PH.D. THESIS

FRAGILE and SEMI-FRAGILE WATERMARKING TECHNIQUES for IMAGE AUTHENTICATION

Ph.D. degree of Electronic Engineering, 2008
Michele De Santis
University of ROMA TRE, Rome Itay.

Abstract

Techniques to establish the authenticity and integrity of digital images are becoming increasly essential over the Intermet. The authentication algorithm should distinguish malicious manipulations from the original watermarked image[1]. Fragile watermarking is recomended for a very sensible image authentication scheme, able to localize the slightest changes occured to the marked image. The fragile watermarking shows the limits of its applicability, when secure transacting of digital content is required. In this case, the authentication algorithm should be able to distinguish incidental integrity maintaining distortions such as lossy compression from malicious manipulations. This has motivated further researches into semi-fragile watermarking.

Two novel watermarking algorithms are proposed in this thesis, both robust to self-authenticating. The first algorithm is a fragile spatial parity-checking watermarking method that employs a public-key scheme for still image authentication and integrity verification. The use of the parity-checking scheme and the public-key encryption enables the localization of tampered areas and the authentication of the watermarked image respectevely.

The second algorithm is a semi-fragile holographic watermarking system. The previous algorithm is a really powerfull system to be a fragile technique but still not usefull for secure transacting over the Internet. The holographic technique enables good perfomances against distortions and is still able to guarantee the authenticity and the integrity of digital images. The authenticity is achieved by a public-key encryption of the mark and the integrity is achieved by the holography introduced in this technique.

In addition, the thesis provides theoretical analysis for the perfomance and the feasibility of both schemes.We also present experimental results to verify the theoretical observations and the comparison results for both algorithms to popular techniques.

Acknowledgements

I am grateful to the Department of Electronic Engineering at the University of “Roma Tre” for providing me the opportunity to continue my studies in the digital image watermarking area.

I would like to express my sencere gratitude to my advisor, Prof. Giuseppe Schirripa Spagnolo for his guidance and support at every stage of my Ph.D. study. This thesis would not be possible without his guidance.

I am grateful to the Prof. Francesco Bella, for his continous interest in my research studies and for much more.

I am grateful to Prof. Marco Carli for giving mr the opportunity of teaching the subject, I was studying.

A special thought goes to all my friends for their company during all these years passed together and especially to my girlfriend who did support and abide me always even in the rushing hours of my writing.

This thesis is dedicated to my parents, my brothers, my grandmother and my enlarged family. There are very special thanks for my father, who have been the very important source of my faith in science and in life and even if I cannot see you no more in this earthly life, I will alwats keep you locked in my heart and every day that passes, is just another day that I get closer to you.

Contents

CHAPTER 1

INTRODUCTION TO THE IMAGE AUTHENTICATION

1.1 Introduction

In recent years, the rapid expansion of the interconnected networks and the never-ending development of digital technologies have facilitated instant multimedia transmission and the creation of large-scale digital image databases[2]. Digital images are gradually replacing their classical analog counterparts[1]. The advantages of digitized images are that images can be easily manipulated and reproduced without significant loss of quality. Digital images are easy to edit, modify and exploit.

The Internet has become the most important information source, and offers ubiquitous channels to deliver and to exchange information. Digital images can be easily shared via the Internet and conveniently processed for queries in databases. However, these also imply that images can be modified easily and imperceptibly with malicious intentions. With some powerful image processing softwares, such as Adobe PhotoShop one can remove/replace some features in a picture easily without any detectable trace. These kinds of operations are regarded as tamper. But in some cases, the images are not allowed to be done such operations, such as images for military, medical, and judicative use. The validity of the image is of most importance in these conditions[3]. The production of ownership and prevention of unauthorized manipulation of digital images are becoming an important issue[4]. So some effective ways are needed to guarantee integrity of the image. There is a film which shows of how badly image authentication are needed[5]. In the film Rising Sun, evidence of the crime is digital imagery of a murder which is captured by security camera. In the digital image, the face of the perpetrator was digitally replaced with that of another person. Sean Connery and friends prove that the image had been manipulated, thus leading back toward the truth. In reality, similar manipulations can be made nearly undetectable.If considering that at least one time a day our images are captured by security camera and what if those images contained key legal evidence in a murder trialor a journalist's photographs? This is a serious problem in our society. Techniques securing information flowing on the networks are therefore essential for protecting intellectual properties and fostering the development of electronic commerce.

A number of technologies have been developed to provide data protection including cryptography and steganography. Cryptography is the traditional method udes for authentication. It protects data by encrypting them so that the encrypted data are meaningless to an attacker, so preserving confidentiality[6][7]. The popular method of criptography is public key encryption, which encrypts data using a private key, and an associated public key is used for decryption of the secret message. The problems that might be arisen in this method are difficulty of maintenance and distribution of public key. It has great strength in confidentiality, but when the data is revealed to unauthorized personnel, there is no-protection for the content it-self, wich is integrity control. On the othe hand, steganography conceals the very existence of the data by hiding them in cover data. The problem is that, it cannot extract the hidden data if the stego data undergo some distortions.

A new emerging technology, digital watermaking, complements cryptography and steganography by embedding an invisible signal directly into the data, thus providing a promising way to protect digital data from illicit copying and manipulation. After the embedding, the watermark and the data are inseparable. There is a wide range of applications of digital watermarking including copyright protection, authentication, fingerprinting, copy control, and broadcast monitoring, etc. For different kind of applications, digital watermarking should satisfy different properties[8][9].

A digital watermark is a visually imperceptible, information-carrying signal embedded in a digital image. A watermarking scheme can be classified as either *robust* or *fragile*. Robust watermarks are generally used for copyright and ownership verification. High robustness is a requirement for copyright protection to provide ownership in any kind of attacks. On the other hand, a fragile watermark is a watermark that is readily altered or destroyed when the host image is modified through a linear or non-linear transformation. The sensitivity of fragile marks to modification leads to their being used in image authentication[10][11][5]. Fragile watermarks are useful for purposes of authentication and integrity attestation[12]. A secure authentication system is useful in showing that no tampering has occured during situations where the credibility of an image may be questioned[13]. It provides a guarantee that the image has not been tampered with and comes from the right source. The fragile watermark method is useful to the area where content is so important that it needs to be verified for it being edited, damaged or altered such as medical images. There exist also so called semi-fragile watermarking techniques, where some manipulations are allowed (for example JPEG compression to a predifined quality factor) but other data manipulatioons are detected as tampering[14].

Many fragile watermarking schemes have ben proposed in recent years, for example[12][1][15-18]. Among them, Wong in[15] has proposed a blockwise fragile authentication watermarking and in[16] has improved it by using a public-key based scheme. Since then, some number of public-key based fragile watermarks have appeared in the litterature [19-21]. Using a public-key cipher, claims of image authenticity can be judged without the necessity of disclosing any private information. Moreover, solid cryptgraphy theory makes this scheme reliable, when due cares are taken into account.

A digital signature[22] is an algorithm for ensuring integrity and authenticity of sensitive digital data. It computes a fingerprint of the data by using a hashing function, and then employs an asymmetric (public-key) cipher to encrypt the fingerprint with the originator's private-key. In the signature verification step, the hashing function is applied on the received data and the accompanying signature is decrypted using the signer's public-key. The results are expetected to match, unless the data or signature are corrupted or faked.

Classical digital signatures are able to detect alterations in signed data but no to locate them. In contrast, most fragile watermarking techniques provide the ability lo localise where the alterations have taken place.

Schneider and Chang[23] proposed a method to embed content-based signature using private key as a watermark. This authentication scheme also requires distribution of public key to verify the watermarked image. But the system proposed in this paper uses client-server model that server holds a watermark detection method internally and client can access to the server using Internet to verify the image, which does not require distribution of public key that maybe the major problem of using public key encryption[6].

Fragile watermarking systems are categorized into two categories according to the working domain. First, fragile watermarking that works directly in the spatial domain. Second, fragile watermarking that works in a transformed domain.

Most fragile watermarking systems embed the mark directly through the spatial domain of a Work, such as techniques described in [10][24] and [5]. These techniques embed the mark in the least significant bit (LSB) plane for perceptual transparency. Their significant disavantages include the ease of bypassing the security they provide [13] and [24].

Wong [15] decribed another fragile marking technique which obtains a digest using a hash function. The image, image dimensions, and marking key are hashed during the embedding and are used to modify the least significant bitplane of the original image. This is done in such a way that when the correct detection side information and unaltered marked image are provided to the detector, a bi-level image chosen by the owner (such as company logo or insignia), is observed. This technique has localization properties and can identify regions of modified pixels within a marked image. The technique of Yeung and Mintzer [25] is also one where the correct detection information results in a bi-level image. However, the embedding technique is more extensive than inserting a binary value into the least-significant bit plane. The marking key is used to generate several pseudo-random look-up tables (one for each channel or color component) that control how subsequent modifications of the pixel data will occur. Then, after the insertion process is completed, a modified error difffusion process can be used to spread the effects of altering the pixels making the mark more difficult to see. On the other hand, various transformations, such as: the Discrete Fourier Transform, the Discrete Cosine Transform (DCT) and Wavelet Transforms are used for authentication systems. Usually those systems are semi-fragile since they are almost all robust to Lossy Compression. DCT based watermarking systems are usually robust to Joint Photographic Experts Group (JPEG) lossy compression while those work in the Wavelet Domain are robust to Joint Photographic Experts Group 2000 (JPEG2000) Lossy Compression.

Wu and Liu described a technique in [18] which is based on a modified JPEG encoder. The watermark is inserted – by changing the quantized DCT coefficients – before entropy encoding. A special lookup table of binary values (whose design is constrained to ensure mark invisibility) is used to partition the space of all possible DCT coefficient values into two sets. The two sets are then used to modify the image coefficients in order to encode a bi-level image (such as a logo). In order to reduce the blocking effects of altering coefficients, it is suggested that the DC coefficient – and any coefficients with low energy – is not marked. Kundur and Hatzinakos in [26] embed a mark by modifying the quantization process of the Haar wavelet transform coefficients. While Xie and Arce in [27] selectivily inserts watermark bits by processing the image after it is in a compressed form. A wavelet decomposition of an image contains both frequency and spatial information about the image hence watermarks embedde in the wavelet domain have the advantage of being able to locate and characterize the tampering of a marked image.

Two types of authentication systems are currently being investigated: global and local authentication. As the naming implies, global authentication system considers the Work as a whole (i.e., either the Work is authentic or not). The other type of systems is local, (i.e., the authentication is based on local regions in the Work). So the authentication system outputs the regions in the work as authentic regions while others are not[28].

Two different kinds of fragile watermarking systems are proposed in this thesis. The first one is a transformed domain fragile watermarking. The digital image goes through the Discrete Fourier Transform and then, the hologram of the image is made by a particular kind of coding. The hologram is then embedded into the digital image. Even though the system works in the Fourier domain, it results resistant to cropping attacks. This is the reason to use the hologram of the image as mark. The idea to bring the cropping resistant property in a transformed domain watermarking, is completely new. It allows the technique, to be used to authenticate digital images and to be resistant to Lossy Compression[29].

The second one is a spatial fragile watermarking devoted to digital image authentication and tamper localization. The main novelty of the system is its capability to resist to any kind of cropping attack, being able to authenticate not only the whole image but even very tiny pieces of it. As far as the author knows, this is one of the first systems if not the first one to show such a property[30].

1.2 Image authentication motivation and application

Image authentication assures communication receivers that the received image is in fact from the authorized source, and that the image content is identical to the original one sent. The past few years have witnessed an increasing used of digitally stored information and the development of new multimedia digital servers. Digital images are gradually replacing their classical analog counter parts, since the digital image is easy to edit, modify and exploit. With the rapid development of the Internet, digital images can be easily shared via computer networks and conveniently processed for queries in databases. At the same time, image editing programs are becoming more powerful, so that even an amateur can maliciously modify digital images and create perfect forgeries without leaving any trace on the original image. So techniques to establish the authenticity and integrity of digital images are essential. Especially the image content is used for the content sensitive fields such as photojournalism, courtromm evidence,medical applications, or commercial transaction, the originator of the content has to be verified while ensuring the content has not been changed, manipulated or falsified [10][11].

Image authentication systems have also applicability in law,commerce,defense, and journalism. Since digital images are easy to modify, a secure authentication system is useful in showing that no tampering has occured during situations where the credibility of an image may be questioned. Common examples are the marking of images ina database to detect tampering[24][31][32], the use in a “trustworthy camera” so news agencies can ensure an image is not fabricated or edited to falsify events [33], and the marking of images in commerce so a buyer can be assured that the images bought are authentic upon receipt [16]. Other situations include images used in courtroom evidence, journalistic photography, or images involved in espionage.

Another method to verify the authenticity of a digital work is the use of a signature system [34]. In a signature system, a digest of the data to be authenticated is obtained by the use of cryptographic hash functions [34][35]. A cryptographic hash function is a procedure that takes an arbitrary block of data and returns a fixed-size bit string, the hash value, such taht an accidental or intentional change to the data will almost certainly chenge the hash value. The ideal hash function has four main properties:

- It is easy to compute the hash for ny given data.
- It is extremely difficult to construct a text that has a given hash.
- It is extremely difficult to modify a given text without changing its hash.
- It is extremely unlikely that two different messages will have the same hash.

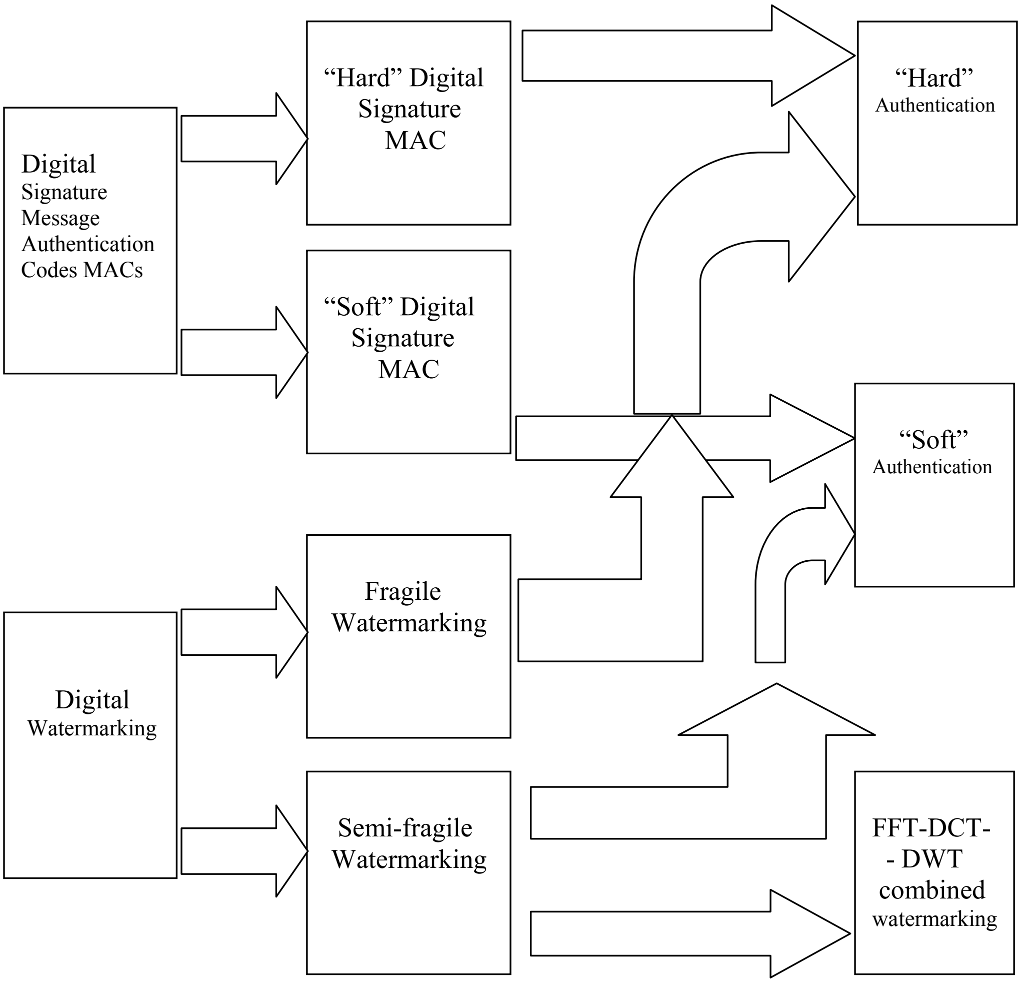
These requirements call for the use of advanced cryptography techniques, hence the name. In various standards and applications, the two most commonly used hash functions are MD5 and SHA-1.

The digest is then cryptographically signed to produce the signature that is bound to the original data. Later, a recipient verifies the signature by examining the digest of the (possibly modified) data and using a verification algorithm determines if the data is authentic. While the purpose of fragile watermarking and digital signature systems are similar, watermarking systems offer several advantages compared to signature systems [36] at the expense of requiring some modification (watermark insertion) of the image data. Therefore the critical information needed in the authenticity testing process is discreetly hidden and more difficult to remove than a digital signature. Also, digital signature systems view an image as an arbitrary bit stream and do not exploit its unique structure. Therefore a signature system may be able to detect that an image had been modified but cannot characterise the alterations. Many watermarking systems can determine which areas of a marked image have been altered and which areas have not, as well as estimate the nature of the alterations.

1.3 Fragile Image Authentication as a solution.

To achieve image auhentication, fragile watermarking techniques are commonly used. “Fragile” authentication is more sensible than the “semi-fragile” or “robust” counterparts. The digital watermark must be fragile to any kind of distortion, the watermarked image may go through. For example, the image after lossy processing such as JPEG could be found to be authentic by “robust” image authentication, but it would fail “fragile” image authentication. For a “fragile” image authentication, one bit error in the message leads to a totally different authenticator, however, for a “semi-fragile” image authentication, such an error does not necessarily alter the authenticator. Fragile image authentication is higly sensitive and dependent on the exact value of image pixels, whereas “semi-fragile” image authentication is sensitive just to content modification and serious image quality tampering. “Semi-fragile” image authentication is ideally indipendent on the logical content-based, non-variant relation among image pixels.

The existing “semi-fragile” image authentication techniques are divided into two categories, that is, **Digital Signature and MAC** and **Semi-fragile Digital Watermarking**. A digital signature is attached to images as a tag or a header.



Digital Signature Message Authentication Codes MACs
“Hard” Digital Signature MAC
“Soft” Digital Signature MAC
“Hard” Authentication
“Soft” Authentication
Digital Watermarking
Fragile Watermarking
Semi-fragile Watermarking
FFT-DCT-- DWT combined watermarking

Figure 1.1: Classification of Authentication systems

The attacker can detach the signature on the Internet and create a denial of the service attack by breaking the logic link. To apply a new digital signature algorithm, the compability with the present

Internet security protocol and standard is very crucial. Since the currently used protocols and standards are designed for the conventional “hard” MACs and digital signatures, and the proposed “soft” digital signature schemes[11,37,38,39,40,41] generate a new format signature with new length and format, there are complex and costly modifications needed to implement the new “soft” digital signature algorithm. These changes include a new type of detector, and a new format for the authentication header or tag. The overhead incurred due to the presence of the added signatures reduces the transmission rate.

In contrast, the authenticator of a semi-fragile system is inseparably bound to the content, hence simplifying the logistical problem of data handling. Most current semi-fragile watermarking schemes must transmit secret information such as the length or location of the original watermark to allow authentication by a receiver[42,43,44,45,46,47,48,49,50,51,52,53]. Unfortunately, the necessity for this confidential information makes such techniques vulnerable to passive attacks and replay attacks common to Internet environments, because the attacker can eavesdrop to obtain these pieces of information and analyze them to get embedded watermark, and then rembed watermark to a new image.

For the above reasons, this thesis presents a novel fragile watermarking algorithm robust to self-authenticating and subsequently a semi-fragile watermarking robust to lossy compression and self-authenticating. Both methods employ a public key scheme for still image authentication and integrity verification, without using explicit information derived from the original image or watermark. The algorithm can be implemented simply and practically like a small,individual image processing software associated with all kinds of network software, standards and protocols. The users just need keys used in the public-key system for content secured image transmission on the Internet.

Desirable features of fragile watermarking systems, noting that the relative importance of these features will depend on the application. Applications may have requirements other than the ones mentioned. In addition to the features described below and the desired properties previously mentioned, other properties can be found in [31][26][18]:

1. **Detect Tampering**. A fragile marking system should detect (with high probability) any tampering in a marked image. This is the most fundamental property of a fragile mark and is a requirement to reliably test image authenticity. In many applications it is also desirable to provide an indication of how much alteration or damage has occured and where it is located (see figure ...)

2. **Perceptual Transparency.** An embedded watermark should not be visible under normal observation or interfere with the functionality of the image[54]. In most cases this refers to preserving the aesthetic qualities of an image, however if an application also performs other operations on marked images (such as feature extraction) then these operations must not be affected. Unfortunately there is not a lot of information how the “noise” introduced by marking process, affects other image processing operations[55]. This is an open research problem. Also, transparency may be a subjective issue in certain applications and finding measures, which correlate well with perceived image quality, may be difficult.

3. **Detection should not require the original image**. As mentioned above the original image may not exist or the owner may have good reason not to trust a third party with the original (since the party could then place their own mark on the original and claim it as their own).

4. **Detector should be able to locate and characterise alterations made to a marked image**. This includes the ability to locate spatial regions within an altered image which are authentic or corrupt. The detector should also be able to estimate what kind of modification had occured.

5. **The watermark detectable after image cropping.** In some applications, the ability for the mark to be detected after cropping may be desirable. For example, a party may be interested in portions (faces, people, etc.) of a larger, marked image. In other applications, this featureis not desired since cropping is treated as a modification.

6. **The watermarks generated by different marking keys should be “orthogonal” during watermark detection.** The mark embedded in an image generated by using a particular marking key must be detected only by providing the corresponding detection side information to the detector. All other side information provided to the detector should fail to detect the mark.

7. **The marking key spaces should be large.** This is to accomodate many users and to hinder the exhaustive search for a particular marking key even if hostile parites are somehow able to obtain both unmarked and marked versions of a particular image.

8. **The marking key should be difficult to deduce from the detection side information.** This is particularly important in systems that have distinct marking and detection keys. Usually in such systems the marking key is kept private and the corresponding detection side information may

be provided by other parties. If the other parties can deduce the marking key from the detection information then they may be able to embed the owner's mark in images that the owner never intended to mark.

9. **The insertion of a mark by unauthorised parties should be difficult.** A particular attack mentioned in [31] is the removal of the watermark from a marked image and subsequently inserting it into another image.

10. **The watermark should be capable of being embedded in the compressed domain.** This is not the same as saying the watermark should survive compression, which can be viewed as an attack. The ability to insert the mark in the compressed domain has significant advantage in many applications.

The framework for embedding and detecting a fragile mark is similar to that of any watermarking system. An owner (or an indipendent third party authority) embeds the mark into an original image (see figure 1.2).

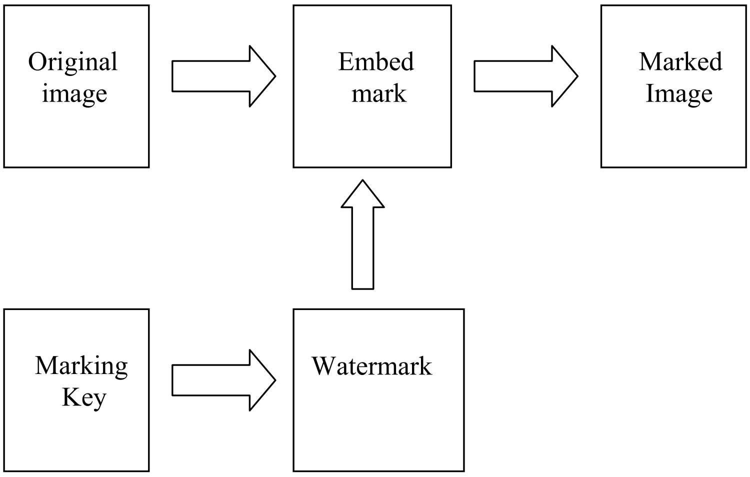


Figure 1.2: Watermark embedding

The marking key is used to generate the watermark and is tipically an identifier assigned to the owner or image. The original image is kept secret or may not even be available in some applications such as digital camera. The marked image may be transmitted, presented, or distributed.

The marked image is perceptually identical to the original image under normal observation. See Figure 1.3 and Figure 1.4, for an example of original and marked images using the fragile marking technique described in [35][56].

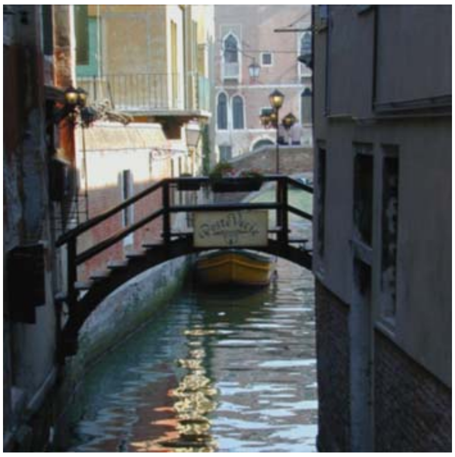


figure 1.3: Original Image

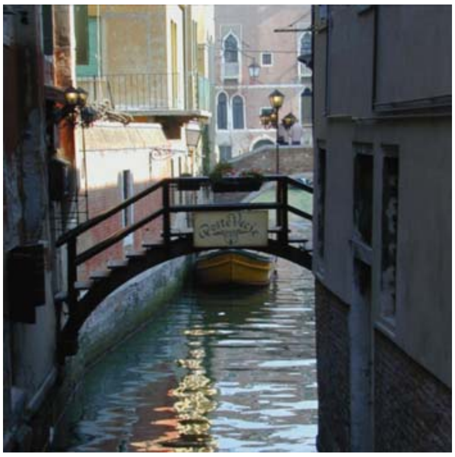
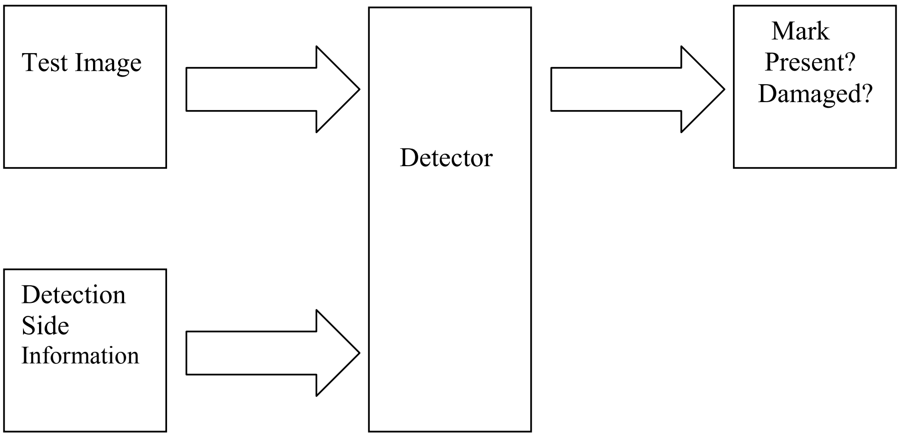


figure 1.4: Watermarked Image

When a user receives an image, he uses the detector to evaluate the authenticity of the received image (see figure 1.5). The detection process also requires knowledge of “side information”. This side information may be the marking key, the watermark, the original image, or other information. The detector is usually based on statistical detection theory whereby a test statistic is generated and from that test statistic the image is determined to be authentic. If it is not authentic then it would be desirable for the detector to determine where the image has not been modified.



1.5: Figure Watermark Detection

The side information used by the detector is very important in the overall use of a fragile watermark. Techniques that require that the detector have the original image are known as private watermarks while techniques that do require the detector to have the original image are known as public watermarks. To be effective a fragile watermarking system must be a public technique. In many applications the original image may never be available since it might have been watermarked immediately upon creation.

In database applications the owner or authority who marks the images is often the party interested in verifying that they have not been altered a subsequent time. For example, in a medical database it is important that any modifications to images, be detected. In other applications, such as commerce, the verifying parties are distinct from the marking entity. In these cases, it is desirable to choose a system where the marking and detection information are distinct. In such a system, the ability to determine the authenticity of images does not also grant the ability to mark images. The vast majority of fragile systems described in the current literature do not implement this approach.

1.4 Attacks on Fragile Marks

One must be mindful of potential attacks by malicious parties during the design and evaluation of marking systems. It may be practically impossible impossible to design a system impervious to all forms of attack, and new mwthods to defeat marking systems will be invented in time. But certainly knowledge of common attack modes is a requirement for the design of improved systems.

The first type of attack is blind modification of a marked image (that is, arbitrarly changing the image assuming no mark is present). This form of attack should be readily recognized by any fragile mark, yet it is mentioned because it may be the most common type of attack that a marking system is to defeat. Variations of this attack include cropping and localized replacement (such as substituting one person's face with another). The latter type of modification is a significant reason why an application may want to be able to indicate the damaged regions within an altered image.

Another type of attack is to attempt to modify the marked image itself, without affecting the embedded mark or creating a new mark that the detector accepts as authentic. Some weak fragile marks easily detect random changes to an image but may fail to detect a carefully constructed modification. An example is a fragile mark embedded in the least-significant bit plane of an image. An attempt to modify the image without realizing that a mark is expressed in the LSB is very likely to disturb the mark and be detected. However, an attacker that may attempt to modify the image without disturbing any LSBs or substitute a new set of LSBs on a modified image tat the detector classifies as authentic.

Attacks may also involve using a known valid mark from a marked image as the mark for another, arbitrary image[24][31]. The mark-transfer attack is easier if it is possible to deduce how a mark is inserted. This type of attack can also be performed on the same image; the mark is first removed, then the image is modified, and finally the mark is re-inserted.

An attacker may be intersted in completely removing the mark and leaving no remnants of its existence (perhaps so they can deny ever bearing witness to an image which has their mark embedded in it). To do so, an attacker may attempt adding random noise to the image, using techniques designed to destroy marks (such as Stirmark[57]), or using statistical analysis or collusion to estimate the original image.

An attacker may also attempt the deduction of the marking key used to generate the mark. The marking key is intimately associated with an embedded mark, so if it is possible to isolate the mark, the attacker can then study it in an attempt to deduce the key (or reduce the search space for the marking key). Once the key is deduced, the attacker can then forge the mark into an arbitrary image.

There are also known attacks that involve the authentication model itself and not so much on the specific mark in an image. Attacks on authentication systems over insecure channels are also possible in [58] and similar vulnerabilities can apply to watermarking systems.

Removal attacks achieve complete removal of the watermark information from the watermarked data without cracking the security of the watermarking algorithm. This category includes denoising, quantization, remodulation, averaging and collusion attacks.
Not all of these methods always come close to complete watermark removal, but they may damage the watermark information significantly.

Geometrical attacks do not remove the embedded watermark itself, but intend to distort the watermark detector synchronization with the embedded information.
To this category, there belong the cropping, flip, rotation and synchronization removal attacks too.

Cryptographic attacks aim at cracking the security methods in watermarking schemes and thus finding a way to remove the embedded watermark information or to embed misleading watermarks. One such technique is brute-force search for the embedded secret information. Practically, application of these attacks is restricted due to their high computational complexity.

Protocol attacks aim at attacking the entire concept of the watermarking application. One type of protocol attack is the copy attack. The main idea of a copy attack is to copy a watermark from one image to another image without knowledge of the key used for the watermark embedding to create ambiguity with respect to the real ownership of data.

This thesis presents a fragile holographic watermaking technique, able to authenticate full digital images and even small pieces of it.
The authentication problem concerned in this thesis is represented by the following scenario. Alice sends an image message I to Bob through an open channel. Bob receives the corresponding message I^r from the channel. Then, Bob must ensure that:

- I^r was actually sent by Alice.
- The content of I^r is exactly the same as I sent by Alice.
- If the content of I^r has been modified, the modification can be located.

1.5 Authentication and Privacy

It is well known that the concealment of information or protection of *privacy* is as old as writing itself. Human ingenuity, found many ways to conceal information: steganography, i.e., the hiding of the mere existence of a message, codes, where words or combinations of words are replaced by fixed symbols, and cryptology or ciphers, where information is transformed to render ti useless for the opponent. The distinction between the latter two is rather subtle, and can be made on the fact that codes split up information according to semantic borders, while ciphers operate on chunks of information independently of the linguistic interpretation. The technological evolution from handwritten messages on paper sent by courier to the communication of information through both local and worldwide communication networks and the storage and processing in a variety of computer systems cetainly has increased the vulnerability of information to eavsdropping. Cryptography was the only solution that was able to make the leap from the closed world of generals and diplomats to worldwide commercial applications.

Apart from concealment or privacy protection, it is equally important that both the contents and the originator of the information are not modified. Both requirements are captured in the term *authentication*. The fact that the relative importance of this threat has increased can be illustrated by the emergence of malicious software programs. The best known examples of this group are certainly the computer viruses[59,60]. Others include worms [61]. Trojan horses, and logical bombs. Every effective solution will have to be based on a verification of the authenticity of the software when it is loaded on the hard disk and when it is loaded by the CPU. The latter application will require very high throughput of 100 Mbytes per second and even more. A second illustration is situated in the banking world. The authenticity of financial transactions is generally considered more important than the secrecy, as one successful fraud can result in a considerable benefit for the attacker. The problem here is not only the economical value of a single attack, but the fact that the trust in the system can be lost[62]. A third application that will become more and more important is the protection of the authenticity of pictures and moving images (e.g. videoconferencing). As one can expect that it will become feasible to “edit” moving pictures and make a person say and do things he or she never said or did, it is required that one can guarantee the authenticity of moving images. This will impose even higher requirements on the throughput. Other applications where authentication is important are alarm systems, satellite control systems, distributed control systems, and systems for access control[63].

Authentication is the protection of the communicating parties against attacks of a third party. However, a different threat emerges when the two communicating parties are mutually distrustful and try to perform a *repudiation*. This means that sender or receiver will try to modify a message and/or deny to have sent or received a particular message. In paper documents, protection against this type of attack is offered by a handwritten signature. It is clear that in case of electronic messages, a simple name at the end of the message offers no protection at all. This is analogous to the fact that a photocopy of a document with a manual signature has no value, as one can always produce a bogus document with cut and paste operations. A typical example of this fraud is the electronic communication of orders to a stockbroker. The customer can place an order to buy shares of company X. If some delays later the transaction turns out badly, he will try to deny his order. If on the other hand, the transaction is successful, the stockbroker might claim that he never received the order with the intention to keep the profit. An elegant technical solution to this problem was offered by the concept of digital signature schemes based on trapdoor one-way functions[64].

The concepts of authentication and privacy, that were at first closely related, grew more and more apart. First a model has to be given for a cypher system. Hereto, the for a symmetric or conventional cipher system introduced by C.Shannon in 1949[65] will be extended to include asymmetric or public-key cipher systems (figure 1.6).

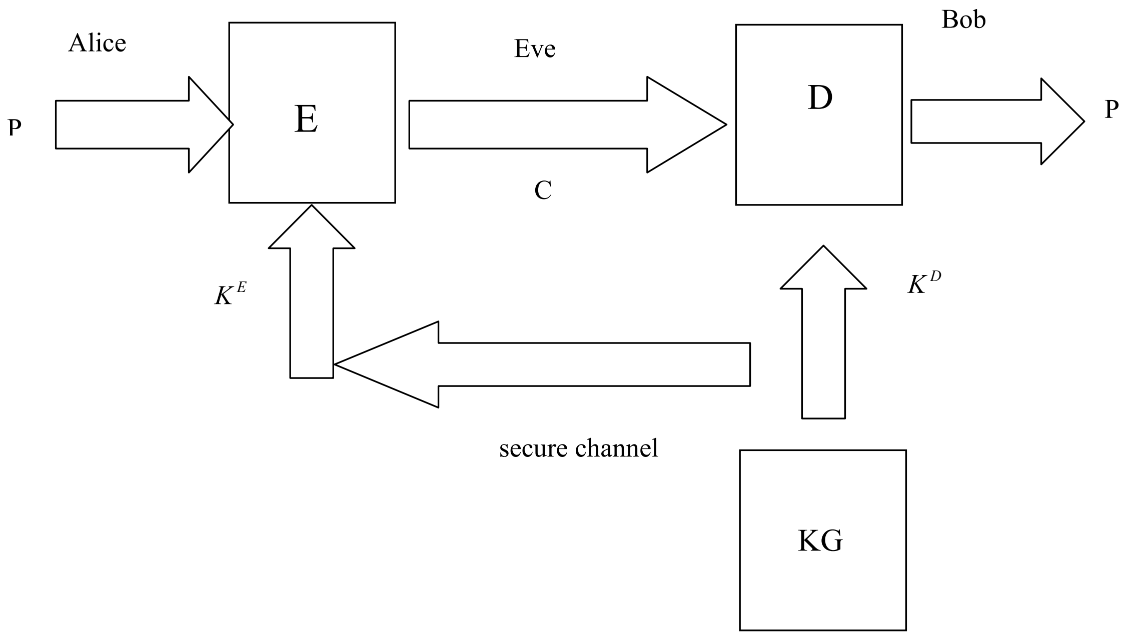


figure 1.6: Model of cipher system

The sender Alice wants to transmit the plaintext *P* to the receiver. She will transform the plaintext *P* into the ciphertext *C* with the encryption algorithm *E*. The encryption algorithm *E*, is actually a family of transformations parameterized by an ecryption key K^E. The receiver Bob will recover the plaintext *P* by applying the decryption algorithm *D*, This algorithm is in the same way parameterized by a decryption key K^D. The model also contains a key generation algorithm *KG*, that produces corresponding pairs K^E and K^D. For simplicity it will be assumed that the generation of keys is controlled by Bob. The key K^E has to be sent to Alice through a secure channel. The eavesdropper Eve, who also will be called cryptanalyst or opponent, knows the desription of E and D, and has access to C. She will try to derive information on the plaintext P.

In case of a symmetric cipher, K^E is made public, and therefore this type of cipher is also called public-key algorithm. The channel to distribute K^E only has to protect the authenticity, while Bob has to protect both the authenticity and the secrecy of K^D. The assumptions underlying this cipher are that knowledge of E,D and C does not allow to derive Pand that knowledge of K^E does not allow to derive K^D. The concept invented to achieve these properties is trapdoor one-way permutation (in fact a trapdoor one-way function suffices). This is a permutation that is hard to invert unless one knows some secret trapdoor information.

A beatiful property of the public-key algorithms is that if the encryption function is a trapdoor one-way permutation, they can be turned easily into a digital signature scheme. If Bob transforms a plaintext P with his secret key K^D, he obtains the ciphertext $C^{'}$. It is possible for Alice – in fact for everyone who knows the corresponding public key K^E - to encipher $C^{'}$ with K^E and to verify that P is obtained. Note that here the implicit assumption is made that P contains some verifiable redudancy. Because Bob is the only person who knows K^D, it is clear that he is the only person who can possibly have generated $C^{'}$, and hence he ca not deny to have sent P. If both Alice and Bob generate their own key pair and exchange their respective public keys, a superposition of both operations will guarantee both privacy and authentication (figure 1.7).

Alice will decipher P with her secret key K_A^D, subsequently encipher the result with the public key K_B^E of Bob, and send the resulting ciphertext $C^{''}$ to Bob. Bob can obtain the corresponding plaintext and verify its authenticity by deciphering $C^{''}$ with his secret key K_B^D and subsequently encrypting the result with the public key K_A^E of Alice.

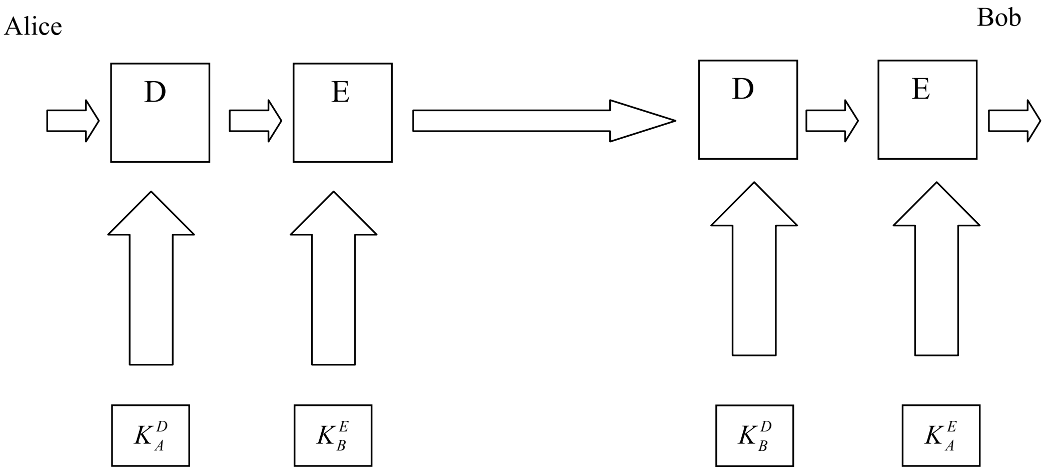


figure 1.7: Protection of both authenticity and privacy with a public key system

It is clear that the extension of this model to a model with central key generation and distribution is straightforward. In this case a hierarchical approach is possible based on master keys.

1.6 Information authentication and digital signatures

This section aims to illustrate briefly how cryptographic hash functions can be used to protect the authenticity of information and to improve digital signature schemes.
The protection of the authenticity of information includes to aspects:

- The protection of the originator of the information, or in ISO terminology[59][66] data origin authentication.
- The fact that the information has not been modified, or in ISO terminology [66] the integrity of the information.

There are two basic methods for protecting the authenticity of information.

- The first approach is analogous to the approach of a symmetric or asymmetric cipher, where the secrecy of large data quantities is based on the secrecy and authenticity of a short key. In this case the authentication of the information will also rely on the secrecy and authenticity of a key. To achieve this goal, the information is compressed to a quantity of fixed length, which is called a *hashcode*. Subsequently the hash code is appended to the information. The function that performs this compression operation is called a *hash function*. The basic idea of the rotection of the integrety is to *add redundancy* to the information. The presence of this redundancy allows the receiver to make the distinction between authentic information and bogus information.

Commento [m1]: h

In order to guarantee the origin of the data, a secret key that can be associated to the origin has to intervene in the process. The secret key can be involved in the compression process or can be used to protect the hashcode and/or the information. In the first case the hashcode is called a Message Authentication Code or MAC, while in the latter case the hashcode is called a Munipulaion Detection Code or MDC.

- The second approach consists of basing the authenticity (both integrity and origin authentication) of the information on the authenticity of a Manipuation Detection Code or MDC. A typical example for this approach is a computer user who will calculate an MDC for all its important files. He can store this collection of MDC's on a floppy, that is locked in his safe, or he can write them down on a piece of paper. If he has to transfer the files to a remote friend, he can simply send the files and communicate the MDC's via telephne. The authenticity of the telephone chnnel is offered here by voice identification.

The second application of cryptographically secure hash functions is the optimization of digital signature schemes and the construction of digital signature schemes that are not based on a trapdoor one-way permutation. The optimization is obtained through signing the MDC of a message instead of every bit or block. The description of the hash function can be public and it does not depend on any secret parameter. The advanteges of this approach are that the signature has a fixed short length and that the computational requirements are minimized. In some cases the security level of the signature scheme can be increased. For some signature schemes based on one-way functions are in general less practical, but can be an alternative if one is not allowed or willing to use a scheme based on a trapdoor one-way permutation.

1.7 Privacy and authentication: two different concepts

Until recently, it was generally believed that encryption of information suffices to protect its authenticity. The argument was that if a ciphertext resulted after derryption in meaningful information, it should be originated with someone who knows the secret key, guaranteeing authenticity of message and sender. As a consequence, if an opponent wants to modify an enciphered message or to send a fake message, he has to know the secret key and hence to break the encryption algorithm. The opposite observation is clearly true: someone who has broken the encryption algorithm can easily modify messages or send bogus messages. One of the famous persons who experienced this was Mary, Queen of Scotland, who organized with A. Babington in 1586 a plot to assassinate Elizabeth. Her encrypted communications were deciphered by Phelippes. This enables Phellipes to add a correctly enciphered postscript to one of Mary's letters asking for *"the names and the qualities of the six gentlemen which are to accomplish the designment"*,[67],pp.122-123. Althuogh the conspirators were caught before the answer could be intercepted, the forgery clearly illustrates the point.

Several counterexamples will show that is not necessary to break the cipher system in reder to be able to falsify messages. At first, the protection of integrity is strongly dependent on the encryption algorithm and on the mode in which the algorithm is used.

A famous ciphers that offers unconditional secrecy is the Vernam cipher or modulo 2 one-time pad [68]. It was invented in 1917 by G. Vernam to encipher printed telegraph communications. The ciphertext is obtained through an addition modulo 2 or exor of the key. Its security relies on the fact that ciphertext and plaintext are statistically independent. The disadvantage of the system is that the size of the key is as large as the the size of the plaintext. However, C.Shannon showed in 1949 [65] that this is optimal to achieve unconditional or perfect secrecy. Nevertheless an active attacker can change any bit of thr plaintext by simply flipping the corresponding bit of the ciphertext, as was remarked by Feistel [69,70]. This obsrvation is also valid for any additive stream cipher and for OFB mode of any block cipher. It holds partially if a block cipher in CFB or CBC mode [71].

If a plaintext longer than one block is enciphered with a block cipher in ECB mode, an active attacker can easily reorder the blocks. Another example is the vulnerability to active attacks of a plaintext encrypted in Cipher Feedback Mode (CFB). Due to the self-syncronization properties, any modification in the ciphertexe will cause a corresponding modification to the plaintext and will subsequently garble the next part of the plaintext. When the error is shifted out of the feedback register, the ciphertext will be deciphered correctly again. If the last part of the ciphertext is modified, it is completely impossible to detect this. If the garbling occurs in the middle of the plaintext, it can only be detedted based on redundancy.

In other modes (e.g. CBC) every ciphertext bit is a complex function of the previous plaintext bits and an initial value. If the modification of a single ciphertext bit results in *t* bits of the plaintext being garbled, the probability that the new plaintext will be accepted as meaningful equals 2^{-tD}, where D is the redundancy in the information. In case of natural language D ≈ 3.7, and this probability is negligible for t>8. Howevwr, if D=0 all messages are meaningful, and encryption offers no authentication, independently of the encryption algorithm or of the mode. This means that the attacker can modify messages or forge messages of his choice. This limitation is that he does not know on beforehand what the corresponding plaintext will be, but many applications can be considered where such an attack would cause serious problems. Note that even if redundancy is present , a human checker or a designated computer program is required to check its presence.

The second illustration of the independence of privacy and authentication is given by the use of *public-key algorithms*. From the previous sections it is clear that two independent operations and two independent key pairs are necessary to protect both privacy and authenticity.

A third example is the use of a Message Authentication Code or MAC to protect the authenticity. A widely accepted and standardized way to compute a MAC is to encrypt the plaintext in CBC mode. The ciphertext corresponding to the last block depends on the secret key, on the initial value, and on all bits of the plaintext, hence it can be used as a MAC. In case the plaintext has to be encrypted, it is very appealing to use the same key for the encryption and for the calculation of the MAC, but a different initil value. However, it can be shown that this approach is insecure [72], and that a different key should be used for authentication and encryption purposes. A similar observations is made in [71] for an authentication schemebased on a stream cipher.

The last argument is an interesting series of toy examples with a two bit key and one bit plaintext [73] illustrating that a cipher can offer either perfecr secrecy, or perfect authenticity or both. The conclusion is that *"....secrecy and authenticity are independent attributes of a cryptographic system...."*.

1.8 Three Approaches to the authentication problem

In present day cryptography, three approaches can be identified to solve most problems compromising information secrecy and information authenticity. These approaches differ in the assumptions about the capabilities of an opponent, in the definition of a cryptanalitic success, and in the notion of security. This taxonomy is based on the taxonomy that was developed for stream ciphers by R. Rueppel [74], and deviates from the taxonomy for authentication developed by G. Simmons [75].

A first method is based on information theory, and it offers unconditional security, i.e., security independent of the computing power of an adversary. The complexity theoretic approac starts from an abstract model for computation, and assumes that the opponent has limited computing power. The system based approach tries to produce practical solutions, and the security estimates are based on the best algorithm known to break the system and on realistic estimates of the necessary computing power or dedicated hardware to carry put the algorithm. In [75] the second and third approach are lumped together as computationally secure, and in [74] a fourth approach is considered, in which the opponent has to solve a problem with a large size (namely examining a huge publicly accessible random string); it can be considered as both computationally secure and information theoretically secure.

The properties of the three approaches are compared in Table 1.1. It should be noted that the information theoretic approach is impratical for most applications because of the size of the secret key. Sometimes the complexity theoretic approach allows for efficient constructions, but in a practical scheme the dimensions of the scheme are fixed and the proof of security has only a limited value.

	Computing of power opponent	security	practicality

Information theoretic
complexity theoretic
system based
	unlimited
	polynomial
	fixed
		provable(unconditional)
		asymptotic(assumptions)
		no proof
			impractical
			impractical
			efficient

Table 1.1: Comparison of the three approaches in cryptography

1.9 Information theoretic approach

This approach results in a characterization of unconditionaly secure solutions, which implies that the security of the system is independent of computing power of the opponent. E.g., in case of privacy protction, it has been shown by C.Shannon that unconditional privacy protection requires that the entropy of the key is lower bounded by the entropy of the plaintext. It should be remarked that both unconditional authenticity are only probabilistic: even if the system is in optimal with respect to some definition, the opponent has always a non zero probability to cheat. However, this probability can be made exponentially small. The advantage of this approach lies in the unconditional security. Like in the case of the Vernam scheme, the price paid for this is that these schemes are rather impratical.,

The cryptographer considers a game-theoretic model, in which the opponent observes l messages and subsequently tries to impersonate or substitute messages. The cryptographer will encipher his messages under control of a secret key. Because the gol of the cryptographer is now the protection of the authenticity (or the combination of secrecy and authenticity), the transformation will be called an authentication code. The information theoretic study of authentication has now been reduced to the design of authentication codes, that are in some sense dual to error correcting codes [75]. In both cases redundant information is introduced: in case of error correcting codes the purpose of this redundancy is to allow the receiver to reconstruct the actual message from the received codeword, and to facilitate this the most likely alterations are in some metric close to the original codeword; in case of authentication codes, the goal of the redundancy is to allow the receiver to detect substitutions or impersonations by an active eavesdropper, and this is obtained by spreading altered or substituted messages as uniformly as possible.

The advantage of this approach lies in the unconditional security. Like in case of the Vernam scheme, the price paid for this is that these schemes are rather impractical.

1.10 Complexity theoretic approach

The approach taken here is to define at first a model of computation, like a Turing machine [76] or a Boolean circuit [77]. All computations in this model are parameterized by a security parameter, and only algorithms or circuits that require asymptotically polynomial time and space in terms of the size of the input are considered feasible. The next step is then to design cryptographic systems that are provably secure with respect to this model. This research program has been initiated in 1982 by A. Yao [78,79] and tries to base cryptographic primitives on general asumptions. Examples of *cryptographic primitives* are: secure message sending, cryptographically secure pseudo-random generation, general pseudo-random generation, general zero-knowledge interactive proofs, Universal Way Hash Functions (UOWHF), Collision Resistant Hash Functions (CRHF), and digital signatures. It will be shown that the latter three are relevant for information authentication. Examples of general assumtios to which these primitives can be reduced are the exsistence of one-way functions, injections, or permutations, and the existence of trapdoor one-way permutations. A third aspect is the efficiency of the reduction,i.e., the number of executions of the basic function to achieve a cryptographic primitive, and the number of interactions between the players in the protocol.

Several lines of research have been followed. A forst goal is to reduce cryptographic primitives to weaker assumptions, with as final goal to prove that the reduction is best possible. A different approach is to produce statements about the impossibility of basing a cryptographic primitive on a certain assumption [80]. One can also try to improve the efficiency of a reduction, possibly at the cost of a sronger assumption. If someone wants to to build a concrete implementation, he will have to choose a particular one-way function, permutation, etc. The properties of a particular problem that is believed to be hard can be used to increase the efficiency of the solutions. Examples of problems that have been intensively used are the factoring of a product of two large primes, the discrete logarithm problem modulo a prime and modulo a composite that is the product of two large primes, and the quadraic residuosity problem.

The complexity theoretic approach has several advantages:

- It results in *provable secure* systems, based on a number of assumptions.

- The constructions of such proofs requires formal definitions of th cryptographic primitives and of the security of a cryptographic primitive.

- The *assumptions* on which the security of the systems is based are also defined formally.

The disadvantage is that the complexity theoretic approach has only a limited impact on practical implemetations, due to limittions that are inherently present in the models.

- In complexity theory, a number of operations that is polunomial in the size of the input is considered to be feasible, while a superpolynomial or exponetial number of operations in the size of the input is infeasible. In an asymptotic setting, abstraction is made from both constant factors and degrees of the polynomials. This implies that this approach gives no information on the security of concrete istances (a practical problem has a finite size). Secondly, the scheme might be impractical because the number of operations to be carried out in polynomial in the size of the input but impractically large.

- The complexity theoretic approach yields only results on the worst case or average case problems in a general class of problems. However, cryptographers studying the security of a scheme are more interested in the subset of problems that is easy.

- Complexity usually deals with single isolated instances of a problem. A crypt-analyst often has a large collection of statistically related problems to solve.

It is interesting to remark that the starting point of this approch was the informal definition of a one-way function and a trpdoor one-way permutation in the seminal paper of W.Diffle and M.Hellman [95]. The first practical public-key cryptosystems were based on the hardness of factoring the product of two large primes (the RSA system proposed by R.Rivest, A. Shamir and L. Adleman [81]) and on the subset sum or knapsack problem (proposed by R. Merkle and M. Hellman [82]). However, it is not possible to show that the security of these systems is equivalent to solving the underlying hard problem. The best illustration of this fact is the fate of the knapsack public-key cryptosystems, that are almost completely broken [83,84]. Although no one has been able to show that the security of RSA public-key cryptosystem is equivalent to factoring, no attack on the RSA scheme has been proposed that is more efficient than factoring the modulus. Historically, the attempts to prove the security of RSA resulted in the construction of new schemes for wich it was possible to prove that breaking the scheme is equivalent to factoring. The next step was to design systems based on other assumptions and finally to generalize these assumptions.

1.11 System based or practical approach

In this approach schemes with fixed dimensions are designed and studied, paying special attention to the efficiency of software and hardware implemetations. The objective of this approach is to make sure that breaking a cryptosystem is a difficult problem for the cryptanalist,
By trial and error procedures, several cryptanalytic principles have emerged, and it is the goal of the designer t avoid attacks based on these principles. Typical examples are statistical attacks and meet in the middle attacks.

The second aspect is to design buiding blocks with provable properties. These building blocks re not only useful for cryptographic hash functions, but also for the design of block ciphers and stream ciphers. Typical examples are statistical riteria, diffusion and confusion, correlation, and non-linearity criteria.

Thirdly, the assembly of basic building blocks to design a cryptographic hash functions can be based on theorems. Results of this type are often formulated proven in a complexity theoretic setting, but can easily be adopted for a more practical definition of “hardness” that is useful in a system based approach. A typical example is the theorem discovered independently in [85] and [86], stating that a collision-resistant hash function can always be constructed if a collision-resistant function exists, where the first reference uses a complexity theoretic approach and the second a more pracical definition. A similar observation holds for the theorem tht hash functions can be parallelized efficiently under certain assumptions, where a complexity theoretic approach was proposed in [85] and a practical approach independently by the author in [87]. But even if the results are formulated directly in a practical approach, interesting results cane be produced.. A nice example is the design of a collision resistent hash function [86] based on the assumption thtat the underlying block cipher is random.

1.12 Message Authentication Code (MAC)

It was explained how the authenticity of informationcan be verified through the protection of the secrecy and/or the authenticity of a short imprint or hashcode. In this section informal definitions will be given for a hash function that uses a secret key (Message Authentication Code or MAC).

The term hash functions originates historically from computer science, where it denotes a function that compresses a string of arbitrary input to a string of fixed length. Hash functions are used to allocate as uniformly as possible storagefor the records of a file. The name hash function has been also widely adopted for cryptographic hash functions or cryptographically strong compression functions but the result of the hash function has been given a wide variety of names in the cryptographic literature: hashcode, hash total, hash result, imprint, (cryptoographic) fingerprint, test key , condensation, Message Integrity Code (MIC), message digest, etc.

Message Authentication Codes have been used for a long time in the banking community and are thus older than the open research in cryptology that started in the mid seventies. However, MAC’s good cryptographic properties were only introduced after the start of open cryptologic research.

Definition 1. *A* **MAC** *is a function satisfying the following conditions*:

1. The description of h must be publicly known and the only secret information lies in the key (extension of Kerchoff's principle).

2. The argument X can be of arbitrary length and the result h(K,X) has a fixed length of n bits (with n≥ 32 . . . 64).

3. Given h, X and K, the computation of h(K,X) must be easy.

4. *Given h and X, it is "hard" to determine h(K,X) with a probability of success "significantly higher" than 1/2^n . Even when a large set of pairs{ X_i,h(K, X_i)} is known, where the X_i have been selected by the opponent, it is "hard" to determine the key K or to compute h(K, $X^{'}$) for any $X^{'} \neq X_i$. This last attack is called an adaptive chosen test attack.*

Note that this last property implies that the MAC should be both one-way and collision resistant for someone who does not know the secret key K. This definition leaves openthe problem whether or not a MAC should be one-way or collision resistant for someone who knows K.

It has been explained that cryptographic hash functions can be used to protect information authenticity and to protect againstthe threat of repudiation.

The simplest approach is certainly the use of a Message Authentication Code or Mac. In order to protect the authenticity of information. The authenticity of the information now depends on the secrecy and authenticity of the secret key and can be protected and verified by anyone who is privy to this key. Several options can be considered , but all share the problem of a double key management : one for the authentication and one for the encryption. It is tempting to use the same key twice, but this has to be discouraged strongly: not only are there dangerous interactions possible between the encryption scheme and the MAC, but the management of both keys should be different. The advantage of this approach is a high security level, owing to the complete separation of protection of privacy and authentication.

The most straightforward option is to calculate the MAC, append it to the information and subsequently encrypt the new message. An alternative is to omit the encryption of the MAC. The third solution is to calculate the MAC on the enciphered message. The advantage is that the authenticity can be verified without knowing the plaintext or the secret key of the encryption algorithm, but in gneral it is preferable to protect the authenticity of the plaintext instead of the authenticity of the ciphertext.

One of th most important issues of a MAC is the capability to guarantee the non-repudiation of origin [88]. The technical term non-repudiation denotes a service wherebythe recipient is given guarantee of the message's authenticity, in the sense that the recipient can subsequently prove a third party that the message is authentic even if its originator subsequently revokes it. The need for a "purely digital, unforgeable, message dependent signature" has been identified by W.Diffie and M.Hellmann in their 1976 paper [89]. In the same paper the authors propose an elegant solution based on trapdoor one-way permutations. The first practical proposal of a public-key system with digital signature capability as suggested by the title of the original paper was the RSA cryptosystem [81]. Its security is based on the fact that it is "easy" to find two large primes, but "hard" to factor this product. Subsequently new schemes appeared, based on the other number theoretic problems like the descrete log problem [90]. The complexity theoretic approach has resulted in provably secure digital signature schemes based on claw-free pairs of pemutations [91,92], one-way permutations [93], and finally on one-way functions [94], which can be shown to be optimal. A remarkable evolution here is the intertwining in some schemes between the signature and the hashing operation. The 1988 CCITT X.400 and X.500 recommendations [95,96] and in particular CCITT X.500 [97] are an example of recent standards that offer security facilities that are based on digital signatures. Digital signature schemes based on the practical approach were further optimized but have not received too much attention.

1.13 Contribution of the thesis.

The general objective of this dissertation, is to advance the research in the area of fragile watermarking techniques.

The two proposed algorithms are both self-authenticating (authentication need do not involve explicit information derived from the original image or watermark) image integrity verification systems achieved by using two "orthogonal" digital information for the watermarking process. This makes authentication more practical for Internet multimedia applications.

The contributions of this thesis can be grouped into the following three tasks:

- Proposal of two novel algorithms in different domains, both realized for image authentication.

- Comprehensive analysis of algorithm feasibility, overall performance, and the capability of both techniques, even though in two different ways, to be resistant to cropping attacks and lossy communication channels.

- Simulation of the scheme along with comparisons between the two fragile watermarking techniques.

In Chapter 2, the existing work on image authentication is dicussed. Chapter 3 looks at the preliminaries helpful to elucidate the novel ideas of the thesis. In Chapter 4-5, the holographic and the parity-checking watermarking algorithms are proposed. Chapter 6 includes conclusions and comparison results.

CHAPTER 2

EXISISTING WORK

In this chapter, we review the topic of the image authentication. We classify the different algorithm.

2.1 Classical Image Authentication

The classical scenario for image authentication can be described as follows [98]: a sender S wants to transmit a digital image I to a receiver R. When the image I^r is eventually delivered to R, by means of a network facility or any other media capable of storing digital data, an effective authentication scheme must ensure with high probability that:

- The image I^r received by R is exactly the same as the I the sender S has sent. (integrity verification).

- The receiver R can verify the alleged source of the image I^r, i.e., R can determine whether I^r has been actually sent by S, or has been forged by a pirate. (Alleged source verification).

- R can demonstrate that I^r was actually sent by S, and S cannot deny having sent I^r. (Non repudiation-property).

2.2 Hashed Message Authentication Code HMAC

A hash function[1], such as MD5 [99] and SHA-1[100], produces a one-way message digest, a "fingerprint" of a file, message, or other block of data. The hash based MAC[2] [101] encrypts the hash value of the message with a secret key shared by the sender and the receiver. This technique assumes that two communicating parties, say A and B, share a common secret key K_{AB}. When A has a message M to send to B, it calculates the message authentication code as a function of the message and the key: $MAC_M = F(K_{AB}, M)$, where $F(:)$ must fulfill several properties to be considered as secure [98]. The message plus the MAC code are transmitted to the intended recipient. The recipient performs the same calculation on the received message, using the same secret key, to generate a new message authentication code. The received code is compared to the calculated code. If we assume that only the receiver and the sender know the identity of the secret key, and if the received code matches the calculated code, then

1. The receiver is assured that the message has not been altered. If an attacker alters the message but does not alter the code, then the receiver's calculation of the code will differ from the received code. Because the attacker is assumed not to know the secret key, the attacker cannot alter the code to correspond to the alterations in the message.

2. The receiver is assured that the message is from the alleged sender. Because non one else knows the secret key, no one else could prepare a message with a proper code.

2.3 Digital Signature Algorithms

The asymmetric encryption algorithms published in the late 1970s, such as RSA [102], in conjunction with secure hash functions, are digital signature algorithms[3], which allow the sender to associate its unforgeable imprint with the digital image, so that the receiver, R, can check its integrity and its source. Non-repudiation is also guaranteed. The asymmetric encryption involved the use of two separate keys: a public key made public for others to decrypt a received message, and a private key is known only to its owner to encrypt the original. When A has a message M to send to B, it calculates the digital signature sig_M as a function of the hashed message $H(M)$ and the private key $K_{private} : sig_M = F(K_{private}, H(M))$. The message plus digital signature are transmitted to the intended receiver. The receiver performs the same Hash calculation on the received message to generate a new hashed message. The receiver also decrypts the received signature sig_M, using the public key K_{public}, to get the received hashed message. The received hashed message is compared to the new hashed message. If we assume that only the sender knows the identity of the secret key, and if the received hashed message is identical to the new hashed message, then

1. The receiver is assured that the message has not been altered. If an attacker alters the message but does not alter the code, then the receiver's calculation of the hashed message will differ from the new hashed message. Because the attacker is assumed not to know the private key, the attacker cannot alter the code to correspond to the alterations in the message.

2. The receiver is assured that the message is from the alleged sender. Because no one else knows the private key, no one else could prepare a message with a proper digital signature.

Digital signature algorithms are very popular in image security applications. In 1993, G.L.Friedman proposed a “trustworthy digital camera” [103]. An embedded microprocessor in the camera authenticates the output files. The private key is securely stored in the microprocessor, while the public one is printed like a serial number on the camera itself and can be used to prove that the digital photographic images have not been subjected to tampering.

The classical authentication techniques are well suited for assessing the credibility of images to be published in newspapers or magazines, or to be presented as evidence in a court of law. The classical techniques are fragile to content preserving modification, due to the properties of secure hash functions and of encryption algorithm. If x and y are images such that $x \neq y$, the hash and encryption output of x is different from that of y with high probability. In the Internet environment, lossy compression, file format conversion, and undesired erroneous modification of single pixels is common. Thus, the classical techniques fail in authentication even if the semnatic content is preserved. Moreover, the integrity verification mechanism of classical techniques is not capable of locating the modified regions of received images. Since the classical techniques that involve “hard” authentication, are inappropriate for the unavoidably lossy Internet, a more flexible definition of “integrity”, which is capable of discriminating content preserving manipulations from malicious tampering, is needed.

2.4 MAC and digital signatures

For a MAC and digital signature, modifications in the image do not necessarily alter the authenticator. Ideally, a MAC and a digital signature techniques tolerate small or neglegible distortions caused by compression or other standard manipulations performed in image communications, but they do not tolerate other malicious tampering that modifies the image content. The MAC and Digital Signatures could be further classified into two categories: signature generation dependent on bit representation of image coefficients, and signature dependent on image features. The following two subsections discuss some current work under the above two categories, respectively.

2.5 Signature Dependence on Bit Representations of Image Coefficients.

For this class of the techniques the Message Authentication Code (MAC) or digital signature is computed from the image bit plane of an arbitary (more general) domain. In [104], Xie and Arce re-formulate the binary message information in the spatial domain with a key by zero padding and row permutation. Then they employ the XORing key operation, row binary calculation, and column binary calculation to obtain a new kind of Approximate Message Authentication Code (AMAC). In contrast to a conventional MAC, the proposed AMAC produces “similar” outputs for “similar” inputs which allows images with low levels of distortion to be authenticated.

Xie and Arce improved upon the work of [104] in [37] by taking the most significant bits of the mean of each DCT 8×8 block, then composes a bit matrix with these significant bits. The same row and column calculations as [104] are taken on the bit matrix to get Image Message Authentication Code (IMAC). The IMAC is better than the AMAC for tolerance to image compression, and allows easier location of tampering. In [105], Fridirch takes random scan of DCT 64×64 blocks to get a sequence, quantizes the sequence by a threshold, and then extracts robust bits to form a digital signature, since the robust bits can stay invariant to small image changes. This kind of digital signature can be used on JPEG format original images, but fails if the image is compressed after the signature is obtained.

Moulin and Koon take random tiling of each DWT subband of the image, compute mean and variance, and then round the calculated mean by a randomized quantization seed, with a Reed-Muller error correcting decoder to produce a binary hash value in [106]. The hashed value is used for image authentication. Since they take the mean of tiles as input, which will not change greatly if there are small changes on the pixels in the tiles, the hashed values should produce the same result for small image distortions.

2.6 Signature Dependence on Image Features.

Invariant features are extracted from the images, then encrypted by a private key (from PKI) to get a new Message Authentication Code (MAC) or digital signature. Algorithmas work in one of three common domains:spatial,FFT,DCT or DWT.

In [38], Liu and Lou quantize and encode with less code the mean of randomly chosen 8×8 blocks in the spatial domain to get a digital signature, since they believe that the mean of random blocks should be more steady against small distortion of images than every pixel value.

Lin and Chang look at an invariant feature in the DCT domain – the distance between pixels at corresponding locations in different blocks in [39][107][108][109]. The algorithm forms 8×8 blocks into pairs, selects coefficients in low frequency, and iteratively compares the distance of corresponding coefficients δ in blocks with a set of ks^4 to get digital signature. The digital signature gets a better result on compression than conventional digital signatures.

In [40][41], Liao and Lu focus on the distance of corresponding DWT coefficients in different scales and orientations. They quantize and encode the distance of corresponding DWT coefficients in different scales and orietations to create a digital signature. Their algorithm has a similar resulat on compression as [39][107][108][109].

2.7 Watermarking

An authentication watermark can be a logo or a digital signal with well defined spatial and frequency domain properties, or a binary feature extracted from the image content. The watermark is superimposed on the image so that no perceptible distortion is introduced. The receiver extracts the embedded signal from the waterrked image and compares it to the original one, or the one generated from the received image, using the same method as the sender. The authenticator of watermarking algorithms is inseparably bound to the content, hence simplifying the logistical problem of data handling. The authentication information is lumped and localized in the image, so the modifications can be well located.

The security of watermarking algorithms relies on some secret information that is available to the sender and the receiver, such as keys, watermark location, watermak length, etc., for generating, embedding, and extracting the mark. The scheme is secure as long as it is impossible or at least computationally unfeasible to tamper with the embedded watermark or to re-embed a proper watermark without the knowledge of above mentioned secret information. In the following two subsections, we discuss current existing fragile watermarking schemes and also semi-fragile schemes.

2.8 Fragile Watermarking

The fragile watermarking algorithms are concerned with comlete integrity verification . The slightest modification of the host image will alter or destroy the fragile watermark.

In [110], Yueng embeds a binary logo of the same size as the host image by means of a key dependent look-up table (LUT) that maps every possible pixel luminance value to either 0 or 1. The watermark is inserted by adjusting the Least Significant Bit (LSB) value of each image pixel in the spatial domain to match its corresponding LUT value. At the receiving side, the LUT can be reconstructed due to the knowledge of the secret. The integrity verification can be performed either by simple visula inspection of the extracted mark, or by automated comparison with the original one. The watermarking is very sensitive to any distortion on the image, but it is very vulnerable to block analysis attacks.

Delp and Wolfgang arrange M-sequences of -1's and 1's into 8×8 block in the spatial domain of image in [56]. M-sequences are inserted to corresponding image LSB blocks. The receiver would extract new M-sequences from the LSB, and compare them with the original one for authentication. Their work is very sensitive to any change, since the LSB of an image is very vulnerable.

In [15][16], Wong lets the watermark be an arbitrary binary pattern or logo of the same size as the host image. The mark is embedded in a block style in the LSB of the image: for every block, the LSB plane is zeroed out, and a secure hash function (MD5) is computed with image height and length; its output is then XORed with the corresponding block of the mark, then encrypted with the sender's private key and the result embedded. To perform the integrity check, the LSB plane is extracted from image and decrypted by the sender's public key; then MD5 is applied to each block of the image, with setting the LSB plane to zero. Finally, the result of the hash function is XORed in a block style with the corresponding block of the extracted binary signal. The PKI makes this algorithm very practical and easy for authentication, but the scheme is too sensitive, since the watermark is embedded into the LSB plane of the image.

In [111], Wu takes a very similar LUT as Yueng did in [], but she embeds the watermark in the DCT domain. The watermark is a binary logo or pattern with the same size as the number of 8×8 blocks in an image. The watermark is embedded into the quantized DCT coefficients via a LUT. To authenticate an image, the receiver extracts the watermark with a LUT, and compares the extracted watermark with the original logo.

2.9 Semi-fragile Watermarking

The semi-fragile watermaking algorithms are concerned with integrity of content verification. The semi-fragile watermaking can discriminate common image processing and small content-preserving noise, such as lossy compression, bit error, salt and pepper noise, etc., from malicious content modification.

In [42], Quelez computes the rank order relationship of image projections on three secret directions to an image authenticator. The image authenticator is embedded by proper pixel modifications, so that desired values for the projection are obtained. The receiver computes the rank order relantionship of received image projections on the three same secret directions as the sender, and gets a new image authenticator. The new image authenticator is compared with the authenticator extracted from the image for authentication. Since the rank order relantionship of image projections would not change greatly under some mild distortion, the algorithm can tolerate some common image processing. So the three secret directions should be kept secure for the senders and the receivers.

In [112][113], Marvel and Bancelet take a thumbnail image via successive filtering and sub-sampling of the image and quantizing the outcome to 4 bits. Information bits of the thumbnail image are embedded in each 64×64 block by mixing Gaussian noise samples and using 15-fold repetition coding t pixels with a specific strength. The extraction of the authentication sequence is based on detection of the signs of the estimated watermark samples. The watermark is estimated by using an alpha trimmed mean filter. The verification of the image authentication is based on the hamming distance of the original authentication sequence and on the one extracted from the image after repetition decoding. The original thumbnail image should be available to the receiver.

Lin and Chang use non over-lapping zones to generate and embed watermarking, and the division method of zones is indicated by a secret mapping method using a seed in [43]. Watermarking is extracted in 8×8 DCT domain of an image, based on the invariant distance relantionships between two DCT coefficients of the same position in two separate 8×8 blocks. Watermarking extracted from each 8×8 DCT block pair is embedded back to the same blocks but in different positions. In the authentication process, the system extracts the authentication bits from the watermarking embedding zone of the received image, and compares them with DCT coefficient relationships in the watermarking generation zone. If they match, the image is authentic. Otherwise, the image has been modified. The map of watermarking generation and embedding should be available to the sender.

In [44][45][46][47], Fridrich looks at a robust hash function for watermark generation. The authentication bits are generated as a robust hash of the 64×64 image blocks, which are projected onto 32 basis vectors, and their inner product is quantized to one bit. Some 32 sets of random 64×64 block size noise patterns are generated.

These 32, 64×64 random patterns are summed to form a spread-spectrum signal and mixed to the middle one third of the block DCT coefficients, with DC coefficients free. The received image is divided into blocks of the same size, and the spread-spectrum signal is generated in the same way as in the embedding stage. This spread-spectrum signal is correlated with the middle third of DCT coefficients and compared with a threshold, which is adjusted to render the number of 1's and 0's equal. The tampering decision is based on the probability of obtaining correct symbol out of 32. In this scheme, 32 basis vectors should be available for the receiver.

Fridich and Goljan take the concatenation of the compressed LSB of visited DCT coefficients and the hash of DCT coefficients as the watermark in [48]. The watermark is embedded in the quantized DCT coefficients, which are different for different JPEG quality factors in an invertible way, so that anyone who possesses the authentication key can revert to the exact copy of the original image before authentication occurred. To authenticate the image, the receiver extracts the watermark for the DCT domain by quantization, breaks the watermark into two parts: the compressed LSB of visited DCT coefficients and the hash of DCT coefficients, and compares the hashed first parts with the second parts. The length of the watermark should be known by the receiver.

In [49], Delp and Lin take a pseudo-random, zero-mean unit variance Gaussian noise sequence with a key controlled seed as a watermark. The watermark is placed in the upper triangular positions except the DC component of an "empty" 8×8 DCT matrix. Then the inverse DCT of the matrix is calculated to get a 8×8 spatial pattern, and it is mixed with the image DCT block at a given strength. The watermark is extracted by suppressing the image spectral components in the block. The verification is based on the correlation of the extracted data with different watermark patterns. Since the watermark is embedded into low frequency coefficients in the DCT domain, the scheme could accept mild distortions of the image. The receiver should have the watermark patterns.

Eggers et al. take a binary sequence as authenticator in [50][114]. The authenticator is embedded with a secret dither sequence. The authentication message is embedded in cover image DCT coefficients by dithered quantization rules. The cover coefficients are the 2nd through the 8^{th} coefficients in zigzag order of the 8×8 block. The receiver will extract the authenticator with a secret dither sequence, and compare the extracted one with the original binary sequence. The original binary sequence should be sent to the receiver with the watermark image.

In [52][26], Kundur takes a key dependent random sequence as a watermak, the watermark is embedded in the fourth level Haar DWT domain by quantizing the DWT coefficient to even or odd multiples of a step size. The decision to map the wavelets to odd or even multiples, is randomized via a key. The watermark extraction is the same as its embedding. The four level tampering result could be obtained by comparing the extracted sequence with the original one. The receiver should own the same key dependent random sequence as the sender for the authentication.

Liao and Lu embed two complementary watermarks in the DWT domain by using cocktail watermarking – one positive modulation (PM) and the other negative modulation (NM) based on wavelet coefficients quantization in [53]. There is a random mapping function p to determine the position of NM and PM, which must be stored for watermark detection and should be kept secret such that pirates cannot easily remove the hidden watermarks. The watermark is not a conventional noise-like watermark or bipolar watermark, but a number watermark. The watermark can be extracted from the received image without access to the host image, but needs a mapping function to distinguish the PM and NM. The modification detection is made by comparing the extracted watermark with the original hidden watermark.

In [115], Xie and Arce improve their work [104] in [37] by embedding their signature AMAC or IMAC back to the image with a private key. The watermark is embedded in the LL component of the DWT domain by a non-overlapping 1×3 window: sort the three components in the window, then split the range between the largest and the smallest one into intervals, and finally, change the median according to the watermark as well as according to the region in which the median falls. To authenticate the received image, the received watermark is extracted by a non-overlapping 1×3 window, decrypted with the public key of the sender and compared to the signature of the received image. The length of the watermark should be known by the receiver.

Xie and Arce improve the algorithm in [115] by taking a significant block choosing before watermarking embedding in the DWT domain in [27]. The significant 2×2 blocks are chosen under the SPIHT algorithm by a non-overlapping 2×2 window. The watermark is embedded in the chosen blocks. Since the robust blocks have comparatively smaller difference after common image processing, the improved scheme is better against common compression.

In [116], Xie and Arce improve the work in [27] by taking the binary Sobel edge map of the LL band of the DWT as the watermark. Since the new watermark contains more image content information, the new scheme is more fragile to small content modification.

2.10 Considerations

From the above discussion of current existing work on image authentication, we can see thet there are several limitations to these techniques.

- Since the Internet is a lossy channel, practical image authentication should be more robust to normal noises and lossy compression. Thus, the classical authentication in Section (2.1) and fragile watermarking scheme [110][13][56][15][16] are not appropriate for Internet image authentication.

- Watermarking has advantages over traditional signature-based authentication because it allows for localization of tampering and involves security structures that are inseparably bound to the image.

- The methods that embed/attach image-indipendent feature for authentication [104][106][112][56][113][52][50][114][26] are robust to minor distortions, but fragile to lossy compression because the authentication code is not dependent on compression –invariant image features.

- In contrast, Fridrich’s watermark [44][45][46][47], which focuses on choosing robust coefficients for hash computation and embedding, is very fragile to normal image processing. The use of hash makes the image-dependent watermark highly sensible to changes.

- The exisisting techniques that make use of hashing or AMAC such as [104][106], make determination of the exact locations of tampering infeasible.

- The exisisting authentication schemes, such as [39][107][108][109][40][41][115][27], can tolerate at best mild compression, but cannot suffer moderate compression and common image processing[5].

- Some authentication techniques are robust for common image processing, but may miss small content modification. For example, Xie’s IMAC [] and Liu’s tolerant digital signature[37], which take advantage of the fact that the mean of a block is more steady under common image processing, are robust for common image processing, miss small content modification, since the mean of each block cannot include the same relationship among blocks. Xie’s DWT watermarking[115][27], which takes the IMAC or Sobel edge map of images as a watermark and embeds it to the significant blocks, misses some small modification, since her IMAC and Sobel edge map is too tolerant to small content modification.

- All current semi-fragile watermarking schemes must transmit secret information such as the length or location of the original watermark to allow the authentication by a receiver. Unfortunately, the necessity of this confidential information makes such techniques vulnerable to eavesdropping attacks.

The existing work also gives us three helpful principles for designing an effective image authentication scheme:

1. Embedding the authenticator in the image using watermarking, simplifies the logistical problem of digital signature data handling.

2. Employing a proper content-based adaptive authenticator, which is based on invariant image features, can increase the robustness to common image processing and the fragility to content modification.

3. The embedding method for watermarking should be robust enough against moderate lossy compression.

Based on the above discussion, we conclude that practical semi-fragile watermarking should embed a well-designed adaptive authenticator, which depends on invariant image features, into images robustly, so that mild image processing only affects part of the watermark to help isolate the distortion. Furthermore, the authentication verification should not use explicit information derived from the original image.

CHAPTER 3

THE WATERMARKING

In this chapter, we give some insight about holographic watermarking techniques.

3.1 Digital Watermarking

Digital watermarking is a rather new technique since the 1990s [117][118][119][120][121]. It is associated with the ancient technique of information hiding known as “steganography” or “covered writing”[122]. It involves the *hiding* of a signal in another signal called the host.

A digital watermark is information (digital mark) that is embedded possibly using a secret key in the host data, in a statistically and perceptually invisible way. Depending on the application, a watermark typically may contain information about the origin, status, or recipient of the host data. The watermark signal is unobtrusive and secure in the composite signal mixture called the “watermarked signal”, but can partly or fully be recovered from the signal mixture with the use of the secret key, if necessary. The watermark signal can serve various purposes, including: ownership assertion, fingerprinting, authentication and integrity verification, content labelling, usage control, and content protection [123][124].

3.2 Three Main Components for a Watermarking System

A watermarking system can be broken into three main components:

1. The generating function, f_g, of the watermark signal, W, to be added to the host signal. Typically, the watermark signal depends on a key, k, and on watermark information, i

$$W = f_g(i,k).$$

Possibly, it may also depend on the host data, X, into which it is embedded

$$W = f_g(i,k,X).$$

2. The embedding function, f_m, which incorporates the watermark signal, W, into the host data, X, yielding the watermaked data X_{w}. Typically, the watermark signal depends on a key, K

$$X_{\mathrm{w}} = f_{\mathrm{m}}(X, W, K)$$

3. The extracting function, f_{x}, which recovers the watermark information, W, from the received watermarked data, $\bar{X}_{\mathrm{w}}$, using the key corresponding to embedding and the help of the original host data, X

$$W = f_{\mathrm{x}}(X, \bar{X}_{\mathrm{w}}, K)$$

or without the original host data, X

$$W = f_{\mathrm{x}}(\bar{X}_{\mathrm{w}}, K)$$

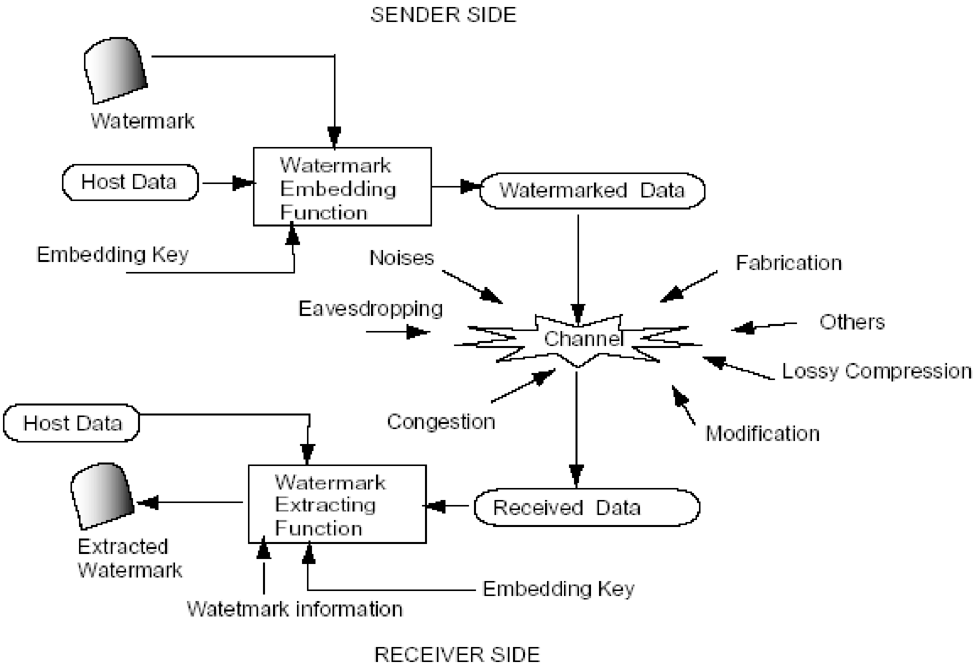


Figure 3.1: Generic watermark scheme

The first two components, watermark generating and wateramark embedding, are often regarded as one, especially for methods in which the embedded watermark is indipendent of the host signal. We separate them for a better analysis of the watermarking algorithms, since some of the watermark is host signal content dependent, with the watermark generating from the host signal and being embedded back to the host signal.

Fig.3.1 shows the generic watermarking scheme. The inputs to the embedding process are the watermark, the host data, and an optional key. The watermark can be of any nature, such as a number, text,binary sequence,or image. The key is used to enforce security and to protect the watermark. The output of the watermarking scheme is the watermarked data. The channel for the watermarked image could be a lossy, noisy, unreliable channel. So the received data may be different from the original marked image. The inputs for extraction are the received watermarked data, the key corresponding to the embedding key, and, depending on the method, the original data and/or watermark information. The confidence measure, indicating how likely it is for the given watermark at the input to be present in the data under inspection, can tell what happened on the image, and provide some information about the image, such as copyright status.

3.3 Image Watermarking Classification

For image watermarking, the host data, X , is in the form of an image. Image watermarking can be classified as visible or invisible. A visible watermarking typically contains a visual message or a company logo indicating the ownership of the image. An invisible watermarked image is visually very similar but not necessarily identical to the original unmarked image. The invisible watermark's existence should be determined only through a watermark extraction or detection algorithm. There are three categories for the invisible watermarking, according to the robustness of watermarking to attacks.

- *Robust Watermarking*: The embedded watermark should be resistant to any processing and/or attack that does not seriously affect the quality and value of the host image. Robust watermarking is used for copyright protection and access control [125][54].

- *Fragile Watermarking*: The watermark should not tolerate any tampering that modifies the complete integrity of the image. Fragile watermarking is used for strict image authentication and integrity verification.

- *Semi-fragile Watermarking*: The watermark should tolerate occasional noise and common image processing such as lossy compression, but be fragile to any malicious tampering that modifies image content. Semi-fragile watermarking is used for soft image authentication and integrity verification.

We focused on fragile watermarking for this thesis.

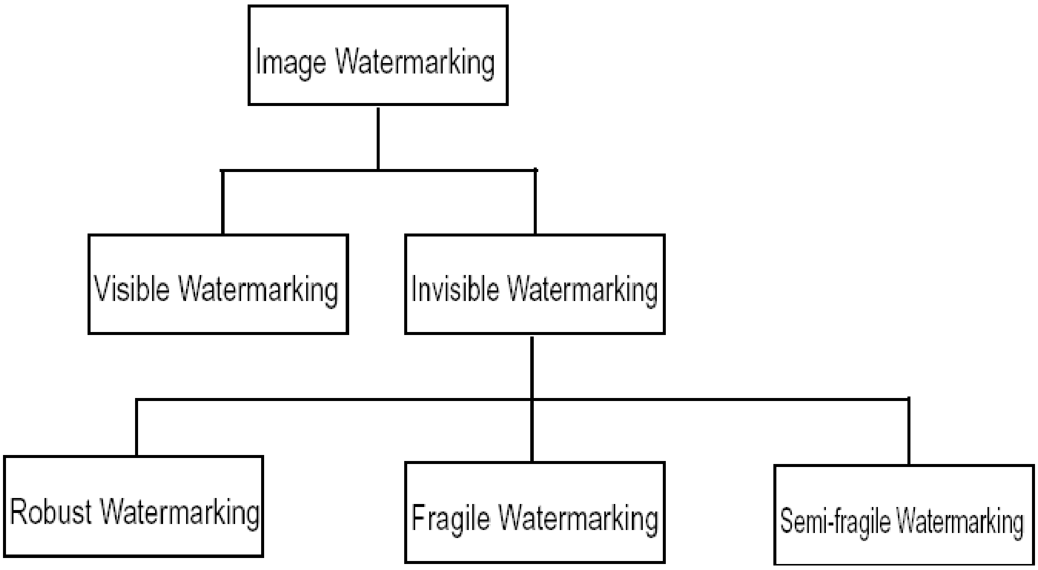


Figure 3.2: Image watermarking classification

3.4 Necessary Properties of Semi-fragile Watermarking

The semi-fragile watermark should satisfy the following specific conditions for soft image authentication and integrity verification:

1. *Perceptual Invisibility*: The watermark embedding should not produce perceivable data alteration. X_w should not contain any perceptual distortion that reduces the quality of the original image X.

2. *Key Uniqueness*: Different keys should not produce equivalent watermarks for the same Image.

3. *Non-invertibility*: The key must not be known if the embedding method and watermark are known to the unauthorized parties.

4. *Product Dependency*: When the watermark generating method is applied to different products with the same key, different watermarks should be produced.

5. *Reliable Detection*: The positive detector output should have an acceptable minimal degree of certainty.

6. *Robustness*: The watermark should be robust to content preserving alteration.

- Mild compression
- Histogram equalization
- Low-pass filtering
- Gaussian noise
- Salt and pepper noise
- Random bit errors in transmission and storage
- Sharpening

7. *Fragility*: The watermark should be fragile to malicious content changing attacks.

8. *Computational efficiency*: The watermarking algorithm should be effectively implemented by hardware or software. Especially, the watermark detection algorithm should be fast enough for multimedia data monitoring in the distribution network.

3.5 Image Authentication

The image authentication in this thesis is to detect *modification* and *fabrication*, and to distinguish the content tampered images from the credible images, in order to deter active attacks on Internet images. As shown by Fig. 3.2, a general image authentication system[104]-[116] must consider three factors: the sender, the transmission channel, and the receiver. The sender creates an authenticator, encrypts it, and sends it with the image. Then the image, with the authenticator, is sent via the unreliable channel. The receiver obtains the possibly corrupt image and authenticator. Using the authenticator generating algoritnm, she/he constructs another authenticator from the received image. By comparing the locally generated authenticator with the received one, a decision can be made about whether or not and where the image has been modified.

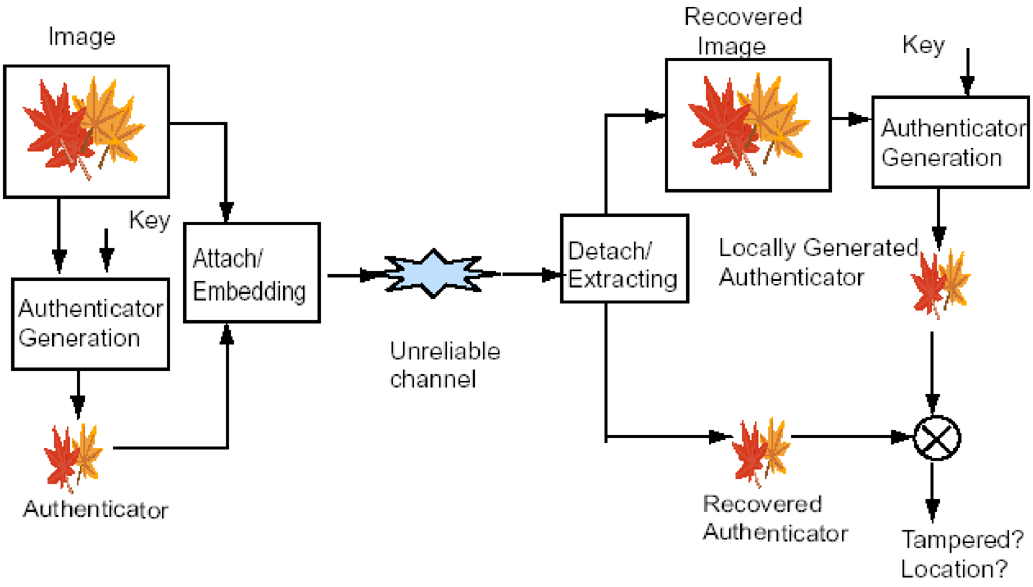


Figure 3.3: Image authentication system

We define I as the original image with dimension $N_x \times N_y$, and $N_x \in Z^+, N_y \in Z^+$, and I^r as the recovered image. Let F be the function to generate the authenticator from images. We also let $\{A\}$

be the authenticator of *original image*, $\left\{A^{re}\right\}$be the received authenticator, and $\left\{A^{r}\right\}$be the authenticator generated from the *recovered image*. Therefore, we have:

$$\{A\} = F(I)$$

$$\left\{A^{'}\right\} = F\left(I^{'}\right)$$

To keep $\left\{A^{\mathrm{re}}\right\}$equal to $\{A\}$is very important in conventional "hard" image authentication, because a single message bit difference affects the calculation of the checksum bits and changes each one in roughly half of tne cases. If $\left\{A^{\mathrm{re}}\right\} \neq \{A\}$, the decision is made on the *Hamming Distance* between $\left\{A^{\mathrm{re}}\right\}$and $\left\{A^{\mathrm{r}}\right\}$ in the majority of algorithms [104]-[116] as follows:

$$D^{\mathrm{H}}\left(\left\{A^{\mathrm{re}}\right\},\left\{A^{\mathrm{r}}\right\}\right) < \tau \text{ authentic}$$

$$D^{\mathrm{H}}\left(\left\{A^{\mathrm{re}}\right\},\left\{A^{\mathrm{r}}\right\}\right) \geq \tau \text{ tampered}$$

where τ is threshold, and $D^{\mathrm{H}}(...)$is the Hamming distance operation.

Since the Internet, channel is an open, unreliable lossy communal channel, and traffic and user data are possibly modified from attacks in various forms of eavesdropping and packet sniffing, the received authenticator $\left\{A^{\mathrm{re}}\right\}$is hardly the same (bit equality) as the original authenticator $\{A\}$, and the sent image I is not the same as the recovered image I^{r}. For traditional "hard" image authentication, a modification of a single message bit affects the calculation of the checksum bits and changes each one in roughly half of the cases[110]-[111]. Images data can tolerate minor changes, due to the existence of irrelevant signal information. The "loss tolerant" feature of an image is exploited in lossy compression for the reduction in file size. In the likely event of "lossy com pression", "occasional" or "low priority" bit losses during transmission, a conventional digital signature and MAC would fail, since the received image data and the signed data are not identical, but the content of images is still the same[109]. So "soft" image authentication is desired for lossy image communication on the Internet. We attempt to achieve a practical semi-watermarking for “soft” image authentication on the Internet by using an holographic watermarking technique based on Fourier transform domain.

3.6 Visibility of Image Distortion

The watermarking embedding is a kind of distortion of the host image. The perceptual invisibility of the distortion of the watermarking is the preliminary condition for common watermarking schemes. The watermarking capacity, which is the amount of embedded information in the host image, is affected by invisibility constraints, robustness requirements, and security issues. In general, if we want to make the watermark bits more robust against attacks, a longer or larger amplitude will be necessary for the embedded watermark. The visual detection thresholds give the maximum space for the watermarking, without noticeable distortion.

Masking occurs when the presence of a signal hides the existence of another. Human Visual System (HVS) models [126][127] indicate that masking effects have different influences in different positions in the spatial pixel domain, the frequency domain, or the frequency-orientation domain. General human vision mechanisms show that masking effects are decided by luminance, contrast and orientation. Luminance masking, with its basic form of Weber’s effect, refers to the fact that the brighter the background is, the higher the luminance masking threshold will be. The visual threshold for a luminance pattern typically depends upon the mean luminance of the local image region. Contrast masking refers to the reduction in the visibility of one image component occasioned by the presence of another component. This masking is also sharply tuned to orientation[128].

Since the watermark in this thesis is embedded in the FFT domain, the visibility of image distortion in the FFT domain is very important.

3.7 PSNR Measure Bound

Assume a $N_x \times N_y$ digital image $I_{x,y}$ and its changed image $\bar{I}_{x,y}$, then Peak Signal Noise Ratio (PSNR) is

$$PSNR = 10\log\frac{(\max I_{x,y})^2}{\frac{1}{N_x \times N_y}\sum_{x-1}^{N_y}\sum_{y-1}^{N_y}(I_{x,y} - \bar{I}_{x,y})^2}$$

The PSNR of an image is a typical measure used for assessing image quality by considering that the just noticeable distortions are uniform in all coefficients in a specific domain, such as special domain, frequency domain, or some transform domain. Since PSNR is more computable and can be used to provide a generic bound for the watermarking capacity. So, we use the PSNR to analyze the watermark embedding distortions on images.

CHAPTER 4

HOLOGRAPHIC WATERMARKING

4.1 Introduction to the holographic watermarking

In this chapter a semi-fragile watermarking is proposed, based on a Computer Generated Hologram coding techniques, adopted and designed to detect any unauthorized modification for the purpose of image authentication.

A semi-fragile watermark is a mark that is readily altered or destroyed when the host image is modified through a linear or nonlinear transformation [25][24]. A semi-fragile watermark monitors the integrity of the content of the image but not its numerical representation. Therefore the watermark is designed so that the integrity is proven if the content of the image has not been tampered with. However if parts of the image are replaced, the watermark information should indicate evidence of forgery.

Image authentication systems have applicability in law, commerce, defense, and journalism [31][32][33][129].

The scope of this chapter is to present an holographic semi-fragile invisible watermarking for digital image authentication. If [130] shows the idea of using holography to watermark an image and [131][132] implement the synthetic hologram with a new encoding technique, this chapter goes through the holographic technique and improve it to achieve the authentication of the whole image and even of small pieces of it. Here, the secret key verification watermarking, proposed in [131][132], is extended into a public key scheme so that the integrity and ownership of the image can be verified using a public key. In such a system, the owner of the image inserts a watermark using a private key (SK). In the watermark extraction procedure, any person can use the public key (PK). A possible scheme to encode the watermark is shown in Figure 4.1(a). In the encoding process a content creator/owner inserts a watermark into an original image.

When a user receives a test image, he uses the detector to evaluate the authenticity of the received image.

The detection process requires knowledge of the "side information". The side information is the public key and the image of the mark. A possible scheme to decode a watermark is shown in Figure 4.1(b).

To compare the recovered watermark with the original inserted one, generally, statistical tests are used. In this work the normalized correlation coefficient is used as statistic test.

The proposed technique has some relevant advantages compared to other semi-fragile schemes [25][5][133][56]. First of all it is cropping resistant, due to the particular intrinsic characteristics of the used holographic watermarking. In addition, the proposed cryptographic scheme allows to have a secure watermarking resistant to data loss; in fact it is possible to extract the mark information also from a watermarked image, transmitted over a noise channel. This is the novelty presented in this method that differs from the usual semi-fragile authentication schemes involving parity-checkings in the spatial domain as [134], which would certainly detect malicious changes but would not withstand data losses and cropping attacks. Furthermore, this technique is addressed to any kind of image not only binary ones.

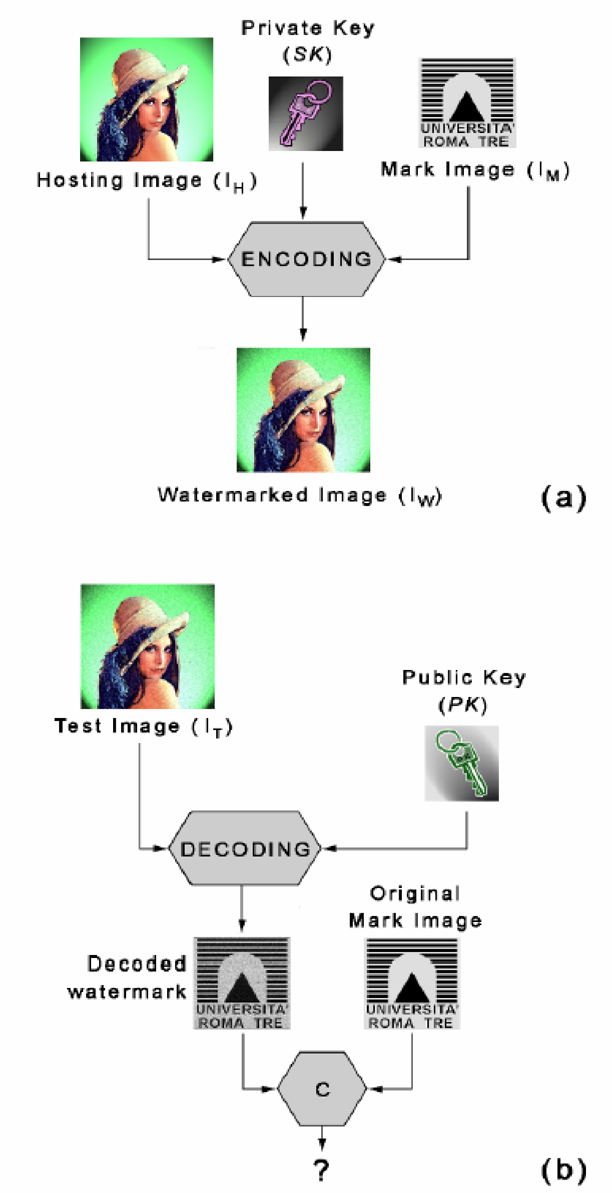


Figure 4.1: Encoding, decoding, and comparing embedded watermarks in a digital image. In the encoding process (a), a content creator/owner inserts a mark into an original digital image. In the decoding process (b), a content owner checks a test image trying to recover a mark, and then compares the recovered image with the original inserted one.

4.2 Fragile Watermarking Based on CGH Coding Techniques.

The proposed technique, of watermarking based on Computer Generated Hologram (CGH), has the important advantage that a content-fragile scheme can be defined: any manipulation changing the visual content of the cover can be detected. Another important characteristic of hologram watermarks is that, even if the embedded data are content-related to the cover, they are noise-like and therefore difficult to be detected and removed.

4.3 Overview on Computer Generated Hologram construction

Optical holography [135] is a technique by which both the amplitude and phase of optical field diffracted from an object of interest are recorded as a hologram in the form of interference fringes. When the optical field from the diffuse object is recorded as hologram, the hologram becomes very similar to a random pattern, because the interference fringes of randomly phase-modulated waves are recorded in the hologram with high density. Nevertheless, the original image of the object can be recovered from the hologram. A Computer Generated Hologram (CGH) image is a hologram computed by numerically simulating the physical phenomena of light diffraction and interference. Also the CGH is similar to a random pattern. Therefore, if a CGH of a mark image is used as input data of the watermarking algorithm, it can be considered as pseudo-noise mask.

The production of hologram using a computer has been discussed in detail in [136, 137], here only the necessary content is presented to understand the following discussion.

In this algorithm, the mask image is hidden in a form of a Fourier-transformed digital hologram of diffused type. Unfortunately, this type of hologram produces, in reconstruction, twin effect (these copies, superimposing themselves, provoke loss of information). To avoid this problem, we simulate an off-axis hologram by means of the following procedure. First of all, the image of the mark ($\mathbf{I}_M$) is resized to 1/64 (in area) of the dimensions of the hosting image (e.g. if host image is 1024 × 1024 the image of mark is resized to 128 × 128). Subsequently, it is duplicated and positioned inside a structure with the same dimensions as the hosting image which must be watermarked. In this way the modified mark image is obtained ($\mathbf{I}_{MM}$), as that shown in Figure 4.2, used to construct the CGH.

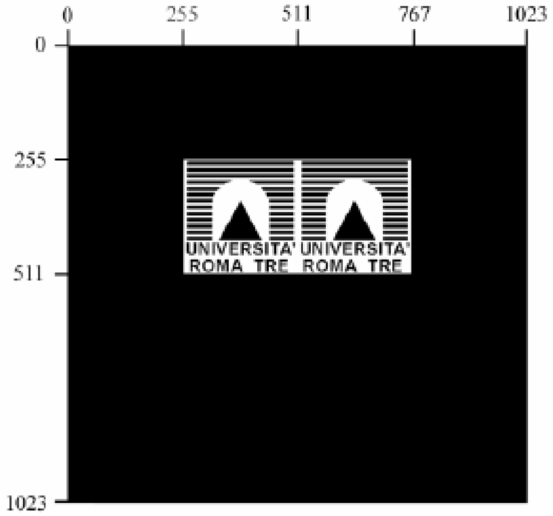


Figure 4.2: Resizing and zero padding of mark image. Mark image at 128 x 128 pixels, padded in a 1024 x 1024 zero matrix.

To make the Fourier-transformed digital hologram, the $\mathbf{I}_{MM}$ image is modulated by a random phase mask $\exp[i\phi(\xi,\eta)]$. The two-dimensional phase $\phi(\xi,\eta)$ is given by random numbers. The $\mathbf{I}_{MM}$ image modulated by the random phase is subsequently numerically Fourier transformed:

$$T(x,y) = \mathbf{FFT}\left\{I_{MM}(\xi,\eta)\exp\left[i\phi(\xi,\eta)\right]\right\} \ . \qquad (1)$$

Now, each element (x,y) of the matrix $\mathbf{T}$ is divided in four sub-elements. The first sub-element represents the real and positive part of $T(x,y)$ (0° angle in the corresponding phasor notation); the second one represents the imaginary and positive part of $T(x,y)$ (90° angle in the corresponding phasor notation); the third the real and negative part of $T(x,y)$ (180° angle in the corresponding phasor notation); eventually the last one represents the imaginary and negative part of $T(x,y)$ (270° angle in the corresponding phasor notation).

After this procedure, the resulting matrix has a dimension four time greater than the original one, due to the fact that each original pixel is now represented by four values. To obtain the CGH with the same dimension of the original image, we have substituted each set of four values, with the related average, made by linear interpolation. In this way we obtain the matrix **SH** which represents the Computer Generated Hologram (also called Synthetic Hologram).

4.4 Watermarking technique – CGH insertion procedure

The matrix **SH** of the mark image is embedded into a hosting matrix image $\mathbf{I}_H$, resulting in the fragile watermarking by means of CGH.

Before embedding the **SH**, the hosting image $\mathbf{I}_H$ is filtered, using the FFT domain, by a Hamming circular filter. In this way all high frequency information is eliminated from the image. This operation is necessary, because the watermarking scheme foresees that the mark can be extracted from the marked image spectra. For this reason the spectra of the obtained filtered image $(\mathbf{I}_F)$ and the spectra of the **SH** have to be spatially separated. By means of the high frequency filtering the hosting image spectra are concentrated only in the low and medium frequencies, whereas the **SH** spectrum is only in high frequency. The Hamming Filter used is:

$$\begin{cases} \text{if} \quad (x-x_c)^2+(y-y_c)^2 \le R & \Rightarrow \quad 0.08+0.46\cdot\left\{1-\cos\left[\dfrac{\pi\sqrt{(x-x_c)^2+(y-y_c)^2}}{R}-\pi\right]\right\} \\ \qquad\quad \text{Otherwise} & \Rightarrow \qquad\qquad\qquad 0 \end{cases} . \qquad (2)$$

Where R is the Filter radius and (x_c, y_c) is the filter center position. In this case (x_c, y_c)correspond to the center of the $\mathbf{I}_H$ spectra, and R is chosen so that the frequency information of the image is not overlapped with **SH** information. Using a mark resized to 1/64 of the host image, it is possible using a radius R equal to the linear dimension of the host image. Figure 4.3 shows a parrot image, with 1024×1024 pixels, filtered with the radius R = 1024. In particular, in Figure 4.3(a) is shown the original image, while Figure 4.3(b) shows the filtered images. Eventually in Figure 4.3(c) is shown a detail of original and filtered image. The Figure 4.3(c) shows that the usage of this kind of filter does not decrease the detail level of the original image.

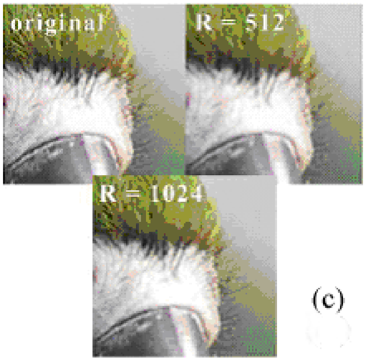


Figure 4.3: The “parrot” images. (a) Original Image; (b) Filtered Image used for inserting a mark of an 1/64 (in area) of the image dimensions – using host image of 1024 x 1024 and using an R = 1024; (c) Detail of original images and filtered ones highlighting that the used filter does not damage the image quality.

Before introducing **SH** inside the filtered image, this is modified to take into account the human eye contrast response. The previous transformation makes possible to apply a brightness-dependent attenuation, in which is applied a greater attenuation value to the lighter image pixels than to the darker ones:

$$SH_{MD}(x,y)=\alpha\left[SH(x,y)\cdot 2I_F(x,y)+SH(x,y)\right] \tag{3}$$

In equation (3) the **SH** and the $\mathbf{I}_F$ are both considered normalized between zero to one in the image space. In this way, the weight of the **SH** is three time greater on the high intensity level pixels. This value is evaluated in empirical way. In other words, during the experiments different weight values have been used between dark and light pixels. The choice of the value “three” is due to optimize the ratio between content information inside the marked image and invisibility.

From now on $\mathbf{SH}_{MD}$ indicates the synthetic hologram utilized in the CGH watermarking scheme. This value is subtracted to the filtered image $\mathbf{I}_F$, obtaining the watermarked image $\mathbf{I}_W$.

The parameter α controls the fragility of the content insert in the host image. It is possible to adjust the α parameter according to the required application. For example, in military images the best value of α is 0.004. On the contrary, for images of reporting news it is possible to use α = 0.1 or more; in this way the content is more resistant to accidental distortions related to the use (e.g. transmission over a noisy channel) of the images.

The resulting image has a difference according to the initial one. To obtain the marked image with the same dynamic of the hosting image, the marked image is modified, according to the following equation:

$$\mathbf{I}_W = \left\{ \frac{(\mathbf{I}_F - \mathbf{SH}_{MD}) - \min[\mathbf{I}_F - \mathbf{SH}_{MD}]}{\max[\mathbf{I}_F - \mathbf{SH}_{MD}] - \min[\mathbf{I}_F - \mathbf{SH}_{MD}]} \cdot (\max[\mathbf{I}_H] - \min[\mathbf{I}_H]) \right\} + \min[\mathbf{I}_H] \ . \qquad (4)$$

The equation (4) shows that the watermarked image $(\mathbf{I}_F - \mathbf{SH}_{MD})$ is firstly normalized between the values 0 and 1, and then its dynamic range is equalized to the $\mathbf{I}_F$-one. Finally the obtained values are shifted to the minimum value of the original image. In this way the marked image $\mathbf{I}_W$ is statistically similar to the original one, with an increasing in mark invisibility.

The pipeline used to obtain the watermarked image by means of this procedure is synthesized in Figure 4. This figure shows as the semi-fragile image watermarking by Computer Generated Holograms method can be applied either to Gray Scale hosting images, or to RGB color ones, working on each color channel.

Different procedures to create and insert holographic watermarking can be found in literature [138][139][140][141][130]. The proposed one has the advantage to be a very performing fragile watermarking scheme, due to the high quality of the inserted synthetic hologram.

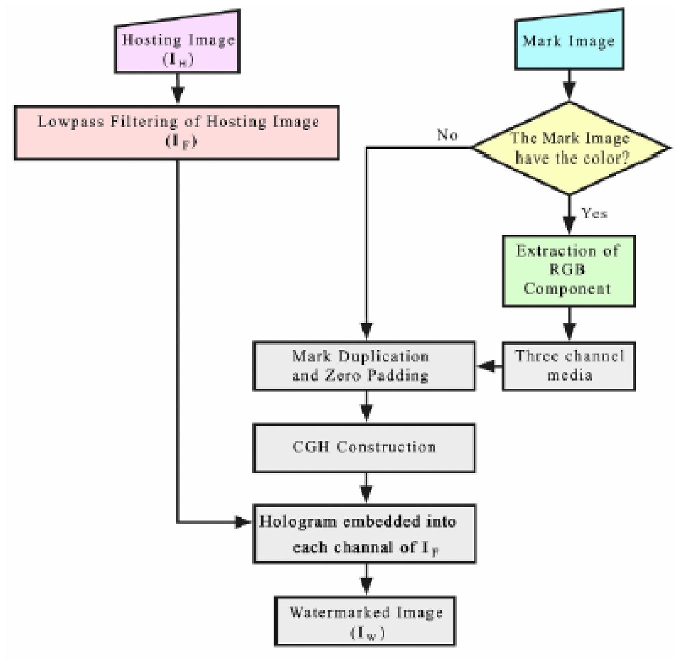

Hosting Image ($\mathbf{I}_H$)
Mark Image
Lowpass Filtering of Hosting Image ($\mathbf{I}_F$)
No
The Mark Image have the color?
Yes
Extraction of RGB Component
Mark Duplication and Zero Padding
Three channel media
CGH Construction
Hologram embedded into each channal of $\mathbf{I}_F$
Watermarked Image ($\mathbf{I}_W$)

Figure 4.4: Pipeline for the production of the CGH-Watermarking.

4.5 Watermarking technique – Mark Detection Procedure

To recover the mark from the watermarked images the procedure is very easy. The FFT is performed on the watermarked image $\mathbf{I}_W$, obtaining four copies of the mark image positioned on the four corners of the frame (see Figure 4.5).

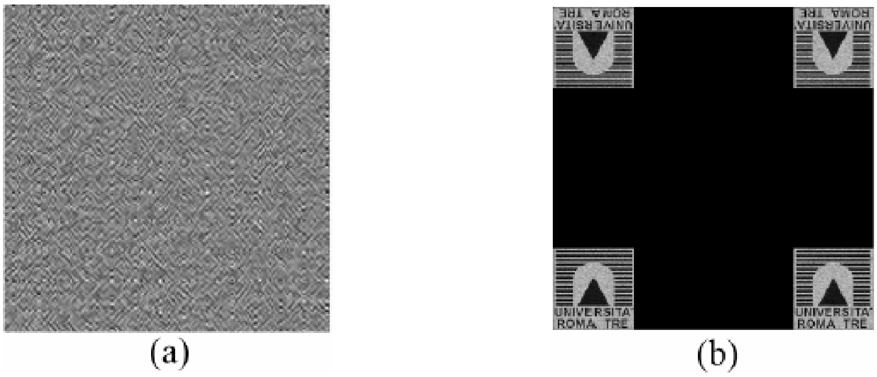


Figure 4.5: (a) Computer Generated Hologram of mark image. (b) Recovered mark copies. Each copy is positioned in one corner and has dimension of 128 x 128.

These four copies are due to the twin image effect, which reproduce, in reconstruction phase, two copies, symmetric with respect to the center of the image, for each image presents into the original mark. Unfortunately, due to the necessity to use a random phase for spreading in all the numerical FFT range the information related to the mark, a reconstructed image affected by a speckle pattern, is obtained. To mitigate this problem, the four extracted copies of the mark are averaged.
The so obtained extracted mark is compared to the original one, by a Threshold Correlation.
If the recovered mark is extracted from a cropped image (with different dimensions from the original image), an averaged mark is obtained with size not equal to the originally inserted one. For this reason, to correctly apply the threshold correlation, the recovered mark is resized to the original one before the comparison.

Cryptographical enhancement

4.6 Overview of the RSA Algorithm

The concept of an asymmetric cryptography (also called Public-Key Cryptography) was developed by Diffe and Hellman [64] and the first practical algorithm was published by River, Shamir and Adleman (RSA) [81]. The RSA algorithm is widely used in Public-Key Cryptography. It is based on the following property of numbers: it is easy to multiply two integers while is very difficult to factor a number that is a product of two large primes. Even with the recent advances in computational number theory and in computer technology, it is,in general, impossible to factor a 1024-bit integer, which is the minimal size recommended by the current standards, within any reasonable amount of time. Like any other public key algorithm, RSA begins with the key generation procedure.

It is supposed to generate the needed asymmetric keys, to be used for signing a message.

Two random large prime numbers, p and q, are chosen. It is computed $n = p \bullet q$. The factors p and q will remain secret. The product $n = p \bullet q$ is made public.

It is reckoned, $\phi(n) = (p-1) \bullet (q-1)$.

It is chosen a small odd number, e, that is relatively prime to $\phi(n)$.

The number e (with $e < n$) has no common factors with $\phi(n)$ $\left[e, \phi(n) \text{ } are \text{ "} relatively\ prime \text{"}\right]$.

It is drawn d such that $e \bullet d - 1$ is exactly divisible by $\phi(n)$. In other words: $d = e - 1 \bullet \text{mod}\left[\phi(n)\right]$.

The public key is $PK = (n, e)$, while $SK = (n, d)$ is the private key.

RSA can also be used to sign a message. Suppose Alice wishes to send a signed message to Bob. She produces a hash value of the message, encodes it with her secret key, and attaches it as a "signature" to the message. This signature can only be decoded with her public key. When Bob receives the signed message, he decodes the signature with Alice's public key, and compares the resulting hash value with the message's actual hash value. If the two agree, he knows that the author of the message was in possesion of Alice's secret key, and that the message has not been tampered with.

To sign a message M (must be smaller than n), Alice computes the signature $S = M^{\text{d}} \bullet (\text{mod}\, n)$. Anyone that knows the corresponding public key, can verify the signature by checking whether $M = S^{\text{e}} \bullet (\text{mod}\, n)$.

Figure 4.6 shows the use of a digital signature realized by means of the RSA algorithm.

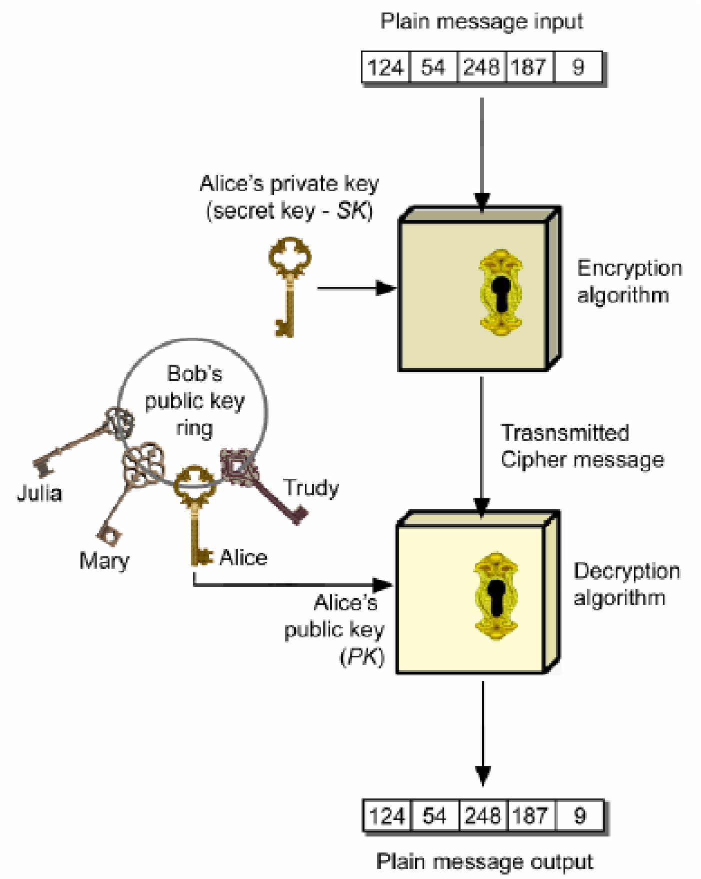


Figure 4.6: Sender (Alice) digitally signs the document, establishing she is the owner/creator. Recipient can prove to someone that Alice, and no one else (including recipient), signed the document.

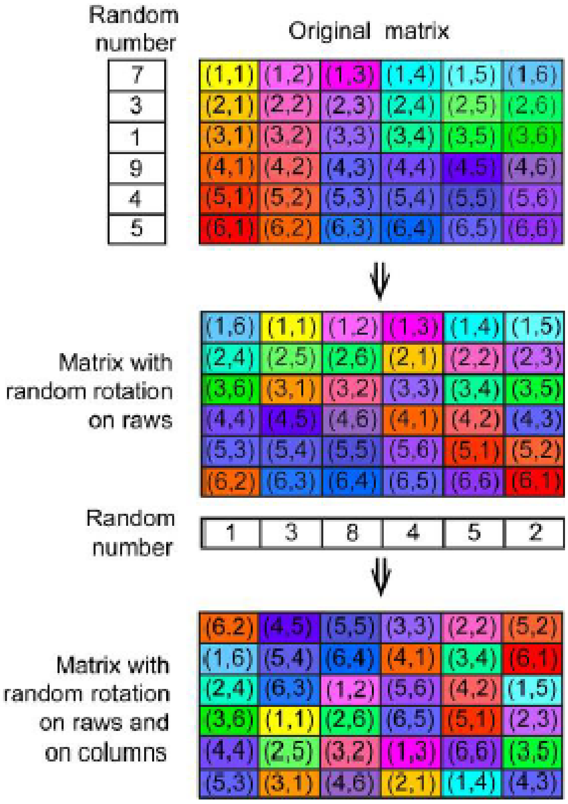
To create a Fragile Watermarking scheme useful for image authentication, the mark has been encrypted with an appropriate cryptographic signature. Because, a digital signature is used, it is possible to verify that the image has not been tampered with, and is also possible to identify the origination of the image.

The used cryptographic signature is derived from the AES [142] and from RSA cryptosystem [64] [81].

Two different vectors are created using a pseudo-random number generator (one for rows and one for columns), called $\mathbf{Rand}_{\mathrm{ROW}}$ and $\mathbf{Rand}_{\mathrm{COL}}$, with dimensions equal to the number of rows and columns of the mark image respectively. After a shift rotation operation is applied to each pixel of each row, using as offset the related $\mathbf{Rand}_{\mathrm{ROW}}$ element value (i.e. to shift the i-th row pixels, the i-th

Rand$_{\mathrm{ROW}}$ element is used). The same approach is repeated also for each pixels of each column, using the other random vector, **Rand**$_{\mathrm{COL}}$.

In the following Figure 4.7 the complete path applied to a 6×6 matrix is shown.



Random number	Original matrix					
7	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
3	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
1	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
9	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
4	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
5	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

⇓

Matrix with random rotation on raws

(1,6)	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)
(2,4)	(2,5)	(2,6)	(2,1)	(2,2)	(2,3)
(3,6)	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)
(4,4)	(4,5)	(4,6)	(4,1)	(4,2)	(4,3)
(5,3)	(5,4)	(5,5)	(5,6)	(5,1)	(5,2)
(6,2)	(6,3)	(6,4)	(6,5)	(6,6)	(6,1)

Random number						
	1	3	8	4	5	2

⇓

Matrix with random rotation on raws and on columns

(6,2)	(4,5)	(5,5)	(3,3)	(2,2)	(5,2)
(1,6)	(5,4)	(6,4)	(4,1)	(3,4)	(6,1)
(2,4)	(6,3)	(1,2)	(5,6)	(4,2)	(1,5)
(3,6)	(1,1)	(2,6)	(6,5)	(5,1)	(2,3)
(4,4)	(2,5)	(3,2)	(1,3)	(6,6)	(3,5)
(5,3)	(3,1)	(4,6)	(2,1)	(1,4)	(4,3)

Figure 4.7: Example of complete encryption path applied to a 6×6 Image. The resulting Image is completely similar to random noise.

To make the cryptosystem the **Rand**$_{\mathrm{ROW}}$ and **Rand**$_{\mathrm{COL}}$ vectors are encoded by means of an asymmetric cryptographic algorithm (RSA algorithm). In asymmetric cryptography, the key for the encryption is not the same as the key for the decryption. Each user has two keys: a Public Key (PK), which everybody knows, and a Private Key (SK), which is kept secret (private).

The encryption of the mark is realized by means of the encoded vector. The cipher mark is inserted, in the hosting image, using an appropriate weight value. In this way, the CGH watermarking is realized. Subsequently, sender (Alice) encrypts **Rand**$_{ROW}$ and **Rand**$_{COL}$ vectors with the secret key of RSA algorithm obtaining two new vectors **E_Rand**$_{ROW}$ and **E_Rand**$_{COL}$(the data are partitioned into blocks, the number of the vector, and the encryption is applied to those blocks sequentially so that lost of ending blocks will not affect the blocks before them; losing a block would mean to lose only a row or a column of the mark). In this way Alice digitally signs the document, establishing she is the document owner/creator. When recipient (Bob) gets the signed document extracts the mark, embedded in the watermarking image, by means of an appropriate FFT technique. This mark must be decrypted by means of **Rand**$_{\text{ROW}}$ and **Rand**$_{\text{COL}}$ vectors. The **Rand**$_{\text{ROW}}$ and **Rand**$_{\text{COL}}$ vectors can be obtained from **E_Rand**$_{\text{ROW}}$ and **E_Rand**$_{\text{COL}}$ using the public key of Alice. Bob obtains **Rand**$_{\text{ROW}}$ and **Rand**$_{\text{COL}}$ vectors signed by Alice by applying Alice's public key to **E_Rand**$_{\text{ROW}}$ and **E_Rand**$_{\text{COL}}$ (see equation 5).

$$\underbrace{\left.\begin{array}{l}\mathbf{E\_Rand}_{\text{ROW}}=\left(\mathbf{Rand}_{\text{ROW}}\right)^{d}\bmod(n)\\ \mathbf{E\_Rand}_{\text{COL}}=\left(\mathbf{Rand}_{\text{COL}}\right)^{d}\bmod(n)\end{array}\right\}}_{Alice}\overset{\text{sending to Bob}}{\Rightarrow}\underbrace{\left\{\begin{array}{l}\mathbf{Rand}_{\text{ROW}}=\left(\mathbf{E\_Rand}_{\text{ROW}}\right)^{e}\bmod(n)\\ \mathbf{Rand}_{\text{COL}}=\left(\mathbf{E\_Rand}_{\text{COL}}\right)^{e}\bmod(n)\end{array}\right.}_{Bob} \tag{5}$$

$$\text{Alice's key}\begin{cases}(n,e)\text{ public}\\(n,d)\text{ private}\end{cases}$$

Therefore Bob can prove to someone that Alice, and no one else (including Bob), must have signed the document .

Figure 4.8 shows the complete scheme of the proposed CGH watermarking.

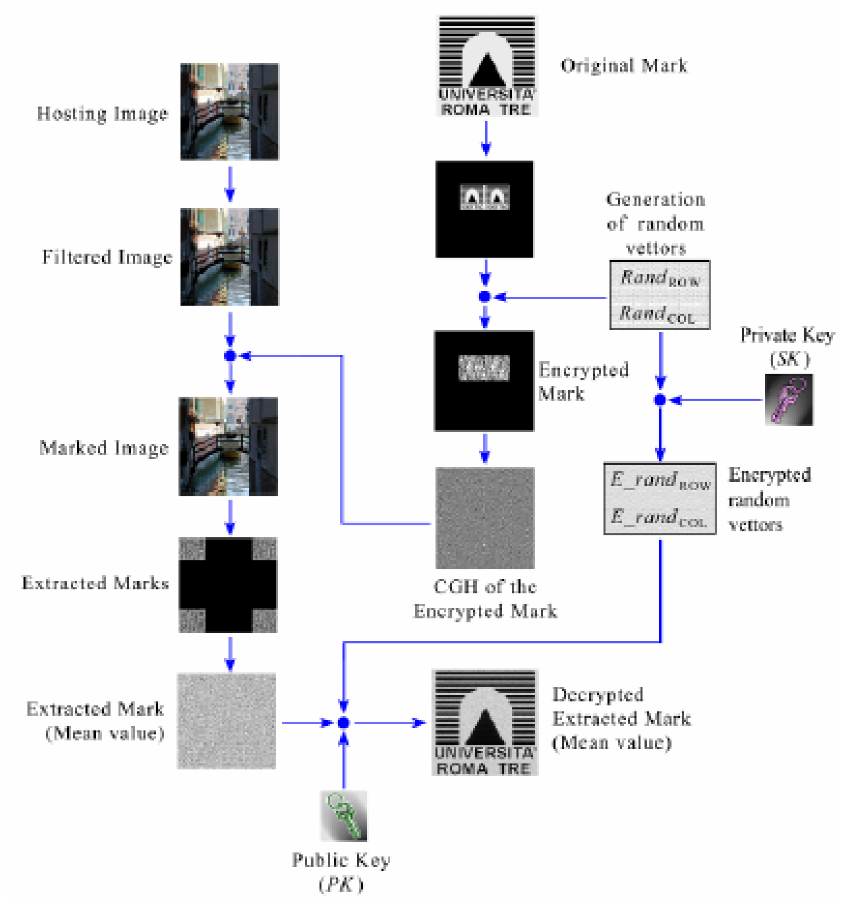


Figure 4.8: Entire sequence of the watermarking algorithm.

The considered holographic technique sticks a speckle noise to the mark, once inserted into the host image. This added noise avoids the possibility of forging watermarked image, by simply copying the watermark from a watermarked image to any arbitrary image with the same size, without leaving a trace of it. In fact, forging an image, in this way, would bring the recovered mark to have a very low cross-correlation (this is because the mark has been inserted twice and this brings to add the speckle noise twice into the recovered mark). So, using one not re-editing image, as a picture, this copying watermarking is not practicable.

4. 7 Experimental Test

To make the tests, hosting images with dimension 1024×1024 pixels have been used. Each one has been filtered to allow the correct mark insertion. The used mark was a gray scale image with dimensions 128×128 pixels. It has to be underlined that there is no limitation in image and mark dimensions; in fact the mark is resized to 1/8 of the width and 1/8 of the height of the hosting image. In this scheme, the mark is not used as it is, but the hosting image is marked by means of the encrypted version of it. To compare the recovered watermark to the originally inserted one, and then verify the presence of forgery and/or tampering, the correlation coefficient is used as statistic test.

Figure 4.9(a) shows an hosting image and a gray scale mark figure 4.9(b). Using the CGH watermarking procedure, described above, it is possible to obtain the watermarked image shown in figure 4.9(c). This watermarked image is realized using a weigh factor α = 0.1. It is easy to see that the watermark is invisible. If one uses the correct keys applied to the watermark extraction procedure to figure 4.9(c), one obtains an output image as in figure 4.9(d), indicating the presence of a proper watermark.

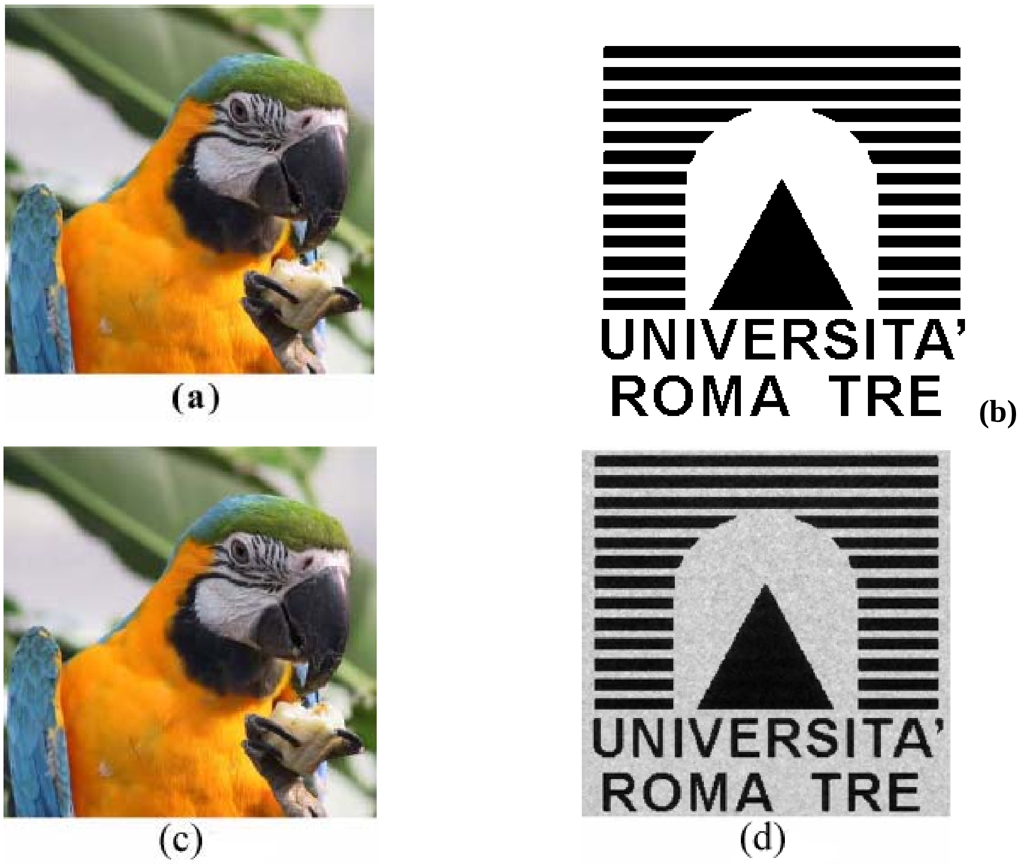
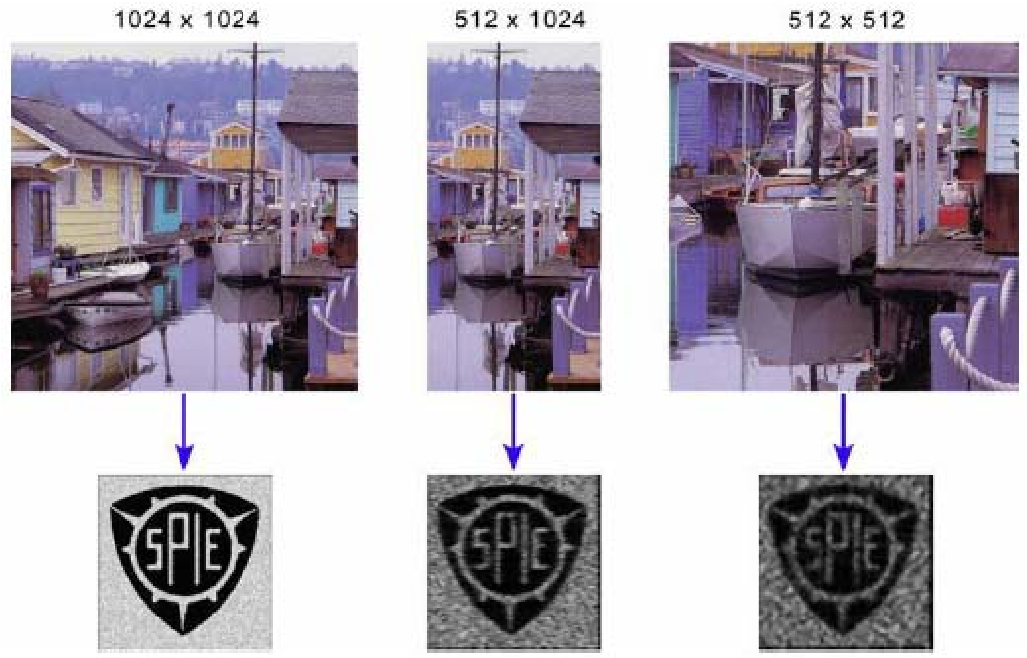


Figure 4.9: Example of watermarking. (a) Hosting image; (b) grayscale mark, (c) watermarked image (α= 0.1); (c) extracted mark.

4.8 Cropping resistant

The CGH watermarking technique is able to detect the presence of the mark not only to full-size image but also to partial ones (cropping resistant). In Figure 4.10 is performed a test of an image of 1024×1024 pixels marked using a weight factor $\alpha = 0.1$.



4.10: Partial images and relative reconstructed watermarked.The "houseboats" image is marked using a coefficient = 0.1. Sized to have the same printed dimension.

This text demonstrates that the reconstructed watermark may be recognized for partial images, although the image quality degrades with a decrease in image size (watermark can still be reconstructed from the partial images by means of CGH watermarking technique). In other words, with the proposed method detecting the parts of the image which were replaced or modified is possible; the watermark information of partial image should indicate evidence of forgery. In fact, dividing the original image in patches, it is possible to identify the presence of tampered zones.

Losses of data in networking diffusion are comparable to cropping attacks. This is the reason why cropping resistance is a great target for a watermarking system. By means of holography this system turns out to be cropping resistant. Figure 4.11 shows the answer of the system to different percentages of cropping of the watermarked image.

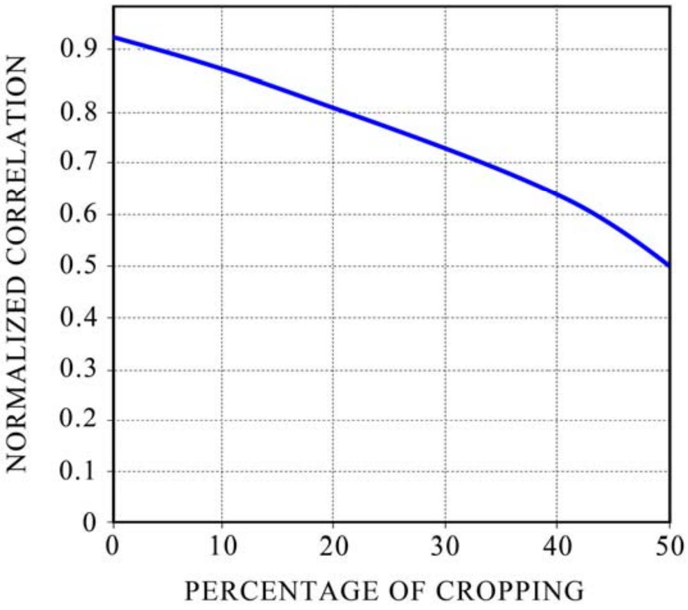


Figure 4.11: Normalized correlation coefficient versus percentage of cropping.

4.9 Robustness to additive noise.

Some processes that might be applied to a Work have the effect of adding a random signal. That is , where $\mathbf{I}_W$ is a watermarked image and **N** is a random vector drawn from some distribution, independently of $\mathbf{I}_W$. Noise processes, where the noise is independent of the watermarked image, are cases of additive noise. Five-hundred images were watermarked with the mentioned additive technique and the corresponding marks were retrieved with normalized correlation coefficients. Figure 4.12 shows the results of these tests. All of these experiments have been carried out using three different weight factors ($\alpha = 0.04$, $\alpha = 0.1$ and $\alpha = 0.2$). The results show that adjusting the α parameter it is possible to change the robustness versus additive noise.

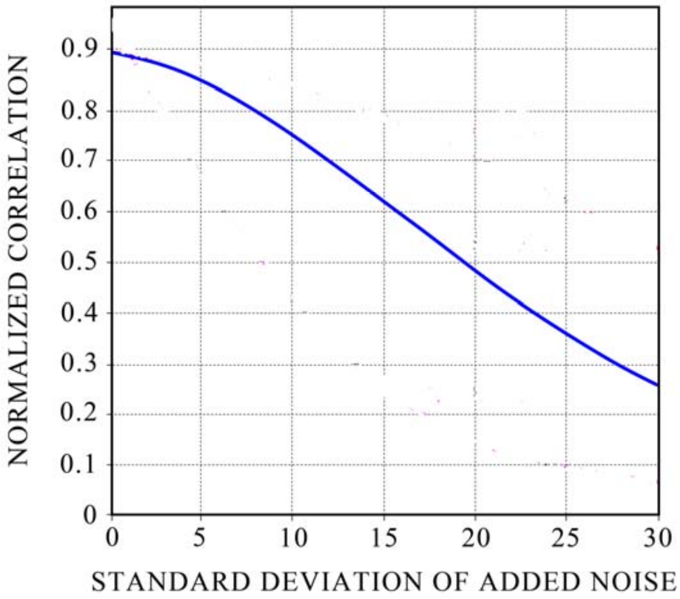


Figure 4.12: Normalized correlation coefficient versus standard deviation of added noise.

4.10 Robustness to Amplitude changes

In reality, many processes applied to watermarked images are not well modelled by additive noise. The change in a watermarked image is usually correlated with the image itself. Many processes are deterministic functions of the image. A simple example is that of changes in amplitude. That is,

$$\mathbf{I}'_W = \beta \mathbf{I}_W , \tag{6}$$

where $\boldsymbol{I_W}$ is a watermarked image and β is a scaling factor (while the scale factor decreases from the severity of the attack). Figure 4.13 shows the normalized correlation coefficient versus scaling factor. These results show that CGH watermarking is relatively robustness versus amplitude change).

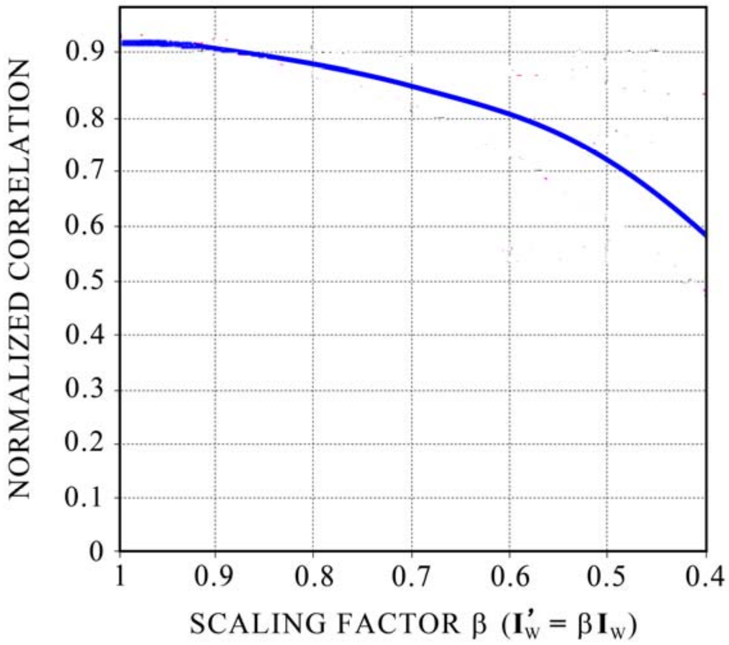


Figure 4.13: Normalized correlation coefficient versus amplitude changes.

4.11 Robustness to linear filtering.

Another common type of signal processing that changes images in a deterministic fashion is linear filtering. That is,

$$\mathbf{I}'_W = \mathbf{I}_W * \mathbf{f} , \tag{7}$$

where $\mathbf{I}_W$ is the watermarked image, **f** is a filter, and $*$ denotes convolution. Five-hundred watermarked images were distorted with Gaussian, low-pass filters of varying standard deviation width. The filters were computed as

$$f[x,y] = \frac{f_0[x,y]}{\sum_{x,y} f_0[x,y]},$$

where

$$f_0[x,y] = \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right),$$

as the width, σ, of the filter increased, more distortion was introduced. Figure 4.14 indicates that the pattern inserted is extremely sensitive to low-pass filtering. The detection values begin falling off sharply when the filter width exceeds 0.25, to show the best result only the weight factor 0.1 has been used.

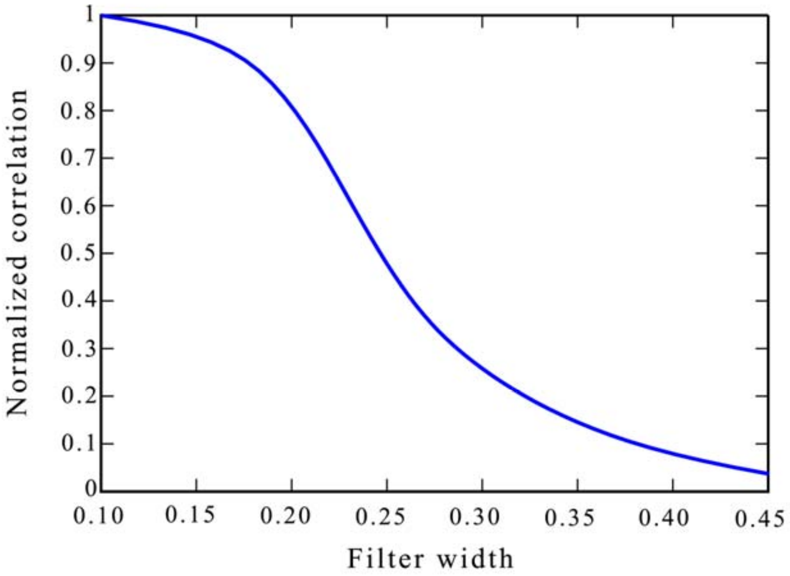


Figure 4.14: Normalized correlation coefficient versus width of linear filtering.

4.12 Robustness to JPEG compression

This technique behaves like a real semi-fragile one when tested with different JPEG compression rates. Five-hundred images were watermarked with the mentioned additive technique and the corresponding marks were retrieved with normalized correlation coefficients. In all experiments, the stronger weight factor (α = 0.1) was used. Figure 4.15 shows the normailzed correlation versus the JPEG compression rate. The plot shows that the CGH mark is quite fragile to JPEG compressions.

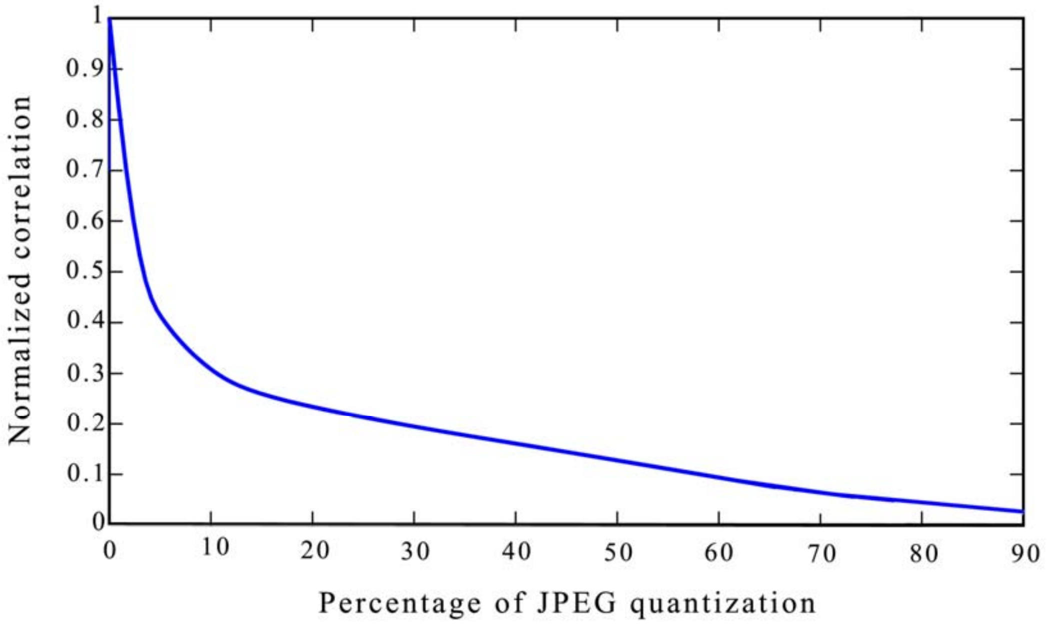


Figure 4.15: Normalized correlation coefficient versus different JPEG compression rates.

4.13 Robustness to copy attack

The novelty of the proposed method is the combined use of RSA-cryptography and holography. Cryptography. It is a way to prevent the possibility of tampering the original watermarked document. The idea of making a forgery might be to destroy the present mark and substituting it with a different one. This is not feasible since the present mark is coded with a private key known only by the owner of the watermarked image. In this way the forgery would go easily detected.

Another possible attack, is the copy attack. A copy attack occurs when an adversary copies a watermark from one watermarked image to another. As such, it is a form of unauthorized embedding. Thinking about it, one might wonder that it should be possible. It suffices to take the two coded vectors, to decode them with the public key and to make the same rotations, as before with the old watermarked image, with the new one. The acquaintance of the private key is not demanded. All has been said is possible but it is not possible to embed the old mark in a new image without this operation going undetected. In fact, each time the mark is recovered from the watermarked image, this holographic technique puts some noise, called speckle noise, into the mark. This noise makes the correlation value of the recovered mark to shrink a little bit. If this operation is carried out twice, the noise is added twice and the final correlation gets a very low value.

4.14 Evaluating Perceptual Impact of Watermarks

The evaluation function is the PSNR (Power Signal Noise Ratio), defined as:

$$D_{PSNR} = -10 \times \log_{10}(D_{NMSE}),$$

where,
$$D_{NMSE}\left(I_W, I_O\right) = \frac{\sum_{x=1}^{W}\sum_{y=1}^{H}\sum_{c=1}^{3}\left[I_F\left(x,y,c\right) - I_H\left(x,y,c\right)\right]^2}{\sum_{x=1}^{W}\sum_{y=1}^{H}\sum_{c=1}^{3}\left[I_H\left(x,y,c\right)\right]^2} \qquad (8)$$

Table 4.1 shows the mean values of the PSNR and the NMSE, obtained from the average of one hundred images watermarked with five hundred different marks.

	mean values
D_{NMSE}	8×10^{-4}
D_{PSNR}	30 dB

Table 4.1:

CHAPTER 5

SPATIAL DOMAIN WATERMARKING

The fragile watermarking technique developed in this chapter combines the theory of bitmap parity checking [143] with asymmetric cryptography [81].

5.1 Watermarking embedding process

The watermark embedding process starts computing the features of the Original Image ($\mathbf{I}_O$) with the aid of a one-way hash function. In this work, the one-way hash function is a resized bitmap version of the Original Image, Figure 5.1. Obviously, different one-way hash function can be used [99, 144].

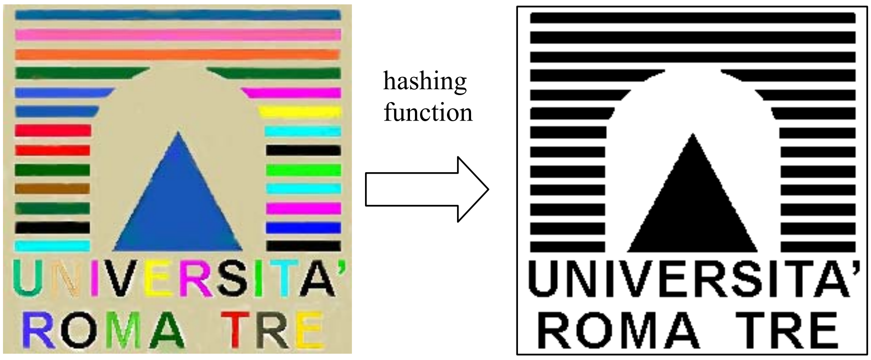


Figure 5.1: Original Image and its hashing function

Subsequently, the extracted feature image is divided in blocks of 32×32 pixels. Each block is surrounded on every side by lines of 1s and 0s, to make the technique resistant to cropping attacks. The final size of the Mark Image will be the same as the Original Image. After that, all the 32×32 blocks are coded with a 1024-bit RSA algorithm, the rows and coloumns of 0s and 1s, are kept uncoded.

The encrypted feature image can be organized as well as a binary matrix. This matrix can be seen as a binary Mark Image ($\mathbf{I}_M$).

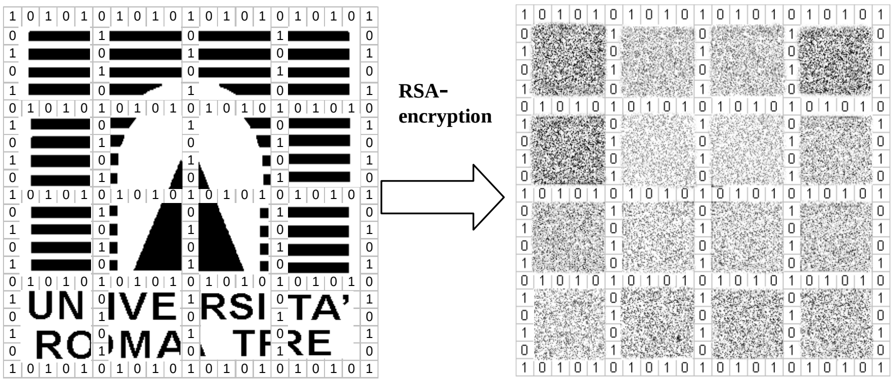


Fig. 5.2: The Mark Image is divided into blocks of 32×32 pixels, in the figure the blocks are much bigger, the Mark Image is usually about 1024×1024 pixels, but that just to show how the system works.

The next step consists to insert, by an invisible way, the Mark Image into the Original Image, to obtain the Watermarked Image $\mathbf{I}_W$. To insert $\mathbf{I}_M$ into $\mathbf{I}_O$, the Original Image is divided in blocks of 32×32 pixels. Each pixel of each block, is represented by the three plane values, Red, Green and Blue. For each triple, we determine the corresponding bitmap triple, by comparing each pixel value of the triple with the mean value of the block of the corresponding color. The system was studied for RGB color images, Figure 5.3.

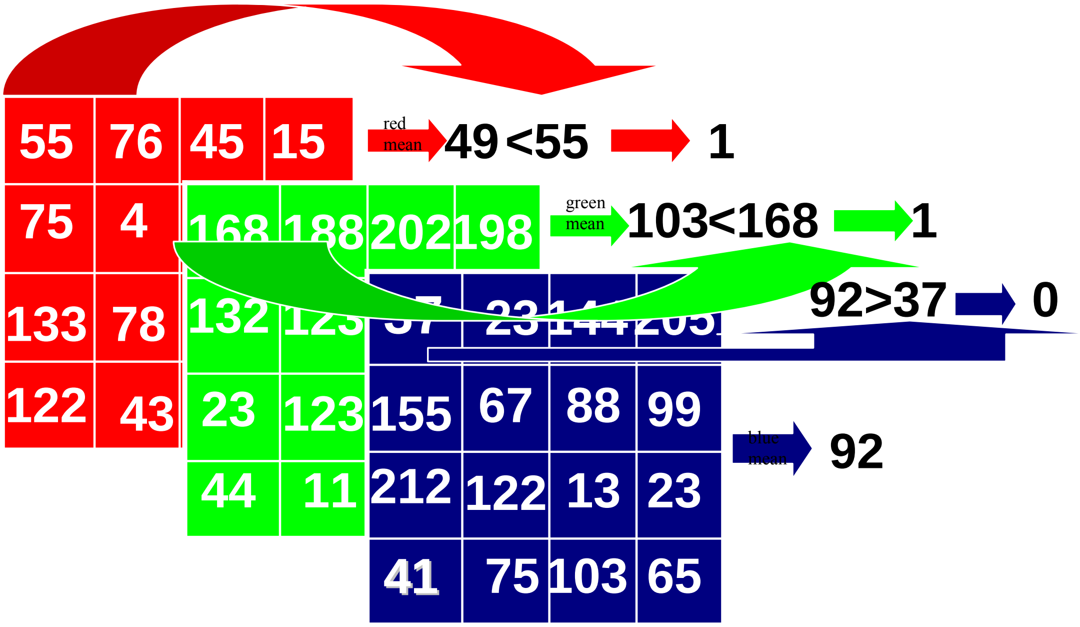


Fig 5.3: In this case, 0 is the result of the parity checking.

To embed the Mark Image, the pixel values of the Original Image may be modified by checking the even or odd parity of the corresponding bitmap triple. To realize a procedure of blind fragile image authentication, the embedding process must not modify the Original Image feature image. By choosing, inside the Original Image triple, the closest pixel to the mean value of its block, the least modification will be made, Figure 5.4.

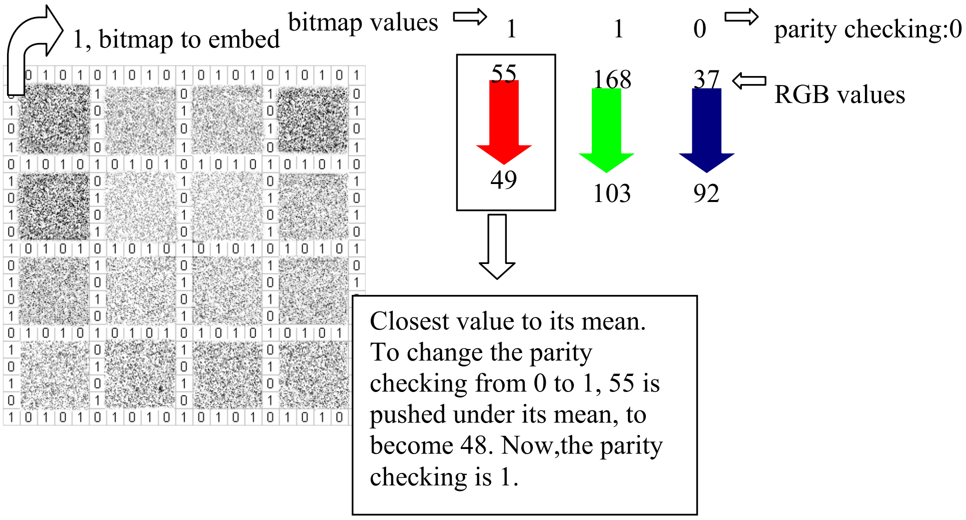
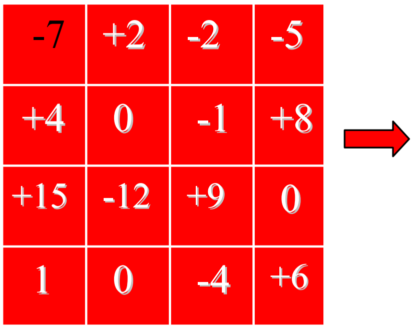

1, bitmap to embed
bitmap values
1
1
0
parity checking:0
55
168
37
RGB values
49
103
92
Closest value to its mean. To change the parity checking from 0 to 1, 55 is pushed under its mean, to become 48. Now,the parity checking is 1.

Figure 5.4: Way to modify the pixels.

A modification has been made within the red layer of the block. To maintain the mean value of this layer, the same as before the modification took place, we make a table 32×32 for each color layer, to memorize how many pixels varied within each layer during the embedding step. In this case, a -7 is jotted down on the first square of the table, Figure 5.5. Once the embedding process for the first block is completed, the sum of the modification values is calculated. The result takes into account, how much the mean value of the red layer has changed after the embedding process.



-7	+2	-2	-5
+4	0	-1	+8
+15	-12	+9	0
1	0	-4	+6

+14, sum of the red layer.

Figure 5.5: Table of the pixel modifications caused by the parity checking algorithm

The sum, to have the mean of the layer to remain the same, must be equal to 0. The value 14 must be brought to 0. To make it, an operation of subtracting is spread all over the block until the value 14 has been subtracted from the block. To increase or shrink, the system chooses those pixels whose variation, would not affect their parity.

This operation is carried out for each color layer, unless one of them has not been changed at all.

In this way, we obtain that the Marked Image ($\mathbf{I}_W$) has the same bitmap triples of the Original Image ($\mathbf{I}_O$). Since the image's features are related to the respective bitmap blocks, $\mathbf{I}_{\mathrm{W}}$ and $\mathbf{I}_{\mathrm{O}}$ have the same feature images, Figure 5.6.

Image authentication is obtained by extracting both feature images from $\mathbf{I}_W$. The first one, by the one-way hash function applied upon $\mathbf{I}_W$ and the second one, by recovering the inserted encrypted one from $\mathbf{I}_W$. The codified feature image can be decrypted by means of the public key of an RSA algorithm. These two bitmap matrices are then compared (by normalized correlation) against each other (by local normalized correlation) to determine if some image blocks are tampered, Fig 5.7.

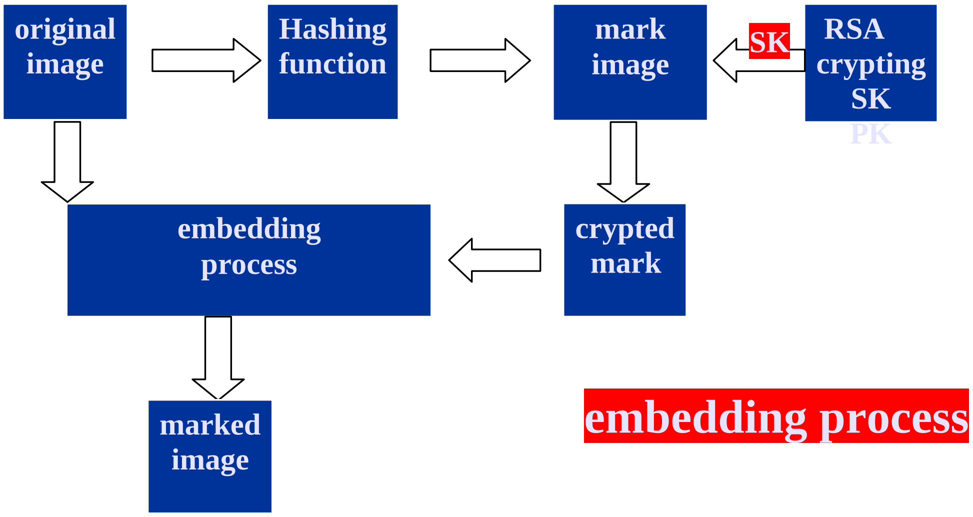


Figure 5.6: It delineates the embedding process.

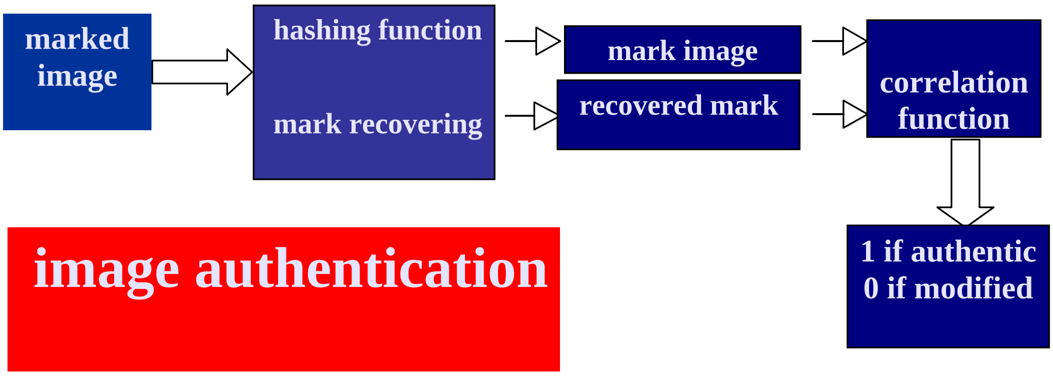


Figure 5.7: the recovering and authentication scheme

5.2 Authentication scheme

To recover the Mark Image from the Watermarked Image, the opposite steps will be taken. Each pixel of the Marked Image is considered as a triple of values. With the same method, used in the insertion section, the bitmap triple is calculated, comparing each pixel, for each color layer, with the mean of its block and then the sum S of the bitmap triple is calculated. The bit of the logo, is obtained by:

$$Bit\ of\ the\ mark = \begin{cases} 0 \ \text{ if } \ (S)\,\text{MOD}\,2 = 0 \\ 1 \ \text{ if } \ (S)\,\text{MOD}\,2 = 1 \end{cases}, \tag{9}$$

Obviously, the recovered Mark Image is coded. To obtain the decoded feature image we use the public key of the RSA algorithm.

PERFORMANCE ANALYSIS

ROBUSTNESS TO VALUMETRIC DISTORTIONS

There are five major types of valumetric [28] distortions on watermark detection: additive noise, amplitude changes, linear filtering, lossy compression and cropping. In our experiments, five hundred different images, of size 1024×1024 pixels, were watermarked. Each $\mathbf{I}_W$ was then distorted with the mentioned attacks.

5.3 Additive noise

Some processes that might be applied to a $\mathbf{I}_W$ (Watermarked Image) have the effect of adding a random signal. That is,

$$\mathbf{I}_{W_{noise}} = \mathbf{I}_W + \mathbf{n}, \tag{10}$$

where **n** is a random vector chosen from some distribution, independently of $\mathbf{I}_W$.

Each $\mathbf{I}_W$ was distorted with additive white Gaussian noise of different powers and the detection values were measured by the normalized correlation detector. Figure 5.8 shows the results of the experiments.

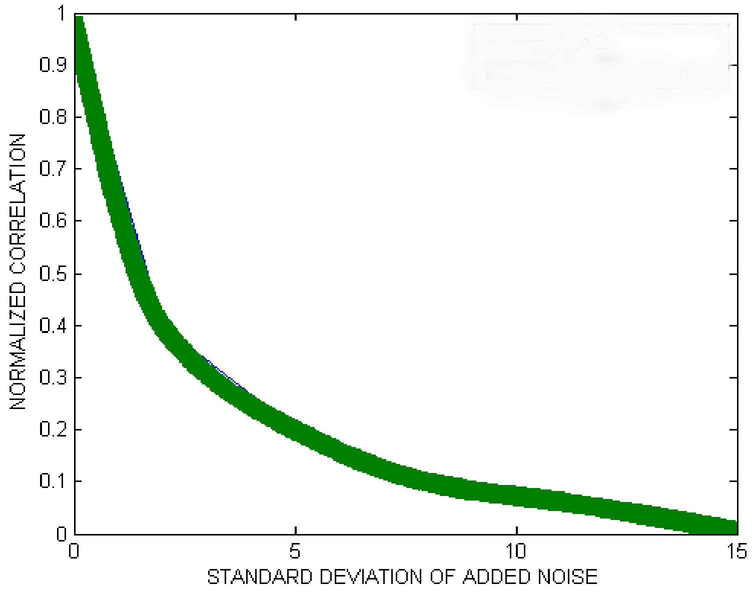


Figure 5.8. Normalized correlation coefficient versus standard deviation of added noise.

5.4 Amplitude Changes

In reality, many processes applied to the Watermarked Image ($\mathbf{I}_W$) are not well modeled by additive noise. The change in $\mathbf{I}_W$ is usually correlated with the watermarked image itself. Many processes are deterministic functions of the $\mathbf{I}_W$. A simple example is that of changes in amplitude. That is,

$$\mathbf{I}'_W = \alpha \mathbf{I}_W , \qquad (11)$$

where α is a scaling factor. In digital images, it represents a change in brightness and contrast. Normalized correlation is specifically designed to be independent of amplitude. Figure 5.9 shows the detection results.

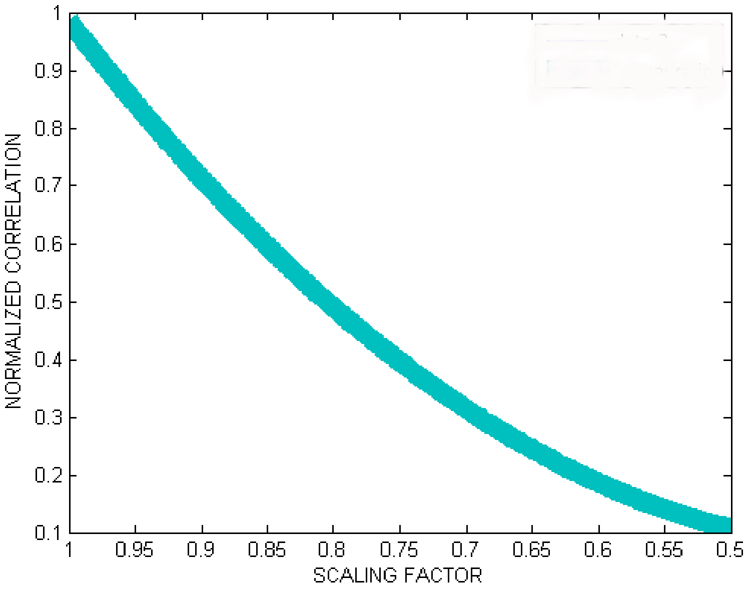


Figure 5.9: The normalized correlation remains high with scale changes.

These results show that normalized correlation remains high with scale changes, except that at very low scaling factors.

5.5 Linear Filtering

Another common type of signal processing that changes $\mathbf{I}_W$ in a deterministic fashion is linear filtering. That is,

$$\mathbf{I}'_W = \mathbf{I}_W * \mathbf{f}\,, \tag{12}$$

f is a filter, and $*$ denotes convolution. The watermarked images were distorted in each color plane, with Gaussian low-pass filters of varying standard deviation width. The filters were computed as

$$f[x,y] = \frac{f_0[x,y]}{\sum_{x,y} f_0[x,y]}, \tag{13}$$

where

$$f_0[x,y] = \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \tag{14}$$

as the width, σ, of the filter increased, more distortion was introduced. Figure 5.10 indicates that the pattern inserted is extremely sensitive to low-pass filtering. The detection values begin falling off sharply when the filter width exceeds 0.25.

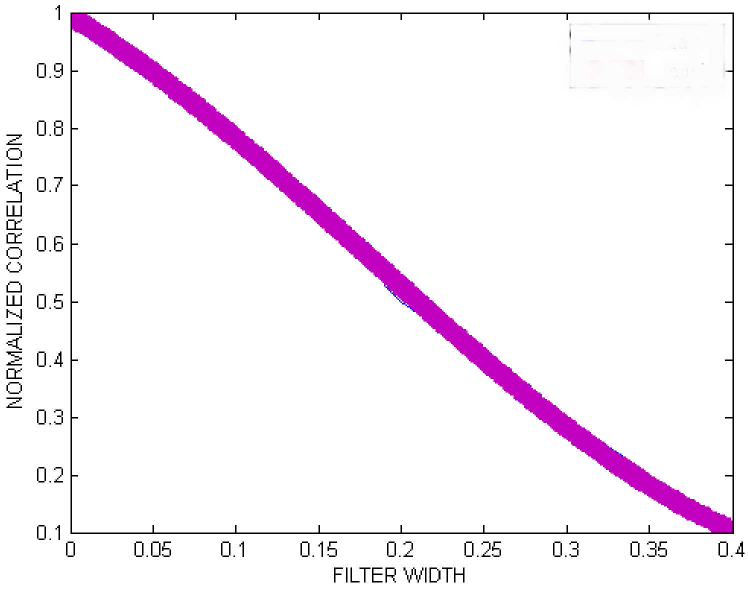


Figure 5.10: Normalized correlation coefficient versus linear filtering.

5.6 Lossy Compression

This simulation is effected using the JPEG compression obtained by applying quantization in the block-DCT domain. Figure 5.11 shows the correlation's values versus the percentage of quantization used.

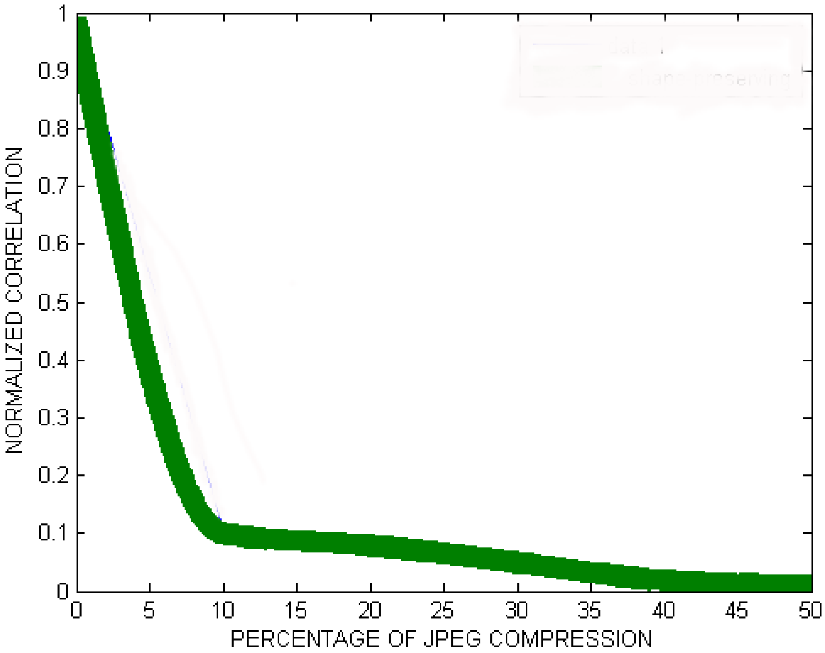


Figure 5.11: Normalized correlation coefficient versus different JPEG compression rates.

5.7 Cropping

To be resistant to cropping attacks, the mark was inserted into the Original Image as shown in fig.5.6. To recover the Mark, before using the procedure shown in section 5.2, the system seeks for the blocks still intact inside the Watermarked Image, they might be all intacts or just fiew of them, it depends on the entity of the cropping. To do that, it starts to make a XOR operation of all lines and columns until it finds the first crossing between one row and one column, both made with all 0s and 1s, Fig. 5.12. Then it recovers the Mark as shown in section 5.2.

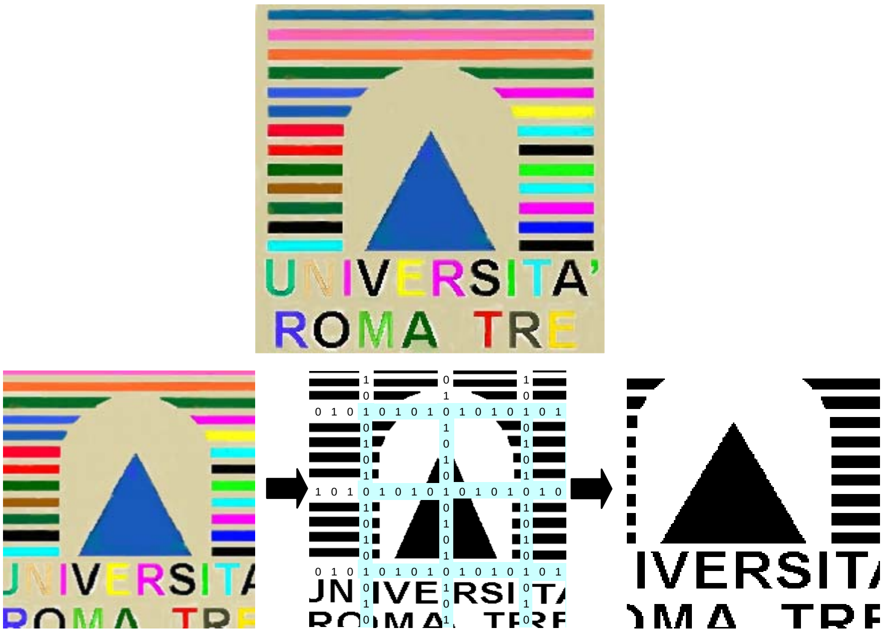
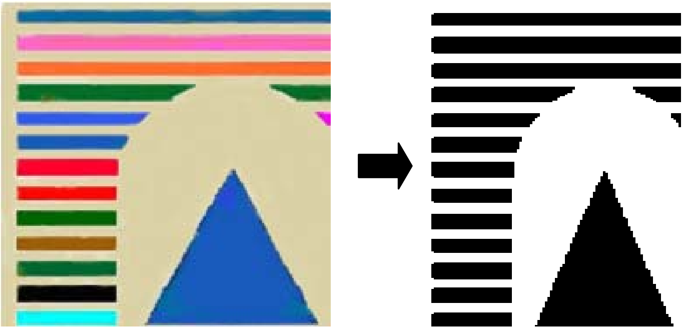


Figure 5.12: Watwermarked Image on top, Cropped Watermarked image at the bottom on the left, mark extracted and decoded from the previous Image, on the middle. The red lines are the first row and column crossing each other, found by the recovering system. From the crossing point on, the mark is decoded and well recovered even after a cropping attack, figure at the bottom on the right.

Here, this technique shows its robustness to cropping attacks.

The system is completely resistant to cropping attacks, figures 5.13 and 5.14. It is able to authenticate the whole watermarked image and even very small pieces of it.



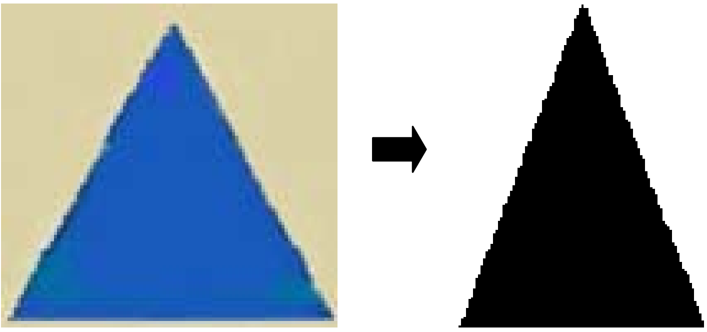


Figure 5.13: It's possible to recover the part of the mark present into the watermarked cropped image, with the same quality as if recovered from the whole watermarked image.

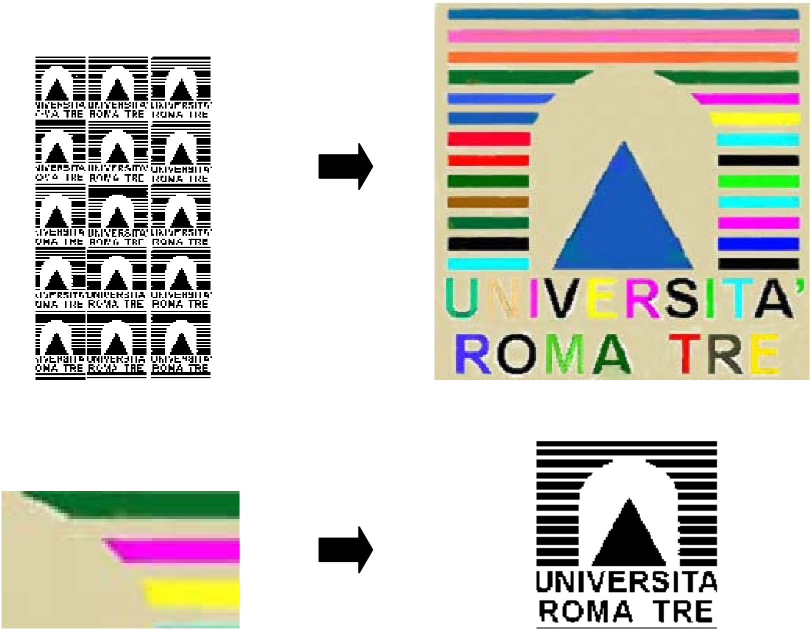


Figure 5.14: The embedded mark can also be made of several tiny "Università ROMA TRE" marks about the same size of a block so that even after a cropping attack, an entire "Università ROMA TRE" mark can be extracted.

This is the first or one of the first systems, as far as the authors know, to show such a capability.

5.8 Localization

If the image has not been modified, the recovered mark exactly matches the feature drawn from the Watermarked Image. However, if a region has been modified, the alterations show up as noise on the recovered binary pattern as on the feature hashed from the watermarked image. In this last case, the normalized correlation between the two images is high on all RSA-coded blocks but the modified ones. On them, the normalized correlation is close to zero since the recovered noise areas are completely uncorrelated.

To illustrate the performance of this algorithm, an experiment over a color image has been performed. Figure 5.15 shows a tampered with Watermarked Image, next to, the Mark recovered from it. Most of this pattern matches the original watermark, but the area containing the white beak is corrupted. Figure 5.15, on the bottom, shows a cropped and tampered with Watermarked Image, next to, the Mark recovered from it. Even after a cropping attack, the system is able to localize the corrupted area.

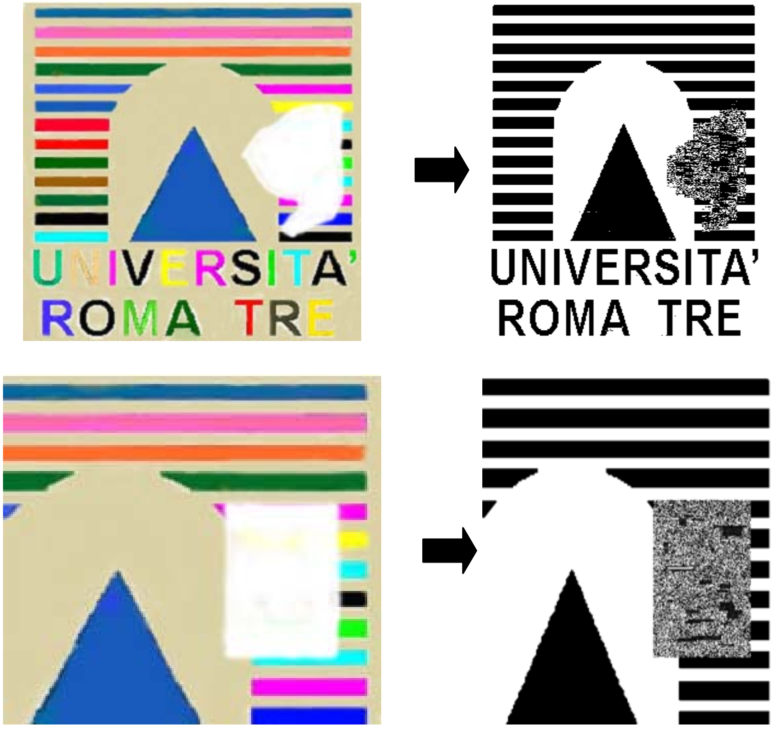


Figure 5.15: On top, tampered with Watermarked Image, next to, the Mark recovered from it- On the bottom, cropped and tampered with Watermarked Image, next to, the Mark recovered from it.

ATTACKS

In this section the use of the RSA-cryptosystem is justified. Let's suppose an adversary wishes to embed a valid watermark into either a modified or counterfeit Watermarked Image. Two basic categories of attack will be examined, search attacks and copy attacks.

5.9 Search attacks

Let's consider the real situation in which everyone has access to the Watermark detector procedure. The Watermarked Image is distributed and the public key is known. This system is an authentication system that must guarantee that each Watermarked Image is delivered without corruption. To embed a forged watermark into a Digital Image ($\mathbf{I}_O$), an adversary can enter slightly modified versions of the Watermarked Image into the detector until one is found that the detector reports as authentic. In this case, it is practically impossible to generate modified versions of the Watermarked Image wich the detector considers as authentic. That is because the system was built with an asymmetric cryptography which is really sensible to any kind of modification: if just one pixel is wrong the detector is not able to correctly recover the mark. The adversary is unable to perfectly make the alteration since he does not know the private key. That makes the system resistant to search attacks.

5.10 Copy Attacks

One of the most difficult attacks to prevent, is the copy attack. The copy attack is an attack in which the adversary copies a legitimate watermark from one Watermarked Image to another. It cannot be prevented from happening but if it happens the system should detect it as a form of unauthorized embedding. In this case what is going to be copied into a different Digital Image, is the coded mark that cannot be changed without knowing the private key. If there is a link between the Digital Image and the Mark, there is no possibility to take the mark from a Watermarked Image to put it into another Digital Image, without leaving trace of it. During the implementation of this watermarking technique, many attempts have been carried out to find out some good relations between pixel's values of the original image and the mark. No one of them turned out to show congruent results. The only way to make a real bond is to use as a mark, the feature extracted from the Original Image with a one-way hash function. This is the way this system was implemented. Several copy attacks were simulated and, in the recovering step, no one of them gave any good correlation value. That makes the system resistant to copy attacks.

5.11 Evaluating Perceptual Impact of Watermarks

How to answer to the question: “how much noisy is the watermark over the Original Image I_O?”. A “perceptual distance” measure is now introduced, the normalized mean squared error (NMSE) function, between the original image I_O and the watermarked image I_W:

$$D_{NMSE}\left(I_W, I_O\right) = \frac{\sum_{x=1}^{W}\sum_{y=1}^{H}\sum_{c=1}^{3}\left[I_F\left(x,y,c\right) - I_H\left(x,y,c\right)\right]^2}{\sum_{x=1}^{W}\sum_{y=1}^{H}\sum_{c=1}^{3}\left[I_H\left(x,y,c\right)\right]^2}, \qquad (15)$$

where *x* and *y* refer to each single colour field and c takes into account the possibility of having three different colour fields, as for an RGB image, red, green and blue.

In this case, the D_{NMSE} gets a value of about 5×10^{-4}, a very low value if two different images can usually get something about 5×10^{-1} [145]. That means that I_W and I_O look very much alike, Fig 5.16.

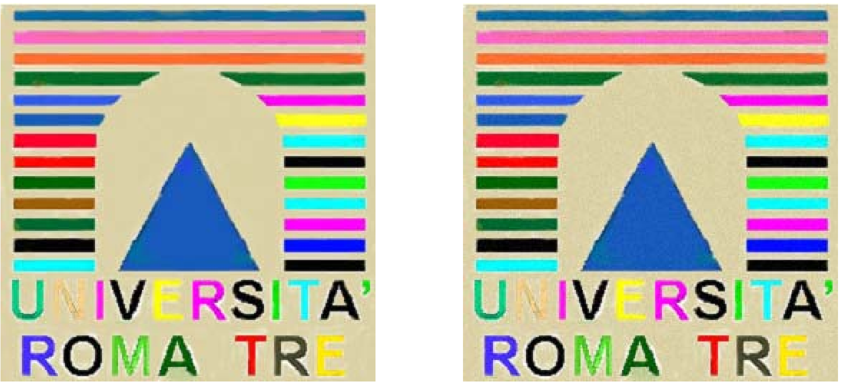


Figure 5.16. Original image and watermarked image.

One more evaluation function is the PSNR (Power Signal Noise Ratio), defined as:

$$D_{PSNR} = -10\times\log_{10}(D_{NMSE}). \qquad (16)$$

In Table 5.1, the mean values are shown for one hundred different images watermarked with five hundred different marks:

	mean values
D_{NMSE}	5×10^{-4}
D_{PSNR}	31dB

Table 5.1

CHAPTER 6

CONCLUSIONS

In Section 4.14, it is shown that the watermark embedding is invisible and has a good PSNR if the embedding factor is around the value of 0.1.So the quality of the image will not be affected by the watermarking embedding. The watermark embedding is very sensible to any distortion, but the error rates for the normal image processing are not above 0.45, except for fabricated image or modified image, since the watermark generation is a flexible scheme that generates a completely different watermark only when the image content is modified.

Section 5.11 shows, for the parity-checking technique, values of PSNR and NMSE close to the ones of the holographic method. As one can see the original and watermarked images look very much the same, with no visible differences.

In a certain way, this is a really good result for both techniques, especially for the holographic one, wich involves the use of the Fourier-transform domain to embedd the mark and quantization for the coding technique, to realize the hologram of the mark. Comparing to the second method, wich is far more linear and it does not involve any kind of transform domain and quantization, the gotten results look pretty much the same, which is a great score more for the holographic than for the spatial technique.

The holographic technique wins also the comparison to valumetric distortions. The resistance to scaling factors, gaussian noise and linear filtering shown by the first technique is in accordance with most of semi-fragile watermarking techniques encountered in the referenced papers, only the resistance to JPEG compression is a little bit less reliable, which classifies the system more fragile than semi-fragile. Linear filtering width of 0.15-0.2 are usually the values introduced by external distortions, as the standard deviation of added noise around 10 and the scaling factors of 0.7, at which values the correlation is about 0.9-0.8. Worse results are achieved in the case of the parity-checking system, which clearly shows is spatial nature. It has less resistance to the valumetric attacks than the first technique, and if the values are not so bad in the case of linear filtering, added noise and standard deviation, they are very bad indeed in the case of JPEG compression rates. The method fits itself very well in the fragile watermarking category.

Nevertheless, the spatial technique finds itself at its own ease when coping with the RSA-cryptosystem. It suits very well in the crypt and decrypt phases and it does not show any kind of problem. The cryptography acts directly on the pixel values of the image sparing the system to the necessity of having “side information” during the recovering step. One cannot say the same thing, in the case of the holographic technique, which acts in the Fourier-transformed domain and then for its own nature a little bit more complicated than the spatial one. As said before, the speckle noise affects the system in the recovering step, that is the reason why the RSA-cryptosystem is not applied directly to the image but to the arrays of rotation, which result to be the “side information” in the recovering step. In any case, good results are achieved for the recovered marks in both cases, correlation of about 0.97-0.96 are the common values for the recovered marks, in the case of the holographic system. Values equal to 1 or really close to it (0.99), are encountered in the case of the spatial technique, which guarantees the best results for the recovering step, at one hundred per cent.

The full resistance to cropping attacks is one of the major if not the major novelty of the two techniques. Even though both resistant to croping attacks, the two systems show two different behaviours to catch the target. The first algorithm is resistant to cropping attacks in the sense that even after the watermarked image was cropped (with high percentage of cropping), the system is still able to recover the entire mark image with high values of correlation with the original one. Even when the 40% of the watermarked image has been removed, the correlation of the recovered mark is still up to 0.7. The second algorithm is always able to recover the mark with correlation equal to 1 but recovers the part of the mark contained in the watermarked image wich survives the cropping attack. Example, if 40% of the watermarked image has been removed, that means that the rest 60% survives the cropping, so the algorithm is able to recover (with the same quality of the original mark) the mark contained in that 60% of the wtermarked image.

Of course, for the two algorithms the resistance to cropping attacks holds a different meaning, still the two techniques show an optimal behaviour in this matter.

As outlined above and through the sections of the thesis, the two techniques are quite different though aiming at the same target. The difference relies on their own nature, on the way they have been realized. One is really sophisticated and makes help of the transform domain to be made, in fact its nature is close to the semi-fragile watermarking systems which are in the most part of cases, realized in transformed domains. Nevertheless, its not fully resistance to compression rates casts the algorithm to be fragile compared to the common semi-fragile algorithms. Of great interest is the capability to recover the whole mark image from only a part of the waterarked image, this is due to the holographic nature of the system. The combined use of the RSA-cryptosystem allows the technique to be used for robust image authentication.

The nature of the second technique is the spatial one, no need to ask for transformed domain. That is the classical spatial algorithm which enriches and improves itself, when finds the way to be fully resistant to cropping attacks. The use of the RSA-cryptosystem directly to the image pixels, raises the system to be a wonderfull robust image authentication method.

References

[1] Y. Zhao, “Dual Domain Semi-fragile Watermarking for Image Authentication”, Degree of Master of Science, Department of Electrical and Computer Engineering, University of Toronto 2003.

[2] C.T. Li and H. Si, “Wavelet-based Fragile Watermarking Scheme for Image Authentication”, Department of Computer Science University of Warwick Coventry, CV 47AL,UK August 2006.

[3] Y. Liu, W. Gao, H.Yao and S. Liu, “A Texture-based Tamper Detection Scheme by Fragile Watermarking”, Computer Science and Technology Department of Harbin Institute of Technology, 2004 IEEE.

[4] M.E. Yalçin and J. Vanderwalle, ”Fragile Watermarking and Unkeyed Hash Function Implementation for Image Authentication on CNN-UM”, Katholicke Universiteit Leuven,Department of Electrical Engineering (ESAT), April 2002.

[5] S. Walton, “Information Authentication for Slippery New Age”, Dr. Dobbs Journal, 20(4), 18--16,1995.

[6] Y. Lin, C. Xu and D.D. Feny, “Web Based Image Authentication Using Invisible Fragile Watermark”, Hong Kong Polytechnic University and University of Sidney NSW 2006.

[7] M. Yeung and F. Mintzer, “An Invisible Watermarking Technique for Image Verification”, Proc. ICIP 1997, Santa Barbara, California.

[8] Y. Li, H. Guo and S. Jajodia, “Camper-Detection and Localization for Categorical Data Using Fragile Watermarks”, University of Singapore and University of Fairfax VA,2004.

[9] I.J. Cox, M. Miller and J. Bloom, “Watermarking Applications and Properties”, in Proc. International Conference on Information Technology: Coding and Computing, 2000.

[10] R.S. Alomari and A. Al-Jaber, “A Fragile Watermarking Algorithm for Content

Authentication", University of Jordan March 2005.

[11] Y. Zaho, "Dual Domain Semi-Fragile Watermarking for Image Authentication", Master Thesis, University of Toronto, 2003.

[12] P.S.L.M. Barreto, H.Y. Kim, V. Rijmen, "Toward a Secure Public-key Blockwise Fragile Authentication Watermarking", University of São Paulo, Brazil and B-3000 Leuven, Belgium.

[13] M. Fridrich and A. Baldoza, "New Fragile Authentication Watermark for Image", ICIP 2000, Vancouver, Canada (2000).

[14] R. Swierczynski, "Fragile Watermarking Using Subband Coding", Institute of Electronics and Telecommunication Poznan University, Sept. 2002

[15] P.W. Wong, "A Watermark for Image Integrity and Ownership Verification", Proceedings of IS&T PIC Conference, (Portland, OR), May 1998 (also available as Hewlett-Packard Labs . Tech. Rep HPL-97-72, May 1997).

[16] P.W. Wong, " A PublicKey Watermark for Image Verification and Authentication" Proceedings of IEEE International Conference on Image Processing,1998, Vol.1, pp. 455-459, (MA 11.07).

[18] M. Wu and B. Liu, "Watermarking for Image Authentication", Proceedings of IEEE International Conference on Image Processing, 1998, vol.2, pp. 437-441, (TA10.11)

[19] C.T. Li, D.C. Lou and T.H. Chen, "Image Authentication and Integrity Verification via Content-Based Watermarks and a Public Key Cryptosystem". Proceedings of IEEE International Conference on Image Processing, 2000, vol.3, pp. 694-697, (WP06.10).

[20] P.S.L.M. Barreto and H.Y. Kim, "Pitfalls in Public Key Watermarking", Proceddings of Sibgrapi-Brazilian Symposium on Computer Graphics and Image Processing, 1992, pp. 241-242.

[21] P.S.L.M. Barreto, H.Y. Kim and V. Rijmen, "Um Modo de Operação de Funções de Hashing

para Localizar Alterações em Dados Digitalmente Assinados", Proceedings of Simpósio Brasileiro de Telecomunicações, 2000, paper #5150124.

[22] A.J. Menez, P.C. Van Oorschot and S.A. Vanstone, "Handbook of Applied Cryptography", (CRC Press, 1997).

[23] M. Scheneider and S-F. Chang, "A Robust Content based Digital Signature for Image Authentication", in Proceeding IEEE International Conference on Image Processing,1996, Lausanne, Switzerland.

[24] E.Liu and E. Delp, "A Review of Fragile Image Watermarks", Proc. Of the Multimedia and Security Workshop (ACM Multimedia 1999), pp. 25-29, 1999.

[25] M. Yeung and F. Mintzer, "Invisible Watermarking for Image Verification", Journal of Electronic Imaging, vol.7, no.3, pp. 578-591, 1998.

[26] D. Kundur and D. Hatzinakos, "Towards a Telltale Watermarking Technique for Tampering -Proofing", Proceedings of the IEEE International Conference on Image Processing,vol.2, pp.409-413, 1998.

[27] L. Xie and G. Arce, "Joint Wavelet Comparison and Authentication Watermarking",Proceedings of the IEEE International Conference on Image Processing,vol.2, pp. 427-431,1998.

[28] J. Cox, L. Miller and A. Bloom, "Digital Watermarking", Morgan Kaufmann Publishers,USA,2002.

[29] G. Schirripa Spagolo and M. De Santis, "Computer Generated Hologram for SemiFragile Watermarking with Encrypted Images", International Journal of Signal Processing 4(2), 133 -141 (2007) ISSN 1304-4478.

[30] M. De Santis and G. Schirripa Spagnolo, "Cropping Resistant Watermarking through Asymmetric Cryptography", Submitted to Hindawi Publishing Corporation.

[31] F. Mintzer, G. Braudaway and M. Yeung, “Effective and Ineffective Digital Watermarking”, Proceedings of the IEEE International Conference on Image Processing, pp. 9-12, Santa Barbara, California, October 1997.

[32] F. Mintzer, G. Braudaway and A. Bell, “Opportunities fo Watermarking Standards”, Communications of the ACM, vol. 41, no. 7, pp. 57-64, July 1998.

[33] G. Friedman, “The Trustworthy Digital Camera: Restoring Credibility to the Photographic Image”, IEEE Transactions on Consumer Electronics, vol.39, pp. 905-910, November 1993.

[34] D. Stinton, “Cryptography Theory and Practice”, CRC Press, Boca Raton, 1995.

[35] R. Wolfgang and E. Delp, “Fragile Watermarking using the VW2D Watermark”, Proceedings of the IS&T/SPIE Conference on Security and Watermarking of Multimedia Constants, pp. 204-213, San Jose, California, January 1999.

[36] N. Memon, S. Shende and P. Wong, “On the Security of the Yueng-Mintzer Authentication Watermarks”, Final Program and Proceedings of the IS&T PICS 99, pp. 301-306, Savanna, Georgia, April 1999.

[37] G.R.A.L. Xie and R.F. Graveman, “Approximate Image Authentication Codes”, IEEE Transactions on Multimedia, 2000.

[38] D.C. Lau and J.L. Liu, “Fault Resilient and Compression Tolerant Digital Signature for Image Authentication”. IEEE, 2000.

[39] C. Lin and S.F. Chang, “A Robust Image Authentication Method Distinguishing JPEG Compression from Malicious Manipulations”, vol. 11 IEEE Transactions on Circuits and Systems of Video Technology, Feb 2001.

[40] C.S. Lu and H.Y.M. Liao, “Structural Digital Signature for Image Authentication: an Incidental Distraction Resistant Scheme”. Los Angeles, California, USA: Proc. Multimedia and Security Workshop at the ACM. Int. Conf. On Multimedia, 2000.

[41] M.K.S. Bhattarcharjee, “Compression Tolerant Image Authentication”, vol.2, pp. 435-439 . Chicago, IL: Proc. IEEE Int. Conf. On Image Processing, Oct. 1998.

[42] P.L.M.P. Quelez, “Spatial Watermrk for Image Verification”, vol. 3971, pp. 120-127, Proc. Of SPIE Security and Watermarking of Multimedia Contents II, Jan 2000.

[43] S.F.C.C.Y. Lin, “Semi-Fragile Watermarking for Authentication JPEG Visual Content”, vol. 9971, pp. 140-151. Proc. Of SPIE Int. Conf. On Security and Watermarking of Multimedia Contents II, Jan. 2000.

[44] J. Fridrich, “Robust Digital Watermarking based on Key-Dependent Basis Function”. Portland, OR: The 2nd Information Hiding Workshop, April 1998.

[45] J. Fridrich, “Combining Low Frequency and Spread Spectrum Watermarking” San Diego: Proc. SPIE Symposium on Optical Science, Engineering and Instrumentation Proceedings, July 1998.

[46] M.G.J. Fridrich, “Protection of Digital Image Using Self Embedding”. New Jersey Institute of Technology, May. 1999.

[47] J. Fridrich, “Image Watermarking for Tamper Detection”, vol.2. Proc. ICIP-98, 1998.

[48] R.D.J. Fridrich, M. Goljan, “Invertible Authentication Watermark for JPEG Images”. IEEE, 2001.

[49] C.E.T. Lin and E.J. Delp, “Detection of Image Alterations Using Semi-Fragile Watermarks”, vol. 3971. Proc. Of SPIE Int. Conf. On Security and Watermarking of Multimedia Contents II, Jan 2000.

[50] B.J.J. Eggers, J.K. Su, “A Blind Watermarks Scheme Based on Structured Code-books”. London: IEEE Conf. Secure Images and Image Authentication, Apr. 2000.

[51] J.C.B.L.M. Marvel, G.W. Hartwig, “Compression Compatible Fragile and Semi-fragile Tamper Detection”, vol. 3971. Proceedings of the SPIE Int. Conf. On Security and Watermarking of Multimedia Content II, Jan. 2000.

[52] D.H.D. Kundur,"Digital Watermarking for Telltale Tamper-proofing and Authentication", vol. 87(7), pp. 1167-1180. Proceedings of the IEEE Special Issue on Identification and Protection of Multimedia Information, Jul. 1999.

[53] C.S.C.S. Lu, H. Mark Liao,"Combined Watermarking for Image Authentication and Protection". IEEE, 2000.

[54] R.B. Wolfgang, C.L. Podilchuck and E.J. Delp, "Perceptual Watermarks for Digital Images and Video", Proceedings of the IEEE, vol. 87, no. 87, no. 7, pp. 1108-1126, July 1999.

[55] S. Parenti and M. Yeung, "Verification Watermarks on Fingerprints Recognition and Retieved", Proceedings of the IS&T/SPIE Conference on Security and Watermarking of Multimedia Contents, pp. 66-78, San Jose, California, January, 1999.

[56] R. Walfgang and E. Delp, "A Watermark for Digital Images", Proceedings of the IEEE International Conference on Image Processing, vol. 3, pp. 219-222, 1996.

[57] Stirmark Software: http:// www.cl.camac.uk/~fapp2.watermarking/stirmark,1997.

[58] D. Stinson, "Cryptography Theory and Practice", CRC Press. Boca Raton, 1995.

[59] B. Freneel, "Analysis and Design of Cryptographic Hah Functions", Ph. D. Thesis, February 2003.

[60] F. Cohen, "Computer Viruses-theory and Experiments", Computer & Security, Vol.6, 1987, pp. 22-35.

[61] J.F. Shoch and J.A. Hupp,"The "worm" Programs-early Experience with a Distributed Computation", Communications ACM, Vol. 25, no. 3, March 1982, pp. 172-180.

[62] R. Garon and R. Outerbridge,"DES watch: an Examination of the Sufficiency of the Data Encryption Standard for Financial Institution Information Security in the 1990's", Cryptologie, Vol.XV, no.3, pp. 177-193.

[63] Y. Desmedt, J. Vanderwalle and R. Govaerts,”The Mathematical Relation between the Economic Cryptographia and Information Theoretical Aspects of Authentication”, Proc. Fourth Symposium on Information Theory in the Benelux, Haasrode, Belgium, 26-27 May 1983, pp. 63-65.

[64] W. Diffie and M.E. Hellman,”New Directions in Cryptography”, IEEE Trans. On Information Theory, Vol. IT-22, No. 6, 1976, pp. 644-654.

[65] C.E. Shannon, “Communication Theory of Secrecy Systems”, Bell System Technical Journal, Vol. 28, 1949, pp. 656-715.

[66] “Information Processing – Open System Interconnection – Basic Reference Model – Part 2: Security Architecture”, IS 7498/2, ISO/IEC, 1987.

[67] D. Kahn, “The Codebreakers. The Story of Secret Writing” Mactillar, New York, 1967.

[68] G.S. Vernam, “Cipher Printing Telegraph System for Secret Wire and Radio Telegraph Communications”, Journal American Institute of Electrical Engineers, Vol. XLV, 1926, pp. 109-115.

[69] H. Feistel,”Cryptography and Computer Privacy”, Scientific American, Vol. 228, No. 5, May 1973, pp. 15-23.

[70] H. Feistel,W.A. Notz and J.L. Smith, “Some Cryptographic Techniques for Machine-to-machine Data Communications”, Prc. IEEE, Vol.63, No. 11, November 1975, pp. 1545-1554.

[71] Y. Desmedt, “Analysis of the Security and New Algorithms for Modern Industrial Cryptography”, Doctoral Dissertation, Katholieke Universiteit Leuven, 1994.

[72] R.R. Jueneman, S.M. Matyas and C.M. Meyer,”Message Authentication with Manipulation Detection Codes”, Proc. 1983 IEEE Symposium on Security and Privacy, 1984, pp.33-54.

[73] J.L. Massey, “An introduction to Contemporary Cryptology”, in “Contemporary

Cryptology:the Science of Information Integrity", G.J. Simmons, Ed., IEEE Press, 1981, pp.3-39.

[74] R.A. Rueppel, "Stream Ciphers", in "Contemporary Cryptology, he Science of Information Integrity", G.J. Simmons, Ed., IEEE Press, 1991, pp. 85-134.

[75] G.J. Simmons, "A Survey of Information Authentication", in "Contemporary Cryptology: the Science of Information Integrity", G.J. Simmons, Ed., IEEE Press, 1991, pp. 381-419.

[76] A.V. Aho, J.E. Hopcroft and J.D. Ullman,"The Design and Analysis of Computer Algorithms", Addison-Wesley, 1974.

[77] P.E. Dunne, "The Complexity of Boolean Networks", A.P.I.C. Studies in Data Processing No. 29, Academic Press, 1988.

[78] A.C. Yao, "Theory and Applications of Trapdoor Functions", Proc. 23rd IEEE Symposium on Foundations of Computer Science, 1982, pp. 80-91.

[79] A.C. Yao, "Computational Information Theory", in "Complexity in Information Theory" Y.S. Abu-Mostafa, Ed., Springer-Verlag, 1988, pp. 1-15.

[80] P. Impagliazzo and S. Rudich,"Limits on the Provable Consequences of One-way Permutations", Proc. 21st ACM Symposium on the Theory of Computing, 1990, pp. 44-61.

[81] R.L. Rivest, A. Shamir and L.Adleman, "A Method for Obtaining Digital Signature and Periodic-key Cryptosystems", Communications ACM, Vol. 21, February 1978, pp. 120-126.

[82] R. Merkle and M. Hellman, "Hiding Information and Signature in Trapdoor Knapsack", IEEE Trans. On Information Theory, Vol. IT-24, No. 5, 1978, pp. 525-530.

[83] E.F. Brickell and A.M. Odlyzko,"Cryptanalysis: a Survey of Recent Results", in "Contemporary Cryptology: the Science of Information Integrity", G.J. Simmons, Ed., IEEE Press, 1991, pp. 501-540.

[85] I.B. Damgard, “A Design Principle for Hash Functions”, Advances in Cryptology, Proc. Crypto ’89, LNCS G.Brassard, Ed., Springer-Verlay, 1990, pp. 416-427.

[86] R. Merkle, “One Way Hash Functions and DES”, Advances in Cryptology, Proc. Crypto ’89, LNCS 435, G. Brassard, Ed., Springer-Verlay, 1990, pp. 428-446.

[87] B. Preneel, A. Bosselaers, R. Govaerts and J. Vanderwalle,”A Chosentext Attack on the Modified Cryptographic Checksum Algorithm of Cohen and Huang”, Advances in Cryptology, Proc. Crypto ’89, LNCS 435, G. Brassard,Ed., Springer-Verlay, 1990, pp.154-163.

[88] “Information Processing – Open Systems Interconnection – Basic Reference Model _ Part 2: Security Architecture”, IS 7498/2, ISO/IEC, 1987.

[89] M. Desoete, K. Vedder and M. Walker, “Cartesian Authentication Schemes”, Advances in Cryptology, Proc. Eurocrypt’89, LNCS 434, J. – J. Quisquarter and J. Vanderwalle, Eds., Springer-Verlay, 1990, pp.476-490.

[90] T. ElGamal, “A Public-key Cryptosystem and a Signature Scheme Based on Discrete Algorithms”, IEEE Trans. On Information Theory, Vol. IT-31, No.4, 1985, pp. 469-472.

[91] S.Goldwasser, S. Micali and R. Rivest, “A “Paradossical” Solution to the Signature Problem”, Proc. 25th IEEE Symposium on Foundations of Computer Science, October 1984, pp. 441-448.

[92] S. Goldwasser, S. Micali and R.L. Rivest,”A Digital Signature Scheme Secure Against Adaptive Chosen-message Attacks”, SIAM Journal on Computing, Vol. 17, No. 2, 1988, pp. 281-308.

[93] M. Naor and M. Yung, “Universal One-way Hash Functions and Their Cryptographic Applications”, Proc. 21st ACM Symposium on the Theory of Computing, 1990, pp. 387-394.

[94] J. Rompel,”One-way Functions Are Necessaryand Sufficient for Secure Signature”, Proc. 22nd ACM Symposium on the Theory of Computing, 1980, pp. 387-394.

[95] “Message Handling/Information Processing Systems”, C.C.I.T.T. Recommendation X.400, 1988

[96] “The Directory-Overview of Concepts”, C.C.I.T.T. Recommendation X.509, 1988, (same as IS 9594-8, 1989).

[98] W. Stallings, “Cryptography and Network Security”, second edition Prentice Hall, 1988.

[99] R.L. Rivest, A. Shamir, “The MD5 Message-digest Algorithm”. Internet Activities Board, Apr. 1992.

[100] NIST., Secure Host Standard. FIPS PUB 180-1, Apr. 1995.

[101] M.H. Krawczyk and R. Canetti,”HMAC: Keyed Hashing for Message Authentication”. Internet 1991. Request for Comments 2104, Feb. 1997.

[102] M.L.R.L. Rivest, A. Shamir,”A Method for Obtaining Digital Signatures and Public Key Cryptosystem”, pp. 120-126 comm. Of ACM, Feb. 1997.

[103] G.L. Friedman, “The Trustworthy Digital Camera: Restoring Credibility to the Photographic Image”, vol. 39, pp. 905-910, IEEE Trans. Consumer Electronics, 1993.

[104] L.X.R.F. Graveman and G.R. Arce, “Approximate Message Authentication Codes”, IEEE Transaction on Image Processing, 2000.

[105] J. Fridrich, “Robust Bit Extraction from Images”, Centro Affati, Florence, Italy: Proceedings of the 1999 IEEE International Conference on Multimedia Computing and Systems.2, June , 1999.

[106] M.H.J.R. Venkatesan, S.M. Koon and P. Moulin, “Robust Image Hashing”, vol.3, pp. 664-666. Proceedings of the IEEE International Conference on Image Processing,Sept. 2000.

[107] C. Lin and S.F. Chang, “Generating ROBUST Digital Signature for Image/Video Authentication”, Bristol, UK: Multimedia and Security Workshop and ACM Multimedia ’98,

Sep. 1998.

[108] C. Lin and S.F. Chang,”A Robust Image Authentication Method Surviving JPEG Lossy Compression”. San Jose: SPIE Storage and Retrieval of Image /Video Database”, EI’98, Jan. 1998.

[109] C. Lin and S.F. Chang, “An Image Authenticator Surviving DCT-based Variable Quantization Table Compressions”. CU/CTR Technical Report 490-98-24, Nov. 1997.

[110] C.M.M. Yeung,”An invisible Watermrking Technique for Image Verification”, pp. 680-683. Proc. Of ICIP 1997.

[111] B.L.M. Wu,”Watermarking for Image Authentication”, vol.2 Proc. ICIP’98, 1998.

[112] J.C.R.L.M. Marvel, C. Bancelet,”Spread Spectrum Image Steganography”, pp. 1075-1083, IEEE Trans. On Image Processing, Aug 1999.

[113] J.C.B.L.M. Marvel, G.W.Hartwig,”Compresion Compatible Fragile and Semi-fragile Tamper Detection”, vol. 3971. Proceedings of the SPIE Int. Conf. On Security and Watermarking of Multimedia Content II, Jan 2000.

[114] B.G.J.J. Eggers,”Blind Watermarking Applied to Image Authentication”, ICASSP’ 2001, May 2000.

[115] G.R.A.L. Xie,”A Class of Authenticatin Digital Signature for Image Authentication”, pp.21-24, Proc. EUSIPCO 98 SIGNAL Processing IX: Theories and applications, 1998.

[116] G.R.A.L. Xie “A Class of Authentication Digital Watermarks for Secure Multimedia Communication”, pp. 1057-7149 Proc. IEEE Int. Conf. On Image Processing, 2002.

[117] R.G.S.A.Z. Tirkel and C.F. Osborne,”A Digital Watermark”, p.86 Proc. ICIP’94.

[118] D.G.W.Bender and N. Morimoto, “Techniques for Data Hiding”, vol. 2420, p. 40. Proc. SPIE, 1995.

[119] K.T.K. Matsui, “Video Steganography: How to Secretly Embed a Signature in a Picture”, vol.1, pp. 187-206, Proc. IMA Intellectual Property Project, 1992.

[120] L.G.H. Berghel, “Protecting Ownership Rights through Digital Watermarking”, vol.29, pp. 101-103. IEEE Comput. , July 1996.

[121] R.J.A.F. Petitcolas and M.G. Kuhn,”Information Hiding. A Survey”, vol.87, No. 7, pp. 1028-1078. Proceedings of the IEEE, July 1999.

[122] S.J.N.F. Johnson, “Exploring Steganography: Seeing the Unseen”, pp. 26-34. IEEE Comput., Feb. 1998.

[123] F.P.- G.J.R. Hernaandez,”Statistical Analysis of Watermarking Schemes for Copyright Protection of Images”, vol. 87, No.7, pp. 1142-1166. Proceedings of the IEEE, Jul.1999.

[124] M.K.F. Harung,”Multimedia Watermarking Techniques”, vol. 87, no.7, pp. 1079-1107. Proceedings of the IEEE, July 1999.

[125] F.V.A. Piva, M. Barni, “DCT-based Waterark Recovering without Restoring to the Uncorrupted Original Image”, pp. 520-523. IEEE, 1997.

[126] S. Daly,”The Visible Difference Predictor: an Algorithm fot the Assessment of Image Fidelity”, pp. 179-206. Digital Images and Human Vision, MIT Press, 1993.

[127] J. Lubin, “The Use of Psycophysical Data and Models in the Analysis of Display System Perfomance”, pp. 163-178. Digital Images and Human Vision, MIT Press, 1993.

[128] G.C. Phillips and H. Wilson, “Orientation Bandwidths of Spatial Mechanism Measured Day Masing”, vol.1, no.2 Journal of Opt. Soc. Of America.

[129] M.M. Yeung, B.L. Yeo and M. Holliman, “Digital Watermarking Shedding Light on the Invisible”, IEEE Micro, pp. 32-41, Nov-Dec. 1998.

[130] N. Takai and Y. Mifune, “Digital Watermarking by a Holographic Technique”, Appl. Opt. Vol. 41, pp. 865-873, 2002.

[131] G. Schirripa Spagnolo, C. Simonetti and L. Cozzella,”Fragile Digital Wtermarking by Synthetic Holograms”, Proc. SPIE vol. 5615, pp. 173-182, 2004.

[132] G. Schirripa Spagnolo, C. Simonetti and L. Cozzella,”Content Fragile Watermarking based on Computer Generated Hologram Coding Technique”, J. Opt. A: Pure Appl. Opt., vol. 7, n.7, pp. 333-342, 2005.

[133] R. Van Schydel, A. Tirkel and C. Osborne, “A Digital Watermark”, Proc. Of the IEEE vol. 2, International Conference on Image Processing (ICIP’94), Austin(Tx) 13-16 November 1994, pp. 86-90.

[134] H.Y.K.M. and R.L. de Queiroz,”A Public-key Authentication Watermarking for Binary Images”, Univerdidade de São Paulo, Brasilia, Brazil, 2004 (ICIP).

[135] P. Hariaharan,”Optical Holography: Principles, Techniques and Applications”, Cambridge University Press, Cambridge 1996.

[136] W.R: Lee, “Sampled Fourier Transform Hologram Generated by Computer”, Appl. Opt. Vol.9, n.3, pp. 639-643, 1970.

[137] W.R. Lee, “Computer Generated Holograms Techniques and Applications”, Progress in Optics Vol. 16 pp. 121-231, North Holland, 1974.

[138] Y. Aoki,”Watermarking Technique Using Computer-generated Holograms”, Electronics and Communications in Japan part 3, vol. 84, n.1, pp. 21-31, 2001.

[139] L. Croce Ferri, “Visualization of 3D Information with Digital Holography using Laser Printers”, Computers & Graphics, vol. 25, pp. 309-321, 2001.

[140] J. Dittmann, L. Croce Ferri, C. Vielhauer,”Hologram Watermarks for Document Authentications”, IEEE International Conference on Information Technology Coding and

Computing. IEEE Computer Society, Las Vegas, NV, USA, April 2-4, pp. 60-64, 2001.

[141] L. Croce Ferri, A. Mayerhőfer, M. Frank, C. Vielhauer, R. Steimetz,"Biometric Authentication for ID Cards with Hologram Watermarks", Proc. SPIE vol. 4675, pp. 240-251, 2002.

[142] NIST FIPS 197 – Advanced Encryption Standard. 26 Nov. 2001.

[143] C.K. Yang and C.S. Huang, "A Novel Watermarking Technique for Tampering Detection Digital Images", Electronics Letters on Computer Vision and Image Analysis 3(1), 1- 12 (2004).

[144] NIST., "Secure Hash Standard", FIPS PUB 180 – 1, Apr. 1995.

[145] S.J. Sangwine and R.E.N. Horne, "The Color Image Processing Handbook", Chapman & Hall, London, UK.