transform (DFT)*: The discrete Fourier transform is widely studied in signal processing. This technique has been considered in the field of watermarking to offer the possibility of controlling the frequencies of the host signal. It is helpful to select adequate parts of the digital data for embedding the watermark in order to obtain the best compromise between visibility and robustness [66, 75].
   Given a two-dimensional signal $f(x, y)$, the DFT is defined to be [23]:

$$F(k_1, k_2) = \beta \sum_{n1=0}^{N_1-1} \sum_{n2=0}^{N_2-1} F(n_1, n_2)\ exp(-i2\pi n_1 k_1/N_1 - i2\pi n_2 k_2/N_2)$$

with $\beta = (N_1, N_2)^{-1/2}$ and $i = \sqrt{-1}$. In addition, the inverse DFT (IDFT) is given by

$$f(n_1, n_2) = \beta \sum_{k1=0}^{N_1-1} \sum_{k2=0}^{N_2-1} F(k_1, k_2)\ exp(i2\pi k_1 n_1/N_1 + i2\pi k_2 n_2/N_2)$$

The DFT is useful for watermarking purposes to perform phase modulation between the watermark and its cover. However, the DFT is used more often in derived forms such as the discrete cosine transform or the Mellin-Fourier transform.

2. *Discrete Cosine Transform (DCT)*: The DCT was widely studied by the source coding community in the context of JPEG's and MPEG's [84, 74]. This techniques was also considered to embed a message inside images [49] and videos [64]. The main arguments for using DCT in watermarking are the following: (a) embedding rules operating in the DCT domain are often more robust to JPEG and MPEG compression so that the watermark designer can prevent JPEG or/and MPEG attacks more easily. (b) studies on visual distortions that were previously conducted in the field of source coding can be reused. These studies help to predict the visible impact of the watermark on the cover-image. (c) watermarking in the DCT domain offers the possibility of directly realizing the embedding operator in the compressed domain to minimize the computation time. In digital image processing, the two-dimensional version of the DCT is given by

$$S(u,v) = \frac{2}{N} C(u)C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} s(x,y)\ cos(\frac{\pi u(2x+1)}{2N})\ cos(\frac{\pi v(2y+1)}{2N})$$

$$s(x,y) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u)C(v)S(u,v)\ cos(\frac{\pi u(2x+1)}{2N})\ cos(\frac{\pi v(2y+1)}{2N})$$

As previously mentioned, JPEG and MPEG use the Discrete Cosine Transform to achieve image compression. The compressed data are stored as integers, but the calculations for the quantization process require floating point calculations which are rounded. Information can be hidden by manipulating the rounding values (either up or down) of the JPEG/MPEG coefficients. The advantage of this techniques is that it minimizes the block-like appearance resulting when the boundaries between 8 x 8 sub-images become visible (known as blocking artifacts), and the disadvantage is that the computational complexity of DCT is $O(nlogn)$ [42, 47, 29].

3. *Mellin-Fourier Transform*: Most watermarking algorithms have problems in extracting the watermark after a geometric transformation has been applied on the watermarked object. Ó Ruanaidh and T. Pun proposed the use of the Mellin-Fourier transform to overcome this drawback [86]. The transform space of Mellin-Fourier is based on the translation property of the Fourier transform:

$$f(x_1 + a, x_2 + b) \leftrightarrow F(k_1, k_2)\ exp[-i(ak_1 + bk_2)]$$

In this case, only the phase is altered by a translation. As a consequence, if the workspace (i.e. the space in which the watermark will be embedded) is limited to the subspace related to the amplitude of the Fourier transform, it will be insensitive to a spatial shift of the picture. In order to become insensitive to rotation and zoom, log-polar mapping (LPM) can be used, as follows [23]:

$$(x,y) \mapsto \begin{cases} x = exp(\rho)\ cos(\theta) \\ y = exp(\rho)\ sin(\theta) \end{cases} \quad with\ \rho \in R\ and\ \theta \in [0, 2\pi]$$

Thus, the rotation of any element (x, y) in the cartesian coordinate system will result in a translation in the logarithmic coordinate system. Similarly, a zoom will result in a translation in the polar coordinate system.

4. *Discrete Wavelet Transform (DWT)*: The multiresolution aspect of wavelets is helpful in managing a good distribution (i.e. location) of the message in the cover in terms of robustness versus visibility. The wavelet transform consists of a multiscale spatial-frequency decomposition of an image [61]. One way to construct these different levels of resolution is to cascade two-channel filter banks and use a down-sampling process. The two-channel filter banks, which must be orthogonal, are defined by the equations:

$$H(\omega) = \sum_k h_k \; exp(-jk\omega) \qquad high \; pass$$

$$G(\omega) = \sum_k g_k \; exp(-jk\omega) \qquad low \; pass$$

Then the iterative process of the decomposition is given by

$$c_{j-1,k} = \sum_n h_{n-2k} \; c_{j,n} \; ; \qquad d_{j-1,k} = \sum_n g_{n-2k} \; c_{j,n}$$

whereas the iterative reconstruction process is defined by

$$c_{j,n} = \sum_k h_{n-2k} \; c_{j-1,k} \; + \; \sum_k g_{n-2k} \; d_{j-1,k}$$

There are some different approaches in using wavelet transform in watermarking. For instance, Wang and Kuo [105] proposed a multitreshold wavelet coding scheme allowing significant coefficient searching. The authors assumed that these coefficients do not change much after several signal processing operations. In addition, if these coefficients lose their fidelity significantly, the reconstructed image could be perceptually different from the original one. In contrast to the methods which select a predefined set of coefficients (e.g. low-frequency coefficients), the resulting method is image dependent. Thus, this method is suitable for textured and smooth images as well. Kundur et al. [50] described a watermarking method using wavelet-based fusion. It consists of adding the wavelet coefficients of the watermark and the image at different resolution levels. Prior to being added, the wavelet coefficients of the watermark are modulated using a human-model constraint based on a measure called saliency. Other schemes based on DWT can be found in [107, 111, 93].

5. *Spread Spectrum*: Spread spectrum techniques for watermarking have received special attention [36]. The reason for this is very similar to the arguments for using spread spectrum techniques in steganography. In general, the message used to watermark is a narrow-band signal, as compared to the wide band of the cover digital data. Spread spectrum techniques applied to a message allow the frequency bands to be matched before transmitting the message (watermark) through the covert channel (digital data). Spread spectrum techniques also offer the possibility of protecting watermark privacy using a secret key to control a pseudonoise generator. The two main spread spectrum methods are *direct-sequence spread spectrum* and *frequency-hopping as a spread spectrum* [80]. The former consists of a time modulation of the original signal using a wide-band pseudonoise signal. The resulting signal looks like the pseudonoise signal. At the receiver, the original signal is reconstructed by demodulation of the receiver signal using the same pseudonoise signal. The latter uses a random process in which the carrier frequency is altered, describing a wide range of frequency values. As a result, the modulated signal has a wide spectrum. In both approaches, the resynchronization between the receiver signal and the random signal is the main issue of the recovery process. This requires knowledge of the random process used during the modulation. Dugelay et al. [23] showed that, in particular, for watermarking issues, synchronization

problems occur after geometrical attacks are performed on digital data. Hartung et al. [36] proposed a practical watermarking scheme that can be applied to images and videos.

Let a_j, $(a_j \in \{-1, 1\})$, be a binary signal. From a_i, a signal b_i that is a temporal stretching of a_i can be derived

$$b_i = a_j, \qquad j\ cr \leq i < (j+1)\ cr$$

where cr is the chip rate. Thus, a modulation is performed between b_i and a pseudonoise p_i in order to obtain the watermark signal which will be directly embedded in the image v_i,

$$w_i = \alpha\ b_i\ p_i$$

where α is a strength factor controlling the robustness versus visibility trade-off. The final embedding formula is

$$\overline{v}_i = v_i + \alpha\ b_i\ p_i$$

where $\overline{v}_i$ denotes the watermarked data. The basic idea behind watermark recovery is to demodulate the received signal and to add each signal component corresponding to each piece of binary information:

$$s_j = \sum_{i=jcr}^{(j+1)cr-1} p_i\ \overline{v}_i = \sum_{i=jcr}^{(j+1)cr-1} p_i\ v_i \quad + \sum_{i=jcr}^{(j+1)cr-1} p_i^2\ \alpha\ b_i$$

Assuming that the p_i signal is zero-mean and is statistically independent of v_i, one can expect s_j to be $s_j \approx cr\ \alpha\ a_j$. Thus, a_j is given by $a_j = sign(s_j)$.

Most of the techniques that we have mentioned previously can be applied to video as well. To do so, the general trade-off (ratio, visibility, and robustness) included in a watermarking algorithm is significantly modified [23]. The modification of the ratio between the information contained in a given watermark and in the host signal becomes less critical. However, due to the temporal dimension, the visual distortion is even more difficult to manage. Consequently, the list of possible attacks increases. Another important issue is that the problem of the complexity of computation is now primordial when real-time watermarking is desired. Most research on watermarking has mainly concerned still images, so that video and audio watermarking remains an open problem. In any case, most of the basic ideas defined for still images can also be used for moving pictures. For instance, Barni et al. [4] introduced a method based on the addition of DCT coefficients from the host signal and the watermark which can be realized on JPEG streams as well as on MPEG streams. However, some problem related to the temporal dimension of videos have to be overcome.

Regarding audio watermarking, signal processing operations are frequently applied to the host audio. Operations that modify the host signal also modify the embedded data. Most audio data embedding techniques are based on spread spectrum methods [92]. Some other approaches replace least significant bits or spectral components of the host signal to embed data.

Several techniques have been proposed. For instance, Bender et al. [7] proposed a techniques using a phase coding approach in which data are embedded by modifying the phase values of the Fourier transform coefficients of audio segments. The authors also proposed embedding data as spread spectrum noise.

Another audio data embedding technique was proposed by Tilki and Beex [96]. In this technique, Fourier transform coefficients over the middle frequency bands (e.g. 2.4 kHz to 6.4 kHz) are replaced with spectral components from a signature sequence. The middle frequency band was selected so that

the data remain outside of the more sensitive low frequency range. The signature is of short time duration and has a low amplitude relative to the local audio signal.

Swanson et al. [92] proposed a perceptual audio data embedding technique designed to be flexible, i.e., it is able to embed data rates that range from low to high, depending on the application. The algorithm employs the projection of audio frequency sub-bands onto a pseudo-random direction dictated by a secret key. The projection is followed by a non-linear quantization step to avoid the need for the original audio signal during the extraction. Moreover, the detection process includes a sophisticated searching mechanism to properly synchronize with the embedded data without access to the original audio signal. The data embedding algorithm is designed to be robust to many distortions.

Other techniques for audio watermarking based on the Fourier transform, wavelet transform, and spread spectrum can be found in [16, 83, 69, 91, 110].

4.3 Types of Existing Algorithms

Watermarking techniques can be also classified into a number of classes according to existing algorithms. This classification, in which some concepts such as blindness, privateness, readability, and reversibility were analyzed in different scenarios, was first proposed by Bartolini et al. [5]. In this section we focus on the main ideas presented in [5].

Blind and Non-Blind: A watermarking algorithm is said to be blind (or oblivious) if it requires neither the secret original nor the embedded watermark. In other words, it does not require the comparison between the original data and the watermarked data. On the contrary, a watermarking algorithm is said to be non-blind if it needs the original data to extract the information contained in the watermark. Some authors refer to blind techniques as oblivious or private techniques [19, 18]. Blind algorithms are, in general, less robust than non-blind ones since the true data in which the watermark is hidden are not known and must be treated as disturbing noise [5]. Considering that in the real world the original data are not always warranted, non-blind algorithms may become useless for certain applications. For this reason, this type of algorithm is not useful to prove rightful ownership [19]. For some authors [52], there is one more algorithm which can be classified between the blind and non-blind types. This the semi-blind algorithm, which does not use the original data for detection and also can be applied to prove ownership and copy control in applications such as DVD's.

Private and Public: A watermark is said to be private if only authorized readers can detect it. This type of watermark has a mechanism that makes it impossible for unauthorized people to extract the watermark. A private watermark uses blind algorithms, since blind techniques are by themselves private and only authorized users can access the original data required to read the watermark [19]. Bartolini et al. [5] extended the concept of privateness to all techniques that use any mechanism to deny the extraction of the watermark to unauthorized users. So privateness may be achieved by assigning to each user a different pseudo-random key, knowledge of which is necessary to extract the watermark from the host document [46, 18]. Unlike private watermarking, public watermarking is a technique that allows anyone to read the watermark. The decision to use private or public watermarks depends on the application they are applied to [57]. For instance, suppose a user downloads data over the Internet and he/she wants to know the owner of this multimedia data. If a private scheme was used to insert the watermark, there is no way to read the owner's name, whereas this would be possible if a public watermark has been used instead.

Readable and Detectable: There is another important distinction between watermarking algorithms used to embed a code that can be read and those in which a mark has been inserted which can

be only detected. In the former case, the bits of information contained in the watermark can be read without knowing them in advance. So this type of watermark is called readable. In the latter case, one can only verify whether a given code is present in the multimedia document. For this reason, this is a detectable watermark [19, 52]. Detectable watermarking techniques are intrinsically private, since it is impossible for an attacker to guess the content of the watermark without knowing anything about it [5].

Reversibility and Invertibility: Watermarking is said to be reversible if once a watermark has been read (detected) it can be removed from the multimedia data. In other words, to successfully demonstrate rightful ownership, non-invertibility of the watermark has to be granted. On the other hand, a watermark is said to be invertible if it is possible to generate a false watermark and a fake original multimedia document which is perceptually similar to the true one, so that by embedding the false watermark in this false multimedia document, a document which is equal (invertibility) or perceptually equal to the true one is obtained. As discussed in [19], invertible watermarking schemes are likely to be of little use in many practical applications.

5 Watermarking: Progress, Limitations, and Challenges

In this section, we investigate the technological advances in digital watermarking techniques and discuss some limitations which demonstrate that some levels of robustness in watermarking techniques are still inadequate when compared with recent attacks. As a result, some open problems arise which remain as challenges and opportunities for further work.

5.1 Technological Progress

Until recently, information hiding techniques, notably steganography and digital watermarking techniques, received much less attention from the research community and from industry than cryptography [77]. However, this scenario has changed rapidly, especially since 1996 when the first conference on digital watermarking was organized.

Today these techniques are also spreading to a number of applications dealing with ownership and authentication for digital documents including not only images, but also audio and video [54]. For this reason, digital watermarking has emerged as the leading candidate to solve the difficult problem of providing guarantees to creators and consumers of digital content concerning the protection of their legal rights [26].

To accomplish this, recent watermarking techniques have provided robust mechanisms which have led to advances in the design and construction of secure systems [40, 42]. Such techniques have been used in great number of commercial applications [19], as well as other scenarios, as discussed in Section 5.1. Therefore, the interest in watermarking technology is high, both in academia and industry.

The interest from academia is reflected in the number of publications on watermarking and in the number of conferences on watermarking and data hiding being held. On the other hand, the interest from industry is easily seen in the large number of companies that have been founded in the last few years [52].

Apart from research and business activities, various international research projects have been carried out [46]. For instance, a goal of the European Community is to develop practical watermarking techniques. This community has conducted relevant projects on digital watermarking, such as TALISMAN[1] (Tracing Authors' Rights by Labelling Image Services and Monitoring Access Network), whose

[1] Information available at http://www.cordis.lu/esprit/src/talisman.htm

purpose is to provide European community service providers with a standard copyright mechanisms to protect digital products against large-scale commercial piracy and illegal copying. The resulting product of this project is a system for protecting video sequences through labeling and watermarking [21]. OCTALIS[2] (Offer of Content Through Trusted Access Link) [81] is the followup project of TALISMAN and OKAPI[3] (Open Kernel for Access to Protected Interoperable Interactive Services) [35], whose main goal is to integrate a global approach to equitable conditional access and efficient copyright protection and to demonstrate its validity on large-scale trials on the Internet and EBU (European Broadcast Union).

International standardization consortia are also interested in watermarking techniques. For example, the emerging video-compressing standard MPEG-X (ISO/IEC 14496) provides a framework which allows integration with encryption and watermarking [52]. The DVD industry standard is looking for a means of holding copy control and copy protection mechanisms that use watermarking to signal the copy status of multimedia data [9].

In spite of these advances and the recent technological progress, digital watermarking is still not a mature and understood technology, so a lot of questions have not been answered yet. In addition, the theoretical fundamentals are still weak, and a large number of systems are designed heuristically [46, 53].

5.2 Watermarking Limitations

Expectations for watermarking techniques should be realistic since watermarking systems deal with a trade-off between robustness, watermark data rate (payload), and imperceptibility. A robust watermark which resists all attacks is not realistic [52, 54]. Even when designed under realistic expectations, watermarking techniques are still vulnerable to attacks by experts. The robustness criteria vary from one system to another, and recent attacks have shown that all the levels of robustness are still inadequate [77, 76]. Attacks on watermarks may not necessarily remove the watermark, but disable its readability. Although several commercially available watermarking scheme are robust to many type of attacks, these are often not robust to combinations of basic transformations, such as scaling, cropping and rotation. JPEG compression, additive Gaussian noise, low-pass filtering, rescaling, cropping and rotation have been addressed in most literature as basic attacks [76].

An attacker may try to estimate the watermarking and subtract it from a marked image. The reason is simple - such an attack is extremely dangerous if the attacker can find a generic watermark, for example, one with $W = F(S_0, W)$ not depending significantly on the image S_0. The estimation W of the watermark can be used to remove a watermark from any arbitrary marked data, without any further effort [66]. For instance, the attacker may separate the watermark W by adding or averaging multiples watermarking images, such as multiple successive marked images $S_0 + W$, $S_1 + W$, ..., $S_N + W$ from a video sequence. The addition of N images results in $NW + \sum S_i$, which tends to NW for a large N and a sufficient number of independent images S_0, S_1, ..., S_N [9, 53].

Craver et al. [20] identified three classes of attacks, taking into account the way in which the attacks affect the watermarking technology. The first class is *robustness attacks* which aim to weaken or remove the presence of the digital watermark. *Presentation attacks* fit in the second class. Attacks categorized in this class do not remove the watermark, but modify the content so that the detector can no longer find or extract it anymore. The last class, in which an attacker tries to forge invalid or multiple interpretations from watermark evidence, is called *interpretation attacks*.

Attacks that try to defeat a watermarking system are described next. Our purpose is to present

[2]Information available at http://www.octalis.com/

[3]Information available at http://www.tele.ucl.ac.be/OKAPI/index.html

the basic concepts behind these attacks. We do not discuss implementation considerations nor the mathematical background of these attacks, but these issues are fully covered in [46, 42, 66, 27, 77, 76].

The Uncorrelated noise attack: this is a robustness attack in which a random noise is added to the content in an attempt to distort the watermark beyond any reasonable level of usefulness [26]. In other words, the watermark is attacked by adding a random value to each pixel. The random values must be small enough so that they do not make the image appear grainy [20].

The Collusion attack: consider a watermarking scheme in which an image has been watermarked many times under different secret keys. The simple average of all these copies starts to resemble the original image, but it does not contain any useful watermarking data [68]. A Collusion attack is also a kind of robustness attack.

The Inversion attack: When an attacker has complete knowledge of how a watermark was embedded into multimedia data, he or she can detect the watermark and reverse the insertion process to remove the watermark entirely. To prevent this kind of attack, various watermarking systems use a secret key that describes exactly the position where the watermark is embedded in the image [26]. An Inversion attack is another type of robustness attack.

The Mosaic attack: this is a presentation attack that is quite general and that possesses the noticeable property that one can remove the watermark from one image and still have it presented in exactly the same way (pixel by pixel) as the image given by a standard browser [76, 77]. This kind of attack is also called a chopping attack in the sense that an image is chopped into a number of distinct sub-images, such as tiles or strips made up of the content of the original image. The original image can be restored by viewing the sub-images one after another in the appropriate configuration, for instance, a web page [26, 27].

The StirMark attack: this attack belongs to the class of robustness attacks. It was designed by the research group at the University of Cambridge. The StirMark simulates image distortions that commonly occur when a picture is printed, photocopied, and rescanned. The image is slightly stretched and compressed by random amounts, and a small amount of noise is added to simulate the quantization errors of A/D (Analogic to Digital) and D/A (Digital to Analogic) conversion. The main part of the SirMark is the small geometrical change which causes loss of synchronization between the watermark detector and the image. For low-frequency watermarks, small geometrical deformations can cause large differences in DCT coefficients [27]. Given a watermarking scheme, one can invert a distortion that will prevent detection of the watermark while leaving the perceptual value of the previously watermarked object undiminished [77].

The classification of these three groups may be not very clear yet. For instance, the StirMark both diminishes the watermark and distorts the content to fool the detector. So it can be defined as a robustness attack and as a presentation attack as well.

None of these classes of attacks is easy to defend against. Nevertheless, research is progressing toward the design of watermarking systems that can be successful in resisting many known attacks. On the other hand, as watermarking technology will evolves, attacks on watermarking will as well. Therefore, protecting unwatermarked content against attacks is a relevant issue for further research.

5.3 Open Problems and Challenges

Although a variety of target applications (copyright, authentication, copy protection, etc) and host signals exist (audio, image and video), there are many instances in which watermarking techniques can

provide working and successful solutions. However, the unlimited nature of attacks and the trade-off between visibility and robustness are major challenges of watermarking.

In this section we enumerate some open problems related to watermarking techniques with the purpose of providing an overview of the most recent challenges of this technology. We classify the challenges into the following groups: image, video, and audio. In addition, we present the fourth group, called multimedia, which contains challenges that fit into one or more the previous groups.

Group 1: Image

Tamper-proofing: the easiest way to implement tamper-proofing is to encode a check-sum of the image within the image. This method is triggered by small changes in images, and one alternative for this involves a overlaying pattern on the image. The challenge is to find a pattern resilient to simple modifications such as filtering. This search for patterns and other methods of detecting tampering remain an active area of research [28, 7].

Oblivious secure watermarking: this technology still presents changes in the gray levels due to filtering, lossy compression, and simple geometric transformations (e.g. shift, scaling, rotation and cropping). In cases of nonlinear geometric transformation such as the StirMark, the synchronization of the watermark detector becomes a difficult problem due to the computationally intensive nature of the task. The challenge is to design a robust watermark with computationally efficient detector that extract watermarks from images which have undergone general geometric distortions [27].

Region-Based Watermarking for Images: another interesting issue in watermarking is to investigate a mechanism for generating regions, and then modifying the distributions of each segment. The key benefits of this type of method are: a) flexibility - watermarks can be inserted in irregularly shaped regions; b) invariance to geometric attacks - watermarks can be recovered with resistance to rotation, scaling and cropping; c) detection without pre-processing - the image can be tested without detecting the amount of change performed. The main problem with this method as presented by Brisbane et al. [12, 13], is the lack of capacity of the watermark. It needs to be improved so as to allow watermarking to be orthogonal, colorless, and intensity invariant.

Group 2: Video

Synchronizing the watermark detector: the complexity of the embedding and detection algorithms is also an important issue, in particular for video watermarking. The development of a theoretical foundation for digital watermarking is necessary to build reliable systems and to understand the fundamental limits of the approach. Synchronizing the watermark detector is still an open problem. Efficient synchronization algorithms must be investigated with the purpose of analyzing their accuracy and influence on detection reliability. Different (coded) modulation schemes for watermark embedding should be compared regarding their complexity and achievable watermark bit rate. The proper choice of the watermarking domain (e.g. spatial, DFT, DCT or wavelet domain) has to analyzed further [25].

Video Watermark: Su et al. [89] proposed a video watermark technique in a novel content-dependent spatially localized framework in which the watermark energy is concentrated in sub-frames with desirable properties, and the sub-frame locations are synchronized using visual content rather than structural markers. They showed that this video watermarking method has a robustness to geometric distortions. The method is distinguished by its ability to be embedded and extracted using

frame-based algorithms, while resisting collusion. There are some directions for future research that are extensions of their work, including improving performance against scaling and aspect ratio changes, enhancing the sub-frame selection algorithm, e.g., incorporating key-dependence to improve secrecy, and applying state-of-the-art image watermarking ideas at the sub-frame level, e.g., turbo codes.

Video and Audio Object Processing and Coding for Databases: Broadcasting systems are making increasing use of computers and networks, and the advantages of computer-supported image processing can be seen in TV productions today. Multicasting on the Internet and video-on-demand systems suggest that future broadcasting stations will be using video servers [31]. In this context, video databases will play an important role in broadcasting systems as an alternative to tapes, and VCR's (Video Cassette Recorder) or as an archive for data storage. To get the most out of a video database, it is necessary to improve the image handling processing. Challenges include: a) studying object-based coding, which works well with a future production system where images will be handled as video components; b) investigating video segmentation methods which make use of information obtained in the studio or program production stage; c) researching digital video and audio watermarking techniques for coding the meta-data of video objects, as well as investigating the technical requirements of watermarking for broadcast use.

Group 3: Audio

Audio Watermarking: many of the current audio data embedding techniques are based on spread spectrum techniques and are inherently projection techniques on a given direction. Swanson et al. [92] showed in their work that, ideally, a larger projection value will indicate the presence of a binary symbol. To reduce the probability of a false detection, the length of the audio segment and pseudo-random direction are increased to reduce the chances of a high correlation between the original host signal and the pseudo-random sequence. As the size increases, the amount of data that can be embedded in the host signal decreases. The challenge here is that future audio data embedding algorithms should avoid the overused spread spectrum/matched filter approach. Another important open issue is that data algorithms are likely to implement active control over audio clips and use more sophisticated signal dependent keys.

Group 4: Multimedia

Robustness requirements: Researchers are always attempting to define the robustness of copyright marking systems, since robustness depends on the application. The goal is to find a marking scheme that discourages piracy operations, by tracking the use of recordings by broadcasters, copy control and other methods. For this problem, the challenge is to achieve a marking scheme which does not affect the sonic quality of the sound recording even after operations such as filtering (conversion D/A and A/D), compression, adding additive noise, and others. This scheme should be resistant so that there is no way to remove or alter the embedded information without degradation of the sound quality [23].

Trade-off between robustness and imperceptibility: the two most important properties of a digital watermarking are robustness and imperceptibility. There have been many efforts aimed at designing a watermarking structure to maximize its robustness and imperceptibility. Currently, the degree of success is based on heuristic arguments, but the results are not convincing so far. This trade-off remains one of the most important technical challenges in the field of watermarking research [88].

Protocols for Multimedia Distribution: the advent of electronic commerce and the creation of electronic distribution channels for multimedia content have brought new challenges regarding the protection of intellectual property. A particularly promising method is the use of digital watermarking to embed additional copyright information within multimedia data. Tomsich and Katzenbeisser [97] pointed out that watermarking alone is not sufficient to resolve rightful ownership of digital data; a protocol relying on the existing public-key infrastructure (which is also used for digital signatures) is necessary. One of the vulnerabilities if the existing protocols for multimedia distribution relies on the watermarking algorithm, in the sense that most known watermarking systems are sensitive to intentional distortions of the digital data and do not merge the digital data and the watermark completely, as copy attacks show.

Watermark data rate (payload): the data payload of a watermarking method has two important aspects: the number of distinct watermarks that may be inserted and the number of watermarks that may be detected by single iteration with a given watermark detector. In many watermarking applications, each detector does not need to test for all the watermarks that might possibly be present. For instance, many companies are interested in setting up web-crawlers that look for their watermarks on the web. Although the number of distinct possible watermarks has to be at least equal to the number of companies, each crawler could test for as few as one single watermark. So the challenge is that a watermarking for such an application is supposed to have a payload of many bits, but not all bits are available from any given detector [66].

6 Related Work

A number of efforts have been made to investigate state-of-art in information hiding, such as [77, 58, 7]. Roche and Dugelay mentioned in [85] that only two papers dealing with watermarking appeared in 1992. Six years after, this number increased to 106. The information available today is spread over countless papers and conference proceedings. In this paper, for instance, we have collected over 100 references as well.

As information hiding is evolving rapidly, new attempts are constantly being made to dscribe the current state of research. For instance, Cox et al. [18] outlined the desirable properties of watermarks whose purpose is to encode copyright information. These properties depend on the intended use of the watermark: copyright control, authentication, etc, as can be seen in the Section 2.3. After emphasizing the importance of perceptual modeling, and introducing a framework for watermarking, they then discussed some techniques presented in early papers.

Voyatzis et al. [99] described a general framework for image copyright protection through digital watermarking. In particular, they presented the main features of an efficient watermarking scheme, discussed robustness issues, and described the three main stages of a watermarking algorithm, namely watermark generation, embedding, and detection.

The work of Nikolaidis and Pitas [72] is an extension of [99]. The authors identified a number of distinct watermarking application areas and presented an overview of typical watermarking schemes (embedding and extraction). Moreover, they outlined some attacks and described some research results related to watermarking technology.

In the same direction, Ferrill and Moyer [26] outlined digital watermarking and surveyed the state of digital watermarking research, highlighting the technical and legal problems that must be solved before digital watermarking can be widely used.

Lin and Delp [58] also studied the use of data hiding techniques in digital images. In particular, they described how one can use steganography to hide information in a digital image. In this paper, the

authors investigated recent developments in data hiding, specifically pertaining to copyright protection of digital images.

A classification of digital watermarking techniques based on existing algorithms was proposed by Bartolini et al. [5]. This classification was aimed at describing the way the decoding process works. In the second part of the paper, the authors presented an object-oriented Electronic Copyright Management System (ECMS) relying on the state-of-the-art technology, with the purpose of providing enforcement of copyright laws in an open-network environment.

Petitcolas et al. [77] gave an overview of information hiding with a focus on steganography. They looked at a range of applications and tried to place the various techniques in an historical context. The authors described a number of attacks on information hiding systems and some of the problems in constructing a general theory and the practical requirements that marking schemes and steganography systems may have to meet.

In our work, we investigate the status, limitations, and prospects of digital watermarking, and we place watermarking technology in the context of hiding information. Our work differs from related work in the following aspects. First, we present a watermarking taxonomy based on the domain of insertion, area of applications, and types of existing algorithms. Second, we discuss the advances concerning watermarking techniques and investigate the recent technical progress. Finally, we categorize some watermarking attacks and discuss how they can affect modern watermarking systems. We believe that this discussion is useful in that it points out some limitations of watermarking and directions for future work, besides providing a wealth of material one could use as an introduction to this area of research.

7 Conclusions

In the first part of this paper, we presented a brief overview of information hiding, describing its main disciplines (covert channels, steganography, anonymity, and digital watermarking) and some areas in which this technology can be applied. In the rest of our work, we concentrated our attention on digital watermarking in particular. We provided an overview of the current status of and prospects for digital watermarking, emphasizing its basic principles. In addition, we described a taxonomy based on domain of insertion, applications area, and types of existing algorithms. We also offered an overview of the advances and limitations, which led us to some interesting topics for future research.

Our work showed that even though a wide variety of watermarking techniques have been proposed in the last few years, it is quite difficult to classify the approaches and assess their quality. This is due to the fact that there are various watermarking methods which provide robustness, but different levels of required robustness can be identified, depending on the application. As a result, there is no generic watermarking method which fits all kind of systems. Furthermore, none of the techniques proposed so far seem to be robust to all possible attacks and image operations.

This study also revealed that current watermarking technologies face three important drawbacks that prevent them from becoming the best solutions for protecting multimedia data. The first drawback was discussed previously. It is clear that watermarking techniques are still vulnerable to attacks. The second problem is that there are no watermarking technology standards in place so the community has not accepted this technology as a legitimate form of multimedia data protection. The third drawback concerns legal problems for managing content distribution through the Internet. On the other hand, there have been many efforts to create applications in which watermarking can provide successful solutions, such as copy protection for Internet audio distribution, broadcast monitoring and others. This fact may assure that watermarking technology will become widely deployed in the near future.

In terms of technical progress, we can see some great advances. The recent explosion in research in watermarking to protect intellectual property is evident. As a consequence, a large amount of money has

been spent on copyright protection systems, so there is a continuous increase in watermarking research. This implies that one can expect that sooner or later, digital watermarking will provide efficient solutions to the difficult problem of providing security to creators and consumers of digital content regarding their legal rights.

Given all these facts it seems that digital watermarking will play an increasing role in the future for security in the digital world. The progress in watermarking may suggest that this technology will achieve the goals that have been set for it.

8 Acknowledgments

Stanley Oliveira was partially supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) of Ministry for Science and Technology of Brazil, under Grant No. 200077/00-7. Mario Nascimento and Osmar Zaïane were partially supported by a Research Grant from NSERC, Canada. We wish to acknowledge the comments and suggestions of Dr. Xiaobo Li that helped us improve this article.

References

[1] R. J. Anderson, R. Needham, and A. Shamir. The Steganographic File System. In *IWIH: International Workshop on Information Hiding*, April 1998.

[2] R. J. Anderson and F. A. P. Petitcolas. Information Hiding and Digital Watermarking: An Annotated Bibliography. Last update: August 1999. Computer Laboratory at the University of Cambridge, UK. Availabel at. http://www.cl.cam.ac.uk/~fapp2/steganography/bibliography/.

[3] R. J. Anderson and F. A. P. Petitcolas. On The Limits of Steganography. *IEEE Journal on Selected Areas in Communications (J-SAC), Special Issue on Copyright and Privacy Protection*, 16(4):474–481, 1998.

[4] M. Barni, F. Bartolini, V. Cappellini, and A. Piva. Robust Watermarking of Still Images for Copyright Protection . In *Proceedings 13th International Conference on Digital Signal Processing*, volume 2, pages 499–502, Santorini, Greece, July 1997.

[5] F. Bartolini, G. Bini, V. Cappellini, A. Fringuelli, G. Meucci, and A. Piva. Enforcement of Copyright Laws for Multimedia Through Blind, Detectable, Reversible Watermarking. In *IEEE International Conference on Multimedia Computing and Systems, Vol. 2*, pages 199–203, Florence, Italy, June 1999.

[6] F. L. Bauer. *Decrypted Secrets - Methods and Maxims of Cryptology.* Berlin, Heidelberg, Germany: Springer-Verlag, 1997.

[7] W. Bender, D. Gruhl, N. Morimoto, and A. Lu. Techniques for Data Hiding. *IBM Systems Journal*, 35(3 and 4):313–336, 1996.

[8] D. Blagden and N. Johnson. Broadcast Monitoring: a Practical Application of Audio Watermarking. In *Proceedings of SPIE Vol. 3657, Security and Watermarking of Multimedia Contents but Withdrawn*, pages 3657–20, January 1999.

[9] J. A. Bloom, I. J. Cox, T. Kalker, J. P. M. G. Linnartz, M. L. Miller, and C. B. S. Traw. Copy Protection for DVD Video. *Proceedings of the IEEE (USA)*, 87(7):1267–1276, July 1999.

[10] A. G. Bors and I. Pitas. Image Watermarking using DCT Domain Constraints. In *Proceedings of the IEEE International Conference on Image Processing*, volume 3, pages 231–234, Lausanne, Switzerland, September 1996.

[11] G. Braudaway, K. Magerlein, and F. Mintzer. Protecting Publicly-Available Images with a Visible Image Watermark, Proc. SPIE, vol. 2659, pp.126-133, 1996.

[12] G. Brisbane, R. Safavi-Naini, and P. Ogunbona. Region-Based Watermarking for Images. In *International Workshop on Information Security, M. Mambo & Y.Zheng (eds)* , pages 154–166, Kuala Lumpur, Malaysia, November 1999.

[13] G. Brisbane, R. Safavi-Naini, and P. Ogunbona. Region-Based Watermarking by Distribution Adjustment. In *International Workshop on Information Security*, pages 54–68, Wollongong, NSW, Australia, December 2000.

[14] D. Chaum. Untraceable Electronic Mail, Return Addresses and Digital Pseudonyms. *Communications of the ACM*, 24(2):84–88, 1981.

[15] D. Chaum. The Dining Cryptography Problem: Unconditional Sender and Recipient Untraceability. *Journal of Cryptography*, 1(1):65–75, 1988.

[16] M. Cooperman and S. Moskowitz. Steganographic Method and Device. U.S. Patent no. 5,613,004, 1997.

[17] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon. Secure Spread Spectrum Watermarking for Multimedia. *IEEE Transactions on Image Processing*, 6(12):1673–1687, 1997.

[18] I. J. Cox and M. L. Miller. A review of watermarking and the importance of perceptual modeling. Vol. 3016. In *Proc. SPIE Human Vision and Elect. Imaging II, vol. SPIE, vol. 3016*, pages 92–99, San Jose, CA, February 1997.

[19] S. Craver, N. Memon, B. L. Yeo, and M. M. Yeung. Resolving Rightful Ownerships with Invisible Watermarking Techniques: Limitations, Attacks, and Implications. *IEEE Journal of Selected Areas in Communications*, 16(4):573–586, 1998.

[20] S. Craver, B. Yeo, and M. Yeung. Technical Trials and Legal Tribulations. *Communications of the ACM*, 41(7):45–54, 1998.

[21] J. F. Delaigle, J. M. Boucqueau, J. J. Quisquater, and B. Macq. Digital Images Protection Techniques in a Broadcast Framework: An Overview. Technical report, TALISMAN Project. ACTS project AC019, Catholic University of Louvain, Brussels, Belgium, 1996.

[22] J. Dittmann, A. Steinmetz, and R. Steinmetz. Content-Based Digital Signature for Motion Pictures Authentication and Content-Fragile Watermarking. In *IEEE International Conference on Multimedia Computing and Systems, Vol. 2*, pages 209–213, Florence, Italy, June 1999.

[23] J. L. Dugelay and S. Roche. A Survey of Current Watermarking Techniques, Chapter 6 of "Information Hiding: Techniques for Steganography and Digital Watermarking", S. Katzenbeisser and F. A. P. Petitcolas (eds.), Norwood, MA: Artech House, 2000, pp.121-148.

[24] D. L. Schilling (ed.). *Meteor Burst Communications: Theory and Practice.* Wiley Series in Telecommunications, New York: J. Wiley and Sons, 1993.

[25] J. J. Eggers and B. Girod. Watermark Detection after Quanization Attacks. In *Proceedings Third International Workshop on Information Hiding (Lecture Notes in Computer Science, Springer. 4). Andreas Pfitzmann, Ed*, pages 172–186, Dresden, Germany, September 1999.

[26] E. Ferrill and M. Moyer. A Survey of Digital Watermarking. February 1999. Available at. http://www.cc.gatech.edu/~mjm/dw/watermarking.html.

[27] J. Fridrich. Applications of Data Hiding In Digital Images. In *Tutorial for The ISPACS Conference*, Melbourne, Australia, November 1998.

[28] J. Fridrich. Image Watermarking for Tamper Detection. In *Proceedings of the IEEE International Conf. on Image Processing (ICIP '98), vol. 2*, pages 404–408, Chicago, Illinois, USA, October 1998.

[29] J. Fridrich, A. C. Baldoza, and R. J. Simard. Robust Digital Watermarking Based on Key-Dependent Basis Functions. In *Proceedings of the 2nd Information Hiding Workshop*, pages 15–17, Portland, Oregon, April 1998.

[30] G. L. Friedman. The Trustworthy Digital Camera: Restoring Credibility to the Photographic Image. *IEEE Transactions on Consumer Electronics*, 39(4):905–910, November 1993.

[31] K. Fukui. The future evolution of digital broadcasting, 2001.

[32] M. Gerzon and P. Graven. A High-Rate Buried-Data Channel for Audio CD. Journal of the Audio Engineering Society, vol. 43, no. 1/2, Jan.-Feb. 1995, pp.3-22.

[33] V. Glicor. A Guide to Understanding Covert Channels Analysis of Trusted Systems. Technical report, NCSC-TG-030, National Computer Security Center, Fl. George G. Meade, Maryland, USA, November 1997.

[34] D. M. Goldschlag. Hiding Routing Information. In *Proceedings of the 1st International Workshop (Lectures Notes in Computer Science, 1174)*, pages 137–150, Isaac Newton Institute, Cambridge, England. Ed. Berlin, Germany: Springer-Verlag, May 1996.

[35] J. Guimaraes, J. M. Boucqueau, and B. Macq. OKAPI : A kernel for access control to MM services based on Trutesd Third Parties. In *Proceedings of European Conference on Multimedia Applications, Services and Techniques (ECMAST'96)*, pages 783–798. Louvain-la-Neuve, Belgium, May 1996.

[36] F. Hartung and B. Girod. Digital Watermarking of MPEG-2 Coded Video in the Bitstream Domain. In *Proc. ICASSP 97*, volume 4, pages 2621–2624, Munich, Germany, April 1997.

[37] F. Hartung and B. Girod. Fast Public-Key Watermarking of Compressed Video. In *International Conference on Image Processing (ICIP'97)*, volume I, pages 528–531, Santa Barbara, California, U.S.A., 1997.

[38] Herodotus. *The Histories*. London, England: J. M. Dent and Sons, Ltd, 1992.

[39] A. Herrigel, J. O' Ruanaidh, H. Petersen, S. Pereira, and T. Pun. Secure Copyright Protection Techniques for Digital Image. In *Information Hiding*, pages 169–190, (D. Aucsmith, ed.), vol. 1525 of Lecture Notes in Computer Science, (Berlin), April 1998.

[40] N. F. Johnson. An Introduction to Watermark Recovery from Images. In *Proceedings of the SANS Intrusion Detection and Response Conference (IDR'99)*, San Diego, CA, February 1999.

[41] N. F. Johnson. In Search of the Right Image: Recognition and Tracking of Images in Image Databases, Collections, and the Internet. Technical report, George Mason University, Center for Secure Information System, June 1999.

[42] N. F. Johnson, Z. Duric, and S. Jajodia. *Information Hiding: Steganography and Watermarking - Attacks and Countermeasures.* Kluwer Academic Publishers, 2000.

[43] N. F. Johnson and S. Jajodia. Exploring Steganography: Seeing the Unseen. *Computer*, 31(2):26–34, 1998.

[44] D. Kahn. *The Codebreakers - The Story of Secret Writing.* New York, NY, USA: Scribner, 1996.

[45] S. Kang and Y. Aoki. Image Data Embedding System for Watermarking Using Fresnel Transform. In *IEEE International Conference on Multimedia Computing and Systems, Vol. 1*, pages 885–889, Florence, Italy, June 1999.

[46] S. C. Katzenbeisser and F. A. P. Petitcolas (eds.). *Information Hiding Techniques for Steganography and Digital Watermarking.* Norwood, MA: Artech House, 2000.

[47] H. Kii, J. Onishi, and S. Ozawa. The Digital Watermarking Method by Using both Patchwork and DCT. In *IEEE International Conference on Multimedia Computing and Systems, Vol. 1*, pages 895–899, Florence, Italy, June 1999.

[48] M. Kobayashi. Digital Watermarking: Historical Roots. Technical report, RT0199, IBM Research, Tokyo Research Laboratories, Japan, April 1997.

[49] E. Koch and J. Zhao. Towards Robust and Hidden Image Copyright Labeling. In *Proc. of the IEEE Workshop on Nonlinear Signal and Image Processing, IEEE Computer Society Press*, pages 452–455, Los Alamitos, California, 1995.

[50] D. Kundur and D. Hatzinakos. A Robust Digital Image Watermarking Method using Wavelet-Based Fusion. In *IEEE Signal Processing Society 1997 International Conference on Image Processing (ICIP'97)*, pages 544–547, Santa Barbara, California, October 1997.

[51] D. Kundur and D. Hatzinakos. Digital Watermarking Using Multiresolution Wavelet Decomposition. In *International Conference on Acoustic, Speech and Signal Processing (ICASP)*, volume 5, pages 2969–2972, Seattle, Washington, USA, 1998.

[52] M. Kutter and F. Hartung. Introduction to Watermarking Techniques, Chapter 5 of "Information Hiding: Techniques for Steganography and Digital Watermarking", S. Katzenbeisser and F. A. P. Petitcolas (eds.), Norwood, MA: Artech House, 2000, pp.97-120.

[53] M. Kutter and F. A. P. Petitcolas. Fair Benchmarking for Image Watermarking Systems. In *Proceedings of the SPIE, Security and Watermarking of Multimedia Contents*, volume 3657, pages 226–239, San Jose, CA, USA, January 1999.

[54] Jack Lacy, Schuyler R. Quackenbush, Amy R. Reibman, and James H. Snyder. Intellectual Property Protection Systems and Digital Watermarking. In *Information Hiding*, pages 158–168, 1998.

[55] B. W. Lampson. A Note on the Confinement Problem. *Communications of the ACM*, 16(10):613–615, October 1973.

[56] C. Y. Lin and S. F. Chang. Issues for Authenticating MPEG Video. In *Proceedings of the SPIE 3657, Security and Watermarking of Multimedia Contents*, pages 54–65, November 1999.

[57] C.Y. Lin, J. A. Bloom, I. J. Cox, M. L. Miller, and Y. M. Lui. Rotation, Scale, and Translation Resilient Public Watermarking for Images. In Ping Wah Wong and Edward J. Delp, editors, *Security and Watermarking of Multimedia Contents II*, volume 3971, pages 90–98. Society of Photo-optical Instrumentation Engineers (SPIE), 2000.

[58] E. T. Lin and E. J. Delp. A Review of Data Hiding in Digital Images. In *Proceedings of the conference on Image processing, image quality, image capture systems PICS '99*, pages 274–278, Savannah, Georgia, USA, April 1999.

[59] J. P. M. G. Linnartz. The "Ticket" Concept for Copy Control Based on Embedded Signalling. *Computer Security - 5th European Symposium os Research in Computer Security (Lecture Notes in Computer Science)*, 1485:257–274, Springer 1998.

[60] M. Maes, T. Kalker, J. Haitsma, and G. Depovere. Exploiting Shift Invariance to Obtain a High Payload in Digital Image Watermarking. In *Proceedings of the IEEE International Conference on Multimedia Computing and Systems, Vol. 1*, pages 7–12, Florence, Italy, June 1999.

[61] S. Mallat. *A Wavelet Tour of Signal Processing*. London: Academic Press, 1998.

[62] C. G. Martin. Digital Images Watermarking Techniques. M.Sc. Thesis, Department of Computer Science at the Rochester Institute of Technology, Rochester, NY, USA. May 2000.

[63] L. M. Marvel. Image Steganography For Hidden Communication. Ph.D. Thesis, Department of Computer and Electrical Engineering at the University of Delaware, USA. May 1999.

[64] K. Matsui and K. Tanaka. Video-Steganography: How to Secretly Embed a Signature in a Picture. In *IMA Intellectual Property Project Proceedings*, volume 1, pages 187–205, January 1994.

[65] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. *Handbook of Applied Cryptography*. Fourth Printing. Boca Raton, Florida: CRC Press, 1999.

[66] M. Miller, I.J. Cox, J.P.M.G. Linnartz, and A.C.C. Kalker. A Review of Watermarking Principles and Practices", Chapter 18 of "Digital Signal Processing for Multimedia Systems", K.K. Parhi and T. Nishitani (eds.), Marcel Dekker, Inc., New York, March 1999, pp.461-486.

[67] F. Mintzer, G. W. Braudaway, and A. E. Bell. Opportunities for Watermarking Standards. *Communications of the ACM*, 41(7):57–64, 1998.

[68] F. Mintzer, J. Lotspiech, and N. Morimoto. Safeguarding Digital Library Contents and Users: Digital Watermarking, D-LIB on-line magazine, December 1997.

[69] D. Moses. Simultaneous Transmission of Data and Audio Signals by Means of Perceptual Coding. U.S. Patent no. 5,473,631, 1995.

[70] A. H. Murray and R. W. Burchfield (eds.). The Oxford English Dictionary: being a corrected re-issue, 1933.

[71] B. Newman. *Secrets of German Espionage*. London, England: Robert Hale Ltd, 1940.

[72] N. Nikolaidis and I. Pitas. Digital Image Watermarking: an Overview. In *IEEE International Conference on Multimedia Computing and Systems, Vol. 1*, pages 1–6, Florence, Italy, June 1999.

[73] J. Ohnishi and K. Matsui. Embedding a Seal into a Picture under Orthogonal Wavelet Transform. In *Proceedings of the IEEE International Conference on Multimedia Computing and Systems*, pages 514–521, Japan, June 1996.

[74] W. B. Pennebaker and J. L. Mitchell. *JPEG Still Image Data Compression Standard*. Van Nostrand Reinhold, New York, NY, USA, 1993.

[75] S. Pereira, J. O. Ruanaidh, F. Deguillaume, G. Csurka, and T. Pun. Template Based Recovery of Fourier-Based Watermarks Using Log-polar and Log-log Maps. In *IEEE International Conference on Multimedia Computing and Systems, Vol. 1*, pages 870–874, Florence, Italy, June 1999.

[76] F. A. P. Petitcolas and R. J. Anderson. Attacks on Copyright Marking Systems. In *Proceedings of the 2nd International Information Hiding Workshop*, pages 219–239, Portland, Oregon, USA, April 1998.

[77] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn. Information Hiding - A Survey. *Proceedings of the IEEE*, 87(7):1062–1078, 1999.

[78] Fabien A. P. Petitcolas. Introduction to Information Hiding, Chapter 1 of "Information Hiding: Techniques for Steganography and Digital Watermarking", S. Katzenbeisser and F. A. P. Petitcolas (eds.), Norwood, MA: Artech House, 2000, pp.1-14.

[79] B. Pfitzman. Information Hiding Terminology. In *Proceedings of the 1st International Workshop (Lectures Notes in Computer Science, 1174)*, pages 347–350, Isaac Newton Institute, Cambridge, England. Ed. Berlin, Germany: Springer-Verlag, May 1996.

[80] R. Pickholtz, D. Schilling, and L. Milstein. Theory of spread-spectrum communications- A Tutorial. *IEEE Transactions on Communications*, 30(5):855–884, 1982.

[81] L. Piron, M. Arnold, M. Kutter, W. Funk, and J. M. Boucqueau. OCTALIS Benchmarking: Comparison of Four Watermarking Techniques, in Security and Watermarking of Multimedia Contents, vol. 3657, P. W. Wong and E. J. Delp (eds.), San Jose, California, USA, January 2000, pp.240-250.

[82] C. Podilchuk and W. Zeng. Watermarking of the JPEG Bitstream. In Proc. of the International Conference on Imaging Science, Systems, and Technology, Las Vegas, Nevada, USA, pp.253–260, June 30 - July 3, 1997.

[83] R. D. Preuss, S. E. Roukos, A. W. F. Huggins, H. Gish, M. A. Bergamo, and P. M. Peterson. Embedded Signalling. U.S. Patent no. 5,319,735, 1994.

[84] K. Rao and P. Yip. Discrete Cosine Transform: Algorithms, Advantages, Applications. New York: Academic Press. 1990, 1990.

[85] S. Roche and J.-L. Dugelay. Image Waternarking Based on the Fractal Transform. In *Proceedings of the Workshop on Multimedia Signal Processing, IEEE*, pages 358–363, Los Angeles, CA, USA, December 1998.

[86] J.J.K. Ó Ruanaidh and T. Pun. Rotation, Translation and Scale Invariant Digital Image Watermarking. In *Proceesings of the International Conference Image Processing (IEEE ICIP)*, pages 536–539, Santa Barbara, California, October 1997.

[87] M. Schneider and S. F. Chang. A Robust Content Based Digital Signature for Image Authentication. In *Proceedings of the IEEE International Conference on Image Processing (ICIP'96)*, pages 227–230, Lausanne, Switzerland, September 1996.

[88] J. K. Su and B. Girod. On the Robustness and Imperceptibility of Digital Fingerprints. In *IEEE International Conference on Multimedia Computing and Systems, Vol. 2*, pages 530–535, Florence, Italy, June 1999.

[89] K. Su, D. Kundur, and D. Hatzinakos. A Content-dependent Spatially Localized Video Watermark for Resistance to Collusion and Interpolation Attacks. In *Proceedings of the International Conference on Image Processing (ICIP01)*, page Applications ii, Thessaloniki, Greece, October 2001.

[90] M. Swanson, M. Kobayashi, and A. Tewfik. Multimedia Data Embedding and Watermarking Technologies. *Proceedings of the IEEE*, 86(6):1064–1087, June 1998.

[91] M. Swanson, B. Zhu, A. Tewfik, and L. Boney. Robust Audio Watermarking Using Perceptual Masking. *Signal Processing*, 66(3):337–355, 1998.

[92] M. D. Swanson, B. Zhu, and A. H. Tewfik. Current State of the Art, Challenges and Future Directions for Audio Watermarking. In *Proceedings of ICMCS'99 - IEEE International Conference on Multimedia Computing and Systems*, pages 19–24, Florence, Italy, June 1999.

[93] M.D. Swanson, B. Zhu, and A.H. Tewfik. Multiresolution scene-based video watermarking using perceptual models. *IEEE Journal on Special Areas in Communications*, 16(4):540–550, 1998.

[94] A. Tacticus. *How to Survive Under Siege - Aineias the Tactician.* Oxford, England: Clarendon Press, pp.84-90 and 183-193, Clarendon Ancient History Series, 1990.

[95] P. Termont, L. De Strycker, J. Vandewege, J. Haitsma, T. Kalker, M. Maes, A. Langell G. Depovere, C. Alm, and P. Norman. Performance Measurements of a Real-Time Digital Watermarking System for Broadcast Monitoring. In *IEEE International Conference on Multimedia Computing and Systems, Vol. 2*, pages 220–224, Florence, Italy, June 1999.

[96] J. Tilki and A. Beex. Encoding a Hidden Digital Signature onto an Audio Signal using Psychoacoustic Masking . In *Proceedings fo the 7th International Conference on Signal Processing Applications & Technology*, pages 476–480, Boston, USA, October 1996.

[97] P. Tomsich and S. Katzenbeisser. Copyright Protection Protocols for Multimedia Distribution Based on Trusted Hardware. In *Proceedings of (PROMS 2000) Protocols for Multimedia Systems*, pages 249–256, Cracow, Poland, October 2000.

[98] R. G. van Schyndel, A. Z. Tirkel, and C. F. Osborne. A Digital Watermark. In *Proceedings of the International Conference on Image Processing, IEEE, vol. 2*, pages 86–90, Austin, Texas, USA, 1994.

[99] G. Voyatzis, N. Nikolaidis, and I. Pitas. Digital Watermarking: an Overview. In *9th European Signal Processing Conference (EUSIPCO'98)*, pages 9–12, Island of Rhodes, Greece, September 1998.

[100] G. Voyatzis and I. Pitas. Chaotic Mixing of Digital Images and Applications to Watermarking. In *Proceedings of the European Conference on Multimedia Applications, Services and Techniques (ECMAST'96)*, volume 2, pages 687–695, Louvain-la-Neuve, Belgium, May 1996.

[101] G. Voyatzis and I. Pitas. Embedding Robust Logo Watermarks in Digital Images. In *Proceedings of the 13th International Conference on Digital Signal Processing (DSP'97)*, volume 1, pages 213–216, Louvain-la-Neuve, Belgium, July 1997.

[102] G. Voyatzis and I. Pitas. Protecting Digital Image Copyrights: a Framework. *IEEE Computer Graphics and Applications*, 19(1):18–24, Jan.-Feb 1999.

[103] M. Waidner. Unconditional Sender and Recipient Untraceability in Spite of Active Attacks. In *J.-J. Quisquater and J. Vandewalle, editors, Advances in Cryptology - EUROCRYPT 89, volume 434 of Lecture Notes in Computer Science*, pages 302–319, Springer-Verlag, 1990, 10-13 April 1989, April 1990.

[104] C. Wang, J. Hill, J. Knight, and J. Davidson. Software Tamper Resistance: Obstructing Static Analysis of Programs. Technical Report CS-2000-12, DEpartment of Computer Science, University of Virginia, USA, 12 2000.

[105] H.-J. Wang and C.-C.J. Kuo. Image Protection via Watermarking on Perceptually Significant Wavelet Coefficients. In *Proceedings of the IEEE Signal Processing Society 1998 Workshop on Multimedia Signal Processing*, pages 279–284, Los Angeles, California, USA, December 1998.

[106] J. Wilkins. *Mercury: on the Secret and Swift Messenger: Shewing, How a Man May With Privacy and Speed Communicate His Thoughts to a Friend at Any Distance.* London, England: Printed for Rich Baldwin, 2nd ed., 1694.

[107] T. Wilson, S. Rogers, and L. Meyers. Perceptual-Based Hyperspectral Image Fusion Using Multiresolution Analysis. *Optical Engineering*, 34(11):3154–3164, 1995.

[108] R. Wolfgang, C. Podilchuk, and E. Delp. Perceptual Watermarks for Images and Video. To appear in Proceedings of the IEEE, May 1999. A copy of this paper is available at http://www.ece.purdue.edu/ ace.

[109] R. B. Wolfgang and E. J. Delp. A Watermark for Digital Images. In *Proceedings of the IEEE International Conference on Image Processing (ICIP'96), vol. 3*, pages 219–222, Lausanne, Switzerland, September 1996.

[110] J. Wolosewicz. Apparatus and Method for Encoding and Decoding Information in Audio Signals. U.S. Patent no. 5,774,452, 1998.

[111] X.-G. Xia, C. G. Boncelet, and G. R. Arce. Wavelet Transform Based Watermark for Digital Images. *Optics Express*, 3(12):497–511, 1998.

[112] L. Xie and G. R. Arce. A Blind Wavelet Based Digital Signature for Image Authentication. In *Proceedings of the 9th European Signal Processing Conference (EUSIPCO'98)*, pages 21–24, Island of Rhodes, Greece, September 1998.

9 Glossary

Some key terms used in the description of digital copyright protection schemes are explained here.

Blind (public) scheme: Only the watermarking-key and the signal to be tested are required for the detection.

Cover-data: Implies that data are unmarked (i.e. it has no IAD embedded in it). The cover-data are also referred to as the cover data set. See cover-signal.

Cover-signal : the audio-visual signal (still image, audio track, video) in which one wishes to hide information - the work.

Cropping (or zooming): The act of cutting away and discarding the unnecessary portions of the picture. Most photo editing applications include a cropping tool for this purpose.

DCT: Discrete Cosine Transform.

Detection Watermark: An image independent watermark that can be detected using a secret key.

DVD: Digital Versatile (Video) Disk.

Embedding key: A secret key used to embed the mark.

Extraction key: A key used to detect or extract a watermark. Symmetric watermarking algorithms require the use of the same secret key for embedding and extraction. Asymmetric algorithms use a secret key for embedding and a public key for extraction. Keys are built in such a way that the private key cannot be computed from the public one.

FFT: Fast Fourier Transform.

Image: An image is in either digital or physical form. It may be a still image or a video frame. It can also refer to other types of data, such as audio data.

Image Authentication Data (IAD): the authentication data used in the image authentication process.

key: Any watermark signal is associated (one by one) with an integer number (or a set of integer numbers) which is the watermark key. This key is used to produce, embed and detect a watermark. The key is private and exclusively identifies the legal owner of the digital product.

Low-pass filtering: This includes linear and non-linear filters. Low-pass filters and median filters are used most often for noise suppression or smoothing, while high-pass filters are typically used for images.

Low-pass pseudo-random or chaotic signals: They are applied to still images and audio data. Their robustness is satisfactory under filtering and JPEG compression.

Pseudo-random noisy binary patterns: They are applied to still images. The watermark embedding and detection algorithms are very fast; however, such watermarks are not resistant to high JPEG compression on ratio and low-pass filtering.

Non-blind (private) scheme: The original non-watermarked cover-signal, the extraction key and the signal to be tested are required for detection.

Oblivious: A watermark technique which does not require the cover-image to extract the mark. In this case, only the stego-image is required to extract the mark when using an oblivious marking scheme.

Payload: Message or sequence of information bits to be hidden in the cover-signal, that is, the hidden data.

Private Watermark: A watermark-dependent watermark that can only be detected using a private key. It is not possible for an unauthorized third party to overwrite or delete the private watermark without the cryptographic secret keying information.

Projections Operations: In audio watermarking, projection techniques are used for data extraction procedures. A larger projection value will indicate the presence of one type of data, e.g., a binary symbol or a watermark that represents an author. A segment of the original host signal that is highly correlated with the projection direction will provide a false detection. The projection direction cannot be easily changed since the decoder does not have access to the original host signal.

Public Watermark: A watermark that can be detected using a publicly available key.

Rotation: An image can be rotated by rotating the points in the image through an angle of q.

Scaling: Points can be scaled along the x axis and the along the y axis into new points through the multiplication of the coordinates.

Semi-blind scheme: The published watermarked audio-visual signal, the extraction key and the signal to be tested are required for detection.

Signal: A signal is in either digital or physical form. It may refer to one dimensional or multidimensional signals such as image and video signals.

Statistical efficiency: The detection of a particular watermark is successful when a suitable key is used. Each watermark corresponds to a unique key.

Statistical invisibility: The possession of a great number of digital products, watermarked by the same key, does not dispose the watermark. Different products watermarked by the same key transfer different watermark signals. We can claim that the extraction of the owner's key by a third person is impossible. Counterfeit watermark keys cannot be determined, i.e. keys, which detect a watermark that had never been embedded in the digital image before.

Stego-image: Implies that an image or piece of data is marked (i.e. it has an IAD embedded in it). The stego-image is also referred to as the stego data set.

Translation: We can translate points in the (x, y) plane to new positions by adding translation amounts to the coordinates of the points.

Watermark-access-unit: Smallest part of a cover-signal in which a watermark can be reliably detected and the payload extracted.

Watermark or mark: What is actually imperceptibly added to the cover-signal in order to convey the hidden data. This process is comparable to placing an electronic stamp on a document.

Watermarking scheme: The set algorithms required for embedding and extraction.