Digital Video Watermarking Techniques for Secure Multimedia Creation and Delivery

CHAN Pik-Wah

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Master of Philosophy
in
Computer Science and Engineering

Supervised by

Prof. Michael R. Lyu

July 2004

Abstract of thesis entitled:

Digital Video Watermarking Techniques for Secure Multimedia Creation and Delivery

Submitted by CHAN Pik-Wah

for the degree of Master of Philosophy

at The Chinese University of Hong Kong in July 2004

There is an explosion of data exchange on the Internet and the extensive use of digital media. Consequently, digital data owners can quickly and massively transfer multimedia documents through the Internet. It has aroused intense interest in multimedia security and multimedia copyright protection. In this paper, a comprehensive approach for protecting and managing video copyrights with watermarking techniques is introduced. We propose a novel hybrid digital video watermarking scheme based on the scene change analysis, error correction code and genetic algorithm. Our video watermarking algorithm is robust against the attacks of frame dropping, averaging and statistical analysis, which were not solved effectively in the past. It started with a complete survey about current watermarking technologies. We have discovered that none of the existing schemes was capable of resisting all attacks. Accordingly, we came up with the idea of embedding a single watermark of different parts into different scenes of a video. Then we analyze the strength of dif-

ferent watermarking schemes. A hybrid approach is applied to form a super watermarking scheme which can resist most of the attacks. In order to reinforce the robustness of the scheme, the watermark is refined by an error correcting code, while the correcting code is embedded as a watermark in the audio channel. Furthermore, the fidelity of the scheme is enhanced by applying genetic algorithm. It optimizes the quality of the watermarked video. Also, our scheme allows blind retrieval of the embedded watermark, which does not need the original video and the watermark is perceptually invisible. The effectiveness of the scheme is verified through a series of experiments, in which a number of standard image processing attacks are conducted, and the robustness of our approach is demonstrated by using the criteria of the latest StirMark test.

摘要

論文題目：多媒體製造及安全傳遞的數碼影像水印技術

作者：陳碧華
修讀學位：　哲學碩士
香港中文大學計算機科學及工程學部
日期：　二零零四年六月

現今互聯網上的數據互換急速上升及數碼化媒體的應用不斷增加，使數碼數據的用家可透過互聯網高速及大量地傳送多媒體檔案。因此，多媒體的安全性及版權保護引起廣大的關注及興趣。在這篇論文中，我們發表一個完整策略，利用水印技術去保護及管理影像版權。

我們提出一個利用場境改變分析及改正規則的創新混合式的數碼影像水印方案。此方案可抵抗多類不同的攻擊，例如部份畫面被刪減，透過數據分析而移取水印等等。過去，這些問數都缺乏有效的解決方法。

首先我們就現今的水印技術作詳細調查，我們發現沒有一個現存的水印方案可抵抗所有攻擊。因此，我們提出一個新的方案將一個水印分成數份後，把他們分別埋入不同的畫面之中。其後透過分析不同水印方案的強項，運用混合策略去構成一個可抵抗大部分攻擊的超級方案。爲提升方案的防禦度，我們將水印的改正規則埋入聲音頻道裏。在提取水印過程中，這些改正規則可改良提取的水印。此外，我們亦用遺傳學方法加強方案的逼真度。這方法可提升影像被印後的質素。與此同時，我們的方案容許“瞎提取水印”方式，即在提取水印時無須原本的影像及水印。我們的水印在埋入後是不會被察覺的。

透過一系列的實驗，本方案的有效性已被驗證。當中有標準的圖像加工及一些針對影像特質的攻擊。本方案的防禦性亦已被最新的“StriMark”標準測試認證。

Acknowledgement

I would like to take this opportunity to express my gratitude to my supervisor Prof. Michael R. Lyu, for his generous guidance and patience given to me in the past two years. His numerous support and encouragement, as well as his inspiring advice are extremely essential and valuable in my research papers (conference papers published in ICICS'2003 and WWW'2004 and journal paper submitted to IEEE transaction TCSVT) and my thesis.

I am also grateful for the time and valuable suggestion that Prof. Irwin King, Prof. Tien-Tsin Wong and Prof. Roland Chin have given in marking my term paper. Without their effort, I will not be able to strengthen and improve my research project and papers.

I would also like to show my gratitude to the Department of Computer Science and Engineering, CUHK, for the provision of the best equipment and pleasant office environment required for high quality research.

Special thanks should be given to Mr. Edward Yau and Mr. Sam Sze who have given me valuable suggestions, encouragement and supports. And I would like to give my thanks to my fellow colleagues, H. Y. Chan, C. H. Hoi, K. Z. Huang, K. Y.

Lee, C. H. Chan, C. Y. IP, Y. Lam, C. H. Law, W. Hung, C. W. Leung, C. W. Wong, J. Y. Zheng, Y. K. Yu, N. S. Lau and T. H. Ng. They have given me support, and a joyful and wonderful university life.

Finally, I am deeply indebted to my family for their unconditional love and support over the years.

This work is dedicated to my family for the support and patience

Contents

List of Figures

List of Tables

Chapter 1

Introduction

With the rapid growth of the Internet and multimedia systems in distributed environments, it is easier for digital data owners to transfer multimedia documents across the Internet. Therefore, there is an increase in concern over copyright protection of digital contents [1, 2, 3, 4]. Traditionally, encryption and control access techniques were employed to protect the ownership of media. These techniques, however, do not protect against unauthorized copying after the media have been successfully transmitted and decrypted. Recently, watermark techniques are utilized to maintain the copyright [4, 5, 6, 7]. In this paper, we focus on engaging the digital watermarking techniques to protect digital multimedia intellectual copyright, and propose a new algorithm particularly for video watermarking purpose.

1.1 Background

Multimedia and network security issues are classically handled through cryptography, however, cryptography ensures confiden-

tiality, authenticity, and integrity only when a message is transmitted through a public channel such as an open network. It does not protect against unauthorized copying after the message has been successfully transmitted. Digital watermarking is an effective way to protect copyright of multimedia data even after its transmission. Watermarking is a concept of embedding a special pattern, watermark, into a multimedia document so that a given piece of copyright information is permanently tied to the data. This information can later prove the ownership, identify a misappropriating person, trace the marked document's dissemination through the network, or simply inform users about the rights-holder or the permitted use of the data [6].

Digital watermarking remains a largely untested field. There is only a very few number of industrial associations have published the requirements for testing watermarking algorithms [8]. Numerous inventive watermarking approaches have been proposed in these few years and most of them focus on digital image watermarking. In recent years, image watermarking technique becomes mature, thus researcher starts to explore a more challenging research topic – digital video watermarking. Most of the proposed video watermarking schemes are based on the techniques of image watermarking and directly applied to raw video or compressed video. However, current image watermarking schemes are not capable of adequately protecting video data [9].

Video watermarking introduces some issues which is not present in image watermarking. Due to large amounts of data and inherent redundancy between frames, video signals are highly suscep-

tible to pirate attacks, including frame averaging, frame dropping, frame swapping, statistical analysis, etc. Applying a fixed image watermark to each frame in the video leads to problems of maintaining statistical and perceptual invisibility. Furthermore, such an approach is necessarily video independent; as the watermark is fixed while the frame changes. Applying independent watermarks to each frame also presents a problem. Regions in each video frame with little or no motion remain the same frame after frame. Motionless regions may be statistically compared or averaged to remove independent watermarks [10]. In addition, video watermarking schemes must not use the original video during watermark detection as the video usually is in very large size and it is inconvenient to store it twice. We propose a new video watermarking scheme to overcome these problems.

1.2 Research Objective

As video copyright protection is strongly concerned, a robust video watermarking scheme is necessary. In order to design a robust, invisible, blind and not removable video watermarking scheme, a survey and investigation has been done on multimedia security issues and multimedia watermarking scheme. Various watermarking scheme schemes are compared and evaluated. Base on these, a new approach and procedures for multimedia security based on watermarking are proposed [11, 12].

A Hybrid scene-based video watermarking scheme with error correcting code and genetic algorithm is proposed. In this scheme, the watermark is decomposed into different parts and

embedded in the frames of different scenes in the video with hybrid approaches. As identical watermark is used within each motionless scene and independent watermarks are used for successive different scenes, the proposed method is robust against the attack of frame dropping, averaging, swapping, interpolation and lossy compression. At the same time, error correcting code is extracted from the video channel and embedded into the audio channel, which provides extra information for recovery of extracted watermark. The performance of the watermarking scheme is enhanced by applying the genetic algorithm (GA) optimization. Moreover, the scheme allows blind retrieval of embedded watermark which does not need the original video.

Video watermarking is not a stand alone technology. It can be associated with different applications to achieve a sophisticated system. This research can be continuous by applying this new developed scheme to specific environment or application and examine its usefulness.

1.3 Contributions

Our research work has the following contributions:

- We have performed a complete survey on the current watermarking technologies. It is noticed that none of the current watermarking schemes can resist all attacks. With this finding, we propose a hybrid watermarking scheme based on scene change analyze [11], error correction codes and genetic algorithm [12, 13].

- We proposed a new scheme which applies scene change detections and scrambled watermarks in a video. The scheme is robust against frame dropping, as the same part of the watermark is embedded into the frames of a scene. For different scenes, different parts of the watermark are used, making the scheme robust against frame averaging and statistical analysis [11]. This scheme is innovative in attacking the problems that are not solved effectively in the past.

- To increase the robustness, the watermark strength of the scheme, we propose several hybrid approaches. The first one is visual-audio hybrid watermarking scheme. As videos consist of both video and audio channels, the robustness of our scheme can be enhanced by including an audio watermark. Consequently, we embed error correcting codes of a video watermark as an audio watermark, which can refine the retrieved watermark during watermark detection [12].

- The second approach is another hybrid with different watermarking schemes. As no existing scheme is resistant against all attacks, we employ the hybrid scheme to embed different parts of a watermark into different scenes. Thus, the proposed scheme is capable of resisting most of the common attacks [12].

- To increase the fidelity, the media quality, of the watermarking scheme, we propose a GA-based watermarking scheme. By employing GA, the quality of the watermarked video is enhanced.

- Experiments have been done on these novel video watermarking schemes to test and show its performance. The robustness of our approach is demonstrated using the criteria of the latest StirMark test [14].
- We compare the proposed scheme with the existing scheme in different aspects and discuss the advantages and the disadvantages of our scheme.

Our approach cultivates an innovative idea in embedding different parts of a watermark according to scene changes, embedding its error correcting codes as an audio watermark, applying a hybrid approach to the proposed scheme and employing GA to optimize the fidelity of the scheme. This approach is never explored in the literature, and its advantages are clear and significant. The effectiveness of this scheme is verified through a number of experiments.

1.4 The Structure of this Thesis

This paper is organized as 5 chapters. The next chapter introduces the issues related to multimedia security and different multimedia watermarking techniques, and a survey on current watermark techniques and video watermarking scheme are provided. Novel video watermarking scheme is described in chapter 3 and the experimental results in Chapter 4 are followed by. Finally, a conclusion would be given in chapter 5.

□ **End of chapter.**

Chapter 2

Literature Review

Nowadays the digital media is easily to be reproduced due to the rapidly growth of internet and the multimedia technologies, this drives to urgent need to resolve the security and copyright protection issues. Therefore, the field of digital watermarking grows extremely fast in these few years [15].

The purpose of a digital watermark is to embed auxiliary information into a digital signal by making small changes that are not perceptible to its intended recipient. For instance, in the case of multimedia watermarking, the hidden signal should not result in any visible or audible distortions. Because the embedded signals enable invisible tags to be attached to digital documents, watermarks are powerful tools that will play a role in solving the growing digital property identification problem [16].

This chapter overviews previous work in digital video watermarking and related fields. We first have a look of two popular security tools, digital signatures and cryptography. Then the principle and the techniques of digital watermarking are dis-

cussed. Besides, the essential differences and the advantages that watermarking techniques provide over these existing technologies are explained. Moreover, the essential ingredients of video watermarking are presented. The following section reviews a number of techniques proposed in the literature. Then different watermarking algorithms are implemented and evaluated. Finally, a comparison among different video watermarking is given.

2.1 Security in Multimedia Communications

In a decade years ago, multimedia documents are rarely available to the mass consumer market. However, as the rapidly development of the pervasive digital information technology, everyone's computer can have high quality video compression, increasing network bandwidth and accessibility, dense portable storage media, and compounding processing power. Nevertheless, these technological advances lead to another crisis. Multimedia users had the ability to tamper with, produce copies of, and illegally redistribute digital content. Without solving this security issue, digital multimedia products and services cannot take-off in an e-commerce setting [17].

Digital signature and cryptography are currently two standardized approaches to protect the digital contents. Digital signature is commonly used to authenticate digital transmissions. It is based on public key cryptography and one-way hash functions. By passing the document through a publicly available one-way hash function, a unique identifier is generated, which is

signed with the owner's private key. Then, a string is produced and it is referred to as the digital signature. In addition to the signed document, the intended recipients obtain public keys from certification authorities [16]. The document is authentic only when it matches with the decrypted signature by applying the hash function.

However, the document and signature are not bound in any noteworthy manner. When transmitting the multimedia documents, they may become separated accidentally in transit or intentionally by a malicious party. Thus, the receiver may not able to verify the authentic document. In addition, this method of tamper detection is too strict for multimedia objects. It dose not allow the document to undergo compression and format changes while still maintaining their authenticity. If just one bit differs from the original, for instance due to lossy compression for efficient network transfer, the hash identifier test will fail.

Use of cryptographically secure license keys is another method for protecting digital intellectual property. The content of the documents are protected form manipulation and stealing during delivery as the assessment of the document is only permitted to those who possess the appropriate key. However the critical flaw in this solution is that after transmitted and delivery the document [17], the permitted recipient is able to access to original proprietary data, which can then be reproduced perfectly and redistributed inexpensively. Thus, This technique is not effective because it does not provide permanent protection for the multimedia content after delivery. Moreover, with this scheme, the intellectual property owner is not able to trace the

responsibilities of pirating the properties.

According to the findings, we notice that an ideal solution must somehow integrate security information directly into the content of the multimedia document and the security information should be inseparable from the document during its useful lifespan. Moreover, the additional information should be perceptually invisible as the multimedia documents are ultimately processed by human viewers or listeners and the contents should not be affected. Finally is the flexibility of the scheme. It should be able to support identification of different copies of the document.

Digital watermarking may be one of the suitable solutions. It is an analogous techniques that have been used to protect valuable hardcopy documents, such as money, cheques and official correspondence, for long time ago. Paper watermarks are faint designs that are embedded by the manufacturer into the paper used to produce such hardcopies. These marks are convincingly hard to fake, and at the same time they do not obstruct the normal processing, i.e. reading, and are impossible to be removed without leaving any engram or causing severe damage to the contents of the document. Digital watermarking technologies strive to achieve these goals in a digital environment by inserting a retrievable watermark directly into the softcopy data stream [17].

2.2 Cryptography

Cryptography is the first technology that content owners would turn to. It is probably the most common method of protecting digital documents and certainly one of the best developed as a science. Before delivery, the content is encrypted and the a decryption key is provided only to those who have permission to access the legitimate copies of the content. Then, the encrypted file can be made available through the Internet, but would be useless to a pirate without appropriate key. After encrypted, the structure of the message is changed. It is meaningless and unintelligible unless it is decrypted [18].

There are two kinds of cryptosystems: symmetric and asymmetric [19]. Symmetric cryptosystems use the same key, known as the secret key, to encrypt and decrypt a message, and asymmetric cryptosystems use one key, named as public key, to encrypt a message and a different key, named as private key, to decrypt it. Asymmetric cryptosystems are also called public key cryptosystems .

Symmetric cryptosystems have a problem: "how do you transport the secret key from the sender to the recipient securely and in a tamper proof fashion?" [19]. If you could send the secret key securely, in theory, you then would simply use that secure channel to send your message instead of encrypting your message with symmetric cryptosystem. Commonly, trusted couriers are used as a solution to this problem.

One example using symmetric cryptosystem is shown in Figure 2.1. Alice and Bob want to communicate in secret, while

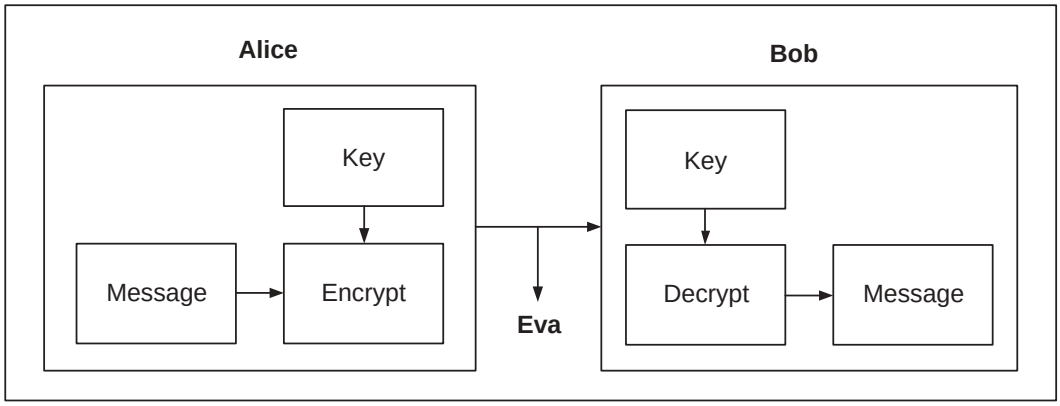

Figure 2.1: Symmetric Cryptosystem

Eve wants to eavesdrop. Alice and Bob could be military jets, on-line businesses or just friends trying to have a private conversation. They cannot stop Eve listening to their radio signals, so they can keep communication by using cryptography.

Alice and Bob exchange a digital key, so they both know it, but it is otherwise secret [20]. Alice uses this key to encrypt messages she sends, and Bob reconstructs the original messages by decrypting with the same key. The encrypted messages are useless to Eve, who does not know the key, and so cannot reconstruct the original messages. With a good encryption algorithm, this scheme can work well, but exchanging the key while keeping it secret from Eve is a problem.

Asymmetric cryptosystem is another more efficient and reliable solution, such as RSA, which is the popular security tool [20]. Asymmetric cryptosystems is different, because it splits the key up into a public key for encryption and a secret key for decryption. It's not possible to determine the secret key from the public key.

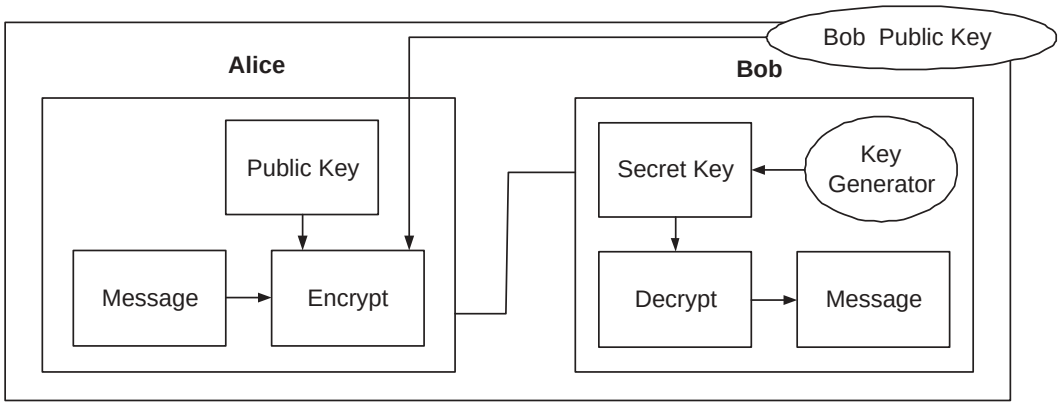

Figure 2.2: Asymmetric Cryptosystems

In the Figure 2.2, Bob generates a pair of keys and tells everybody, including Eve, his public key, while only he knows his secret key. Anyone can use Bob's public key to send him an encrypted message, but only Bob knows the secret key to decrypt it. This scheme allows Alice and Bob to communicate in secret without having to meet.

2.3 Digital Watermarking

Watermarking technique is a particular embodiment of multimedia security. Digital Watermark is defined as a digital signal or pattern inserted into a digital data, which can also be referred to copyright information. Watermarking is a key process in the protecting copyright ownership of electronic data, including image, videos, audio ... etc. The term watermarking comes from using the invisible ink to write secret messages [18]. The additional requirement for watermarking is robustness. Even if the existence of a watermark is known, such as the case in public watermarking schemes, it should be ideally impossible for an attacker to remove or destroy the embedded watermark without rendering the cover object unusable. Generally, watermark has three distinct properties imperceptible, inseparable from the work, and undergoes the same transformation as the work [21].

A simple watermarking idea is shown in Figure 2.3. Watermark is a design of the watermark signal $\boldsymbol{W}$ to be added to the host signal. The watermark signal, apart from depending on the watermark information $\boldsymbol{W'}$, may also depend on a key $\boldsymbol{K}$ and the host data $\boldsymbol{I}$ into which it is embedded, shown in Equation 2.1

$$W = f_0(I, K, W') \tag{2.1}$$

In watermarking algorithm, the host data $\boldsymbol{I}$, such as stego-image, is input to the watermarking algorithm and the algorithm watermarks the image with a watermark $\boldsymbol{W}$ and output the watermarked image $\boldsymbol{I'}$ with the Equation 2.2:

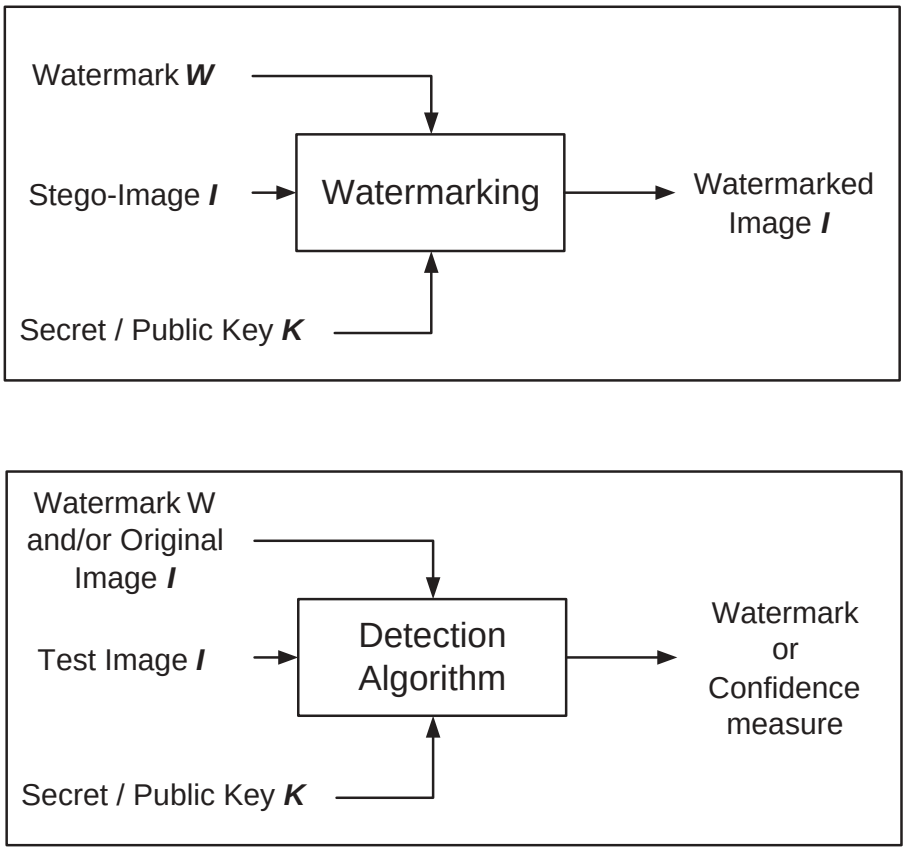

Figure 2.3: Watermarking Embedding and Detection Scenario

$$I \bigoplus W \longrightarrow I' \tag{2.2}$$

Verification algorithm is a design of the corresponding extraction method that recovers the watermark information from the signal mixture, perhaps with the help of the key and the original, shown in Equation 2.3 .

$$I' = g(I, I', K) \tag{2.3}$$

2.4 Essential Ingredients for Video Watermarking

Watermarking systems can be characterized by a number of defining properties including embedding effectiveness, fidelity, data payload, blind or informed detection, false positive rate, robustness, security, cipher and watermark keys, modification and multiple watermark, cost, tamper resistance, unobtrusiveness, ready detection, unambiguous, sensitivity, and scalability. The relative importance of each property is dependent on the requirement of the application and the role the watermark will play. Some of them are common to most practical applications. In this section, such general requirements are listed and briefly discussed. The analysis focuses on image and video watermarking.

2.4.1 Fidelity

What requirements should an ideal watermarking system have? The first requirement would clearly be that of Fidelity [22]. A watermarking system is of no use to anyone if it distorts the cover image to the point of being useless, or even highly distracting. Ideally, the watermarked image should be perceptually invisible even on the highest quality equipment.

Although visible watermarks tend to be more robust, for general purpose applications it is desirable for the embedded mark to be imperceptible to the human eye or ear. Invisibility is that degree that an embedded watermark remains unnoticeable when

a user views the watermarked contents. So far researchers have tried to hide the watermark in such a way that it is impossible to be noticed. However this requirement conflicts with other requirements such as tamper resistance and robustness.

2.4.2 Robustness

The ideal watermark must also be highly robust, entirely resistant to distortion introduced during either normal use, i.e. unintentional attack, or a deliberate attempt to disable or remove the watermark present, i.e. intentional, or malicious attack. Unintentional attacks involve transforms that are commonly applied to images during normal use, such as cropping, resizing, contrast enhancement... etc.

Robustness is the resilience of an embedded watermark against removal by signal processing. The use of music, images and video signals in digital form, commonly involves many types of distortions, such as lossy compression, or, in the image case, filtering, resizing, contrast enhancement, cropping, rotation and so on. For watermarking to be useful, the mark should be detectable even after such distortions occurred. It is a common opinion [18] that robustness against signal distortion is better achieved if the watermark is placed in perceptually significant parts of the signal. This depends on the behavior of lossy compression algorithms, which operate by discarding perceptually insignificant data not to affect the quality of the compressed image, audio or video.

A particularly interesting form of unintentional attack is that

of image compression. Meerwald [23] points out that lossy compression and watermarking are inherently at odds; watermarking seeks to encode information in extra bits that compression hopes to remove. Thus, ideal watermarking and compression systems are most likely inherently exclusive.

In malicious attacks, an attacker deliberately tries to disable the watermark, often through a geometric distortion or the addition of noise. A final note is that robustness can include either resilience to attack, or complete fragility. It may be the case that some watermarking systems may require the watermark to totally destroy the cover object if any tapering is present [24].

Consequently, watermarks hidden among perceptually insignificant data are likely not to survive compression. In the image watermarking case, the resistance to geometric manipulations, such as translation, resizing, rotation and cropping is still an open issue, yet such operations are very common and a solution needs to be found before watermarking techniques are successfully applied to image copyright protection.

Most of Video watermarking scheme base on the techniques of the image watermarking. But video watermarking introduces some issues not present in image watermarking. Video signals are highly susceptible to pirate attacks, including frame averaging, frame dropping, frame swapping, statistical analysis, interpolation etc.

Petitcolas [25] provides us with a rough set of reliability or robustness metrics, shown below in Table 2.1.

Table 2.1: Basic Robustness Requirements

	Level Zero	Low Level	Moderate
Standard JPEG Compression Quality	100-90	100-75	100-50
Color Correction (GIF)	256	256	16
Cropping	100-90	100-75	100-50
Gamma Correction		0.7-1.2	0.5-1.5
Scaling		1/2-3/2	1/3-2
Rotation		± 0-2 deg.	± 0-5 deg., 90 deg.
Horizontal Flip		yes	yes
Uniform Noise		1-5	1-15
Contrast		± 0-10	± 0-25
Brightness		± 0-10	± 0-25
Median Filter			3×3

2.4.3 Use of Keys

Another property of an ideal watermarking system is that it implement the use of keys to ensure that the approach is not rendered useless the moment that the algorithm becomes known [22]. It may also be a goal that the system utilizes an asymmetric key system such as in public/private key cryptographic systems. Although private key systems are fairly easy to implement in watermarking, asymmetric key pairs are generally not. The risk here is that embedded watermarking systems might have their private key discovered, ruining security of the entire system. This was exactly the case when a single DVD decoder implementation left it's secret key unencrypted, breaching the entire DVD copy protection mechanism.

2.4.4 Blind Detection

Blind detection refers to the ability to detect the watermark without access to the original document. Because of the immense size of uncompressed video files and the difficulty of indexing them to search for a specific frame, it is an especially important requirement in video watermarking.

2.4.5 Capacity and Speed

Slightly less important requirements of an ideal watermarking system might be capacity, and speed. A watermarking system must allow for a useful amount of information to be embedded into the image. This can range from a single bit all the way up to multiple paragraphs of text. Furthermore, in watermarking systems destined for embedded applications, the watermark detection (or embedding) may not be overly computationally intensive as to preclude its use on low cost micro-controllers

Capacity is that amount of information that can be expressed by an embedded watermark. Theoretical capacity of embedded watermarks has been examined using information-theoretic concepts. Depending on the application at hand, the watermarking algorithm should allow a predefined number of bits to be hidden. General rules do no exist here, however, in the image case, the possibility of embedding into the image at least 300-400 bits should be granted. In any case, system designers should keep well in mind that the number of bits could be hidden into data is not unlimited; but very often is fairly small.

2.4.6 Statistical Imperceptibility

The last possible requirement of an ideal watermarking system is that of statistical imperceptibility [25]. The watermarking algorithm must modify the bits of the cover in such a way that the statistics of the image are not modified in any telltale fashion that may betray the presence of a watermark. This requirement is not quite as important here as it is in steganography, but some applications may require it.

2.4.7 Low Error Probability

Even in the absence of attacks or signal distortions, the probability of failing to detect the watermark, i.e. false-negative, and of detecting a watermark when, in fact, one does not exist, i.e. false-positive, must be very small. Usually, statistically based algorithms have no problem in satisfying this requirement; however such ability must be demonstrated, if watermarking is to be legally credible.

2.4.8 Real-time Detector Complexity

For consumer-oriented watermarking applications, it is important that the complexity of the detection and extraction algorithms be low enough to execute within the specified real-time deadlines.

2.5 Review on Video Watermarking Techniques

As a method of intellectual property protection, digital watermarks have recently stimulated significant interest and become a very active area of research. Although watermarking is a recent field of research, many techniques have already been proposed both in the academic as well as in the industry. Various techniques are applied in watermarking algorithms. They can be classified into different types based on the offered functionalities. In this section, a brief review of the current video watermarking technologies is presented.

A digital document can be authenticated with what is known as a digital watermark. A watermark is a secret code or image incorporated into an original content, which acts to verify both the owner and content of the document. The use of perceptually invisible watermark is one of the copyright protection. A watermarking algorithm consists of three parts: watermark, marking algorithm and verification algorithm.

Each owner has a unique watermark. The marking algorithm incorporates the watermark into the image or video. The verification algorithm authenticates the content, determining both the owner and the integrity of the content. A variety of imperceptible watermarking schemes have been proposed over the past few years. Numerous watermarking research tasks have proposed many watermark techniques in terms of various application areas. Moreover, different insertion and extraction methods can be found. We can classify the watermarking techniques according to several points of view, as shown in Table 2.2. In this

section, we focus on analyzing the video watermark processing methods.

Digital watermarking can be applied to many different types of documents, including text [26], audio [27, 28, 29, 30, 31, 32], image [33, 34, 35, 36, 37, 38, 39] and video [40, 41, 42, 43, 44, 45]. Watermark techniques can be classified into visible [46, 47] and invisible [29, 30, 40, 41] watermarks. In general, invisible watermarks are mostly used . The oblivious meaningful video watermarking remains a challenging problem since the original video is often unavailable due to videos' bulky volume. Watermarks, on the other hand, need robustness to protect the ownership from various attacks. They can be classified into three categories, robust [48, 49, 50, 51], semi-fragile [52] and fragile [53] watermarks. Different applications would be chosen for different levels of robustness according to the requirement. Applications for copyright protection would require to use a robust watermark. Applications for proving integrity would employ a fragile or semi-fragile watermark. Watermarks to be inserted can also be classified into two types: noise type [54] and image type [41, 42, 43, 44]. A noise type includes pseudo noise, Gaussian random and chaotic sequence. A watermark can be a random sequence with one information bit or multiple-bit meaningful information. The random sequence watermark is more robust in general; however, embedding meaningful watermark is more important in some applications. For image types, there are binary image, stamp, logo and label. Moreover, watermark processing methods are classified into four categories: spatial domain, frequency domain, compression domain and hybrid. Finally,

watermark extraction methods can be classified as private [55], semi-private [56] and public [57] watermarking, according to the necessity of the original media.

Table 2.2: Classification of watermarking according to several viewpoints

Classification		Contents	
Inserted media category		Text [26] Image [27, 28, 29, 30, 31, 32] Audio [33, 34, 35, 36, 37, 38, 39] Video[40, 41, 42, 43, 44, 45]	
Perceptivity of watermark		Visible [46, 47] Invisible[29, 30, 40, 41]	
Robustness of watermark		Robust [48, 49, 50, 51] Semi-fragile [52] Fragile [53]	
Inserting watermark type		Noise [54] Information tagging Image [41, 42, 43, 44]	
Processing method	Spatial domain	LSB Image checksum [58] Patchwork [59] Random function[60]	
	Frquency domain	Look-up table	
		Spread spectrum	DCT [61, 62] Wavelet (DWT) [63, 64, 65, 66] Fourier (DFT)[67, 68]
	Compression domain	MPEG1[42, 69, 70] MPEG2[71, 72] MPEG4[26, 73] JPEG2000 [74]	
	Hybrid	Visual-audio [75, 76] Different watermarks [77] Different watermarking scheme [78]	
Necessary data for extraction		Private [55] Semi-private [56] Public [57]	

2.5.1 Video Watermarking

Many digital watermarking schemes have been proposed in the literature for still images and videos. Most of them operate on uncompressed videos [10, 79, 80], while others embed watermarks directly into compressed videos [72, 81].

Recently, researchers tend to investigate video watermarking techniques that is robust and invisible. These schemes can be distinguished in terms of the domain that the watermark being embedded or detected, their capacity, real-time performance, the degree to which all three axes are incorporated, and their resistance to particular types of attacks [17]. A classification map of the existing video watermarking techniques is presented in Figure 2.4. They can be divided into 3 main groups based on the domain that the watermark is embedded, they are spatial domain, frequency domain and MPEG coding structure based. Most of the proposed video watermarking scheme based on the techniques of the image watermarking and applied to raw video or the compressed video. As some issue in video watermarking is not present in image watermarking, such as video object and redundancy of the large amount video data, researchers have make use of those characteristics to develop different schemes. In the following sections, each class of algorithms is briefly described. Besides, we present the important idea, strength and limitation introduced by these schemes.

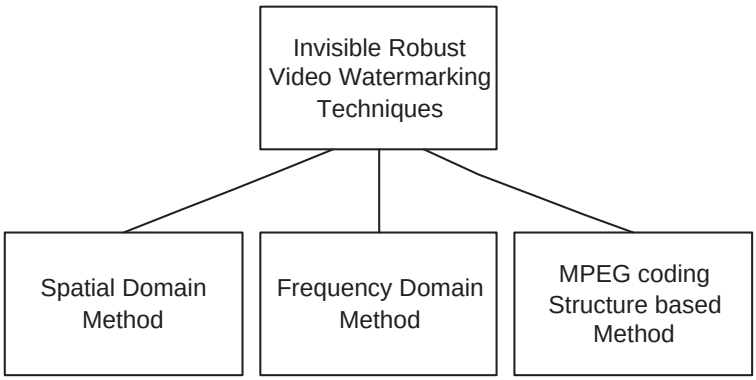

Figure 2.4: Classification map of existing digital video watermark techniques

2.5.2 Spatial Domain Watermarks

We first review the video watermarking techniques in the spatial domain. Algorithms in this class generally share the following characteristics:

- The watermark is applied in the pixel or coordinate domain.
- No transforms are applied to the host signal during watermark embedding.
- The watermark is derived from the message data via spread spectrum modulation.
- Combination with the host signal is based on simple operations, in the pixel domain.
- The watermark can be detected by correlating the expected pattern with the received signal.

The main strengths of pixel domain methods are that they are conceptually simple and have very low computational complexities. As a result they have proven to be most attractive for

video watermarking applications where real-time performance is a primary concern. However, they also exhibit some major limitations: The need for absolute spatial synchronization leads to high susceptibility to de-synchronization attacks; lack of consideration of the temporal axis results in vulnerability to video processing and multiple frame collusion; and watermark optimization is difficult using only spatial analysis techniques. The three methods that fall into this class can be distinguished by the dimensionality of the watermark pattern. Techniques based on 1D and 2D spread spectrum modulation, and 3D Code Division Multiple Access (CDMA) modulation have been proposed.

Several different methods enable watermarking in the spatial domain. The simplest is to just flip the lowest-order bit of chosen pixels in a grey scale or colour image. This will work well only if the image is subjected to any human or noisy modification. A more robust watermark can be embedded in an image in the same way that a watermark is added to paper. Such techniques may superimpose a watermark symbol over an area of the picture and then add some fixed intensity value for the watermark to the varied pixel values of the image. The resulting watermark may be visible or invisible depending upon the value (large or small, respectively) of the watermark intensity. One disadvantage of spatial domain watermarks is that picture cropping, which is a common operation of image editors, can be used to eliminate the watermark.

Spatial watermarking can also be applied using colour separation. In this way, the watermark appears in only one of the colour bands. This renders the watermark visibly subtle so that

it is difficult to detect under regular viewing. However, the watermark appears immediately when the colours are separated for printing or xerography. This renders the document useless to the printer unless the watermark can be removed from the colour band. This approach is used commercially for journalists to inspect digital pictures from a photo-stockhouse before buying non-watermarked versions.

Least Significant Bit Modification

The most straight-forward method of watermark embedding, would be to embed the watermark into the least-significant-bits of the cover object [82]. Given the extraordinarily high channel capacity of using the entire cover for transmission in this method, a smaller object may be embedded multiple times. Even if most of these are lost due to attacks, a single surviving watermark would be considered a success.

LSB substitution however despite its simplicity brings a host of drawbacks. Although it may survive transformations such as cropping, any addition of noise or lossy compression is likely to defeat the watermark. An even better attack would be to simply set the LSB bits of each pixel to one ... fully defeating the watermark with negligible impact on the cover object. Furthermore, once the algorithm is discovered, the embedded watermark could be easily modified by an intermediate party.

An improvement on basic LSB substitution would be to use a pseudo-random number generator to determine the pixels to be used for embedding based on a given "seed" or key [82]. Security

of the watermark would be improved as the watermark could no longer be easily viewed by intermediate parties. The algorithm however would still be vulnerable to replacing the LSB's with a constant. Even in locations that were not used for watermarking bits, the impact of the substitution on the cover image would be negligible. LSB modification proves to be a simple and fairly powerful tool for stenography, however lacks the basic robustness that watermarking applications require.

Correlation-Based Techniques

Another technique for watermark embedding is to exploit the correlation properties of additive pseudo-random noise patterns as applied to an image [83]. A pseudo-random noise (PN) pattern W(x,y) is added to the cover image I(x,y), according to the equation shown below in Equation 2.4.

$$I_w(x, y) = I(x, y) + k \times W(x, y) \tag{2.4}$$

In Equation 2.4, k denotes a gain factor, and IW the resulting watermarked image. Increasing k increases the robustness of the watermark at the expense of the quality of the watermarked image.

To retrieve the watermark, the same pseudo-random noise generator algorithm is seeded with the same key, and the correlation between the noise pattern and possibly watermarked image computed. If the correlation exceeds a certain threshold T, the watermark is detected, and a single bit is set. This method can easily be extended to a multiple-bit watermark by

dividing the image up into blocks, and performing the above procedure independently on each block.

This basic algorithm can be improved in a number of ways. First, the notion of a threshold being used for determining a logical "1" or "0" can be eliminated by using two separate pseudo-random noise patterns. One pattern is designated a logical "1" and the other a "0". The above procedure is then performed once for each pattern, and the pattern with the higher resulting correlation is used. This increases the probability of a correct detection, even after the image has been subject to attack [83].

2.5.3 Frequency Domain Watermarks

Generally DCT, FFT and wavelet transform are used as the methods of data transformation. In these methods, a watermark that one wishes to embed distributively in overall domain of an original data, and the watermark, is hardly to be deleted once embedded. For transformed domain techniques, they have hierarchical watermarking with Discrete Cosine Transform, subband watermarking techniques, Discrete Wavelet Transform or Discrete Fourier Transform.

The main strength offered by transform domain techniques is that they can take advantage of special properties of alternate domains to address the limitations of pixel-based methods or to support additional features. For instance, designing a watermarking scheme in the Discrete Cosine Transform (DCT) domain leads to better implementation compatibility with popular video coding algorithms such as Moving Pictures Experts Group

(MPEG)-2, and in the shift and rotation-invariant Fourier domains facilitates the design of watermarks that inherit these attractive properties. Besides, analysis of the host signal in a frequency domain is a prerequisite for applying more advanced masking properties of the HVS to enhance watermark robustness and imperceptibility. Generally, the main drawback of transform domain methods is their higher computational requirement. We discuss the details of three methods here: Discrete Cosine Transform, Discrete Wavelet Transform, and Discrete Fourier Transform.

Discrete Cosine Transform

The classic and still most popular domain for image processing is that of the Discrete Cosine Transform, or DCT. The DCT allows an image to be broken up into different frequency bands, making it much easier to embed watermarking information into the middle frequency bands of an image. The middle frequency bands are chosen such that they have minimize they avoid the most visual important parts of the image (low frequencies) without over-exposing themselves to removal through compression and noise attacks (high frequencies) [83].

One such technique utilizes the comparison of middle-band DCT coefficients to encode a single bit into a DCT block. To begin, we define the middle-band frequencies (FM) of an 8×8 DCT block

Discrete Wavelet Transform

Another possible domain for watermark embedding is that of the wavelet domain. The DWT (Discrete Wavelet Transform) separates an image into a lower resolution approximation image (LL) as well as horizontal (HL), vertical (LH) and diagonal (HH) detail components. The process can then be repeated to computes multiple "scale" wavelet decomposition, as in the 2 scale wavelet transform shown below in Figure 2.5.

One of the many advantages over the wavelet transform is that that it is believed to more accurately model aspects of the HVS as compared to the FFT or DCT. This allows us to use higher energy watermarks in regions that the HVS is known to be less sensitive to, such as the high resolution detail bands LH,HL,HH). Embedding watermarks in these regions allow us to increase the robustness of our watermark, at little to no additional impact on image quality [83].

One of the most straightforward techniques is to use a similar

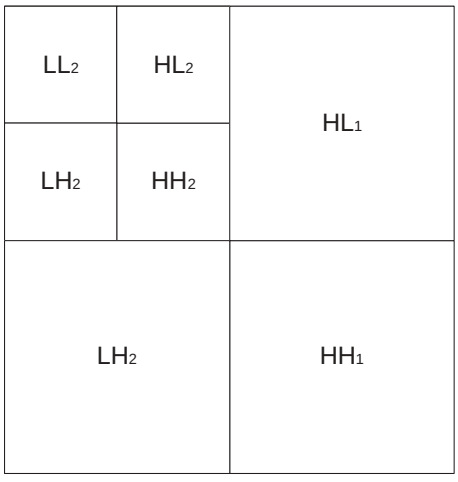

Figure 2.5: 2 Scale 2-Dimensional Discrete Wavelet Transform

embedding technique to that used in the DCT. This can be easily extended to multiple bit messages by embedding multiple watermarks into the image. As in the spatial version, a separate seed is used for each PN sequence, which are then added to the detail coefficients. During detection, if the correlation exceeds threshold for a particular sequence a "1" is recovered; otherwise a zero. The recovery process then iterates through the entire PN sequence until all the bits of the watermark have been recovered.

Furthermore, as the embedding uses the values of the transformed value in embedded, the embedding process should be rather adaptive; storing the majority of the watermark in the larger coefficients. The author [23] claims that the technique should prove resistant to JPEG compression, cropping, and other typical attacks.

Niu and Sun proposed a new wavelet-based digital watermarking in [84]. This paper proposes a method of embedding a digital watermark image in video. In the watermarking, the decomposed watermark image with different resolution is embedding in the corresponding resolution of the decomposed video by means of multiresolution signal decomposing. The proposed method is robust against the attack of frame dropping, averaging and lossy compression. In [85], Serdean et al proposed another DWT-base scheme. This scheme is a high capacity blind video watermarking system invariant to geometrical attacks such as shift, rotation, scaling and cropping. A spatial domain reference watermark is used to obtain invariance to geometric attacks by employing image registration techniques to determine and invert the attacks. A second, high capacity watermark, which carries

the data payload, is embedded in the wavelet domain according to a human visual system model.

Mitchell et al proposed multiresolution video watermarking using perceptual models and scene segmentation in [86]. The watermark consists of static and dynamic temporal components that are generated from a temporal wavelet transform of the video scenes. To generate the watermark, the resulting wavelet coefficient frames are modified by a perceptually shaped pseudo-random sequence representing the author. The noise-like watermark is statically undetectable to thwart unauthorized removal. Furthermore, the author representation resolves the deadlock problem. The multiresolution watermark may be detected on single frames without knowledge of the location of the frames in the video scene.

Discrete Fourier Transform

M. Barni et al proposed a robust watermarking approach for raw video in [87]. This approach first extracts the brightness of the to-be-marked frame, computing its full-frame DFT and then taking the magnitude of the coefficients. The watermark is composed of two alphanumerical strings. The DFT coefficient is altered, then IDFT. Only the first frame of each GOP is watermarked, which was composed of twelve frames, leaving the other ones uncorrupted. It is good robustness to the usual image processing as linear/non-linear filtering, sharpening, JPEG compression and resist to geometric transformations as scaling, rotation and cropping. Decide to watermark one or more frames

in GOP, a trade-off between time spent for marking and the degree of robustness needed for the sequence can be achieve

2.5.4 Watermarks Based on MPEG Coding Structures

Video watermarking techniques that use MPEG-1, -2 and -4 coding structures as primitive components are primarily motivated by the goal of integrating watermarking and compression to reduce overall real-time video processing complexity. Compression in block-based schemes like MPEG-2 is achieved by using forward and bi-directional motion prediction to remove temporal redundancy, and statistical methods to remove spatial redundancy. One of the major drawbacks of schemes based on MPEG coding structures is that they can be highly susceptible to re-compression with different parameters, as well as conversion to formats other than MPEG. There are a number of MPEG-2 and -4-based techniques that have been proposed, including approaches based on GOP modification [69], high frequency DCT coefficient manipulation, DCT block classification [71], [88] and three more robust and general algorithms that will be discussed in detail in this section. The two MPEG-2 watermarking methods considered here embed hidden data by swapping level-adjacent Variable Length Codeword (VLC) codeword and manipulating mean luminance over regions of pixels.

Vassaux et al proposed a video object watermarking which is based on the structure of MPEG-4 in [89]. This paper presents a so-called scrambling technique which allows adapting any classical spread spectrum watermarking scheme operating in the

spatial domain to the Mpeg-4 requirements concerning VO manipulation. This technique can be easily added to the embedding and detection schemes without changing the watermarking algorithm. It modifies some predefined pairs of quantized DCT coefficient in the luminance block of pseudo-randomly selected MBs. It is based on spread-spectrum techniques. Dividing the image into blocks of equal size, them binary sequence is generated using secret key, and then adds to the image. Special decomposition of mpeg-4 include the fact that VO have significant value Watermark information has to present in each VO

In [90], Swanson, et al. presents a watermarking procedure to embed copyright protection into video sequences which is object-based transparent. To address issues associated with video motion and redundancy, individual watermarks are created for objects within the video. Each watermark is created by shaping an author and video dependent pseudo-random sequence according to the perceptual masking characteristics of the video. As a result, the watermark adapts to each video to ensure invisibility and robustness. Furthermore, the noise like watermark is statistically undetectable to prevent unauthorized removal.

Lu and Liao proposed another video object-based watermarking in [91] which is resist to rotation and flipping. Video object is very important concept in Mpeg4 standard. Video object may be purposely cut and pasted for illegal use. In this paper, a robust watermarking scheme foe video object protection is proposed. For each segmented video object, a watermark is embedded by a new technology designed based in the concept of communication with side information. To solve the asyn-

chronous problem caused by object placement, it proposes to use eigenvectors of a video object for synchronization of rotation and flipping.

In [92], Mobasseri proposed direct sequence watermarking using m-frames This scheme applies a direct sequence spread spectrum model to the watermarking of the digital video. First, video signal is modeled as a sequence of bit planes arranged along the time axis. Watermarking of this sequence is a two layer operation. A controlling m-sequence first establishes a pseudorandom order in the bit plane stream for later watermarking. Watermark, defined as m-frames, supplant the tagged bit planes. Moreover, attempts in corrupting the image to destroy the watermark render the video useless before damaging the seal itself. The watermarked video is also robust to video editing attempts such as subsampling, frame reordering etc. The watermark is also identifiable from very short segment of video. Individual frames extracted from the video will also contain the copyright information.

2.6 Comparison between Different Watermarking Schemes

In general, watermarking schemes can be roughly divided into two categories: spatial domain watermark, and transformed domain watermark. We have chosen some representative watermarking schemes in each category for implementation and performed experiments to compare their robustness. They are: Least Significant Bit (LSB) based watermarking scheme [93], threshold-based correlation watermarking scheme [83], Direct sequence watermark using m-frame [94, 95], DFT with template matching [96], Discrete Wavelet Transform (DWT) based watermarking scheme [97], Discrete Cosine Transform (DCT) based watermarking scheme [98] and spread spectrum [99] watermarking scheme. To evaluate the algorithms, the StirMark 4.0 benchmark program [100, 101] and 30 different images are used. The tests are divided into the following sections: PSNR, compression, scaling, cropping, shearing, rotation, row/column removal, and random geometric distortions. Each attack is considered by itself and it is applicable after watermarking. For each image, we assign a score of 1 if the watermark is correctly decoded in the case. A value of zero is assigned if the watermark is incorrect. The comparison is shown in Table 2.3.

From the result, we find that the watermarking schemes in spatial domain are less robust than those in frequency domain. LSB, threshold-based correlation and m-sequence watermarks are perform worse than the other five implemented watermarking algorithms. Therefore, these watermarking algorithms are

Table 2.3: Comparison between different watermarking schemes

Attack Class	LSB	Threshold - based Correlation	m-sequence / m-frame	Spread Spectrum
JPEG Lossy Compression	0.20	0.75	0.7	0.85
PSNR	0.13	0.82	0.89	0.9
Add Noise	0.10	0.7	0.75	0.89
Median Filter	0.21	0.4	0.4	0.35
Row / Column Removal	0.4	0.63	0.7	0.69
Cropping	0.49	0.65	0.75	0.78
Rescale	0.22	0.5	0.62	0.83
Rotation	0.14	0.52	0.61	0.85
Affine	0.15	0.46	0.56	0.76
Geometrical Distortions	0.25	0.42	0.5	0.62
Shearing	0.27	0.3	0.54	0.85

Attack Class	Mid-band DCT	Mid-band DWT	DFT template Matching	Radon Transform
JPEG Lossy Compression	1	0.75	0.74	0.83
PSNR	0.98	1	0.81	0.78
Add Noise	0.95	0.73	0.86	0.75
Median Filter	0.4	0.3	0.25	0.3
Row / Column Removal	0.65	0.5	1	0.75
Cropping	0.62	0.76	0.89	0.85
Rescale	0.53	0.75	0.78	1
Rotation	0.5	0.52	1	0.98
Affine	0.35	0.45	0.98	0.83
Geometrical Distortions	0.64	0.75	0.37	0.75
Shearing	0.35	0.42	1	0.6

classified as fragile or semi-fragile watermarking. They can be applied for the purpose of proving the integrity of a document.

The frequency domain watermarking schemes are relatively more robust than the spatial domain watermarking schemes, particularly in lossy compression, noise addition, pixel removal, rescaling, rotation and shearing. DCT-based watermarking scheme is the most robust to lossy compression. In this approach, an image is broken up into different frequency bands by DCT, making it much more easier to embed watermarking information into the middle frequency bands of the image. The middle frequency

bands are chosen to minimize the change of the most visually important parts of the image (low frequencies) without over-exposing themselves to the removal through compression and noise attacks (high frequencies). Moreover, DWT-based watermarking scheme is the most robust to noise addition.

DFT-based watermarking scheme with template matching can resist a number of attacks, including pixel removal, rotation and shearing. The purpose of the template is to enable resynchronization of the watermark payload spreading sequence. It is a key dependent pattern of peaks, which is also embedded into DFT magnitude representation of the frame. The peaks are not embedded by addition, but rather by modifying the value of the target coefficient, such that it is at least two standard deviations above the mean.

Radon transformation resists attacks by resealing and geometric distortion. In the scheme, invariant watermarks use the Radon transform and higher order spectra. A bispectrum feature vector of the image is used as the watermark. This approach differs from the previous methods in that it embeds watermarks into the phase of the higher order spectra. The Radon embedding grid also outperforms the Fourier-Mellin based methods.

The weakness of the existing algorithms, however, includes: i) The watermark is not robust to attacks which are specifically targeted at to videos, such as frame dropping, averaging and statistical analysis. ii) The bit rate of the watermark is low. Some algorithms embed only one bit information as the watermark. iii) Existing techniques are not aware of the usefulness of the audio channel in a video. iv) None of the existing watermarking

schemes resists to all the attacks. v) A frequency domain watermark is more robust than a spatial domain watermark. To tackle these problems, in this paper, we propose a novel watermarking scheme based on scene changes with a hybrid approach.

□ **End of chapter.**

Chapter 3

Novel Watermarking Schemes

In this chapter, we present the proposed innovative digital video watermarking scheme. A scene-based video watermarking scheme is proposed, which is robust against the attacks of frame dropping, averaging and statistical analysis, which were not solved effectively in the past [11]. Moreover, a hybrid approach is proposed, which can improve the robustness of the watermarking scheme [12, 13]. To enhance the fidelity of the scheme, a GA-based watermarking scheme is presented. In the following sections, the detail of each algorithm is described.

3.1 A Scene-based Video Watermarking Scheme

The new watermarking scheme we propose is based on scene change analysis. Figure 3.1 shows an overview of our watermarking process. In our scheme, a video and a watermark are taken as the input, and the watermark is then decomposed into different parts which are embedded in corresponding frames of different scenes in the original video.

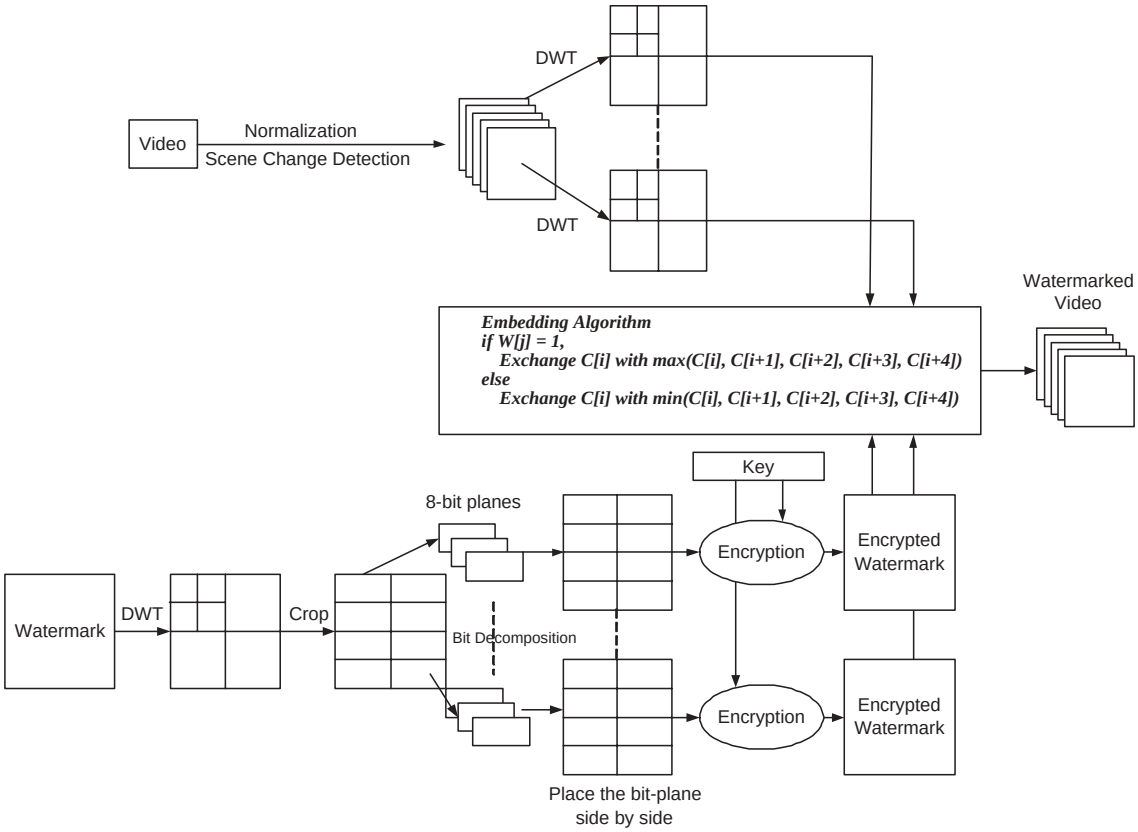

Figure 3.1: Overview of the watermarking process

As applying a fixed image watermark to each frame in the video leads to the problem of maintaining statistical and perceptual invisibility [102], our scheme employs independent watermarks for successive but different scenes. However, applying independent watermarks to each frame also presents a problem if regions in each video frame remain little or no motion frame after frame. These motionless regions may be statistically compared or averaged to remove the independent watermarks [10, 103]. Consequencely, we use an identical watermark within each motionless scene. With these mechanisms, the proposed method is robust against the attacks of frame dropping, averaging, swapping, and statistical analysis. This newly proposed scheme consists of four parts, including: watermark preprocess, video preprocess, watermark embedding, and watermark detection. Details are described in the following sections.

3.1.1 Watermark Preprocess

A watermark is scrambled into small parts in a preprocess, and they are embedded into different scenes so that the scheme can resist a number of attacks towards to the video. A 256-grey-level image is used as the watermark, so 8 bits can represent each pixel. The watermark is first scaled to a particular size as follows:.

$$2^n \leq m, \quad n > 0 \tag{3.1}$$

$$p + q = n, \quad p, q > 0 \tag{3.2}$$

where m is the number of scene changes and n, p, q are positive integers. The size of the watermark is represented as

$$64 \cdot 2^p \times 64 \cdot 2^q \tag{3.3}$$

Then the watermark is divided into 2^n small images with size 64×64. Figure 3.2 shows the procedure of the watermark preprocess with m = 10, n = 3, p = 1, and q = 2.

In the next step, each small image is decomposed into 8 bit-planes, and a large image m_n can be obtained by placing the bit-planes side by side only consisting of 0's and 1's. These processed images are used as watermarks, and totally 2^n independent watermarks are obtained. To make the scheme more robust, the processed watermarks m are transformed to the wavelet domain and encrypted [104]. Sample preprocessed watermarks are shown in Figure 3.3, where (a) is the original watermark, (b)-(i) represent the scrambled watermarks in the spatial domain, and (j) shows the encrypted watermark of (b), i.e., m'_0.

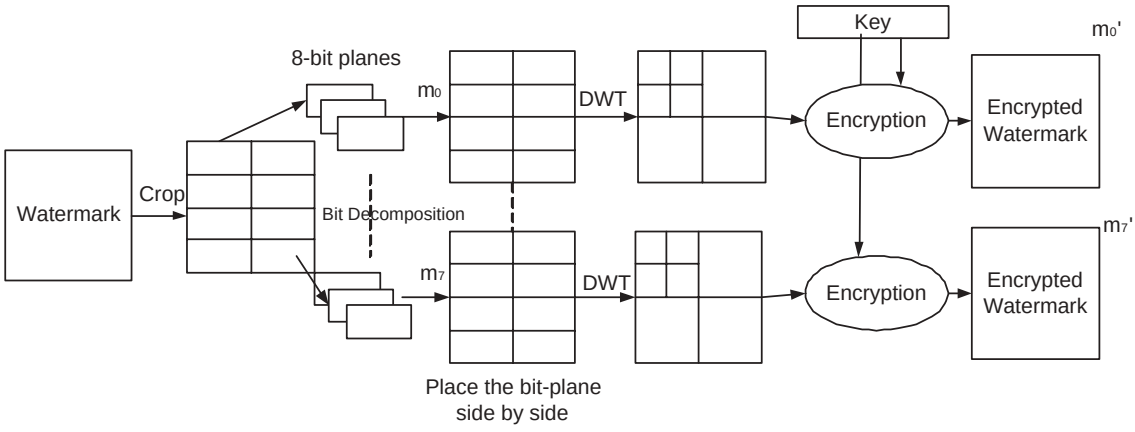

Figure 3.2: Preprocessing the watermark

3.1.2 Video Preprocess

Our watermark scheme is based on 4 levels Discrete Wavelet Transform (DWT). All frames in the video are transformed to the wavelet domain. The frames are decomposed in 4 level sub-band frames by separable 2-D wavelet transform. It produces a low-frequency sub-band LL_4, and three series of high-frequency subbands LH_j, HL_j, HH_j, where $j < 4$. The low frequency sub-band is a lowpass approximation of the original frame, and con-

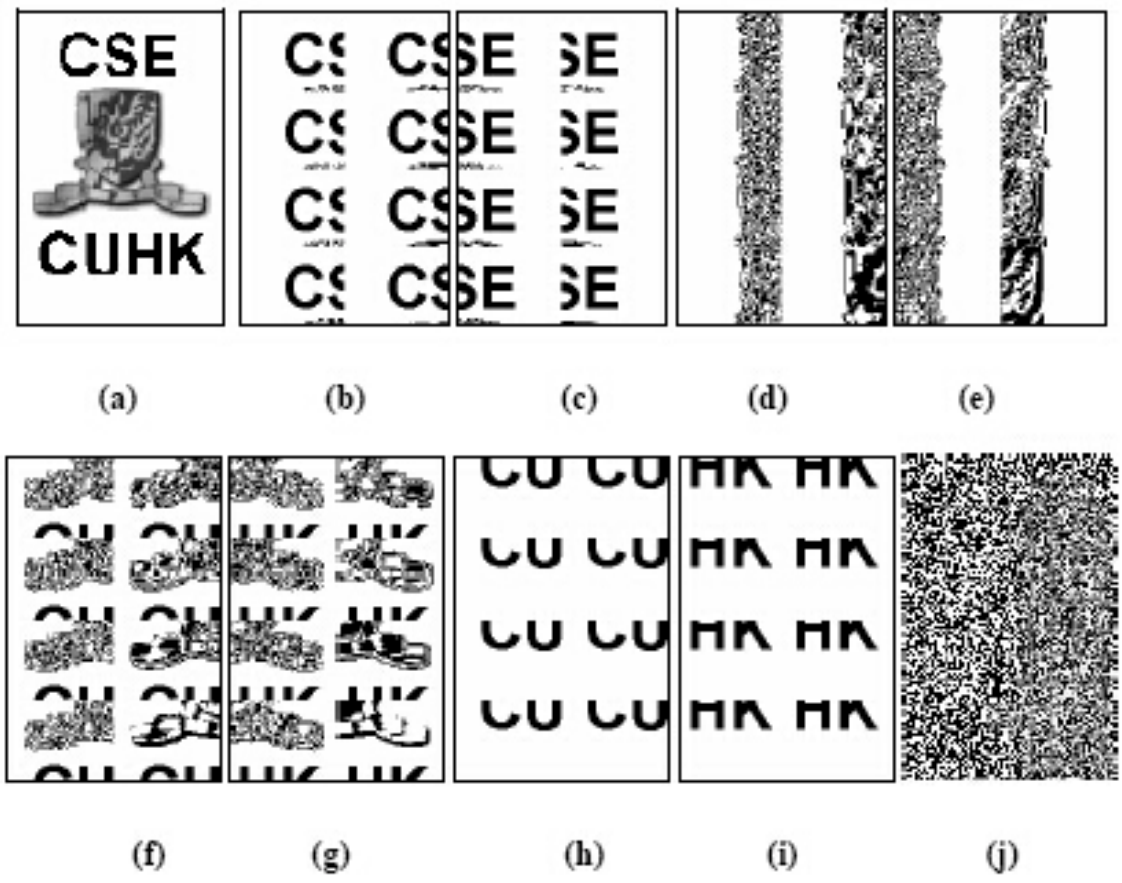

Figure 3.3: (a) Original watermark (b-i) Preprocessed watermark $m_0 - m_7$ (j) Encrypted watermark m'_0

tains most of energy of the frame. The other subbands include edge components of horizontal, vertical and diagonal directions at different scale and resolution, respectively. According to the energy distribution, LL_4 is the most important then LH_j, HL_j, and HH_j. For different levels, the higher the level, the more important the subbands. In our scheme, we only embed the watermark in the middle frequency subbands. Our scheme is based on 4 levels DWT, which is determined by experiments. If less than 4 levels is applied, the capacity of the scheme would be decreased; if larger than 4 levels is applied, the quality of the watermark video is affected.

Moreover, scene changes are detected from the video by applying the histogram difference method on the video stream. The histogram difference method is used for scene change detection. Each frame is coded in 24-bit image, eight bits for each color (red R, green G, blue B). Consequently, each pixel is checked and classified into different classes. For efficiency purpose, only the most significant two bits for each color are considered. Then, the total difference of the whole histogram (H) is calculated as:

$$H = \sum_{i,j=0}^{63} [P_a(i) - P_b(j)]^2 \tag{3.4}$$

where P_a and P_b are the frequency distribution of the pixel level of two images, a and b, i and j are two successive columns in the color histogram. If $H > threshold(T)$, we consider there is a scene change. The threshold is again determined by experiments.

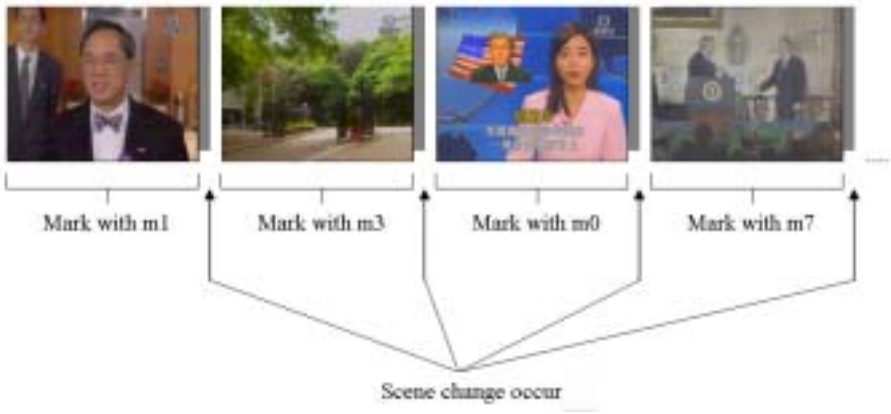

Figure 3.4: Scene change detection

Independent watermarks are embedded in frames of different scenes. Within a motionless scene, an identical watermark is used for each frame. As shown in Figure 3.4, watermark m_1 is used for the first scene. When there is a scene change, another watermark m_3 is used for the next scene. The watermark for each scene can be chosen with a pseudo-random permutation such that only a legitimate watermark detector can reassemble the original watermark.

3.1.3 Watermark Embedding

Watermark is then embedded to the video frames by changing position of some DWT coefficients with the following condition:

$$
\begin{aligned}
&if \;\; W_j \;\; = 1 \\
&\qquad Exchange \; C_i \; with \; max(C_i, C_{i+1}, C_{i+2}, C_{i+3}, C_{i+4}) \\
&else \\
&\qquad Exchange \; C_i \; with \; min(C_i, C_{i+1}, C_{i+2}, C_{i+3}, C_{i+4})
\end{aligned}
$$

(3.5)

where C_i is the i^{th} DWT coefficient of a video frame, and W_j is the j^{th} pixel of a corresponding watermark image [105]. When the watermark $W_j = 1$, we perform an exchange of the C_i with the maximum value among C_i, C_{i+1}, C_{i+2}, C_{i+3}, C_{i+4}. When $W_j = 0$, we perform an exchange of the C_i with the minimum value among C_i, C_{i+1}, C_{i+2}, C_{i+3}, C_{i+4}. With this algorithm, the retrieval of the embedded watermark dose not need the original image. The sequence of watermark coefficients used is depicted in Figure 3.5, where higher frequency coefficients are embedded to higher frequency parts of the video frame, and only the middle frequency wavelet coefficient of the frame (middle frequency sub-band) is watermarked [10].

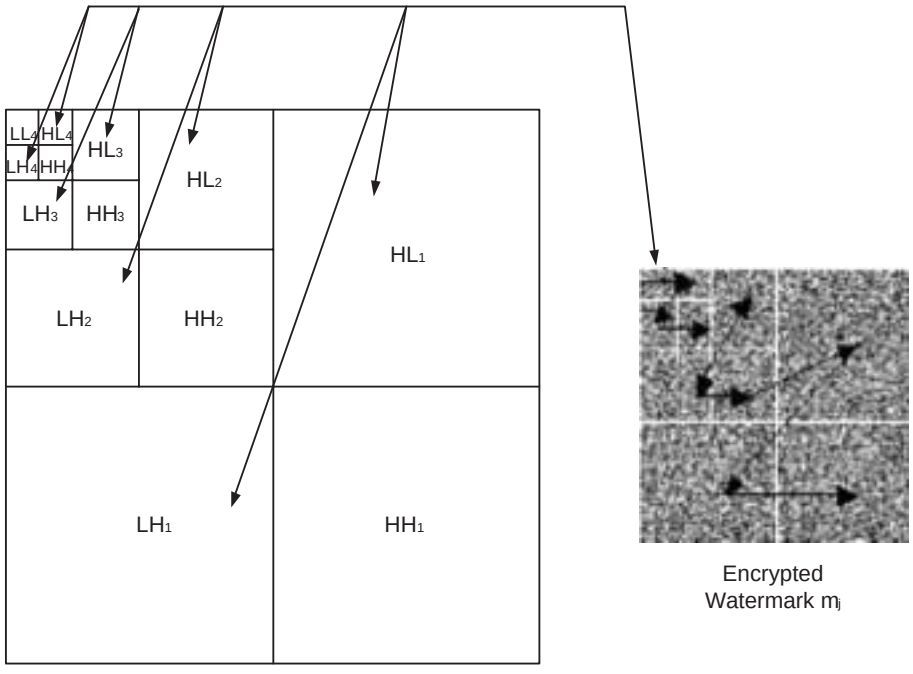

Figure 3.5: Embedding watermarks in a frame

3.1.4 Watermark Detection

The video is processed to detect the video watermark. In this step, scene changes are detected from the tested video. Also, each video frame is transformed to the wavelet domain with 4 levels. Then the watermark is extracted with the following condition :

$$\begin{aligned} if \quad & WC_i > \quad median(WC_i, WC_{i+1}, WC_{i+2}, WC_{i+3}, WC_{i+4}) \\ & EW_j = 1 \\ else & \\ & EW_j = 0 \end{aligned} \tag{3.6}$$

where WC_i is the i^{th} DWT coefficient of a watermarked video frame, and EW_j is the j^{th} pixel of the extracted watermark [106]. When the watermark WC_j is greater than median value among $WC_i, WC_{i+1}, WC_{i+2}, WC_{i+3}, WC_{i+4}$, the extracted watermark is considered as one, i.e., $EW_j = 1$; otherwise, it is considered as zero, i.e., $EW_j = 0$. With this algorithm, the retrieval of the embedded watermark dose not need the original image. This is an important property to video watermarking.

As an identical watermark is used for all frames within a scene, multiple copies of each part of the watermark may be obtained. The watermark is recovered by averaging the watermarks extracted from different frames. This reduces the effect if the attack is carried out at some designated frames. Thus we can combine the 8 bit-planes and recover the 64×64 size image,

i.e., $1/2^n$ part of the original watermark.

If enough scenes are found and all parts of the watermark are collected, the original large watermark image can be reconstructed. This can be shown in Figure 3.6, where the original frame, the watermarked frame, and the extracted watermark are depicted.

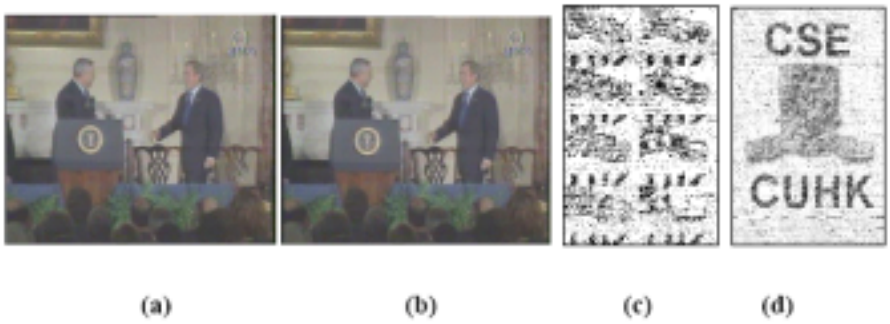

Figure 3.6: (a) Original frame (b) Watermarked frame (c) Extracted watermark corresponding to Figure 3.3(g) (d) Recovered watermark.

3.2 Theoretical Analysis

In this section, performance and the capacity of the proposed watermarking scheme will be calculated.

3.2.1 Performance

In this section, the performance of the proposed algorithm is calculated. Let T be the total number of frame in a video and $n_1 \times n_2$ be the size of the video frame and m the total number of scene change in the video.
Size of the video frame = $n_1 \times n_2$
Size of the watermark = $m_1 \times m_2$
Number of frames = T
Number of scene change = m

Prepare Watermark

To prepare the watermark for the scheme, a watermark is scrambled into small parts in preprocess, and they are embedded into different scenes. The watermark is first scaled to a particular size with the Equation 3.1.

$$2^n \leq m, \quad n > 0 \tag{3.7}$$

$$p + q = n, \quad p, q > 0 \tag{3.8}$$

where m is the number of scene changes and n, p, q are positive integers. The size of the watermark should be

$$64 \cdot 2^p \times 64 \cdot 2^q \tag{3.9}$$

Then the watermark is divided into 2^n small images with size 64×64.

Find N in Equation 3.7 $= O(\log m)$
Find (q, p) in Equation 3.8= $(\frac{n}{2}, n - \frac{n}{2}) = O(1)$

Resize the watermark:
if size of watermark is smaller than video frame $= O(m_1 m_2)$
if size of watermark is greater than video frame $= O(64 \times 64 \times 2^n)$
$= O(m) = O(max[(m_1, m_2),\ m])$

Generating different part of watermark $= 2^n \times 64 \times 64 = O(m)$
Total running time $= O(\log m) + O(1) + O(max[(m_1, m_2),\ m])$
$+ O(m) = O(max[(m_1, m_2),\ m])$

Scene Change

Scene changes are detected from the video by applying the histogram difference method on the video stream. The histogram difference method is used for scene change detection. Each frame is coded in 24-bit image, eight bits for each color (red R, green G, blue B). Consequently, each pixel is checked and classified into different classes. For efficiency purpose, only the most significant two bits for each color are considered. Then, the total difference of the whole histogram (H) is calculated by Equation 3.10:

$$H = \sum_{i,j=0}^{63} [P_a(i) - P_b(j)]^2 \tag{3.10}$$

Scanning to generate the histogram for 1 frame $= n_1 \times n_2$
Create histogram $= n_1 \times n_2$
Compare the histogram $= 64 \times 64$
Total running time $= [2(n_1 \times n_2) + 64] \times T = O(n_1 n_2 T)$

Embedding watermark

Our watermark scheme is based on 4 levels Discrete Wavelet Transform (DWT). All frames in the video are transformed to the wavelet domain. The frame is decomposed in 4 level sub-band frame by separable 2-D wavelet transform. It produces a low frequency sub-band LL_4, and three series of high frequency subbands LH_j, HL_j, HH_j, where $j < 4$.

Running time for DWT
$= 2[n_1 \times n_2 + n_1 \times \frac{n_2}{2} \times 2] + 2[\frac{n_1}{2} \times \frac{n_2}{2} + \frac{n_1}{2} \times \frac{n_2}{4} \times 2] + 2[\frac{n_1}{4} \times \frac{n_2}{4} + \frac{n_1}{4} \times \frac{n_2}{8} \times 2] + 2[\frac{n_1}{8} \times \frac{n_2}{8} + \frac{n_1}{8} \times \frac{n_2}{16} \times 2]$
$= 4n_1 n_2 + 2n_1 n_2 + n_1 n_2 + \frac{n_1 n_2}{2}$
$= \frac{15 n_1 n_2}{2}$
$= O(n_1 n_2)$

When embedding the watermark, only the middle frequency wavelet coefficient of the frame (middle frequency sub-band) is watermarked, i.e., DWT coefficients of HL_1, LH_1, HL_2, LH_2, HL_3, LH_3, HL_4 and LH_4 are watermarked [10].

Total number of pixel to watermark

$= \frac{n_1 \times n_2}{2} + \frac{\frac{n_1}{2} \times \frac{n_2}{2}}{2} + \frac{\frac{n_1}{4} \times \frac{n_2}{4}}{2} + \frac{\frac{n_1}{8} \times \frac{n_2}{8}}{2}$
$= \frac{n_1 \times n_2}{2} + \frac{n_1 \times n_2}{8} + \frac{n_1 \times n_2}{32} + \frac{n_1 \times n_2}{128}$
$= \frac{85 n_1 n_2}{128}$
$= O(n_1 n_2)$

Number of operation for watermark
$= \frac{85 n_1 n_2}{128} \times T$
$= O(n_1 n_2 T)$

After the watermark is embedded, the video frame is inverse-DWT. Running time for IDWT
$= 2[n_1 \times n_2 + n_1 \times \frac{n_2}{2} \times 2] + 2[\frac{n_1}{2} \times \frac{n_2}{2} + \frac{n_1}{2} \times \frac{n_2}{4} \times 2] + 2[\frac{n_1}{4} \times \frac{n_2}{4} + \frac{n_1}{4} \times \frac{n_2}{8} \times 2] + 2[\frac{n_1}{8} \times \frac{n_2}{8} + \frac{n_1}{8} \times \frac{n_2}{16} \times 2]$
$= 4 n_1 n_2 + 2 n_1 n_2 + n_1 n_2 + \frac{n_1 n_2}{2}$
$= \frac{15 n_1 n_2}{2}$
$= O(n_1 n_2)$

Total running time for embedding watermark
$= O(n_1 n_2 T) + 2O(n_1 n_2 T)$
$= O(n_1 n_2 T)$

Finally, Running Time
$= O(max[(m_1, m-2), m]) + O(n_1 n_2 T) + O(n_1 n_2 T) \quad n_1 n_2 \geq m_1, m_2$
$= O(n_1 n_2 T)$

3.2.2 Capacity

Watermarking can be viewed as a communication problem with side information available at the encoder and the decoder. The problem is mathematically defined by distortion constraints, by statistical models for the host signal, and by the information available in the game between the information hider, the attacker, and the decoder. Capacity of the watermark is defined as how much information can be carried by the watermark when it is embedded in an image. In particular, information theory explains why the performance of watermark decoders that do not have access to the host signal may surprisingly be as good as the performance of decoders that know the host signal. Capacity expressions are derived under a parallel-Gaussian model for the host-image source. [107]

In this section, we investigate the watermarking capacity based domain-specified masking effects. We derive the capacity when that power and noise constraints are not uniform across sample, ie., the capacity issue in a variant state channel.

We consider an video as a channel with spatial-variant states, which the power constraint of each state is determined by HSV model or masking effect in some special domains. In this way, each coefficient is considered as an independent random variable with its own noise distribution. We will not consider a coefficient as a communication channel [108, 109] because a channel usually incites its reuse temporally, spatially, or in other domain.

Here we first define the symbols that will be used in this section. Let $X_1, X_2 \ldots X_N$ be the changes of the coefficients in a

discrete video frame due to watermarking. The power constrain of these values are the asking bounds determined by the source coefficient values $S_1, S_2 \ldots S_N$. We define a masking function f s.t. $E(\mathbf{X}\mathbf{X}^T) \leq f(\mathbf{S})$ where $\mathbf{X} = [X_1, X_2 \ldots X_N]^T$ and $\mathbf{S} = [S_1, S_2 \ldots S_N]^T$. In the receiver ed, consider $\mathbf{Y} = S_W$ **-** $\mathbf{S} =$ $\mathbf{X}$ **-** $\mathbf{Z}$ where $\mathbf{Z}$ are the noises added to the coefficients during transmission.

Capacity = C
Host data = S = $[S_1, S_2 \ldots S_N]^T$
Watermark = X = $[X_1, X_2 \ldots X_N]^T$
Power constrain = $E(XX^T) \leq f(S)$
Noise = Z

$$Y = S_W - S = X - Z \tag{3.11}$$

Then, the maximum capacity of these multi-variant symbols in Equation 3.12. We can assume $\mathbf{X}$ and $\mathbf{Z}$ are independent.

$$\begin{aligned} C &= Max_{p(x):E(XX^T)\leq f(s)} I(X;Y) && (3.12)\\ &= Max_{p(x)}[h(Y) - h(Y|X)] && (3.13)\\ &= Max_{p(x)}[h(Y) - h(Z)] && (3.14) \end{aligned}$$

where p(.) represents any probability distribution, I(.;.) represents mutual information and h(.) represents the differential entropy.

According to Theorem 9.6.5 in *Elements of Information Theory* [110], $\mathbf{Y}$ has zero mean and covariance $K = E(XX^T)$, the differential entropy of $\mathbf{Y}$, *i.e.*, h($\mathbf{Y}$) satisfies the following

$$h(Y) \leq \frac{1}{2}\log(2\Pi \exp)^n |K| \tag{3.15}$$

with equality iff $Y \sim N(0, K)$ and $|.|$ is the absolute value of the determinant. Here, this theorem is valid no matter what the range of $\mathbf{K}$ is.

Therefore, from 3.12 and $|E(XX^T) + E(ZZ^T)| = |f(s) + E(ZZ^T)|$, we can see that

$$C = \frac{1}{2}\log(2\Pi\exp)^n|f(S) + E(XX^T)| - h(Z) \tag{3.16}$$

where we assume $f(\boldsymbol{S})$ is diagonal and nonnegative $s.t.$ $|E(XX^T)+E(ZZ^T)| \leq |f(s) + E(ZZ^T)|$. This assumption means that embedded watermark values are mutually independent.

Equation 3.16 is the watermarking capacity in a variant-state channel without specifying any type of noise. It is he capacity given a noise distribution. If we look at Equation 3.16 and Theorem 9.6.5 in [110] again, for all type of noise, we can find that C will be at least

$$\begin{aligned} C_{min} &= \frac{1}{2}\log(2\Pi\exp)^n|f(S) + E(ZZ^T)| && (3.17)\\ &\quad -\frac{1}{2}\log(2\Pi\exp)^n|E(ZZ^T)| && (3.18)\\ &= \frac{1}{2}|f(S) + E(ZZ^T)^{-1} + I| && (3.19) \end{aligned}$$

When noise is Gaussian distribution. If we further assume that noise are also independent in samples, then the watermarking capacity will be:

$$\begin{aligned} C_{min} &= \sum_{i=1}^{n}\frac{1}{2}\log(1 + \frac{P_i}{N_i}) && (3.20)\\ &= \sum_{i=1}^{n}\frac{1}{2}\log(1 + \frac{P_i}{\sigma_n^2}) && (3.21) \end{aligned}$$

where P_i and N_i are the power constrains in the i^{th} coefficient, respectively. It is inserting that even though we use the multi-

variants to derive 3.20 instead of using Parallel Gaussian Channels, their results are the same in this special case.

3.3 A Hybrid Watermarking Scheme

In the previous section, a novel scene-based watermarking scheme is proposed, which is resistant to the attacks of the video properties, including frame averaging, frame dropping, and statistical analysis. However, the scheme dose not improve the robustness against the attacks by image processing on the video frames. Therefore, we propose a hybrid approach to improve the performance and the robustness of the watermarking scheme based on the conclusion drawn from the survey and the properties of a video.

The scene-based watermarking scheme can be improved with two type of hybrid approaches: visual-audio hybrid watermarking and hybrid with different watermarking schemes. Figure 3.7 shows the overall framework of the proposed scheme.

The visual-audio hybrid watermarking scheme applies both video and audio watermarks in a video. Error correcting codes are extracted from the video watermark and embedded as audio watermark in the audio stream. This approach takes the advantage of watermarking the audio channel, because it provides an independent means for embedding the error correcting codes, which carry extra information for watermark extraction. Therefore, the scheme is more robust than other schemes which only use video channel alone. The hybrid approach with different watermarking schemes can further be divided into to two classes: independent scheme and dependent scheme. From the survey, we find that no watermarking scheme can resist all watermark attacks; hybrid with different watermarking schemes can be one

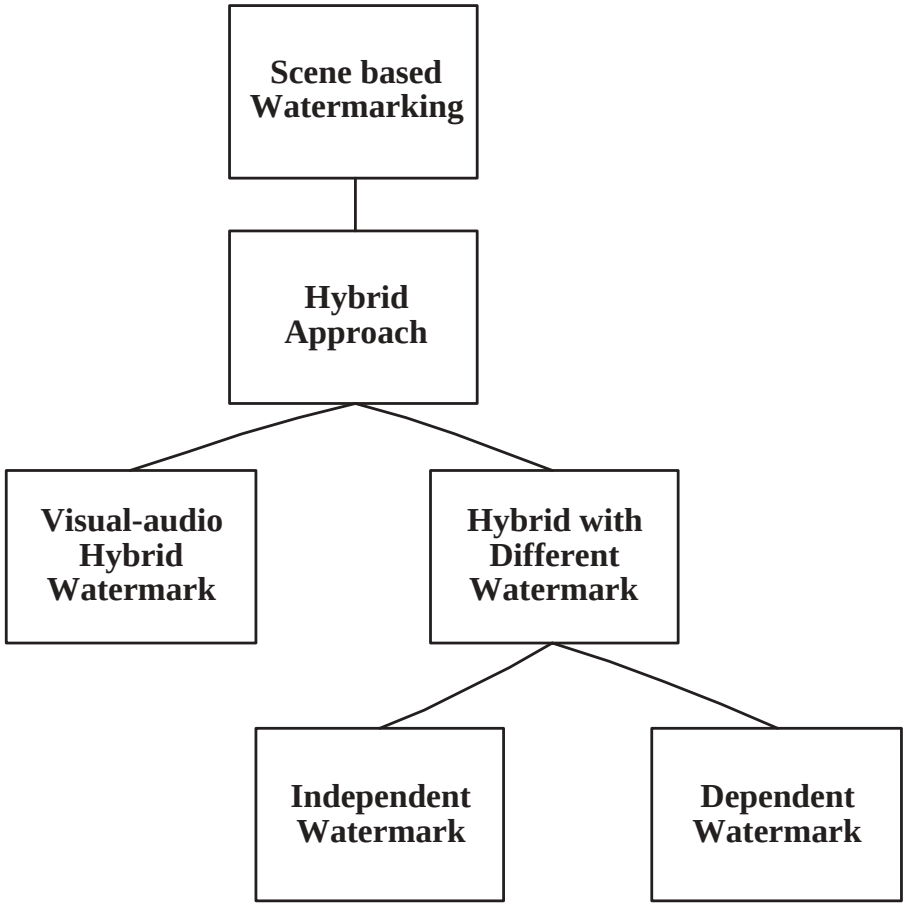

Figure 3.7: Possible improvement for scene based watermarking scheme

of the solutions. It takes advantages of various watermarking schemes by combining them in different ways.

3.3.1 Visual-audio Hybrid Watermarking

The visual-audio watermarking scheme combines a video watermark and an audio watermark. We embed error correcting codes of a video watermark as an audio watermark and refine the retrieved video watermark during detection [11]. Figure 3.8 shows an overview of our visual-audio watermarking process. In

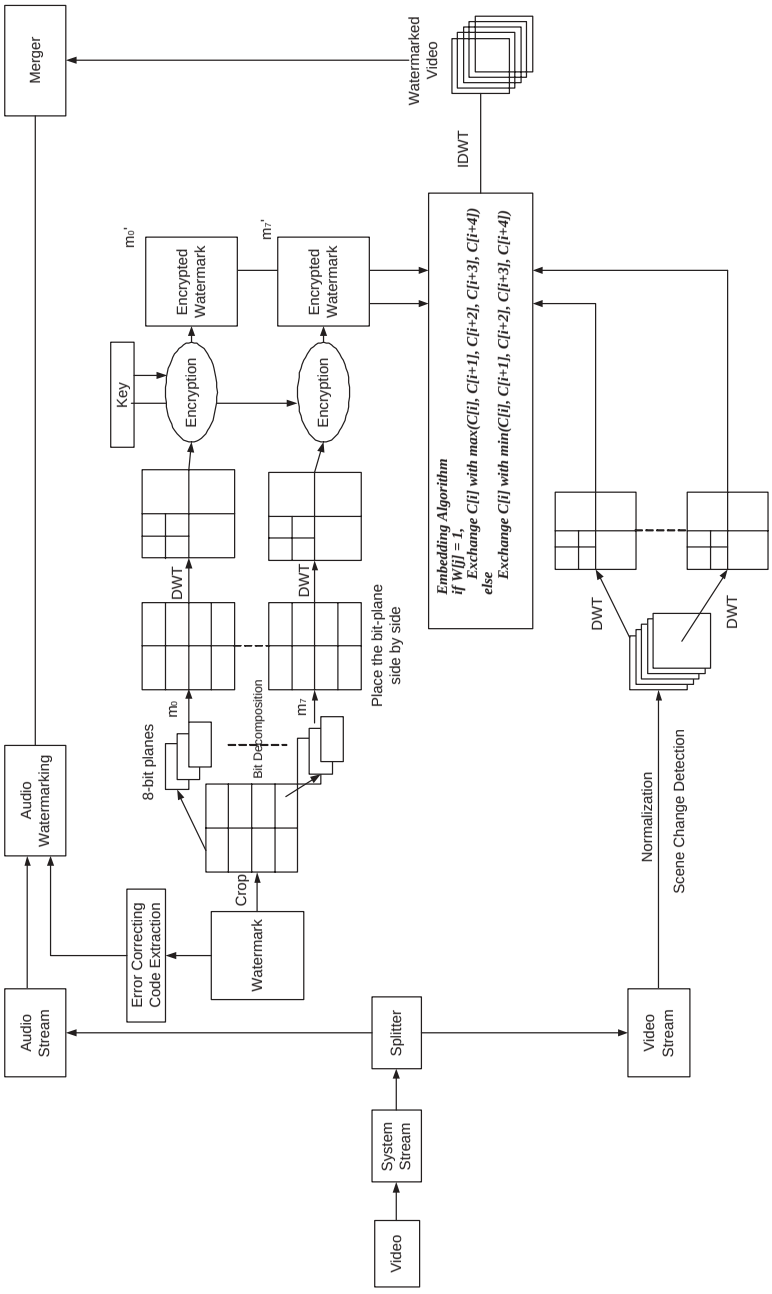

Figure 3.8: Overview of visual-audio hybrid watermarking scheme

our scheme, an input video is split into audio and video streams, which undergo separate watermarking procedures. On the one hand, a video watermark is decomposed into various parts, embedded in corresponding frames of different scenes in the original video. On the other hand, error correcting codes are extracted from the watermarks and embedded as an audio watermark in the audio channel, which in turn makes it possible to correct and detect the changes from the extracted video watermarks. This additional protection mechanism enables our scheme to overcome the corruption of a video watermark, thus the robustness of the scheme is preserved under certain attacks.

Audio Watermark

The watermark embedded in the audio channel provides the error correction and detection capability for the video watermark. In the detection phase, it would be extracted and used for refining the video watermark. Disparate error correction coding techniques can be applied, such as Reed-Solomon coding techniques [103] and Turbo coding [111].

Error correcting codes play an important role in watermarking, especially when the watermark is damaged significantly. Error correcting codes overcome the corruption of a watermark, and make the watermark survive through serious attacks. Moreover, our scheme benefits from audio watermarking as it provides an independent channel for embedding the error correcting codes, which carry extra information for video watermark extraction.

The key to error correction is redundancy. The simplest error correcting code is repeating everything several times. However, in order to keep the audio watermark inaudible, we cannot embed too much information into an audio channel. In our scheme, we apply averaging to obtain the error correction code. Within a small region of an image, the pixels are similar. Hence, an average value of a small region can be fully utilized to estimate the pixels within the particular region. The average value of the pixels in each region is calculated as Equation 3.22

$$\sum_{i=0}^{x}\sum_{j=0}^{y} W_{j\times z+r\times x+s\times y\times z+i} \tag{3.22}$$

where W is a pixel in the image, k is the k^{th} block of the average image, (r, s) is coordinate of region k, z is the width of the image, (x, y) is the coordinate of the pixel in region k, and $x \times y$ is the size of the block. A sample is shown in Figure 3.9.

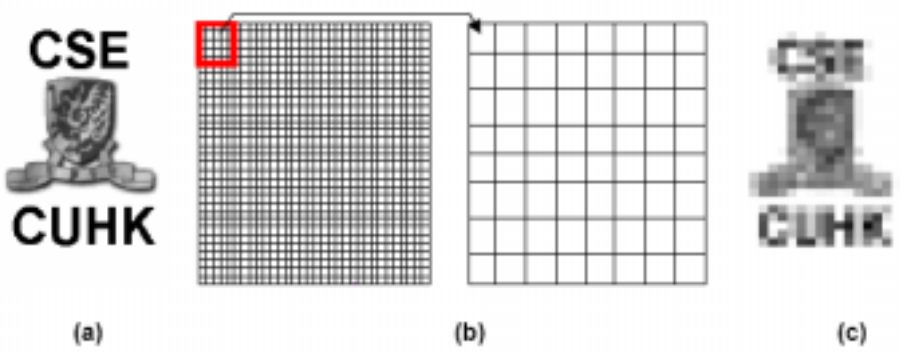

Figure 3.9: (a) Original video watermark (b) Visualization of averaging (c) Audio watermark (average of a)

Audio Watermarking Embedding

In the audio channel, the audio watermarking is based on Modulated Complex Lapped Transform (MCLT) [112]. The MCLT is a 2x over sampled DFT filter bank, used in conjunction with analysis and synthesis windows that provide perfect reconstruction. The MCLT is well suited for noise suppression and echo cancellation.

The MCLT is based on the oddly-stacked time-domain aliasing cancellation (TDAC) filter bank introduced by Princen, Johnson, and Bradley [113]. Its basis functions can be obtained by cosine modulation of smooth windows, in the form [114].

After the wave is extracted from the audio channel, it is transformed from original time domain to frequency (MCLT) domain. The magnitude then is modulated according to the prepared watermark in the previous section.

After addition of the watermark, we generate the time-domain watermarked audio signal by combining the marked vector $\tilde{y}$ with the original phase of $\tilde{x}$, and passing those modified frames through the inverse MCLT. Figure 3.10 shows the detail. For the typical 44.1 kHz sampling, we use a length-2048 MCLT. Only the MCLT coefficients within the 2-7 kHz subband are modified and considered in the detection process, to minimize carrier noise effects as well as sensitivity to downsampling and compression.

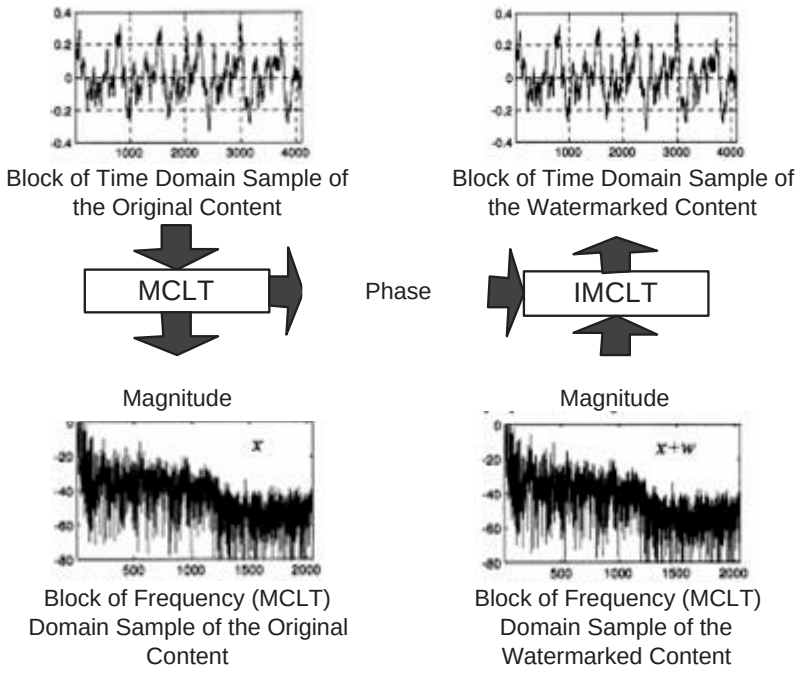

Figure 3.10: Audio watermark embedding with MCLT

Watermarked Frame and Wave

After applying the algorithm described in the previous sections on both video and audio channel, the watermarking process for a video is completed. Figure 3.11 (a) shown the one of the original frame, and Figure 3.11 (b) shown the watermarked frame. Figure 3.12 shown the part of the original wave form, and Figure 3.13 shown the watermarked wave form.

Watermark detection

The watermark is detected through the process whose overview is shown in Figure 3.14. A test video is split into video stream and audio stream, and watermarks are extracted separately by audio watermark extraction and video watermark extraction.

(a) (b)

Figure 3.11: One of the (a) original video frame and (b) watermarked video frame

Then the extracted watermarks undergo a refining process.

The video stream is processed to get the video watermark. At the same time, error correcting codes are extracted from the audio stream and the video watermark extracted is refined by this information with the Equation 3.23:

$$RW_{ij} = \frac{EW_{ij}f + Avg_k \times g}{f + g} \tag{3.23}$$

where RW_{ij} is the refined watermark, EW_{ij} is the extracted video watermark from Equation (7), Avg_k is the extracted audio watermark, k is the k^{th} block of the average image, (i, j) is coordinate of the video watermark, and $f : g$ is a ratio of importance of the extracted video watermark to the audio watermark. In all the subsequent experiments, we assume $f = 0.5$ and $g = 0.5$, $f + g = 1$.

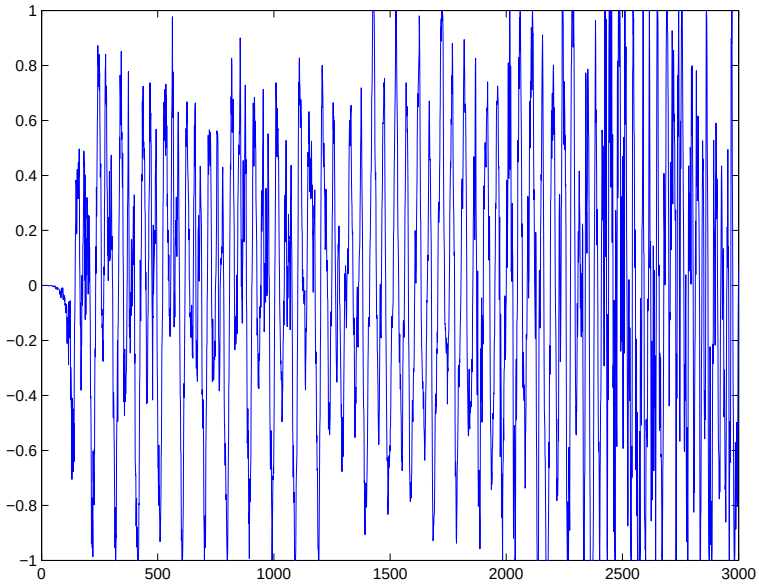

Figure 3.12: Block of samples of the original wave content

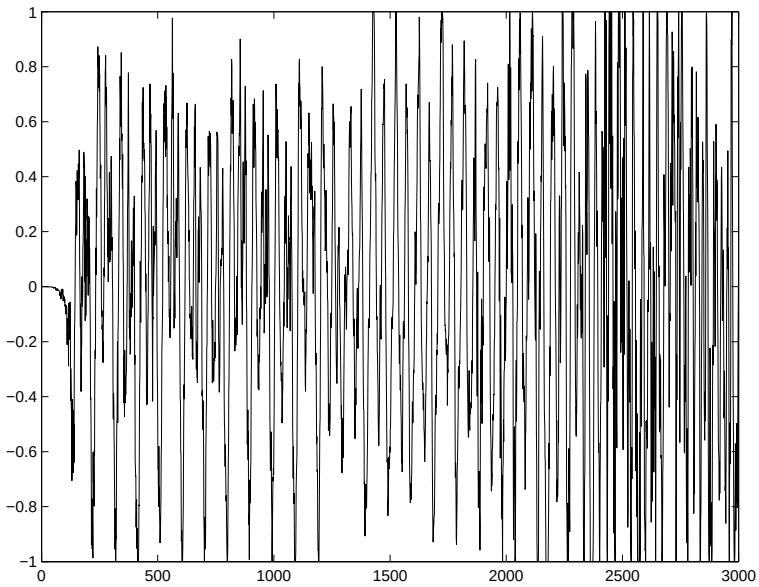

Figure 3.13: Block of samples of watermarked wave content

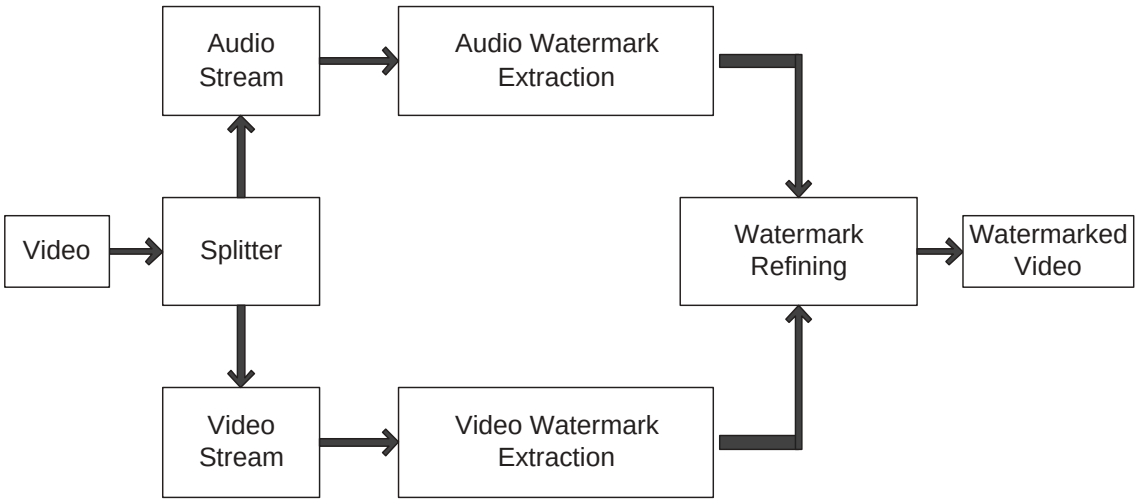

Figure 3.14: Overview of detection of the watermark

3.3.2 Hybrid Approach with Different Watermarking Schemes

No watermarking scheme is found in the current literature to be capable of resisting all watermark attacks. The hybrid approach can be a possible solution. As stated earlier, it can be classified into independent schemes and dependent schemes. Independent watermarking schemes include either different schemes for different scenes or different schemes for different parts of the frame. Dependent watermarking schemes embed a watermark in each frame with several different schemes.

We propose three approaches for the hybrid watermarking schemes. They combine alien schemes in disparate ways. Four watermarking schemes are chosen, each of which strives a different set of attacks. These four schemes are: Discrete Wavelet Transform based watermarking (DWT), Discrete Cosine Transform based watermarking, (DCT), Discrete Fourier Transform

based watermarking (DFT), and Radon Transform based watermarking (RADON). As they embed the watermark in various domains, their robustness properties are preserved. By combining the advantages of these watermarking schemes systematically, various kinds of attacks can be resisted altogether.

Independent Hybrid Watermarking

In the independent hybrid watermarking approach, the applied watermarking schemes do not affect each other. We propose two approaches to combine the employed watermarking schemes: Different schemes for different scenes, and different schemes for different parts of each frame.

Different schemes for different scenes In this approach, a watermark is still decomposed into different parts which are embedded in the corresponding frames of different scenes in the original video. Each part of the watermark, however, is embedded with a different watermarking scheme. Figure 3.15 illustrates the idea. Within a scene, all the video frames are watermarked with the same part of a watermark by the same watermarking scheme.

When there is an attack on the watermarked video, different watermarking schemes are resistant against it. Consequently, some parts of the watermark still survive after the attack. This approach thus enhances the chance of survival under several attacks, and raises its robustness. The merit is that only part of the watermark is damaged if the watermarked video is attacked, provided that at least one of the watermarking schemes is resistant against the attack. The disadvantage of this approach is

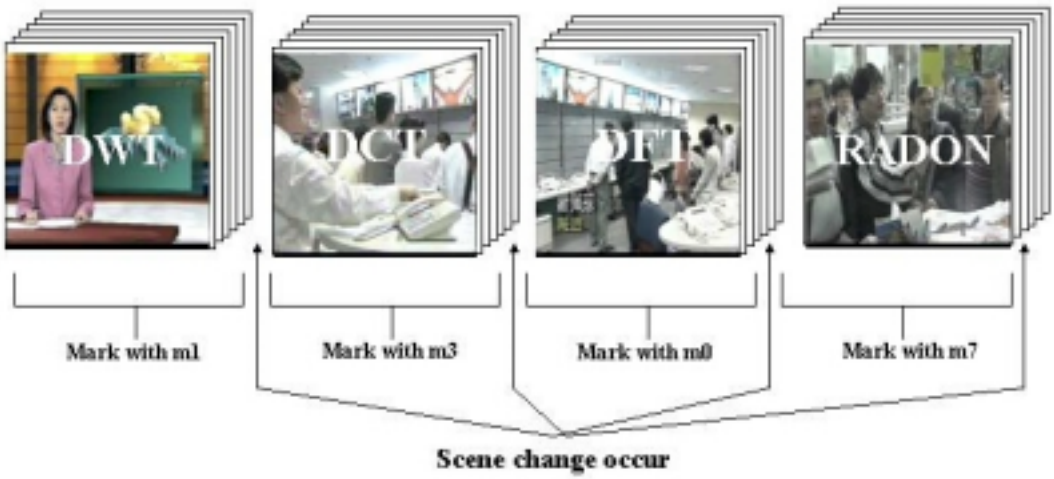

Figure 3.15: Hybrid approach with different scheme for different scene

that the accuracy of the extracted watermark is lower, compared with the other schemes specified to a particular attack.

Different schemes for different parts of each frame This approach is similar to the previous approach. However, four different watermarking schemes are applied to each frame instead of different schemes for different scenes. The idea is illustrated in Figure 3.16, where each video frame is divided into four parts, and the watermark for that frame is also divided into four parts. Then, each part of the watermark is embedded into the frame in different domains.

When a watermarked video is attacked, part of the watermark in each frame may still survive. Therefore, information for every part of the watermark can be retrieved, and the watermark can be approximately estimated. Although the accuracy of the extracted watermark is reduced, it is more resistant against attacks.

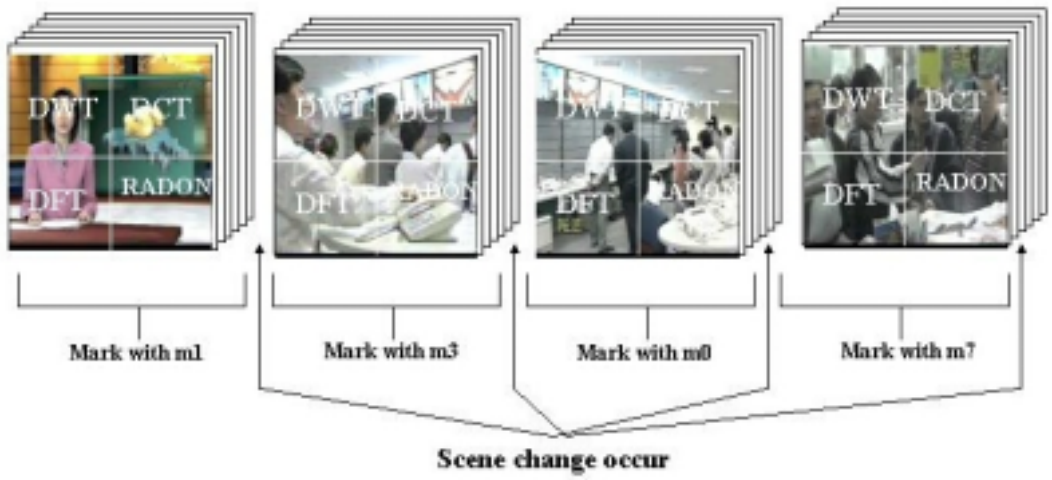

Figure 3.16: Hybrid approach with different scheme for different part of frame

Dependent Hybrid Watermarking

In the dependent hybrid watermarking approach, each frame is embedded with the same watermark serially by different watermarking schemes in various domains. In our approach, again four different watermarking schemes are applied. The frame is first transformed to wavelet domain, and the watermark is embedded. Then the frame is inversely transformed to frequency domain by DCT to embed the same watermark, so on and so forth for DFT and Radon transform, subsequently.

As the watermarks are embedded in different domains, they can compensate each other in resisting against different attacks. Nevertheless, different watermarking schemes may also affect each other when under attacks.

Experiments are done to evaluate the effectiveness of different approach and it is presented in Chapter 4.

3.4 A Genetic Algorithm-based Video Watermarking Scheme

The problem of designing a feasible watermarking scheme can be viewed as an optimization problem with three conflicting goals: higher fidelity (media quality index), better robustness (watermark strength) and larger data capacity. In the previous section, we discussed applying a hybrid watermark approach to improve the robustness of the scheme. In this section, genetic algorithm (GA) is employed to enhance the fidelity [115].

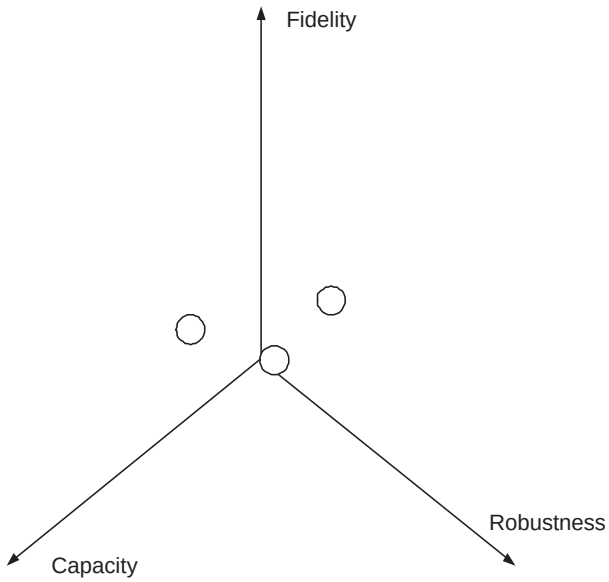

Figure 3.17: The graph of three mutually orthogonal axes representing the capacity, robustness and fidelity of the watermarking scheme

In general, higher fidelity, better robustness, and larger data capacity is three goals that most watermarking scheme would like to achieve. However, since these requirements are conflicting with each other, designing an optimal watermarking scheme

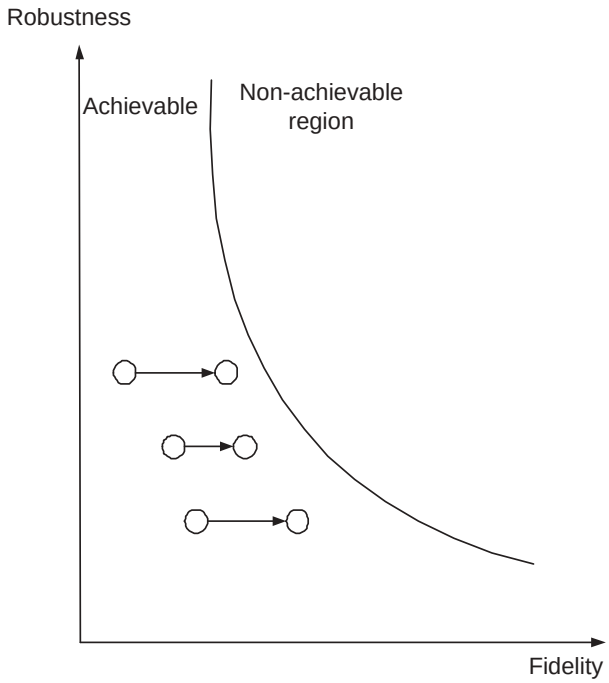

Figure 3.18: The graph of two mutually orthogonal axes representing the robustness and fidelity of the watermarking scheme

has become an inherently difficult and interesting problem. For instant, the fidelity requirement often limits the strength of embedded signals, which consequently constraints the robustness of a watermarking scheme against common or malicious manipulations.

In the existing watermarking schemes, taking the application-specific characteristics into consideration makes reasonable trade-offs among the three requirements [115]. In other words, one can view the embedding process as a selection of feasible embedding points within an acceptable region in the space spanned by three mutually orthogonal axes representing the prescribed requirements, respectively as shown in Figure 3.17.

However, in most watermarking schemes, the process of de-

ciding the trade-off is always done in a heuristic manner, lacking systematic explorations or even optimizations. In this section, a GA-based watermark fidelity enhancing method is proposed, which optimize the performance towards the non-achievable region as show in Figure 3.18

3.4.1 Watermarking Scheme

In the area of evolutionary computation, GA is an important optimization technique [115]. Here we employ this techniques to optimize the performance of our proposed scheme. To model the problem as a GA problem, the fitness function, chromosome and GA operators should be defined. In GA-based optimizations, the problem to be addressed is defined as an objective function that indicates the fitness of any possible solution. According to the problem specific constraints, a population of candidate initialized, named as chromosome, which is a finite-length strings.

During practical GA-based optimization processes, three GA operators: reproduction, crossover, and mutation are applied to the chromosomes repeatedly. Reproduction is a process in which individual chromosomes are copied according to their objective function values, that is, the fitness values. Chromosomes with higher fitness values have higher probability to contribute more offspring in the next generation. The objective function decides the probability of chromosomes' survival or removal during the competition. Crossover is the procedure that pairs of chromosomes exchange portions of their genes to form new chromo-

somes, and consequently, new parameters other than the initial ones can be produced for evaluation. Mutation is the occasional random alternation of the value in some positions of chromosomes. Mutation can be regarded as a random walk through the parameter space. By sparingly using mutation, chromosomes with good performance but not obtainable by reproduction and crossover may also have the chance to be selected. These operators are used repeatedly to obtain successive generations of chromosomes. [116]

Within a generation, only the chromosomes with the higher fitness values can survive. This portion of those will be pass as parent chromosomes to the next generation. After numbers of generation, the chromosomes are optimized. We can obtain the near-optimal solutions of modelled problem. [115].

3.4.2 Problem Modelling

By applying the GA to the scene-based video watermarking scheme, the watermarked video quality is improved while still keep the robustness of the watermark against image manipulation.

The embedding position of the different parts of watermark within the video are simulated as chromosomes. Then several genetic-algorithmic operations are applied to optimize the video frame quality after embedding. In our experiments, we use image quality indicators, Mean Absolute Difference (MAD) to measure the objective function values during optimization.

We choose the watermark embedding positions within a video

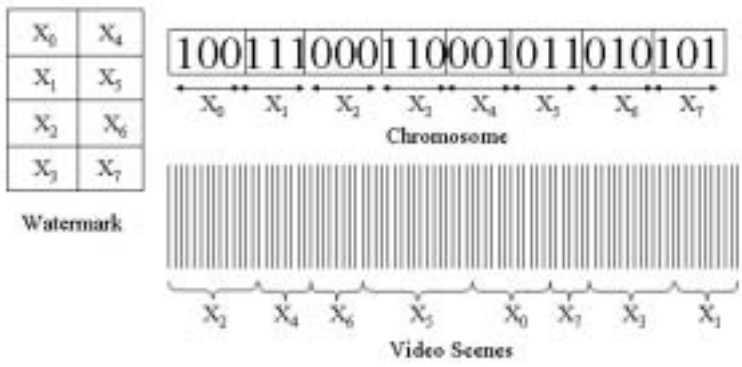

Figure 3.19: A illustrative diagram for GA-based optimization process

as our search space, i.e., search a scene that a particular part of watermark is best to embed to. Then apply the genetic-algorithmic operators to find the best combination of watermark and video scene. A illustrative diagram is shown in the Figure 3.19. Assume a video is consists 8 scenes. The watermark is scrambled into 8 parts and embedded into different scenes. The optimal combination of watermark and scene are shown in the chromosome.

By repeatedly applying the GA operations to each original video frame and watermark image, we can get near-optimal embedding positions for the original frame and the watermark images. Figure 3.20 shows the details of the GA-based optimization process for each video scene.

In the watermark scheme, the input video frames are transformed to frequency domain and the middle-frequency range coefficients of the watermark are modified according to the watermark. The basic idea is that the human eyes are sensitive to the low frequency noise and the quantization step of lossy compression may discard the high frequency components; therefore,

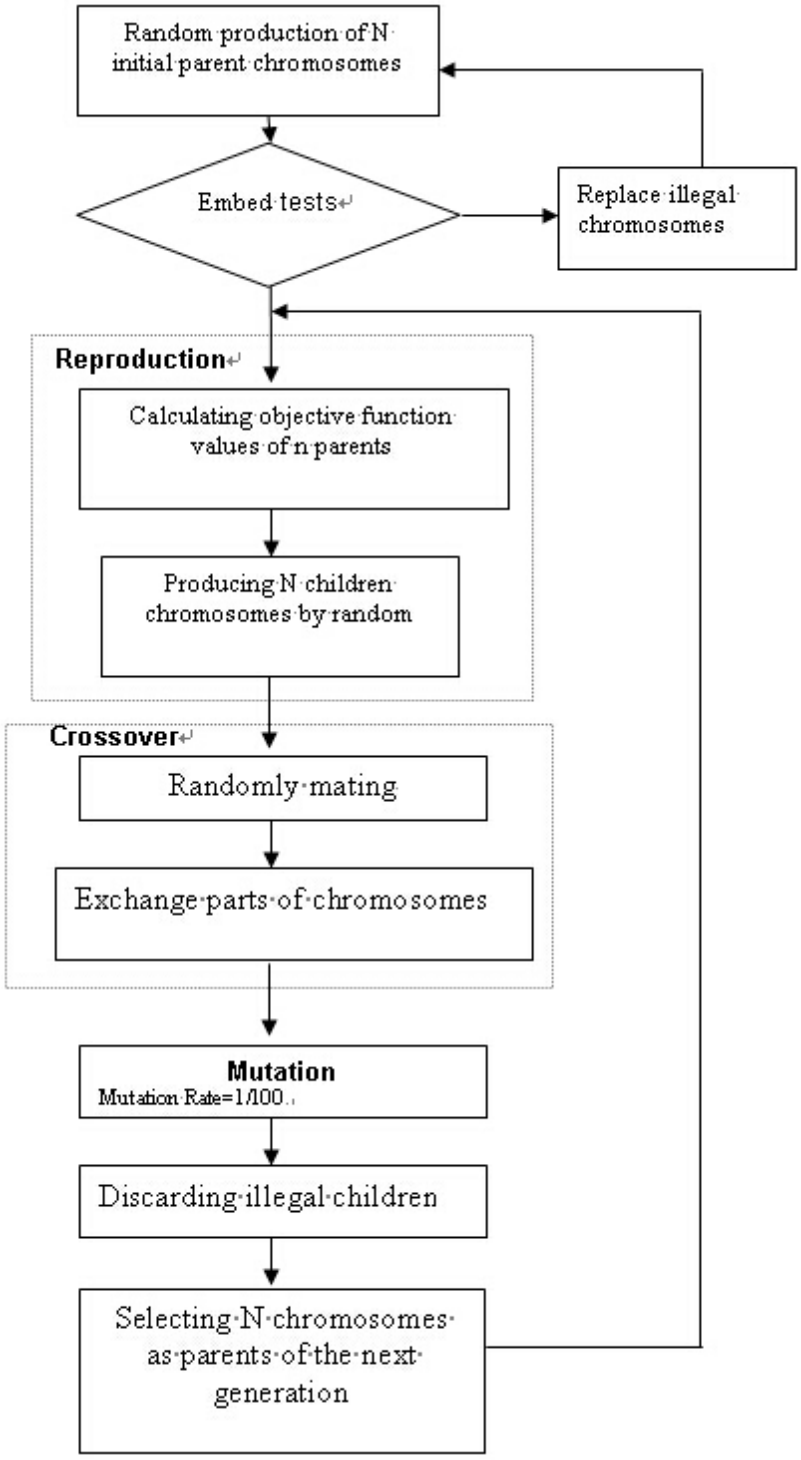

Figure 3.20: The GA-based optimization process for part of watermark

the reasonable trade-off is to embed the watermark into the middle-frequency range of the video frames, i.e. watermarks are embedded by modifying the middle-frequency coefficients within each frame block of the original frame.

3.4.3 Chromosome Encoding

Assume we want to embed a part of the watermark image into one of the scene in a video. The position of the scenes are encode as a chromosome. The number of scene in the video is M and we can user $log_2 M$ bit to represent the position of the scene. The scene of the i^{th} watermark is best to be embedded to X_i in the video can be defined as:

$$\{(X_i) \mid 0 \leq X_i < M, X_i \neq X_j \ if \ i \neq j \ and \ M > 1\} \quad (3.24)$$

The last two constraints of Equation 3.24 imply:

1. There are at least two scene-changes in the video, otherwise, the search space of the GA is 1 and no optimization can be done.

2. In a video, the video frames that have been embedded should not be embedded again.

Figure 3.21 shows the sample chromosome which represents the positions of the watermark that should be embedded to. There are 8 scene changes in the video and 3-bit number need to be used to represent the position of video scenes; therefore, a 24-bit chromosome represents the sequence of video scene to be used.

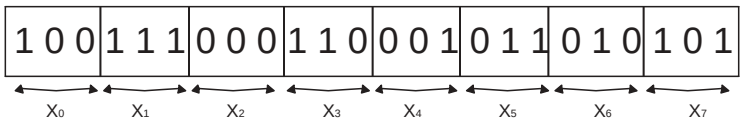

Figure 3.21: A 24-bit chromosome represents the sequence of the scenes to embed

3.4.4 Genetic Operators

The initial chromosomes are generated with the above specification. Then, the genetic-algorithmic operators are apply to the chromosomes for optimization with the fitness function, they are: reproduction, crossover, and mutation.

Fitness function

The fitness function f is a measure of profit we need during optimization. This application of GA aims to improve the fidelity, i.e. the image quality, of the video frames, the reciprocal of image similarity indicators (maximum absolute different MAD) is chosen to be the fitness function. The definition of fitness function f will be:

$$f = \frac{1}{\sum_{x=0}^{7} \sum_{y=0}^{7} |I'(x,y) - I(x,y)|} \tag{3.25}$$

Where I' and I are the intensity values of the same pixel position within a video frame after and before embedding, respectively.

After calculation of the fitness function value (defined in Equation 3.25) for each parent chromosome, we employ the biased-roulette-wheel method [refGA] in GA to generate N children. The basic idea is that the higher a parent chromosome's

fitness function value is, the higher probability it has to contribute one or more offspring in the next generation.

Reproduction

Reproduction is a process that the children chromosomes are generated according to the fitness function values of their parents.

Crossover

After N legal parent chromosomes and M children chromosomes are generated, crossover operator is applied to on these children chromosomes. Firstly, the recently reproduced chromosomes are randomly mated in a pair. Then an integer position p between 1 and the chromosome length (L) minus 1 is selected at random for each pair where $p \leq L - 1$. Finally, each chromosome swaps all the bits between the location of $p+1$ and chromosome length with its mate.

Mutation

The last operator is mutation. Mutation is the random change of bit values with small probability within a chromosome. Mutation introduces some randomness into the optimization, thus, new combination of watermark and video scene are generated. This increases the chance of approaching the optimal, otherwise, the optimization process is very slow. In our experiment, we use 0.05 as the mutation rate, i.e. a bit may change polarity (take complement) at a probability of 0.05.

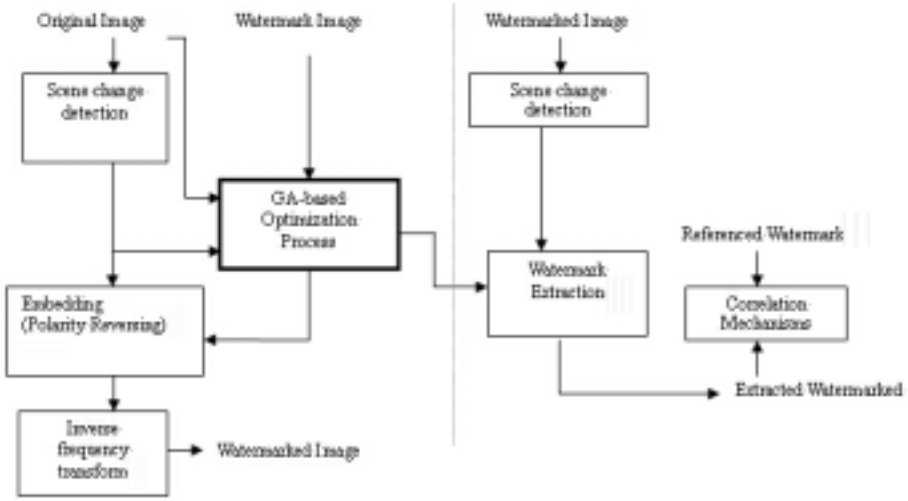

Figure 3.22: The GA-based watermarking algorithm

After performing all the GA operations, more children chromosomes may not satisfy the constraints of embedding positions. After the checking and discarding process, we can get M legal children chromosomes ($M \leq N$). From the $M + N$ chromosomes (M children and N parents), select N chromosomes with the larger fitness function values to be the next parent generations.

Repeat all the operations mentioned above until the number of generations specified has been done, and choose the best chromosome, i.e. the one with the largest fitness function value, to be the sequence of video scenes to be used. Figure 3.22 depicts the block diagram of the proposed watermarking algorithm with GA optimization.

The original video and the watermark are input to the system. Firstly, the scene change analysis is applied to the video and the watermark is preprocessed. Then, the GA-optimization process will find out the almost-optimal combination of water-

mark and video scene. The watermarks are embedded in to the video frames according to the result of the GA optimization. For detection phase, the GA information should be passed to the detector. The watermarked video is passed in to the detection system; the watermark is extracted with the GA information provided by the embedding phase. The original video frame, the watermarked video frame with scene-based watermarking scheme, and the the watermarked video frame with GA-based watermarking scheme are shown in Figure 3.23. There is no significant perceptual different between in Figure 3.23 (b) and (c). However, the enhancement can be indicated by measuring with the quality index and this experimental results will be shown in the following section.

□ **End of chapter.**

(a)

(b)

(c)

Figure 3.23: Comparison between watermarked video with and without GA optimization a) Original video frame (b) Video frame watermarked with scene-based scheme (c) Video frame watermarked with GA-based scheme

Chapter 4

Experimental Results

In this chapter, we present our experimental results on the scene-based watermarking scheme, the hybrid watermarking scheme and the GA-based watermarking scheme. The experiments are basically divided into 2 types: test on robustness and test on fidelity. In the following sections, we present the implementation detail of the proposed schemes and the experimental results.

4.1 Test on Robustness

In this section, the robustness of the scene-based watermarking scheme and the hybrid watermarking scheme is tested. To implement the proposed watermarking scheme, the software VirtualDub [117] is employed. The performance of the new video watermarking scheme is evaluated through several experiments: the experiment with various dropping ratio, the experiment with various number of frame colluded, the experiment with various quality factor of MPEG, and the test of Robustness with StirMark 4.0. Another DWT-based watermarking scheme, which

embeds an identical watermark in all frames, is implemented to compare with the proposed scheme. A video clip with 1526 frames of size 352×288 is used in our experiments. The video consists of 10 scene changes. The experiments are done on a desktop computer with Pentium 4 CPU 2.00GHz and 512MB RAM.

Distinguishable attacks, including frame dropping, frame averaging, lossy compression, and StirMark 4.0 [14], are carried out to the watermarked video to test the robustness of our scheme. The audio channel is also attacked by adding some noises into it. After extracting and refining the watermarks, a quantitative measurement is required to provide an objective judgment of the extraction fidelity. Therefore, a similarity measurement of the extracted and the referenced watermarks can be defined as Equation 4.1, Normalized Correlation:

$$NC = \frac{\sum_i \sum_j W_{ij} \times RW_{ij}}{\sum_i \sum_j W_{ij}^2} \tag{4.1}$$

which is the cross-correlation normalized by the reference watermark energy giving unity as the peak correlation [118], where W_{ij} is the original watermark and RW_{ij} is the refined watermark from Equation 3.23. We use this measurement to evaluate our scheme in the experiments.

The NC values are retrieved when the watermarked video is facing different attacks. The experimental results are described in detail in the following sections.

4.1.1 Experiment with Frame Dropping

As a video contains a large amount of redundancies between frames, it may suffer attacks by frame dropping. This experiment is aimed at examining the robustness of the scheme under the frame dropping attack. Different percentages of frames are dropped and the obtained results are shown in Figure 4.1.

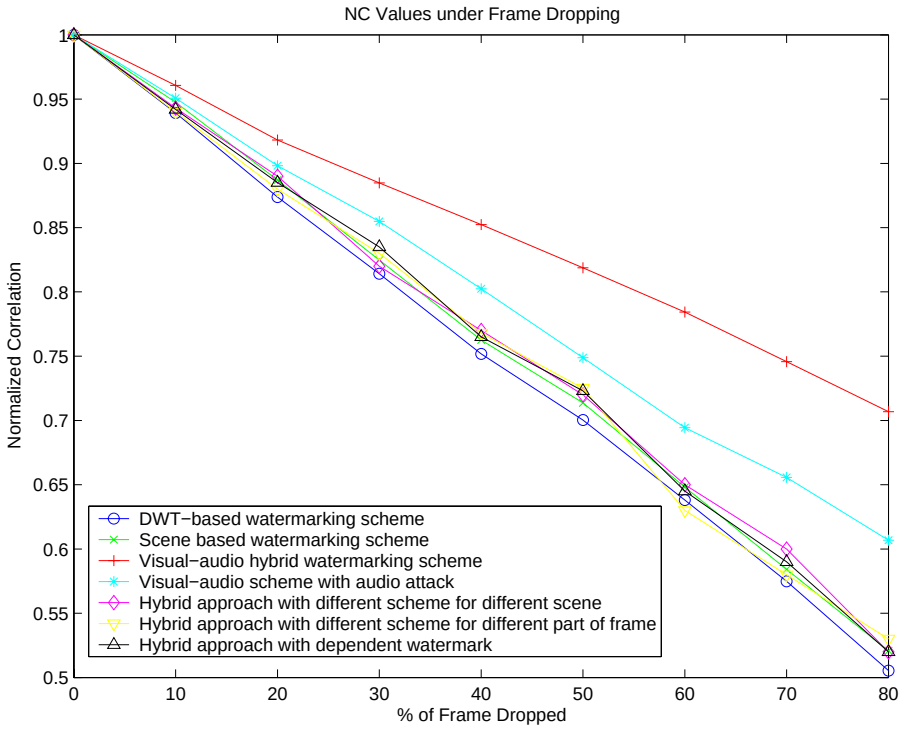

Figure 4.1: NC values under frame dropping

From the experiment, we find that the scheme achieves better performance than the DWT-based scheme without scene-based watermarks [119, 120]. It is because in each scene, all frames are embedded with the same watermark. It prevents the attackers from removing the watermark by frame dropping. If they try to remove one part of the watermark, they need to remove

the whole trunk of frames (i.e., the whole scene), leading to a significant damage to the video. In addition, when the frames are dropped, the error is only introduced to a corresponding small part of the watermark. For the DWT-based scheme (i.e., non-scene-based), however, the error is introduced to the whole watermark, making the performance worse.

The performance of the scheme is significantly improved by combining with an audio watermark, the visual-audio watermarking scheme, especially when the dropping rate of video frame is high. Due to the increased dropping rate, the error of the extracted watermark is increased, which significantly damages the watermark. The error correcting codes from the audio watermark provide information to correct the error and recover the corruption of the video watermark. Consequently, the NC values of the watermark are higher than these without the error correcting codes. Moreover, the error correcting codes are embedded in the audio channel. As frame dropping would not affect the audio channel much, our scheme benefits by allowing uninterrupted error correcting codes to refine the watermark and improve the NC values.

When the audio channel is also attacked, the NC values of the extracted watermark are decreased. The error correcting codes in the audio channel are altered by the attack. Although the capability to recover the error in the video watermark is dropped, the result is still better than the scheme without an audio watermark, as the attacked audio watermark still contains some information to recover the watermark in the video channel.

4.1.2 Experiment with Frame Averaging and Statistical Analysis

Frame averaging and statistical analysis is another common attack to the video watermark. When attackers collect a number of watermarked frames, they can estimate the watermark by statistical averaging and remove it from the watermarked video [121, 122]. The scenario is shown in Figure 4.2.

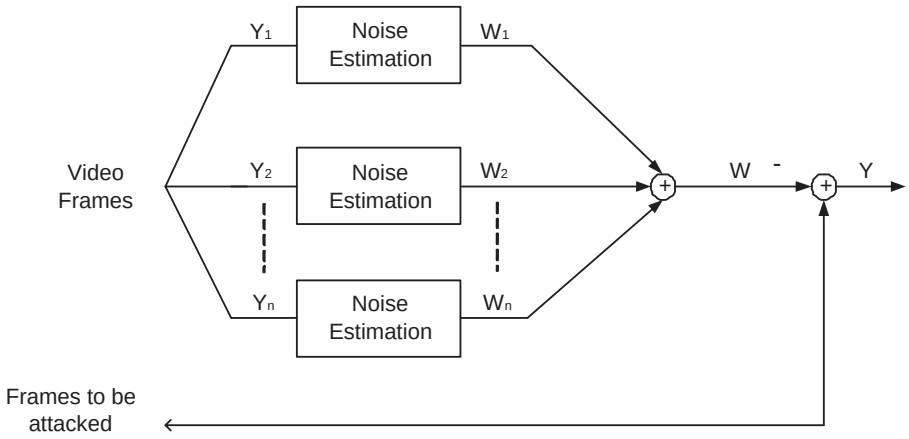

Figure 4.2: Scenario of statistical averaging attack

Experiments have been conducted to evaluate the proposed scheme under this attack, and the results are shown in Figure 4.3. It is found that the proposed scheme can resist to statistical averaging quite well. This is because our scheme crops a watermark into pieces and embeds them into different frames, making the watermarks resistant to attacks by frame averaging for the watermark extraction. The identical watermark used within a scene can prevent attackers from taking the advantage of motionless regions in successive frames and removing the watermark by comparing and averaging the frames statisti-

cally [123]. On the other hand, independent watermarks used for successive, yet different scenes can prevent the attackers from colluding with frames from completely different scenes to extract the watermark.

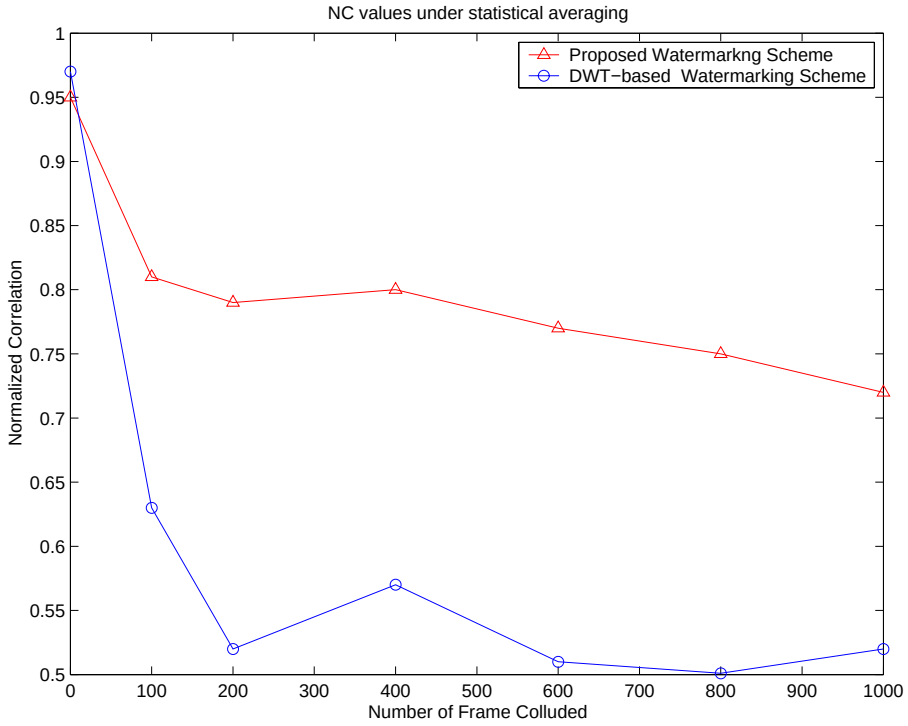

Figure 4.3: NC values under statistical averaging

4.1.3 Experiment with Lossy Compression

This experiment is aimed at testing the robustness of the scheme under attack by lossy compression. Figure 4.4 shows the NC values of the extracted watermarks with different quality factors of MPEG.

From the experiment, we note that the proposed scheme improves the robustness for watermark protection. The perfor-

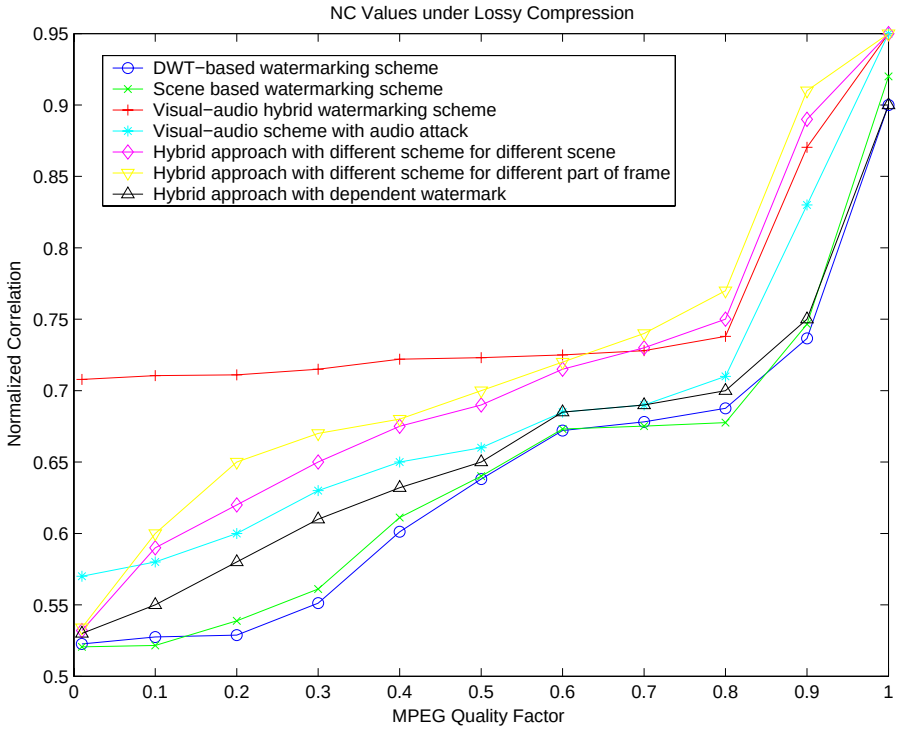

Figure 4.4: NC values under lossy compression

mance of the scheme is significantly improved by combining with audio watermark again, especially when the quality factor of MPEG is low. This is because when the quality factor of MPEG is low, the error of the extracted watermark is increased and the watermark is damaged significantly. As the error correcting codes are provided from the audio watermark, they are not affected by the lossy compression attack applied to the video channel. Consequently, the error correcting codes can overcome the corruption of the video watermark, achiving higher NC values.

The proposed scheme without audio watermark has similar performance with the other DWT-based scheme because both

of them satisfy the following condition. Higher frequency DWT coefficients of the watermark are embedded in the higher frequency part of the video frame and high frequency sub-band DWT coefficients (HH) of video frame are not watermarked. This approach makes the watermark survive MPEG lossy compression, as lossy compression removes the details (i.e. the high frequency components) of the image [124].

The performance of the scheme is also improved by the hybrid approach with different watermarking schemes. The NC values of the hybrid approach under lossy compression are higher than the scheme which only applying single watermarking scheme in a video. From the survey, we find that the DCT-based watermarking scheme is the most resistant one against lossy compression. MPEG is based on DCT and lossy compression mostly only modifies the highest frequency coefficients; therefore, modifying the mid-band coefficients in this domain during watermarking a video would reduce the effect of compression on the watermark. When compression is applied to the watermarked video, the watermark embedded in the video with DCT-based watermarking scheme survives. Therefore, at least one forth of the watermark can be retrieved from the video. This increases the robustness of the scheme. However, the robustness of the scheme is not improved by the hybrid approach with dependent watermark.

4.1.4 Test of Robustness with StirMark 4.0

StirMark 4.0 [100, 101] is a benchmark for testing robustness of a watermarking scheme. In this experiment, we employ Stir-

Mark 4.0 to test the robustness of the proposed schemes when image processing is applied, including cropping, Peak Signal to Noise Ratio (PSNR), adding noise, median filter, row/column removal, rescaling, rotation and affine, and compare them with the current techniques that exist in the literature. Figures 4.5, 4.8, 4.7 and Table 4.1 show the result of the watermarked video under different attacks from StirMark.

Figure 4.5(d) shows the result of the NC values of the watermark under different cropping ratios. The performance of the watermarking schemes without audio watermark decreases greatly with the increasing cropping ratio. When the cropping ratio is greater than 60 present, the NC values of the retrieved watermarks are less than 0.6. The visual-audio watermarking scheme, on the other hand, gives better performance. The NC values are greater than 0.7 even when the cropping ratio is 90 present. This shows that the audio watermark significantly improves the robustness of the watermarking scheme. However, the result shows that the performance of scheme is not improved by the hybrid approach with different watermarking schemes. The NC values of the hybrid approaches are similar to those of the DWT-based watermarking scheme.

Figure 4.9 shows the result of the watermarking scheme under different PSNR. The NC values of the watermarks are decreased with the PSNR. The proposed scheme reveals improvement in this experiment as well. The NC values of the DWT-based watermarking scheme slightly higher than the hybrid approach. It is because the watermark in the wavelet domain is not affected by PSNR too much, but other watermarking schemes are af-

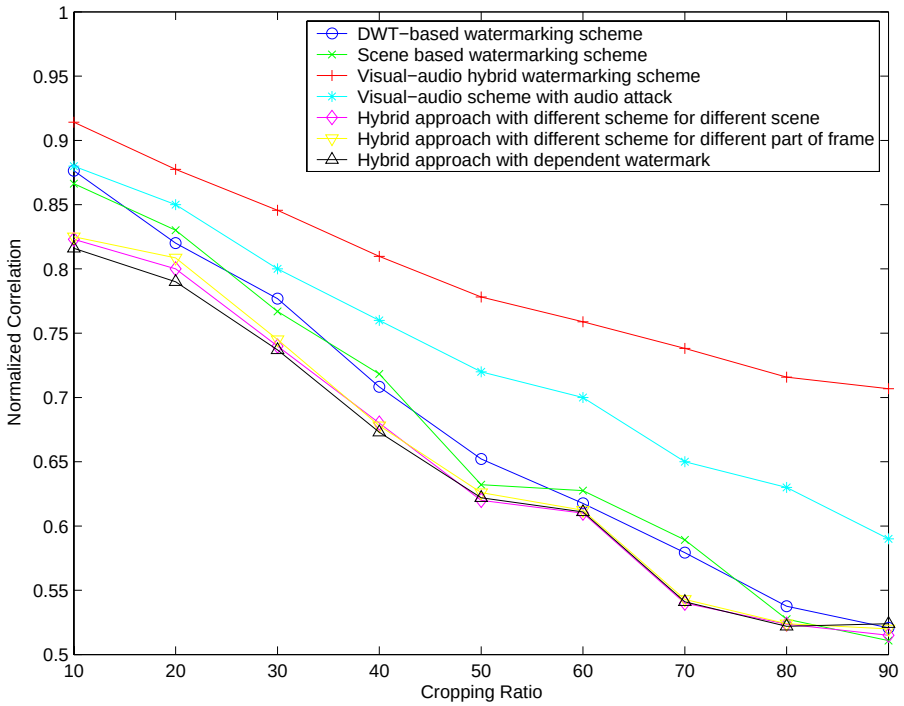

Figure 4.5: NC values under cropping

fected. Thus, the overall performance the hybrid approach is decreased. This is one of the disadvantages of the hybrid approaches.

When the watermarked video is rescaled, the proposed scheme also portrays improvement. Figure 4.7 depicts the NC values when the watermarked video is rescaled with various factors. The performance of the scheme is significantly improved again by visual-audio watermarking scheme, especially when the rescaling factor is large. Furthermore, the improvement becomes more evident with the increase of the rescaling factor. This is because when the rescaling factor increases, the error of the extracted watermark is increased, which significantly damages the watermark. The error correcting codes from the audio water-

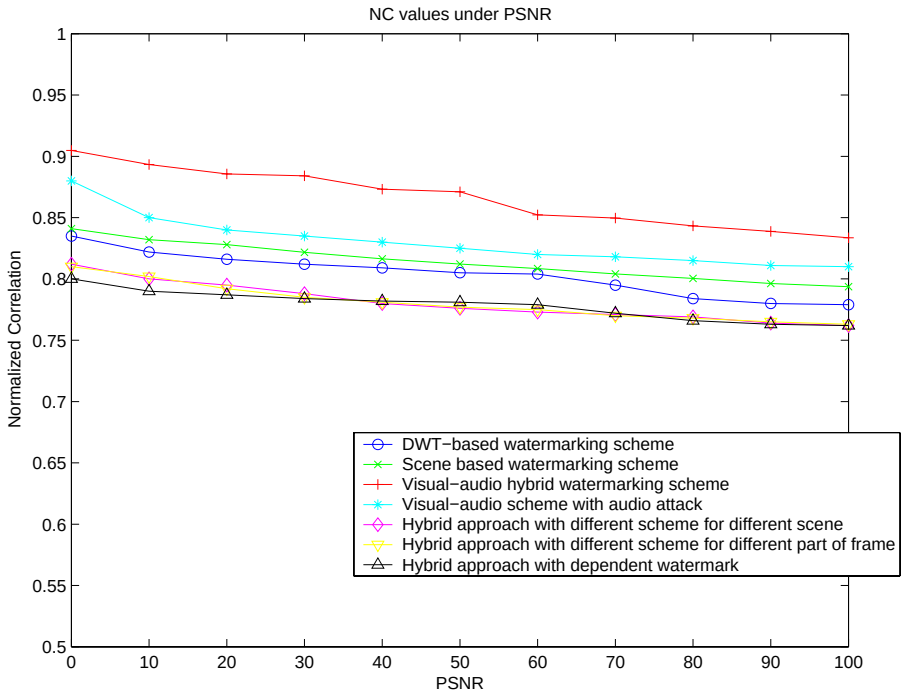

Figure 4.6: NC values under PSNR

mark, on the other hand, provide information to correct the error and overcome part of the corruption in the video watermark and produces higher NC values of the recovered watermark.

With the hybrid approaches, the robustness of the scheme is increased. The NC values of the extracted watermark are higher than the DWT-based and scene-based watermarking schemes. In wavelet domain, the coefficients vary when the size of frame or image is different. The coefficients in the Randon transformed domain, however, do not vary too much when rescaling. As shown in Figure 4.7, the hybrid approaches with different schemes perform better than the scene-based watermarking scheme. The hybrid approach with dependent watermark, nevertheless, performs worse than other hybrid approaches. It is

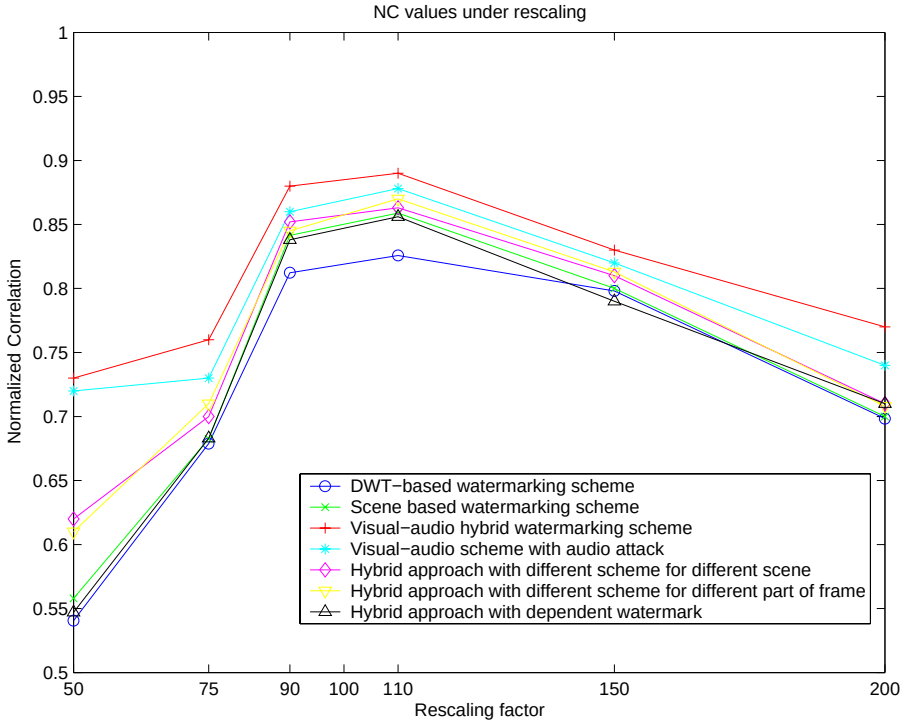

Figure 4.7: NC values under different rescaling factor

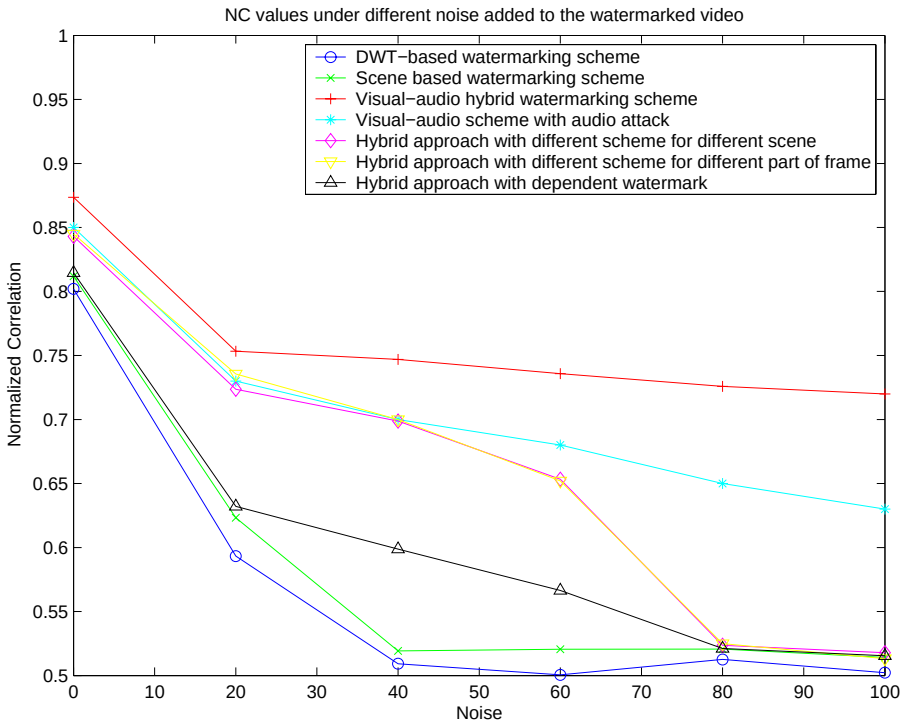

Figure 4.8: NC values under different noise added to the watermarked video

caused by the dependent properties of the watermarking scheme. When one watermark is corrupted, it may affect the extraction of other watermarks in the same frame. Therefore, the NC values of the extracted watermark are lower than those of other hybrid approaches.

When additional noise is added to the watermarked video, the proposed scheme also shows improvement. Figure 4.8 depicts the NC values when different noises are added to the watermarked video. The performance of the scheme is significantly improved by combining with an audio watermark, especially when more noises are added.

There are several tests from the StirMark 4.0. The result is summarized in Table 4.1. The proposed video watermarking

scheme also shows improvement when the videos are under other attacks, including: row removal, rotation, PSNR, affine. . . etc.

Attack Class	DWT-based watermarking scheme	Scene-based watermarking scheme	Visual-audio hybrid watermarking scheme	Visual-audio hybrid watermarking scheme with audio attack
Lossy Compression	0.61	0.62	0.82	0.69
PSNR	0.72	0.76	0.86	0.80
Add Noise	0.63	0.60	0.76	0.67
Median Filter	0.54	0.54	0.74	0.60
Row / Column Removal	0.69	0.71	0.85	0.75
Cropping	0.68	0.66	0.78	0.70
Rescale	0.63	0.62	0.75	0.69
Rotation	0.60	0.61	0.73	0.67
Affine	0.55	0.55	0.78	0.70
Overall	0.62	0.63	0.78	0.69

Attack Class	Hybrid approach with different scheme for different scene	Hybrid approach with different scheme for different part of frame	Hybrid approach with dependent watermark
Lossy Compression	0.71	0.72	0.68
PSNR	0.82	0.81	0.81
Add Noise	0.70	0.69	0.64
Median Filter	0.55	0.52	0.52
Row / Column Removal	0.77	0.78	0.74
Cropping	0.72	0.69	0.67
Rescale	0.71	0.68	0.63
Rotation	0.69	0.66	0.64
Affine	0.73	0.71	0.63
Overall	0.71	0.70	0.66

Table 4.1: Robustness comparison between different watermarking schemes

4.1.5 Overall Comparison

From the above results, the effectiveness of the scene-based hybrid schemes are demonstrated. The scene-based watermarking scheme achieves higher NC values when attacks based on video properties are launched. This indicates that the water-

marking scheme work well by applying scene change detection with scrambled watermarks. The performance of the scheme is further improved by combining with an audio watermark, especially when the video watermark is corrupted, such as the attack by lossy compression. When audio channel is also attacked, the error correction information is altered. The overall performance, however, still shows improvement. The robustness of the scheme is also raised by engaging other hybrid approaches.

4.2 Test on Fidelity

In this section, we focus on evaluating the performance of the GA-based Watermarking Scheme. In the experiment, we prove the fidelity enhancing effectiveness of the proposed optimization process.

The GA-based watermarking scheme is implemented with the GAlib [125]. To evaluate the fidelity of the watermarking scheme, the peak signal-to-noise ratio (PSNR) and maximum absolute difference (MAD) is used.

PSNR is given by

$$PSNR = 10 \log_{10} \frac{255^2}{\sigma_q^2} [dB] \tag{4.2}$$

where σ_q^2 is the mean square of the difference between the original video frame and the watermarked one [126].

MAD is given by

$$MAD = \sum_{x=0}^{7} \sum_{y=0}^{7} |I'(x,y) - I(x,y)| \tag{4.3}$$

Where I' and I are the intensity values of the same pixel position within a video frame after and before embedding, respectively.

The performance of the GA-based video watermarking scheme is evaluated through several experiments with different number of generation in the GA-optimization process. Then, the quality of the video is evaluated with PSNR and MAD. The quality of the video watermarked by the scene-based watermarking scheme and the hybrid watermarking scheme are also evaluated, and compared with the GA-based scheme. In the experiment, two

video clips are used. One of the video clips has 1526 frames of size 352×288 and it consists of 10 scene changes. Another video clip has 4236 frames of size 352×288 and it consists of 22 scene changes. The experiments are done on a desktop computer with Pentium 4 CPU 2.00GHz and 512MB RAM.

4.2.1 Parameter(s) Setting

Table 4.2 shows the parameters setting for the GA-based video watermarking experiments.

Table 4.2: Parameters Setting for GA-based experiment

Parameter	Value
Population size	100
Mutation probability	0.05
Crossover probability	0.9
Score frequency	1
Flush frequency	25
Number of generations	In the experiment,we have used 0, 1, 5, 10, 20, 40, 100, 200, 400 and 600

Table 4.3 shows the computation time of the GA-based scheme.

4.2.2 Evaluate with PSNR

PSNR measures the signal to noise ratio of the watermarked video, thus, we can evaluate the its fidelity. From the graph 4.9, it's clear that the GA-based algorithm successfully reduces the video frame distortion due to watermark embedding. As

Table 4.3: The computation time of the GA-based scheme

Number of Generation	Computation time of Video 1 (s)	Computation time of Video 2 (s)
0	50	141
5	62	179
10	80	234
20	103	289
40	124	354
100	156	423
200	185	534
400	226	620
600	259	702

the number of generations increases, the improvement of video quality gradually approaches to a saturation value.

Table 4.4 depicts a comparison of PSNR with different watermarking scheme. The PSNR of the video is reduced to 3/4 after GA is applied in optimizing the fidelity of the scheme. It shows that the GA-based optimization effectively improves the performance of the scheme.

4.2.3 Evaluate with MAD

MAD measures the different between the original video and the watermarked video, thus, we can evaluate the quality of the watermarked video. From the graph 4.10, it shows how the iterative generation numbers affect the optimizing performance. The MAD of the watermarked video is decreased with the GA

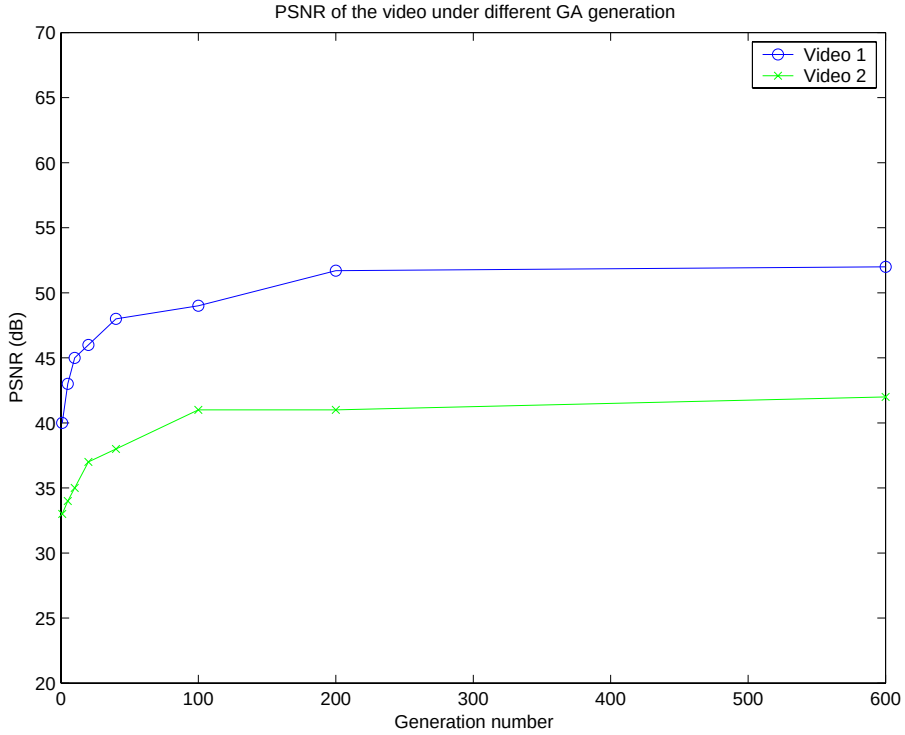

Figure 4.9: PSNR of the video under different GA generations

Table 4.4: PSNR comparison between different watermarking schemes

Watermarking scheme	PSNR of Video 1	PSNR of Video 2
Scene-based Watermarking scheme	40	33
Visual-audio hybrid watermarking	41	33
Hybrid approach with different scheme for different scene	43	34
Hybrid approach with different scheme for different part of frame	42	36
Hybrid approach with dependent watermark	35	27
GA-based watermarking scheme watermark scheme	52	42

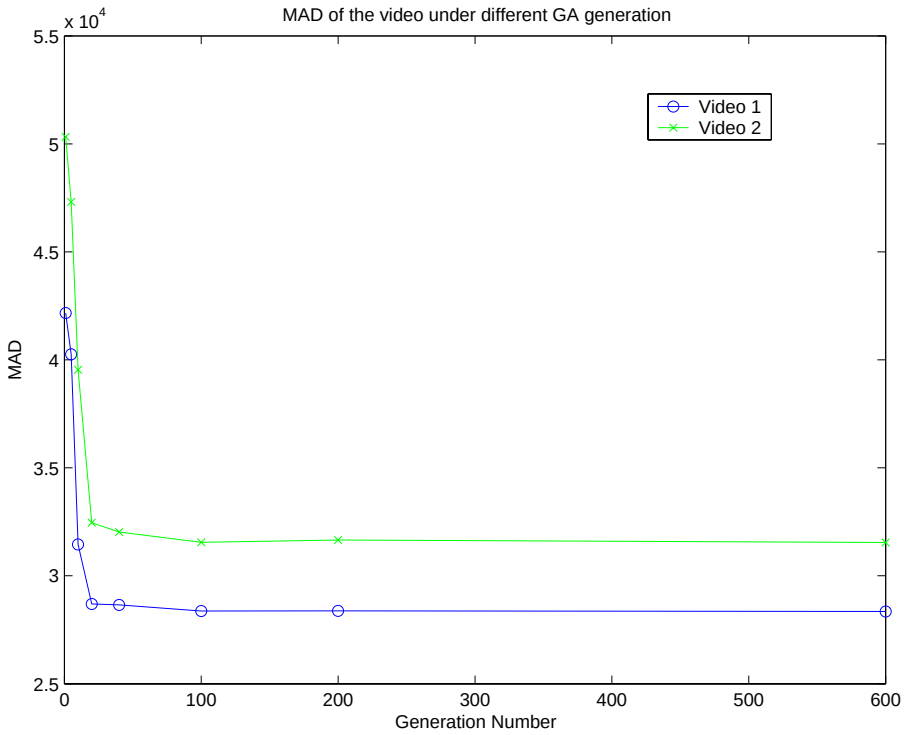

Figure 4.10: MAD of the video under different GA generations

generation number. The optimization performance saturate after about 200 generations. Therefore, the performance of the proposed GA-based watermarking scheme is converged to optimal.

A comparison of MAD with different watermarking scheme is shown in Table 4.5. It shows the enhancing effectiveness of the GA-based watermarking scheme. The MAD of the video is reduced to 3/5 after GA is applied to optimize the fidelity of the scheme.

Table 4.5: MAD comparison between different watermarking schemes

Watermarking scheme	MAD of Video 1	MAD of Video 2
Scene-based Watermarking scheme	42168	50325
Visual-audio hybrid watermarking	42189	50489
Hybrid approach with different scheme for different scene	43984	52695
Hybrid approach with different scheme for different part of frame	43798	52786
Hybrid approach with dependent watermark	48652	55785
GA-based watermarking scheme watermark scheme	28346	31546

4.3 Other Features of the Scheme

Our proposed scheme is an invisible watermarking scheme. In the scene-based watermarking scheme, as low frequency sub-band DWT coefficients (LL) are not watermarked and image energy is concentrated on the lower frequency wavelet coefficients, the watermark is perceptually invisible. If these coefficients are altered, however, the perceptual quality will be affected [123]. Additionally, retrieval of the embedded watermarks does not need the original video, i.e. a blind watermarking scheme. This is an important performance feature of the scheme since it takes a long time to transmit, store, and process the original video.

The experiments show that the proposed scheme is robust to most of the existing attacks, however, there are still some weaknesses in our scheme. The computation time of the GA-based scheme is rather long if the number of GA generation applied is large or the number of the scene change of the video increases. Also, when the encoding method is applied again with another watermark, the proposed scheme is not robust against

it efficiently. The original watermark can only be extracted if the second watermark is removed first.

4.4 Conclusion

From the experiment, we prove that our proposed scheme enhances two of the three prescribed watermarking requirements, robustness and fidelity. The robustness-enhancement provided by hybrid scene-based watermarking scheme and the fidelity-enhancement provided by the GA-based watermarking scheme are important steps toward a prefect watermarking scheme.

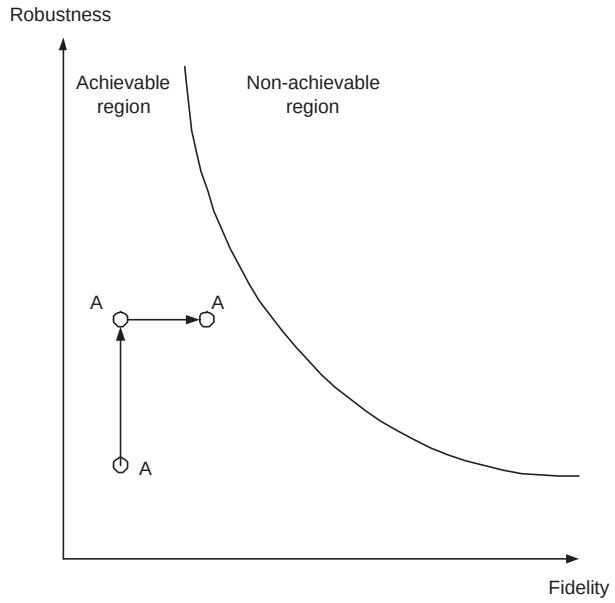

Figure 4.11: A conceptual illustration on the performance of the proposed scheme

Figure 4.11 shows a conceptual illustration. When there are two orthogonal axes, robustness and fidelity. The robustness is indicated by the strength of the watermark, and the fidelity can

be represented by quality index. Moreover, the curve represent the best robustness under the fidelity performance constrain. To optimize the performance of a watermarking scheme, we try to move the point towards the curve.

The point **A** represents the performance of a scene-based watermarking scheme. By applying the hybrid approaches, the robustness of the scheme is improved and moves the point to **A'**. When the watermarking scheme is further enhanced by GA, i.e. increase the fidelity the scheme, the point moves along the fidelity axes. Thus, it moves towards **A"**. Therefore, our proposed scheme is approaching to the "optimal" embedding configuration.

□ **End of chapter.**

Chapter 5

Conclusion

This thesis investigates the knowledge of digital video watermarking techniques for secure multimedia creation and delivery. After noticing the importance of the multimedia security and video watermarking in nowadays Internet world and reviewing the state-of-the-arts technologies of the audio watermarking, image watermarking and video watermarking, an innovative hybrid digital video watermarking scheme with scene change analysis, error correcting code and GA optimization is proposed. The process of this comprehensive video watermarking scheme, including watermark preprocessing, video preprocessing, watermark embedding, and watermark detection, is described in detail. Experiments are conducted to demonstrate that our scheme is robust against attacks by frame dropping, frame averaging, and statistical analysis, and the robustness against the image processing attacks is tested with StirMark benchmark. Moreover, the fidelity of the scheme is evaluated.

Our approach cultivates an innovative idea in: (1) embedding different parts of a watermark according to scene changes, (2)

embedding its error correcting codes as an audio watermark, (3) applying a hybrid approach to the proposed scheme, and (4) employing the GA algorithm to enhance the fidelity. This approach is never explored in the literature, and its advantages are clear and significant. The effectiveness of this scheme is verified through a number of experiments.

To conclude our work, we contribute on the followings:

- We have performed a complete survey on the current watermarking technologies.
- We propose a scene-based watermarking scheme. The scheme is robust against frame averaging, frame dropping, frame swapping, statistical analysis, etc
- We propose a visual-audio hybrid watermarking scheme. The robustness of our scheme can be enhanced by including an audio watermark. We embed error correcting codes of a video watermark as an audio watermark and refine the retrieved watermark during watermark detection.
- We propose a hybrid approach with different watermarking schemes. We employ the hybrid scheme to embed different parts of a watermark into different scenes. There are different ways to embed the watermarks.
- We propose a GA-based watermarking scheme increase the fidelity, i.e. the media quality index, of the watermarking scheme. By employing GA, we can optimize the combination of the watermark and the scenes in the video.

- Experiment has been done on these novel video watermarking schemes to test and show its performance. The robustness of our approach is demonstrated using the criteria of the latest StirMark test.

□ **End of chapter.**

Bibliography

[1] A. Piva, F. Bartolini, and M. Barni, "Managing copyright in open networks," *IEEE Transactions on Internet Computing*, Vol. 6, Issue. 3, pp. 18-26, May-June 2002.

[2] C. Lu, H. Yuan, and M. Liao, "Multipurpose Watermarking for Image Authentication and Protection," *IEEE Transactions on Image Processing*, Vol. 10, Issue. 10, pp. 1579-1592, Oct. 2001.

[3] C. Lu, S. Huang, C. Sze, and H. Y. M. Liao, "Cocktail watermarking for digital image protection," *IEEE Transactions on Multimedia*, Vol. 2, pp. 209-224, Dec. 2000.

[4] J. Lee and S. Jung, "A survey of watermarking techniques applied to multimedia," *Proceedings 2001 IEEE International Symposium on Industrial Electronics (ISIE2001)*, Vol. 1, pp. 272-277, 2001.

[5] M. Barni, F. Bartolini, R. Caldelli, A. De Rosa, and A. Piva, "A Robust Watermarking Approach for Raw Video," *Proceedings 10th International Packet Video Workshop* PV2000, Cagliari, Italy, May 1-2 2000.

[6] F. Petitcolas (Eds)"Information hiding techniques for steganography and digital watermarking Stefan Katzenbeisser," Artech House Books, Dec. 1999, ISBN 1-58053-035-4.

[7] A. Eskicioglu and E. Delp, "An overview of multimedia content protection in consumer electronics devices," *Proceedings Signal Processing Image Communication 16 (2001)*, pp. 681-699, 2001.

[8] N. Fates and F. Petitcolas, "Watermarking schemes evaluation tool," *Proceedings of IEEE Multimedia Software Engineering*, pp. 328-331, Taiwan, Dec. 2000.

[9] A. Herrigel and J. Ruanaidh, "Secure Copyright Protection Techniques for Digital Images," *Processing in Workshop on Information Hiding*, LNCS, Springer Verlag.

[10] M. Swanson, B. Zhu, and A. Tewfik, "Multiresolution Video Watermarking using Perceptual Models and Scene Segmentation," *Proceedings International Conference on Image Processing (ICIP '97)*, 3-Volume Set-Volume 2, Washington, DC, Oct. 26-29, 1997.

[11] P.W. Chan and M. Lyu, "A DWT-based Digital Video Watermarking Scheme with Error Correcting Code," *Proceedings Fifth International Conference on Information and Communications Security (ICICS2003)*, Lecture Notes in Computer Science, Springer, Vol. 2836, pp. 202-213, Huhehaote City, Inner-Mongolia, China, Oct. 10-13, 2003.

[12] P.W. Chan, M.R. Lyu, R. Chin, "Copyright Protection on the Web: A Hybrid Digital Video Watermarking Scheme," *Poster Proceedings 13th International World Wide Web Conference (WWW'2004)*, pp. 354-355, New York, May 17-22, 2004.

[13] P.W. Chan, M.R. Lyu and R.T. Chin "A Novel Scheme for Hybrid Digital Video Watermarking: Approach, Evaluation and Experimentation," submitted to *IEEE Transactions on Circuits and Systems for Video Technology.*

[14] F. Petitcolas, M. Frederic Raynal, J. Dittmann, C. Fontaine, N. Fates, "A public automated web-based evaluation service for watermarking schemes: StirMark Benchmark," In Ping Wah Wong and Edward J. Delp, editors, *proceedings of electronic imaging, security and watermarking of multimedia contents III*, Vol. 4314, San Jose, California, U.S.A., Jan. 20V26, 2001. The Society for imaging science and technology (I.S. and T.) and the international Society for optical engineering (SPIE). ISSN 0277-786X, http://www.petitcolas.net/fabien/watermarking/stirmark/.

[15] K. Su, D. Kundur, and D. Hatzinakos, "Statistical Invisibility for Collusion-resistant Digital Video Watermarking," *IEEE Transactions on Multimedia 2004*

[16] K. Su, D. Kundur, and D. Hatzinakos, "Spatially Localized Image-dependent Watermarking for Statistical Invisibility and Collusion Resistance," *IEEE Transactions on Multimedia 2004*

[17] K. Su, "Digital Video Watermarking Principles for Resistance to Collusion and Interpolation Attacks," Master of Applied Science thesis, University of Toronto, Sept. 2001.

[18] I. Cox, M. Miller, and J. Bloom, "Digital watermarking," Morgan Kaufmann Publishers, Oct. 2001, ISBN 1-55860-714-5.

[19] F. Litterio, http://world.std.com/ franl/crypto/.

[20] P. Johnston, http://pajhome.org.uk/crypt/rsa/intro.html.

[21] Watermarking World, http://www.watermarkingworld.org/.

[22] M. Kutter and F. Hartung, "Introduction to Watermarking Techniques," *Proceedings Information Techniques for Steganography and Digital Watermarking*, S.C. Katzenbeisser et al., Eds. Northwood, MA: Artec House, pp. 97-119, Dec. 1999.

[23] H. Inoue, A. Miyazaki, and T. Katsura "An Image Watermarking Method Based on the Wavelet Transform", Kyushu Multimedia System Research Laboratory.

[24] I. Cox, M. Miller, J. Linnartz, and T. Kalker, "A Review of Watermarking Principles and Practices" *Proceedings Digital Signal Processing for Multimedia Systems*, K.K. Parhi, T. Nishitani, eds., New York, New York, Marcel Dekker, Inc., pp. 461-482, 1999.

[25] F. Petitcolas, "Watermarking Schemes Evaluation", *IEEE Signal Processing Magazine*, Vol. 17, pp. 58-64, Sept. 2000.

[26] Y. Kim, K. Moon, and I. Oh, "A text watermarking algorithm based on word classification and inter-word space statistics," *Proceedings Seventh International Conference on Document Analysis and Recognition 2003*, pp. 775 -779, Aug. 3-6, 2003 .

[27] R. Wolfgang, C. Podilchuk, and E. Delp, "Perceptual Watermarks for Digital Images and Video," *Proceedings of the SPIE/IS and T International Conference on Security and Watermarking of Multimedia Contents*, Vol. 3657, pp. 40-51, San Jose, CA, Jan. 25 - 27, 1999.

[28] R. Wolfgang and E. Delp, "A Watermarking Technique for Digital Imagery: Further Studies," *Proceedings of the International Conference on Imaging Science, Systems, and Technology*, pp. 279-287, Las Vegas, Jun. 30 - Jul. 3, 1997.

[29] D. Sachs, R. Anand, and K. Ramchandran, "Wireless image transmission using multiple-description based concatenated codes," *Proceedings Data Compression Conference DCC 2000*, pp. 569, 2000.

[30] A. Doufexi, A. Nix, D. Bull, "Robust wireless image transmission using jointly-optimized modulation and source coding," *Proceedings IEEE 51st Vehicular Technology Conference 2000, VTC 2000-Spring* Vol. 3, pp. 2039-2043, Tokyo, 2000.

[31] M. Buckley, M. Ramos, S. Hemami, and S. Wicker, "Perceptually-based robust image transmission over wire-

less channels," *Proceedings 2000 International Conference on Image Processing*, Vol. 2, pp. 128-131, 2000.

[32] H. Bassali, J. Chhugani, S. Agarwal, A. Aggarwal, and P. Dubey, "Compression tolerant watermarking for image verification," *Proceedings 2000 International Conference on Image Processing*, Vol. 1, pp. 430-433, 2000.

[33] L. Boney, A. Tewfik, and K. Hamdy, "Digital watermarks for audio signals," *Proceedings Third IEEE International Conference on Multimedia Computing and Systems*, pp. 473-480, Jun. 17-23, 1996.

[34] C. Lu, M. Liao, and L. Chen, "Multipurpose audio watermarking," *Proceedings 15th International Conference on Pattern Recognition 2000*, Vol. 3, pp. 282-285, 2000.

[35] S. Craver, M. Wu, and B. Liu, "What can we reasonably expect from watermarks?," *Proceedings IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics 2001*, pp. 223-226, 2001.

[36] C. Xu, J. Wu and Q. Sun "Digital audio watermarking and its application in multimedia database," *Proceedings the Fifth International Symposium on Signal Processing and Its Applications ISSPA'99*, Vol. 1, pp. 91-94, 1999.

[37] S. Foo, T. Yeo, D. Huang, "An adaptive audio watermarking system," *Proceedings IEEE Region 10 International Conference on Electrical and Electronic Technology 2001*, Vol. 2, pp. 509-513, TENCON, 2001.

[38] K. Kaabneh and A. Youssef, "Muteness-based audio watermarking technique," *Proceedings International Conference on Distributed Computing Systems Workshop, 2001*, pp. 379-383, Apr. 2001.

[39] C. Xu, J. Wu, and Q. Sun, "A robust digital audio watermarking technique" *Proceedings the Fifth International Symposium on Signal Processing and Its Applications ISSPA '99*, Vol. 1, pp. 95-98, 1999.

[40] X. Wang, Y. Liu, and S. Blostein, "Video image transmission via mobile satellite channels," *Proceedings IEEE International Conference on Communications, ICC 95*, Vol. 1, pp. 352-356, Seattle, 18-22 Jun. 1995.

[41] N. Checcacci, M. Barni, F. Bartolini, and S. Basagni, "Robust video watermarking for wireless multimedia communications ," *Proceedings IEEEE Wireless Communications and Networking Confernce 2000, WCNC. 2000*, Vol. 3, pp. 1530-1535, 2000.

[42] C. Lin and S. Chang, "Issues and Solutions for Authenticating MPEG Video," *Proceedings IEEE International Conference on Acoustics, Speach and Signal Processing*, pp. 54-65, Mar. 15-19, 1999.

[43] C. Lu and M. Liao, "Video object-based watermarking: a rotation and flipping resilient scheme," *Proceedings 2001 International Conference on Image Processing*, Vol. 2, pp. 483-486, 2001.

[44] E. Lin, C. Podilchuk, T. Kalker, and E. Delp, "Streaming Video and Rate Scalable Compression: What Are the Challenges for Watermarking?," *Proceedings the SPIE International Conference on Security and Watermarking of Multimedia Contents III*, Vol. 4314, San Jose, CA, Jan. 22-25, 2001.

[45] F. Hartung and B. Girod, "Watermarking of Uncompressed and Compressed Video", *IEEE Transaction Signal Processing*, Vol. 66, no. 3 (Special issue on Watermarking), pp. 283-301, May 1998.

[46] R. Wolfgang and E. Delp, "Overview of Image Security Techniques with Applications in Multimedia Systems," *Proceedings the SPIE Conference on Multimedia Networks: Security, Displays, Terminals, and Gatweys*, Vol. 3228, pp. 297-308, Dallas, Texa, Nov. 2-5, 1997.

[47] R. Wolfgang, C. Podilchuk, and E. Delp, "Perceptual Watermarks for Digital Images and Video," *Proceedings the SPIE/IST International Conference on Security and Watermarking of Multimedia Contents*, Vol. 3657, pp. 40-51, San Jose, CA, Jan. 25-27, 1999.

[48] T. Furon, and P. Duhamel, "Robustness of asymmetric watermarking technique", *Proceedings International Conference onImage Processing, 2000*, Vol.3, pp. 21-24.

[49] R. Lancini, F. Mapelli, and S. Tubaro, "A robust video watermarking technique in the spatial domain," *Processing*

and Multimedia Communications 4th EURASIP-IEEE Region 8 International Symposium on Video/Image VIProm-Com, pp. 251-256, 2002.

[50] M. Ramkumar and A. Akansu, "Robust Protocols for Proving Ownership of Image," *IEEE Transactions on Multimedia.*

[51] P. Lee and M. Chen, "Robust error concealment algorithm for video decoder," *IEEE Transactions on Consumer Electronics*, Vol. 45, Issue. 3, pp. 851 -859, Aug. 1999.

[52] D. He, Q. Sun and Q. Tian, "A semi-fragile object based video authentication system," *Proceedings of the 2003 International Symposium on Circuits and Systems ISCAS '03*, Vol.3, pp. 814-817, May 25-28 2003.

[53] J. Fridrich, M. Goljan and A. Baldoza, "New fragile authentication watermark for images," *Proceedings. 2000 International Conference on Image Processing*, Vol. 1, pp. 446-449, Sept. 10-13, 2000.

[54] M. Swanson, B. Zhu, and A. Tewfik, "Transparent robust image watermarking," *Proceedings International Conference on Image Processing, 1996*, Vol. 3, pp. 211-214, Sept. 16-19, 1996.

[55] Y. Steinberg, "Watermarking identification for private and public users: the broadcast channel approach," *Proceedings the 2002 IEEE Information Theory Workshop*, pp. 5-7, Oct. 20-25, 2002.

[56] P. Wong and O. An, "A novel semi-private watermarking technique," *Proceedings IEEE International Symposium on Circuits and Systems 2002 (ISCAS 2002)*, Vol. 2, pp. 448-451, May 26-29, 2002.

[57] T. P.-C. Chen and T. Chen, "A Framework for Optimal Public Watermark Detection", Carnegie Mellon University Technical Report: AMP01-03.

[58] R. Wolfgang and E. Delp. "A watermark for digital images," *Proceedings International Conference on Images Processing*, pp. 219-222, Lausanne, Switzerland, Sept. 1996.

[59] H. Kii, J. Onishi, and S. Ozawa, "The Digital Watermarking Method by Using both Patchwork and DCT," *IEEE International Conference on Multimedia Computing and Systems*, Vol. 1, Florence, Italy, Jun. 07-11, 1999.

[60] N. Merhav, "On random coding error exponents of watermarking systems," *IEEE Transactions on Information Theory*, Vol. 46 Issue. 2, pp. 420-430, Mar. 2000.

[61] M. Barni, F. Bartolini, A. Rosa and A. Piva, "Capacity of full frame DCT image watermarks," *IEEE Transactions on Image Processing*, Vol. 9 Issue. 8, pp. 1450-1455, Aug. 2000.

[62] A. Bors and I. Pitas, "Image watermarking using DCT domain constraints," *Proceedings International Conference on Image Processing 1996*, Vol. 3, pp. 231-234, Sept. 16-19, 1996.

[63] I. Hong, I. Kim and S. Han, "A blind watermarking technique using wavelet transform," *Proceedings IEEE International Symposium on Industrial Electronics ISIE 2001*, Vol. 3, pp. 1946-1950, 2001.

[64] X. Niu and S. Sun, "A New Wavelet-Based Digital Watermarking for Video," *Proceedings 9th IEEE Digital Signal Processing Workshop*, Texas, USA, Oct. 2000.

[65] K. Leung and B. Zeng, "Wavelet-based digital watermarking with halftoning technique," *Proceedings The 2001 IEEE International Symposium on Circuits and Systems, 2001. ISCAS 2001*, Vol. 5, pp. 235-238, 2001.

[66] Y. Wang, J. Doherty, and R. Dyck, "A wavelet-based watermarking algorithm for ownership verification of digital image," *IEEE Transactions on Image Processing*, Vol. 11, No. 2, Feb. 2002.

[67] P. Premaratne and C. Ko, "A novel watermark embedding and detection scheme for images in DFT domain," *Seventh International Conference on Image Processing and Its Applications 1999*, Vol. 2, pp. 780-783, Jul. 13-15, 1999.

[68] Q. Cheng and T. Huang, "Optimum detection and decoding of multiplicative watermarks in DFT domain," *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '02) 2002*, Vol. 4, pp. 3477-3480, May 13-17, 2002.

[69] J. Linnartz and J. Talstra, "MPEG PTY-marks: Cheap detection of embedded copyright data in DVD video," *Proceedings 5th European Symposium on Research in Computer Security*, pp. 221-240, 1998.

[70] L. Qiao and K. Nahrstedt, "Watermarking method for mpeg encoded video: Towards resolving rightful ownership," *IEEE Multimedia Computing and Systems*, Jun. 1998.

[71] T. Chung, M. Hong, Y oung-Nam Oh, Dong-Ho Shin, and Sang-Hui Park, "Digital watermarking for copyright protection of MPEG2 compressed video," *IEEE Transactions on Consumer Electronics*, Vol. 44, Issue. 3, pp. 895-901, August 1998.

[72] S. Arena and M. Caramma, "Digital watermarking applied to MPEG2 coded video sequence exploiting space and frequency masking," *Proceedings International Conference on Image Processing (ICIP-2000)*, Vol. 3, pp. 438-441, Vancouver, Canada, 2000

[73] A. Piva, R. Caldelli and A. De Rosa, "A DWT-based object watermarking system for MPEG-4 video streams," *Proceedings 2000 International Conference on Image Processing*, Vol. 3, pp. 5-8, 2000.

[74] M. Suhail and M. Obaidat, "On the digital watermarking in JPEG 2000," *Proceedings The 8th IEEE International Conference on Electronics, Circuits and Systems, 2001 (ICECS 2001)*, Vol. 2, pp. 871-874, Sept. 2-5, 2001.

[75] J. Dittmann, A. Mukherje, and M. Steinebach, "Media-independent watermarking classification and the need for combining digital video and audio watermarking for media authentication," *Proceedings International Conference on Information Technology: Coding and Computing 2000* pp. 62-67, 2000.

[76] J. Dittmann and M. Steinebach, "Joint watermarking of audio-visual data," *Processing 2001 IEEE Fourth Workshop on Multimedia Signal*, pp. 601-606, 2001.

[77] J. Fridrich, "A hybrid watermark for tamper detection in digital images," *Proceedings the Fifth International Symposium onSignal Processing and Its Applications ISSPA '99, 1999*, Vol. 1, pp. 301-304, 1999.

[78] X. Kang, J. Huang, Y. Shi, and Y. Lin, "A DWT-DFT composite watermarking scheme robust to both affine transform and JPEG compression," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 13, Issue: 8, pp. 776-786, Aug. 2003.

[79] F. Deguillaume, G. Csurka, J. Ruanaidh, and T. Pun, "Robust 3D DFT video watermarking," *Proceedings Electronic Imaging' 99: Security and Watermarking of Multimedia Contents*, Vol. 3657, San Jose, CA, Jan. 1999.

[80] M. Ejima and A. Miyazaki, "A wavelet-based watermarking for digital images and video," *Proceedings International Conference on Image Processing (ICIP-2000)*, Vol. 3, pp. 678-681, Vancouver, Canada, 2001.

[81] F. Hartung and B. Girod, "Watermarking of uncompressed and compressed video," *Proceedings Signal Processing*, Vol. 66, pp. 283-301, 1998.

[82] N. Johnson and S. Katezenbeisser, "A Survey of Steganographic Techniques in Information Techniques for Steganography and Digital Watermarking," S.C. Katzenbeisser et al., Eds. Northwood, MA: Artec House, pp. 43-75, Dec. 1999.

[83] G. Langelaar, I. Setyawan, R. Lagendijk, "Watermarking Digital Image and Video Data," *IEEE Signal Processing Magazine*, Vol 17, pp. 20-43, Sept. 2000.

[84] X. Niu and S. Sun, "A New Wavelet-Based Digital Watermarking for Video", *Proceedings 9th IEEE Digital Signal Processing Workshop*, Texas, USA, Oct. 2000.

[85] C. Serdean, M. Ambroze, M. Tomlinson, and G. Wade, "Combating geometrical attacks in a dwt based blind video watermarking system", *Proceedings Video/Image Processing and Multimedia Communications 4th EURASIP-IEEE Region 8 International Symposium on VIPromCom*, pp. 263-266, 2002.

[86] D. Swanson, B. Zhu, and A. Tewfik, "Multiresolution Video Watermarking using Perceptual Models and Scene Segmentation," *Proceedings International Conference on Image Processing (ICIP '97)*, Vol. 2, Washington, DC, Oct. 26-29, 1997.

[87] M. Barni, F. Bartolini, R. Caldelli, A. De Rosa, and A. Piva, "A Robust Watermarking Approach for Raw Video", *Proceedings 10th International Packet Video Workshop PV2000*, Cagliari, Italy, May 1-2, 2000.

[88] M. Holliman, N. Memon, B. Yeo, and M. Yeung, "Adaptive public watermarking of DCT-based compressed images," *Proceedings SPIE*, Vol. 3312, pp. 284-295, 1997.

[89] B. Vassaux, P. Nguyen, S. Baudry, P. Bas, and J. Chassery, "Scrambling technique for video object watermarking resisting to mpeg-4," *Proceedings Video/Image Processing and Multimedia Communications 4th EURASIP-IEEE Region 8 International Symposium on VIPromCom*, pp. 239-244, 2002.

[90] M. Swanson, B. Zhu, B. Chau, and A. Tewfik, "Object-Based Transparent Video Watermarking," *Proceedings IEEE Signal Processing Society 1997 Workshop on Multimedia Signal Processing*, Princeton, New Jersey, USA, Jun. 23-25, 1997.

[91] C. Lu and M. Liao, "Video object-based watermarking: a rotation and flipping resilient scheme," *Proceedings 2001 International Conference on Image Processing*, Vol. 2, pp. 483-486, 2001.

[92] B. Mobasseri, "Direct sequence watermarking of digital video using m-frames," *Proceedings International Conference on Image Processing (ICIP-98)*, Vol. 3, pp. 399-403, Chicago, Illinois, Oct. 4-7, 1998.

[93] N. Memon, "Analysis of LSB based image steganography techniques Chandramouli," *Proceedings 2001 International Conference on Image*, Vol. 3. pp. 1019-1022, Oct. 7-10, 2001.

[94] B. Mobasseri, "Direct sequence watermarking of digital video using m-frames," *Proceedings 1998 International Conference on Image Processing ICIP 98*, Vol. 2, pp. 399-403, Oct. 4-7, 1998.

[95] B. Mobasseri, "Direct sequence watermarking of digital video using m-frames," *Proceedings International Conference on Image Processing (ICIP-98)*, Vol. 3, pp. 399-403, Chicago, Illinois, Oct. 4-7, 1998.

[96] S. Pereira and T. Pun, "Robust template matching for affine resistant image watermarks, " *IEEE Transactions on Image Processing*, Vol.9 Issue. 6, pp. 1123-1129, Jun. 2000.

[97] I. Hong, I. Kim, and S. Han, "A blind watermarking technique using wavelet transform" *Proceedings IEEE International Symposium on Industrial Electronics ISIE 2001*, Vol. 3, pp. 1946-1950, 2001.

[98] F. Duan, I. King, L. Xu, and L. Chan, "Intra-block algorithm for digital watermarking", *Proceedings IEEE 14th International Conference on Pattern Recognition (ICPR'98)*, Vol. 2, pp. 1589-1591, Aug. 17-20, 1998.

[99] M. George, J. Chouinard and N. Georganas, "Digital watermarking of images and video using direct sequence spread

spectrum techniques," *Proceedings 1999 IEEE Canadian Conference on Electrical and Computer Engineering*, Vol. 1, pp. 116 121, May 9-12, 1999.

[100] F. Petitcolas and R. Anderson, "Evaluation of Copyright Marking Systems," *Proceedings IEEE multimedia Systems (ICMCS'99)*, Florence, Italy, Jun. 7-11, 1999.

[101] M. Kutter and F. Petitcolas, "A Fair Benchmark for Image Watermarking Systems," *Proceedings of Electronic Imaging '99, Security and Watermarking of Multimedia Contents*, Vol. 3657, pp. 226-239.

[102] N. Checcacci, M. Barni, F. Bartolini, and S. Basagni, "Robust video watermarking for wireless multimedia communications," *Proceedings 2000 IEEE Wireless Communications and Networking Conference (WCNC 2000)*, Vol. 3, pp. 1530-1535, 2000.

[103] L. Zhang, Z. Cao, and C. Gao, "Application of RS-coded MPSK modulation scenarios to compressed image communication in mobile fading channel," *Proceedings 2000 52nd IEEE Vehicular Technology Conference*, VTS-Fall VTC.2000, Vol. 3, pp. 1198-1203, 2000.

[104] P. Dang and P. Chau, "Image encryption for secure Internet multimedia applications," *IEEE Transactions on Consumer Electronics*, Vol. 46, Issue. 3, pp. 395-403, Aug. 2000.

[105] F. Duan, I. King, L. Xu, and L. Chan, "Intra-block max-min algorithm for embedding robust digital watermark into

images," In Horace H.S. Ip and Arnold W.M. Smeulders, editors, *Proceedings of the IAPR International Workshop on Multimedia Information Analysis and Retrieval, MINAR'98*, Lecture Notes in Computer Science, Vol. 1464, pp. 255-264, Berlin Heidelberg, Germany, 1998. Springer-Verlag.

[106] J. Cox, J. Kilian, F. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Transactions on Image Processing*, Vol. 612, pp. 1973-1987, Dec. 1997.

[107] A. Moulin, "The role of information theory in watermarking and its application to image watermarking," *Signal Processing 2001*, Vol. 81, pp. 1121-1139, 2001.

[108] M. Ramkumar and A. Akansu, "A Capacity for Data Hidig in Internet Multimedia," *Symposium on Content Security and Data Hiding in Digital Media*, NJIT, Jersey City, May 1999.

[109] S. Servetto, C. Podilchu and K. Ramchadra, "Capacity Issues in Digital Image Waterarmking," *IEE International Conference on Image Processing*, Chicago, Oct. 1998.

[110] T. Cover and J. Thomas, "Elements of Information Teory," John Wiley and Sones, Lnc., 1991.

[111] A. Ambroze, G. Wade, C. Serdean, M. Tomlinson, J. Stander, and M. Borda, "Turbo code protection of video

watermark channel," *Proceedings IEEE Vision, Image and Signal Processing*, Vol. 148, Issue. 1, pp. 54-58, Feb. 2001.

[112] D. Kirovski and H. Malvar, "Robust covert communication over a public audio channel using spread spectrum," *Proceedings First Information Hiding Workshop*, Pittsburgh, PA, USA, 2001.

[113] J. Princen, A. Johnson, and A. Bradley, "Subband/ transform coding using filter bank designs based on time domain aliasing cancellation," *Proceedings IEEE ICASSP*, pp. 2161-2164, Dallas, TX, Apr. 1987.

[114] H. Malvar, "A Modulated Complex Lapped Transform and its Applications to Audio Processing," *Proceedings IEEE ICASSP*, Mar. 1999.

[115] C. Huang and J. Wu, "A watermark optimization technique based on genetic algorithms," *SPIE Electronic Imaging 2000*, San Jose, Jan. 2000.

[116] C. Huang and J. Wu, "Using Genetic Algorithms as Watermarking Performance Optimizers," *IEEE Signal Processing Letters*.

[117] VirtualDub is distributed under the GNU General Public License, written by Avery Lee, virtualdub.org / phaeron (at), http://www.virtualdub.org/.

[118] C. Hzu and J. Wu, "Digital watermarking for video," *Proceedings 1997 13th International Conference on Digital Signal Processing, DSP 97*, Vol. 1, pp. 217-220, Jul, 2-4, 1997.

[119] X. Niu and S. Sun, "A New Wavelet-Based Digital Watermarking for Video," *9th IEEE Digital Signal Processing Workshop*, Texas, USA, Oct. 2000.

[120] D. Kirovski, and H. Malvar, "Robust spread-spectrum audio watermarking," *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001*, Vol. 3, pp. 1345-1348, 2001.

[121] K. Su, D. Kundur and D. Hatzinakos, "A Novel Approach to Collusion-Resistant Video Watermarking," *Proceedings Security and Watermarking of Multimedia Contents IV SPIE*, E. J. Delp and P. W. Wong, eds., Vol. 4675, pp. 12, San Jose, California, Jan. 2002.

[122] K. Su, D. Kundur and D. Hatzinakos, "A Content-Dependent Spatially Localized Video Watermarked for Resistance to Collusion and Interpolation Attacks," *Proceedings IEEE International Conference on Image Processing*, Oct. 2001.

[123] Y. Wang, J. Doherty, and R. Dyck, "A wavelet-based watermarking algorithm for ownership verification of digital image," *IEEE Transactions on Image Processing*, Vol. 11, No. 2, Feb. 2002.

[124] R. Wolfgang, C. Podilchuk, and E. Delp, "The effect of matching watermark and compression transforms in compressed color images," *Proceedings International Conference on Image Processing (ICIP-98)*, Vol. 1, pp. 440-444, Chicago (IL, USA), Oct. 1998.

[125] The software for this work used the GAlib genetic algorithm package, written by Matthew Wall at the Massachusetts Institute of Technology, http://lancet.mit.edu/ga/.

[126] E. Gilmore, M. Chouikha, "Dictionary Approaches To Image Compression and Reconstruction," *Proceedings of the IASTED International Conference on Signal and Image*, Las Vegas, USA, Sept. 12-16, 1998.