CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Repository@USM

IMPROVEMENT OF HYBRID DIGITAL IMAGE WATERMARKING SCHEMES BASED ON SVD IN WAVELET TRANSFORM DOMAIN

NASRIN MOHAMED HASSAN MAKBOL

UNIVERSITI SAINS MALAYSIA

2015

IMPROVEMENT OF HYBRID DIGITAL IMAGE WATERMARKING SCHEMES BASED ON SVD IN WAVELET TRANSFORM DOMAIN

by

NASRIN MOHAMED HASSAN MAKBOL

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

February 2015

ACKNOWLEDGEMENTS

First and foremost, All Praises be to Allah the Almighty, for delivering me the patience, the strength and the guidance in completing this thesis successfully.

I wish to express my sincere appreciation to those who have supported me in one way or the other during my study and life.

I am extremely grateful to my supervisor, Dr Khoo Bee Ee, for her invaluable guidance, support, motivation, comments and all the useful discussions and brainstorming sessions that have resulted in a lot of improvement in this research. Her deep insights helped me at various stages of my research. I also remain indebted for her understanding and support during the times when I was really down and depressed due to personal family problems. Without her efforts, I would not able to bring this research to a completion.

My warmest acknowledgements to my lovely husband, Taha and my lovely sons, Mohammed and Majd for their endless love and encourage me all along without hesitation. They have been a constant source of strength and inspiration.

My special thanks to the best mother in the world, my father and to all my brothers and sisters for their concern and moral support.

Finally, I would like to thank all staff at the School of Electrical and Electronic Engineering, Universiti Sains Malaysia (USM) for their valuable support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

1-D	One Dimensional
ACO	Ant Colony Optimisation
AES-192	Advanced Encryption Standard with 192 bits key length
AT	Arnold Transform
BCR	Bit Correction Rate
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DE	Differential Evolution
DS	Digital Signature
DWT	Digital Wavelet Transform
FPP	False Positive Problem
FTP	File Transfer Protocol
GA	Genetic Algorithm
HVS	Human Visual System
IPR	Intellectual Property Rights
IWT	Integer Wavelet Transform
JPEG	Joint Photographic Experts Group

LWT	Lifting Wavelet Transform
MOACO	Multi-objective Ant Colony Optimisation
MSF	Multiple Scaling Factor
MSE	Mean Square Error
MZF	Multiple Zooming Factor
NC	Normalised Correlation
RDWT	Redundancy Digital Wavelet Transform
SHA-1	Secure Hash Algorithm-1
SSF	Single Scaling Factor
SVD	Singular Value Decomposition
PSNR	Peak Signal to Noise Ratio
PSO	Particle Swarm Optimisation
PSPC	Print-Scan and Print-Cam

LIST OF SYMBOLS

A	Matrix of an image
A_{cf}	The column flipped matrix A
A_e	The matrix A after translate process
A_r	The rotated matrix A
A_{rf}	The raw-flipped matrix A
A^T	The transpose of matrix A
B	Row-scaled version of matrix A
C	Column-scaled version of matrix A
C_{itr}	Number of iteration within AT
E	The principal component
O	The owner host image
O_W	The obtained watermarked image
O_W^1	Watermarked image of the owner after embed the watermark W^1
O_W^2	Watermarked image of the owner after embed the watermark W^2
O_{WF}	The forged watermarked image
S_{new}	Singular values after watermark embedding
U^1	The singular vector U results of applying SVD after embed W^1 into O

U^2	The singular vector U results of applying SVD after embed W^2 into O
U_{WF}	The singular vector U of the forged watermarked image
V^1	The singular vector V results of applying SVD after embed W^1 into O
V^2	The singular vector V results of applying SVD after embed W^2 into O
V_{WF}	The singular vector V of the forged watermarked image
W^1	The first embedded watermark
W^2	The second embedded watermark
W_F	The forged watermark
W_X	The extracted watermark from an arbitrary image X
c_j	Low-band output coefficients at level j
c_j^i	The original signal in IWT
d_j	High-band output coefficients at level j
$D_{watermark}$	Watermark extraction function
e_j^{i-1}	Old even sets in IWT
E_j^{i-1}	New even sets in IWT
$E_{watermark}$	Watermark embedding function
$f_{pheromone}$	Pheromone communication model
$f\acute{[n]}$	The reconstructed signal in RDWT
HH	Diagonal sub-band

HL	Horizontal sub-band
I	Original/Host data
$\bar{I}$	Watermarked data
Key	Secret key
LH	Vertical sub-band
LL	Approximation sub-band
M	Number of pixels/One of image dimensions
m	Number of pixels/One of image dimensions
N	Number of pixels/One of image dimensions
n	Number of pixels/One of image dimensions
o_j^{i-1}	Old odd sets in IWT
O_j^{i-1}	New odd sets in IWT
P	Prediction operation in IWT
S	Singular values
S_A	Singular values of matrix A
S_W	Singular values of the watermark
$S_{Watermarked}$	The singular values after applying watermark embedding process
S_{WHost}	The obtained singular values after applying SVD on the result of the watermark embedding process

T	The Arnold Transform periodic
U	Singular vector U
U_A	Singular vector U of matrix A
U_W	Singular vector the watermark
U_{WHost}	The obtained singular vector U after applying SVD on the result of the watermark embedding process
$U()$	Update operator in IWT
V^T	Singular vector V
V_A^T	Singular vector V of matrix A
V_{WHost}	The obtained singular vector V after applying SVD on the result of the watermark embedding process
V_W	Singular vector V of the watermark
W	Watermark
$\bar{W}$	Extracted watermark
(x, y)	The original image pixel
(x', y')	The transformed image pixel
x_{min}	The optimal point found in the design space of ACO
$x(i, j)$	The pixel value in the host image
$y(i, j)$	The pixel value in the watermarked image

y_{high}	Output of high-pass filter in DWT
y_{low}	Output of low-pass filter in DWT
α	Scaling factor

PENAMBAHBAIKAN SKIM HIBRID TERA AIR IMEJ DIGIT BERDASARKAN SVD DALAM DOMAIN PENJELMAAN WAVELET

ABSTRAK

Teknik tera air imej digital telah membolehkan maklumat yang tidak jelas pada imej dapat disembunyikan untuk memastikan maklumat tersebut dapat diekstrak kelak. Keteguhan, ketidakjelasan, kapasiti dan keselamatan adalah keperluan yang paling penting dalam mana-mana skema tera air. Teknik tera air imej digital menjadi lebih mencabar apabila keseimbangan di antara keteguhan ketidakjelasan dan keupayaan perlu dicapai. Baru-baru ini skema tera air berasaskan penguraian campuran nilai tunggal (SVD) di dalam domain wavelet telah mendapat banyak perhatian. Tujuan kajian ini adalah untuk membangunkan campuran skema tera air yang mencapai keteguhan dan ketidakjelasan yang tinggi dan juga mengekalkan keseimbangan di antara keteguhan, ketidakjelasan dan kapasiti. Objektif ini dicapai dengan menggabungkan ciri-ciri SVD dan penjelmaan wavelet. Isu keselamatan disebabkan masalah positif palsu (FPP) yang boleh berlaku dalam sebahagian besar skim tera air berasaskan SVD telah dibincangkan dan ditangani. Kajian ini mencadangkan lima gabungan skema tera air berasaskan SVD di dalam domain wavelet. Dalam skim yang pertama, tera air imej kelabu iaitu RDWT-SVD telah digabungkan secara langsung dengan nilai tunggal (S) bagi setiap sub-jalur penjelmaan *wavelet diskrit image* asal. Skim cadangan yang kedua iaitu IWT-SVD-AT pula menggunakan penjelmaan wavelet integer (IWT) yang berbeza daripada RDWT kerana ciri-cirinya. Tera air ini dikarau menggunakan penjelmaan Arnold (AT) sebelum digabungkan ke dalam S bagi setiap sub-jalur yang asal. Walaupun

keputusan yang memberangsangkan oleh skim pertama dan kedua, mereka terdedah kepada FPP. Oleh itu, mereka gagal untuk menyelesaikan pemilikan yang sah. Dalam skim yang ketiga, gabungan skim IWT-SVD telah dicadangkan dengan mekanisme pengesahan berasaskan Tandatangan Digital (DS) untuk menyelesaikan FPP. Skim ini mengatasi skim sebelumnya dari segi keteguhan, kapasiti, keselamatan, pengiraan masa dan pencapaian ketidakjelasan yang tinggi. Untuk dua skim lagi yang dicadangkan, skim keempat dan kelima, FPP sama sekali dapat dielakkan menggunakan strategi penggabungan baru yang berbeza. Dalam skim keempat iaitu IWT-SVD-MOACO, vektor U tunggal tera air digabungkan dengan S daripada sub-jalur IWT LL. Pengoptimuman multi-objektif koloni semut (MOACO)digunakan untuk mencari pelbagai pemfokusan / faktor penskalaan (MZF) yang optimum berbanding menggunakan faktor penskalaan tunggal (SSF) untuk mencapai keseimbangan yang optimum di antara ketidakjelasan dan keteguhan. Akhir sekali, satu skim berdasarkan blok-SVD hibrid iaitu DWT-SVD-HVS menggunakan penjelmaan wavelet diskret (DWT) telah dibangunkan. Tera air binari telah digabungkan dengan beberapa blok yang dipilih bedasarkan kriteria sistem visual manusia (HVS). Selain itu, semua skim yang dicadangkan diuji dengan pelbagai imej berwarna. Semua skim telah menunjukkan prestasi yang baik terhadap pelbagai imej berwarna.

IMPROVEMENT OF HYBRID DIGITAL IMAGE WATERMARKING SCHEMES BASED ON SVD IN WAVELET TRANSFORM DOMAIN

ABSTRACT

Digital image watermarking techniques have enabled imperceptible information in images to be hidden to ensure the information can be extracted later from those images. Robustness, imperceptibility, capacity and security are the most important requirements of any watermarking scheme. Recently, hybrid Singular Value Decomposition (SVD)-based watermarking schemes in the wavelet domain have significantly gained a lot of attention. The aim of this study is to develop hybrid digital image watermarking schemes by combining the properties of SVD and the chosen wavelet transforms to achieve high robustness and imperceptibility, as well as maintaining the trade-off between robustness, imperceptibility and capacity. The security issue due to the false positive problem (FPP) that may be occurring in most of SVD-based watermarking schemes, has been covered and addressed. This study proposes five hybrid robust SVD-based image watermarking schemes in the wavelet domain. In the first scheme, a grey image watermark is embedded directly into the singular values (S) of each redundant discrete wavelet transform transform (RDWT) sub-band of the host image. The scheme is named RDWT-SVD. The second proposed scheme, namely IWT-SVD-AT, utilised the integer wavelet transform (IWT) instead of RDWT due to its properties. The watermark is scrambled using Arnold Transform (AT) before being embedded into the S of each IWT sub-band host. Despite the impressive results by the first and the second schemes, they were vulnerable to the FPP. Thus, they have failed to resolve the rightful ownership. In

the third scheme, a hybrid IWT-SVD scheme is proposed with a novel Digital Signature (DS)-based authentication mechanism to solve the FPP. The scheme outperforms the previous schemes in terms of robustness, capacity, security, computation time and attains high imperceptibility. In the remaining two proposed schemes; the fourth and fifth schemes, the FPP is totally avoided using new different embedding strategies. In the fourth scheme namely IWT-SVD-MOACO, the singular vector U of the watermark is embedded into the S of IWT LL sub-band. Multi-objective ant colony optimisation (MOACO) is used to find the optimal multiple zooming/scaling factor (MZF) instead of the single scaling factor (SSF) to achieve the optimal trade-off between imperceptibility and robustness. Finally, a hybrid SVD block-based scheme namely DWT-SVD-HVS using discrete wavelet transform (DWT) is developed. A binary watermark is embedded into a number of blocks which is selected based on some human visual system (HVS) criterion. The scheme shows a high imperceptibility and good robustness. Finally, all the proposed schemes are evaluated with different colour images and had been shown a successful applicability with colour images.

CHAPTER 1

INTRODUCTION

1.1 Introduction

Diverse computer and communication technologies have resulted in new opportunities to process and distribute multimedia information. These technologies include powerful software and devices (e.g., scanners, printers, and digital cameras) that enable users to create, manipulate, duplicate, and distribute images easily and economically. Amid growing interest in digital media, this field faces many challenges, one of which is the protection and integrity of multimedia content. The development of multimedia networks and the availability of numerous content distribution applications on the Internet (e.g., peer-to-peer file sharing and file transfer protocol) significantly contribute to facilitating the exchange and distribution of multimedia content, even for unauthorised users. Focus should be given to ensure the secure circulation and successful commercialisation of multimedia content and guarantee that multimedia information is used only by authorised users (Khan *et al.*, 2014). The intellectual property rights (IPR) protection of content, verification of origin of content, and identification of authorised parties who initially distribute images to unauthorised parties are gaining attention. These issues are especially important when multimedia content is used in sensitive fields where error is not tolerated, such as in courtroom proceedings, photojournalism, commercial transactions, and medical applications. Thus, several techniques have been generated to address these issues (Singh & Chadha, 2013; Vellasques *et al.*, 2010).

One of the principal technologies designed and used to protect data and secure systems is cryptography. According to traditional cryptographic systems, data can be protected from unauthorised users by using a cryptographic algorithm that allows only the person with the key (or keys) to encrypt or decrypt data. The disadvantage of this strategy for the protection of the data is that once this data is decrypted by pirates, there is no way to protect data and keep track of illegal distribution. Furthermore, legally proving ownership is impossible, which leads to illegal copying and distribution or abuse of information (Terzija, 2008).

Watermarking technology is an alternative approach to IPR protection. The technology embeds imperceptible information (i.e., copyright protection information) into content such that the information cannot be removed during the normal use, which provides an indicator of digital data ownership. Digital watermarking is a reliable technique to safeguard digital content. It imperceptibly alters the original digital content to embed a message on the content, which can be used later for authentication. The technique is employed in numerous applications such as copy protection, authentication, and broadcast monitoring.

1.2 Motivation and Problem Statement

Images constitute a major component of the multimedia content; these images include illustrative diagrams, digital arts, and legacy cultural panels in digital photographs and digitised form. Nowadays, the advances in computer hardware, software, and networks have resulted in threats to copyright infringement and content integrity. Copying, modifying, and distributing images over the Web have become easy. Therefore, technologies to protect digital content are necessary.

Image watermarking is the most researched and published over the last few years. The reason may be due to the large demand on image watermarking products due to the availability of so many images at no cost on the World Wide Web that should be protected (Hartung & Kutter, 1999). Digital image watermarking protects content by embedding a signal (i.e., owner information) into the host image without noticeable degradation in visual quality. Consequently, a watermarked image is developed and marked as public or sent to end users. The extracted or detected watermarks can be used for copyright protection and content authentication purposes (Pérez-Freire *et al.*, 2006). Researchers interested in digital image watermarking face challenges in creating new algorithms with suitable properties (requirements) to serve these intended applications. The most attractive properties that are essential requirements for any watermarking technique are robustness, imperceptibility, security, and capacity.

A trade-off always exists among robustness, capacity, and imperceptibility (Cox *et al.*, 2007; Bhatnagar, 2012). For instance, enhancing the watermark robustness would in turn reduce its imperceptibility because of the higher watermark energy placed on the cover image (Aslantas, 2009). Moreover, higher capacity would compromise its imperceptibility because more modifications the cover image are needed to embed the watermark. Hence, developing any watermarking technique typically requires finding a balance among these conflicting requirements (Pérez-Freire *et al.*, 2006; Singh & Chadha, 2013). The fourth essential property which is security, refers to scheme resistance against hostile attacks. Invisible watermarks ensure that attackers cannot access secured data to remove or alter them.

The current challenge is achieving the trade-off among the most important watermarking requirements (i.e., robustness, capacity, and imperceptibility). High robustness against attacks and maintenance of good visual quality for the watermarked image, which is the core motivation for most existing watermarking schemes. Watermarking technologies prioritize robustness and imperceptibility, which are the major requirements that differentiate watermarking from other data protection technologies (Cox *et al.*, 2007; Bhatnagar, 2012). Various image watermarking techniques have been established to address related problems. These techniques are categorised into two sets according to embedding domain: spatial domain techniques and transform domain techniques. The wavelet technique under transform domain techniques have gained popularity because of its properties. The wavelet transform has accurate model aspects HVS because of multi-resolution analysis (Chang *et al.*, 2005; Maity & Kundu, 2011).

Most image watermarking schemes improve their performance by combining two or more transforms, which are referred to as hybrid schemes. The idea emerged based on the assumption that combining two or more transforms can make up for the defects of an individual transform and result in an effective scheme (Ganic & Eskicioglu, 2005; Lai & Tsai, 2010; Loukhaoukha, 2011; Ali, Ahn & Pant, 2014). The incentive for developing hybrid schemes is to use the properties of the incorporated transforms and achieve the required goals. The success of hybrid schemes in achieving the desired goals depends on the successful selection of the involved transforms. The transforms are selected according to their properties, and these properties are used to achieve a compromise between watermarking properties. Several robust hybrid digital-image watermarking schemes based on singular value decomposition (SVD) in the wavelet domain were developed a few years ago (Liu & Tan, 2002; Ganic & Eskicioglu, 2005;

Loukhaoukha & Chouinard, 2009; Lai & Tsai, 2010; Lagzian *et al.*, 2011*a*; Rastegar *et al.*, 2011; Lai, 2011*a*). SVD is a numerical analysis tool used to analyse matrices. A matrix in SVD is decomposed into three matrices that have the same size as the original matrix; U, S and V, where S represents singular values, and U and V represent the left and right singular vectors.

SVD has many important mathematical properties that are useful in a lot of applications. Newly developed SVD-based watermarking schemes perform effectively in keeping minor changes with largely altered singular value S, which are caused by the attacks. Some researchers have demonstrated and analysed the S of the image under geometrical distortions (Chung *et al.*, 2007; Lai, 2011*b*). Most SVD-based watermarking schemes display high robustness against image processing attacks and geometrical attacks while maintaining good imperceptibility, which is the main goal of any watermarking scheme. SVD is preferred for implementation with other transforms because it requires extensive computations when applied separately. The various embedding strategies of SVD-based watermarking schemes are based on the U, S, and V matrices involved. Each embedding type has advantages and disadvantages. Due to the stability and the properties of S, most of the researchers prefer to embed into S. Despite the stability and the robustness of the SVD-based image watermarking when embedding is performed in S, these schemes are subjected to high probability of the false positive problem (FPP) (Ling *et al.*, 2011; Guo & Prasetyo, 2014*a*). Recently, avoiding FPP is one of the active research topics in the SVD-based image watermarking area.

Satisfying all the requirements in addition to avoiding the FPP are important in any SVD-based image watermarking schemes to serve in some important applications such as copyright protection and authentication. Furthermore, because of the importance and widespread use of the colour images, developing colour image watermarking schemes become an important issue.

1.3 Research Aim and Objectives

This study focuses on developing new hybrid digital-image watermarking schemes based on SVD in the wavelet transform domain. This study aims to combine SVD properties and the properties of selected wavelet transforms to develop new schemes that satisfy the most important watermarking requirements. This study targets high robustness while maintaining good imperceptibility and high embedding capacity. Moreover, this study aims to solve or prevent security issues caused by FPP, which occurs in many SVD-based watermarking schemes based on embedding into S and adopting both U and V as secret keys.

Several objectives are identified to be achieved in this research. They are listed as follows:

i. To study and develop hybrid SVD-based image watermarking schemes using different wavelet transforms.

ii. To study and develop SVD-based embedding strategies to solve/avoid the FPP.

iii. To assess the developed watermarking schemes by conducting a comprehensive analysis about their feasibility, robustness, and performance.

iv. To modify the developed watermarking schemes to be suitable for colour images.

1.4 Research Scope

The scope of this study includes developing SVD-based digital-image watermarking schemes that can be used in various critical and attractive applications in the digital domain (e.g., copyright protection). The targeted schemes use SVD properties and chosen wavelet transforms to successfully achieve the main requirements of any watermarking scheme, which are robustness, imperceptibility, capacity, and security. Furthermore, maintaining the trade-off between robustness, imperceptibility and capacity, which is considered as a challenge. The targeted schemes must prove their reliability by overcoming or avoiding the drawbacks caused by the high probability of FPP, which occur in most SVD-based watermarking schemes. This will serve for the copyright protection and authentication applications. Furthermore, a suitable modification on the proposed schemes is done to make them suitable for colour images which are commonly used nowadays. The schemes must possess the following specific properties:

- Robustness: to resist all possible image processing distortions, especially JPEG compression attack, which is considered one of the most common attacks on digital images. Furthermore, to resist all geometrical attacks which represent high challenging attacks.
- Imperceptibility: to maintain good visual quality of the image after embedding process (watermarked image) which is the main target of any watermarking scheme. This helps to add more security and preserve the commercial value of the host image.
- Capacity: to embed a large amount of watermark without affective the imperceptibility in addition to achieve high robustness against attacks. Embedding

multiple watermarks in different sub-bands is strongly contributed to improve the robustness where each sub-band has its different resistance against different attacks. Besides, it ensures to recover at least one watermark from any sub-band.

- Security: to provide high level of security such as blind extraction/detection, resistance to malicious attacks, encrypt/scramble the watermark, and solve/avoid the FPP. These security measure help to provide high reliability to serve number of significant applications.

1.5 Thesis Outlines

The remainder of this thesis is organised as follows: Chapter 2 presents an overview of topics related to digital watermarking, which begins by introducing the differences among the available security techniques (i.e., cryptography, steganography, and watermarking) and explaining why digital watermarking is preferred. An overview of the basic principles of digital watermarking (i.e., framework, properties, applications, and classifications) is presented. Other related topics are also presented, including Arnold Transform (AT), Ant Colony Optimisation principle (ACO), and reviews of SVD and SVD-based digital image watermarking techniques. The chapter ends with a comparison of some previous SVD-based watermarking schemes.

In Chapter 3, five different SVD-based image watermarking techniques in different wavelet domains are proposed, analysed, and studied. All of these schemes processes (e.g., embedding and extraction processes) are provided.

In Chapter 4, the experimental setups and results of all proposed schemes are presented. The proposed schemes are examined and evaluated using different test images. Comparative analyses and results with previous schemes are explained.

Chapter 5 presents the applicability of all proposed schemes for colour image application. All proposed schemes are tested on colour images in addition to grey images. Finally, Chapter 6 provides the conclusions, contributions, and suggestions for future research work.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

The widespread distribution of digital data (e.g., images, text, video or audio) over the Internet has highlighted the importance of increasing knowledge about the protection of Intellectual Property Rights (IPR) (Singh & Chadha, 2013). Various types of data are stored and converted to digital format so that they can be copied easily without loss of quality and be efficiently distributed thereafter. Pirates exploit the reproduction, manipulation, and retransmission of digital data and violate the real owner's copyright. Accordingly, designing techniques that preserve the ownership of digital information is the motivation for developing advanced multimedia services. A few approaches have been formulated to protect such data, and one of them is encryption (cryptography) (Khan *et al.*, 2014).

Cryptography is the most common method used to protect digital content. In this method, the message content is encrypted using a secret key, which is also called the decryption key. Only authorised parties who have purchased legitimate copies of the content are given access to the key. However, this method fails to help the seller monitor how a legitimate customer handles the content after decryption (Cox *et al.*, 2007).

Data hiding methods (Steganography and Watermarking) have been examined to protect the digital media content and overcome the challenges posed by the adoption of cryptography. Data hiding is a general term that refers to techniques of hiding messages in the content; hiding means either saving the content with the presence of secret information or making the information imperceptible (as in watermarking) (Cox *et al.*, 2001, 2007). Watermarking is an alternative and complementary technology to cryptography that can protect the content even after decryption (Cox *et al.*, 2007; Khan *et al.*, 2014). Steganography is an alternative tool to cryptography in terms of providing privacy and security. Unlike cryptography that encrypts messages, steganography hides the messages into the content so that their existence is not revealed (Cox *et al.*, 2007).

2.2 Watermarking and Steganography

Cryptography is concerned with hiding the message content itself and not the existence of the message, whereas steganography and watermarking conceal the existence of the message. The fundamental differences between watermarking and steganography can be summarised as follows:

1. Watermarking is the process of hiding a message (called the watermark) into a cover medium (called the host) such that the message cannot be removed or replaced (Wolfgang & Delp, 1996; Nikolaidis & Pitas, 1996; Cox *et al.*, 1997; Swanson *et al.*, 1998; Wolfgang *et al.*, 1999; Hsu & Wu, 1999; Langelaar *et al.*, 2000; Cox *et al.*, 2007; Keskinarkaus, 2012; Ali, Ahn & Pant, 2014). Steganography is also the process of hiding a message in a way that an eavesdropper cannot detect it (Johnson & Katzenbeisser, 2000; Johnson *et al.*, 2001; Provos & Honeyman, 2003).

2. In watermarking systems, the hidden information is usually related to either the protected object or its owner, and no relationship exists between the hidden

data and the cover work as in steganographic systems (Petitcolas *et al.*, 1999; Katzenbeisser *et al.*, 2000).

3. In considering watermarking techniques, priority is given to robustness. Although robustness is also required in some steganography applications, capacity is generally given priority (Petitcolas *et al.*, 1999; Nikolaidis *et al.*, 2001; Cox *et al.*, 2007).

4. Preserving the visual quality of the host is an important requirement in watermarking applications, but it is not a major concern in steganography where the cover signal serves as the carrier.

Watermarking still considered as an active research field. It has three attributes that distinguish it from other related technologies and make it invaluable to some applications. These attributes are as follows (Cox *et al.*, 2007):

1. Watermarks are imperceptible, and therefore do not detract from the aesthetics of the image.

2. Embedded watermarks, unlike header fields, cannot be removed even if the hosts are displayed in or transformed into another format, because watermarks are inseparable from the host where they are embedded.

3. Finally, watermarks undergo the same transformations as the host image itself. This means that it is sometimes possible to learn something about those those transformations by looking at the resulting watermarks. This means that it is sometimes possible to learn something about those transformations by looking at the resulting watermarks.

2.3 Digital Watermarking Framework

A typical greyscale image of 8 bits per pixel (each pixel has $2^8 = 256$ grey levels) is mainly used for experimentation in the research community, but few researchers use other image types such as colour and halftone images. The general model of any watermarking system consists of two main phases: embedding and extraction (Fridrich, 1999; Hartung & Kutter, 1999; Cox *et al.*, 2001; Meerwald & Uhl, 2001; Keskinarkaus, 2012; Ali, Ahn & Pant, 2014). These consecutive phases comprise the embedding, distribution, extraction, and decision phases. These phases are explained in the following subsections.

2.3.1 Embedding Phase

The first step in any digital watermarking system is the embedding process. In this process, a digital datum called "watermark" (W) is embedded into the original/host data (I) to obtain a watermarked datum indicated as ($\bar{I}$). Watermarking approaches vary according to certain criteria to satisfy the specified requirements. These requirements are explained in Section 2.4. The type of embedding domain is one of these criteria. Two embedding domains exist, namely, spatial and transform. In spatial domain techniques, the watermark is embedded by directly changing the pixel values of the original data (Van Schyndel *et al.*, 1994; Lu *et al.*, 2000; Keskinarkaus, 2012). In transform domain techniques, the original data are converted to coefficients by using a transform, such as discrete cosine transform (DCT) (Cox *et al.*, 1997; Lin & Chen, 2000; Al-Haj, 2007; Lai & Tsai, 2010; Lai, 2011*b*), discrete wavelet transform (DWT) (Al-Haj, 2007; Lai & Tsai, 2010; Gupta & Raval, 2012; Run *et al.*, 2012), redundant discrete wavelet transform (RDWT) (Lagzian *et al.*, 2011*a*), Radon transform (Zhu

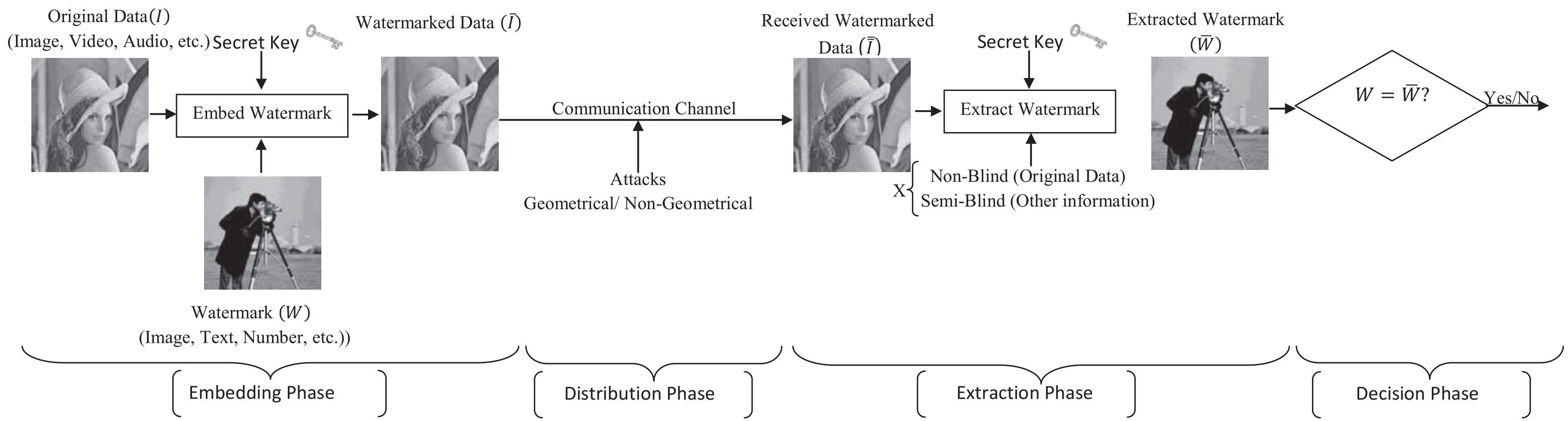

Figure 2.1: General model of watermarking system.

et al., 2010; Rastegar *et al.*, 2011), and lifting wavelet transform (LWT) (Loukhaoukha & Chouinard, 2009; Loukhaoukha, 2011). The embedding process occurs by altering the coefficients of these transforms. Different encoding functions are conducted to embed the watermark, including additive, multiplicative, and quantisation functions. The mathematical formula of the embedding process is represented as follows:

$$\bar{I} = E_{watermark}(I, Key, W) \tag{2.1}$$

where $E_{watermark}$ denotes the encoding function to embed the watermark, and Key denotes the secret key. The original datum (image), the watermarked datum (watermarked image), and the watermark (also, image) are represented respectively by I, $\bar{I}$ and W as illustrated in Figure 2.1.

2.3.2 Distribution Phase

The next step is distributing or transmitting the obtained watermarked digital image from the previous phase (embedding) through a digital channel. Such a process is called the distribution phase and is achieved by publishing the data on a Web server or selling the data to a customer. The digital data are subjected to many risks that may damage the transmitted data as they travel through the digital channels. These risks, called "attacks", include compression and image processing (non-geometrical) attacks, which are the most commonly applied, and hostile (geometrical) attacks.

2.3.3 Extraction Phase

In the extraction phase, the watermarked image ($\bar{I}$)(may be distorted) is received by the receptor. The watermark (W) is then extracted or detected according to the scheme and the application where the scheme is used. The extraction process, as shown in Figure 2.1, varies according to the information required by the process itself. Extraction can be classified into three types: blind, semi-blind, and non-blind. Blind extraction needs the secret key only and not the original data to complete the extraction process; the schemes supplied by such an extraction process are called blind schemes (Ganic & Eskicioglu, 2005; Lai *et al.*, 2007; Lai & Tsai, 2010; Lagzian *et al.*, 2011*a*; Loukhaoukha *et al.*, 2014). Semi-blind extraction does not rely on the original data but may rely on other information including the secret key (Ni *et al.*, 2005; Chang, Chou & Lu, 2007; Sang & Alam, 2008; Gokhale & Joshi, 2012). Non-blind schemes need the original data and the secret key to complete the extraction process (Zaboli & Moin, 2007; Dharwadkar *et al.*, 2011; Minamoto & Ohura, 2011; Pradhan *et al.*, 2012).

Similar to embedding process (Equation (2.1)), the extraction process can be expressed mathematically as follows:

$$\bar{W} = D_{watermark}(\bar{I}, Key, X) \tag{2.2}$$

where $D_{watermark}$ represents the decoding function to extract the desired watermark. $\bar{I}$ and *Key* denote the watermarked image (may be distorted) and the secret key, respectively. The X relies on the scheme type, where X is the original data if the scheme is non-blind, otherwise it indicates other information such as the watermark if the scheme is semi-blind. $\bar{W}$ denotes the extracted watermark and $\bar{W}$, *Key* and X are illustrated in Figure 2.1. Similar to the embedding function, the extraction function can be an additive, multiplicative or quantisation.

2.3.4 Decision Phase

The final stage in any watermarking system is the decision phase. In this phase, the correlation of the extracted watermark with the original watermark, or the similarity between ($\bar{W}$) and (W), is analysed. Several similarity measures are employed to evaluate the correlation between ($\bar{W}$) and (W), such as the normalised correlation (NC) and bit correction rate (BCR). NC is mostly used for a greyscale watermark logo, whereas BCR is mostly used for a binary watermark logo (Maity & Kundu, 2011). The NC is defined as:

$$NC(W,\bar{W}) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}[W(i,j)-\mu_W][\bar{W}(i,j)-\mu_{\bar{W}}]}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}[W(i,j)-\mu_W]^2}\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}[\bar{W}(i,j)-\mu_{\bar{W}}]^2}} \tag{2.3}$$

where $M \times N$ represent the number of pixels in the watermark, W and $\bar{W}$ indicate the original watermark and the extracted watermark, respectively, while the μ_W and $\mu_{\bar{W}}$ indicate the mean of the original watermark and the mean of the extracted watermark respectively.

The correlation coefficient between W and $\bar{W}$ can be between -1 and 1. If the NC value is 1, then the extracted and original watermarks are identical (Langelaar *et al.*, 2000; Mabtoul *et al.*, 2006). If it is near 1, then the extracted and original watermarks are strongly correlated. If it is near -1, then the extracted watermark remains strongly correlated with the original watermark but looks similar to the negative film. If it is near 0, the extracted watermark is totally uncorrelated with the original watermark (Bhatnagar, 2012). Generally, *NC* is considered acceptable if its value is 0.75 or higher (Al-Haj, 2007).

The other criteria which is *BCR* can be defined as:

$$BCR(W,\bar{W}) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\overline{W_{ij}\otimes\bar{W}_{ij}}}{M\times N} \tag{2.4}$$

where N and M represent the number of pixels in the watermark, and W and $\bar{W}$ indicate the original and the extracted watermark, respectively. The $\otimes$ denotes the XOR logical operation. When the *BCR* value is near 1 under applicable attacks, the scheme is robust against these attacks (Maity & Kundu, 2011).

2.4 Digital Watermarking Properties

Any watermarking system ought to satisfy some properties when it is implemented. The desired properties of a watermarking system are dictated by the application in which the system is used. Robustness, capacity, imperceptibility, and security are the basic and most important properties of any watermarking system. A given watermarking system performance can be assessed based on these properties. Robustness refers to the property of watermarks to resist signal processing operations, whereas imperceptibility refers to the property of watermarks to be imperceptible. Following is the illustration of these properties separately.

2.4.1 Robustness

Robustness is the ability of a scheme to maintain the validity of including a watermark even after being subjected to geometrical or non-geometrical attacks. A robust watermark must be resilient against potential attacks and can be retrieved or detected even after all the attacks occur. For example, a watermark is robust against JPEG compression, if it can be detected even after exposure to lossy compression operation. Filtering, noise insertion, and smoothness are examples of signal processing operations that are non-geometrical attacks. Rotating, translation, and scaling are examples of the geometrical attacks. Robustness varies from one operation to another as well as from one scheme to another. On the one hand, there is no watermarking scheme can probably resist all kinds of attacks; on the other hand, not all applications require robustness against all attacks. Thus, the robustness of a scheme against attacks is application dependent (Cox *et al.*, 2001; Koz, 2002).

In television and radio broadcasting, for instance, the watermarking system should provide good robustness against lossy compression, digital-to-analogue/analogue-to-digital conversion (applied on the transmitter and receiver sides), additive noise during the transmission, and small horizontal and vertical translations. On the other hand, watermarks for this application need not survive rotation, scaling,high-pass filtering, or any of a wide variety of degradations that occur only prior to the embedding of the watermark or after its detection. (Cox *et al.*, 2001). In Web publishing, where compression techniques are required for images and videos that may include watermarks, and those watermarks must resist compressions. However, when only parts of the multimedia object are needed, robustness against cropping attacks is required (Ming *et al.*, 2008). For fingerprinting applications, robustness and security are the two essential requirements to withstand malicious attacks. Moreover, different levels of resistance against attacks can be achieved by different sub-bands, and the robustness strength varies according to these sub-bands. In image watermarking, a higher level of robustness can be achieved by embedding the watermark into the perceptually important parts of an image located in the LL sub-band if a transform domain is adopted (Vellasques *et al.*, 2010). *NC* and *BCR* are proposed to assess the robustness of a watermark after the watermark has been exposed to attacks. These measures evaluate the similarity between the original watermark W and the extracted watermark $\bar{W}$ after applying attacks. As mentioned, *NC* is mainly used when the watermark is a grey image or logo, and *BCR* is mainly used when the watermark is a binary image or logo (Maity & Kundu, 2011). *NC* and *BCR* can be obtained using Equations (2.3) and (2.4), respectively. Usually, higher *NC* or *BCR* indicates increased robustness against the encountered attacks.

Notably, robustness may be undesirable in certain cases, such as in authentication applications where fragile watermarks are required. In fact, fragile watermarks form an important branch of watermarking that has attracted significant attention from researchers. A fragile watermark is designed not to be robust; thus, applying any signal processing operation to the watermarked image causes the loss of the watermark (Cox *et al.*, 2001; Piper & Safavi-Naini, 2013).

2.4.2 Imperceptibility

Imperceptibility or perceptual transparency is one of the most important requirements of any watermarking system regardless of its purpose or application. Artefacts, which occur as a result of the watermarking process, are not only annoying but may also reduce or eliminate the commercial value of the watermarked data (Katzenbeisser *et al.*, 2000). In the digital watermarking field, watermarked data (image, video, and audio) that look similar to the original/host data are preferred. In case of distortions caused by embedding the watermark, the aesthetic value of the watermarked data may be degraded. Moreover, they raise doubts and endanger watermark security.

The quality of the digital image after embedding the watermark should be preserved during the design of any image watermarking system. Increasing robustness by embedding a high-capacity watermark may distort imperceptibility; thus, a trade-off should be made between these two properties to set them to the required level (Haque, 2008). Some watermarking schemes apply human visual system (HVS) models and employ the properties of these models to embed the watermark into imperceptible regions in the host object while enhancing the robustness of the watermark through additive embedding of a large part of the watermark into image regions that have a complex texture. The HVS model claims that the human eyes are less sensitive to changes that occur in high-texture regions compared with those in flat regions, so increasing the embedding capacity in the high-texture regions does not significantly affect the imperceptibility. The criteria to calculate the imperceptibility of the watermarked image are classified as either objective or subjective. Both types of criteria should be good for imperceptibility to be considered good (Cox *et al.*, 2007). The objective criteria to measure the imperceptibility is the Peak Signal-to-Noise Ratio (*PSNR*).

The *PSNR* can be calculated as follows:

$$PSNR = 10log_{10}[\frac{max(x(i,j))^2}{MSE}] \tag{2.5}$$

(a) Host Lena image (512×512). (b) Watermarked Lena image (*PSNR* 54.0353 dB).

Figure 2.2: Example of host and watermarked images.

where $max(x(i,j))$ is the maximum possible pixel value in the image, while the mean square error (*MSE*) between the host image x and the watermarked image y can be defined as:

$$MSE = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} [x(i,j) - y(i,j)]^2 \tag{2.6}$$

where m and n are the image dimensions. Equations (2.5) and (2.6)are used for greyscale images. For colour image watermarking, the MSE equation (Equation (2.6)) is modified as follows:

$$MSE = \frac{1}{3 * m * n} \sum_{i=1}^{m} \sum_{j=1}^{n} ([x_{C_1}(i,j) - y_{C_1}(i,j)]^2 + [x_{C_2}(i,j) - y_{C_2}(i,j)]^2 + [x_{C_3}(i,j) - y_{C_3}(i,j)]^2) \tag{2.7}$$

where x_{C_1} and y_{C_1}, x_{C_2} and y_{C_2}, and x_{C_3} and y_{C_3} indicate the values of C_1, C_2, and C_2 colour channels of the host image x and the watermarked image y with size of $m \times n$.

For good imperceptibility, the watermarked image should look nearly the same as the host image. In other words, the host image is not significantly affected by the embedding process. Some researchers believe that an imperceptibility of 38 dB is the minimum acceptable value of *PSNR* (Lee *et al.*, 2012), whereas others declare that 30 dB is the minimum acceptable value for perceptual fidelity (Chen *et al.*, 2005; Chang, Lin, Tseng & Tai, 2007; Maeder *et al.*, 2008). Subjective criteria can be evaluated visually by comparing the watermarked image against its host image. Figure 2.2(a) shows the host Lena image (512×512), and Figure 2.2(b) shows the corresponding watermarked Lena image (*PSNR*=54.0353 dB) after a watermarking scheme is applied (an example from the proposed schemes in the next chapter). This value is considered high based on the objective criteria, while Figures 2.2(a) and 2.2(b) show that the two images look the same based on the subjective criteria.

2.4.3 Capacity and Data Payload

In general, data payload indicates the number of watermark bits encoded and embedded within a unit of time or message (Cox *et al.*, 2007). The watermark capacity refers to the maximum repetition of data payload inside an image (Pérez-Freire *et al.*, 2006). A watermark may have a high capacity but a low data payload. For instance, embedding a 1-bit watermark many times across the image may be necessary. In other words, the capacity of a watermark is the amount of watermark information that can be included into the image, whereas the capacity to embed multiple watermarks into an image is the sum of the data payloads of all individual watermarks (Alattar *et al.*, 2003).

Capacity has attracted the attention of researchers because of its importance. An increase in the embedding capacity improves system robustness but reduces the viewing quality. However, this condition depends on the application. For instance, the presence or absence of one bit is sufficient for a copy control application, whereas other applications, such as copyright protection or fingerprinting, require considerably more data (Cox *et al.*, 2007).

Table 2.1 presents an example of rough estimates of low, medium, and high payloads, particularly for images (Zeki, 2009).

Table 2.1: The Payload categorisation based on the message size (Zeki, 2009).

Message Size % of the Host Message	The Embedding Capacity
0 - 2 %	Low
2 - 10%	Medium
10 - 20%	High
$>$ 20%	Very High

2.4.4 Security

Security in watermarking techniques can be assessed in the same way as in encryption techniques. Only the authorised party should be able to access the watermark even if some pieces of information are available. For instance, even if the watermark embedding or extraction algorithm is known by an unauthorised party, this piece of information (i.e., the algorithm) cannot help the attacker to recover the watermark because the security lies in the secret key selection. Therefore, such a technique is considered truly secure (Swanson *et al.*, 1999; Xie *et al.*, 2006; Yen & Tsai, 2008).

Oostveen *et al.* (2000) and Barni *et al.* (2003) consider security as the inability of unauthorised users to access the raw watermarking channel. Access refers to the ability to write, modify, remove, and detect the raw watermark bits. In other words, the security