الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Ferhat ABBAS Sétif 1

UNIVERSITÉ FERHAT ABBAS - SETIF1

FACULTÉ DE TECHNOLOGIE

THÈSE

Présentée au Département d'Electronique

Pour l'obtention du diplôme de

DOCTORAT

Domaine : Sciences et Technologie

Filière : Electronique **Option : Traitement de signal**

Par

NOUIOUA Imen

THÈME

Développement et Implémentation d'Algorithmes de Tatouage Numérique des Données Multimédia

Soutenue le 04/07/2019 devant le Jury:

FERHAT-HAMIDA Abdelhak	Professeur	Univ. Ferhat Abbas Sétif 1	Président
AMARDJIA Nourredine	Professeur	Univ. Ferhat Abbas Sétif 1	Directeur de thèse
BOUGHEZEL Saad	Professeur	Univ. Ferhat Abbas Sétif 1	Examinateur
HOCINI Abdesselam	Professeur	Univ. Md Boudiaf M'sila	Examinateur

Table of contents

Acknowledgements

First of all, I would like to express my sincere gratitude to my advisor Prof. Nourredine AMARDJIA for the continuous support he gave me throughout my PhD. studies and related research, for his patience, motivation, and immense knowledge. His guidance helped me during all the time of my research and also during the writing of this thesis. I could not have imagined having a better advisor and mentor for my PhD. studies.

Besides my advisor, I would like to thank Prof. Abdelhak FERHAT HAMIDA for his support, advices and for accepting to be the president of the jury.

I also want to express my deep gratitude to Prof. Saad BOUGUZEL and to Prof. Abdesselam HOCINI for taking the time and effort to read and examine my thesis.

Last but not the least, I would like to thank my family: my parents, my husband ,my brothers and sister for supporting me spiritually throughout the writing of this thesis and my life in general.

List of figures

List of tables

List of acronyms

AVC	Advanced Video Coding
BER	Bit Error Rate
CIF	Common Intermediate Format
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DWT	Discrete Wavelet transform
FA	Frame Averaging
FD	Frame dropping
FR	Frame Rate
FMFE	Fast Motion Frame Extraction
FS	Frame swapping
GIF	Graphics Interchange Format
HEVC	High Efficiency Video *Coding*
HVS	Human Visual System
JPEG	Joint Photographic Experts Group
JND	Just Noticeable Distortion
LSB	Least Significant Bit
MC	Motion Compensation
ME	Motion Estimation
MMMV	Mean Magnitudes of Motion Vectors
MPEG	Moving Picture Experts
MR-SVD	Multiresolution Singular Value Decomposition
MSE	Mean Square Error
MSSIM	Mean Structure Similarity Index Measure
NC	Normalized coefficient
NTSC	National Television System Committee
PAL	Phase Alternation Line
PSNR	Peak Signal to Noise Ratio
QCIF	Quart de Common Intermediate Format
QIM	Quantization Index Modulation
SBD	Shot Boundary Detection

S-JND	Spatial Just Noticeable Distortion
SS	Spread spectrum
ST-JND	Spatiotemporal Just Noticeable Distortion
SVD	Singular Value Decomposition
CSF	Contrast Sensitivity Function

General introduction

General context

In the recent years, the fabulous growth of the internet technology and the expansion of powerful computing devices have not only boosted the multimedia electronic commerce up but also incited artists to share and promote their work online. This obviously implied a massive presence on the web of digital multimedia data such as audio, image and video. However, with the spread out and ease of use of powerful multimedia dedicated processing tools these data can be downloaded, easily modified, illicitly appropriated and then largely redistributed or commercialized on the internet. Protecting intellectual property rights of owners has then become a major concern. Consequently, a solution to this problem is provided by digital watermarking, which consist of embedding a message named "watermark" into multimedia elements that can be detected or extracted later without changing the original content of the digital media [1]. However, to develop efficient watermarking schemas, two important properties have to be taken into account [2,3]: (1) Imperceptibility : for an invisible watermarking scheme there must be no discernible difference between the original and the watermarked contents, (2) Robustness: the embedded watermark should be able to survive, to some extent, intentional and unintentional content manipulations.

A secure digital watermarking technique comprises two procedures: an embedding procedure and an extraction procedure. The embedding procedure consists of inserting in the host multimedia content (usually called the cover) a watermark which is a digital signature that holds copyright information exclusively limited to the owner. Subsequently, by means of the given secret keys, the extraction procedure permits solely to the owner or to an authorized recipient of the digital content to retrieve the watermark from the watermarked content [2].

Digital watermarking can be done either in the spatial domain or in a transform domain. A spatial domain technique works directly on pixels: the watermark is embedded by usually modifying directly the pixels values such as least significant bits (LSBs) [4], whereas a transform domain technique embeds the watermark by adjusting the transform domain coefficients. Popular transforms that have been frequently used are the Discrete Cosine Transform (DCT) [5], the Discrete Wavelet Transform (DWT) [6], the Discrete Fourier Transform (DFT) [4] and the Singular Value Decomposition (SVD) [7, 8]. Many combinations between these transforms have also been investigated in the literature to accomplish better results [9, 10]. Compared to spatial domain techniques, transform domain ones have shown to achieve better robustness and imperceptibility [9]. Furthermore, the extracting process can be blind, semi-blind or non-blind. In a blind watermarking scheme, neither the original cover nor the embedded watermarks are required for detection but just the secret keys [2, 7, 11]. In a semi-blind watermarking scheme, only some information from the original cover and the secret keys are needed [2, 12]. A non-blind watermarking scheme requires the original cover, the original watermark and the secret keys [2, 9]. This makes the blind watermarking schemes the most challenging ones to develop.

Initially, digital watermarking has been mainly studied for still images but in recent few years a considerable number of techniques dealing with video watermarking have been considered. However, one must say that video watermarking algorithms are more difficult to develop than those operating on images. This is essentially due to the temporal dimension which necessitates some specific requirements [13]: (1) The robustness of the watermark should not only deal with common image processing attacks such as noise adding, JPEG compression, etc., but also with video processing attacks such as MPEG compression and frame synchronization attacks. (2) The imperceptibility in video watermarking is more difficult to achieve due to motion of objects in video sequences, so the temporal dimension should be taken into account in order to avoid distortion between frames. (3) The complexity of the watermarking scheme should be low because of the significant number of frames to be processed in a video signal.

Given that a digital video sequence is considered basically as a collection of sequential images [14], many of the image watermarking techniques that are present in the literature were extended to video [6, 9, 15, 16], as they embed the watermark in all frames of the video sequences without taking the temporal dimension into account. Thus, these algorithms are robust to frame dropping and frame swapping, but in return they are time consuming and also affect the perceptibly of the video quality.

Contributions

To overcome the problem in frame-by-frame embedding techniques, we should not neglect the temporal dimension of a video to make the schema more performing. Therefore, in this thesis we will propose two video watermarking techniques in the uncompressed domain:

- A robust video watermarking using spatiotemporal just noticeable distortion: In this work, we will present a robust video watermarking scheme based on Spatiotemporal Just Noticeable Distortion (ST-JND) in the DCT domain. ST-JND, which refers to the maximum distortion threshold that the HVS cannot perceive, is employed in order to obtain a trade-off between watermark robustness and imperceptibility. The watermark sequence is embedded and detected using a quantization index modulation (QIM) algorithm with adaptive (dynamic) step size that is calculated using the ST-JND. By this way, we will consider the temporal dimension of video sequences.
- A Novel Blind and Robust Video Watermarking Technique in fast motion frame using SVD in MRSVD domain: In this second work, we will propose a novel and efficient digital video watermarking technique based on the Singular Value Decomposition performed in the Multiresolution Singular Value Decomposition domain. The proposed method chooses only the fast motion frames in each shot of the video to host the watermark. In doing so, the number of frames to be processed is consequently reduced and a better quality of the watermarked video is ensured since the human visual system cannot notice the variations in fast moving regions.

Chapters’ organization

The remaining parts of this thesis are organized as follows:

- Chapter 1 provides an overview of digital watermarking by including the historical aspect and also its applications and properties. Moreover, it explains the basic scheme of a watermarking system and gives a summary of the classification of different watermarking schemas.

- Chapter 2 exhibits a specific overview of video watermarking by giving a fundamental view on video signal. Furthermore, it talks about specific challenges that we have to take into account in video watermarking. It also provides the reader by a global view of the scientific production in video watermarking.
- Chapter 3 presents a robust video watermarking using spatiotemporal just noticeable distortion, which is a representation of the spatial and temporal human visual system. By this way, we overcome the problem in frame-by-frame watermarking technique, which ignores the temporal characteristic of a video. We get then a more efficient watermarking scheme.
- Chapter 4 exposes a novel video watermarking scheme in fast motion frames using Singular Value Decomposition in the Multi Resolution Singular Value. As in chapter 3, to solve the problem of a frame-by-frame watermarking technique, the proposed method chooses only the fast motion frames in each shot to host the watermark. By doing so, the number of frames to be processed is therefore reduced and a better quality of the watermarked video is ensured since the human visual system cannot notice the variations in fast moving regions.

Chapter 1: Digital watermarking – an overview

Chapter organisation

This chapter aims to provide an overview of digital watermarking. Section 1.2 reviews the historical aspect. Section 1.3 discusses several applications in which digital watermarking are already being used. Section 1.4 lists important properties for watermarking. Section 1.5 gives the basic framework of a digital watermarking system. Section 1.6 illustrates different aspects used in watermarking classification. Finally, Section 1.7 describes how a digital watermarking algorithm can be evaluated.

1.1. Introduction

With the rapid development of the Internet, digital documents such as pictures, video and music became more available to users with high quality and low cost, which lead to the problem of unauthorized copying and redistribution through the network. Therefore, owners and creators of digital products are searching for reliable solutions to protect the copyright of their digital contents against piracy and malicious manipulation. In the late 1990, digital Watermarking has been established as an effective solution to these concerns by embedding a message named “watermark” into multimedia elements that can be detected or extracted later without changing the original content of digital media [17, 18].

In this chapter, we intend to provide an overview of digital watermarking. First, we will review the historical aspect of digital watermarking, give some of its applications and provide its important properties. We will then explain the basic scheme of a watermarking system. After that, we will give an overview on the classification of the different watermarking techniques and finally we will speak about the performance evaluation of a watermarking algorithm.

1.2. History

Paper watermarks appeared in the art of papermaking about 1282, in Italy.This has played a major role in the evolution of the papermaking industry. The competition between professionals and the quantity of paper processed made it difficult to keep track of the paper source. The introduction of watermarks was a perfect method to eliminate any possibility of confusion. By the eighteenth century, the technique spread in Europe and America and was basically used to indicate the brand or manufacturer of the paper. Watermarks later served as an indication of paper format, quality, strength and also for dating and identification.

Other interests for watermarks, such as bank notes or stamps, show that the technique is a straightforward and quite secure mean for identification since it is still used nowadays. From paper watermarks to digital watermarks, the gap narrowed. The latest term deals with identification of digital documents instead of paper. The first research papers on watermarking of digital images were published by Tanaka et al. in

1990 and later by Tirkel et al. in 1993. Around 1995, the interest in the digital watermarking started to increase and the subject began to stimulate the intensification of research activities [1].

1.3. Applications

Watermarking can be used in a wide variety of applications. We can cite copyright protection, broadcast monitoring, transaction tracking, authentication, copy control …etc [19].

1.3.1. Copyright protection

Historically, copyright protection is the first targeted application for digital watermarking. The data owner can embed a watermark representing the copyright information in his data. This watermark can prove the ownership if an illegal copy is found. Muzak's original interest in watermarking was to distinguish between theirs and similar recordings. The most ambitious form of such an application, which has received much attention in the watermarking literature, is the use of watermarks to actually prove ownership in a court of law. In 1996, Craver et al. pointed out that there is an inherent problem in using watermarks for proof of ownership. Specifically, with many watermarking methods, it is possible for adversaries to make it appear as though all distributed copies of a Work contain their watermarks, even though those marks were never actually embedded [19,20].

1.3.2. Broadcast monitoring

Broadcast monitoring is the process of tracking activities on broadcasting channels in compliance with intellectual rights and other broadcasting laws. By embedding watermarks in commercial advertisements an automated monitoring system can verify whether or not advertisements are broadcasted as contracted [1,19].

1.3.3. Transaction tracking

In transaction tracking or fingerprinting, a unique watermark is embedded into each copy of a work. Typically, the watermark identifies the legal recipient of the copy and can be used to trace the source of illegally redistributed content. One common alternative to watermarking is to use visible marks. For example, highly sensitive business documents, such as business plans, are sometimes printed on backgrounds containing large gray digits (see fig. 1.1), with a different number for each copy. Records are then kept about who has which copy. These marks are often referred to as “watermarks” [1,19].

Proceedings of The first International Conference on Nanoelectronics, Communications and Renewable Energy 2013 222

A Novel Video Watermarking Technique Based On Multiresolution Singular Value Decomposition

NOUIOUA Imen
Laboratoire Croissance et Caractérisation de Nouveaux Semi-conducteurs (LCCNS) – Département d'Electronique - Université FERHAT Abbas de Sétif
imen_bou84@hotmail.fr

AMARDJIA Nourredine
Laboratoire Instrumentation Scientifique (LIS) – Département d'Electronique - Université FERHAT Abbas de Sétif
amardjia_nour@yahoo.fr

Abstract— In this paper, a new robust video watermarking technique is proposed which combines Singular Value Decomposition (SVD) and its multiresolution form (MR-SVD). The experimental results obtained through the proposed scheme demonstrate that this technique provides imperceptible watermarks and is robust against several attacks like noise addition, rotation and MPEG compression.

Keywords- Video watermarking; Singular Value Decomposition (SVD); Multiresolution-SVD (MR-SVD).

I. INTRODUCTION

Due to the extraordinary revolution of internet, the digital multimedia (movie, music, and image) can be easily obtained with low cost and high quality. Therefore, owners and creators of digital products are concerned about illegal copying of their products. The Digital Watermarking has been proposed as a solution to the problem of copyright protection [1].

The basic idea of the “watermarking” is to embed a robust and subliminal (invisible or inaudible) data into multimedia elements [2].

Watermarking techniques can be classified into two classes: spatial domain and frequency domain. The spatial domain methods embed the watermark by modifying the pixel values of the host directly, but in the frequency domain techniques, the watermark is embedded by modifying the transform coefficients of the host. The most commonly used transforms are the Discrete Fourier Transform (DFT), the Discrete Cosine Transform (DCT), the Discrete Wavelet Transform (DWT) and the singular value decomposition (SVD) [3, 4].

Kakarala and Ogunbona [5] have proposed a multiresolution form of the SVD (MR-SVD) and showed how it could be used for signal analysis. A new video watermarking scheme is presented here, where the SVD is combined with the MR-SVD.

This paper is organized as follow. To understand our scheme, we introduce in the next section the concept of SVD and MR-SVD transformation techniques. In section three, we propose our video watermarking method where we describe the watermark embedding and extraction processes. In section four, we present the experimental results and in section five, we compare our results with existing work done by previous researchers. The conclusions of our study are stated in section six.

II. THEORETICAL FUNDAMENTALS

A. Singular Value Decomposition

Singular Value Decomposition is said to be a significant topic in linear algebra by many renowned mathematicians. Also SVD has shown its usefulness in variety of applications including image processing and watermarking.

If A is a matrix representing for example an image of size m*n, then the SVD of A is given by:

$$A = U*S*V^T \quad (1)$$

Where U and V are orthogonal matrices and S is a diagonal matrix $S = diag(\{s_1, s_2 \dots s_n\})$

$s_1, s_2 \dots s_n$ are the singular values (SV) of A, which observe the condition $s_1 \geq s_2 \geq \dots \geq s_n$; These SV represent the energy of the image and have very good stability [6].

B. Multiresolution singular value decomposition

The MR-SVD represents a signal as a series of an approximation and details like the DWT. This section explains how the MR-SVD, introduced in [5], works.

1) 1D Multiresolution Singular Value Decomposition

Let X=[x(1) ... x(N)] represent a finite extent 1-D signal. Assume that N is divisible by 2L for some L ≥1. Let the data matrix at the first level, denoted X1, be constructed so that its top row contains the odd-numbered samples and the bottom row contains the even-numbered samples:

$$X_1 = \begin{pmatrix} x(1) & x(3) & \dots & x(N-1) \\ x(2) & x(4) & \dots & x(N) \end{pmatrix} \quad (2)$$

Let U_1 be the eigenvector matrix bringing the scatter matrix $T_1 = X_1 X_1^T$ into diagonal form:

$$U_1^T T_1 U_1 = S_1^2 \quad (3)$$

Where $S_1^2 = diag\{s_1(1)^2, s_1(2)^2\}$ contains the squares of the two singular values, with $s_1(1) \geq s_1(2)$.
Now let:

ICNCRE '13 ISBN : 978-81-925233-8-5 www.edlib.asdf.res.in

Figure 1.1 Example of a unique identification code printed in the background of a text document.

1.3.4. Content Authentication

The purpose of content authentication is to verify the integrity of the test data and detect any possible manipulation (i.e. tamper detection) as in fig. 1.2. This is done by embedding fragile or semi fragile watermarks. When the data is corrupted, the watermark will also be changed [21].

Figure 1.2 Left is the original image, right is the tampered image [21].

1.3.5. Copy Control

Copy control is essentially subjected to protect digital information from illegal copying. This can be done by using a recording device integrated with a watermark detector which determines whether the data offered to the recorder may be stored or not [22].

1.4. Properties of watermarking systems

We can identify some general properties common to existing watermarking systems.

1.4.1. Imperceptibility (transparency or fidelity)

One of the most important requirements in digital watermarking is the perceptual transparency of the watermark. Generally, the imperceptibility of a watermarking system refers to the perceptual similarity between the original and watermarked versions of the cover (i.e. the watermarks do not create visible artifacts in still images, alter the bit rate of video or introduce audible artifacts in audio signals). For this reason, the most of existing algorithms, in image or video watermarking, take the human visual system HVS into account to perform the embedding of watermarks [1].

According to imperceptivity property, we can find two kind of algorithms in digital watermarking: visible watermarking and invisible watermarking.

- ❖ Visible watermarking: In perceptible or visible watermarking systems, the embedded watermark or logo is visible to human eye without degrading the readability of the digital media as shown in fig. 1.1. In general they are applicable to images and video only. Examples of visible watermarks are logos on TV visibly superimposed on the corner of the video frames [1, 23].

- ❖ Imperceptible watermarking: In imperceptible watermarking, the watermarked document and the original one are extremely similar. Generally they are useful for complex applications such as document identification in which watermarked content must appear unchanged compared to the original.

1.4.2. Robustness

Robustness refers to the ability to recover the watermark after transmission of the watermarked host in a noisy channel and a distortion introduced by standard attacks which try to remove the watermark. Depending on the application of a watermarking method, the required robustness must influence the design process. For example, in television broadcast monitoring, the watermarks should be robust to lossy compression, digital-to-analog conversion, lowpass filtering, additive noise and some small amount of horizontal and vertical translation. However, watermarks for this application need not survive rotation, scaling, high-pass filtering, or any of a wide variety of degradations that occur only prior to the embedding of the watermark or after its detection [1].

According to robustness property we can find three classes of digital watermarks dealt in the literature: robust, fragile and semi fragile.

- ❖ Robust: Robust watermarking is mainly used to sign copyright information of the digital works where the embedded watermark must resist the common edit processing, image processing and lossy compression.
- ❖ Fragile: Fragile watermarking is largely used to check accuracy authentication as in fig. 1.2, which must be very sensitive to the changes of signal. We can determine whether the data has been tampered or not according to the state of a fragile watermarking [21].

- Semi fragile: Semi fragile watermarking is capable of tolerating some degree of change to a watermarked image such as the addition of noise from lossy compression. It is primarily used to certificate the integrity and authenticity of image data [24].

1.4.3. Capacity

The capacity or data payload is the maximum number of bits that can be hidden in a given cover Work [1]. Capacity property depends on the purpose of application. Thus different applications may require very different data payloads. For example, in television broadcast, monitoring might need at least 24 bits of information to identify all commercials, whereas in copy protection purposes, a payload of 4 bits is usually sufficient (see table 1.1) [1, 25].

1.4.4. Security

Security property of watermarking schemes may not be always necessary in practice. Cox et al. [1] referred to security as the ability to resist hostile attacks. A hostile attack is any process specifically intended to thwart the watermark's purpose. The security of a watermark influences the robustness enormously. If a watermark is not secure, it cannot be very robust.

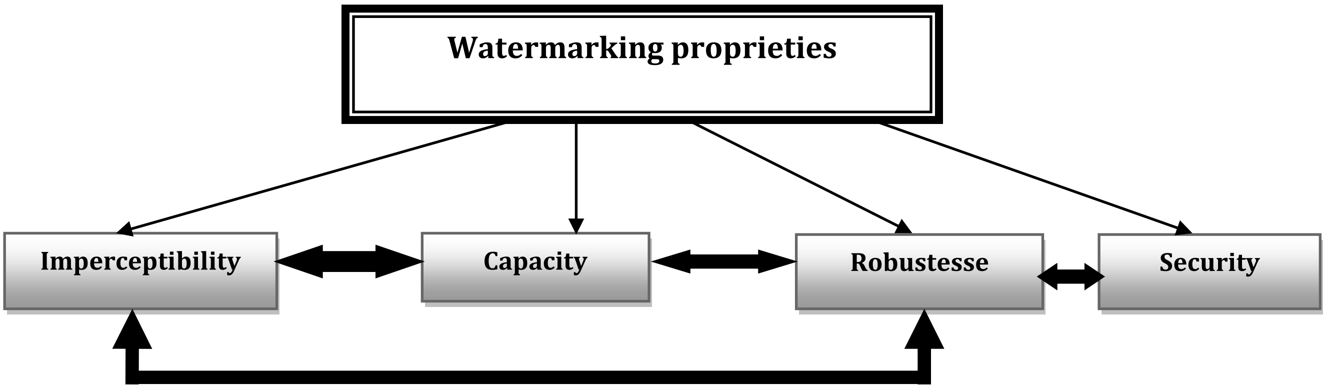


Figure 1.3 Dependencies between basics propreties of digital watermarking.

The relation between the basic properties is presented in fig. 1.3. If a watermarking system has better performance in one property, it may be weak in other properties. For example, a very robust watermark can be obtained by large modifications. Consequently, large modifications in the host will be noticeable (i.e.

low perceptible quality). However, different applications of digital watermarking require different level of such properties. Table 1.1. Presents digital watermarking properties versus applications.

Table 1.1. Digital watermarking properties versus applications.

Application	imperceptibility	Robustness	capacity	Comment
Copy control	High	High	Low	Copy control applications may require just 4 to 8 bits of information [1].
Television broadcast monitoring	High	Medium	Medium	Television broadcast monitoring might require at least 24 bits of information to identify all commercials [1].
Copyright protection	Medium	Low	High	Copyright protection requires from 64 bits [38].

1.5. Basic watermarking scheme

A basic watermarking scheme is resumed in fig. 1.4. Every watermarking system share the same basic model: a watermark embedding system, transmission over a channel and a watermark recover system (also called watermark extraction system).

1.5.1. Watermark embedding system

The watermark embedding system takes three inputs as shows fig. 1.4: a watermark, a cover data and a secret or public key. This phase can be modeled by the following function:

$$C_W = E(C, W, K) \quad (1.1)$$

where E is the embedding function, C the original cover, W the original watermark and K a secret or public key.

The watermark can be any kind of data: number, text or image. The cover may be in theory any kind of digital document. However, the most appropriate for digital watermarking are images, video and audio. In the embedding process, the watermark is inserted into the content in various methods that we can classify according to the embedding domain (spatial and transform) and the insertion technique (additive and substitution).

1.5.1.1. Embedding domain

1.5.1.1.1. Spatial domain

In the spatial domain, the watermark is embedded by usually modifying directly the pixels values such as the least significant bits (LSBs) method and the patchwork method [2, 4]. These methods are simple and inexpensive in term of computing time. Some spatial domain techniques can be robust to geometric attacks.

1.5.1.1.2. Transform domain

Transform domain techniques embed the watermark by adjusting the transform domain coefficients of the host. Three main steps must be specified: host transformation, watermark embedding, and watermark recovery. Compared to spatial domain techniques, transform domain ones have shown to achieve better robustness and imperceptibility [9].

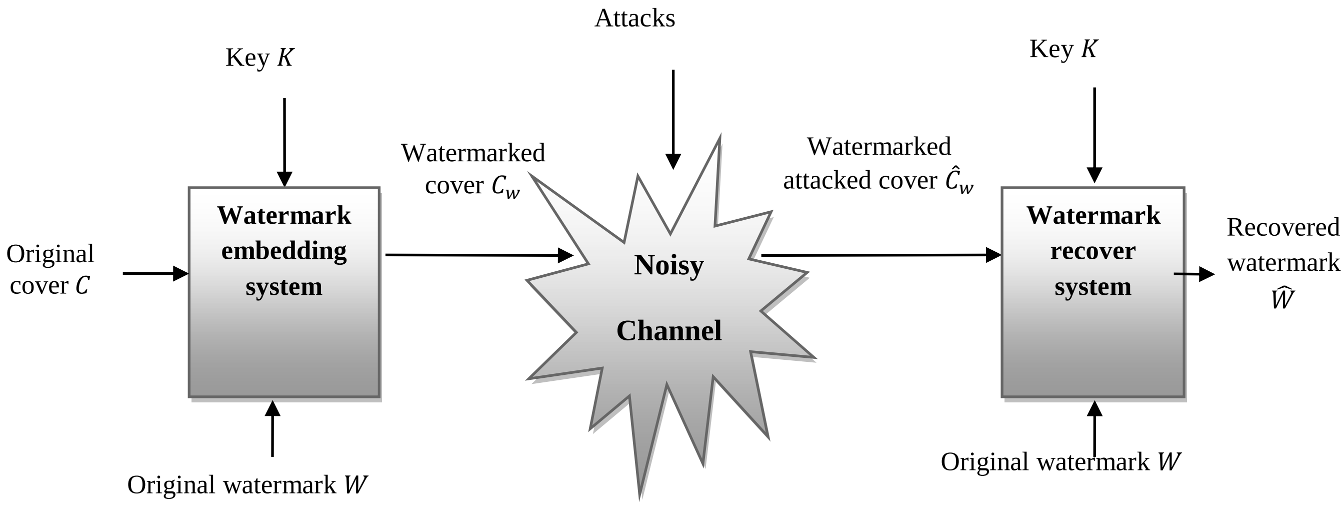


Figure 1.4 Basic model of a digital watermarking system.

Popular transforms that have been frequently used are the Discrete Cosine Transform (DCT), the Discrete Wavelet Transform (DWT) and the Singular Value Decomposition (SVD).

- ***The Discrete Cosine Transform (DCT):*** The DCT is one of the most popular transforms used in digital image processing and signal processing. Many of the compression techniques are developed in the DCT domain (JPEG, MPEG, MPEG1, and MPEG2) because of its advantage of concentrating the energy of the transformed signal in low frequency range and also its low complexity implementation. For this reasons, many researchers have used the DCT in digital watermarking which provides robustness to JPEG and MPEG compression attacks. Moreover, watermarking in the DCT domain offers the possibility of directly realizing the embedding operator inside a JPEG or MPEG encoder [5, 26, 27].

- ***The Discrete Wavelet Transform (DWT):*** The DWT has been successfully applied for the new compression standard JPEG-2000. In several recent publications, this technique has been useful to image watermarking because it has excellent spatial localization, and multi-resolution characteristics [6].

- ***The Singular value decomposition (SVD)***: In recent years, SVD has been proposed as a powerful watermarking technique for copyright protection due to its attractive mathematical features. In chapter 3, we detail the principle of this decomposition [7, 8].

1.5.1.2. Embedding techniques

1.5.1.2.1. Additive watermarking

In the additive watermarking, the message to be added is not correlated with the host image. The main reason for the popularity of additive watermarking is its simplicity, for which:

$$Y_i = X_i + \alpha W_i \qquad (1.2)$$

where X_i is the i-th component of the original signal, W_i the i-th sample of the watermark signal, α a scaling factor controlling the watermark strength, and Y_i the i-th component of the embedded signal.

In the following we will give some examples of additive watermarking schemes.

- ***Patchwork [28]:***

The patchwork algorithm is one of the earliest watermarking algorithms which appeared in the scientific literature. It is a typical spatial domain, additive algorithm.

The patchwork algorithm can be summarized as follows. The embedding process is achieved by randomly selecting, according to a secret key K, a subset S of image pixels, and then dividing S into two equal subparts S_1 and S_2. Then, pixels belonging to S_1 are increased by a small quantity d, whereas pixels in S_2 are decreased by the same amount using the following formula:

$$X_w(i,j) = X(i,j) + dW_i(i,j) \quad (1.3)$$

where $X(i,j)$ is the image pixel at position (i,j) and $W(i,j)$ is a white, pseudo-random signal taking values +1 or -1 with equal probability.

$$W = \begin{cases} 1, & If\ X(i,j) \in S1 \\ -1, & If\ X(i,j) \in S2 \end{cases} \quad (1.4)$$

by replacing (1.4) in (1.3) we find

$$X_w(i,j) = \begin{cases} X(i,j) + \alpha, & If X(i,j) \in S1 \\ X(i,j) - \alpha, & If X(i,j) \in S2 \end{cases} \quad (1.5)$$

For a watermarked image, the average difference between pixels in S1 and S2 should approach 2d, conversely for non watermarked image the average difference should be close to zero.

The most significant problem with this scheme is the fragility in the face of synchronization attacks (e.g. geometrical attacks) [29].

- ***Additive Spread spectrum(SS)[1]:***

Cox et al. introduced the concept of Spread Spectrum (SS) watermarking in [30]. The basic idea is to spread one message bit over many samples of the host data. This is done by modulating the host data with a sequence obtained from a pseudo-random generator.

Fig. 1.5 shows a generic model for a spread spectrum watermark embedding system. The SS watermark embedding in the sample domain can be achieved by additive embedding as:

$$y_i = x_i + \alpha w_i \tag{1.6}$$

$$w_i = b_i \cdot p_i \tag{1.7}$$

where x is the vector representation of the M samples of the host data, b the vector representation of the spread sequence of size M, α a scaling parameter, p pseudo-noise sequences with $p_i \in \{-1,+1\}$, and y the vector representation of the resulting watermarked data of size M.

The embedding can be applied in spatial domain or transform domain using wavelet transform, discrete cosine transform, Fourier transform, or another transform.

In an SS technique, the watermark can be extracted without using the original host, by correlating the watermarked signal y with the same pseudo-random sequence p used in the embedding. However, the recovery of watermark is more robust if the original host is available at the decoder [31].

The SS watermarking scheme is robust to signal processing operations (such as lossy compression, filtering, etc.) and common geometric transformations (such as cropping, scaling, translation, and rotation). On the other hand, a major drawback of SS watermarking methods is their low embedding capacity due to the spreading of a single bit in larger cover sample in order to achieve higher robustness and security [32].

1.5.1.2.2. Substitute watermarking

Substitute watermarking is based on the substitution of some proprieties of the host in order to embed the watermark. The most popular substitute method is the Quantization Index Modulation (QIM) method [29].

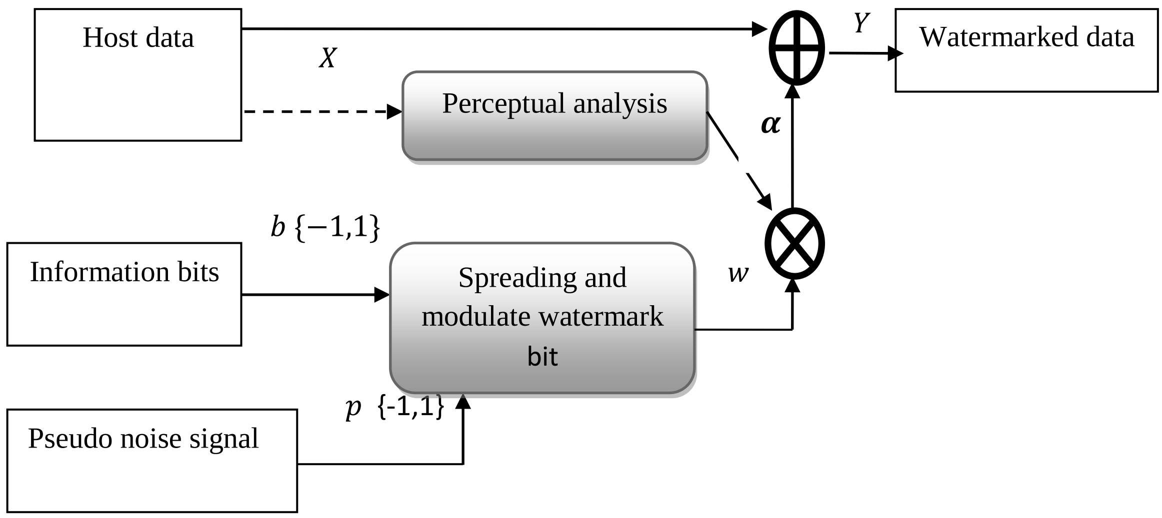


Figure 1.5 A generic model of a spread spectrum watermark embedding mechanism.

❖ ***Quantization index modulation (QIM)[33]:***

To overcome the problem in additive spread spectrum watermarking technique Chen and Wornell proposed a new class of embedding methods, termed Quantization Index Modulation (QIM). QIM is accomplished by modulating a signal with the embedded information. Then, the Quantization is performed using the associated quantizer.

One of the QIM techniques presented in [33] is the dither modulation (DM). The basic DM algorithm quantizes the feature vector x using a quantizer q chosen from a family of quantizes based on the message bit w to be embedded. The watermarked featured vector y is then obtained by:

$$y = s(x; w) = q\big(x + d(w)\big) - d(w) \tag{1.8}$$

where q (.) is a quantization function with the step Δ defined as :

$$q(x) = round\left(\frac{x}{\Delta}\right)\Delta \tag{1.9}$$

d(w) is a dither value corresponding to the bit of w

$$d(1) = \begin{cases} d(0) + \Delta/2\,, & d(0) < 0 \\ d(0) - \Delta/2\,, & d(0) \geq 0 \end{cases} \tag{1.10}$$

Fig. 1.6 illustrates how a scalar host signal x can be dither modulated to embed either $w = 0$ or $w = 1$. We simply choose $d(0) = -\Delta/4$ and $d(1) = +\Delta/4$

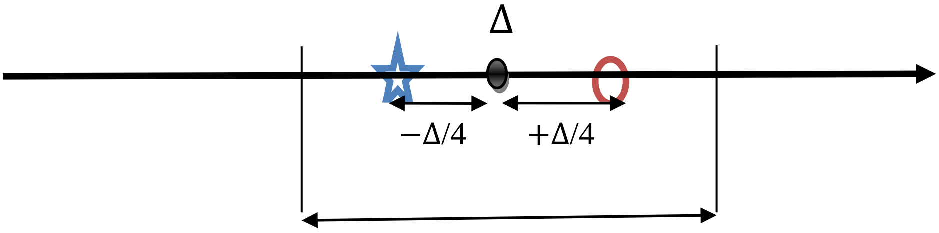


Figure 1.6 Illustration of dither modulation of the scalar host signal x, using a quantiser step size ofΔ, $\mathrm{d}(0) = -\Delta/4$, and $d(1) = +\Delta$.

To determine the embedded message bit $\widetilde{W}$ at the decoder, the watermarked signal y is requantized using the same family of quantizer as presented in fig. 1.7. Then, the minimum distance decoder is performed as:

$$\widetilde{W} = argmin_{w \in \{0.1\}} |y - s(y; w)| \qquad (1.11)$$

QIM achieves provably a high embedding capacity with low-complexity realizations . Furthermore , the extraction procedure in QIM is blind which makes it suitable for robust watermarking.

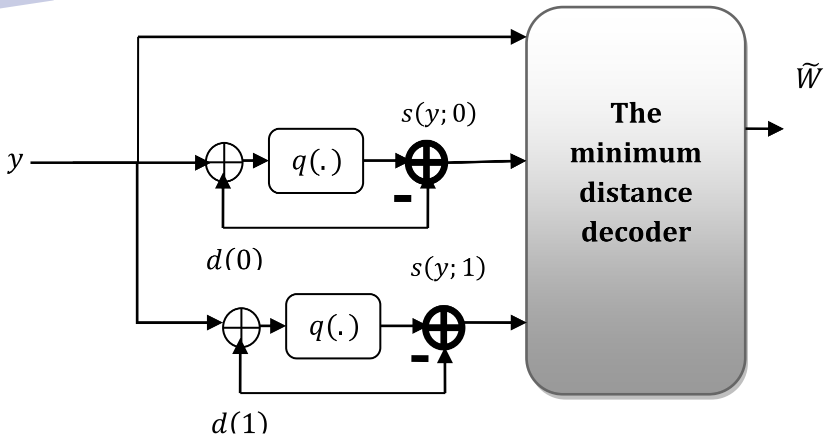


Figure 1.7 Dither Modulation recover system.

1.5.2. Transmission over a noisy channel

After the embedding of the watermark, the watermarked document C_W is broadcast in a lossy channel (i.e. undergoes a series of manipulations or modifications). Stefan et al. [2] classified the standard distortions or attacks as: Destruction attacks and Synchronization attacks.

1.5.2.1. Destruction attacks

Destruction attacks are groups of distortions which can be considered as additive noise to the data. These are:

- Additive and multiplicative noise (Gaussian, uniform, speckle, mosquito).
- Filtering (median filtering, morphological filtering).
- Lossy compression images: JPEG. Video: H.261, H.263, MPEG-2, MPEG-4. Audio: MPEG-2 audio, MP3, MPEG-4 audio, G.723).
- Transcoding: (H.263 → MPEG-2, GIF → JPEG).

1.5.2.2. Synchronization attacks

The synchronization attacks are distortion due to the spatial and temporal geometry modifications of the data like:

- Local and global affine transforms: translation, rotation, scaling, shearing.
- Data reduction: cropping, clipping, histogram modification.

1.5.3. Watermark recovery system

Generally, the recovery scheme takes two inputs, the watermarked attacked cover and the key, and in some cases, an additional one such as the watermark or the original cover. It is modeled by the following function:

$$\widehat{W} = R(\hat{C}_w, K, \dots) \tag{1.12}$$

where R is the recover function, $\widehat{W}$ is the recover watermark and $\hat{C}_w$ is the the watermarked attacked cover

Depending on the information needed to extract the watermark, digital watermarking systems are categorized in three schemes [2]: blind, semi-blind and non-blind.

1.5.3.1. Blind watermarking

This kind of system only requires the watermarked data to extract the watermark. For this reason, it is the most challenging one to develop [2, 7, 11, 35].

1.5.3.2. Semi-blind watermarking

It does not use the original data but only the copy of the watermark and answers whether it is present or not [2, 12, 35].

1.5.3.3. Non- blind watermarking

A non-blind watermarking scheme requires the original cover. It can only be used in those applications where the original Work is available (private watermarking applications). Non-blind or Semi-blind can be used for evidence in court to prove ownership, copy control or fingerprinting (where the owner should still have an unwatermarked version of the Work) [2, 9, 19].

1.6. Classifications of watermarking techniques

From the previous sections, we can make a classification of the watermarking techniques. As show fig. 1.8, it is possible to group them according to different criteria [36]:

- ❖ Type of documents (audio, image and video);
- ❖ Embedding domain (spatial, spectral);
- ❖ Embedding technique (additive and substitution);
- ❖ Information needed for extraction (blind, semi-blind and non-blind);
- ❖ Preservation of the original image (invertible and non-invertible)
- ❖ Robustness of watermark (robust, fragile and semi-fragile).

Each class of watermarking technique has a different purpose, but in this thesis, we will focus on the imperceptible and robust video watermarking in the transform domain.

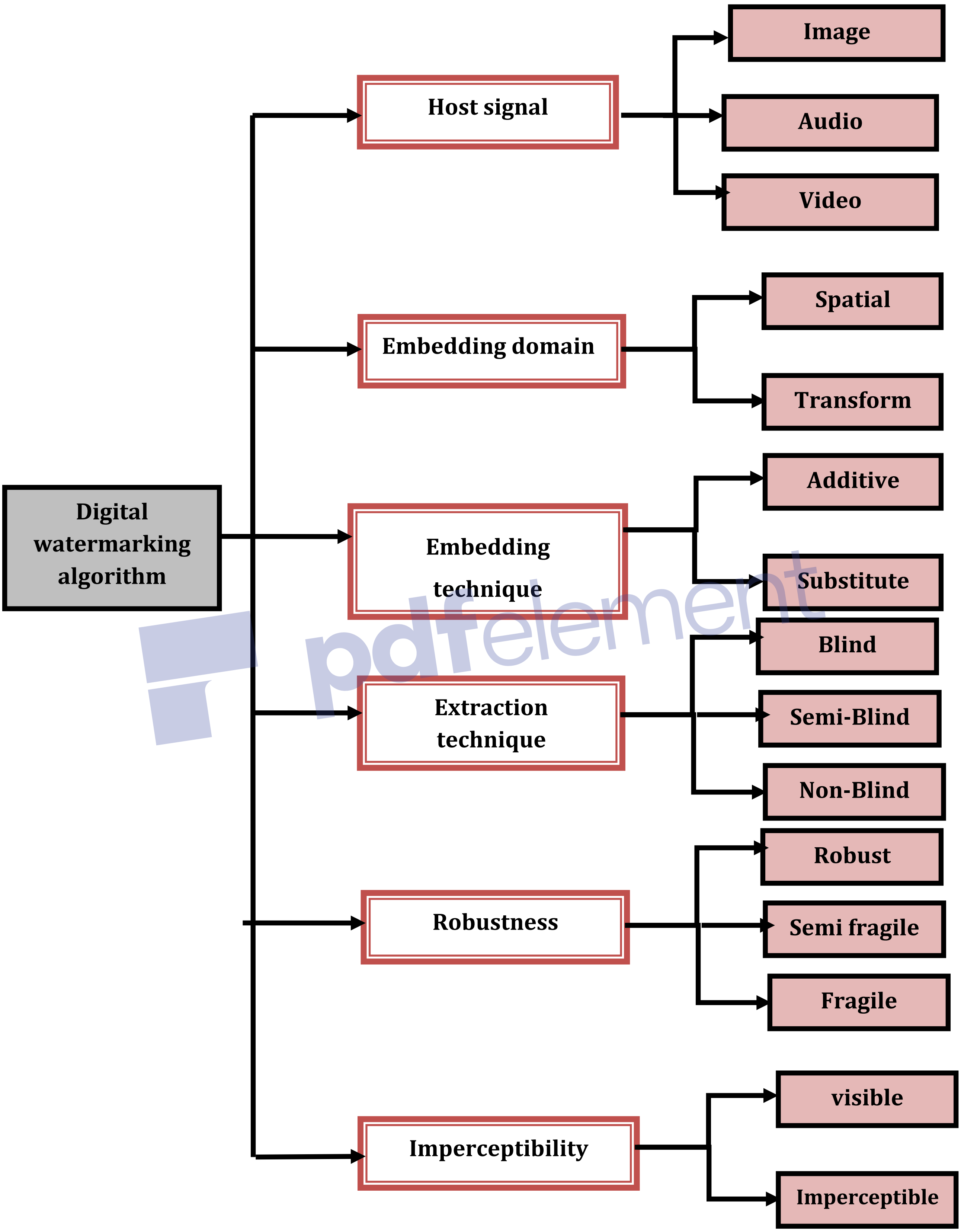


Figure 1.8 Classifications of the watermarking techniques.

1.7. Performance evaluation of image and video watermarking algorithms

1.7.1. Imperceptibility performance

The imperceptibility of the watermark is estimated by measuring the PSNR (Peak Signal to Noise Ratio) and Mean Structure Similarity Index Measure (MSSIM) [37]. The PSNR is calculated as follows:

$$PSNR = 10log_{10}\left(\frac{256^2}{MSE}\right) \tag{1.13}$$

where the Mean Square Error (MSE) between the host luminance Y and the watermarked luminance Y' is defined as:

$$MSE = \frac{1}{M \times N}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}|Y(i,j) - Y'(i,j)|^2 \tag{1.14}$$

with M and N respectively being the height and width of the video frame.

The MSSIM is defined as follows:

$$MSSIM(Y,Y') = \frac{1}{M}\sum_{j=1}^{M} SSIM(Y_J, Y_j') \tag{1.15}$$

where Y_J and Y_j' are the image contents at the j local window, and M is the number of local windows of the image.

$$SSIM(Y,Y') = [l(Y,Y')]^{\alpha} \cdot [c(Y,Y')]^{\beta} \cdot s[Y,Y')]^{\gamma} \tag{1.16}$$

where l, c and s are the luminance comparison, the contrast comparison and the structure comparison functions respectively. α, β and γ are parameters used to adjust the relative importance of the three components.

Generally ,the imperceptibility of a watermarking algorithm will be desirable when the PSNR is greater than 36 dB and the MSSIM close to 1 [1].

1.7.2. Robustness performance

The robustness for any watermarking system is a very important requirement. To compare the similarities between the original watermark W and the extracted watermark $\widehat{W}$, we use the Normalized Coefficient (NC) and the Bit Error Rate (BER) [34].

The NC and BER are respectively calculated as:

$$NC(W,\widehat{W}) = \sum\sum \frac{W(i,j)\widehat{W}(i,j)}{|W(i,j)|^2} \quad (1.17)$$

$$BER = \frac{1}{P}\sum_{j=1}^{p}\left|\widehat{W}(j) - W(j)\right| \quad (1.18)$$

where W, $\widehat{W}$ and P are respectively the original watermark, the extracted watermark and the size of the watermark. The correlation between W and $\hat{W}$ is very high when NC is close to 1

1.8. Conclusion

In this chapter, we provided an overview of digital watermarking including its properties, applications and basic scheme. We have also presented a classification of watermarking techniques according to different criteria such as: host signal, imperceptibility, robustness, extraction technique, embedding domain and embedding technique. Each class of watermarking techniques has a different purpose. However, in this thesis, we will focus on the imperceptible and robust video watermarking techniques in the transform domain. Therefore, in the next chapter, we will present a detailed overview of video watermarking with its specific challenges. Then, we will expose the different watermarking techniques intended to video.

Chapter 2: Video watermarking

Chapter organisation

In this chapter, we aim to provide an overview of digital video watermarking. Sections 2.2 gives fundamental view on video signal concepts. Section 2.3 lists the main requirements that have to be taken up when designing a new video watermarking system. Section 2.4 gives a classification of existing video watermarking. Finally, conclusions are given in section 2.5.

2.1. Introduction

Nowadays, online video has become one of the top activities for users. For example, the YouTube web site contains more than 1.9 billion users and every day, users watch over a billion hours of video. Despite this, research in digital video watermarking is still somewhat an unexplored field of research compared to image watermarking [39, 40] and in which the most of existing image watermarking techniques that are present in the literature were extended to video by embedding the watermark in all frames of the video sequences without regard to the temporal structure of the video.

Therefore, in this chapter we will present a specific overview of video watermarking. First, we will give a fundamental view on video signal. Then, we will talk about specific challenges that have to be taken into account in video watermarking and finally, we will provide the reader by a global view of the scientific production in video watermarking.

2.2. video concept

Before discussing the details of video watermarking challenges and techniques, we provide in the following a fundamental view on video.

2.2.1. Video structure

The term video refers to the visual information captured by a camera, and it is usually applied to a time-varying sequence of pictures. A video signal comprises a sequence of frames (i.e. images) that move rapidly in succession with a fixed frame rate (typical frame rates are 30 and 25 frames per second as seen in the various video formats NTSC, PAL, etc.) [41].

As illustrated in fig.2.1, a video can be broken down into a hierarchy of three units: frames, shots, and scenes. A shot is a sequence of frames recorded in a single-camera operation and a scene is a collection of consecutive shots that have semantic similarity in objects, persons, space and time [42].

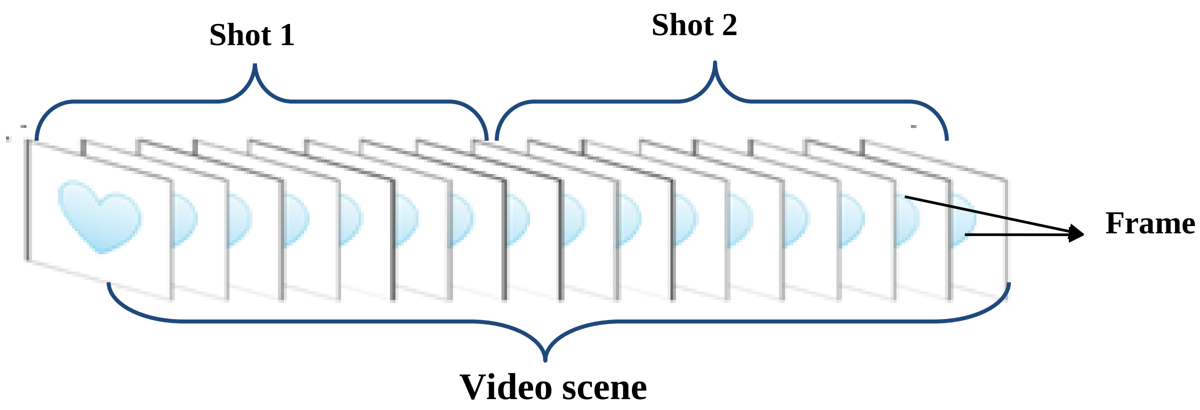


Figure 2.1 Video structure.

2.2.2. Video proprieties

Digital video can be characterized by four different properties: frame rate, frame resolution, pixel depth and bit rate.

- ***Frame rate (FR):*** is the number of frames captured per second. All analytic profiles providing the illusion of motion necessitate a minimum of 12 frames per second (fps). If the frame rate is lower than the minimum requirement, objects are not detected and tracked effectively. In an NTSC system the frame rate is 29.97 frames per second. For the PAL system, the frame rate is 25 frames per second [41, 43].

- ***Frame resolution:*** is the number of "pixels" (picture elements) within each picture frame. The more pixels, the sharper the image. As shown in fig. 2.2, there are different video formats, such as CIF(352x288), QCIF (176x144) ITU-R709 (1920x1080) and etch format have different resolution and different application. For example, the format comparable to laptop requires dimensions of around 352 x 288 pixels (CIF) [44].

- ***Pixel depth****:* Also known as Color depth, is the number of bits used to represent the color of a single pixel in video frame ,whose unit is bits per pixel (bpp) [45].

- ***Bit rate:*** is the number of bits that are conveyed or processed per unit of time. The higher the bit rate, the better the quality of the video. As an example, a YouTube video with a resolution of 854x480 has a bit rate range between 500 and 2000 Kbps [39, 41].

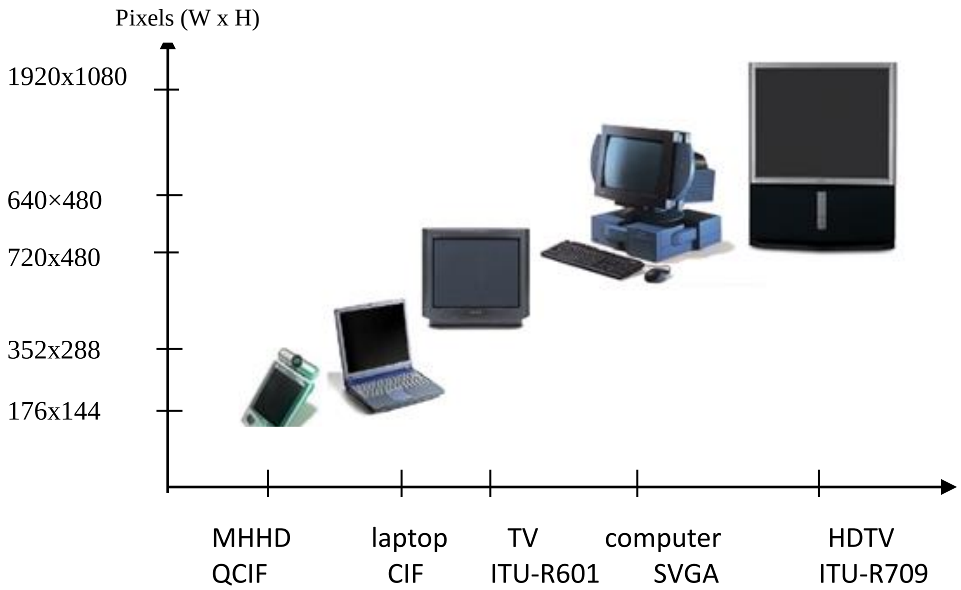


Figure 2.2 Digital video resoltion and applications [44].

2.2.3. Video Coding

Video coding is an important requirement for most video appliances such as mobile video, digital TV, internet video streaming and video conferencing. It is the process of compressing and decompressing a raw digital video sequence without retaining as much of the originals quality as possible in order to reduce the quantity of data. Video compression is performed by removing the spatial redundancies (i.e. image compression) and temporal redundancies (using motion compensation MC and motion estimation ME). The general video compression block diagram is shown in fig.2.6. The most common video coding are: MPEG (Moving Pictures Expert Group) and H26x [46, 47].

2.3. Specific challenges in video watermarking

One must say that video watermarking algorithms are more difficult to develop than those operating on images. This is essentially due to the temporal dimension, which necessitates some specific requirements. This section points out three major challenges for digital video watermarking [40, 48].

2.3.1. The robustness

In video watermarking techniques, the robustness should not only deal with common image processing attacks such as noise adding, JPEG compression, etc., but also must resist to no-hostile video processing attacks.

2.3.1.1. Image processing attacks

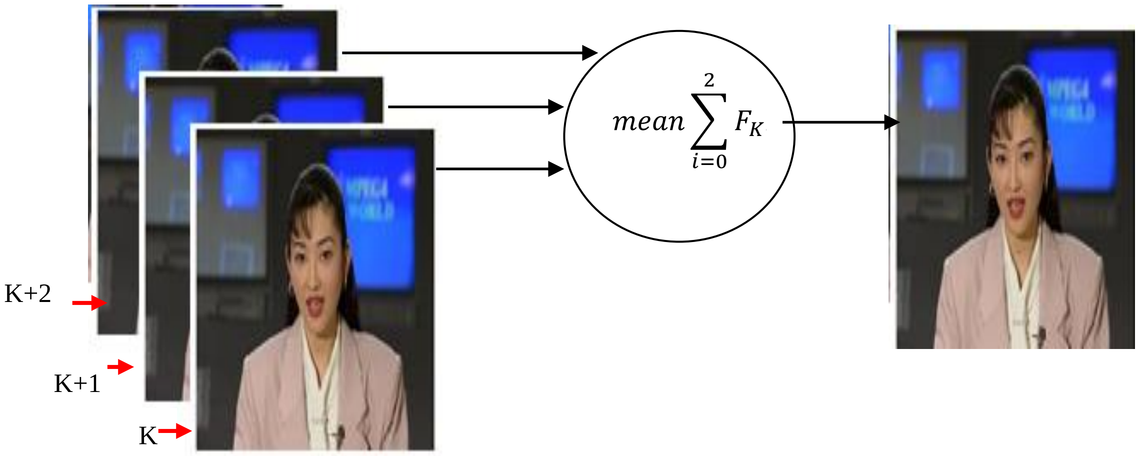
Considering a video as a sequence of images, the attacks applied to images can then be applied to the video sequences. The common image processing attacks are destruction attacks and special synchronization attacks (see chapter 1 section 1.5.2.1) However, such applications do not require robustness to spatial synchronization attacks. For example, in the case of broadcast video, watermarks need not survive rotation, scaling or high-pass filtering [1].

2.3.1.2. Temporal synchronizations attacks

Because contents in the consecutive frames of a video are almost identical, it makes the video sequences susceptible to temporal synchronization attacks. Temporal synchronization is the process of changing the temporal structures in watermarked video frames. These transformations include frame averaging (FA), frame swapping (FS) and frame dropping (FD) as shown in fig 2.3. In FA, the value of each pixel in the attacked video is obtained by averaging the pixel values of correlated video frames. However, in FD some frames from a watermarked video sequence are randomly removed and replaced with correspondent in the original video. On the other hand, FS modifies the order of some frames in a watermarked video and may not be perceptible to the human eye [49].

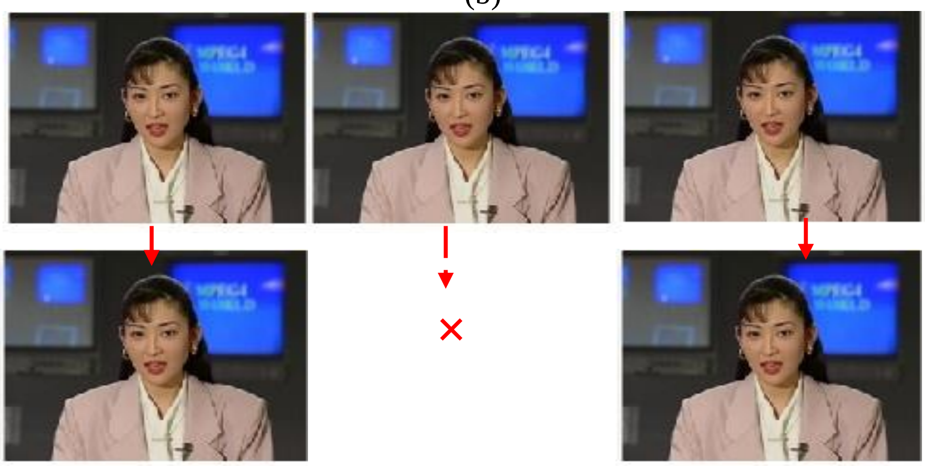
2.3.1.3. Video compression attacks

In order to reduce the storage needs, content owners often encode the video files with a different lossy compression as MPEG-1, MPEG-2, MPEG-4, H.264/AVC and H.265/HEVC. This compression may degrade the perceptual quality of the video because it removes the spatial and temporal redundancy in a video and as a result, removes the watermark even if this is not the primary goal. In case when the watermark is embed directly in a compressed video, users may transcode video (format conversion e.g. MPEG1→H.264), which may also remove the watermark.



(a)

(b)



(c)

Figure 2.3 (a) Temporal Frame Averaging (TFA): Similar video frames carrying uncorrelated watermarks are averaged to produce unwatermarked content;
(b) Temporal Frame swapping (TFS): two frames are swapped;
(c) Temporal Frame dropping (TFD): one frame is dropped.

2.3.2. The imperceptibility

The imperceptibility in video watermarking is more difficult to achieve than in image watermarking. This is due to motion of objects in video sequences, which makes the visibility of the watermark more intense; thus, a fixed watermark added to a moving object will be more perceptible than if the object is static [50]. Therefore, the temporal dimension of video should be taken into account in order to avoid distortion between frames.

2.3.3. The complexity

In image watermarking, embedding and extraction of a watermark requires just a few seconds. However, such a delay is unrealistic in the context of the video because of the significant number of frames to be processed in a video signal (25 frames/s). For this reason, the complexity of the watermarking algorithm should obviously be as low as possible in order to be implemented in real time applications with low cost. One of the most important way of achieving real-time is “Blind watermarking schemes” (i.e. which do not consider the original data) [48].

2.4. Video watermarking techniques

According to Gwenael et al. [48], video watermarking algorithms proposed in the literature can be classified into three main categories: (1) Frame-by-frame; (2) Integration of the Temporal Dimension; (3) Exploiting the Video Compression Format.

2.4.1. Frame-by-frame

Given that a digital video sequence is considered as a collection of sequential images [14], many of the images watermarking techniques that are present in the literature were extended to video. Frame by frame watermarking techniques use two major embedding strategies: (1) embed the same watermark signal into each frame, or (2) embed different (uncorrelated) watermarks into each video frame. The positive of this method is the simple implantation. But in return, such a scheme often suffers from poor robustness performance against various video processing attacks due to large amounts of data and inherent redundancy between frames (i.e. many frames are visually similar to each other) [51]. For example when the watermark to be embedded is not the same for all frames of a video sequence, the hidden data can be desynchronized with a simple operation such as frame dropping or frame averaging. In addition, these techniques are computationally intensive due to the redundancy of the same process to all frame of video. Fig. 2.4 presents the two frame-by-frame embedding strategies using additive embedding technique.

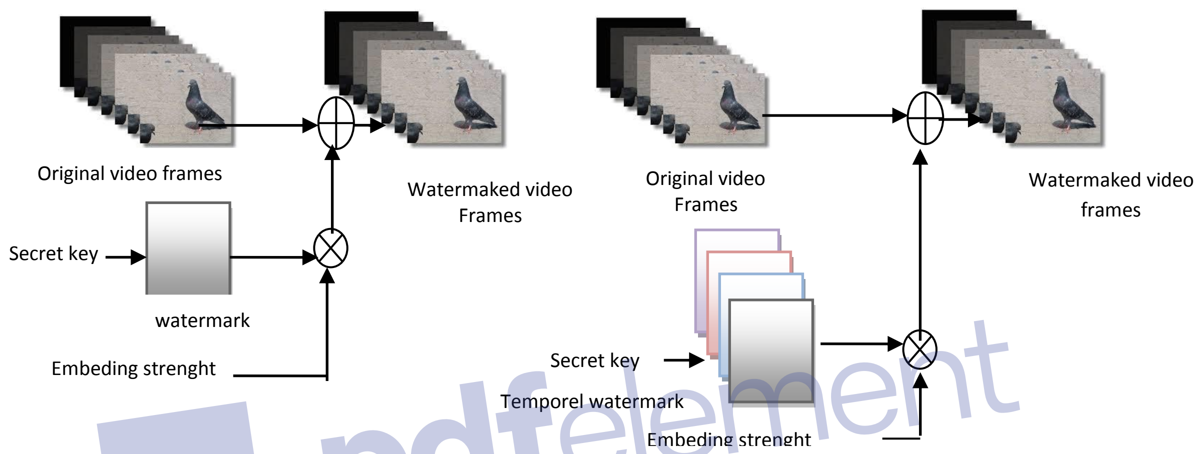


Figure 2.4 Frame by frame video watermarking: (a) the same watermark is embedded in each video frame with additive technique; (b) different watermarks are embedded in each video frame with additive technique.

2.4.2. Exploiting the Video Compression Format

In general, a watermark can be inserted into a video signal, (1) before the compression, (2) during the compression process, or (3) after the compression process [52].

2.4.2.1. Watermark embedding before the compression

The digital watermark is embed in an uncompressed digital video with the intention that the watermark signals remain present after the watermarked videos are compressed using any compression algorithm. The naive frame-by-frame techniques mentioned previously are examples of uncompressed embedding. The main advantage of this technique is that the video data can be compressed with different standards and data rates as presented in fig. 2.5, in condition that the embedded watermark has to be robust to the compression. However, it has high computational cost [53].

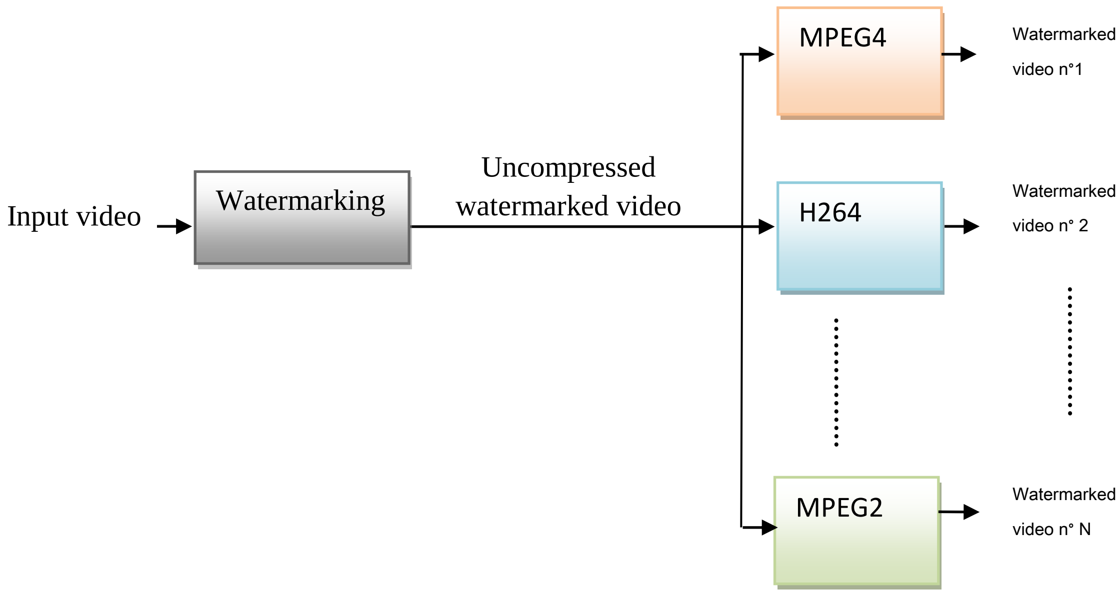


Figure 2.5 Architecture for uncompressed-domain watermarking algorithms.

2.4.2.2. Watermark embedding during the compression

In this technique, compression and watermark embedding are combined. The watermark is embedded during an encoded bit stream generated using encoders conforming to MPEG2, MPEG4, H.264…etc. standards. It should be noted that most of these techniques are employed in the discrete cosine transform (DCT) domain as presented in fig. 2.6. In these algorithms, the watermark embedding is usually performed either (1) on motion vector, (2) on VLC Code words, or (3) by modifying the DCT coefficients (see fig. 2.6) [45-56].

These techniques provide a high level of flexibility and therefore result in robust and imperceptible watermarks. However, they are not scalable to a large number of users because each user requires an individual encoding [40, 53].

2.4.2.3. Watermark embedding after the compression

In order to reduce the storage needs, content owners often re-encode the video files with a different compression ratio or to a different compression format. Embedding the watermark directly in the compressed video stream often allows real-time processing of the video. However, introducing a single change in the compressed domain may make quality preservation an issue in such methods. In addition, this technique is inherently tied to a video compression standard and may not be robust in video format conversion (transcoding) [48].

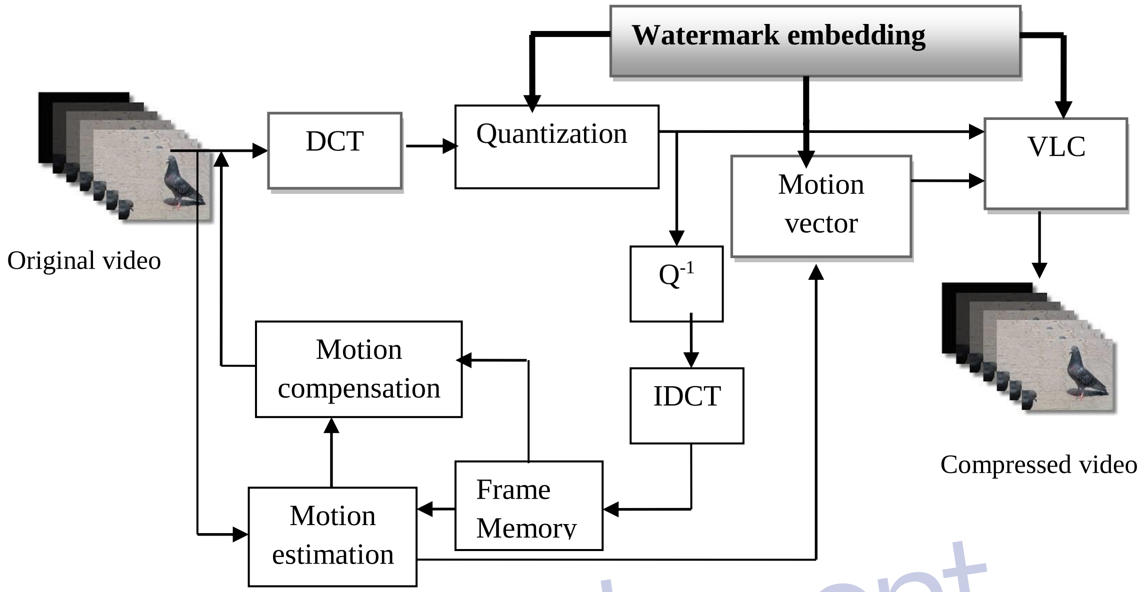


Figure 2.6 Generic diagram of embedding watermark during the compression.

2.4.3. Integration of the Temporal Dimension

The major problems of frame-by-frame video watermarking techniques are because the new temporal dimension is not satisfactorily taken into account. Solutions to this problem are proposed in the literature by (1) Video content regarded as a three-dimensional signal or (2) considering the properties of the Human Visual System (HVS).

2.4.3.1. Video content regarded as a three-dimensional signal:

Considering video as three-dimensional signals, many 3D transformations such as the 3D Discreet Fourier transform (DFT) [57], the 3D wavelet transform (3D-DWT) [58], and the 3D Discrete Cosine Transform (3D-DCT) [59] can be exploited for video watermarking by embedding the watermark into mid-frequency region to provide good robustness and high imperceptibility. The advantages of these techniques are the high robustness to temporal modification. However, the required computational cost may have reduced the research effort in this direction [48].

2.4.3.2. Considering the properties of the Human Visual System (HVS)

HVS is considered on one hand to use the temporal dimension in video watermarking. Many researchers have investigated how to reduce the visual impact of embedding a watermark within still images by considering the properties of the Human Visual System (HVS) such as frequency masking, luminance masking and contrast masking. Such studies are easily exported to video by using frame-by-frame techniques. However, these techniques don't consider the temporal properties of a video. Motion is the most important characteristic in the video, so we must create new conceptual measures that must be designed on this basis for use in a digital video watermark [48].

2.5. Conclusion

In this chapter, we have presented an overview of digital video watermarking and we can sum up that:

- ❖ Video watermarking has specific challenges that should to be considered
- ❖ Video watermarking schemes are broadly categorized in three classes: frame-by-frame, integration of the Temporal Dimension and exploiting the Video Compression Format. Each class has its own advantages and disadvantages. In practical application, different watermark classes are selected according to the different applications and the different requirements.
- ❖ Frame by frame video watermarking is the most one used because it remains relatively simple. However, it neglects the video temporal dimension.
- ❖ The majority of video watermarking methods apply in an uncompressed domain because they are generally independent of any video coding. Consequently, they are more flexible than compressed domain algorithms.

In the next two chapters, we propose two robust and imperceptible video watermarking schemas in the uncompressed domain, in which we will integrate the temporal dimension of video by considering the temporal characteristic of human visual system (HVS).

Chapter 3: A robust video watermarking technique using Spatiotemporal Just Noticeable Distortion profile

Chapter organisation

In this chapter, a robust video watermarking scheme based on spatiotemporal just noticeable distortion (ST-JND is proposed). This chapter is organized in five sections. Section 3.2 and Section 3.3 introduce the preliminaries of our scheme. Section 3.4 gives the details of the proposed video watermarking which includes two parts: the watermark embedding and extracting processes. The experimental results concerning the transparency and robustness against various attacks with comparisons are presented in Section 3.5. Finally, conclusions are given in the section 3.6.

3.1. Introduction

Most of the current video watermarking techniques are based on image watermarking and directly applied to video sequences. However, these methods are not sufficient for copyright protection in video data because video watermarking introduces a number of issues not present in image watermarking due to the temporal dimension. Human visual system (HVS) is considered on one hand to surmount this problem but many researchers have investigated how to reduce the visual impact of embedding a watermark by considering just the spatial properties of the Human Visual System (HVS) such as frequency masking, luminance masking and contrast masking without considering the temporal sensitivity of the human eye. The aim of the work presented in this chapter is to overcome these limitations [40, 60].

In this chapter, we present a robust video watermarking scheme based on spatiotemporal just noticeable distortion in DCT domain [81]. Spatiotemporal just noticeable distortions (ST-JND) which refers to the maximum distortion threshold that the HVS cannot perceive is employed in order to obtain a trade-off between watermark robustness and imperceptibility. The quantization index modulation (QIM) algorithm is used to embed the watermark in 2D-Discrete Cosine Transform (2D-DCT) domain. The dynamic quantization step size of the QIM is adaptively calculated by modeling the spatiotemporal JND with the static quantization step. Thus, this watermarking scheme is capable of relatively shaping lower injected-watermark energy onto the more sensitive regions of the host video and higher energy onto the less perceptually regions.

The remainder of this chapter is organized as follows. First, we introduce the concept of DCT and ST-JND model. Then, we explain our video watermarking method by describing the processes of watermark embedding and extraction. After that, we present the experimental results and discuss them. Finally, we concludes this chapter.

3.2. 2D-Discrect cosine transform

The 2D-DCT, which is a direct extension of the 1D-DCT, is a frequently used transform in image processing. It transforms a 2D signal, an image for example, from its spatial representation into its frequency representation, where most of the energy will be concentrated in the lower frequencies. Actually, the JPEG image compression standards rely on this propriety by discarding the higher frequency coefficients since perceptually significant components normally correspond to low and mid-band frequencies. Bearing this in mind, many of the watermarking techniques that use the 2D-DCT embed the watermark in the mid-band frequencies of the cover to gain robustness against JPEG compression [61, 62].

The 2-D DCT block calculates the two-dimensional discrete cosine transform of the input signal and the 2-D IDCT block calculates the two-dimensional inverse discrete cosine transform of the input signal. The equations for the 2D-DCT and 2D-IDCT are respectively [61]:

$$F(m,n) = \frac{2}{\sqrt{MN}} C(m)C(n) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) cos \frac{(2x+1)m\pi}{2M} cos \frac{(2x+1)n\pi}{2N} \quad (3.1)$$

$$f(x,y) = \frac{2}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} C(m)C(n)F(m,n) cos \frac{(2x+1)m\pi}{2M} cos \frac{(2x+1)n\pi}{2N} \quad (3.2)$$

where C(m), C(n)=1/$\sqrt{2}$ for m,n=0 and C(m), C(n)=1 otherwise.

3.3. Spatiotemporal Just Noticeable Distortion (JND) Profile

Just-noticeable distortion (JND) refers to the maximum distortion threshold that the HVS cannot perceive. Therefore, it is an efficient model to represent the perceptual redundancies. Here we compute a combined spatiotemporal JND estimation proposed in[63], which incorporates the spatiotemporal contrast sensitivity function (CSF), luminance adaptation and contrast masking all together. The corresponding JND can be expressed as:

$$JND(n,k,i,j) = T(n,k,i,j) \times F_M(n,k,i,j) \quad (3.3)$$

where JND is the Spatiotemporal JND threshold , T is Base Distortion Threshold and F_M is the product of the luminance adaptation factor F_{lum} and the contrast-masking factor $F_{contrast}$.

$$F_M(n,k,i,j) = F_{lum}(n,k) \times F_{contrast}(n,k,i,j) \tag{3.4}$$

where n is the index of a block in the video frame, (i, j) are the DCT coefficients indices and k is the frame position along the video sequence.

3.3.1. Base Distortion Threshold

The base threshold for a DCT subband is determined by:

$$T(n,k,i,j,)= \frac{M}{G(n,i,j,k).\emptyset_i\emptyset_j(L_{max}-L_{min})} \cdot \frac{1}{r+(1-r)cos^2\theta_{i,j}} \tag{3.5}$$

where G is the the spatiotemporal CSF, M is the number of gray levels, L_{min} and L_{max} are the maximum and minimum gray intensities on the frame, r is set to 0.6 and $\emptyset$ is the DCT normalization factor which is defined as:

$$\emptyset_m = \begin{cases} \sqrt{\frac{1}{N}}, & u=0 \\ \sqrt{\frac{2}{N}}, & u \geq 0 \end{cases} \qquad u=0,1,2\ldots\ldots N\text{-}1 \tag{3.6}$$

Here θ accounts for the effect of an arbitrary subband

$$\theta_{ij} = arcsin\left(\frac{2\rho_{i,0}.\rho_{o,j}}{\rho_{i,j}^2}\right) \tag{3.7}$$

$$\rho_{i,j} = \frac{1}{2N}\sqrt{\left(\frac{i}{\omega_x}\right)^2 + \left(\frac{j}{\omega_y}\right)^2} \tag{3.8}$$

$$\omega_x = \omega_y = 2arctan\left(\frac{1}{2R_{vd}.Pich}\right) \tag{3.9}$$

where N is the dimension of the DCT block, i and j are the spatial frequency, ω_x and ω_y are the horizontal and vertical visual angles of a pixel.

Here R_{vd} stands for the ratio of viewing distance to picture height, and Pich is the number of pixels in picture height.

3.3.2. The spatiotemporal CSF Model:

Spatiotemporal CSF refers to the spatial acuity of the HVS depending on the velocity of the image traveling across the retina (i.e. thespatiotemporal contrast sensitivity consider the relation of visibility threshold with spatial and temporal frequencies of a visual signal). The spatiotemporal CSF is given as follows:

$$G(k,n,i,j) = c_0(k_1 + k_2|log(\varepsilon.v(n,k)/3|^3.v(n,k) \\ .\left(2\pi\rho_{i,j}\right)^2.exp(-2\pi\rho_{i,j}.c_1.(\varepsilon.v(n,k) + \frac{2}{k_3}) \quad (3.10)$$

where $c_0 = 7.12,\ c_1 = 0.56,\ k_1 = 6.1,\ k_2 = 7.3,\ k_3 = 23$, and ε=1.7 .

For a real formulation of the spatiotemporal CSF, the influence of observers' eye movement needs to be considered for motion imagery and the retinal velocity in (3.10) has to be expressed by easily measurable variables from the moving images. The retinal velocity(v) presented in fig. 3.1 is the difference between the speed of an object within an image when eye movement is not involved (v_I) and the eye movement velocity (v_E) [65, 66]:

$$v(n,k) = v_I(n,k) - v_E(n,k) \quad (3.11)$$

$$v_I(n,k) = f.\sqrt{(MV_x(n,k).\theta_x)^2 + (MV_y(n,k).\theta_y)^2} \quad (3.12)$$

$$v_E(n,k) = \min[\beta.v_I(n,k) + V_{min},V_{max}] \quad (3.13)$$

where β is the gain of the spontaneous smooth-pursuit eye movement (SPEM) which indicates the efficiency of object tracking, V_{min} is the minimum eye velocity due to the drift movement, V_{max} is the maximum eye velocity before saccadic movement, MV_x and MV_y are the horizontal and vertical motion vectors for the block in the frame and f is the frame rate of video.

β is assigned with a value of 0.92, and V_{min} and V_{max} are set to 0.15 and 80.0 deg/s

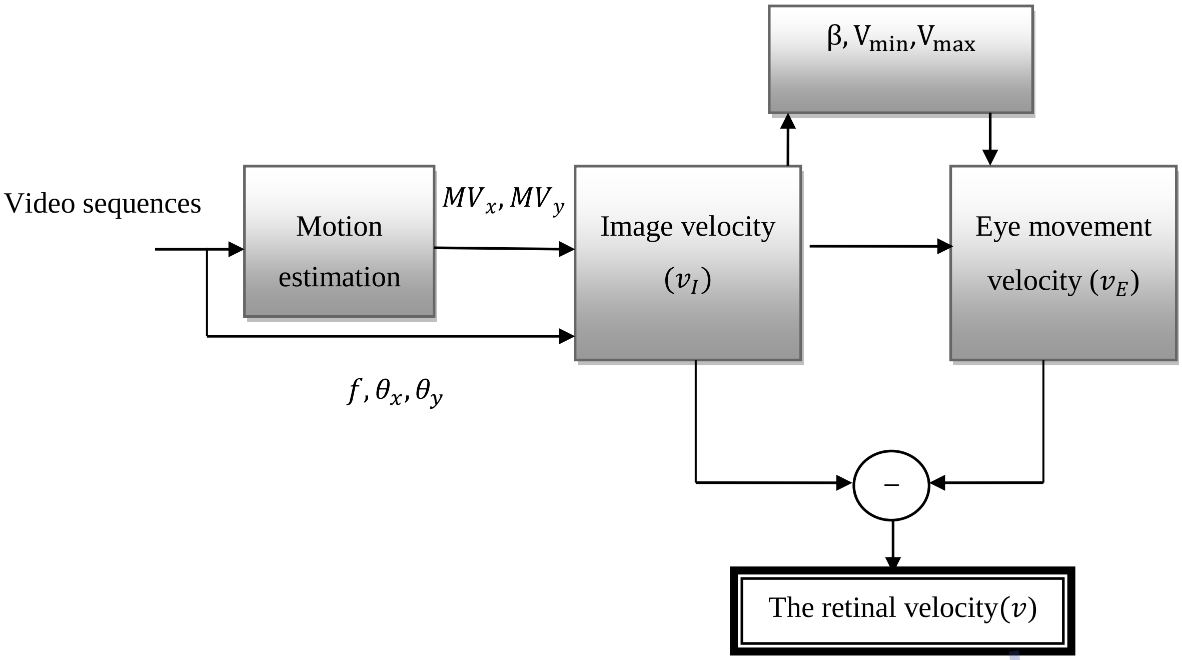


Figure 3.1 Process of the retinal velocity

3.3.3. Luminance adaptation

JND is influenced by the intensity scale of the digital image. As illustrate fig. 3.2, a higher visibility threshold T occurs in either very dark or very bright regions in an image, and a lower T occurs in regions with medium brightness. Therefore, the experimental formula of the luminance adaptation factor is shown as follows [64]:

$$F_{lum}(n,k) = \begin{cases} k_1\left(1 - \dfrac{2.C(n,0,0,k)}{M.N}\right)^{\gamma_1} + 1, & if\ C(n,0,0,k) \leq \dfrac{M.N}{2} \\ k_2\left(\dfrac{2.C(n,0,0,k)}{M.N} - 1\right)^{\gamma_2} + 1, & otherwise \end{cases} \qquad (3.14)$$

where $k_1 = 2$,k_2 =0.8,$\gamma_1 = 3$, and $\gamma_2 = 2$.

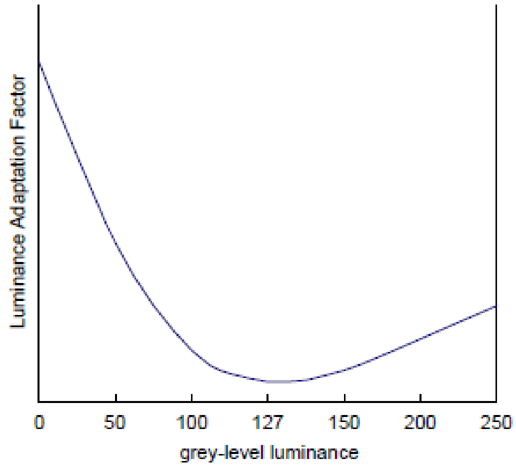


Figure 3.2 Distortion visibility as a function of background brightness [63,64, 67].

3.3.4. Contrast masking

Contrast masking is an important phenomenon in the HVS perception and is referred to as the reduction in the visibility of one visual component in the presence of another. Usually noise is less visible in the regions where texture energy is high, whilst noise is easily observed in the smooth and edge areas. An accurate block classification method where DCT blocks are classified into Texture (T), Edge (E) and Plain (P) blocks is given in [67]. In this method, DCT sub-bands, which are denoted as C(i,j), are divided in low (L), middle (M) and high (H) frequencies as presented in fig.3.3.

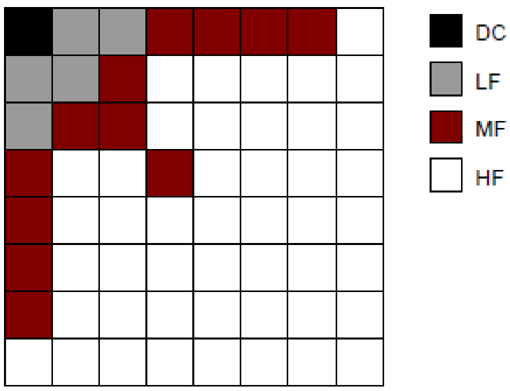


Figure 3.3 Block classification scheme for a DCT block [64]

The contrast masking is calculated as follows:

$$F_{contrast}(n1,n2,i,j) = \begin{cases} \xi(n1,n2) & for\ (i,j)\epsilon LF \cup MF\ in\ Edge\ block \\ \xi(n1,n2)\cdot max\left\{1,\left(\frac{C(n1,n2,i,j)}{T(n,i,j,k).F_{lum}(n,k)}\right)^{\varepsilon}\right\} & otherwise \end{cases} \tag{3.15}$$

where ξ is the inter-band masking and can be represented as:

$$\xi(n1,n2) = \begin{cases} 1+\left[\frac{(TexE(n1,n2)-\mu_2)}{2\mu_{3}-\mu_2}\right].\delta_1 & forTEXTUREblock \\ \delta_1 & for\ EDGE\ block\ and\ L+M>400 \\ \delta_1 & for\ EDGE\ block\ and\ L+M\leq 400 \\ \delta_1 & for\ PLAIN\ block \end{cases} \tag{3.16}$$

3.4. The proposed watermarking scheme

In this section, we describe the proposed video watermarking technique using the spatiotemporal Just Noticeable Distortion in DCT domain.

3.4.1. Watermark embedding process

Figure 3.4 shows the watermark embedding process, which can be summarized as follows:

(1) The original video is partitioned into groups of k frames and every frame of the group is converted from the RGB color space to the YCbCr one.

(2) Every luminance frame is decomposed into blocks of size [u × u]. Then the2D-DCT is applied to each block.

(3) Calculate spatial Spatiotemporal Just Noticeable of DCT sub-bands C(1,2) for all DCT blocks of a video frame by performing the process described in section 3.3.

(4) The QIM embedding process is applied to embed a watermark bit W_t into the sub-band C(1,2)according to the following equation[re]:

$$C_W(n,1,2,k) = \begin{cases} C(n,1,2,k).\left\lfloor\frac{|C(n,1,2,k)|}{2D}\right\rfloor.2D & if\ W_t=0 \\ C(n,1,2,k).\left(\left\lfloor\frac{|C(n,1,2,k)|}{2D}\right\rfloor.2D+D\right) & if\ W_t=1 \end{cases} \tag{3.17}$$

where D represents the dynamic quantization step size which is defined as:

$$D = Q + Q\left(\frac{JND(n,1,2,k)}{10}\right) \quad (3.18)$$

where Q represents the static quantization step size, (Q step) determined through the trade-off between robustness and imperceptibility, which is described later.

(5) To build the watermarked luminance, the 2D-IDCT is applied

(6) Reconstruction of the watermarked video frame is done using the watermarked luminance part and the original frame chrominance parts Cb and Cr, by conversion from the YCbCr into the RGB color space.

3.4.2. Watermark extraction process

The watermark extraction process is shown in fig. 3.5 and is described as follows:

(1) The watermarked video is divided into groups of k frames.

(2) Every frame of the group is converted from the RGB color space to the YCbCr one.

(3) The watermarked luminance is decomposed into blocks of size [u × u]. Then, the 2D-DCT is applied to each block.

(4) Segment the watermarked luminance frames into non-overlapping blocks of size [u × u] and apply the 2D-DCT transform to each block.

(5) Extract the watermark using the QIM extraction process:

$$\widetilde{W} = \begin{cases} 0, & if \lfloor C_w(n,2,1,k)/D \rceil even \\ 1, & if \lfloor C_w(n,2,1,k)/D \rceil odd \end{cases} \quad (3.19)$$

where C_w is the watermarked coefficients of DCT coefficient and D represents the dynamic quantization step size which was calculated in the embedding process and stored to be used as a secret key in the extraction process.

(6) Since a video clip contains several frames, in which the same watermark is embedded, we calculate the final recovered watermark by averaging the watermarks extracted from these different frames.

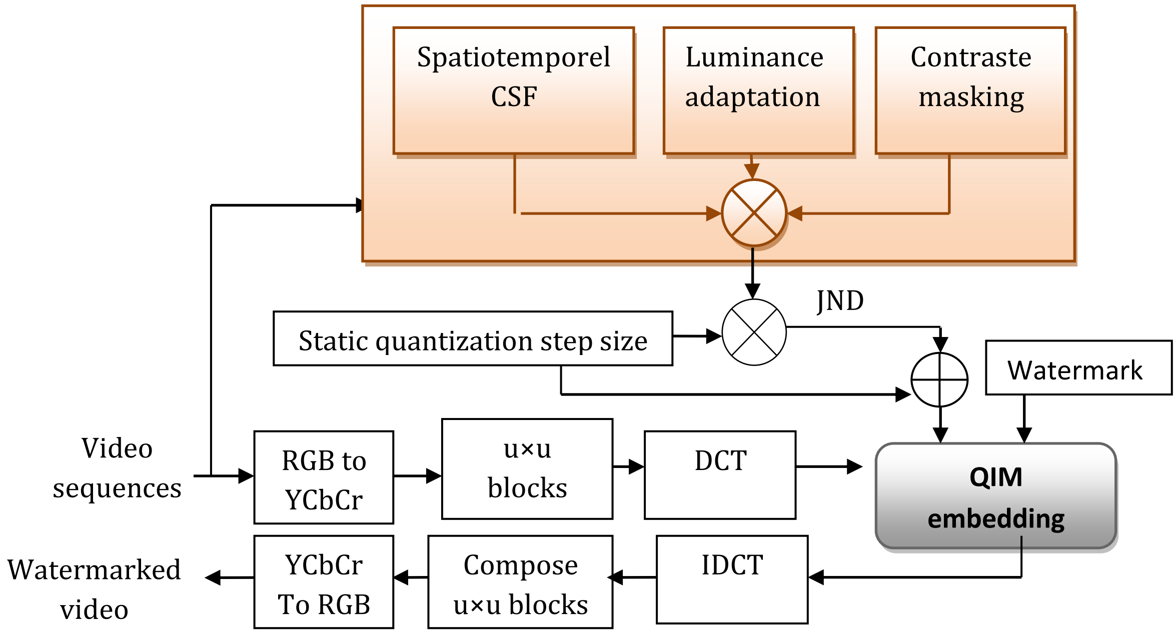


Figure 3.4 Watermark embedding process.

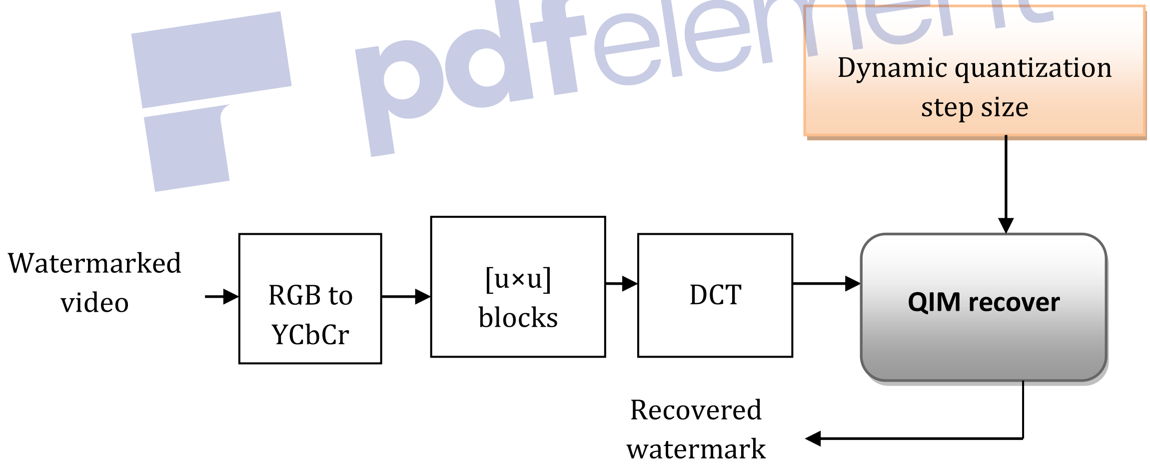


Figure 3.5 Watermark extraction process.

3.5. Experimental results

In this section, we evaluate the performances of the proposed video watermarking technique in terms of imperceptibility and robustness. We use QCIF (352 × 288) videos sequences in RGB uncompressed avi format and with frame rate of 30 fps [68]. The binary image used as a watermark is shown in fig. 3.6. The resolution of the image depends on the frames resolution and blocks' size. The experiments are approved out to demonstrate the effect of the proposed watermarking scheme using spatiotemporal Just Noticeable Distortion (ST-JND) and to compare it to a similar without JND and with spatial JND (S-JND) proposed in [64].

3.5.1. The choice of static quantization step size Q

Determining an appropriate static quantization step size Q is crucial to obtain a good performance of the proposed scheme from the watermark imperceptibility and robustness points of view. First, to evaluate the relationship between static quantization step size Q and watermark imperceptibility, the PSNR and MSSIM are obtained varying this value. In addition, the relationship between this factor and watermark robustness is evaluated using NC and BER of the extracted watermark without any attacks. From table 3.1, we determine that the most adequate value of Q is equal to 12 because using this value makes the PSNR and MSSIM between the watermarked video and the original one respectively equal to 44.27 dB and 0.9878. In addition, the NC and BER between the original watermark and extracted watermark are respectively equal to 1 and 0.

Table 3.1 PSNR, MSSIM, NC and BER for different quantization step Q.

Q	PSNR	MMSSIM	NC	BER
3	55.2225	0.9989	0.6545	0.1542
6	49.8681	0.9964	0.9963	0.0013
12	**44.2778**	**0.9878**	**1**	**0**
15	42.4499	0.9822	1	0
18	41.0068	0.9761	1	0

3.5.2. Imperceptibility tests

The imperceptibility of the watermark is estimated by measuring the PSNR (Peak Signal to Noise Ratio) and Mean Structure Similarity Index Measure (MSSIM) which are calculated using the luminance space Y of the original and watermarked frames. Fig. 3.8 shows the PSNR and MSSIM of all frames of the video after watermark embedding in the absence of any attacks. The watermarked videos without considering HVS criteria and with considering S-JND have larger PSNR and MSSIM than those with the ST-JND, because it takes advantage of the spatial and temporal HVS characteristics. We can see no obvious degradation in fig. 3.7 where PSNR is greater than 44dB and MMSIM is near to one.

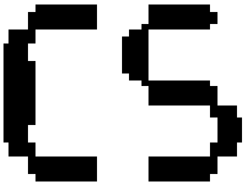


Figure 3.6 Binary image used as watermark

Figure 3.7 Watermarked video frames from 'foreman' video sequences.

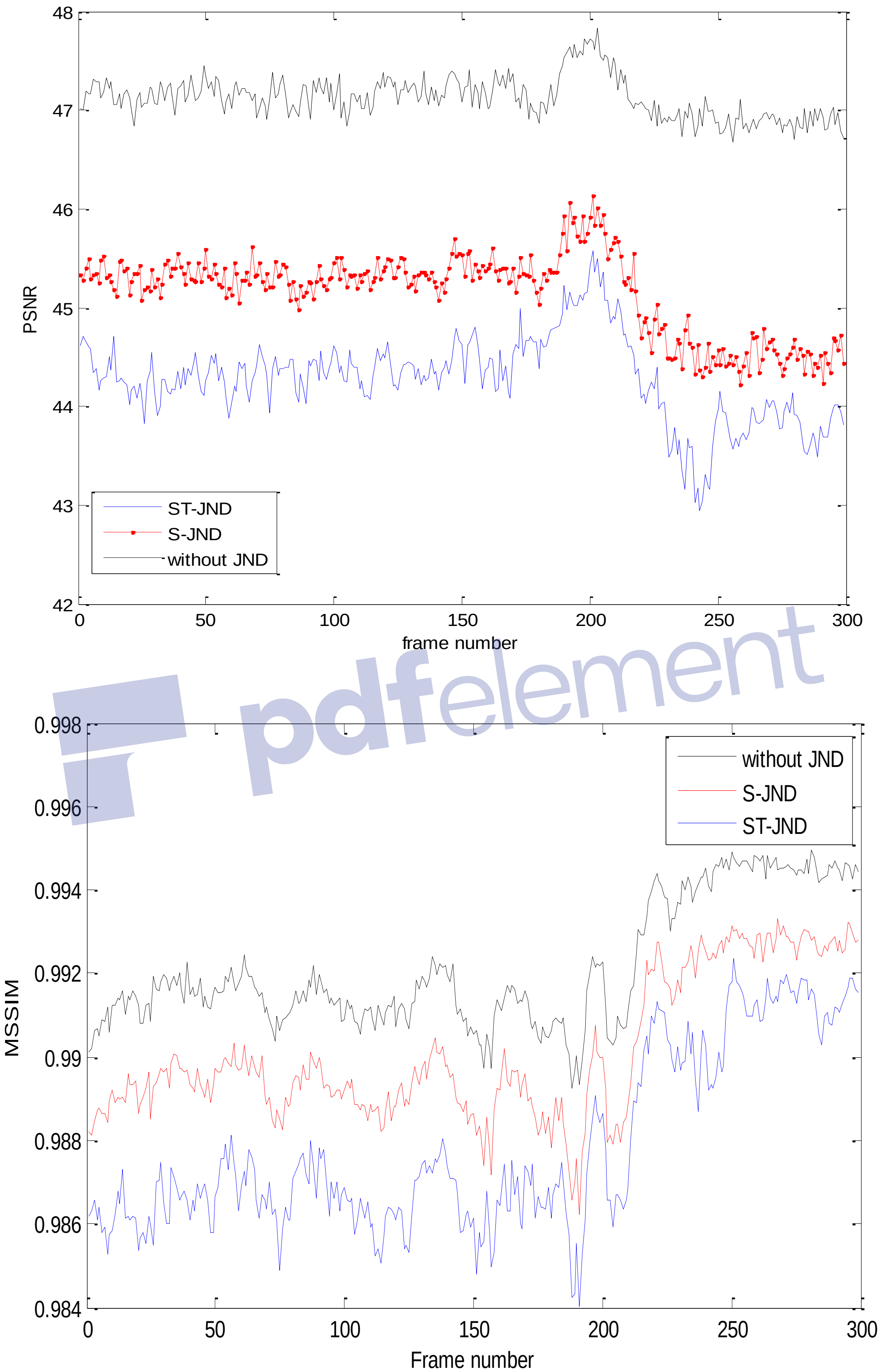


Figure 3.8 The PSNR and MSSIM of watermarked “foreman” video sequences.

3.5.3. Robustness tests

The robustness for any watermarking system is a very important requirement. To verify it, we apply to the watermarked video various types of attacks and we use the Normalized Coefficient (NC) and the Bit Error Rate (BER) to compare the similarities between the original watermark W and the extracted watermark $\widehat{W}$.

As we cited in chapter 2, video watermarking attacks are divided into three categories: image processing attacks, temporal synchronization attacks and video compression attacks.

3.5.3.1. Image processing attacks

The common image processing used as attacks are**:**

- *JPEG compression:* The watermarked video is compressed with different quality factors ranging from 20 to 100 and experimental results are presented in fig. 3.9.
- *Adding a noise:* Two kinds of noises, Gaussian noise and Salt & pepper noise with mean 0 and standard deviation δ are added to the watermarked video. The experimental results are presented in table 3.2.
- *Filtering:* 3x3 median filters and 3x3 Gaussian filters are applied separately to the watermarked videos. The experimental results are presented in table 3.3.

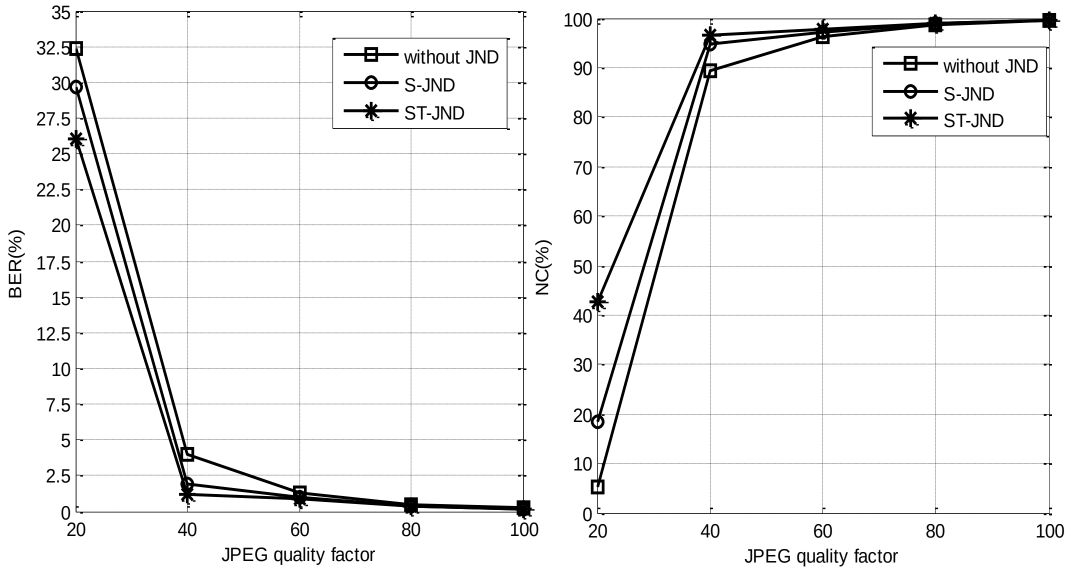


Figure 3.9 Robustness versus JPEG compression.

Table 3.2 Robustness versus adding a noise.

Methode	Gaussian noise			Salt and pepper noise		
	δ	NC	BER	ρ	NC	BER
Without JND	0.3	0.9293	0.0255	0.3	0.9629	0.0134
	0.6	0.7726	0.0873	0.6	0.9302	0.0255
	1	0.5870	0.1710	1	0.8914	0.0404
	1.2	0.5145	0.2067	1.2	0.8745	0.0470
S-JND	0.3	0.9565	0.0155	0.3	0.9694	0.0111
	0.6	0.8530	0.0546	0.6	0.9420	0.0212
	1	0.6977	0.1196	1	0.9073	0.0343
	1.2	0.6291	0.1509	1.2	0.8910	0.0406
ST-JND	0.3	0.9743	0.0091	0.3	0.9745	0.0092
	0.6	0.9079	0.0335	0.6	0.9512	0.0178
	1	0.7827	0.0833	1	0.9211	0.0291
	1.2	0.7199	0.1101	1.2	0.9059	0.0349

Table 3.3 Robustness versus Filtering

Method	Median filter 3×3		Gaussian low pass filter 3×3	
	NC	BER	NC	BER
Without JND	0.4257	0.2332	0.9305	0.0254
S-JND	0.4766	0.2064	0.9695	0.0110
ST-JND	0.5147	0.1880	0.9726	0.0099

3.5.3.2. Temporal synchronization attacks

The video sequences are susceptible to temporal synchronization attacks such as frame averaging and frame dropping

- Frame averaging (FA): In frame averaging, we replace selected watermarked frames by the average of their previous, current and next frames. As presented in table 3.4, watermarked video is averaged for various averaging rates and then we tried to extract the watermark.
- Frame dropping (FD): In frame dropping, selected watermarked frames are replaced by their corresponding original frames. Table 3.4 shows the average NC and BER values given at different frame dropping rate.

3.5.3.3. Video compression attacks

Video compression is a fundamental attack in video watermarking that should be verified as video sequences are stored and transmitted in compressed format. Here we use a tool for video processing named VirtualDub to compress the videos sequences with two different lossy compressions [78]: H264 and MPEG4 coding with a different bit rate and the experimental results are presented in table 3.5.

Table 3.4 Robustness versus temporal frame synchronization

Methode	FD			FA		
	FD rate	NC	BER	FA Rate	NC	BER
Without JND	50%	0.5122	0.2355	50%	0.5246	0.1441
	25%	0.7558	0.1178	25%	0.7613	0.0726
S-JND	50%	0.5127	0.2344	50%	0.5274	0.1440
	25%	0.7570	0.1171	25%	0.7639	0.0720
ST-JND	50%	0.5149	0.2315	50%	0.5409	0.1342
	25%	0.7596	0.1155	25%	0.7712	0.0668

Table 3.5 Robustness versus video compression.

Methode	MPEG4			H264		
	Bitrate	NC	BER	Bitrate	NC	BER
Without JND	2Mbps	0.5071	0.1866	1Mbps	0.6060	0.1617
	3Mbps	0.6202	0.1428	2Mbps	0.8597	0.0543
	4Mbps	0.7428	0.0966	3Mbps	0.9421	0.0213
S-JND	2Mbps	0.5920	0.1495	1Mbps	0.6306	0.1506
	3Mbps	0.6940	0.1120	2Mbps	0.8997	0.0382
	4Mbps	0.7880	0.0775	3Mbps	0.9623	0.0138
ST-JND	2Mbps	0.6738	0.1170	1Mbps	0.7238	0.1099
	3Mbps	0.7596	0.0863	2Mbps	0.9308	0.0259
	4Mbps	0.8532	0.0528	3Mbps	0.9766	0.0085

3.5.4. Discussion

The experimental results presented in tables 3.1 to 3.5 and illustrated in fig .3.9 show that watermarking scheme using ST-JND is robust to almost all attacks (JPEG compression, adding noise, filtering, frame averaging, frame dropping, MPEG4 compression, and H264 compression) and achieves better performance than the corresponding without using JND and with using just S-JND approach.

3.6. Conclusion

In this chapter, we presented a semi blind video watermarking scheme using spatiotemporal just noticeable distortion (ST-JND) in the DCT domain. ST-JND is used in order to improve robustness in the detection, and at the same time obtain a good quality of the watermarked video. The watermark sequence was embedded and detected using a quantization index modulation (QIM) algorithm with adaptive step size, which is calculated using spatiotemporal just noticeable distortion (JND). The experimental results show that the proposed scheme not only improves the perceptual quality, but also was robust against various attacks. Compared to a non-HVS approach and S-JND approach, the ST-JND based video watermarking algorithm achieves better performance.

Chapter 4: A novel blind and robust video watermarking technique in fast motion frames based on SVD and MR-SVD

Chapter organization

In this chapter, a novel and efficient digital video watermarking technique based on the Singular Value Decomposition performed in the Multiresolution Singular Value Decomposition domain is proposed. This chapter is organized in five sections. Section 4.2 introduces the preliminaries of our scheme. Section 4.3 gives the details of the proposed video watermarking which include four parts: the fast motion frames extraction, the watermark pre-processing and the watermark embedding and extracting processes. The experimental results concerning the transparency and robustness against various attacks with comparisons with other previous algorithms found in the literature are presented in Section 4.4. Finally, conclusions are given in the last section.

4.1. Introduction

Given that a digital video sequence is considered basically as a collection of sequential images [14], many of the image watermarking techniques that are present in the literature were extended to video [6, 9, 15, 16] , as they embed the watermark in all frames of the video sequences. Thus, these algorithms are robust to frame dropping and frame swapping, but in return they are time consuming and also affect the perceptibly of the video quality. To solve this problem of frame by frame embedding, an answer to the following key question should be found: What are the preferred frames to host the watermark without degrading the visual quality of the watermarked video while maintaining the robustness reasonably unaffected? The answer is to adaptively embed the watermark in selected frames.

In this direction, very few video watermarking schemes were considered. Tabassum and Islam [69] proposed a digital video watermarking technique based on identical frame extraction. In this method, the host video is initially divided into video shots. Then from each video shot one video frame called identical frame is selected for watermark embedding. In [70], Agilandeeswari and Ganesan developed an approach for video watermarking using SVD and DWT. In their algorithm they extracted the non motion frames from the video using histogram difference based scene change detection algorithm, and then they embedded in them the same watermark. However, the problem in these techniques is the small number of watermarked video frames. So if those embedded frames are lost, the scheme becomes unreliable. In [71], Jiang Xuemei et al. developed an approach for video watermarking based on shot segmentation and block classification. They selected the frames with the biggest luminance value in every shot to be the host frames. The watermark signal is cropped into small watermarks according to the number of host frames in the host video. These small watermarks are then respectively embedded into the different selected host frames. In addition, Chetan et al. [72] proposed a robust video watermarking scheme based on scene changes which embed different parts of a single watermark into different scenes of a video. These frames are selected based on scene change detection. In these two last cited techniques, if one watermarked frame is lost, the watermark cannot be extracted completely.

Chapter 4: A novel blind and robust video watermarking technique in fast motion frames based on SVD and MR-SVD

In this chapter, we present a novel video watermarking scheme in fast motion frames using Singular Value Decomposition in the Multi Resolution Singular Value Decomposition (MR-SVD) domain [82]. The main contribution in this work is as follows:

(i) In order to avoid embedding the watermark in all the frames of the video sequences, we first segment the video into temporally stationary signals using shot boundary detection. Then, from each shot we choose the frames with big motion energy (fast motion frames) to embed the watermark. This is done because the human visual system (HVS) cannot notice the details of fast moving regions [64] and thus the perceptual invisibility of the watermark is guaranteed.

(ii) Because of their relevant advantages, we use a combination of the SVD and MR-SVD transforms. SVD, with its attractive mathematical properties, has been broadly applied in image compression and image watermarking and proved to be an efficient technique in both domains [73]. Most existing SVD based watermarking techniques combine the SVD transform with the multi resolution 2D-DWT [9,10], as they showed to be reliable and provide high robustness and better perceptual image quality. However, one of the drawbacks of the DWT is its huge resources consummation and high computation cost due to the convolutions carried out in each of the filters. To overcome this issue, Kakarala and Ogunbona [74] proposed the idea of the MR-SVD which performs multiresolution decomposition similar to that of the DWT, has perfect reconstruction, and above all is a matrix based operation like the SVD. Therefore a hybrid SVD MR-SVD watermarking technique is based only on matrix operations which make it well suited for real time applications and simple for hardware implementation.

(iii) Also, we embed watermark information by quantization index modulation (QIM) which has been shown to be host interference free, and provably optimal in terms of channel capacity under an additive white Gaussian noise attack. Furthermore, the extraction procedure in QIM is blind which makes it suitable for robust watermarking [33].

(iv) Moreover, to embed the watermark in a secure manner, we encrypt the watermark using a logistic map based encryption [75].

The remainder of this paper is organized as follows: First, we introduce the preliminaries of our scheme. Then we will give the details of the proposed video watermarking which include four parts: the fast motion frames extraction, the watermark pre-processing and the watermark embedding and extracting processes. After that, we will present the experimental results concerning the transparency and robustness against various attacks with comparisons with other previous algorithms and finally we will give conclusions to this chapter.

4.2. Singular Value Decomposition

In linear algebra, Singular Value Decomposition (SVD) is a numerical technique that decomposes a matrix into three matrices with valuable properties when applied in digital image processing [73].

If a matrix A represents for example an image of size N ×N, then the SVD of A is given by:

$$A = USV^T = \begin{pmatrix} u_{11} & \cdots & u_{1N} \\ \vdots & \ddots & \vdots \\ u_{N1} & \cdots & u_{NN} \end{pmatrix} \begin{pmatrix} s_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_N \end{pmatrix} \begin{pmatrix} v_{11} & \cdots & v_{1N} \\ \vdots & \ddots & \vdots \\ v_{N1} & \cdots & v_{NN} \end{pmatrix} \tag{4.1}$$

where U and V are orthogonal matrices representing respectively, the horizontal and vertical details (edges) of the image and S is a diagonal matrix, where the diagonal elements $S_i(1 \leq i \leq N)$ with $S_1 \geq S_2 \geq \ldots \geq S_N$, are the singular values (SVs) of A.

Two main properties, related to the SVs, make the SVD appropriate for watermarking when the matrix S is utilized [8, 73].

- ❖ The energy content (luminance) of the image A is located in the SVs.
- ❖ The SVs have very good stability, i.e., a small perturbation added (a watermark for example) to the image does not change significantly the SVs.

4.3. Multiresolution Singular Value Decomposition(MR-SVD)

As stated in the introduction, the MR-SVD, initially introduced in [74], is a matrix based operation.

4.3.1. 1-D Multiresolution Singular Value Decomposition

Let $X = [x(1) \ \dots \ x(N)]$ represent a finite extent 1-D signal and assume that N is divisible by 2L for some L ≥1. Let the data matrix at the first level, denoted X1, be constructed with its top and bottom rows containing respectively the odd-numbered and even-numbered samples of X:

$$X_1 = \begin{pmatrix} x(1) & x(3) & \dots & x(N-1) \\ x(2) & x(4) & \dots & x(N) \end{pmatrix} \tag{4.2}$$

The corresponding centered matrix is $\bar{X}_1 = X_1(I_N - \left(\frac{1}{N}\right) e_N e_N^T)$, where I_N is the identity, and e_N is the vector containing all ones.

Let U1 be the eigenvector matrix bringing the scatter matrix $T1 = \bar{X}_1 \bar{X}_1^T$ into diagonal form:

$$U_1^T T_1 U_1 = S_1^2 \tag{4.3}$$

where $S_1^2 = \mathrm{diag}\{s_1(1)^2, s_1(2)^2\}$ contains the squares of the two singular values, with $s_1(1) \geq s_1(2)$.

Now let: $\hat{X}_1 = U_1^T \bar{X}_1$

The top row of $\widehat{X}_1$, denoted by $\emptyset_1 = \hat{X}_1(1,:)$, corresponds to the largest eigenvalue and represents the approximation component. The bottom row of $\widehat{X}_1$, designated by $\varphi_1 = \hat{X}_1(2,:)$, corresponds to the smallest eigenvalue and contains the detail component. The successive levels of decomposition repeat the procedure described above by placing the approximation component $\emptyset_1$ in place of X. Hence the MR-SVD can be written as:

$$X \rightarrow \{\emptyset_L, \{\varphi_l\}_{l=1}^{L}, \{U_l\}_{l=1}^{L}\} \tag{4.4}$$

where L is the desired level of decomposition.

4.3.2. 2-D Multiresolution Singular Value Decomposition

We briefly describe here the 2-D MR-SVD. The first level decomposition of the image proceeds as follows. The M × N image X is divided into non-overlapping 2×2 blocks and each block is arranged into a 4×1 vector by stacking columns to form the data matrix X_1. The Eigen-decomposition of the 4×4 scatter matrix is:

$$T_1 = X_1 X_1^T = U_1 S_1^2 U_1^T \quad (4.5)$$

Let

$$\hat{X}_1 = U_1^T \bar{X}_1 \quad (4.6)$$

The top row of the resulting matrix $\hat{X}_1(1,:)$ is rearranged to form an M/2 × N/2 matrix which is considered as the smooth (approximation) components of the image. The remaining rows $\hat{X}_1(2,:)$, $\hat{X}_1(3,:)$ and $\hat{X}_1(4,:)$, contain the detail components, which are denoted φ_l^1, φ_l^2, φ_l^3 respectively. The complete transform can be represented as follows:

$$X \rightarrow \{\emptyset_L, \{\varphi_l^1, \varphi_l^2, \varphi_l^3\}_{l=1}^{L}, \{U_l\}_{l=1}^{L} \quad (4.7)$$

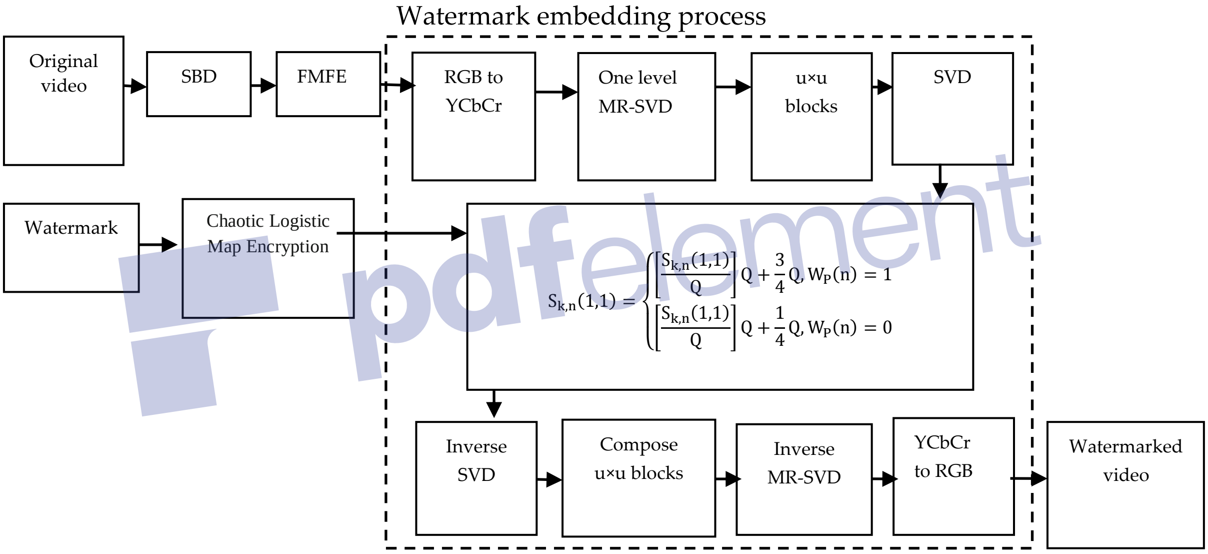
The original image X can be reconstructed from the right hand side, since the steps are reversible. As an example, the one level MR-SVD decomposition of the video frame “Foreman” is depicted in fig. 4.1.

4.4. Proposed method

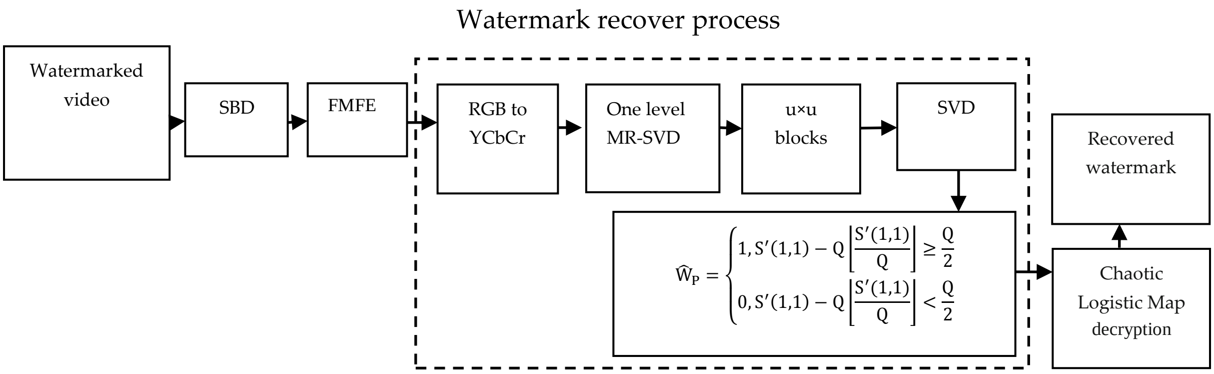
Our proposed algorithm encompasses four consecutive parts:

- ❖ Fast motion frames where to embed the watermark
- ❖ Watermark pre-processing
- ❖ Watermark embedding process
- ❖ Watermark extraction process

Figure 4.1 Original video frame Foreman image and its 1-level MR-SVD.



(a)



(b)

Figure 4.2 Proposed video watermarking technique: (a) embedding process and (b) recover process.

4.4.1. Fast motion frames extraction for embedding (FMFE)

In order to increase the quality of the watermarked video, the proposed system exploits the characteristics of the human visual System (HVS) to select frames in which the watermark is embedded effectively. Because the HVS is less sensitive to errors in regions with great motion, we select the frames that have big motion energy to be the host frames [64]. In our scheme, to extract the fast motion frames, the cover video V is first converted into individual frames, then shot boundaries (SBD) are detected using the algorithm proposed in [76] to obtain temporally stationary signals because frames within the same shot have a strong correlation. Afterwards, we measure the average motion energy of each frame using the mean magnitudes of motion vectors (MMMV). Let $S = \{F_1, F_2, \ldots, F_k\}$ be a shot of length k, where F_k, $(k = 1, 2, \ldots, k)$, represents the k^{th} frame in the shot. The MMMV of the frame can be calculated as follows:

$$MMMV(F_j) = \frac{1}{N} \sum_{\substack{1 \le i \le N \\ 1 \le j \le k}} MV_x^2(i,j) + MV_y^2(i,j) \tag{4.8}$$

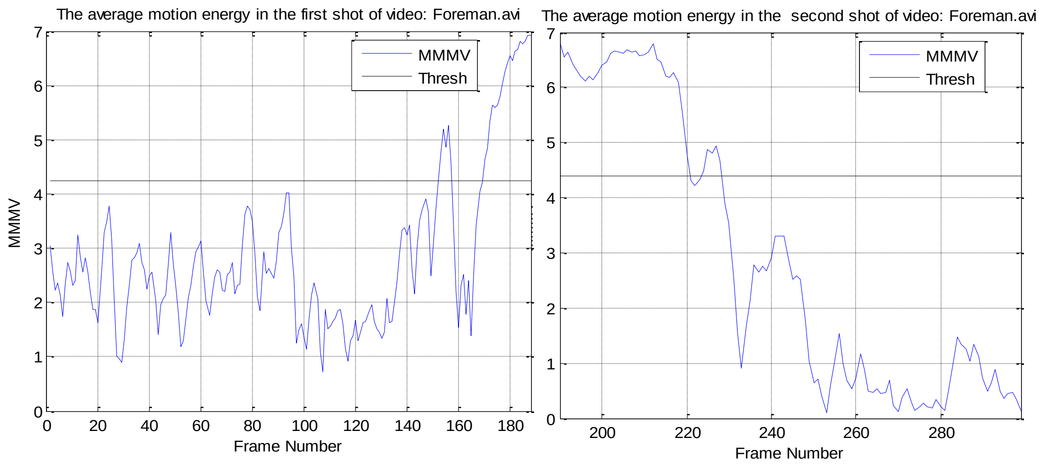
where N is the number of macro blocks in a frame and MV_x and MV_y represent the components of the motion vector (MV) in respectively the X axis direction and the Y axis direction.

$$Thresh = \frac{\alpha}{k} \sum_{j=1}^{k} MMMV(F_j)) \tag{4.9}$$

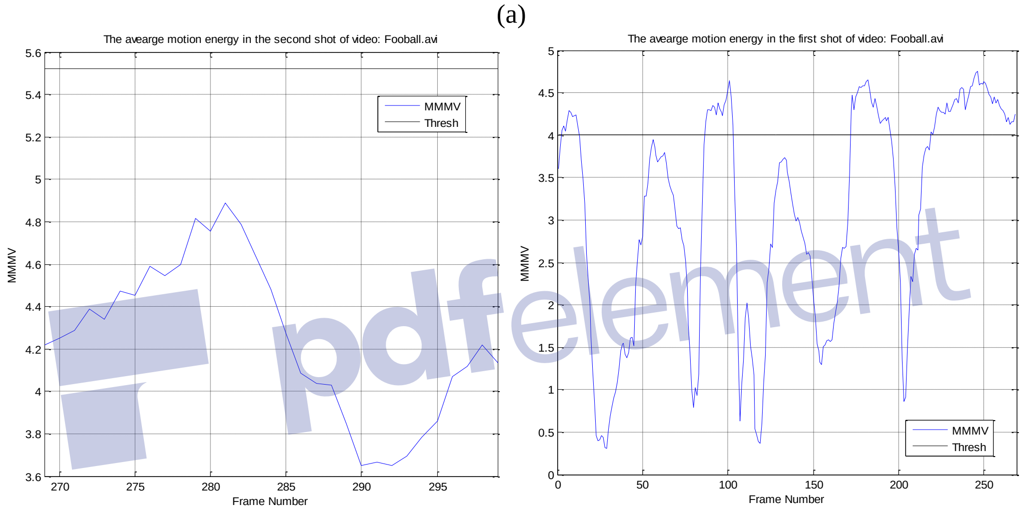
The above threshold is used as a decision rule to distinguish between fast and slow motion frames. Furthermore, we use a constant α to adjust the decision rule:

$$\begin{cases} F_j is\ faste\ motion\ frame, & if\ MMMV(F_j) > thresh \\ F_j is\ no\ motionf\ rame, & if\ MMMV(F_j) \le thresh \end{cases} \tag{4.10}$$

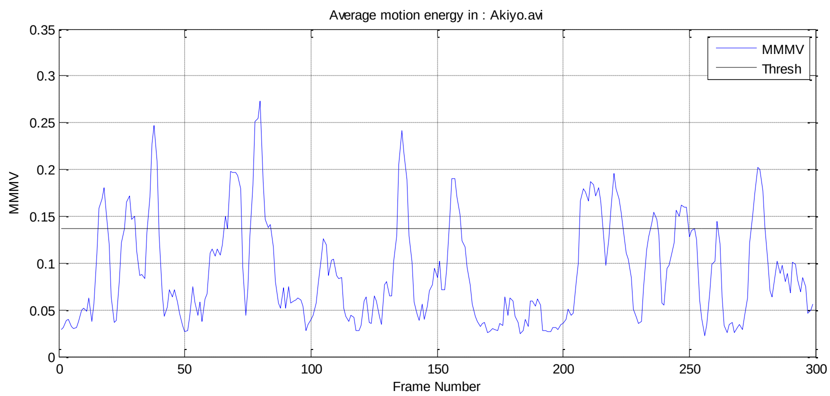
Figure 4.3 shows the average motion energy of some frames in Foreman, Football and Akiyo video shots. In the case of the Forman video, we see that the fast motion frames take place between the frames 150 and 228.



(a)



(b)



(c)

Figure 4.3 Average motion energy in (a) Foreman, (b) Football and (c) Akiyo videos.

4.4.2. Watermark preprocessing

In order to improve the security of the proposed algorithm, the binary watermark should be first pre-processed before embedded. Here, the watermark is scrambled by using the chaotic logistic map which is determined by the following equation [75, 77]:

$$X_{n+1} = \lambda X_n(1 - X_n) \tag{4.11}$$

where $3.75 < \lambda < 4$ is the system parameter. The initial value $X_0 \in [0, 1]$ is adopted as a key. Then, the binary image logo or signature $W(n)$ is scrambled by $X(n)$ with the following rule:

$$W_P\ (n) = W(n) \oplus X(n), 0 \leq n \leq N \tag{4.12}$$

with N being the total number of bits in the watermark and $\oplus$ is the binary Exclusive Or (XOR) operation.

4.4.3. Watermark embedding process

In this section we describe the proposed video watermarking scheme. Fig. 4.2 (a) shows the block diagram of the proposed video watermark embedding procedure, which is described as follows:

(1) The fast motion frames are extracted from the original color video. Only these frames are watermarked. This makes the watermarked video quality good because the watermark is not embedded in all the frames as it is done in other watermarking schemes [6, 9].

(2) Every fast motion frame is converted from the RGB to the YCbCrcolor space.

(3) Every luminance frame is transformed with the 1-Level MR-SVD decomposition to get approximation and detail components {Ø, φ 1, φ2, φ3}.

(4) The approximation component Φ is decomposed into blocks of size [u × u]. The SVD is applied to each block of the approximation component Ø (i.e. the low frequency sub band), which contains the major video frame energy.

$$\emptyset_{k,n} = U_{k,n} \times S_{k,n} \times V_{k,n}^{T} \quad (4.13)$$

where k is the fast motion frame index and n is the location of the block. We use the SVD due to its suitable properties discussed earlier.

(5) The watermark is encrypted using chaotic logistic map.

(6) In order to guarantee robustness to our watermarking scheme, the watermarking system inserts the watermark bits in the largest singular value using the QIM method as:

$$S'_{k,n}(1,1) = \begin{cases} \left[\frac{S_{k,n}(1,1)}{Q}\right] Q + \frac{3}{4} Q, & W_P(n) = 1 \\ \left[\frac{S_{k,n}(1,1)}{Q}\right] Q + \frac{1}{4} Q, & W_P(n) = 0 \end{cases} \quad (4.14)$$

where Q is the quantization step and [•] stands for the rounding operation.

(7) Inverse SVD transformation is conducted to obtain the watermarked block:

$$\emptyset'_{k,n} = U_{k,n} \times S'_{k,n} \times V_{k,n}^{T} \quad (4.15)$$

(8) To build the watermarked luminance, inverse MR-SVD is applied to the modified approximation component $\emptyset'$.

(9) Reconstruction of the watermarked video frame is done using the watermarked luminance part and the original frame chrominance parts Cb and Cr, by conversion from the YCbCr into the RGB color space.

4.4.4. Watermark extraction process

The watermark extraction process is shown in fig. 4.2 (b) and is described as follows:

(1) The watermarked fast motion frames are extracted from the watermarked color video.

(2) Every watermarked fast motion frame is converted from the RGB to the YCbCr.

(3) Every watermarked luminance frame is transformed into 1-Level MR-SVD decomposition to get the watermarked approximation $\emptyset'$.

(4) The watermarked approximation component is decomposed into blocks of size $[u \times u]$.

(5) The SVD is applied to each block of the approximation component.

$$\emptyset'_{k,n} = U_{k,n} \times S'_{k,n} \times V_{k,n}^{T} \tag{4.16}$$

(6) The embedded watermark is extracted by the following rule:

$$\widehat{W}_P = \begin{cases} 1, & S'(1,1) - Q\left\lfloor \frac{S'(1,1)}{Q} \right\rfloor \geq \frac{Q}{2} \\ 0, & S'(1,1) - Q\left\lfloor \frac{S'(1,1)}{Q} \right\rfloor < \frac{Q}{2} \end{cases} \tag{4.17}$$

where $\lfloor \bullet \rfloor$ is the floor function.

(7) Decryption with the same chaotic sequence is performed to get the hidden binary watermark $\widehat{W}$.

(8) Since a video clip contains several fast motion frames, in which the same watermark is embedded, we calculate the final recovered watermark by averaging the watermarks extracted from these different frames.

4.5. Results and Discussion

The proposed algorithm is implemented using MATLAB. We used three CIF (352×288) standard sequences shown in fig. 4.4 (Foreman: 300 frames, Akiyo: 300 frames and Football: 260 frames). The test videos are in RGB uncompressed avi format, with a frame rate of 30 fps [68]. The watermark is a binary image. Its resolution is 9×11.

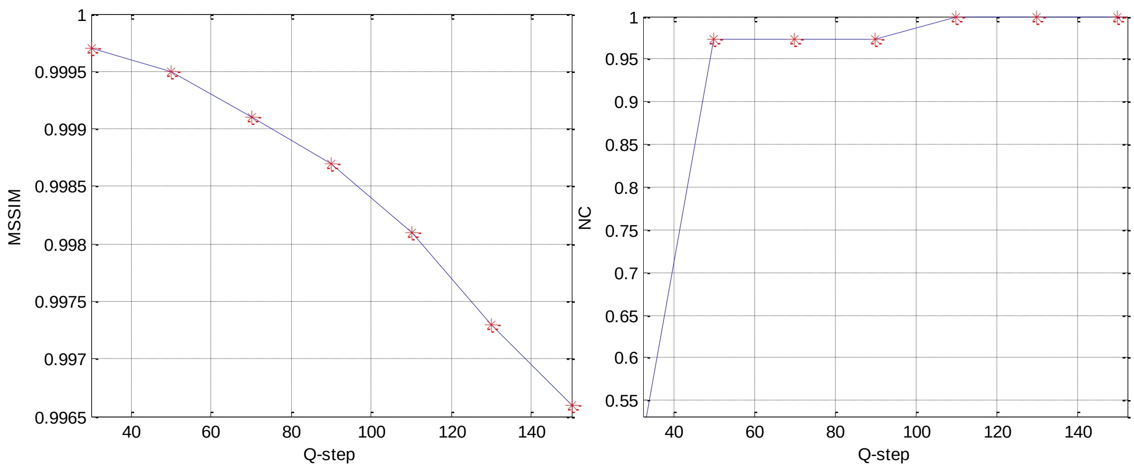
4.5.1. Parameter setting

The block size in the proposed algorithm can be set to any desired value. However, a small block size leads to block effects problem, whereas a large block size reduces the total number of the watermark bits. We carried out various experiments and found that the block size of 16x16 allows the embedding of an acceptable number of watermark bits without causing noticeable block effect.

As shown in fig 4.5, small values of the quantization step yield good transparency at the expense of poor robustness and vice versa. From this figure, it can be observed that the value of the quantization step Q=110 gives a good compromise between robustness and imperceptibility.

(a) (b) (c)

Figure 4.4 Video sequences used for testing. (a) Foreman, (b) Akiyo, (c) Football



(a) (b)

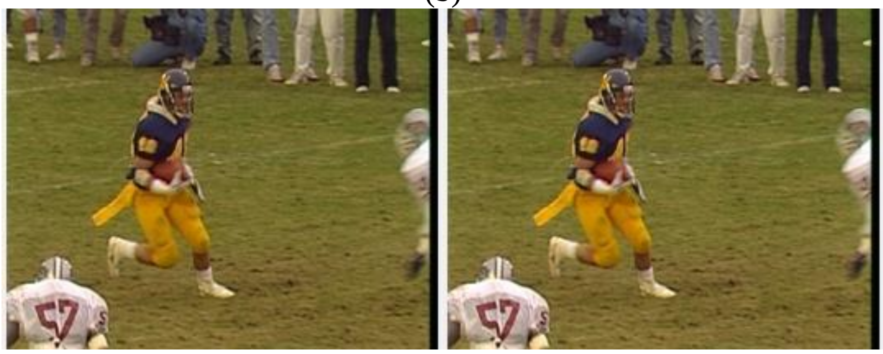
Figure 4.5 (a) MSSIM of watermarked video for different quantization step Q, (b) NC of extracted watermark with JPEG compression attack for different quantization step.

4.5.2. Imperceptibility tests

As shown in Table 4.1, the average PSNR values of the watermarked videos are higher than 40 dB and the corresponding MSSIM values are very close to one. This indicates the invisibility of the watermark which means that the watermarked videos appear visually identical to the original ones as shown in fig 4.6.

(a)

(b)



(c)

Figure 4.6 Original and watermarked frames: (a) frame 228 of Foreman (b) frame 16 of Akiyo (c) frame 128 of Football.

Table 4. 1 The PSNR and MSSIM of the watermarked videos

Video	Foreman	Akiyo	Football
Average PSNR	40.20	40.45	40.26
Average MSSIM	0.9981	0.9988	0.9977

4.5.3. Robustness tests

4.5.3.1. Image processing attacks

- Adding a noise: Three kinds of noises are added to the watermarked video: Gaussian noise, Salt & pepper noise and speckle noise with density of 1%. It can be seen from table 4.2 that the watermark is always detectable with NC and BER values respectively close to 1 and 0, especially, for Foreman and Football video sequences.
- Filtering: 3x3 median filter and 5x5 Gaussian filter are applied separately to the watermarked videos and we can see from table 4.2 that the proposed method is robust against median and Gaussian filtering.
- JPEG compression: The watermarked video is compressed with different quality factors ranging from 10 to 100. Fig. 4.7 shows the results for the JPEG compression. For instance, if the watermarked video is compressed with a quality factor of 40, the obtained NC is greater than 95% and the BER value is lower than 0.2%. This confirms the robustness of the proposed scheme to the JPEG compression attack.

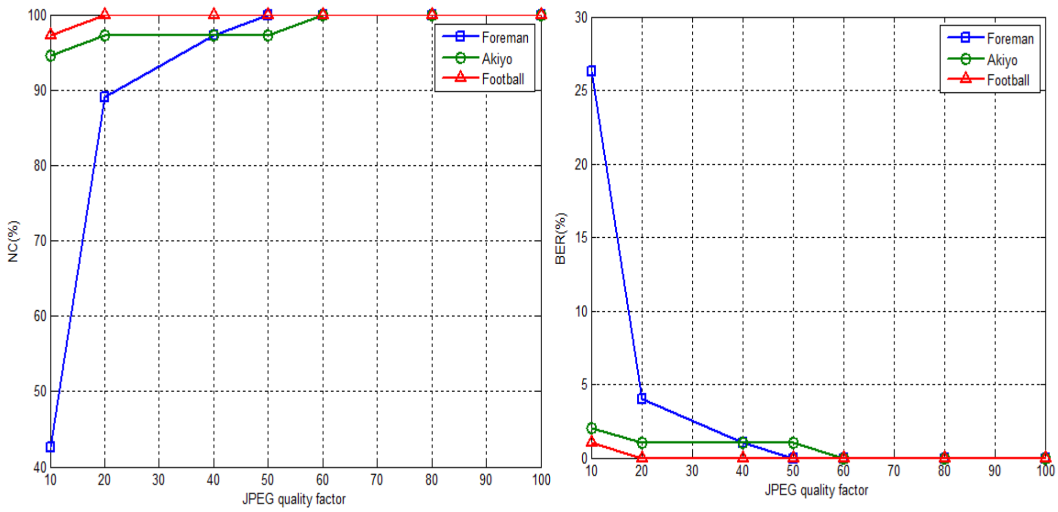
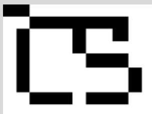
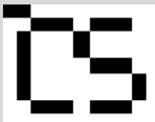
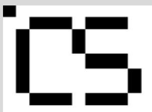


Figure 4.7 NC and BER values of the extracted watermarks under the JPEG compression attack.

Table 4. 2 Extracted watermark (with NC and BER) under image processing attack.

Attacks	Foreman	Akiyo	Football
No attack	 NC= 1 BER=0	 NC= 1 BER=0	 NC= 1 BER=0
Gaussian noise	 NC=0.9008 BER=0.0404	 NC=0.6674 BER=0.1414	 NC= 1 BER=0
Salt& Pepper noise	 NC= 1 BER=0	 NC= 0.9238 BER= .0303	 NC= 1 BER=0
Speckle noise	 NC=0.9479 BER=0.0202	 NC=0.7870 BER=0.0909	 NC= 1 BER=0
Median filter	 NC=0.9732 BER=0.101	 NC= 1 BER=0	 NC= 1 BER=0
Gaussian low pass filter	 NC= 1 BER=0	 NC= 1 BER=0	 NC= 1 BER=0

4.5.3.2. Frame synchronization attacks

- Frame dropping: Table 4.3 shows the average NC and BER values given at different frame dropping rates. Our scheme achieves strong robustness against frame dropping even for the case of high rates (i.e. 80%).
- Frame averaging: The watermarked video is averaged for various averaging rates and then we tried to extract the watermark. Table 4.4 shows that the watermark can be recovered at frame averaging rates up to 50%.
- Frame swapping: The results presented in Table 4.5 prove the robustness of our watermarking scheme against frame swapping because when all watermarked frames are swapped the NC is 1 and the BER is 0.

4.5.3.3. Video compression attacks

We compress the videos sequences with two different lossy compressions [78]: H264 coding with a bit rate of 512 kbps and MPEG4 coding with bit rates of 1500 kbps and 1000 kbps. The NC and BER are depicted in table 4.6 and we see that the algorithm can resist to video compression attacks.

Table 4. 3 Extracted watermark (with NC and BER) under frame dropping attack.

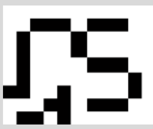
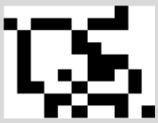
Frame dropping rate	Foreman	Akiyo	Football
50%	 NC= 1 BER=0	 NC= 1 BER=0	 NC= 1 BER=0
80%	 NC=0.8485 BER=0.0606	 NC=0.7822 BER=0.101	 NC=0.7156 BER= 0.1313

Table 4. 4 Extracted watermark (with NC and BER) under frame averaging attack.

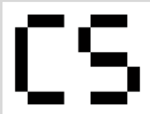
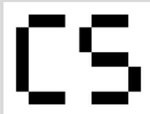
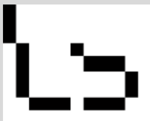
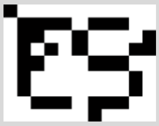
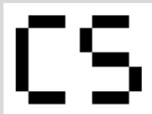
Frame averaging rate	Foreman	Akiyo	Football
25%	 NC= 1 BER=0	 NC= 1 BER=0	 NC= 1 BER=0
50%	 NC=0.6819 BER=0.1111	 NC=0.800 BER=0.0909	 NC= 1 BER=0

Table 4. 5 Extracted watermark (with NC and BER) under frame swapping attack

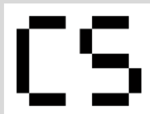
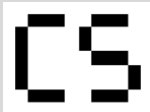
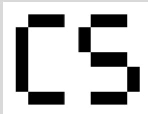
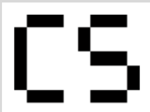
Frame Swapping rate	Foreman	Akiyo	Football
50%	 NC= 1 BER=0	 NC= 1 BER=0	 NC= 1 BER=0
100%	 NC= 1 BER=0	 NC= 1 BER=0	 NC= 1 BER=0

Table 4. 6 Extracted watermark with (with NC and BER) under different video compression attacks.

Type of compression	Foreman	Akiyo	Football
H264 compression (512kbps)	 NC= 1 BER=0	 NC= 1 BER=0	 NC= 1 BER=0
MPEG4 compression (1500kbps)	 NC= 1 BER=0	 NC=0.8708 BER=0.0505	 NC= 1 BER=0
MPEG4 compression (1000kbps)	 NC= 1 BER=0	 NC=0.848 BER=0.0606	 NC=0.9238 BER=0.0303

4.5.3.4. Results with the Stirmark benchmark

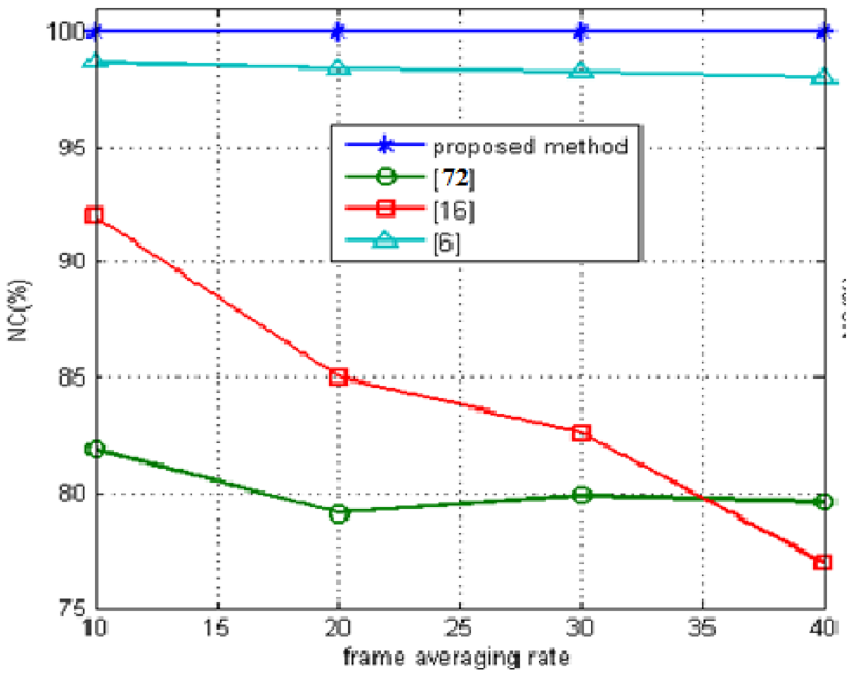
We also evaluated the proposed method with Stirmark 3.1, which is a well-known evaluation tool for watermarking robustness of watermarked video frames under image processing attacks [79, 80]. Table 4.7 shows the evaluation of the proposed method with the Stirmark benchmark under JPEG compression, median filter with JPEG compression, Gaussian filter with JPEG compression, sharpening with JPEG compression and removing of rows and columns with JPEG compression. As can be observed, the proposed watermarking technique can resist to all these named attacks.

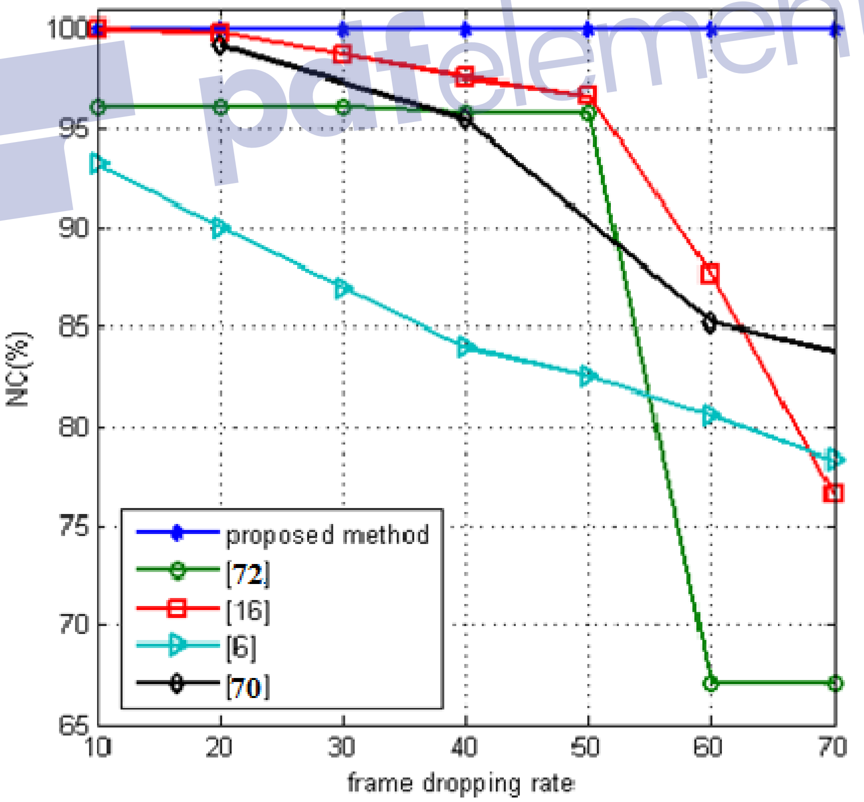
Table 4. 7 Extracted watermark (with NC and BER) under image processing attacks with Stirmark benchmark.

Attacks	NC	BER
JPEG 10	0.5677	0.1515
JPEG 30	0.9479	0.0202
JPEG 60	0.8485	0.0606
JPEG 90	1.0000	0.0000
3x3 median filter+ JPEG	0.9732	0.0101
Gaussian filter +JPEG	0.7681	0.1010
Sharpening +JPEG	0.5127	0.2626
removing one row and one column +JPEG	0.5523	0.2020

4.5.4. Comparison with some previously reported algorithms

In order to evaluate the performance of our algorithm, we compared the results of the proposed video watermarking scheme to the results of related video watermarking schemes given in [6], [16], [70] and [72], that we introduced and discussed in the introduction. Figure 4.8 displays these results of comparison under frame dropping, frame averaging and JPEG compression attacks. It is clear that the proposed scheme achieves a better robustness.


100
95
90
85
80
75
NC(%)
10
15
20
25
30
35
40
frame averaging rate
proposed method
[72]
[16]
[6]


100
95
90
85
80
75
70
65
NC(%)
10
20
30
40
50
60
70
frame dropping rate
proposed method
[72]
[16]
[6]
[70]

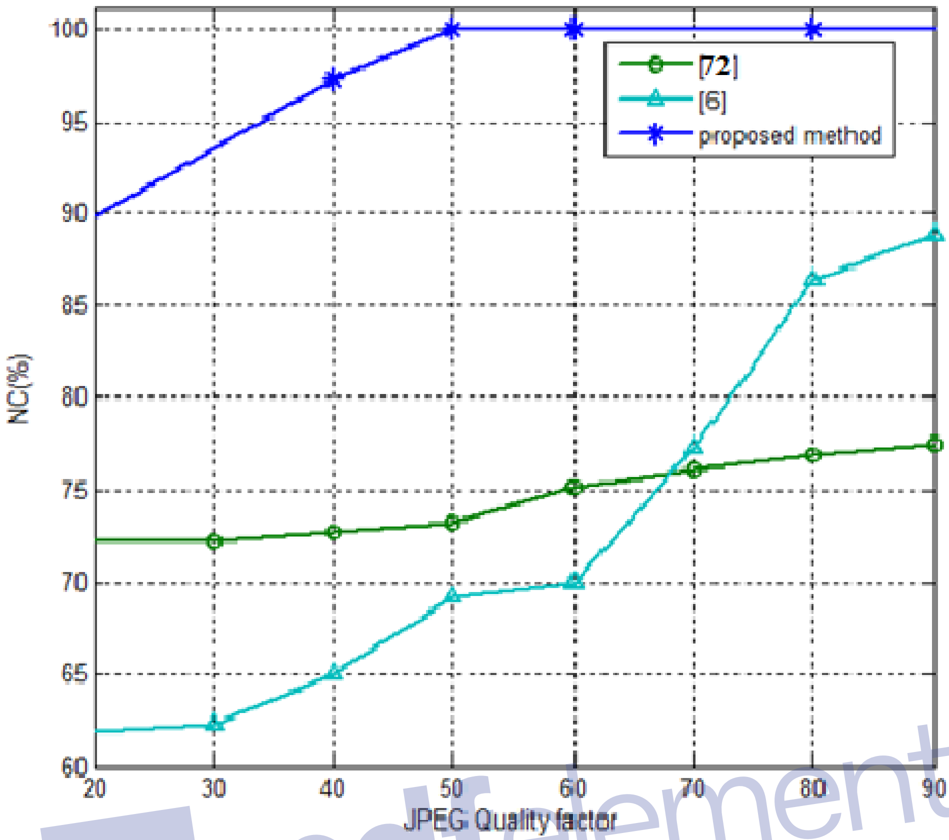


Figure 4.8 Comparison of experimental results between the proposed technique and other schemes under frame averaging attack, frame dropping attack and JPEG compression attack.

4.5.5. Complexity of the proposed algorithm

The proposed watermarking schema has a low complexity because:

- The algorithm is done in the MR-SVD domain which has an overall complexity of $O(n)$, where n indicates the signal length [74].
- To extract fast motion frames we use a simple threshold decision method based on motion energy.
- The embedding and extraction are not applied to all the frames of the video but just to fast motion frames.

In fact, the execution times of the proposed scheme required in watermark embedding and extraction processes are in the following results: watermark embedding: 0.36 s per frame; watermark extraction: 0.19 s per frame. The simulation was conducted on Intel(R) Core™ i5, 1.80 GHz processor, with 4 GB RAM using MATLAB ver.R2013b.

4.6. Conclusion:

In this chapter, we presented a novel blind video watermarking scheme in fast motion frames using the SVD, the MR-SVD, the QIM and the Logistic Map encryption. In the proposed technique, we solved the problem of embedding the watermark in all frames by choosing only the frames that have big motion energy to be the host frames, suitable to the HVS. Using this approach, we have limited the number of frames to be processed and also assured a higher imperceptibility of the watermark. In addition, the combination of the characteristics of the SVD, MR-SVD, QIM and Logistic Map Encryption make our scheme secure and robust to a variety of attacks. The experimental results confirmed that the proposed watermarking scheme has good imperceptibility with PSNR greater than 40 dB. The proposed scheme is not only robust to image processing attacks, but also to frame synchronization and video compression attacks. The comparison results with other algorithms related to video watermarking indicated the superiority of our scheme.

General conclusion and future work

The subject of our thesis was to develop and implement digital watermarking algorithms for multimedia data and precisely for video sequences. After doing an overview of video watermarking, we have found that frame-by-frame embedding watermarking techniques were the most popular video watermarking developed because they remained relatively simple and any image watermarking method that have been presented in the literature can be applied directly to video. However, the major drawback of these techniques was that they did not consider the temporal dimension of a video. To overcome this problem of frame-by-frame embedding techniques, we have proposed two contributions:

In the first contribution, we have proposed a semi blind video watermarking scheme using spatiotemporal just noticeable distortion (ST-JND) in DCT domain. ST-JND is a way to model the spatial and temporal properties of the Human Visual System and use it in order to improve robustness in the detection, and at the same time obtain a good quality of the watermarked video. The watermark sequence was embedded and detected using a quantization index modulation (QIM) algorithm with adaptive step size, which is calculated using spatiotemporal just noticeable distortion (ST-JND). The experimental results showed that the proposed scheme not only improved the perceptual quality but also gained robustness against various attacks. Compared to a non-HVS approach and S-JND approach, the ST-JND based video watermarking algorithm achieved better performance.

In the second contribution, we have proposed a novel blind video watermarking scheme in fast motion frames using the SVD, the MR-SVD, the QIM and the Logistic Map encryption. In the proposed technique, we solved the problem of embedding the watermark in all frames by choosing only the frames that have big motion energy to be the host frames, suitable to the HSV. By using this approach, we have limited the number of frames to be processed and assured a higher imperceptibility of the watermark. In addition, the combination of the characteristics of the SVD, MR-SVD, QIM and Logistic Map Encryption made our scheme secure and robust to a variety of attacks. The experimental results confirmed that the proposed watermarking scheme had good imperceptibility with PSNR greater than 40 db. Furthermore, the proposed scheme was not only robust to image processing attacks, but also to frame synchronization and video compression attacks. The comparison of our results with the ones of other algorithms related to video watermarking indicated the superiority of our scheme.

As a future work, we will improve the ST-JND model to be compatible with the quantization index modulation watermarking technique, because the effect of video motion is not clear as in additive embedding technique.

References

[1] Cox I. Cox, M. Miller, J. Bloom, J. Fridrich, T. Kalker, "Digital Watermarking and Steganography," Morgan Kaufmann Publishers Inc., 2008.

[2] S. Katzenbeisser, F. Petitcolas, "Information Hiding Techniques for Steganography and Digital Watermarking," Artech House, 2000

[3] H. Tao, L. Chongmin, J. Mohamad Zain, A. N. Abdalla, "Robust image watermarking theories and techniques: A review," Journal of Applied Research and Technology, vol. 12, no. 1, pp. 122-138, 2014.

[4] M. Kutter, F. Bossen, "Digital watermarking of color images using amplitude modulation", Journal of Electronic imaging, vol.7, no. 2, pp. 326-332, 1998.

[5] A. Karmakar, A. Phadikar, B. S. Phadikar, G. K. Maity,, "A blind video watermarking scheme resistant to rotation and collusion attacks," Journal of King Saud University - Computer and Information Sciences, vol. 28, no. 2, pp. 199-210, 2016.

[6] T. R. Singh, K. M. Singh, S. Roy, "Video watermarking scheme based on visual cryptography and scene change detection, " AEU-International Journal of Electronics and Communications, vol. 67, no. 8, pp. 645-651,2013.

[7] W. Kong, B. Yang, D. Wu, X. Wu D., Niu X., "Svd based blind video watermarking algorithm", in proceedings of International Conference on Innovative Computing, Information and Control (ICICIC'06) , Vol. 1, pp. 265-268, Beijing, China, 30 Augst-1 September 2006.

[8] R. Liu, T. Tan , "An svd-based watermarking scheme for protecting rightful ownership," IEEE transactions on multimedia, vol. 4, no. 1,pp. 121-128, 2002.

[9] O. S. Faragallah, "Efficient video watermarking based on singular value decomposition in the discrete wavelet transform domain," AEU-International Journal of Electronics and Communications, vol. 67, no. 3, pp.189-196, 2013.

[10] K. Niu, X. Yang, L. Xiang., "Hybrid quasi-3d dwt/dct and svd video watermarking," in proceedings of International Conference on Software Engineering and Service Sciences (ICSESS), pp. 588-591, Beijing, China, 16-18 July 2010.

[11] N. I. Yassin, N. M. Salem, M. I. El Adawy, "Qim blind video watermarking scheme based on wavelet transform and principal component analysis," Alexandria Engineering Journal, vol. 53, no. 4 , pp. 833-842,2014.

[12] D. K. Thind, S. Jindal, "A semi blind dwt-svd video watermarking," Procedia Computer Science, vol. 46, no. Supplement C, pp. 1661- 1667, 2015.

[13] N. Leelavathy, E. Prasad, S. S. Kumar, "Video watermarking techniques: A review," International Journal of Computer Applications, vol. 104, no.7.

[14] F. Hartung, M. Kutter, "Multimedia watermarking techniques", in proceedings of the IEEE, vol. 87, pp.1079-1107, july 1999.

[15] P. Rasti, S. Samiei, M. Agoyi, S. Escalera, G. Anbarjafari, "Robust non blind color video watermarking using QR decomposition and entropy analysis, " Journal of Visual Communication and Image Representation, vol. 38, pp. 838-847, 2016.

[16] S. M. Youssef, A. A. El Farag, N. M. Ghatwary, "Adaptive video watermarking integrating a fuzzy wavelet-based human visual system perceptual model," Multimedia tools and applications, vol. 73, no. 3, pp. 1545-1573, 2014.

[17] S.A.K.Mostafa, A. S. Tolba , F. M. Abdelkader, and H. M. Elhin, "Video Watermarking Scheme Based on Principal Component Analysis and Wavelet Transform," IJCSNS International Journal of Computer Science and Network Security, vol.9, no.8, August 2009.

[18] C. Rey, and J. Dugelay, "Un panorama des méthodes de tatouage permettant d'assurer un service d'intégrité pour les images," Traitement du Signal, pp. 283–295, 2001.

[19] Cox I. Cox, M. Miller, J. Bloom, J. Fridrich, T., "The First 50 Years of Electronic Watermarking," EURASIP Journal on Applied Signal Processing, Vol. 2002, No. 2, pp.126–132, 2002.

[20] C.M. Barni and F. Bartolini. , "Watermarking Systems Engineering : Enabling Digital Assets Security and Other Applications". CRC, 2004.

[21] B.Sevinc , H.T. Sencar, N. D. Memon, "An efficient and robust method for detecting copy-move forgery," in proceedings of International of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2009), Taipei, Taiwan, 9-24 April 2009.

[22] M. K. Arnold, M. Schmucker, S. D. Wolthusen, "Techniques and Applications of Digital Watermarking and Content Protection, Artech House," ISBN: 1-58053-111-3,2003.

[23] H. Wei-Fan and L. Pei-Yu, "Imperceptible Visible Watermarking Scheme Using Color Distribution Modulation," in proceedings of. The ninth International Conference on Ubiquitous Intelligence and Computing and 9th International Conference on Autonomic and Trusted Computing, pp.1002-1005, Fukuoka, Japan, 4-7 Sept. 2012.

[24] E.T.Lin, C. I. Podilchuk, and E.J. Delp., "Detection of Image Alterations Using Semi-Fragile Watermarks,"in proceedings of Security and Watermarking of Multimedia Contents II (SPIE), 9 May 2000.

[25] A. Hanjalic, G.C. Langelaar , P.M.B. van Roosmalen , J. Biemond, R.L. Lagendijk, "Image and Video Databases: Restoration, Watermarking and Retrieval," Volume 8 ; 1st Edition chapter 7, Elsevier Science BV,2000.

[26] A.N. Al-Gindy ; A. Tawfik , H. Al Ahmad , R. A. Qahwaji, "A New Blind Image Watermarking Technique for Dual Watermarks Using Low-Frequency Band DCT Coefficients," in processing of 14th IEEE International Conference on Electronics, Circuits and Systems, Marrakech, Morocco, 11-14 Dec. 2007.

[27] A. Singhal and, K. U. Singh, "Digital Watermarking Algorithms: A Review," International Journal of Engineering and Computer Science, Vol.4, Issue 9, pp. 14409-14416, Sep. 2015.

[28] W. Bender, D. Gruhl, N. Morimoto, A. Lu, "Techniques for data hiding," IBM Syst. J., vol. 35,pp.313–336, 1996.

[29] G.J.Durieu, "Survey of watermarking techniques," Theses and Dissertations. 708,2001.

[30] I.J. Cox, J. Kilian, F. Thomson Leighton, T. Shamoon, "Secure spread spectrum for multimedia", IEEE Trans. Image Process., vol. 6, no. 12, pp. 1673-1687, 1997.

[31] M. George, "Spread spectrum watermarking for images and video," master Thesis; 1999.

[32] Y.Xiang, I. Natgunanathan, Y.Rong, and S. Guo, "Spread Spectrum-Based High Embedding Capacity Watermarking Method for Audio Signals," IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, vol. 23, no. 12, December 2015.

[33] B. Chen, G. W. Wornell, "Quantization index modulation methods for digital watermarking and information embedding of multimedia," J. VLSI Signal Process. Syst., vol. 27, no. 1-2, pp. 7-33, 2001.

[34] J. Lim. "Is BER the bit error ratio or the bit error rate?" (Accessed on 16 February2015.)

[35] G. J. Durieu, "Survey of watermarking techniques," Theses and Dissertations. 708,2001.

[36] F. Davoine and S. Pateux, "Tatouage de Documents Audiovisuels Numériques, " Traite IC2: Traitement du Signal ET de l'Image, Hermès Science Publications, Lavoisier, France, 2004.

[37] B. G. Haskell, A. N. Netravali, “Digital Pictures: Representation, Compression, and Standards,” Perseus Publishing, 1997.

[38] M. Chaumont “ tatouage de documents numériques, "cours, 12 novembre 2008

[39] https://www.youtube.com/yt/about/press/ access 17/02/2019

[40] G. Doërr, “Security issue and collusion attacks in video watermarking,” doctoral thesis, 2005.

[41] Glossary of video terms and acronyms : www.tektronix.com

[42] S. Deb, “ Video Data Management and Information Retrieval”, IRM Press pp. 321-346, IRM Press, 2005

[43] D.A. Cook, “A history of narrative film,” second Edition, p2 W.W. Norton & Compnay Inc. New York, 1990.

[44] T. Wiegand , “Digital Image Communication Digital,”Course at Technical University of Berlin,2003.

[45] G.J. Sullivan, J.R. Ohm; W.J. Han; T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," IEEE Transactions on Circuits and Systems for Video Technology, 2013.

[46] H. Koumaras, A. Kourtis, and D. Martakos , “Principles of Digital Video Coding,” , Encyclopedia of Information Technology, 2nd Edition, Idea Group Pub., pp. 3119-3124,October 2008.

[47] M.V. LATHA, A. FATHIMA “A REVIEW ON IMAGE AND VIDEO COMPRESSION STANDARDS,” Asian Journal of Pharmaceutical and Clinical Research, Special Issue (April)
2017.

[48] G. Doërr and J.L. Dugelay, "A Guide Tour of Video Watermarking", Signal Processing: Image Communication, vol.18, no.4,pp. 263-282, 2003.

[49] E.T.Lin, and E.J. Delp,“Temporal synchronization in video watermarking,” IEEE Transactions on Signal Processing ,vol.52 ,pp. 3007-3022,2002.

[50] A. BOUDERBALA, “Implémentation d’un algorithme de tatouage Vidéo robuste dans Le domaine compressé, ” magister thesis.

[51] L. E. A. SAUERBRONN, M. A. DREUX, “Transparent Digital Watermark, ” In Monografias do Departamento de Informática da PUC-Rio, 2000, Rio de Janeiro, 2000.

[52] A. Boho, G. V. Wallendael, A. Dooms, J. D. Cock, G. Braeckman, P. Schelkens, B. Preneel, and R. Van de Walle, "End-to-end security for video distribution," IEEE Signal Process. Mag., vol. 30, no. 2, pp. 97–107, 2013.

[53] S.Lian, Z. Liu, Z. Ren, and H. Wang, "Commutative Encryption and Watermarking in Video Compression," IEEE Transactions on Circuits and Systems for Video Technology, vol.17, no.6, pp. 774–778, 2007.

[54] H. X. Wang, Y. N. Li, Z. M. Lu, and S. H. Sun, "Compressed Domain Video Watermarking in Motion Vector," International Workshop on Intelligent Information Hiding and Multimedia Signal Processing, vol. 3682, pp. 580-586, Melbourne, Australia, 14-16 September, 2005.

[55] C.S. Lu, J.R. Chen, H.Y. Liao, and K.C. Fan, "Real-time MPEG2 video watermarking in the VLC domain," In Proceedings of the IEEE International Conference on Pattern Recognition, vol. II, pp. 552–555, August 2002.

[56] F. Hartung, B. Girod, "Watermarking of uncompressed and compressed video," Signal Processing, vol. 66 no.3,pp. 283–301, (1998).

[57] Y.Y. Lee, H.S. Jung, S.U. Lee, "3D DFT-based video watermarking using perceptual models," IEEE 46th Midwest Symposium on Circuits and Systems, vol. 3, pp. 1579–1582, 2003.

[58] P. Campisi, A. Neri, "Perceptual video watermarking in the 3D-DWT domain using a multiplicative approach, in: Digital Watermarking, "Springer, pp. 432–443, 2005.

[59] H.Park, S.H. Lee, Y.S. Moon, "Adaptive video watermarking utilizing videocharacteristics in 3D-DCT domain", in: Digital Watermarking," Springer, pp. 397–406, 2006.

[60] S. M. Youssef, A. A. El Farag, N. M. Ghatwary, "Adaptive video watermarking integrating a fuzzy wavelet-based human visual system perceptual model," Multimedia tools and applications, vol. 73, no. 3, pp. 1545-1573, 2014.

[61] W.H Chen, C.H. Smith, and S.C. Fralick, "A fast computational algorithm for the discrete cosine transform,"IEEE Trans. Commun., vol. COM-25, pp. 1004-1009. 1977.

[62] A.B. Watson, "Image Compression Using the Discrete Cosine Transform", Mathematica Journal, pp. 81-88, 1994.

[63] Y. Jia, W. Lin, A. A. Kassim, "Estimating just-noticeable distortion for video," IEEE Transactions on Circuits and Systems for Video Technology, vol. 16, no. 7, pp. 820-829, 2006.

[64] X. Zhang, W. S. Lin, and P. Xue, “Improved estimation for just-noticeablevisual distortion,” Signal Process., vol. 85, no. 4, pp. 795–808,Apr. 2005.

[65] S. Daly, “Engineering observations from spatiovelocity and spatiotemporal visual models,” in Vision Models and Applications to Image and Video Processing. Norwell, MA: Kluwer, 2001.

[66] S. J. P. Westen, R. L. Lagendijk, and J. Biemond, “A quality measure for compressed image sequences based on an eye movement compensated spatio-temporal model,” in Proc. Int. Conf. Image Process., vol. 1, pp. 279–282, Oct. 1997.

[67] C.H. Chou and Y.C. Li, “A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile,” IEEE Trans. Circuits Syst. Video Technol., vol. 5, no. 6, pp. 467–476, Dec. 1995.

[68] . http://media.xiph.org/video/derf/. (Accessed on 06 July 2018).

[69] T. Tabassum, S. M. Islam, “A digital video watermarking technique based on identical frame extraction in 3-level dwt,” in proceedings of 15th International Conference on Computer and Information Technology (ICCIT), pp.101-106, Chittagong, Bangladesh, 22-24 December 2012.

[70] L. Agilandeeswari, K. Ganesan, “A robust color video watermarking scheme based on hybrid embedding techniques,” Multimedia Tools and Applications, vol. 75, no. 14, pp. 8745-8780,2016.

[71] J. Xuemei, L. Quan, W. Qiaoyan, “Qiaoyan, A new video watermarking algorithm based on shot segmentation and block classification,” Multimedia tools and applications, vol. 62, no. 3, pp. 545-560, 2013.

[72] K. Chetan, K. Raghavendra, “Dwt based blind digital video watermarking scheme for video authentication,” International journal of computer applications, vol. 4, no. 10, pp. 19-26, 2010.

[73] H. Andrews, C. Patterson, “Singular value decompositions and digital image processing,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 24, no. 1, pp. 26-53, 1976.

[74] R. Kakarala, P. O. Ogunbona, “Signal analysis using a multiresolution form of the singular value decomposition,” IEEE Transactions on Image processing, vol. 10, no. 5, pp. 724-735, 2001

[75] M K. M. S, M. Yaghoobi, "A new approach for image encryption using chaotic logistic map," in proceedings of International Conference on Advanced Computer Theory and Engineering, pp. 585-590 Phuket, Thailand, 20-22 December 2008.

[76] R. Dugad, K. Ratakonda, N. Ahuja, "Robust video shot change detection," ," in proceedings of IEEE Second Workshop on Multimedia Signal Processing, pp.376-381, 1998.

[77] A. Anees, I. Hussain, A. Algarni, M. Aslam, "A Robust Watermarking Scheme for Online Multimedia Copyright Protection Using New Chaotic Map," Security and Communication Networks, 2018.

[78] http://www.virtualdub.org (Accessed on 06 July 2018).

[79] F. Petitcolas, J.A. Ross , G. K. Markus, "Attacks on copyright marking systems," in David Aucsmith (Ed), Information Hiding, Second International Workshop" , ," in proceedings of LNCS 1525 , pp. 219-239, Portland, Oregon, U.S.A., 15-17April 1998.

[80] http://www.petitcolas.net/watermarking/stirmark (Accessed on 19 September 2018).

[81] I. Nouioua, and N. Amardjia, "watermarking for video using spatiotemporal Just Noticeable Distortion," in International Conference on EMBEDDED SYSTEMS in TELECOMMUNICATIONS and INSTRUMENTATION (ICESTI'14) ,annaba, Algeria 27 – 29 OCTOBER 2014.

[82] I. Nouioua, N. Amardjia, and S. Belilita, "A Novel Blind and Robust Video Watermarking Technique in Fast Motion Frames Based on SVD and MR-SVD," Security and Communication Networks, vol. 2018, 17 pages, 2018.

Abstract

In this thesis, we give a deep overview of digital video watermarking and present two contributions in this domain. In the first contribution, we propose a robust video watermarking scheme based on spatiotemporal just noticeable distortion in the DCT domain. Spatiotemporal Just noticeable distortions (ST-JND) which refers to the maximum distortion threshold that the HVS cannot perceive is employed in order to obtain a trade-off between watermark robustness and imperceptibility. The ST-JND based video watermarking algorithm achieves better performance compared to a no-HVS approach and S-JND approach. In the second contribution, we overcome the problem of a frame-by-frame embedding technique by choosing only the frames that have big motion energy to be the host frames. By doing so, the number of frames to be processed is consequently reduced and a better quality of the watermarked video is ensured since the human visual system cannot notice the variations in fast moving regions. The experimental results show that the proposed method can achieve a very good transparency while being robust against various kinds of attacks. Compared to several methods found in the literature, the proposed method gives a better robustness.

Keywords: Video watermarking; Singular Value Decomposition; Human Visual System; Spatiotemporal Just Noticeable Distortion.

Résumé

Dans cette thèse, nous donnons un large aperçu du tatouage numérique de la vidéo et présentons deux contributions dans ce domaine. Dans la première contribution, nous proposons un schéma de tatouage vidéo robuste basé sur la distorsion spatio-temporelle juste perceptible (ST-JND) dans le domaine de la DCT. La ST-JND fait référence au seuil de distorsion maximal que le système visuel humain ne peut pas percevoir et est utilisée afin d'obtenir un compromis entre la robustesse et l’imperceptibilité du tatouage. L'algorithme de tatouage vidéo basé sur la ST-JND a permis d'atteindre de meilleures performances par rapport à une approche non-JND et une approche S-JND. Dans la deuxième contribution, nous résolvons le problème de la technique d’incorporation de la marque image par image en choisissant uniquement les images ayant une grande énergie de mouvement comme images hôtes. Ainsi, le nombre d'images à traiter est réduit et une meilleure qualité de la vidéo tatouée est assurée puisque le système visuel humain ne peut pas remarquer les variations dans les régions à déplacement rapide. Les résultats expérimentaux montrent que la méthode proposée a une très bonne transparence tout en étant robuste contre divers types d’attaques. Comparée à plusieurs méthodes trouvées dans la littérature, la méthode proposée donne une meilleure robustesse.

Mots clés : Tatouage vidéo ; Décomposition en Valeurs Singulières ; Système Visuel Humain ; Distorsion Spatiotemporelle Juste Perceptible.

ملخص

في هذه الأطروحة، نقدم نظرة عامة عميقة حول وضع العلامات المائية للفيديو الرقمي ونأتي بمساهمتين في هذا المجال. في المساهمة الأولى، نقترح خطة قوية للعلامات المائية للفيديو تعتمد على تشوه ملحوظ مكانياً وزمانيا (ST-JND) في مجال DCT. ST-JND و الذي يشير إلى الحد الأقصى للتشويه الذي لا يمكن لنظام البشري البصري إدراكه, استعمل من أجل الحصول على المفاضلة بين متانة العلامة المائية وعدم إدراكها. حققت خوارزمية علامة مائية الفيديو المستندة إلى ST-JND أداء أفضل بالمقارنة مع نهج non-JND ونهج S-JND. في المساهمة الثانية، تغلبنا على مشكلات في تقنية تضمين الإطار تلو الأطر باختيار الإطارات التي لديها طاقة حركة كبيرة لتكون الإطارات المضيفة فقط. وبذلك، يتم تقليل عدد الإطارات التي سيتم معالجتها ويتم ضمان جودة أفضل للفيديو ذي العلامة المائية نظرًا لأن النظام البصري البشري لا يمكنه ملاحظة الاختلافات في المناطق سريعة الحركة. توضح النتائج التجريبية أن الطريقة المقترحة يمكن أن تحقق شفافية جيدة للغاية، بينما تكون قوية ضد أنواع مختلفة من الهجمات. بالمقارنة مع ديد من الطرق الموجودة في ىرخا ه، فإن الطريقة المقترحة تعطي متانة أفضل.

الع

كلمات البحث: وشم الفيديو، تحلل القيمة المفرد، النظام البصري البشري، فقط التشويه الزماني المكاني الملحوظ