42

VIDEO WATERMARKING
OVERVIEW AND CHALLENGES

Gwenaël Doërr and Jean-Luc Dugelay
Multimedia Communications, Image Group
Eurécom Institute
Sophia-Antipolis, France
jean-luc.dugelay@eurecom.fr

1. INTRODUCTION

If you hold a common banknote up to the light, a watermarked drawing appears. This watermark is invisible during normal use and carries some information about the object in which it is embedded. The watermarks of two different kind of banknotes are indeed different. This watermark is directly inserted into the paper during the papermaking process. This very old technique is known to prevent common methods of counterfeiting. In the past few years, the use and distribution of digital multimedia data has exploded. Because it appeared that traditional protection mechanisms were no longer sufficient, content owners requested new means for copyright protection. The previous paper watermark philosophy has been transposed to digital data. Digital watermarking, the art of hiding information in a robust and invisible manner, was born. The recent interest regarding digital watermarking is demonstrated in Table 1, which reports the increasing number of scientific papers dealing with this subject. Today, entire scientific conferences are dedicated to digital watermarking e.g. “SPIE: Security and Watermarking of Multimedia Content”. Moreover, even if it is a relatively new technology, some industries have already commercialised watermarking products e.g. the widespread Digimarc.

Table 1. Number of publications having the keyword *watermarking* as their main subject according to INSPEC database, July 2002.

Year	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**
Publications	2	21	54	127	213	334	376

Digital watermarking has first been extensively studied for still images. Today however, many new watermarking schemes are proposed for other types of digital multimedia data, so called as *new objects*: audio, video, text, 3D meshes... This chapter is completely devoted to digital video watermarking. Since the main subject of this book is video databases, the reader is assumed not to be familiar with the concept of digital watermarking. Consequently, the fundamentals of the theory are presented in Section 2. Many applications of digital watermarking in the context of the video are presented in Section 3 in order to give an overview of the possible benefits that technology can bring. The Section 4 lists the main challenges that have to be taken up when designing a new video watermarking system. Finally, the major trends in the domain, to the best knowledge of the authors, are reported in Section 5.

2. WHAT IS DIGITAL WATERMARKING?

The end of the previous millennium has seen the transition from the analog to the digital world. Nowadays, audio CDs, Internet and DVDs are more and more widespread. However film and music content owners are still reluctant to release digital content. This is mainly due to the fact that if digital content is left unprotected, it can be copied rapidly, perfectly, at large scale, without any limitation on the number of copies and distributed easily e.g. via Internet. Protection of digital content has relied for a long time on encryption but it appeared that encryption alone is not sufficient enough to protect digital data all along its lifetime. Sooner or later, digital content has to be decrypted in order to be eventually presented to the human consumer. At this very moment, the protection offered by encryption no longer exists and a user may duplicate or manipulate it.

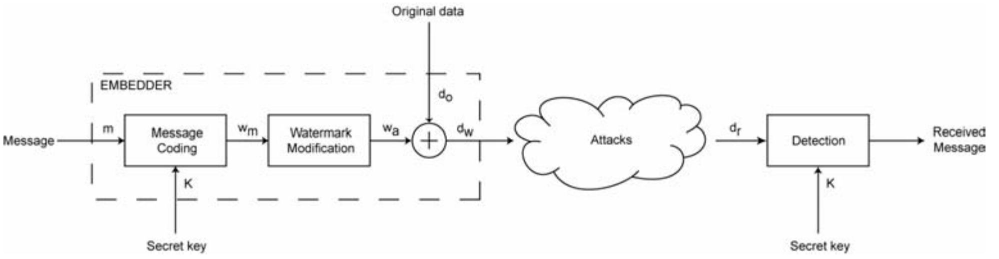

Figure 1. Simple watermarking scheme

Digital watermarking has consequently been introduced as a complementary protection technology. The basic idea consists in hiding information imperceptibly into digital content. This watermarked signal should survive most common signal processings and even malicious ones if possible. The hidden information is inherently tied to digital content and protects it when encryption has disappeared. It is important to understand that digital watermarking does not replace encryption. Those are two complementary techniques. On one hand, encryption prevents an unauthorised user from accessing digital content in clear during its transport. On the other hand, digital watermarking leaves an underlying invisible piece of evidence in digital data if a user, who had access to the data in clear after decryption, starts using digital data illegally (reproduction, alteration).

Depending on what information is available during the extraction process, two separate classes of watermark detectors have been defined. If the detector has access to the original data additionally to the watermarked data, the watermark detector is called non-blind. However this kind of algorithm is less and less represented nowadays. Keeping an original version of each released digital data is indeed a very strong constraint for digital content owners in terms of storage space. As a result, most of the watermark detectors are actually considered as blind: the detector has only access to the watermarked data in order to extract the hidden message.

The Figure 1 depicts a simple watermarking scheme with blind detection. The goal is to embed the message m into some original data d_o. The first step consists in encoding the message to be hidden with a secret key K. Typically the message is over sampled in order to match the dimension of the original data and is XORed with a pseudo-random noise generated thanks to a pseudo-random number generator which takes the secret key K as an input seed. Next, the generated watermark signal w_m is modified e.g. it is scaled by a given watermarking strength. The final step simply adds the obtained watermark w_a to the original data in order to obtain the watermarked data d_w. This watermark embedding could be performed in whatever desired domain (spatial, Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT), Fourier-Mellin). Watermarked data is then transmitted and is likely to be submitted to various processings (lossy compression, noise addition, filtering) which can be seen as attacks altering the watermark signal. If at some moment, someone wants to check if a watermark has been embedded with the secret key K in some received digital data d_r, the data is simply sent through a detector. The majority of the existing detection algorithms can be seen as the computation of a correlation score between received data d_r and the generated watermark w_m. This correlation score is then compared to a threshold in order to assert the presence of the watermark or not.

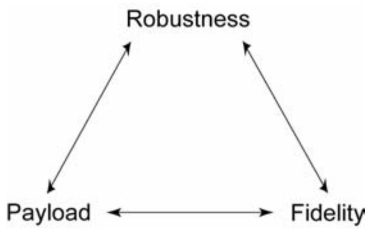

Figure 2. Trade-off in digital watermarking

There exists a complex trade-off in digital watermarking between three parameters: data payload, fidelity and robustness. It is illustrated in Figure 2 and further presented below.

Payload
Data payload can be defined by the number of bits that can be hidden in digital data, which is inherently tied to the number of alternative messages that can be embedded thanks to the watermarking algorithm. It should be noted that, most of the time, data payload depends on the size of the host data. The more host samples are available, the more bits can be hidden. The capacity is consequently often given in terms of bits per sample.

Fidelity

Watermarking digital content can be seen as an insertion of some watermark signal in the original content and this signal is bound to introduce some distortion. As in lossy compression, one of the requirements in digital watermarking is that this distortion should remain imperceptible. In other terms, a human observer should not be able to detect if some digital data has been watermarked or not. The watermarking process should not introduce suspicious perceptible artefacts. The fidelity can also be seen as the perceptual similarity between watermarked and unwatermarked data.

Robustness

The robustness of a watermarking scheme can be defined as the ability of the detector to extract the hidden watermark from some altered watermarked data. The alteration can be malicious or not i.e. the alteration can result from a common processing (filtering, lossy compression, noise addition) or from an attack attempting to remove the watermark (Stirmark [40], dewatermarking attack [44]). As a result, the robustness is evaluated via the survival of the watermark after attacks.

It is quite easy to see that those three parameters are conflicting. One may want to increase the watermarking strength in order to increase the robustness but this results in a more perceptible watermark on the other hand. Similarly, one can increase the data payload by decreasing the number of samples allocated to each hidden bit but this is counterbalanced by a loss of robustness.

As a result, a trade-off has to be found and it is often tied to the targeted application. It is useless to design a high capacity algorithm if there are only a few different messages to be hidden in practice. This is typically the case in a copy control application where two bits are enough to encode the three messages *copy-always*, *copy-once* and *copy-never*. Most of the time, the watermark signal should have a low energy so that the induced distortion remains imperceptible. However in a high degrading environment, it is sometimes necessary to embed a strong watermark in order to survive the transmission. Finally some applications do not require the watermark to be robust. In fact the weakness of a fragile watermark can even be exploited in order to ensure the integrity of digital data [43]. If no watermark is found, digital data is not considered legitimate and is discarded. There is not consequently *one* optimal watermarking algorithm. Each watermarking scheme is based on a different trade-off and one has to be cautious when benchmarking various algorithms. It should be ensured that the methods under investigation are evaluated under similar conditions [29]. In other terms, in order to perform a fair performance comparison in terms of robustness, the evaluated watermarking algorithm should have roughly the same capacity and introduce approximately the same visual distortion.

The last few years have seen the emergence of a new trend in the watermarking community. The watermarking process is now seen as the transmission of a signal through a noisy channel. Original data is then seen as interfering noise which reduces significantly the amount of reliably communicable watermark information. In this new perspective, Chen and Wornell noticed a precious paper written by Costa [8]. He showed that, if a message is sent through a channel corrupted by two successive additive white Gaussian noise sources and if the transmitter knows the first noise source, interference from the first noise source can be entirely cancelled.

From a watermarking point of view, the message can be seen as the watermark, the first known noise source as the original data and the second unknown noise source as the attacks. Even if Costa's model is substantially different from a real watermarking system, it means that side information at the embedder enables to reduce interference from the original data. This implication has received further support from subsequent theoretical work.

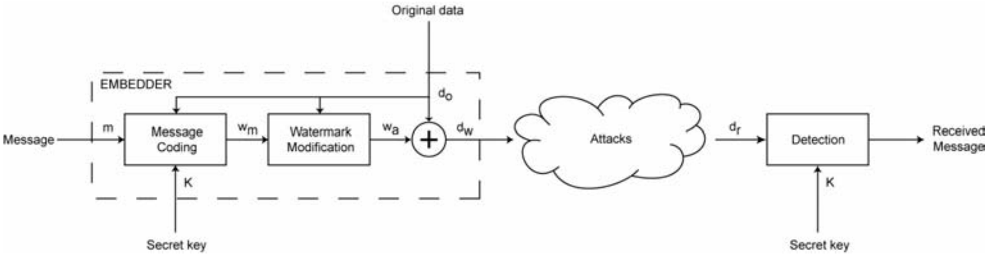

Figure 3. Informed watermarking scheme

In Figure 1, the embedder can be seen as blind. Information contained in the original data is not exploited during the *message coding* and *watermark modification* steps. Costa's work encourages designing new algorithm based on Figure 3 where side information is taken into account during those two steps. Informed watermarking can be done during message coding (informed coding) and/or watermark modification (informed embedding). With informed coding, for a given message, a pool of different alternative watermarks is available and the embedder chooses the one for which the interference introduced by the original data will be minimised. With informed embedding, the goal is to optimally modify the watermark so that the detector extracts the expected message. A typical example is to perceptually shape the watermark accordingly to the original data so that fidelity is increased while robustness is maintained.

Since presenting the whole theory behind digital watermarking is far beyond the scope of this chapter, the interested reader is invited to read the various books devoted to the subject. An introducing overview of digital watermarking can be found in [26]. Further details are developed in [9] where the authors even provide samples of source code. Finally an in depth discussion on informed watermarking is conducted in [19].

3. APPLICATIONS OF WATERMARKING VIDEO CONTENT

If the increasing interest concerning digital watermarking during the last decade is most likely due to the increase in concern over copyright protection of digital content, it is also emphasised by its commercial potential. The following section is consequently completely dedicated to the presentation of various applications in which digital watermarking can bring a valuable support in the context of video. Digital video watermarking may indeed be used in many various applications and some of them are far from the original copyright enforcement context. The applications presented in this section have been gathered in Table 2. This is not an exhaustive list and many applications are still to be imagined.

Table 2. Video watermarking: applications and associated purpose

Applications	**Purpose of the embedded watermark**
Copy control	Prevent unauthorised copying.
Broadcast monitoring	Identify the video item being broadcasted.
Fingerprinting	Trace back a malicious user.
Video authentication	Insure that the original content has not been altered.
Copyright protection	Prove ownership.
Enhanced video coding	Bring additional information e.g. for error correction.

3.1 COPY CONTROL

The Digital Versatile Disk (DVD) and DVD players appeared on the consumer market in late 1996. This new technology was enthusiastically welcomed since DVD players provide a very high-quality video signal. However, the advantages of digital video are counterbalanced by an increased risk of illegal copying. In contrast to traditional VHS tape copying, each copy of digital video data is a perfect reproduction. This raised the concern of copyright owners and Hollywood studios request that several levels of copy protection should be investigated before any device with digital video recording capabilities could be introduced.

The Copy Protection Technical Working Group (CPTWG) has consequently been created in order to work on copy protection issues in DVD. A standard has not been defined yet. However a system, which could become the future specification for DVD copy protection, has been defined [5]. The three first components are already built in consumer devices and the other three are still under development.

- **The Content Scrambling System (CSS)**. This method developed by Matsushita scrambles MPEG-2 video. A pair of keys is required for descrambling: one is unique to the disk and the other is specific to the MPEG file being descrambled. Scrambled content is not viewable.
- **The Analog Protection System (APS)**. This system developed by Macrovision modifies NTSC/PAL. The resulting video signal can be displayed on televisions but cannot be recorded on VCR's. However, the data on a disk are not NTSC/PAL encoded and APS has to be applied after encoding in the DVD player. Some bits are consequently stored in the MPEG stream header and give the information of whether and how APS should be applied.
- **The Copy Generation Management System (CGMS)**. This is a pair of bits stored in the header of an MPEG stream encoding one of three possible rules for copying: *copy-always*, *copy-never* and *copy-once*. The copy-once case is included so that time-shifting is allowed i.e. a copy of broadcast media is made for later viewing.
- **5C**. A coalition of five companies designs this mechanism. It allows several compliant devices, connected to the same digital video bus e.g. IEEE1394 (firewire), to exchange keys in an authenticated manner so that encrypted data can be sent over the bus. Noncompliant devices do not have access to the keys and cannot decrypt the data.
- **Watermarking**. The main purpose of watermarking is to provide a more secure solution than storing bits in the MPEG stream header. In DVD,

digital watermarking is primarily intended for the CGMS bits and secondary for the APS bits.

- **Physical identifiers**. The idea is to design secure physical media identifiers in order to be able to distinguish between original media and copies.

Figure 4 shows how those mechanisms have been put together in the DVD so that copy protection is enforced. The additional performance brought by watermarking is emphasized by the dark walls.

Everything starts when Hollywood studios release a new copyrighted DVD with CGMS bits encoding the message *copy-never*. Both CSS keys are stored on the lead-in area of the DVD. This area is only read by compliant players. This prevents factory-pressed legal disks from being displayed by noncompliant players. Moreover bit-for-bit illegal copies will contain CSS scrambled content, but not the keys. As a result, such illegal copies cannot be displayed by any player, compliant or not. If the output signal given by compliant players is digital, CGMS bits prevent copying in the compliant world while 5C will avoid any communication with any noncompliant devices. However, to date, analog monitors are still widespread and even compliant players output an analog signal for compatibility. Since CGMS bits do not survive digital to analog conversion, watermarking is introduced in order to avoid copying in the compliant world. Unfortunately, in the noncompliant world, APS only disables copying of analog NTSC/PAL signals on VHS tapes. Disks without CSS or CGMS can then be easily generated e.g. thanks to a simple PC with a video capture card.

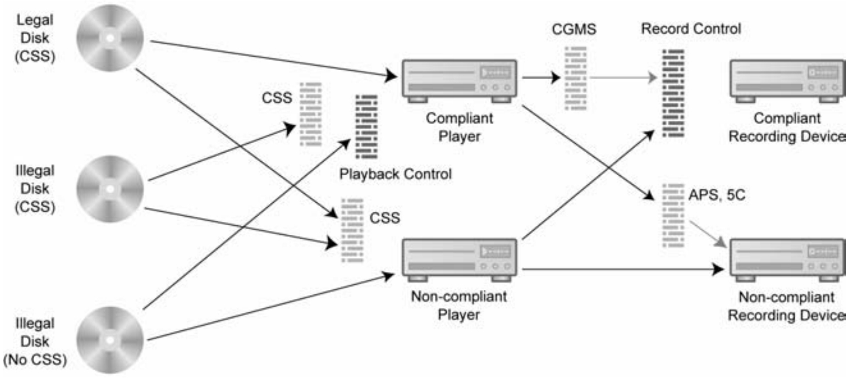

Figure 4. DVD copy-protection system

Now illegal disks containing unscrambled content without CSS or CGMS are available. They may have been generated as described previously. But they can also be generated directly from an original legal disk since CSS was cracked in 1999 [39]. The remaining CGMS bits can then be trivially stripped from the MPEG stream. Such illegal copies can of course be displayed by noncompliant players but watermarking has to be introduced in order to prevent those copies to enter the compliant world. Compliant players will detect the *copy-never* watermark embedded in *unscrambled* DVD-RAM and will refuse playback. The video signal given by a noncompliant player can be recorded by noncompliant recording devices. However watermarking prevents copying with compliant devices. The whole protection system results in two hermetically separated worlds. A consumer should have both types of players in order to display legal and illegal disks. The expense of such a strategy will help to “keep honest people honest”.

It is important for DVD recorders to support the *copy-once* case in order to allow time shifting. When the recorder detects the *copy-once* message, it should modify the stream so that the hidden message becomes *copy-never*. This can be easily done in the case of stored bits in the MPEG header but it is less straightforward when using watermarking. Two proposals are investigated. The first one consists in superimposing a second watermark when a *copy-once* watermark is detected. The two watermarks together will then encode the message *copy-never*. The second proposal avoids remarking and exploits the ticket concept [34]. The idea is to use two hidden signals: an embedded watermark W and a physical ticket T. There exists a relationship between the two signals which can be written $F^n(T)=W$, where $F(.)$ is a one way hash function and n is the number of allowed passages though compliant devices. The ticket is decremented each time the data go through a compliant player or recorder. In other terms, the ticket is modified according to the relation $T'=F(T)$. During playback, the ticket in transit can be embedded in MPEG *user_data* bits or in the blanking intervals of the NTSC/PAL standard. During recording, the ticket can be physically marked in the *wobble*[1] in the lead-in of optical disks.

3.2 BROADCAST MONITORING

Many valuable products are distributed over the television network. News items, such as those sold by companies like Reuters or Associated Press, can be worth over 100,000 USD. In France, during the final of the 2002 FIFA World Cup Korea Japan™, advertisers had to pay 100,000 Euros in order to broadcast a thirty seconds commercial break shot on television. The same commercial would even have been billed 220,000 Euros if the French national team had played during the final. Owners of copyrighted videos want to get their royalties each time their property is broadcasted. The whole television market is worth many billions of dollars and Intellectual Property Rights violations are likely to occur. As a result, a broadcast surveillance system has to be built in order to check all broadcasted channels. This will help verifying that content owners get paid correctly and that advertisers get what they have paid for. Such a mechanism will prevent confidence tricks such as the one discovered in Japan in 1997 when two TV stations were convicted of overbooking air time [27].

The most naive approach of broadcast monitoring consists of a pool of human observers watching the broadcasts and recording of whatever they see. However, this low-cost method is far from being optimal. Human employees are expensive and are not foolproof. As a result, research has been conducted in order to find a way of automating broadcast monitoring. The first approach, referred as *passive monitoring*, basically makes a computer simulate a human observer: it monitors the broadcasts and compares the received signals with a database of known videos. This approach is non intrusive and does not require cooperation from advertisers

[1] The wobble is a radial deviation of the position of pits and lands relative to the ideal spiral. Noncompliant recorders will not insert a ticket and the illegal disk will not enter the compliant world.

or broadcasters. However such a system has two major drawbacks. First, it relies on the comparison between received signals against a large database, which is non trivial in practice. Pertinent signatures, clearly identifying each video, have to be defined and an efficient search for nearest neighbours in a large database has to be designed. This results in a system that is not fully reliable. This may be accurate for acquiring competitive market research data i.e. when a company wants to know how much its competitors spend in advertising. On the contrary, a small error rate (5%) is dramatic for verification services because of the large amount of money at stake. The second con is that the reference database is likely to be large and the storage and management costs might become rapidly prohibitive.

In order to reach the accuracy required for verification services, a new kind of systems, referred as *active monitoring*, has been designed. The underlying idea is to transmit computer-recognizable identification information along with the data. Such identification information is straightforward to decode reliably and to interpret correctly. This approach is known to be simpler to implement than passive monitoring. First implementations of active monitoring placed the identification information in a separate area of the broadcast signal e.g. the Vertical Blanking Interval (VBI) of an analog NTSC/PAL video signal. However dissimulating identification data into other data is exactly the purpose of digital watermarking. Even if watermark embedding is more complicated than storing information in some unused part of a video stream, digital watermarking can be considered as a robust way to implement active monitoring. The European project VIVA (Visual Identity Verification Auditor) proved the feasibility of such a system [14]. The participants used a real-time watermarking scheme which provides active monitoring services over a satellite link. The complexity of the detection algorithm is moderate enough to allow simultaneous monitoring of many channels.

3.3 FINGERPRINTING

The explosion of the Internet has created a new way of acquiring copyrighted content. When a user wants to obtain a new video clip or a new movie, the simplest strategy is to log on Internet and to use one of the popular peer-to-peer systems e.g. Napster, Gnutella, KaZaA, Morpheus. Multimedia digital contents, stored throughout the world on thousands of computers logged on at the same moment, will instantly get accessible. As a result, European engineering students often download and watch the most recent Hollywood films a long time before they are released in their own country. The situation is even worse in audio with the exchange of MP3[1] files. As a result, copyright owners lose a large amount of royalties [32]. Legal action has been taken to ban such distributed systems but, when Napster has been sentenced guilty, two other systems appeared. The basic problem does not come from peer-to-peer systems. It would be a great tool if only legal data was transiting on such distributed networks. The problem is that a *traitor* has made available copyrighted material without any kind of permission. The basic idea would consequently be to be able to identify the traitor when an illegal copy is found in order to sue him in court. This can be done by embedding an indelible and invisible watermark identifying the customer.

[1] The MPEG-1 audio layer 3 (MP3) is a popular audio format for transmitting audio files across the Internet.

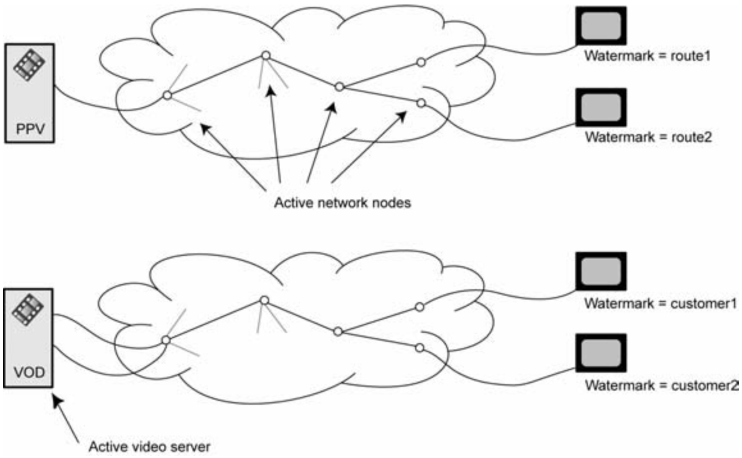

Figure 5. Alternative watermarking strategies for video streaming.

In a near future, the way people are looking at TV will be significantly be modified. Video streaming is indeed likely to become more and more widespread. It is consequently necessary to find a way of protecting digital video content and digital watermarking seems to be a potential candidate [33]. Pay-Per-View (PPV) and Video-On-Demand (VOD) are two real-life applications of video streaming. In both of them, digital watermarking can be used in order to enforce a fingerprinting policy. The customer ID is embedded into the delivered video data in order to trace back any user breaking his/her licence agreement. The main difference resides in the watermarking strategy as depicted in Figure 5. Embedding the watermark on the customer side has been suggested [20] but it should be avoided if possible in order to prevent reverse engineering. In a PPV environment, a video server multicasts some videos and customers have only to connect to the server in order to obtain the video. The video server is passive. At a given moment, it delivers the same video stream to multiple users. In order to enforce fingerprinting, a proposed method [7] is to have each network element (router, node or whatever) embed a piece of watermark as the video stream is relayed. The resulting watermark will contain a trace of the route followed by the video stream. Such a strategy requires support from network providers, who might not be forthcoming about it. In a VOD framework, the video server is active. It receives a request from a customer and sends the requested video. It is a multi-unicast strategy. This time, the video server can insert a watermark identifying the customer since each connection is dedicated to only one customer. The main challenge is then to scale the system to many users.

Another fingerprinting application has been considered with the apparition of a new kind of piracy. Nowadays illegal copying of brand new movies projected onto cinema screen by means of a handheld video camera has become a common practice. The most memorable example is surely when, one week after its US release, the very anticipated “Starwars Episode I: The Phantom Menace” was available on the Internet in a low quality version, with visible head shadows of audience members. Although the quality of such copies is usually very low, their economical impact can be enormous. Moreover, the upcoming digital cinema format to be introduced in theatres raises some concern. With higher visual quality, the threat becomes larger and Hollywood studios want to oblige cinema owners to prevent the presence of video cameras in their premises. Once again, digital watermarking could provide a solution [21]. A watermark can be embedded during show time identifying the cinema, the presentation date and time. If an illegal copy created with a video camera is found, the watermark is extracted and the cinema to blame is identified. After many blames, the cinema is sanctioned with a ban on the availability of content.

3.4 VIDEO AUTHENTICATION

Large amounts of video data are distributed throughout the Internet every day. More and more video cameras are installed in public facilities for surveillance purpose. However, popular video editing softwares permit today to easily tamper with video content, as shown in Figure 6, and video content is no more reliable. For example, in some countries, a video shot from a surveillance camera cannot be used as a piece of evidence in a courtroom because it is not considered trustworthy enough. When someone is emailed a somewhat unusual video, it is quite impossible to determine if it is an original or a hoax. Authentication techniques are consequently needed in order to ensure authenticity of video content. Methods have to be designed for verifying the originality of video content and preventing forgery. When a customer purchases video content via electronic commerce, he wants to be sure that it comes from the alleged producer and that no one has tampered with the content. The very first research efforts for data authentication used cryptography. The major drawback of such an approach is that it provides a *complete verification*. In other terms, the data is considered as untouchable and the data for authentication has to be exactly the same one as the original one. But this strong constraint might be too restricting. One might prefer to allow some distortions on the digital data if the original content has not been significantly modified. This is typically the case in wireless environment where some noise is added to the data. This approach is referred as *content verification*.

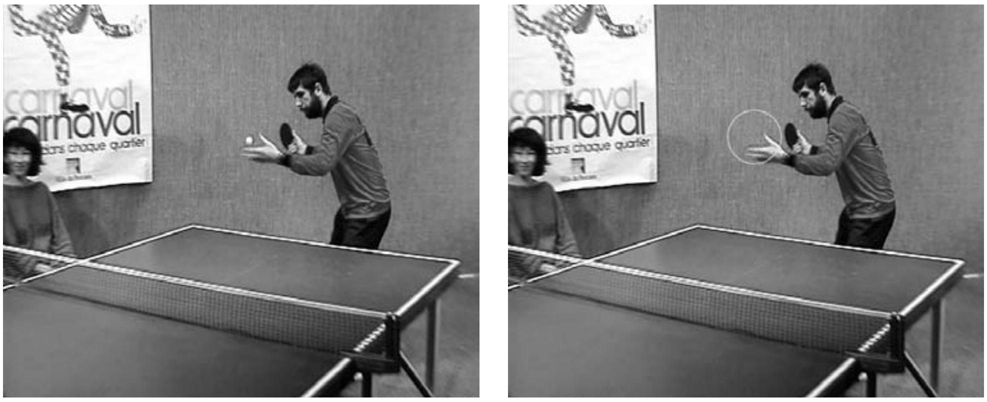

Figure 6. Original and tampered video scenes

Researchers have investigated the use of digital watermarking in order to verify the integrity of digital video content. A basic approach consists in regularly embedding an incremental timestamp in the frames of the video [37]. As a result, frame cuts, foreign frame insertion, frame swapping, and frame rate alteration can be easily detected. This approach is very efficient for detecting temporal alteration of the video stream. However, it might fail in detecting alterations of the content itself e.g. a character is completely removed from a movie. Investigations have consequently been conducted in order to prevent modifications of the content of the video itself. One proposal [17] embeds the edge map of each frame in the video stream. During the verification process, if the video content has been modified, there will be a mismatch between the extracted edge map from the verified video and the watermarked edge map. The detector will consequently report content tampering. Another proposal exploits the idea that a movie is made up of one audio and one video stream and that both need to be protected against unauthorised tampering. The fundamental idea is then to combine video and audio watermarking [18] in order to obtain an efficient authenticating system. Features of both streams are embedded one into another. Modification from either the sound track, or the video track, is immediately spotted by the detector, since the extracted and watermarked features will differ.

3.5 COPYRIGHT PROTECTION

Copyright protection is historically the very first targeted applications for digital watermarking. The underlying strategy consists in embedding a watermark, identifying the copyright owner, in digital multimedia data. If an illegal copy is found, the copyright owner can prove his/her paternity thanks to the embedded watermark and can sue the illegal user in court. This perfect scenario is however likely to be disturbed by malicious users in the real world [10]. If an attacker adds a second watermark into a video clip, both the original owner and the attacker can claim ownership and therefore defeat the purpose of using watermarking. Using the original video clip during the verification procedure happens to prevent the multiple ownership problems in some cases. However, this problem still holds if the watermarking algorithm is invertible because it allows the attacker to produce his/her own counterfeited original video clip. In this case, both the original owner and the attacker have an original video clip which contains the watermark of the other one. As a result, nobody can claim ownership! This situation is referred as the *deadlock* problem in the watermarking community. Watermark algorithms are consequently required to be non-invertible in order to provide copyright protection services and they are often backed up by an elaborated protocol with a trusted third party. Copyright protection has been investigated for video watermarking [42] even if this not the most targeted application.

Instead of protecting the whole video stream, copyright owners might rather want to protect only a part of the video content. The commercial value in a video is indeed often concentrated in a small number of video objects e.g. the face of an actor. Moreover, future video formats will distinguish the different objects in a video. This will be the case with the upcoming MPEG-4 format. Recent research has consequently investigated digital watermarking of video objects [41]. Watermarking video objects prevents unauthorised reuse in other video clips. However video objects are likely to be submitted to various video editing such as scaling, rotation, shifting and flipping. As a result, special care must be taken regarding the resilience of the watermark against such processings. This can be quite easily obtained thanks to a geometrical normalisation [4], according to the moments and axes of the video object, prior to embedding and extraction.

3.6 ENHANCED VIDEO CODING

The attentive reader may have noticed that video watermarking and video coding are two conflicting technologies. A perfect video codec should remove any extra redundant information. In other terms, two visually similar videos should have the same compressed representation. If one day, such an optimal video codec is designed, then video watermarking will disappear since unwatermarked and watermarked data would have the same compressed representation. Digital watermarking can be consequently seen as the exploitation of the features of the compression algorithms in order to hide information. However recent research has shown that digital watermarking can benefit to the coding community. The video coding process can be sequenced in two steps. During *source coding*, any redundant information is removed in order to obtain the most possible compressed representation of the data while keeping its original visual quality. This compressed representation is then submitted to *channel coding*, where extra redundant information is added for error correction. Channel coding is mandatory since errors are likely to occur during the transmission, e.g. in a wireless environment. Digital watermarking can be introduced as an alternative solution for introducing error correcting information after source coding, without inducing any overhead [3]. Experiments have demonstrated the feasibility of such an approach and results are even reported showing that digital watermarking can have better performances than traditional error correction mechanisms [45].

Embedding useful data directly into the video stream can spare much storage space. A typical video stream is made up of two different parallel streams: the audio and video streams. Those two streams need to be synchronised during playback for pleasant viewing, which is difficult to maintain during cropping operations. Hiding the audio stream into the video one [38] will implicitly provide efficient and robust synchronisation, while significantly reducing the required storage need or available bandwidth. In the same fashion, the actual Picture-in-Picture system can be improved by hiding a video stream into another one [48]. This technology, present in many television sets, uses separate data streams in order to superimpose a small video window over the full-size video displayed on the television set. Digital watermarking allows embedding the secondary video stream into the carrier one. During playback, the watermark is extracted and the embedded video is displayed in a window within the host video. With such an approach, only one stream needs to be transmitted. This approach can be extended so that a user can switch to the *PG* version of an *R* rated movie, with alternative dialogs and scenes replacing inappropriate content.

4. CHALLENGES IN VIDEO WATERMARKING

Digital watermarking has focused on still images for a long time but nowadays this trend seems to vanish. More and more watermarking algorithms are proposed for other multimedia data and in particular for video content. However, even if watermarking still images and video is a similar problem, it is not identical. New problems, new challenges show up and have to be addressed. This section points out three major challenges for digital video watermarking. First, there are many non-hostile video processings, which are likely to alter the watermark signal. Second, resilience to collusion is much more critical in the context of video. Third, real-time is often a requirement for digital video watermarking.

4.1 VARIOUS NON-HOSTILE VIDEO PROCESSINGS

Robustness of digital watermarking has always been evaluated via the survival of the embedded watermark after attacks. Benchmarking tools have even been developed in order to automate this process [1]. In the context of video, the possibilities of attacking the video are multiplied. Many different non-hostile video processings are indeed available. Non-hostile refers to the fact that even content provider are likely to process a bit their digital data in order to manage efficiently their resources.

Photometric attacks

This category gathers all the attacks which modify the pixel values in the frames. Those modifications can be due to a wide range of video processings. Data transmission is likely to introduce some noise for example. Similarly, digital to analog and analog to digital conversions introduce some distortions in the video signal. Another common processing is to perform a gamma correction in order to increase the contrast. In order to reduce the storage needs, content owners often transcode, i.e. re-encode with a different compression ratio, their digital data. The induced loss of information is then susceptible to alter the performances of the watermarking algorithm. In the same fashion, customers are likely to convert their videos from a standard video format such as MPEG-1, MPEG-2 or MPEG-4 to a *popular* format e.g. DivX. Here again, the watermark signal is bound to undergo some kind of interferences. Spatial filtering inside each frame is often used to restore a low-quality video. Inter-frames filtering, i.e. filtering between adjacent frames of the video, has to be considered too. Finally, chrominance resampling (4:4:4, 4:2:2, 4:2:0) is a commonly used processing to reduce storage needs.

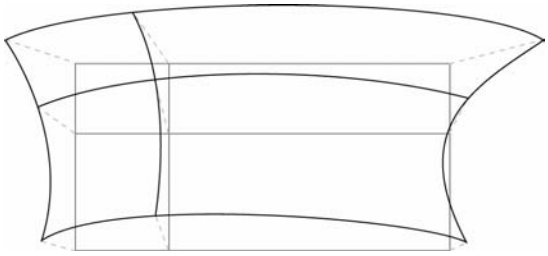

Figure 7. Example of distortion created by a handheld camera (exaggerated)

Spatial desynchronisation

Many watermarking algorithms rely on an implicit spatial synchronisation between the embedder and the detector. A pixel at a given location in the frame is assumed to be associated with a given bit of the watermark. However, many non-hostile video processings introduce spatial desynchronisation which may result in a drastic loss of performance of a watermarking scheme. The most common examples are changes across display formats (4/3, 16/9 and 2.11/1) and changes of spatial resolution (NTSC, PAL, SECAM and usual movies standards). Alternatively the pixel position is susceptible to jitter. In particular, positional jitter occurs for video over poor analog links e.g. broadcasting in a wireless environment. In the digital cinema context, distortions brought by the handheld camera can be considered as non-hostile since the purpose of the camera is not explicitly to remove the embedded watermark. It has been shown that the handheld camera attack can be separated into two geometrical distortions [13]: a bilinear transform, due to the misalignment between the camera and the cinema screen, and a curved transform, because of the lens deformations. This results in a curved –bilinear transform depicted in Figure 7 which can be modelled by twelve parameters.

Temporal desynchronisation

Similarly temporal desynchronisation may affect the watermark signal. For example, if the secret key for embedding is different for each frame, simple frame rate modification would make the detection algorithm fail. Since changing frame rate is a quite common processing, watermarks should be designed so that they survive such an operation.

Video editing

The very last kind of non-hostile attacks gathers all the operation that a video editor may perform. Cut-and-splice and cut-insert-splice are two very common processings used during video editing. Cut-insert-splice is basically what happens when a commercial is inserted in the middle of a movie. Moreover, transition effects, like fade-and-dissolve or wipe-and-matte, can be used in order to smooth the transition between to scenes of the video. Such kind of editing can be seen as temporal editing in contrast to spatial editing. Spatial editing refers to the addition of a visual content in each frame of the video stream. This includes for example graphic overlay, e.g. logos or subtitles insertion, and video stream superimposition, like in the Picture-in-Picture technology. The detector sees such operation as a cropping of some part of the watermark. Such a severe attack is susceptible to induce a high degradation of the detection performances.

Table 3. Examples of non-hostile video processings

Photometric	- Noise addition, DA/AD conversion - Gamma correction - Transcoding and video format conversion - Intra and inter-frames filtering - Chrominance sampling (4:4:4, 4:2:2, 4:2:0)
Spatial desynchronisation	- Changes across display formats (4/3, 16/9, 2.11/1) - Changes of spatial resolution (NTSC, PAL, SECAM) - Positional jitter - Handheld camera attack
Temporal desynchronisation	- Changes of frame rate
Video editing	- Cut-and-splice and cut-insert-splice - Fade-and-dissolve and wipe-and-matte - Graphic overlay (subtitles, logo) - Picture-in-Picture

There are many various attacks to be considered as reminded in Table 3 and it may be useful to insert countermeasures [12] in the video stream in order to cop with the distortions introduced by such video processings. Moreover, the reader should be aware that many other hostile attacks are likely to occur in the real world. Indeed, it is relatively easy today to process a whole movie thanks to the powerful available personal computers. It is virtually possible to do whatever transformation on a video stream. For example, for still images, Stirmark introduces random local geometric distortions which succeed in trapping the synchronisation of the detector. This software has been optimised for still images and, when used on each frame of the video stream, visible artefacts can be spotted when moving objects go through the fixed geometric distortion. However future versions of Stirmark will surely address this visibility issue.

4.2 RESILIENCE AGAINST COLLUSION

Collusion is a problem that has already been pointed out for still images some time ago. It refers to a set of malicious users who merge their knowledge, i.e. different watermarked data, in order to produce illegal content, i.e. unwatermarked data. Such collusion is successful in two different distinct cases.

- **Collusion type I:** The *same watermark* is embedded into different copies of *different data*. The collusion can estimate[1] the watermark from each watermarked data and obtain a refined estimate of the watermark by linear combination, e.g. the average, of the individual estimations. Having a good estimate of the watermark permits to obtain unwatermarked data with a simple subtraction with the watermarked one.
- **Collusion type II:** *Different watermarks* are embedded into different copies of the *same data*. The collusion only has to make a linear combination of the different watermarked data, e.g. the average, to produce unwatermarked data. Indeed, generally, averaging different watermarks converges toward zero.

Collusion is a very important issue in the context of digital video since there are twice more opportunities to design a collusion than with still images. When video is considered, the origin of the collusion can be twofold.

- **Inter-videos collusion:** This is the initial origin considered for still images. A set of users have a watermarked version of a video which they gather in order to produce unwatermarked video content. In the context of copyright protection, the same watermark is embedded in different videos and collusion type I is possible. Alternatively, in a fingerprinting application, the watermark will be different for each user and collusion type II can be considered. Inter-videos collusion requires different watermarked videos to produce unwatermarked video content.
- **Intra-video collusion:** This is a video-specific origin. As will be detailed later, many watermarking algorithms consider a video as a succession of still images. Watermarking video comes then down to watermarking series of still images. Unfortunately this opens new opportunities for collusion. If the same watermark is inserted in each frame, collusion type I can be enforced since different images can be obtained from moving scenes. On the other hand, if alternative watermarks are embedded in each frame, collusion type II becomes a

[1] The watermark is often considered as noise addition. A simple estimation consequently consists in computing the difference between the watermarked data and the low-pass filtered version of it.

danger in static scenes since they produce similar images. As a result, the watermarked video alone permits to remove the watermark from the video stream.

Even if collusion is not really of interest depending on the targeted application e.g. broadcast monitoring, it often raises much concern in digital video watermarking. It gives indeed opportunities for forgery if the watermarking algorithm is weak against intra-video collusion.

The reader will have understood that the main danger is intra-frame collusion i.e. when a watermarked video alone is enough to remove the watermark from the video. It has been shown that both strategies *always insert the same watermark in each frame* and *always insert a different watermark in each frame* make collusion attacks conceivable. As a result, an alternative strategy has to be found. A basic rule [47] has been enounced so that intra-video collusion is prevented, or at least made more difficult. The watermarks inserted into two different frames of a video should be as similar, in terms of correlation, as the two frames are similar. In other terms, if two frames look like quite the same, the embedded watermarks should be highly correlated. On the contrary, if two frames are really different, the watermark inserted into those frames should be unalike. This rule is quite straightforward when regarding attentively the definition of the two types of collusion. This can be seen as a form of informed watermarking since this rule implies a dependency between the watermark and the host frame content. A relatively simple implementation of this approach can be done by embedding a spatially localised watermark according to the content of each frame of the video [46]. A small watermark pattern can be embedded in some key locations of each frame, e.g. salient points. During the extraction process, the detector can easily detect the position of the salient points and look for the presence or the absence of a watermark.

The problem of inter-video collusion still holds. Concerning collusion type I, this issue can be prevented by inserting a Trusted Third Party (TTP) which gives the message to be embedded. This message is often a function of the encrypted message that the copyright owner wants to hide and a hash of the host data. Different videos give different messages to be hidden and consequently different embedded watermarks. The TTP also acts as a repository. When an illegal copy is found, the copyright owner extracts the embedded message and transmits it to the TTP, which in turn gives the associated original encrypted message. If the copyright owner can successfully decrypt it, he can claim ownership. Regarding collusion type II, results obtained for still images can easily be extended to digital video. The problem arises when a coalition of malicious users, having each one a copy of the same data but with a different embedded watermark, colludes in order to produce illegal unwatermarked data. They compare their watermarked data, spot the locations where the different versions differ and modify the data in those locations. A traditional countermeasure [6] consists in designing the set of distributed watermarks so that a coalition, gathering at most c users, will not succeed in removing the whole watermark signal. It should be noted that c is generally very small in comparison with the total number n of users. Moreover the set of watermarks is built in such a way that no coalition of users can produce a document which will make an innocent user, i.e. not in the illegal coalition, be framed. In other terms, colluding creates still watermarked video content and the remaining watermark clearly identifies the malicious colluding users, without ever accusing any innocent customer. Implementations of such set of watermark have already been proposed for still images which are based on the projective geometry [16] or the theory of combinatorial designs [51].

4.3 REAL-TIME WATERMARKING

Real-time can be an additional specification for video watermarking. It was not a real concern with still images. When a person wants to embed a watermark or to check the presence of a watermark in an image, a few seconds is an acceptable delay. However, such a delay is unrealistic in the context of the video. Frames are indeed sent at a fairly high rate, typically 25 frames per second, to obtain a smooth video stream. At least the embedder or the detector, and even sometimes both of them, should be able to handle such a rate. In the context of broadcast monitoring, the detector should be able to detect an embedded watermark in real-time. In a VOD environment, the video server should be able to insert the watermark identifying the customer at the same rate that the video is streamed. In order to meet the real-time requirement, the complexity of the watermarking algorithm should obviously be as low as possible. Moreover, if the watermark can be inserted directly into the compressed stream, this will prevent full decompression and recompression and consequently, it will reduce computational needs. This philosophy has led to the design of very simple watermarking schemes. Exploiting the very specific part of a video compression standard can lead to very efficient algorithms. An MPEG encoded video stream basically consists of a succession of Variable Length Code (VLC). A watermark can consequently be embedded in the stream by modifying those VLC code words [31]. The MPEG standard uses indeed similar VLC code words i.e. with the same run length, the same VLC size and a quantized level difference of one. Such VLC code words can be used alternatively in order to hide a bit.

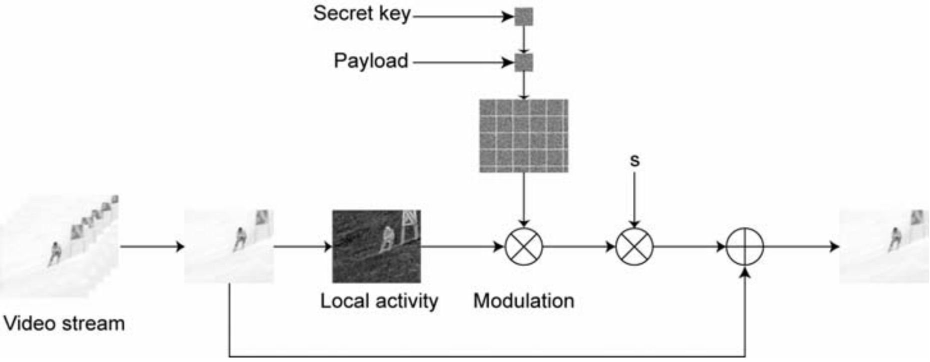

Figure 8. Description of JAWS embedding

Just Another Watermarking System (JAWS)

When considering real-time, the watermarking algorithm designed by Philips Research is often considered as a reference. The JAWS algorithm was originally designed for broadcast monitoring and is actually one of the leading candidates for watermarking in DVD. The real-time requirement is met by using simple operations at video rate and only a few complex ones at a much lower rate [25].

The embedding process is depicted in Figure 8. First of all, an *M*x*M* normally distributed reference pattern p_r is generated with a secret key. In a second step, a reference watermark w_r is created according to the following equation:

$$w_r = p_r - shift(p_r, message) \qquad (1)$$

where the *shift(.)* function returns a cyclically shifted version of the reference pattern p_r. In JAWS, the message is completely encoded by the shift between the two reference patterns. This reference watermark is then tiled, possibly with truncation, to obtain the full-size watermark *w*. For each frame, this watermark is then perceptually shaped so that the watermark insertion remains imperceptible. Each element *i* of the watermark is scaled by the local activity $\lambda(i)$ of the frame, given by Laplacian filtering. The flatter the region is, the lower the local activity is. This is coherent with the fact that the human eye is more sensitive to noise addition in flat regions of an image. Finally, the watermark is scaled by a global factor *s* and added to the frame *F* in order to obtain the watermarked frame F_w. As a result, the overall embedding process can be expressed as:

$$F_w(i) = F(i) + s.\lambda(i).w(i) \qquad (2)$$

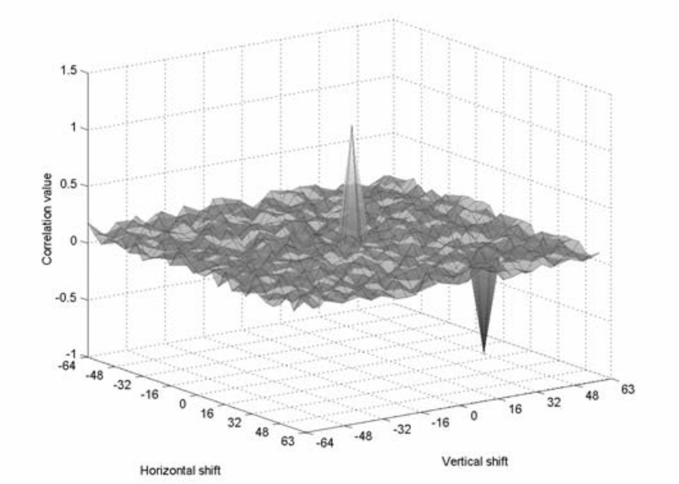

Figure 9. Example of SPOMF detection

On the detector side, the incoming frames are folded, summed and stored in an *M*x*M* buffer *B*. The detectors look then for all the occurrences of the reference pattern p_r in the buffer with a two dimensional cyclic convolution. Since such an operation is most efficiently computed in the frequency domain, this leads to Symmetrical Phase Only Matched Filtering (SPOMF) detection which is given by the following equation:

$$SPOMF(B, p_r) = IFFT\left[\varphi(FFT(B)).\varphi\left(FFT(p_r)^*\right)\right] \quad with\ \varphi(x) = \begin{cases} x/|x| & if\ x \neq 0 \\ 1 & if\ x = 0 \end{cases} \quad (3)$$

where *FFT(.)* and *IFFT(.)* are respectively the forward and inverse Fourier transforms and x^* denotes the complex conjugation. Figure 9 shows the result of such a detection. Two peaks can be isolated which correspond to the two occurrences of p_r in w_r. The peaks are oriented accordingly to the sign before their associated occurrence of p_r in Equation (1). Because of possible positional jitter, all the relative positions between the peaks cannot be used and relative positions are forced to be multiple of a grid size *G*. Once the detector has extracted the peaks, the hidden payload can be easily retrieved. The attentive reader would have noticed that this scheme is inherently shift invariant since a shifting operation does not modify the relative position of the peaks. Significant improvements have been added to this scheme afterwards. For example, shift invariance has been further exploited in order to increase the payload [35] and simple modifications permitted to obtain scale invariance [50].

5. THE MAJOR TRENDS IN VIDEO WATERMARKING

Digital watermarking for video is a fairly new area of research which basically benefits from the results for still images. Many algorithms have been proposed in the scientific literature and three major trends can be isolated. The most simple and straightforward approach is to consider a video as a succession of still images and to reuse an existing watermarking scheme for still images. Another point of view considers and exploits the additional temporal dimension in order to design new robust video watermarking algorithms. The last trend basically considers a video stream as some data compressed according to a specific video compression standard and the characteristics of such a standard can be used to obtain an efficient watermarking scheme. Each of those approaches has its pros and cons as detailed in Table 4.

Table 4. Pros and cons of the different approaches for video watermarking.

	Pros	**Cons**
Adaptation image → video	Inherit from all the results for still images	Computationally intensive
Temporal dimension	Video-driven algorithms which often permit higher robustness	Can be computationally intensive
Compression standard	Simple algorithms which make real-time achievable	Watermark may be inherently tied to the video format

5.1 FROM STILL IMAGE TO VIDEO WATERMARKING

In its very first years, digital watermarking has been extensively investigated for still images. Many interesting results and algorithms were found and when new areas, such as video, were researched, the basic concern was to try to reuse the previously found results. As a result, the watermarking community first considered the video as a succession of still images and adapted existing watermarking schemes for still images to the video. Exactly the same phenomenon occurred when the coding community switched from image coding to video coding. The first proposed algorithm for video coding was indeed Moving JPEG (M-JPEG), which simply compresses each frame of the video with the image compression standard JPEG. The simplest way of extending a watermarking scheme for still images is to embed the same watermark in the frames of the video at a regular rate. On the detector side, the presence of the watermark is checked in every frame. If the video has been watermarked, a regular pulse should be observed in the response of the detector [2]. However, such a scheme has no payload. The detector only tells if a given watermark is present or not but it does not extract any hidden message. On the other hand, the host data is much larger in size than a single still image. Since one should be able to hide more bits in a larger host signal, high payload watermarks for video could be expected. This can be easily done by embedding an independent multi-bits watermark in each frame of the video [15]. However one should be aware that this gain in payload is counterbalanced by a loss of robustness.

Differential Energy Watermarks (DEW)

The DEW method was initially designed for still images and has been extended to video by watermarking the I-frames of an MPEG stream [31]. It is based on selectively discarding high frequency DCT coefficients in the compressed data stream. The embedding process is depicted in Figure 10. The 8x8 pixels blocks of the video frame are first pseudo randomly shuffled. This operation forms the secret key of the algorithm and it spatially randomizes the statistics of pixel blocks i.e. it breaks the correlation between neighbouring blocks. The obtained shuffled frame is then split into *n* 8x8 blocks. In Figure 10, *n* is equal to 16. One bit is embedded into each one of those blocks by introducing an energy difference between the high frequency DCT-coefficients of the top half of the block (region A) and the bottom half (region B). This is the reason why this technique is called a differential energy watermark.

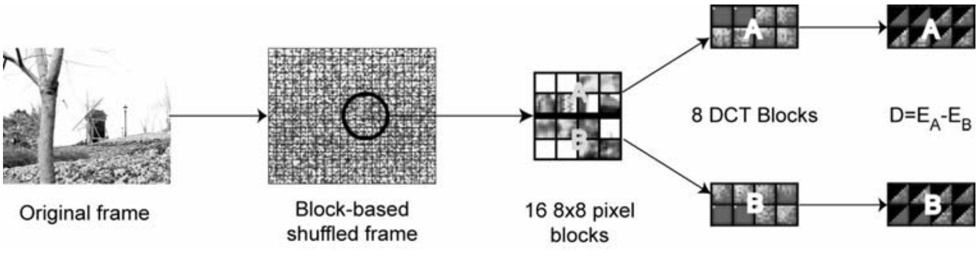

Figure 10. Description of DEW embedding

In order to introduce an energy difference, the block DCT is computed for each *n* 8x8 block and the DCT-coefficients are prequantized with quality factor Q_{jpeg} using the standard JPEG quantization procedure. The obtained coefficients are then separated in two halves and the high frequency energy for each region is computed according to the following equation:

$$E\left(c,n,Q_{jpeg}\right)=\sum_{b=0}^{n/2-1}\sum_{i\in S(c)}\left(\left[\theta_{i,b}\right]_{Q_{jpeg}}\right)^{2} \quad with\ S(c)=\left\{i\in\{0,63\}\mid(i>c)\right\} \tag{4}$$

where $\theta_{i,b}$ is the DCT coefficient with index i in the zig-zag order in the b^{th} DCT block, [.] indicates the prequantization with quality factor Q_{jpeg} and c is a given cut-off index which was fixed to 27 in Figure 10. The value of the embedded bit is encoded as the sign of the energy difference $D=E_A-E_B$ between the two regions A and B. All the energy after the cut-off index c in either region A or region B is eliminated by setting the corresponding DCT coefficients to zero to obtain the appropriate sign for the difference D. It should be noted that this can be easily done directly in the compressed domain by shifting the End Of Block (EOB) marker of the corresponding 8x8 DCT blocks toward the DC-coefficient up to the cut-off index. Finally, the inverse block DCT is computed and the shuffling is inversed in order to obtain the watermarked frame. On the detector side, the energy difference is computed and the embedded bit is determined according to the sign of the difference D. This algorithm has been further improved to adapt the cut-off index c to the frequency content of the considered n 8x8 block and so that the energy difference D is greater than a given threshold D_{target} [30].

5.2 INTEGRATION OF THE TEMPORAL DIMENSION

The main drawback of considering a video as a succession of independent still images is that it does not satisfactorily take into account the new temporal dimension. The coding community has made a big step forward when they decided to incorporate the temporal dimension in their coding schemes and it is quite sure that it is the advantage of the watermarking community to investigate such a path. Many researchers have investigated how to reduce the visual impact of the watermark for still image by considering the properties of the Human Visual System (HVS) such as frequency masking, luminance masking and contrast masking. Such studies can be easily exported to video with a straightforward frame-per-frame adaptation. However, the obtained watermark is not optimal in terms of visibility since it does not consider the temporal sensitivity of the human eye. Motion is indeed a very specific feature of the video and new video-driven perceptual measures need to be designed in order to be exploited in digital watermarking [28]. This simple example shows that the temporal dimension is a crucial point in video and that it should be taken into account to design efficient algorithms.

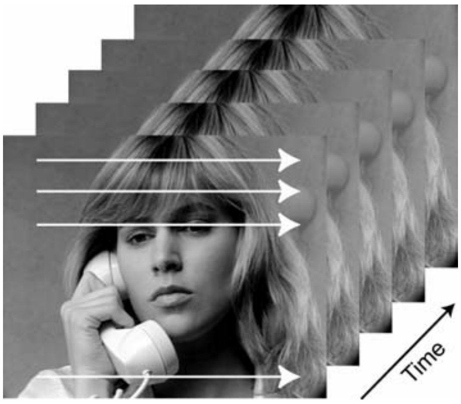

Figure 11. Line scan of a video stream

Spread-Spectrum (SS)

One of the pioneer works in video watermarking considers the video signal as a one dimensional signal [22]. Such a signal is acquired by a simple line-scanning as shown in Figure 11. Let the sequence $a(j)\in\{-1,1\}$ represents the watermark bits to be embedded. This sequence is spread by a chip-rate *cr* according to the following equation:

$$b(i) = a(j), \quad j.cr \le i < (j+1).cr, \quad i \in N \tag{5}$$

The spreading operation permits to add redundancy by embedding one bit of information into *cr* samples of the video signal. The obtained sequence *b(i)* is then amplified locally by an adjustable factor $\lambda(i)\geq 0$ and modulated by a pseudo-random binary sequence $p(i)\in\{-1,1\}$. Finally, the spread spectrum watermark *w(i)* is added to the line-scanned video signal *v(i)*, which gives the watermarked video signal $v_w(i)$. The overall embedding process is consequently described by the following equation:

$$v_w(i) = v(i) + w(i) = v(i) + \lambda(i).b(i).p(i), \quad i \in N \tag{6}$$

The adjustable factor $\lambda(i)$ may be tuned according to local properties of the video signal, e.g. spatial and temporal masking of he HVS, or kept constant depending on the targeted application.

On the detector side, recovery is easily accomplished with a simple correlation. However, in order to reduce cross-talk between watermark and video signals, the watermarked video sequence is high-pass filtered, yielding a filtered watermarked video signal $\underline{v_w(i)}$, so that major components of the video signal itself are isolated and removed. The second step is demodulation. The filtered watermarked video signal is multiplied by the pseudo-random noise *p(i)* used for embedding and summed over the window for each embedded bit. The correlation sum *s(j)* for the j^{th} bit is given by the following equation:

$$s(j) = \sum_{i=j.cr}^{(j+1).cr-1} p(i).\underline{v_w(i)} = \sum_{i=j.cr}^{(j+1).cr-1} p(i).\underline{v(i)} + \sum_{i=j.cr}^{(j+1).cr-1} p(i).\underline{\lambda(i).b(i).p(i)} = \Sigma_1 + \Sigma_2 \tag{7}$$

The correlation consists of two terms Σ_1 and Σ_2. The main purpose of filtering was to leave Σ_2 untouched while reducing Σ_1 down to 0. As a result, the correlation sum becomes:

$$s(j) \approx \Sigma_2 \approx \sum_{i=j.cr}^{(j+1).cr-1} p(i)^2.\lambda(i).b(i) = a(j).cr.mean(\lambda(i)) \tag{8}$$

The hidden bit is then directly given by the sign of *s(j)*. This pioneer method offers a very flexible framework, which can be used as a basic root of a more elaborate video watermarking scheme.

Other approaches have been investigated to integrate the temporal dimension. Temporal wavelet decomposition can be used for example in order to separate static and dynamic components of the video [49]. A watermark is then embedded in each component to protect them separately. The video signal can also be seen as a three dimensional signal. This point of view has already been considered in the coding community and can be extended to video watermarking. 3D DFT can be used as an alternative representation of the video signal [11]. The HVS is considered on one hand to define an embedding area which will not result in a visible watermark. On the other hand, the obtained embedding area is modified so that it becomes immune to MPEG compression. Considering video as a three dimensional signal may be inaccurate. The three considered dimensions are indeed not homogeneous: there are two *spatial* dimensions and one *temporal* one. This consideration and the computational cost may have hampered further work in this direction. However this approach remains pertinent in some specific cases. In medical imaging for example, different slices of a scanner can be seen as different frames of a video. In this case, the three dimensions are homogeneous and a 3D-transform can be used.

5.3 EXPLOITING THE VIDEO COMPRESSION FORMATS

The last trend considers the video data as some data compressed with a video specific compression standard. Indeed, most of the time, a video is stored in a compressed version in order to spare some storage space. As a result, watermarking methods have been designed, which embed the watermark directly into the compressed video stream. The first algorithm presented in Section 4.3 is a very good example. It exploits a very specific part of the video compression standard (run length coding) in order to hide some information.

Watermarking in the compressed stream can be seen as a form of video editing in the compressed domain [36]. Such editing is not trivial in practice and new issues are raised. The previously seen SS algorithm has been adapted so that the watermark can be directly inserted in the non-zero DCT coefficients of an MPEG video stream [22]. The first concern was to ensure that the watermarking embedding process would not increase the output bit-rate. Nothing ensures indeed that a watermarked DCT-coefficient will be VLC-encoded with the same number of bits than when it was unwatermarked. A straightforward strategy consists then to watermark only the DCT coefficients which do not require more bits to be VLC encoded. The second issue was to prevent the introduced distortion with the watermark to propagate from one frame to another one. The MPEG standard relies indeed on motion prediction and any distortion is likely to be propagated to neighbour frames. Since the accumulation of such propagating signals may result in a poor quality video, a drift compensation signal can be added if necessary. In this case, motion compensation can be seen as a constraint. However it could also be exploited so that the motion vectors of the MPEG stream carry the hidden watermark [24]. The components of the motion vector can be quantised according to a rule which depends on the bit to be hidden. For example, the horizontal component of a motion vector can be quantized to an even value if the bit to be hidden is equal to 0 and to an odd value otherwise.

All the frames of an MPEG coded video are not encoded in the same way. The intra-coded (I) frames are basically compressed with the JPEG image compression standard while the inter-coded (B and P) frames are predicted from other frames of the video. As a result, alternative watermarking strategies can be used depending on the type of the frame to be watermarked [23]. Embedding the watermark directly in the compressed video stream often allows real-time processing of the video. However the counterpart is that the watermark is inherently tied to a video compression standard and may not survive video format conversion.

6. CONCLUSION

Digital watermarking has recently been extended from still images to video content. Further research in this area is strongly motivated by an increasing need from the copyright owners to reliably protect their rights. Because of the large economic stakes, digital watermarking is promised to a great future. New applications are likely to emerge and may combine existing approaches. For example, a watermark can be separated into two parts: one for copyright protection and the other for customer fingerprinting. However many challenges have to be taken up. Robustness has to be considered attentively. There are indeed many non-hostile video processings which might alter the watermark signal. It might not even be possible to be immune against all those attacks and detailed constraints has to be defined according to the targeted application. Since collusion is far more critical in the context of video, it must be seriously considered. Finally the real-time constraint has to be met in many applications. In spite of all those challenges, many algorithms have already been proposed in the literature. It goes from the simple adaptation of a watermarking algorithm for still images to the really video specific watermarking scheme.

Open paths still remain in video watermarking. This technology is indeed in its infancy and is far from being as mature as for still images. Quite all possible image processings have been investigated for still images watermarking. On their side, the proposed algorithms for video have remained relatively simple. Many video processings have not been tried and the line is consequently not exhausted. Moreover, introduction of perceptual measures have significantly improved the performances of algorithms for still images. This approach has not been fully extended to video yet. Perceptual measures for video exist but the major challenge consists in being able to exploit them in real-time. Finally, the second generation of watermarking algorithms has only given its first results. Future discoveries in this domain are likely to be of great help for digital video watermarking.

Acknowledgment
The authors acknowledge ... for their pertinent remarks after reviewing this chapter.

REFERENCES

[1]. http://www.cl.cam.ac.uk/~fapp2/watermarking/stirmark/

[2]. M. Barni, F. Bartolini, R. Caldelli, A. De Rosa, and A. Piva, "A Robust Watermarking Approach for Raw Video", in *Proceedings of the Tenth International Packet Video Workshop*, 2000.

[3]. F. Bartolini, A. Manetti, A. Piva, and M. Barni, "A Data Hiding Approach for Correcting Errors in H.263 Video Transmitted Over a Noisy Channel", in *Proceedings of the IEEE Fourth Workshop on Multimedia Signal Processing*, pp. 65-70, 2001.

[4]. P. Bas and B. Macq, "A New Video-Object Watermarking Scheme Robust to Object Manipulation", in *Proceedings of the IEEE International Conference on Image Processing*, 2:526-529, 2001.

[5]. J. Bloom, I. Cox, T. Kalker, J.-P. Linnartz, M. Miller, and C. Traw, "Copy Protection for DVD Video", in *Proceedings of the IEEE*, 87(7):1267-1276, 1999.

[6]. D. Boneh and J. Shaw, "Collusion-Secure Fingerprinting for Digital Data", in *IEEE Transactions on Information Theory*, 44(5):1897-1905, 1998.

[7]. I. Brown, C. Perkins, and J. Crowcroft, "Watercasting: Distributed Watermarking of Multicast Media", in *Proceedings of the First International Workshop on Networked Group Communication*, vol. 1736 of *Lecture Notes in Computer Science*, Springer-Verlag, pp. 286-300, 1999.

[8]. M. Costa, "Writing on Dirty Paper", in *IEEE Transactions on Information Theory*, 29(3):439-441, 1983.

[9]. I. Cox, M. Miller and J. Bloom, *Digital Watermarking*, Morgan Kaufmann Publishers, ISBN 1-55860-714-5, 2001.

[10]. S. Craver, N. Memon, B. Yeo, and M. Yeung, "Can Invisible Watermarks Resolve Rightful Ownerships?", *Technical Report RC 20509*, IBM Research Division, 1996.

[11]. F. Deguillaume, G. Csurka, J. O'Ruanaidh, and T. Pun, "Robust #D DFT Video Watermarking", in *Proceddings of SPIE 3657, Security and Watermarking of Multimedia Content*, pp. 113-124,1999.

[12]. F. Deguillaume, G. Csurka, and T. Pun, "Countermeasures for Unintentionnal and Intentionnal Video Watermarking Attacks", in *Proceedings of SPIE 3971, Security and Watermarking of Multimedia Content II*, pp. 346-357, 2000.

[13]. D. Delannay, J.-F. Delaigle, B. Macq, and M. Barlaud, "Compensation of Geometrical Deformations for Watermark Extraction in the Digital Cinema Application", in *Proceedings of SPIE 4314, Security and Watermarking of Multimedia Content III*, pp. 149-157, 2001.

[14]. G. Depovere, T. Kalker, J. Haitsma, M. Maes, L. De Strycker, P. Termont, J. Vandewege, A. Langell, C. Alm, P. Normann, G. O'Reilly, B. Howes, H. Vaanholt, R. Hintzen, P. Donnely, and A. Hudson, "The VIVA Project: Digital Watermarking for Broadcast Monitoring", in *Proceedings of the IEEE International Conference on Image Processing*, 2:202-205, 1999.

[15]. J. Dittmann, M. Stabenau, and R. Steinmetz, "Robust MPEG Video Watermarking Technologies", in *Proceedings of ACM Multimedia*, pp. 71-80, 1998.

[16]. J. Dittmann, A. Behr, M. Stabenau, P. Schmitt, J. Schwenk, and J. Ueberberg, "Combining Digital Watermarks and Collusion Secure Fingerprints for Digital Images", *Proceedings of SPIE 3657, Security and Watermarking of Multimedia Content*, pp. 171-182, 1999.

[17]. J. Dittmann, A. Steinmetz, and R. Steinmetz, "Content-Based Digital Signature for Motion Pictures Authentication and Content Fragile Watermarking", in *Proceedings of the IEEE International Conference on Multimedia Computing and Systems*, 2:209-213, 1999.

[18]. J. Dittmann, M. Steinebach, I. Rimac, S. Fisher, and R. Steinmetz, "Combined Audio and Video Watermarking: Embedding Content Information in Multimedia Data", in *Proceedings of SPIE 3971*,

Security and Watermarking of Multimedia Content II, pp. 176-185, 2000.

[19]. J. Eggers and B. Girod, *Informed Watermarking*, The Kluwer International Series in Engineering and Computer Science, ISBN 1-4020-7071-3, 2002.

[20]. C. Griwodz, O. Merkel, J. Dittmann, and R. Steinmetz, "Protecting VoD the Easier Way", in *Proceedings of ACM Multimedia*, pp.21-28, 1998.

[21]. J. Haitsma and T. Kalker, "A Watermarking Scheme for Digital Cinema", in *Proceedings of the IEEE International Conference on Image Processing*, 2001.

[22]. F. Hartung and B. Girod, "Watermarking of Uncompressed and Compressed Video", in *Signal Processing*, 66(3):283-301, 1998.

[23]. C.T. Hsu and J.-L. Wu, "DCT-based Watermarking for Video", in *IEEE Transactions on Consumer Electronics*, 44(1):206-216, 1998.

[24]. F. Jordan, M. Kutter, and T. Ebrahimi, "Proposal of Watermarking Technique for Hiding/Retrieving Data in Compressed and Decompressed Video", in *ISO/IEC JTC1/SC29/WG11*, 1997.

[25]. T. Kalker, G. Depovere, J. Haitsma, and M. Maes, "A Video Watermarking System for Broadcast Monitoring", in *Proceedings of SPIE 3657, Security and Watermarking of Multimedia Content*, pp. 103-112, 1999.

[26]. S. Katzenbeisser and F. Petitcolas, *Information Hiding: Techniques for Steganography and Digital Watermarking*, Artech House, ISBN 1-58053-035-4, 1999.

[27]. D. Kilburn, "Dirty Linen, Dark Secrets", *Adweek*, 38(40):35-40, 1997.

[28]. S.-W. Kim, et al., "Perceptually Tuned Robust Watermarking Scheme for Digital Video Using Motion entropy Masking", in *Proceedings of the IEEE International Conference on Consumer Electronics*, pp. 104-105, 1999.

[29]. M. Kutter and F. Petitcolas, "Fair Benchmarking for Image Watermarking Systems", in *Proceedings of SPIE 3657, Security and Watermarking of Multimedia Content*, pp. 226-239, 1999.

[30]. G. Langelaar and R. Lagendijk, "Optimal Differential Energy Watermarking of DCT Encoded Images and Video", in *IEEE Transactions on Image Processing*, 10(1):148-158, 2001.

[31]. G. Langelaar, R. Lagendijk, and J. Biemond, "Real-Time Labelling of MPEG-2 Compressed Video", in *Journal of Visual Communication and Image Representation*, 9(4):256-270, 1998.

[32]. J. Lewis, "Power to the Peer", *LAWeekly*, 2002.

[33]. E. Lin, C. Podilchuk, T. Kalker, and E. Delp, "Streaming Video and Rate Scalable Compression: What Are the Challenges for Watermarking?", in *Proceedings of SPIE 4314, Security and Watermarking of Multimedia Content III*, pp. 116-127, 2001.

[34]. J.-P. Linnartz, "The Ticket Concept for Copy Control Based on Embedded Signalling, in *Proceedings of the Fifth European Symposium on Research in Computer Security*, vol. 1485 of *Lecture Notes in Computer Science*, Springer, pp. 257-274, 1998.

[35]. M. Maes, T. Kalker, J. Haitsma, and G. Depovere, "Exploiting Shift Invariance to Obtain High Payload Digital Watermarking", in

Proceedings of the International Conference on Multimedia Computing and Systems, 1:7-12, 1999.

[36]. J. Meng and S. Chang, “Tools for Compressed-Domain Video Indexing and Editing”, in *Proceedings of SPIE 2670, Storage and Retrieval for Image and Video Database*, pp. 180-191, 1996.

[37]. B. Mobasseri, M. Sieffert, and R. Simard, “Content Authentication and Tamper Detection in Digital Video”, in *Proceedings of the IEEE International Conference on Image Processing*, 1:458-461, 2000.

[38]. D. Mukherjee, J. Chae, and S. Mitra, “A Source and Channel Coding Approach to Data Hiding with Applications to Hiding Speech in Video”, in *Proceedings of the IEEE International Conference on Image Processing*, 1:348-352, 1998.

[39]. A. Patrizio, “Why the DVD Hack was a Cinch”, *Wired*, 1999.

[40]. F. Petitcolas, R. Anderson, and M. Kuhn, “Attacks on Copyright Marking Systems”, in *Proceedings of the Second International Workshop on Information Hiding*, vol. 1525 of *Lecture Notes in Computer Science*, Springer, pp. 218-238, 1999.

[41]. A. Piva, R. Caldelli, and A. De Rosa, “A DWT-Based Object Watermarking System for MPEG-4 Video Streams”, in *Proceedings of the IEEE International Conference on Image Processing*, 3:5-8, 2000.

[42]. L. Qiao and K. Nahrstedt, “Watermarking Methods for MPEG Encoded Video: Toward Resolving Rightful Ownership”, in *Proceedings of the IEEE International Conference on Multimedia Computing and Systems*, pp. 276-285, 1998.

[43]. C. Rey and J.-L. Dugelay, “A survey of Watermarking Algorithms for Image Authentication”, in *EURASIP Journal on Applied Signal Processing*, 6:613-621, 2002.

[44]. C. Rey, G. Doërr, J.-L. Dugelay, and G. Csurka, “Toward Generic Image Dewatermarking?”, in *Proceedings of the IEEE International Conference on Image Processing*, 2002.

[45]. D. Robie and R. Mersereau, “Video Error Correction using Data Hiding Techniques”, in *Proceedings of the IEEE Fourth Workshop on Multimedia Signal Processing*, pp. 59-64, 2001.

[46]. K. Su, D. Kundur, and D. Hatzinakos, “A Content-Dependent Spatially Localized Video Watermark for Resistance to Collusion and Interpolation Attacks”, in *Proceedings of the IEEE International Conference on Image Processing*, 1:818-821, 2001.

[47]. K. Su, D. Kundur, and D. Hatzinakos, “A Novel Approach to Collusion-Resistant Video Watermarking”, in *Proceedings of SPIE 4675, Security and Watermarking of Multimedia Content IV*, pp. 491-502, 2002.

[48]. M. Swanson, B. Zhu, and A. Tewfik, “Data Hiding for Video-in-Video”, in *Proceedings of the IEEE International Conference on Image Processing*, 2:676-679, 1997.

[49]. M. Swanson, B. Zhu, and A. Tewfik, “Multiresolution Scene-Based Video Watermarking Using Perceptual Models”, in *IEEE Journal on Selected Areas in Communications*, 16(4):540-550, 1998.

[50]. P. Termont, L. De Strycker, J. Vandewege, N. Op de Beeck, J. Haitsma, T. Kalker, M. Maes, and G. Depovere, “How to Achieve Robustness Against Scaling in a Real-Time Digital Watermarking

System for Broadcast Monitoring", in *Proceedings of the IEEE International Conference on Image Processing*, 1:407-410, 2000.

[51]. W. Trappe, M. Wu, and K. Ray Liu, "Collusion-Resistant Fingerprinting for Multimedia", in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 4:3309-3312, 2002.