

Contents

1	Biometric Security Systems	3
1.1	“What you have” vs. “What you are”	3
1.2	Overview of the most common biometrics	3
1.2.1	Fingerprints	3
1.2.2	Hand geometry	4
1.2.3	Iris	4
1.2.4	Face recognition	4
1.2.5	Others	4
1.3	Issues and weaknesses of biometrics	5
1.3.1	Ethical issues and privacy concerns	5
1.3.2	Conceptual flaws	6
1.3.3	Technical weaknesses	6
2	Data authentication	10
2.1	Cryptographic digital signature	10
2.1.1	Introduction	10
2.1.2	Application for data origin authentication	11
2.2	Digital watermarking	13
2.2.1	Definition and purposes	13
2.2.2	Why using digital watermarking ?	14
3	Digital watermarking techniques review	16
3.1	Common requirements and techniques	16
3.1.1	General requirements	16
3.1.2	Commonly used techniques	17
3.2	Different domains	18
3.2.1	Spatial domain	19
3.2.2	DCT domain	20
3.2.3	DWT domain	21
3.3	Watermarking evaluation	22

4 Fingerprint images authentication based on digital watermarking	25
4.1 Technical requirements	25
4.1.1 Blind watermark	25
4.1.2 Least consequences on the recognition	26
4.1.3 Key-based mechanism	26
4.1.4 Fully automated verification	27
4.1.5 Detection of <i>replay-attacks</i>	27
4.1.6 Robust vs. fragile	28
4.1.7 Robustness vs.	28

Chapter 1

Biometric Security Systems

1.1 “What you have” vs. “What you are”

Traditional methods for automated personal identification or identity verification mainly rely on “what you have” (a token, such as a key or a card) and/or “what you know” (such as Personal Identification Numbers (PIN) or passwords). However this methods bring a certain level of security, they suffer of conceptual flaws since tokens can be lost or stolen while PIN and passwords can be forgotten, guessed or written down.

Based on the recognition of physical and behavioural characteristics of the individual (“what you are”), biometrics (e.g., fingerprints, iris, face, voice or hand geometry, to cite a few) are tied to their owner (the individual’s body) and can therefore not be lost or forgotten. Their convenience as well as the higher security that this new approach brings is certainly the main reason of its increasing use.

1.2 Overview of the most common biometrics

Although it is extremely difficult to accurately determine nowadays’ market shares of the different biometrics, the figures given below are taken from [Poh] and [FindBiometrics] and are only aimed at giving a rough idea, though most of the results are similar in both sources.

1.2.1 Fingerprints

The extremely high probability of fingerprints uniqueness [Pankanti, 2002] as well as a relatively good acceptance of users certainly explains why finger-

prints recognition is currently the most commonly used biometrics. Hitting 35% of the market, it is mainly applied for law enforcement, civil government, enterprise security, medical and financial transactions.

As far as research is concerned, a majority of the work carried out towards biometrics is concerned by fingerprint matching.

1.2.2 Hand geometry

Unlike fingerprints, the geometry of the hand is not unique [Ross, 2004] because its features are not descriptive enough to allow identification. Although, it remains highly used (27% of the market) because the measurement and analysis of the hands shape is extremely robust in terms of recognition [NCSC, 2002] [Fein, 2004], thus allowing an efficient verification to be applied for time an attendance systems and physical access.

1.2.3 Iris

Although figures showing iris recognition's market share are vague (varying between 9% and 16%), it is being increasingly applied for access control at airports and banking facilities because it achieves a far higher accuracy and reliability than with fingerprints [Daugman, 2003]. Exhaustive information towards iris recognition such as the so-called *Iris-Code*, algorithms for active/passive eye-finding and matching can be found at John Daugman's web page [Daugman, 2004].

1.2.4 Face recognition

Raising 11% of the market share, face recognition, at the exact opposite of fingerprints, is fast but not reliable [Hong, 2004]. This unreliability is explained byt conceptual and technical reasons. First, the effects of aging and facial expression are an issue [SUN, 1999] while external conditions such as the pose and illumination. An interesting (but also often criticised due to privacy and ethical issues) application of facial recognition is for detecting the presence of searched people (e.g, terrorists) in public areas.

1.2.5 Others

Retina scan also occupies a slightly important part of the market with 11%. Although its accuracy is no far below iris, it is difficult to achieve good performances unless the users are well-trained, patient and cooperative,

which is a problem since they tend to find it very intrusive [IBL, 2002]. Voice recognition, with 6% of the market suffers of different issues such as the environmental noise, the mood of the user and is also very demanding in term of training for enrollment. Online signature (5% of the market) is highly accepted but not so accurate [NGUGI, 2004], however, a recent research finding to be published [SPAGNOLO, 2004] and involving 3D Holograms to detect fake signatures seems to enable a chance to inverse this tendance by bringing a far higher accuracy.

1.3 Issues and weaknesses of biometrics

1.3.1 Ethical issues and privacy concerns

An always immediate consequence of governments proposing to apply biometrics into their security systems (border control, national ID, access to sensitive facilities, search for “wanted faces”, etc) is the protest of organisations concerned by privacy and ethics. Indeed, shifting token-based or knowledge-based “keys” by private features of the individual, especially within a digital environment, obviously involves dealing with the individual’s own data. Although it might be argued that getting someone’s fingerprint, iris or face is far from being an indecent intrusion into the private life, the actuals concern in fact go towards answering the following questions:

- What data are to be stored ?
- How are the data stored ? More specifically, is the storage sufficiently safe to effectively avoid any attempts of data theft ?
- How and for which exact purposes are the data going to be used ?

In a worst case scenario, a biometric security system could, on top of its initial purpose, allow data (patterns) theft to lead to identity theft. Also, it could well allow to monitor and trace individual’s activities.

Although, it can be understood that security is an extremely important issue and that it might therefore sometimes require to affect individuals privacy but, nevertheless, an often raised argument against this idea is that none of these would have indeed helped avoiding terrorists attacks such as 9/11.

1.3.2 Conceptual flaws

Shifting to biometrics-based security systems seems quite appealing because it not only adds more security but also brings its collection of advantages in terms of convenience and management costs (i.e, no more password or PIN to remember, regularly change or reset, no more tokens to carry or locks to replace). However, some conceptual flaws come along to balance the opinion on the real efficiency of biometrics.

Philippe Wolf, director of the DCSSI (a governmental agency in charge of dealing with IT Security questions in France and working in collaboration with the Ministry of Defence), has published in October 2003 an article [C'T, 2002] in *Infosecu* (monthly magazine of the French National Center For Scientific Research (CNRS)) where he explains why biometrics are not to be recommended. On top of further arguments to be given in the following section, he explains that, while we can easily reset PIN's or passwords and change keys or locks when they have been lost or compromised, this just does not apply with biometrics.

While the fact that fingerprints or iris patterns normally don't change over the years seems to be a great asset to defend the use of biometrics, it is also probably its main flaw because, once compromised it can not be replaced. To bring a more obvious description of this issue, he brings the example of the movie "*Minority Report*" where the hero needs to have a chirurgical intervention because his iris has been "black-listed".

1.3.3 Technical weaknesses

Further to conceptual flaws, current technical issues also exist and possible attacks can mainly be aimed at three locations [C'T, 2002] [Uludag, 2004] of a typical biometric security system (see 1.3.3 and 1.3.3):

- (1) The sensor that acquires the biometric features
- (3) The system that stores and matches templates
- (2) The medium through the communication of the system happens

1.3.3.1 The sensor

An impressive list of papers have been published towards defeating biometric system by fooling the sensors with artificially created material such as

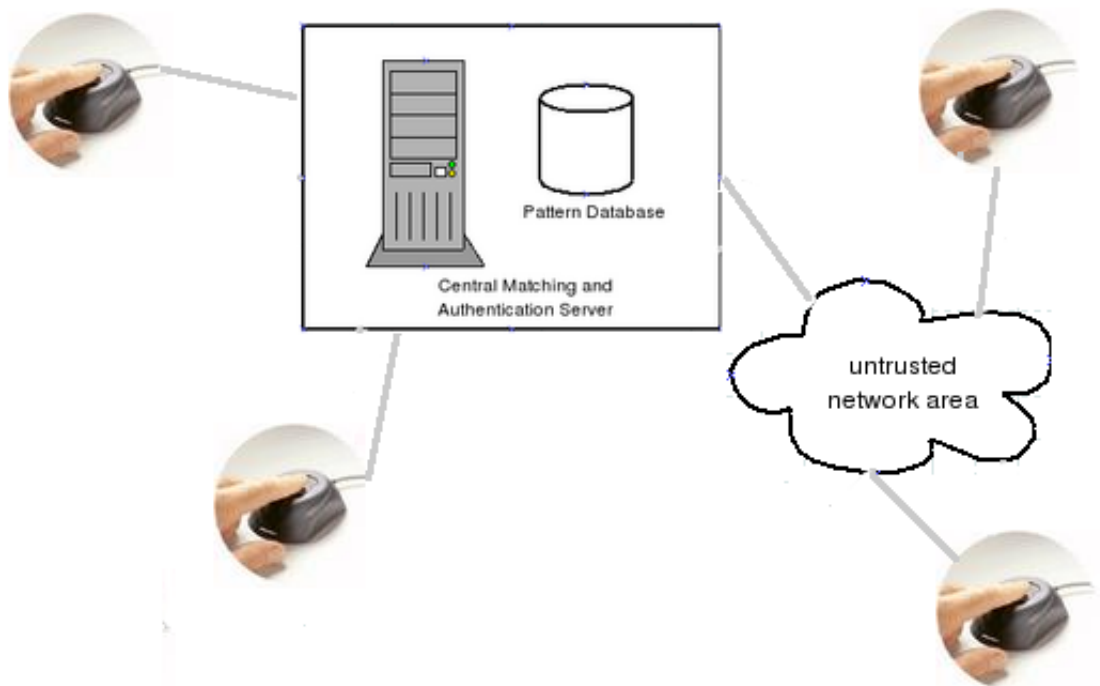


Figure 1.1: Typical biometric-based security system

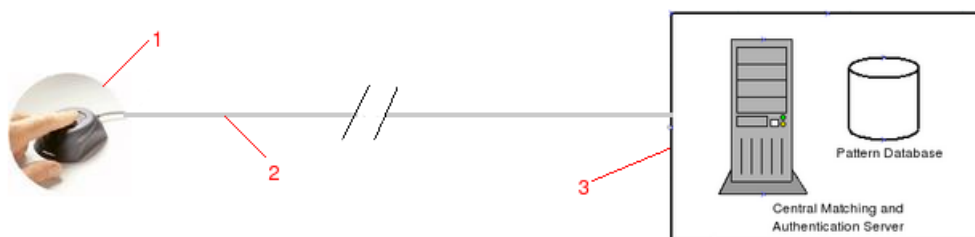


Figure 1.2: The 3 locations of technical weaknesses

gummy fingers [MATSUMO, 2003] [Putte,2004] [Schneier, 2005] or pictures of iris or faces [C'T, 2002]. To address this issue, the sensors need to make the difference between real and fake features, or in other words to implement a so-called *life-detection* mechanism. Such a mechanism can be achieved by measuring the pulse, temperature, blood pressure, electric resistance, etc. However, two German “hackers” claim [SF, 2003], for the case of fingerprints, that the “latex (used as a fake fingerprint) is thin enough to possibly allow this information to pass through the material”.

As part of her Master’s thesis [SF, 2003] on liveness detection and to determine if improvements had been made by the industry after [C'T, 2002] had shown in 2002 how easy their products were to bypassed, a Swedish student has tested fake fingerprints made of gelatin from latent fingerprint at vendors stands during the CeBIT 2004 trade fair in Germany. After each of the nine tested systems had been deceived, the conclusion of the thesis is that, although some systems check for liveness such as the temperature, it appears that the large tolerance of the built-in mechanisms keep them easy to defeat.

1.3.3.2 The storage and authentication system

As mentioned in the previous section about privacy and ethics, the storage and authentication system is also a very sensitive area as it contains the biometric patterns since it contains the data (patterns) and the decision mechanism (matching). While such a central system could be compromised by adding, removing or replacing patterns and permissions settings, it could also be a way to “steal” patterns for later use.

In a networked environment (and especially while connected to the Internet) maintaining an “attack-proof” system is a very hard task in terms of security and makes it therefore practically impossible to guarantee that data won’t be altered. Although vendors claim that, due to their format, pattern would not help any potential attackers in defeating systems, papers have been published demonstrating how indeed it is feasible to reconstruct fingerprints [BOMBA, 2003] [Hill, 2001] and face [Adler, 2003] [Adler-2, 2003] images from their patterns, by using reverse engineering techniques.

1.3.3.3 The communication medium

The scheme used at this location is quite similar to the one applied for the sensors: fooling the devices of the system (sensors and central authentication

system) by reproducing fake data. In practice, by sniffing and recording the communication data stream between the biometric reader and the authentication system and then playing it back at a later stage, thus doing a so-called replay attack [RILA, 2002], we can easily make them “believe” that they are communicating with each other as in a real authentication.

The figure 1.3.3.3 shows the sequences of an extremely simplified biometric system. Within this straightforward scheme, an attacker would listen grab the communication data stream being sent at stage 2, while the fingerprint information is being sent to the authentication system.

However real-life system would be likely to contain more sequences (ie, more “discussion” between the devices) and by thus making the communication harder to replay, as long as the communication is not authenticated between the devices, it still remains possible to engineer a *replay-attack*. The reason is that the authentication system (in the case of our simplified authentication sequences), when receiving data, can’t tell if they are coming from a sensor a genuine sensor of the biometric system or from a *trojan* device.

To address this technical issue and make sure that the received data are truly coming from trusted devices, we need to authenticate them. This can mainly be achieved by two different (but possibly combined) means: cryptographic digital signature and digital watermarking.

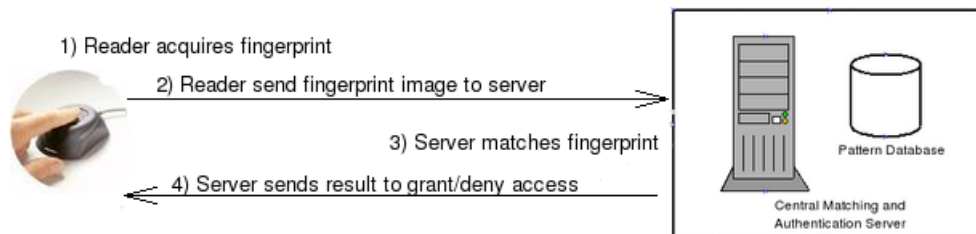


Figure 1.3: Sequences of a very simplified authentication scheme within a biometric system

Chapter 2

Data authentication

2.1 Cryptographic digital signature

2.1.1 Introduction

Cryptography has been used for thousands of years as a mean to safeguard diplomatic and military communications. While sending messages, the Roman emperor Julius Caesar employed a cipher to ensure that nobody else (the enemies) than the intended recipients (his troops) could read them. With the cryptography, aiming at finding methods to ensure safety and security of the conversation on one side, and the cryptanalysis aiming at breaking them on the other side, cryptology can be considered as an empirical science in the sense that a cipher is only considered secure until it has been broken.

Today's main goals and applications of cryptography are toward user authentication, data authentication (data integrity and data origin authentication), non-repudiation of origin, and data confidentiality [VANDENWAUVER, 1997].

Within cryptography, we usually make use of encryption (transforming a plain text into ciphertext) and decryption (the inverse transformation, to retrieve the original plain text from the ciphertext). These two transforms are performed using keys that are to be known the sender and/or the recipient and remain the only required information to encipher and decipher messages.

Two different types of encryptions are usually employed:

- Conventional or symmetric cryptography: the sender and receiver share the same secret key which is used for both encryption and decryption.

Conventional cryptography is computationally very efficient but secure exchange of keys is a major drawback.

- Public key or asymmetric cryptography: different keys are used to cipher and decipher. Each participant in the communication requires a pair of keys which are mathematically related but not derivable from each other [GAO, 1999], one is made secret (or private) and the other one public. If Bob wants to send a message to Alice [WIKIPEDIA, 2004], he needs to encipher it with Alice's public key. Alice's private key, which is kept secret and therefore only accessible by her, is then used to decipher the message from Bob. The major benefit of this scheme regards the key exchange. However, public key cryptography is usually by 1000 times slower than conventional [Mateti, 2000]. Therefore it is often combined with conventional encryption to encrypt random secret keys (then called session keys) [ASPENCRYPT].

A comprehensive introduction to public-key cryptography is given in [Young, 1996] and the Digital Signature Standard (DSS) in [NIST, 1994].

2.1.2 Application for data origin authentication

To recall more precisely what has just been described, the pair of keys used in public key encryption have the following properties:

- A message encrypted by a public key can only be decrypted by its corresponding private key.
- A message encrypted by a private key can only be decrypted by its corresponding public key.
- Although the two keys of a pair are mathematically related by the above cited properties, they are not derivable.

Given these properties, if Bob encrypts a message with Alice's public key prior to sending it to her, the only ways for Eve to read it (after she would have intercepted the ciphertext), since key can't be derived, are either to steal Alice's private key or to break the key by cryptanalysis. Although the first option is quite feasible, the second, thanks to advanced cryptographic algorithm, remains indeed computationally too intensive to be achieved in a reasonable period of time (i.e, a life span, for most of them).

While using cryptography as in the previous example brings confidentiality through the communication channel, it still does not help at authenticating the data origin. Even though Alice can use her private key to decrypt a message apparently received from Bob, nothing does indeed guarantee that it truly came from him since anybody could as well have performed the encryption, by simply using her public key. This is where the second property becomes useful. If Alice receives a message that she can decrypt with Bob's public key, then for sure it must have come from him since he is the only person who has access to his private key. This scheme is called digital signature. To be more accurate, the digital signature is normally produced by encrypting a hash [GAO, 1999] of the message and then attach it to the message. The verification being achieved by decrypting the signature (to retrieve the hash), calculating the hash of the message and then comparing those two values.

Since a key is too computationally intensive to be broken, while the dangerous exchange of secret keys (of symmetric encryption) is avoided, it could be concluded that this cryptography scheme is totally secure. However, there is a last remaining weakness usually referred to as the *man-in-the-middle-attack* [FACTINDEX]. Alice never met Bob in person and the public key that she is using to verify the origin of messages claiming to be from Bob was not been received in a secure manner. In fact when Bob once sent his public key to Alice, Eve intercepted it and replaced it by his own key to Alice instead. Therefore Eve can impersonate Bob by digitally signing messages to Alice.

This obvious glitch can however be solved by using a third-party, called Certificate Authority (CA), in which both Alice and Bob trust. Bob can now go to this authority, prove his identity with relevant documents (e.g, passport) and get his certificate (a document containing personal details and his public key) digitally signed by this authority. When receiving a certificate from Bob, Alice can verify its digital signature Certificate Authority's public key and establish the authenticity of the document, and by thus Bob's public key.

2.2 Digital watermarking

2.2.1 Definition and purposes

In *The Codebreakers* [KAHN, 1996], the author counts an history of Herodotus in which a message was tatoood in the shaven head of a slave, prior to sending him to Aristagoras at Miletus with the instruction to (again) shave his head. Although this clever idea is from another age, it appears that methods haven't changed so much when comparing them to nowadays state of the art digital watermarking. A thorough investigation of watermarking and steganography have been made by Kobayashi [KOBAYASHI, 1997] and Petitcola et al. [PETITCOLAS, 1999]

To avoid confusions, we shall first define the term *steganography*. This technique, which was indeed illustrated above with the history of Herodotus, consists in embedding or hiding secret data in an unsuspected cover object, making the assumption that third parties are unaware of the covert communication.

Digital watermarking, also known as *data hiding*, has the aim (as opposed to steganography) to resist attacks. It consists in embedding information in the media itself, whereas digital signature is a bitstream that is calculated from the media data and then attached to it. While it has been (though, non digitally) used for a long time by governments [BERGHEL, 1996] in currencies, postage stamps and other national ID such as driver licences or passports, digital watermarking can be applied in different areas such as:

- Protection of intellectual property
- Data integrity
- Data authenticity

Here are a few concrete scenarios where watermarking is applied:

- A photographer has sold a picture to a magazine A but notices that this same image has been probably copied because it now also appears in another magazine B to which he didn't sell it. If the hidden copyright information (the watermark) can be recovered from the illegal copy present in magazine B, then proof of authorship can be established and thus used for legal actions.

- An online service sells songs in digital format (such as *MP3* or *OGG Vorbis*) but wishes to track down fraudulent users redistributing the downloaded music via “peer-to-peer” networks. By hiding information identifying the purchase (and therefore the relevant customer) and then monitoring these networks to track songs containing their watermarks, the company can easily find out which customer spread the songs.
- A CCTV camera is connected to the local police station. As in the typical scenario seen in many movies, a burglar unplugs the surveillance camera to replace it by a device that will replay previously recorded video sequences taken by that camera, thus broadcasting fake images while hiding his acts. If the camera embeds a watermark containing a timestamp to be verified at the police station by the system receiving the streaming, then the *replay attack* can be detected

2.2.2 Why using digital watermarking ?

If the level of security achieved by current cryptographic techniques (digital signature in our case) is already sufficient, looking at another solution does not seem to be of an obvious necessity.

A. Tewfik points out in [TEWFIK, 2000] that cryptography usually relies on information being placed in “secure box” and locked with a “key”. This involves that the information itself can not be changed and that only people owning the proper key can gain access to it. However, once open, the security of the entire information protected in that “secure box” is then lost. If you compare this to the digital watermarking where the key is embedded and always resides in the information itself, then you already get an idea of its conceptual advantages.

A concrete illustration is the current DVD encryption methods. The CSS algorithm wraps the video in a container that can only be unlocked and then read if the DVD player provides the proper key. This means that if the video has ever been once decoded, it becomes relatively easy to trans-code its content and redistribute without any remaining mark. This is exactly what happened in 1999 [PATRIZIO, 1999] with a single player of a DVD licensee where the key had been left unencrypted, thus breaking the entire mechanism of DVD copy protection.

If you relate this to an ideal digital watermarking scheme where, despite any applied cryptography, the hidden data remains, even though alteration

has been made to the media, then it should clearly illustrate why approaches combining these two data authentication techniques (cryptography and digital watermarking) are very efficient.

A last argument towards the defence of watermarking versus “all-cryptography” is that “security through obscurity” (given that data hiding is not the only protection being applied) isn’t always a bad thing in the sense that a watermarked message does not appear, at first sight, to be containing any protection systems, thus, being less likely to attract the attention of an attacker, while the presence of encryption is much more obvious.

Chapter 3

Digital watermarking techniques review

3.1 Common requirements and techniques

3.1.1 General requirements

Different applications have different requirements. The following (and sometimes contradicting) adjectives given to different watermarking schemes show different characteristics that may have to be met:

- Blind (or public): the detection of the embedded data does not require the original signal, in opposition to non-blind (or private) watermarking where it is required.
- Semi-blind: some special information is needed to help the detector. In some schemes the “published” signal (i.e, the signal just after it was watermarked) is required to help detecting the presence of the watermark if the signal has been distorted.
- Semi-fragile: the mark is highly sensitive to modifications on the signal and can tell if the signal has been tampered.
- Fragile: in this scheme, it should be possible to detect any modifications but also where they were applied and possibly what the original signal was.
- Robust: the mark is aimed at residing in the host so that any reasonable transformations (i.e, not beyond any recognition of the signal) applied to the signal won't remove it.

Robustness to attacks (aiming at disabling the watermark) and imperceptibility (to the human visual system (HVS)) are at odds. Usually, the more robust the watermark, the more degraded the image. A watermark system is useless if the noise added by the mark results in an image degraded beyond a reasonable extent. Also, an attack is useless if the distortions required to disable the watermark degrade the image beyond any chance of further use. Therefore, a reasonably secure watermark scheme should consider the trade-off issue between robustness and image degradation in a way that any attacks that achieve disabling the watermark can guaranty the resulting image to become useless.

Furthermore, an ideal feature of a watermarking system is the implementation of a key-based mechanism so that finding the used algorithm is not enough to bring the scheme down. Also, a watermarking algorithm should in fact be public because “security through obscurity” is traditionnaly not a good thing.

After all these points have been introduced, it becomes apparent that the design of a “good” watermark scheme remains a quite difficult task. During the past 10 years, interest towards digital watermarking has been increasing, with most of the published papers covering data hiding in multimedia content such as images, video and sound for digital rights management (proof of authorship and ownership, tracking and monitoring of content, data authenticity and integrity, etc). Although, to limit the investigation to the scope of this thesis, this review will focus on the schemes related to image watermarking. Even though, the number of publications on image watermarking techniques is still too large to review them all and make a complete survey, but since they mostly share common principles, we will first show the most common ideas being used and also give more details on some methods.

3.1.2 Commonly used techniques

The embedded message is generally a binary sequence (i.e, a serie of 0’s and 1’s) that may be originating from the conversion of a string of characters (e.g, author or owner name, timestamp, etc), an image (e.g, copyright logo), etc. The reasons for preferring a binary sequence is that it makes it much easier by often embedding in the following way: if the current bit to hide is 1, then put a mark, otherwise leave unchanged.

If a binary sequences is not perfectly recovered, there may be a problem when converting it back to its original form (string, image, etc). For ex-

ample, the ASCII value of the character 'A' is of 65 which is converted the binary sequence 01000001. If a bit is not properly recovered, we may for instance obtain the sequence 01000000 instead, which has a decimal value of 64 and codes the character '@'. Another approach given by [DER-CHYUAN] proposes to embed a binary signal representing a black (bit of value 0) and white (bit value of 1) visually recognizable pattern. The idea is that the artifacts created by the not-properly retrieved bits still permit to some extent the recognition of the hidden pattern, as illustrated in 3.1.2. If this method improves the previous one, it requires a human to determine the validity of the recovered watermark and can therefore not be fully automated.



Figure 3.1: A black and white pattern watermark before and after additive noise

A pseudorandom numbers (PN) generator is almost always used because, once seeded, the same series of numbers can be regenerated. The autocorrelation [Lueke, 1992] is also a very helpful technique for data hiding. Concretely, if a PN signal is added to another signal, it will be correlated with the resulting (the addition) signal, even though other additions have been performed afterwards. So, if among a given set of PN signal, we want to know which ones have (and thus, which haven't) been added to the resulting signal, the coefficients of the calculated correlation between each candidate PN signal and the resulting signal will help determining the answer.

3.2 Different domains

A watermark can be embedded in different domains. Among them are the spatial domain, the DCT domain (either on the full image or block-wise) and the wavelet domain. While embedding an recovering is usually trivial in the spatial domain, it is mostly claimed ?? that more efficiency (great robustness of the mark as well as reduction of the noise added to the signal) can be obtained while performing in the transform domain (especially DCT and wavelet).

3.2.1 Spatial domain

In one of the first published papers on digital watermarking, Caronni proposed a so-called *tagging* method in ?? performing the following steps for the embedding task:

- Define a set of N by N blocks of pixel in the image, with regards to their correlation with neighbouring squares being below an empirically determined threshold. The set of then defined blocks locations determines a key K .
- For each block (and therefore, for each bit the message to hide), tag the blocks, by adding a pseudo random noise, only if the current bit to embed is set to 1.

In the recovery stage, the key is used to know what blocks are concerned. Then for each of this blocks, the correlation between the pseudo random noises (generated with the same seed as during the embedding stage) and the blocks are calculated to work out wheter it had been marked, and therefore determine if the current bit to recover is to be set to either 1 or to 0. Although the key in this scheme is a way to protect the watermark from unauthorized people, since it is determined by the properties of the original image, it is different from a secret key shared by two parties. In fact, such as scheme is a problem if the key can't be communicated through a trusted channel. Furthermore, since each image has its own key, later retrieval involves the management of a keys database, which is not always very convenient.

In [TIRKEL, 1993], Tirkel proposed a first approach embedding a PN signal in the least significant bits (LSB) of an image pixels. In another approach from Tirkel, it is also proposed to add a PN noise to the LSB instead of simply replacing them with message. The great advantages of these schemes is to involve no (for the HVS) distortions and to offer a very trivial recovery. However, both approaches are similarly robust to nothing. Even though these schemes could be used as a semi-fragile watermarking, a successful attack remains very easy to achieve. By taking another image (of at least the same size to ensure the whole message to fit in) and replacing its LSB with those of the originally watermarked image, we can assure (even if the message had before been encrypted, which would bring no real further security) that the expected message will be extracted during the recovery as if the image was legitimate.

A so-called “pathwork” method proposed by Bender et al. in [BENDER, 1995] consists in first randomly selecting a pairs of pixels X_i , Y_i in an image. Then, for each pair (and therefore, for each bit to embed) adding a value k to X_i while subtracting it to Y_i if the bit to embed is 1 (leaving the pair unchanged otherwise). The recovery is then performed by calculating the sum of the difference between respectively X_i and Y_i in the watermarked and original images. A sum equal to $2k$ determines that the recovered bit is to be set to 1, while otherwise (sum equal to zero) to 0.

3.2.2 DCT domain

Koch et al. in [Burgett, 1994], [Koch, 1995] and [Koch, 1996] proposed a watermarking scheme inspired by the JPEG compression scheme. After dividing the image into 8 by 8 pixels squares, the discrete cosine transform (DCT) of each is computed. Among the 12 pairs of pixel located in the middle frequency (noted F_m in figure 3.2.2, one is selected. If the bit to hide has a value of 1, then the coefficient are (if necessary) changed so that their difference becomes positive, or negative otherwise. While proceeding to this alteration, the quantization matrix is taken in consideration to achieve better robustness against JPEG compression. The recovery is done by determining wheter the recovered bits are to be set to either 1 or 0, with regards to the value of the coefficients difference. For this scheme, experimental results claims a robustness to JPEG compression down to 50% quality. Tao and Dickinson in [TAO, 1996] proposed an adaptive DCT-domain scheme following the JPEG compression table and claiming to resist JPEG compression down to 5%, as well as random noise. Bors et al. proposed other similar schemes based on DCT in [BORS, 1996] and [PITAS, 1996].

Swanson et al. proposed another robust scheme in [SWANSON, 1996] and [ZHU] achieved by the following algorithm:

- Divide the image in blocks and compute their DCT
- Compute a frequency mask of each block
- Scale the mask an multiply it by the DCT of a PN sequence
- Add the result to the DCT block and add the mask to make it invisible

The recovery is done by hypothesis testing and for which it requires the original image as well as the original watermarked image. Experimental results

claims a high robustness to different type of distortions (JPEG compression, random noise and cropping).

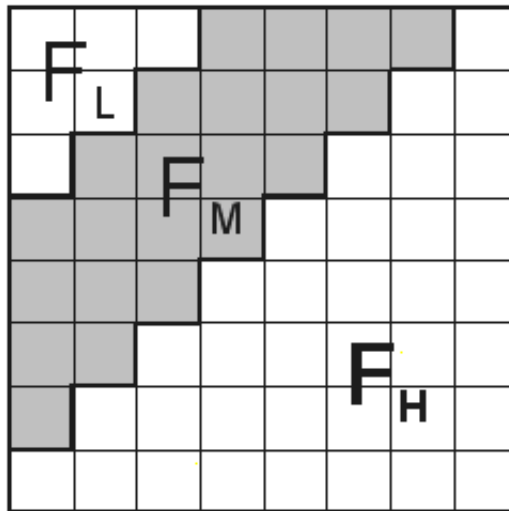


Figure 3.2: The DCT regions (low, middle and high frequencies)

Cox et al proposed in [COX, 1995] and [COX, 1996] a technique based on modulation in the frequency domain. The hidden data are a normal distribution $N(0,1)$, inserted in the perceptually significant spectral components. Detection requires the original image and the original watermarked image. The correlation of the original watermark (the difference between the original watermarked image and the original image) and the (potentially tampered) watermark to verify (the difference between the received watermarked image and the original image) is computed to determine the presence of the watermark. However these scheme is non-blind (i.e, requires the original data of the image and the watermark) the experimental results claim impressive robustness to JPEG compression down to 5% as well as a cycle of faxing-printing-photocopying-scanning.

3.2.3 DWT domain

The discrete wavelet transform (DWT) is used in numerous image watermarking techniques. The wavelet has numerous advantages over other transforms [BD, 2003]. It is sometime illustrated by the sentence “seeing the forest and the trees” as it allows to see both the general overview (the forest) and the specific details hidden in the mass object (the trees). The structure of a 3-level DWT decomposition can be seen in figure 3.2.3.

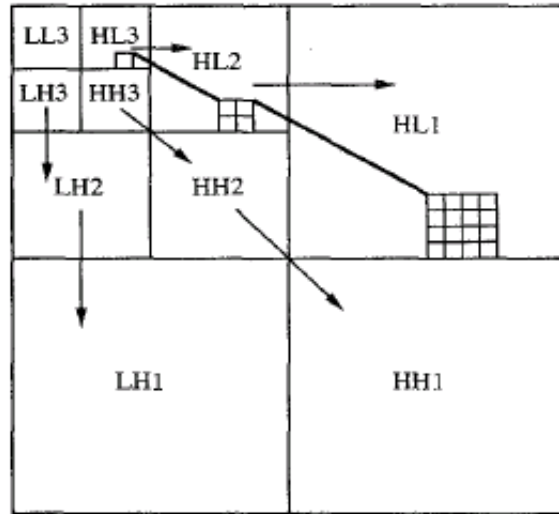


Figure 3.3: Structure of a 3-level 2-D wavelet decomposition

Der-Chyuan et al. [DER-CHYUAN] propose a non-blind watermarking scheme in the wavelet domains where they embed a visually recognisable pattern of black and white pixels. Rajal et al. [BD, 2003] also embed a visually pattern in the wavelet domain by performing the following steps:

- Perform the image into a 2-level two-dimensional wavelet decomposition of the image.
- Given that the visual pattern to embed is of the same size than the LL and HH subbands of the 2-level decomposition, multiply it by a determined factor and cover the LL and HH bands with it

The recovery requires the original image. Both the original image and the watermarked image are decomposed the LL and HH subbands subtracted and then subtracted by the previously used factor. The result should show the initially hidden pattern. Experimental results claim good robustness to JPEG lossy compression.

3.3 Watermarking evaluation

A concern about watermarking schemes is their evaluation. Petitcolas gives in [PETITCOLAS, 2000] a summary of assurance levels (see figure 3.3).

Level of assurance	Criteria
Low	<ul style="list-style-type: none"> - PSNR (when applicable) -A Slightly perceptible but not annoying
Moderate	<ul style="list-style-type: none"> -A Metric based on perceptual model -A Not perceptible under domestic conditions, that is using mass market consumer equipment
Moderate high	Not perceptible in comparison with original under studio conditions
High	Evaluation by a large panel of persons under strict conditions

Figure 3.4: Summary of assurance levels

It must be pointed that the only defined metric in this summary is the Peak Signal-To-Noise Ratio (PSNR). This metric (see figure ??) only gives a rough approximation of the data hiding quality. Other means of evaluation stricly rely on observation under varied conditions because there is no metric to take in consideration the effect on HVS.

$$PSNR = \frac{XY \max_{x,y} p_{x,y}^2}{\sum_{x,y} (P_{x,y} - \bar{P}_{x,y})^2}$$

Figure 3.5: Formulae of the PSNR

A benchmark tool called *stirmark* [PETITCOLAS, 2000] has been developed and is maintained by many researchers. Stirmark, is an engine performing a list of tests and attacks to evaluate watermarking techniques.

Chapter 4

Fingerprint images authentication based on digital watermarking

4.1 Technical requirements

As already explained, digital watermarking requirements depend on the application. If we want to authenticate fingerprint images received from a remote sensor, through an untrusted communication channel, our watermarking-based authentication system should:

- not require the original image or watermark;
- to some extent, not affect the fingerprint recognition;
- feature a key-based mechanism to protect the hidden data;
- allow a fully automated verification of the watermark;
- detect possible *replay attacks*;
- possibly be robust to some extent

4.1.1 Blind watermark

The reasons for choosing a blind watermarking scheme is quite obvious. If watermarking-based copyright proof systems can afford non-blind or semi-blind schemes, it is because they already have a hand to the original image (or original watermarked image) they want to prove authorship of. In our

case, the only object we can get access to, in order to establish the authenticity, is the (possibly tampered) watermarked image. Therefore, a blind watermarking is the only possible options.

4.1.2 Least consequences on the recognition

As already explained, depending on the techniques being used, the noise being brought by digital watermarks, more or less, distorts the signal. The resulting degradation has usually for effect the addition of artifacts to the cover image.

Although this would need to be verified, the degradation of a fingerprint image might affect its recognition during the matching process. We could for example expect the artifacts due to the watermark to result into spurious minutiae being created (and detected) or deleted in the fingerprint, thus bringing the matching score to a lower value.

Therefore, although the authentication of fingerprints images might have a cost on the overall system, if the trade-off between the addition of security and the loss of efficiency is too high, we may have to conclude that the idea of a watermark-based authentication is not well suited for a biometric system.

4.1.3 Key-based mechanism

“Security through obscurity” is traditionnally not a factor to rely on. Therefore, keeping the watermarking scheme secret, as the only way to be protected against attacks, can very hardly guaranty that nobody is ever going to find out the technique being used. Security has always been an empirical activity and usually reaches a higher level when the employed techniques are made public, while the strength of the systems relies on a key which is kept secret and remains the only way to gain access to what is to be protected.

In a system where no key-based mechanism is employed and where the only secret is the technique being used, a potential attacker would first have to find out the specifications of the scheme (which might not be of a great difficulty since they almost all share common principles) and then, without leaving any evidences, he could remove, extract or change the data of an embedded watermark.

Therefore, the use of a key-based mechanism to protect the scheme becomes an urgent requirement for the development of a secure authentication

system. Furthermore, depending on the cipher being used, a mechanism to generate, update and securely populate keys on such a remotely distributed system would be a further enhancement.

4.1.4 Fully automated verification

If the authentication system is to be used within an automated fingerprint identification system (AFIS), it is obvious that the watermark verification also has to be fully automated. As explained in the literature review, if only one bit of a watermark message is not recovered properly, then the message gets wrong and its comparison to the expected message will therefore fail. The visually recognisable *pattern message* technique, presented in the literature review, remains, to some extent robust to a proportion of wrongly recovered bits. However, unless thinking of a pattern recognition system for the recovered message, we can hardly see how this could be automated.

Furthermore, when watermarked images are to be recovered for instance to prove ownership, there is plenty of time and also the judgement of a human being to determine if the tampered watermarked image should be geometrically translated (resizing, rotating, etc) in order to help the recovery. A system that is not only fully automated but also has to give real-time results can not be expected to take decisions helping the recovery in this manner. Therefore it must allow the recovery to be done with the received image “as is”.

4.1.5 Detection of *replay-attacks*

Establishing the authenticity (thus, the origin) not only consists in verifying the presence of authentication data but also involves their validation. In fact, even without knowledge of the implemented authentication system, an attacker could potentially “steal” a fingerprint image from the communication channel to, later on, send it again to the device in charge of matching it against a database, thus doing a so-called *replay-attack*. If the authentication only verifies the presence of the watermark (e.g, extracting it, thanks to the key, and checking that the hidden message corresponds to the serial number of known device), then the authentication system is not worth much more than no authentication at all. The only difference being that the stolen image would need to have been stolen from that same system, which is most likely what an attacker would do even if he was not aware of that small requirement.

This is why the watermarking scheme must embed some data that can not be replayed. A common technique for this purpose is to do a so-called

challenge-response. This challenge is different everytime and usually consists in embedding the current timestamp and check, at the recovery stage, that the difference between the current time and the hidden timestamp is relevant (i.e, does not exceed a treshold determined by the expected time it takes for image to “travel” from the sensor to the recognition device). So, if an attacker steals an image at time t and tries to replay it at $t+N$ (N being above the defined treshold), the deception can be automatically detected and, possibly, recorded to warn the person in charge of the biometric system.

A remaining requirement for this *challenge-response* to ensure detecting *replay-attacks* is obviously that an attacker, aware of the watermarking technique, can not reset the timestamp. This is why the hidden data have to be locked by a key-based mechanism so that only the intended recipient can get access to it, thus avoiding an attacker to get around the protection by updating the timestamp to the current time of the attack.

4.1.6 Robust vs. fragile

Furthermore, an attacker getting hold of a fingerprint image (while going through the communication channel) and aware of the protection could want to disable the embedded watermark (with methods such as additive noise, compression, geometrical translations, etc). Nevertheless, this would be of no use to get around the authentication system since images without a valid watermark (thus, not authenticated), are to be rejected anyway. However, we could imagine that the attacker may reuse this images to impersonate users in a biometric system where images origin is not verified. ideally, the watermark should be robust up to resulting in a fingerprint image too distroted to be matched within a biometric system

require a distortion

a formerly protected image fingerprint image in which the watermark has been disabled

Furthermore, although “attacking” the image (with additive noise, compression, geometrical translations, etc) in order to disable its embedded watermark would be of no use to get around the authentication system, an additional feature could be the robustness of the scheme to possibly track down images reused in a *replay attack*.

4.1.7 Robustness vs.

disabling the digital watermarking Furthermore, even though fingerprint images where the watermark would have been disabled after attacks (additive

noise, compression, geometrical translation, etc) would be rejected by the system and therefore being of no use to deceive the protection

Bibliography

- [Poh] P. Norman, *Introduction to Biometric Authentication*. University of Strasbourg. Retrieved July 12, 2004, from <http://hydria.u-strasbg.fr/~norman/BAS/resources/IntroductionToBiometrics/IntroductionToBiometrics.ppt>
- [FindBiometrics] J. Chang, *Biometrics Industry Report: An Interview with John Chang*. FindBiometrics.com. Retrieved July 12, 2004, from http://www.findbiometrics.com/Pages/feature%20articles/john_chang.html
- [Pankanti, 2002] S. Pankanti et al., *On the Individuality of Fingerprints*. IEEE Transactions on PAMI, Vol. 24, No. 8, pp. 1010-1025, 2002.
- [Ross, 2004] A. Ross, A. Jain, *Hand Geometry*. Biometrics at Michigan State University. Retrieved July 12, 2004, from http://biometrics.cse.msu.edu/hand_geometry.html
- [NCSC, 2002] NSCS, *Hand Geometry*. Individual Biometrics, NSCS Home Page. Retrieved July 12, 2004, from <http://ctl.ncsc.dni.us/biomet%20web/BMHand.html>
- [Fein, 2004] A. Fein, *An introduction to Biometrics*. MGT496a, final paper. March 3, 2004.
- [Daugman, 2003] J. Daugman, *Tests of the Daugman Iris Recognition Algorithms*. Iris Recognition: Reported test results. Retrieved July 12, 2004, from <http://www.cl.cam.ac.uk/users/jgd1000/iristests.pdf>
- [Daugman, 2004] J. Daugman, *John Daugman's webpage*. Cambridge University, Computer Laboratory, Cambridge UK. Retrieved July 12, 2004, from <http://www.cl.cam.ac.uk/users/jgd1000/>

- [Hong, 2004] A. Jong, Y. Kulkarni, A. Ross, A. Jain, *Integrating Faces and Fingerprints for Personal Identification*. Biometrics: Multibiometrics, Biometrics at Michigan State University. Retrieved July 12, 2004, from http://biometrics.cse.msu.edu/multi_bio.html
- [SUN, 1999] W. Sun, *Face Identification*. Shape Analysis in Computer Vision. Final Project Report, Department of Electrical Engineering, McGill University. Retrieved July 12, 2004, from <http://www.cim.mcgill.ca/~wsun/sa/project/node7.html>
- [IBL, 2002] IBL, *Retina recognition*. About Biometrics. Retrieved July 12, 2004, from http://www.eb.uah.edu/ece/biometric/retina_recognition.htm
- [NGUGI, 2004] B. Ngugi, *Fighting Identity Fraud with the Addition of Biometric Techniques*. American Conference on Information Systems, New York, USA.
- [SPAGNOLO, 2004] G.S. Spagnolo et al, *Superposed strokes analysis by conoscopic holography as an aid for a handwriting expert*. Journal of Optics A: Pure and Applied Optics, Volume 6, Number 9, September 2004, Pages 869-874.
- [C'T, 2002] L. Thalheim, J. Krissler, P.M. Ziegler, *Koerperkontrolle: Biometrische Zugangssicherungen auf die Probe gestellt*. C'T Magazine, November, 2002, August 21, p. 114.
- [MATSUMO, 2003] T. Matsumoto, *A Case Study for Identification*. October, 2002. Retrieved July 17, 2004, from <http://www.itu.int/itudoc/itu-t/workshop/security/present/s5p4.pdf>
- [Putte,2004] T. van der Putte, *Biometrical Fingerprint Recognition: Don't Get Your Fingers Burned, Esire, an Origin Extended Enterprise*. . September 21, 2000. Retrieved July 17, 2004, from <http://cryptome.org/fake-prints.htm>
- [Schneier, 2005] T. van der Putte, *Hacking Fingerprint Readers*. Cryptograms, November 15, 2003. Retrieved July 17, 2003, from <http://www.schneier.com/crypto-gram-0311.html>
- [SF, 2003] A. Harrison, *Hackers Claim New Fingerprint Biometric Attack*. Security Focus, August 12, 2003. Retrieved July 17, 2003, from <http://www.securityfocus.com/news/6717>

- [Uludag, 2004] U. Uludag and A.K. Jain, *Attacks on biometric systems: a case study in fingerprints*, Proc. SPIE-EI 2004, Security, Seganography and Watermarking of Multimedia Contents VI, San Jose, CA, January 18-22, 2004.
- [SF, 2003] M. Sanstroem, *Liveness Detection in Fingerprint Recognition Systems*. Master Thesis, June 4, 2004. Retrieved July 17, 2003, from <http://www.ep.liu.se/exjobb/isy/2004/3557/>
- [BOMBA, 2003] M. Bomba, *On the reconstruction of biometric raw data from template data*. Master thesis. July 9, 2003
- [Adler, 2003] A. Adler, *Can images be regenerated from biometric templates*. Biometrics Conference, Sept 22-24, 2003.
- [Adler-2, 2003] A. Adler, *Sample images can be indepedently restored from face recognition templates*. Can. Conf. Electrical Computer Eng. (CCECE). Montreal, Canada, May 2003. pp. 1163-1166.
- [Hill, 2001] C.J. Hill, *Risk of Masquerade Arising from the Storage of Biometrics*. B.S. Thesis. November, 2001
- [RILA, 2002] L. Rila and C. J. Mitchell, *Security analysis of smartcard to card reader communications for biometric cardholder authentication*. Proceedings of CARDIS '02, 5th Smart Card Research and Advanced Application Conference, San Jose, California, November 2002, USENIX Association, Berkeley, CA (2002), pp.19-28.
- [VANDENWAUVER, 1997] M. Vandenwauver, *Introduction to Cryptography*. Laboratorium ESAT-Groep COSIC, Katholieke Universiteit Leuven. Retrieved July 12, 2004, from <http://www.esat.kuleuven.ac.be/cosic/intro/>
- [GAO, 1999] Shuhong Gao, *Theory and models*. Mathematical Models in Public-Key Cryptology, October 17, 1999. Retrieved July 12, 2004, from http://www.math.clemson.edu/faculty/Gao/crypto_mod/node2.html
- [WIKIPEDIA, 2004] Wikipedia, *Alice and Bob*. Wikipedia, the free encyclopedia, August 13, 2004. Retrieved August 23, 2004, from http://en.wikipedia.org/wiki/Alice_and_Bob

- [Mateti, 2000] P. Mateti, *Cryptography Internet Security*. Lectures. Retrieved July 14, 2004, from <http://www.cs.wright.edu/~pmateti/Courses/499/Cryptography/>
- [ASPENCRYPT] ASP Encrypt, *Public-Key Cryptography*. Crypto 101, Chapter 4. Retrieved July 12, 2004, from http://www.aspencrypt.com/crypto101_public.html
- [Young, 1996] D. Young, *An introduction to digital signature*, The Youd Zone, 1996. Retrieved Mai 14, 2004, from <http://www.youdzone.com/signature.html>
- [NIST, 1994] U.S. National Institute of Standards and Technology, *Announcing the Standard for Digital Signature Standard (DSS)*, Federal Information Processing Standards Publication 186, May, 1994. Retrieved Mai 12, 2004, from <http://www.itl.nist.gov/div897/pubs/fip186.htm>
- [FACTINDEX] Fact Index, *Man in the middle attack*. Wikipedia definition. Retrieved Mai 14, 2004, from http://www.fact-index.com/m/ma/man_in_the_middle_attack.html
- [KAHN, 1996] D. Kahn, *The Codebreakers; The Comprehensive History of Secret Communication from Ancient Times to the Internet*. Scribner, December, 1996.
- [BERGHEL, 1996] H. Berghe et al., *Protecting Ownership Rights through Digital Watermarking*. IEEE Computer, 29:7, pp. 101-103 (1996).
- [BERGHEL, 1997] H. Berghel, *Digital Watermarking*. January 2, 1997. Retrieved Mai 12, 2004, from http://www.acm.org/~h1b/publications/dig_wtr/dig_watr.html
- [KOBAYASHI, 1997] M. Kobayashi, *Digital Watermarking: Historical Roots*. IBM Research, Tokyo Research Laboratory, Technical Report, April, 1997
- [PETITCOLAS, 1999] F.A.P Petitcolas, R.J Anderson, and M.G Kuhn, *Information Hiding - A Survey*. source. Proceedings of the IEEE, vol 87, no 7, July, 1999

- [WATERMARKINGWORLD, 2001] F.A.P. Petitcolas, J.H. Lee, M. Brunet, S. Katzenbeisser, M. Kutter, *Watermarking FAQ*. Watermarking World. Retrieved Mai 9, 2004, from <http://www.watermarkingworld.org/faq.html>
- [CHIOU-TING] H. Chiou-Ting, *Digital Watermarking*. Communication and Media Laboratory, National Taiwan University. Retrieved May 2, 2004, from <http://www.cmlab.csie.ntu.edu.tw/~candy/watermark/wat.html>
- [Meerwald, 2001] P. Meerwald, A. Uhl, *A Survey of Wavelet-Domain Watermarking Algorithms*. EI San Jose, CA, USA, 2001.
- [TEWFIK, 2000] A.H. Tewfik, *Digital Watermarking*. IEEE Signal Processing Magazine, vol 17, pp 17-88, September 2000.
- [PATRIZIO, 1999] A. Patrizio, *Why the DVD Hack Was a Cinch*. Wired News, November 2, 1999. Retrieved May 9, 2004, from <http://www.wired.com/news/technology/0,1282,32263,00.html>
- [PETITCOLAS, 2000] F.A.P. Petitcolas, *Watermarking Schemes Evaluation*. IEEE Signal Processing Magazine, Vol 17, pp 58-64, September 2000.
- [Lueke, 1992] H.D. Lueke, *Korrelationssignale*. Berlin, Germany. Springer, 1992.
- [Caronni, 1993] G. Caronni, *Ermitteln unauthorisierter Verteiler von Maschinenlesbar Daten*. ETH, Zurich, Switzerland, Technical Report, August, 1993.
- [Caronni, 1995] G. Carroni, *Assuring ownership right for digital images*. Proc. VIS 1995, Session "Reliable IT Systems", Vieweg, Germany, 1995
- [TIRKEL, 1993] A. Tirkel, G. Rankin, R. van Schyndel, W.Ho, N. Mee, C. Osborne, *Electronic water mark*. Proc. DICTA 1993
- [BENDER, 1995] W. Bender, D. Gruhl, N. Morimoto, *Techniques for data hiding*. Proc. SPIE, vol. 2420, San Jose, February, 1995
- [Burgett, 1994] S. Burgett, E. Koch, J. Zhao, *A novel method for copyright labeling digitized image data*. Fraunhofer Institute, Computer Graphics, Darmstadt, Germany, Technical Report, September, 2004

- [Koch, 1995] E. Koch, J. Zhao, *Toward robust and hidden image copyright labelling*. Proc. Workshop, Nonlinear Signal and Image Processing, Marmaros, Greece, June, 1995
- [Koch, 1996] E. Koch, J. Rindfrey, J. Zhao,, *Copyright protection for multimedia data*. Digital Media and Electronic Publishing, 1996
- [TAO, 1996] B. Tao, B. Dickinson, *Adaptive watermarking in the DCT domain*. Proc. Int. Conf. Image Processing, Lausanne, Switzerland, September, 1996
- [BORS, 1996] A. Bors, I. Pitas, *Embedding parametric digital signatures in images*. EUSIPCO-96, Trieste, Italy, September, 1996
- [PITAS, 1996] A. Bors, I. Pitas, *Image watermarking using DCT domain constraints*. Proc. Int. Conf. Image Processing, Lausanne, Switzerland, September, 1996
- [SWANSON, 1996] M.D. Swanson, B. Zhu, A.H. Tewfik, *Robust Data Hiding for Images*. IEEE Digital Signal Processing, Workshop, Loen, Norway, pp. 37-40
- [ZHU] M.D. Swanson, B. Zhu, A.H. Tewfik, *Transparent Robust Image Watermarking*. Proc. Int. Conf. Image Processing, Lausanne, Switzerland, September, 1996
- [COX, 1995] I. Cox, J. Kilian, T. Leighton, T. Shamoan, *Secure spread spectrum watermarking for images, audio and video*. NEC Res. Inst., Princeton, NJ, Technical Report, 95-10, 1995
- [COX, 1996] I. Cox, J. Kilian, T. Leighton, T. Shamoan, *Secure spread spectrum watermarking for images, audio and video*. Proc. IEEE Int. Conf. Image Processing, Lausanne, Switzerland, September, 1996
- [HSU, 1997] C.T. Hsu, *Digital watermarking for images and video*. Ph.D. dissertation, Communication Multimedia Lab., National Taiwan University, 1997 Security Technology, 2003. Proceedings. IEEE 37th Annual 2003 International Carnahan Conference on , Oct. 14-16, 2003 Pages:370 - 377
- [DER-CHYUAN] L. Der-Chyuan; L. Jiang-Lung; C. Ming-Chang, *Digital watermarking using multiresolution wavelet transform*. Security Technology, 2003. Proceedings. IEEE 37th Annual 2003 International Carnahan Conference, October, 2003 pp. 370-377

- [TACHIBANA, 1996] Y. Tachibana, A. Shima, *On the detection of isolated signals by the partial spectrum of Daubechies wavelet*. Industrial Electronics, Control, and Instrumentation, 1996., Proceedings of the 1996 IEEE IECON 22nd International Conference on , Volume: 3 , 5-10 Aug. 1996 pp.1341-1346 vol.3
- [BD, 2003] Beyond Discovery, *Wavelets: Seeing the Forest and the Trees*. Beyond Discovery. Retrieved June 17, 2004, from <http://www.beyonddiscovery.org/content/view.txt.asp?a=1952>
- [RAVAL] M.S. Raval, P.P. Rege, *Discrete wavelet transform based multiple watermarking scheme*. TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region , Volume: 3 , 15-17 Oct. 2003, pp.935-938 Vol.3