M. Tech (VLSI Design & Embedded System) ABHINAV GUPTA 2018

FEATURE POINT EXTRACTION BASED GEOMETRICALLY RESILIANT DIGITAL IMAGE WATERMARKING

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF

MASTER OF TECHNOLOGY
IN
VLSI Design &Embedded System

Submitted by:

ABHINAV GUPTA

2K16/VLS/01

Under the supervision of

Dr. S. INDU
(PROFESSOR)
(HOD, ECE DEPTT.)

Electronics & Communication Engineering
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

JUNE, 2018

FEATURE POINT EXTRACTION BASED GEOMETRICALLY RESILIANT DIGITAL IMAGE WATERMARKING

A PROJECT REPORT

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF

BACHELOR OF TECHNOLOGY
IN
VLSI Design &Embedded System

Submitted by:

ABHINAV GUPTA

2K16/VLS/01

Under the supervision of

Dr. S. INDU
(PROFESSOR)
(HOD, ECE DEPTT.)

Electronics & Communication Engineering
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

JUNE, 2018

Electronics & Communication Engineering
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I ABHINAV GUPTA, Roll No. 2K16/VLS/01 student of M.Tech (VLSI & Embedded Sytem), hereby declare that the project Dissertation titled "Feature Point Extraction based geometrically resilient digital image watermarking" which is submitted by me to the Department of Electronics & Communication Engineering, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Place: Delhi **ABHINAV GUPTA**
Date:

Electronics & Communication Engineering

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled "Feature Point Extraction based geometrically resilient digital image watermarking" which is submitted by ABHINAV GUPTA, Roll No 2K16/VLS/01 Electronics & Communication Engineering, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the students under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi

Date:

Dr. S. INDU

(SUPERVISOR)

(PROFESSOR)

(HOD, ECE Deptt.)

ACKNOWLEDGEMENT

I have a lot of people to be thankful to who have helped me immensely throughout this journey of mine. This dissertation is the result of work of almost two years, whereby I have been accompanied and supported by many people, to whom I would like to express my gratitude.

I would like to express my deep gratitude to my supervisor, Dr. S. INDU (Professor, HOD, ECE Deptt.) who has provided me with guidelines for my work and has supported me with valuable advice through my study. Without her constant support and motivation this thesis would not have seen the light of this day.

I would like to take this opportunity to express my appreciations to all my friends and colleagues at VLSI Department, Delhi Technological University.

The two people to whom I believe I owe all and saying just ‘Thanks’ would be insufficient, are my parents. I would like to thank them and my siblings for believing in me and supporting me.

ABHINAV GUPTA
Roll no: 2K16/VLS/01
M.TECH. (VLSI Design and Embedded System)
Department of Electronics & Communication Engineering
Delhi Technological University
Delhi – 110042

ABSTRACT

A digital image watermarking scheme based on the features of the original image has been implemented in order to resist geometrical transformations. Scale invariant feature transform(SIFT) has been used for interest point extraction in order to locate the invariant regions for signal embedding. Additionally, the watermarked image was subjected to various image processing attacks such as blurring, rotation, shifting, addition of noise (Gaussian, salt & pepper), histogram equalization, JPEG Compression etc. in order to damage the inserted watermark and produce copyright infringement issues. This thesis work reviews the performance of various watermarking algorithms in the presence of different types of attacks to justify the robustness of the algorithm. The parameters considered to review the performance under these attacks were Peak Signal to Noise Ratio(PSNR), Normalized Correlation(NC) and Tamper Assessment Function(TAF). The performance of this algorithm has been compared with known conventional image watermarking techniques such as Discrete Cosine Transform(DCT), Discreet Wavelet Transform(DWT) and DWT-SVD (DWT Singular Value Decomposition) and verified using MATLAB.

CONTENTS

List of figures

List of tables

List of Abbreviations

HVS	Human Visual System
DCT	Discrete Cosine Transform
SIFT	Scale Invariant Feature Transform
DWT	Discrete Wavelet Transform
SVD	Singular Value Decomposition
JND	Just Noticeable Difference
DFT	Discrete Fourier transform
JPEG	Joint Photographic Experts Group

CHAPTER 1

INTRODUCTION

Over the last couple of decades there has been a tremendous increase in the demand for digital data such as images, audio and videos since these can be made available at a single click of a button at one's own pleasure. As a result, this trend has led to a lot of piracy where illegal and unauthorised duplication of data is carried out without the consent of the owner of that content. Hence, it is of utmost importance that copyright owners must devise a way in order to prove their claim over the digital content they share with their customers. The customer must also be ensured that the content they have paid for is authentic and not pirated. A digital signature to be signed by the owner of the document can act as a means to insure authenticity of the document being sent over the internet but its use is mostly limited to text-based documents. Another emerging technique in this domain is Digital Watermarking. By the application of this technique, digital multimedia can be protected against the threat of piracy. Cox et al. (1997) defined a digital watermark as "a digital signal which contains information about the creator or distributor of media." A digital watermark is supposed to be incorporated into a digital multimedia file in such a fashion that it is almost indistinguishable to the naked eye. Watermarking is used for content authentication, tamper detection, copyright management, and content protection whereas Steganography is used for secret communication. The basic features of digital watermark include imperceptibility, robustness of watermarking algorithm and security of the hiding place.

Improvement in computing and networking technologies have led to concerns regarding the security of digital multimedia against piracy. Unlimited and free access to digital media has proved to be quite convenient for offenders of pirate copyrighted material. Hence, in order to trace and detect copyright infringement the need to develop digital watermarking algorithms significantly increases. At present, copyright protection, data authentication, and integrity verification have become a major concern in the digital multimedia. Watermarking methodology may be classified into three broad categories: robust, fragile, and semi-fragile. Robust watermarks are broadly used for copyright protection and verifying owner credentials, therefore a robust watermark is supposed to be invariant against probable image processing attacks, while on the other hand fragile watermarks are usually employed in those applications where integrity verification is a stringent requirement. Based on the watermark embedding procedure followed, watermarking may be carried out in spatial domain or frequency domain. Watermark embedding in digital images is achieved by embedding a data stream discretely inside a cover image.

A basic watermarking procedure is depicted using Figure 1.1. It majorly consists of two basic steps which are as follows:

(i) Embedding of the watermark

(ii) Extraction of the embedded watermark.

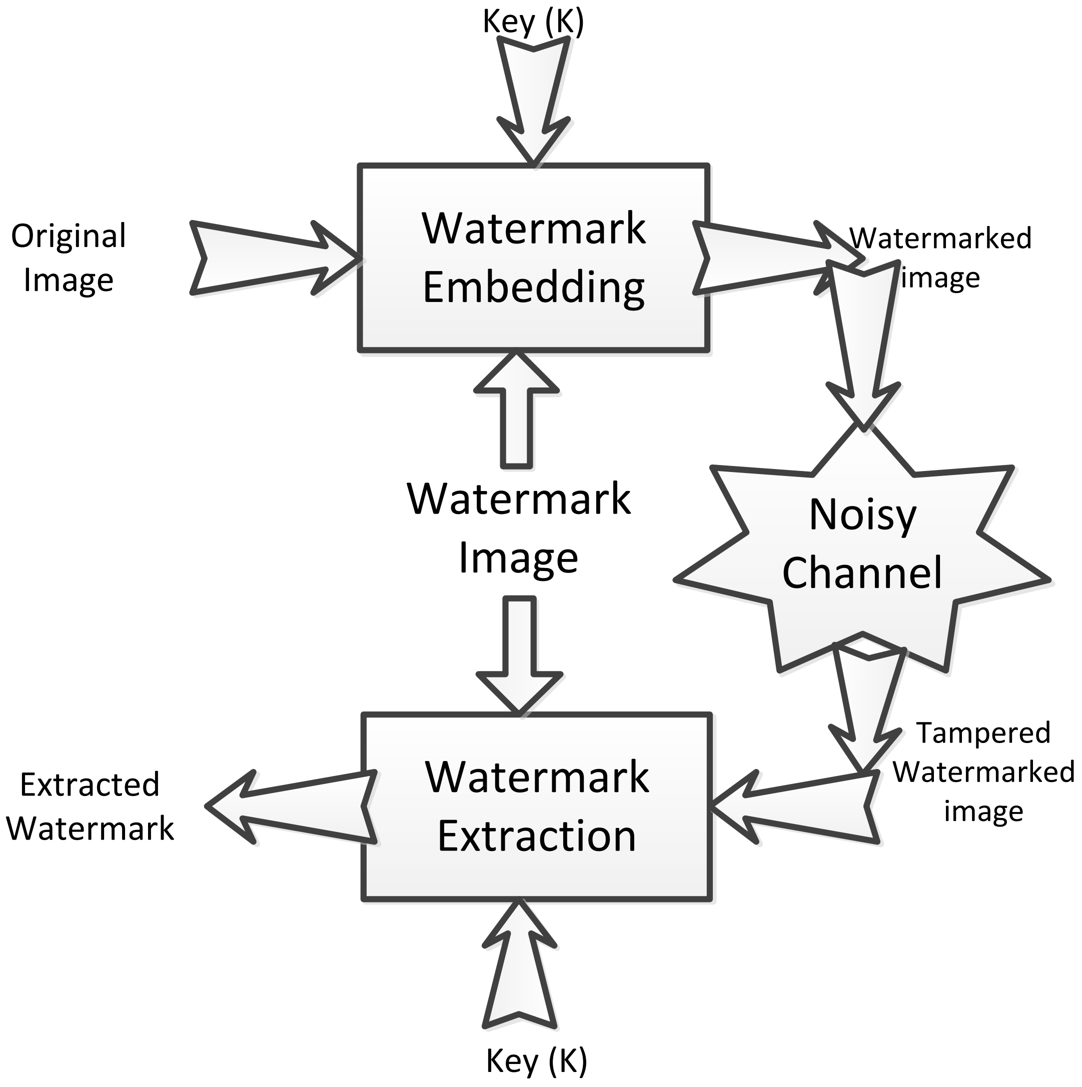


Figure 1.1: The Watermarking Procedure (Non-Blind Image Watermarking)

Watermark embedding is achieved by making use of a private key. The key, here, in fact determines the location in the image where embedding of the watermark is supposed to be carried out. Once embedding of the watermark has been achieved, it's vulnerable to various attacks which might be intentional or even unintentional in nature. The next step is to extract the watermark that has been embedded previously. To achieve this, we take the tampered watermarked image and employ the same key as used earlier in order to determine the location of the watermark, which is then extracted by following the same procedure as carried out during watermark embedding. In order to validate the presence of watermark we compare the original watermark with the extracted watermark. If the comparison leads to a similarity index being above a threshold, the data is acceptable and vice-versa.

The estimation of the performance of a watermarking algorithm is its robustness against various kinds of attacks which have been developed to destroying the watermark that has been embedded. In order to develop watermarking algorithms which are more robust in nature, it is quite essential to become aware of such attacks. In case of image watermarking, a watermark which is resilient to geometrical manipulations is essential because transformations such as scaling, cropping and rotation are common. Nonetheless, these attacks may cause severe synchronization difficulties in the detection of the watermark.

A given watermark may be visible or invisible depending on the application it has been targeted for. The embedding process is directed by use of a private key which decides the locations within the image where the watermark shall be embedded. Once the watermark has been embedded it can experience various attacks since images can be digitally processed. These attacks may be unintentional (in case of images, compression or low pass filtering or gamma correction) or intentional (like cropping, rotation, etc). Hence the watermark is designed to be extremely robust against these possible attacks.
Whenever the owner intends to check whether the watermark has been possibly attacked in the multimedia object, he depends on the private key that had been used to embed the watermark. Using this private key, the embedded watermark sequence may be extracted. This extracted watermark might or might not bear a resemblance to the original watermark since the multimedia object might have been attacked. Therefore, to authenticate the presence of the watermark, either the original watermark is used to find

out the watermark signal (non-blind watermarking) from the watermarked image or the original watermark is compared with the extracted watermark and a statistical correlation measure is used to validate the existence of the watermark (blind watermarking).

Prevailing techniques to oppose geometrical manipulations may be categorised into four classes, that is, an exhaustive search [1][2], watermark implanting in the invariant domain [3][4], interleaving synchronization templates[5][6] and using feature detections to determine watermark location[7]. As the concealed information occurs in the geometrically deformed watermarked image, a thorough geometrical search is a viable solution if a hidden signal is the searched target is already known. Lichtenauer *et al.* [2] inspected the false positive watermark detection of this technique to demonstrate its viability. The main challenge in this technique is to ascertain various geometrical distortions in advance. The searching algorithm should be accompanied with some additional information about the embedded watermark in order to decrease the computational load. Watermarking algorithms may seem elegant in the invariant domain but their performance may hamper in presence of various geometrical attacks. A better approach may be to use synchronization templates. A template is typically a repetitive structure that helps in reducing the manipulations applied on the image. A watermark detector, typically, reverses the geometrical operation in accordance with the extracted template for the watermark detection.

In order to evade attack on the template, feature-based approach has gained much attention in the recent years. Kutter *et al.* [7] in 1999 was the first who argued that feature point extraction-based methods are the second-generation watermarking techniques. They demonstrated this concept by using the Mexican-hat wavelet to extricate features and employing the Voronoi diagram to determine local characteristic regions for watermarking. Numerous techniques have been suggested in recent years [8]–[13]. The main idea behind these methodologies is employing such feature or interest point extraction as Scale-Invariant Feature Transform (SIFT) [15] or Harris corner detection [14] or to determine areas of interest, which may then be transformed into regions having known size, shape, and orientation for the successive watermark implanting and discovery. As the interest point extraction is content based, similar areas will be recognized in both watermarking embedding and detection. However, the feature-based watermarking methodology may encounter various problems. Firstly, to make the watermark detectable in a minor invariant area, such transforms as Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT) or Discrete Cosine Transform (DCT) are typically applied and the mid-frequency coefficients are altered considerably to achieve the required robustness. When compared with the adjoining areas without watermarking, the attributes of the embedded areas might be significantly affected, particularly when the perceptual model isn’t utilised. Secondly, nonconformities in the scale, orientation and position, of the watermarked regions might emerge under attacks. Consequently, almost all techniques have to incorporate different shapes as the invariant region for the watermark detection. False positive detection may also pose a serious problem. Thirdly, there are typically a lot many feature points are found in an image by the frequently employed interest point extraction algorithms. Furthermore, the watermark embedding and detecting algorithms have to choose the same points for further processing and it is a much significant issue. A powerful searching algorithm will have to be applied in order to make these feature-based approaches practicable. Moreover, if a particular transform, such as the DFT or DWT, is employed in the watermarking procedure, it has ought to be computed several times in a single detection. This is due to the fact that several different areas will have to be tested, and thereby leads to increased computation load.

1.1 CHARACTERISTICS OF DIGITAL IMAGE WATERMARKING

Three essential characteristics of digital image watermarking are robustness, capacity and transparency.

1.1.1 ROBUSTNESS

Robustness as defined by Cox et al. (2002) is the "ability to faithfully detect the watermark after common image processing operations"[17]. A watermark inserted in an image may be altered by different digital image processing operations like contrast adjustment, histogram equalization etc. Hence, it is essential a watermarking algorithm be designed such that it is robust against such attacks.

1.1.2 CAPACITY OR DATA PAYLOAD

Capacity or data payload is defined by Cox et al. (2002) as “the greatest amount of information that the inserted watermark is capable of hiding such that the embedded watermark may be extracted convincingly for the purposes of authentication, and copyright safeguards.” Capacity depend on on the size of the original data. However, interleaving as much watermark information as imaginable is a rather difficult task in digital image watermarking. Very often, a pre-requisite for capacity relies on the feasible application for which watermarking is used. For instance, in audio, the capacity would relate to the number of bits inserted every second. For images, the capacity might refer to the number of bits embedded into pixels of the image.

1.1.3. FIDELITY OR TRANSPARENCY

Fidelity as described by Cox et al. (2002) is the “perceptual similarity between the original and the watermarked versions of the original image” [17]. In applications requiring the use of non-perceptible watermarks, the watermark image should be such so as to not affect the cover image after being watermarked. Hence, in such applications, watermarking algorithms must not essentially produce any distortions visible to the naked eye because such distortions lead to a reduction in the commercial value of such images.

The conditions listed above namely robustness, fidelity and capacity are essentially conflicting to each other in nature. In order to increase robustness, one might increase the strength of watermarking resulting into a more perceptible watermark. Hence, one might achieve robustness but that would be at the expense of the fidelity of the image. Similarly, in order to increase capacity, we can reduce the number of samples assigned to each concealed bit but this is followed by a loss in robustness. Additionally, for any digital watermarking algorithm, it’s not possible to meet all three requirements simultaneously rather, a suitable trade-off amongst these characteristics must be achieved.

1.2 CLASSIFICATION OF DIGITAL WATERMARKING

1.2.1 In accordance with its characteristics digital watermarking may be partitioned into robust and fragile watermarking

Fragile watermarking is primarily employed in those algorithms where we need to establish the identity of the sender i.e. to verify the source of the image. Hence, the watermark so implanted should very sensitive against any changes of the signal. By doing this, we can establish whether or not the data has been tampered with in accordance with the state of the extracted watermark. Robust watermarking is primarily employed in those applications whereby the embedded watermark is supposed to resist conventional image processing operations, and lossy compression etc. such that the implanted watermark may not be completely destroyed and can still be detected in order to provide certification of the sender.

1.2.2 In accordance with the implanted media digital watermarking may be categorised into image watermarking, text watermarking, audio watermarking, graphic watermarking and video watermarking

Image watermarking implies addition of the watermark in a still cover image. In text-based watermarking a watermark is added to PDF, or other textual formats in order to avoid changes in text. Audio watermarking intends to insert the watermark in the audio signal. Graphical watermarking intends to implant the watermark in 2-D or 3-D computer-generated graphics in order to indicate the copyright rights of the owner. In case of video watermarking a digital watermark is implanted into the video stream.

1.2.3 In accordance with the detection procedure, digital watermarking may be categorised into blind watermarking and non-blind watermarking

In non-blind watermarking one requires the original watermark in the extraction procedure. It is considered to be more robust as compared to blind watermarking, but its application is limited in literature. On the other hand, blind watermarking does not require the original watermark in the extraction procedure, which has wider applications available in literature but requires stringent extraction algorithms.

1.2.4 In accordance with the purpose of watermarking digital watermarking may be classified as copyright protection watermarking, anonymous mark watermarking, note anti-counterfeiting watermarking, and tampering tip watermarking:

In Copyright protection based watermarking algorithm, the owner intends for others to see the watermark. Anonymous mark watermarking is aimed at hiding important and confidential data of a user and confine illegal users to obtain access to this confidential data. Note anti-counterfeiting watermarking is embedded in currency notes and is detectable after printing, scanning, and other processes. Tampering tip watermarking is intended for protecting the integrity of labels, image content, and opposes the common lossy compression formats.

1.3 APPLICATIONS OF DIGITAL WATERMARKING

As mentioned previously, watermarking has proved to be very useful or rather essential in various areas of interest pertaining to digital images. Various applications involving invisible watermarks have been listed below:

1.3.1 FINGERPRINTING

An owner in order to trace the source of duplicate copies can embed various watermarking keys supplied to respective customers where each and every copy is uniquely identified, as a fingerprint identifies an individual. By means of this technique, the owner can detect those customers who break their licence agreement by illegal duplication of data.

1.3.2 INDEXING

Watermarking may also be employed in order to achieve indexing of data by inserting comments in video contents as well as in news items or movies. This form of indexing can be exploited intensively by search engines. In videos, a specific frame may be displayed by the search engine by putting in appropriate query. Since, the number of images available online grows at a much faster pace as compared with the present-day capability of search engines, it’s quite essential to look for newer means that allows rapid access to online data.

1.3.3 OWNER IDENTIFICATION

In order to prove ownership of a document, an owner may easily recover the watermark he had embedded from the digital content. An inseparable and imperceptible watermark is a good alternative as compared with textual watermark for owner identification.

This application can prove to be very beneficial in settling of copyright disputes in courts

1.3.4 BROADCAST MONITORING

Watermark may also be embedded in the data broadcasted to various locations in order to achieve automated identification of broadcasted programs and a count may be maintained each time a particular watermark is detected so as to assure the advertisers of the correct airtime that they have paid for. Audio watermarking may be employed by FM broadcasters so as to prevent illegal duplication of latest songs being made available by them for the general public.

1.3.5 DATA AUTHENTICATION

Digital watermarking provides a means to authenticate data by ensuring that the data has not been tampered with or not by quantitative measurements. Many organizations in the world devote a considerable amount of time and money on research for newer technologies that facilitate documentation of images. These organizations must also certify the authenticity of the art forms they possess. When these are digitized and published on the internet, various problems emerge. Various images of the same art may be found on the internet having minute differences. Here, the image consisting of the watermark will be regarded as authentic and the rest shall be disposed of. Fragile watermarks are of more importance here since they can be employed to detect tampering of data. If watermark is detection is carried out with sufficient confidence, the data is acceptable, otherwise the data is deemed as corrupted.

1.4 ORGANIZATION OF THESIS

In this thesis work, a feature point extraction based geometrically resilient watermarking algorithm has been presented by employing interest point extraction in order to resist geometrical attacks. This work may be regarded as a further extension of [2] and [16] with the difference that the embedding procedure is facilitated by the interest point extraction. This issue will be discussed in Chapter 3 and then the details of the approach followed will be described in detail, including the signal embedding/detection processes and determination of the invariant areas. Subsequently, the watermarked image was subjected to various attacks in order to test the performance of the algorithm are described in Chapter 4. Further, a detailed performance review has been made with conventional watermarking algorithms such as DCT, DWT, DWT-SVD etc. The results, obtained using MATLAB will be displayed in Chapter 5, followed subsequently by some discussions. The conclusion and future scope of this work shall be presented in Chapter 6.

CHAPTER 2
LITERATURE REVIEW

Image processing operations may be carried out in transform domain or spatial domain. In transform domain, we represent the image in terms of its frequencies; whereas, in the spatial domain we represent it with its pixel values. In simple terms image processing in transform domain implies that the image has been segmented into its constituent frequency bands. In order to translate an image to its frequency domain representation, we can use several reversible transforms like Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), or Discrete Wavelet Transform (DWT). Each of these transforms have their own characteristics and the image is represented in different ways.

A watermark may be embedded within an image by modifying either the pixel values or the coefficients in the transform domain. Simpler watermarks can be embedded relatively easily in the spatial domain by modification of the least significant pixel values but these are very fragile in nature and not robust against image processing attacks; however, much more robust watermarks may be embedded by implanting the watermark in the transform domain by modification of coefficients of transform domain image.

2.1 DCT DOMAIN WATERMARKING

Discreet Cosine Transform based watermarking procedures are much more robust when compared with conventional spatial domain watermarking procedures. Such algorithms are robust against conventional image processing operations like low pass filtering, brightness and contrast adjustment, blurring etc. However, they are rather difficult to implement and are computationally more expensive. At the same time, they are weak against geometrical attacks like rotation, scaling, cropping etc.

Watermarking in the DCT domain may be categorized into Block based DCT watermarking and Global DCT watermarking. One of the first algorithms [18] was presented by Cox et al. (1997) used global DCT method to implant a robust watermark in the perceptually significant portion of the Human Visual System (HVS). Implanting watermark in the perceptually significant portion of the image has its own benefits because most compression methodologies remove the perceptually trivial portion of the image. In spatial domain it denotes the Least Significant Bit (LSB), however in the frequency domain it denotes the high frequency components [19]. The key steps of any block based DCT algorithm are shown in Fig.2.1.

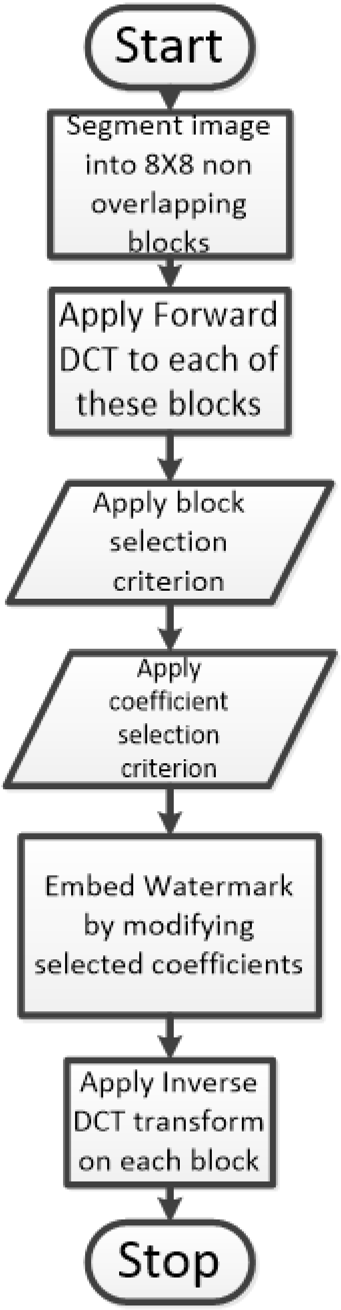

Start
Segment image into 8X8 non overlapping blocks
Apply Forward DCT to each of these blocks
Apply block selection criterion
Apply coefficient selection criterion
Embed Watermark by modifying selected coefficients
Apply Inverse DCT transform on each block
Stop

Figure 2.1: Flowchart for watermarking in the DCT domain

The key difference amongst most of the algorithms in literature is that they vary either in coefficient selection criterion or block selection criterion. Founded on perceptual modelling approach incorporated by the watermarking algorithms they could be categorised as algorithms with:

2.1.1 NO PERCEPTUAL MODELLING

While embedding a watermark, these algorithms do not include any perceptual modelling approach.

2.1.2 IMPLICIT PERCEPTUAL MODELLING

Such algorithms include the transform domain characteristics for perceptual modelling. The coefficient selection conditions are as follows:

1) Those transform coefficients are selected which have larger perceptual capacity. This is so because they allow robust watermarks to be embedded in the image and they result in the least perceptual distortion [42]. These criteria are satisfied by DC components and henceforth they can be selected.

2) Those coefficients are selected which are slightly changed by conventional image processing attacks like rotation, low-pass filtering, scaling, addition of noise etc [42]. Low frequency AC components along with higher magnitude DC components satisfy this criterion and hence they may be selected.

3) Higher frequency components are altered by conventional image processing operations. Therefore, these are not a viable choice for watermarking.

2.1.3 EXPLICIT PERCEPTUAL MODELLING

These algorithms include the Human Visual System (HVS) properties for perceptual modelling. HVS models permit us to increase or decrease the strength of the watermark since it takes into consideration conventional image characteristics like variance, contrast, brightness, etc. Therefore, only those coefficients are adopted which satisfy the HVS criteria.

2.2 DWT DOMAIN WATERMARKING

Wavelet transform has been of great interest over the past few years, particularly in image compression and signal processing in general. In a few applications wavelet-based watermarking algorithms have outperformed DCT based algorithms.

2.2.1 PROPERTIES OF DWT

1) The wavelet transform divides the image into three spatial components, that is horizontal, vertical and diagonal components. Hence wavelet domain mirrors the characteristics of HVS model more precisely [41].

2) Wavelet Transform is computationally efficient and can be realized by using conventional filter convolution.

3) Magnitude of DWT coefficients is larger in the lowest band (LL) at individual level of decomposition and is smaller for other bands (LH, HL, HH) [43].

4) The larger the magnitude of the wavelet coefficient the more vital it is.

5) Watermark recognition at lower resolutions is computationally effective since at every successive resolution level there are limited frequency bands involved.

6) High resolution sub-bands aid to effortlessly locate edge and texture patterns in an image.

2.2.2 ADVANTAGES OF DWT OVER DCT

1) Wavelet transform mimics the Human Visual System model more efficiently than DCT.

2) Wavelet coded image is a multi-resolution portrayal of the image. Therefore, an image can be displayed at different resolutions and may be subsequently processed from lower resolution to higher resolutions.

3) Visual artefacts produced by wavelet coded images are less apparent as compared with DCT because wavelet transform does not divide the image into blocks for processing. At higher compression ratios blocking artefacts are noticed in DCT; but, in wavelet coded images it is much cleaner.

4) DFT and DCT are frame transforms, and therefore any variation in transform coefficients affects the image globally except if DCT has been implemented using a block-based approach. Nevertheless, DWT has spatial frequency locality, which means that the variations introduced will reflect in the image locally [43]. Therefore, a wavelet transform offers both spatial and frequency description of an image.

2.2.3 DISADVANTAGES OF DWT OVER DCT

DWT involves more computational complexity as compared to that of DCT [44]. As Feig (1990) pointed out that it only took 54 complex multiplications to calculate DCT for a block of size 8x8, while in wavelet transform the calculation depends on the span of the filter being used, which is at the very least 1 complex multiplication per coefficient [45].

2.2.4 DWT WATERMARKING

Discreet Wavelet Transform (DWT) based watermarking algorithms follow similar rules as those followed by DCT based algorithms, that is the basic idea is the same; however, the process required to convert the image into its frequency domain representation varies considerably and thereby the subsequent coefficients are different. Wavelet transform employ wavelet filters to convert the image. Various filters are available, though the most commonly utilized filters for watermarking include Haar Wavelet Filter, Daubechies Orthogonal Filters and Daubechies Bi-Orthogonal Filters. Each of these filters divides the image into several frequencies. Singular level decomposition outcomes four frequency representations of the images. These four are referred as the LL, LH, HL, HH sub bands as shown in Fig. 3.

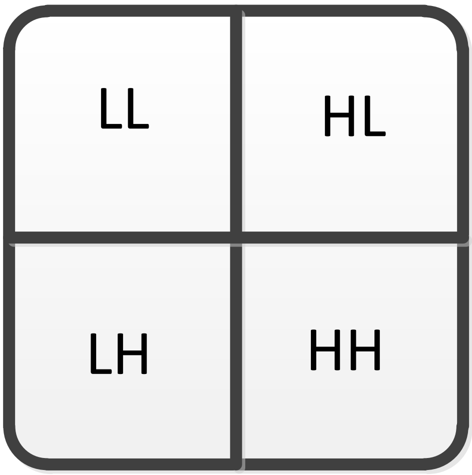


Figure 2.2: Singular level decomposition (SVD) of image

2.2.5 DWT BASED BLIND WATERMARK DETECTION

Lu et al. (1999) presented a novel watermarking procedure known as "Cocktail Watermarking". Using this procedure, they embedded dual watermarks which were complement each other. This scheme was resilient to numerous attacks, and no matter the type of attack performed, one of the watermark was detected. Moreover, they enhanced this technique for image protection and authentication by using the wavelet transform based Just Noticeable Distortion (JND) values. Therefore, this technique achieved content authentication as well as copyright protection simultaneously [20].

Zhu et al.(1999) presented a "multi-resolution watermarking technique" for watermarking images and videos. The watermark is embedded in all the high pass bands in a nested manner at multiple resolutions. This technique does not consider the Human Visual System modelling aspect; however, Kaewkamnerd and Rao [21][22] enhanced this method by taking HVS factor in account.

Pitas and Voyatzis (1999), provided a procedure to implant a binary logo as a watermark, which may be detected using statistical methods. So, in case that the image is degraded significantly and the logo is invisible, correlation may be used for statistical detection. Watermark implanting was based on a chaotic mixing system. The original watermark image was not needed for watermark extraction. Nevertheless, the watermark was implanted in the spatial domain by modifying the pixel values. A similar approach was presented in the wavelet domain [11], whereby the authors proposed a watermarking embedding procedure based on chaotic encryption.

Zhao et al. (2004) presented a dual transform watermark embedding technique for image compression and image authentication. They used the DCT transform for watermark generation and DWT transform for watermark insertion. They exploited the orthogonality of DCT-DWT transform for image watermarking [25].

2.2.6 DWT BASED NON-BLIND WATERMARK DETECTION

This procedure requires the original watermark image for detection of the watermark. Most of the methods found in literature use a smaller image as a watermark and thereby cannot use correlation-based statistical detection for detection of the watermark; Therefore, they rely on the original image for fruitful detection. The dimensions of the watermark image are usually much smaller when compared with the host image.

Lu et al.(2001) presented a robust watermarking procedure based on fusion of the image. They implanted a grayscale and binary watermark which is controlled using the "toral automorphism" described in [23]. Watermark was implanted additively. The originality of this method lies in the usage of a secret image in place of a host image for watermark extraction [40]. Rege and Raval (2003) presented a multiple watermarking procedure. The authors argued that if the watermark had been embedded in lower frequency component of the image it shall be robust against geometric distortions, low pass filtering, and lossy compression. While, if the watermark had been embedded in high frequency component of the image, it shall be robust against brightness and contrast adjustment, cropping, histogram equalization, and gamma correction, and vice-versa. Therefore, in order to attain overall robustness against a variety of attacks, the authors proposed to implant multiple watermarks in low frequency bands and high frequency bands of DWT based watermarking [26].

Hatzinakos and Kundur (1997) presented an image fusion watermarking method. They used the prominent features of the image to embed the watermark. They used a prominence measure to recognize the watermark strength and later implant the watermark additively. Normalized correlation (NC) was used to assess the robustness of the extracted watermark. Later the authors proposed another procedure termed as FuseMark [27], which comprises minimum variance fusion for watermark extraction.

Eskicioglu and Tao (2004) presented an optimum wavelet based watermarking method. They implanted binary watermark (logo) in each of the four bands. However, they implanted the watermarks with different scaling factors in different bands. The scaling factor was larger for the LL sub-band but for the other three bands it was smaller. The superiority of the extracted watermark was determined by resemblance measurement for objective calculation [28].

Eskicioglu and Ganic (2005) enthused by Rege and Raval(2003) proposed a watermarking procedure based on DWT and SVD (Singular Value Decomposition). They claimed that the watermark implanted by Rege and Raval (2003) scheme is perceptible in various portions of the image particularly in the lower frequency regions, thereby reducing the marketable price of the image. Therefore, they generalized their procedure by making use of all four sub-bands and implanting the watermark in the SVD domain. The fundamental method is to divide the image into its four sub-bands and then apply SVD to each band separately. The watermark was essentially implanted by adjusting the particular values from SVD [29]. Fig. 2.2 gives a complete description of the parameters used in these algorithms.

2.3 DFT DOMAIN WATERMARKING

Watermarking in the Discrete Fourier Transform domain has been exploited a lot by researchers due to its robustness against geometrical transformations like cropping, scaling, rotation, translation etc.

2.3.1 CHARACTERISTICS OF DFT

1) Discrete Fourier Transform of a real image is largely complex valued, which is evident in the magnitude and phase representation of the image.

2) DFT watermarking technique is invariant to translation. Spatial alterations in the image disturbs the phase representation of the image but not the magnitude representation [30], even circular shifts in the spatial domain do not affect the magnitude of the Fourier transform coefficients.

3) DFT domain watermarking is also resilient to cropping because cropping gives rise to distortion of the spectrum [30].

4) The stronger components of DFT are the core components which contain lower frequencies.

5) Variation in the spatial domain causes consequential similar variation in the transform domain [22].

6) Scaling in the spatial domain results in inverse scaling in the frequency domain.

2.3.2 COEFFICIENT SELECTION CRITERION

1) Alteration in the lower frequency coefficients is unsuitable since it can result in noticeable artefacts in the spatial domain [21][22]. Therefore, lower frequency coefficients should not be altered.

2) Alteration in the higher frequency coefficients too is unsuitable since they may be eliminated by JPEG compression attack [21][22] leading to a decrease in robustness.

3) Therefore, the most suitable location to implant the watermark is the mid-band frequencies [21][22].

2.3.3 ADVANTAGES OF DFT OVER DWT AND DCT

Watermarking in the DFT domain is invariant to geometrical transformations such as translation, scaling and rotation. Therefore, it may be used to recover from geometric alterations, while watermarking in the spatial domain, DCT and the DWT based watermarking in the frequency domain are not invariant to geometrical distortions and therefore, it is problematic to overcome geometric alterations. Two types of DFT based watermark implanting techniques are available in literature. One is template-based embedding and the other is direct embedding in which the watermark is directly implanted.

2.3.4 DIRECT EMBEDDING

Ruanaidh et al. (1996) suggested a DFT based watermarking procedure in which watermark is implanted by adjusting the phase statistics in the DFT. It has been proved that watermarking based on phase statistics was more robust against image contrast adjustment operations [32]. Later Pun and Ruanaidh (1998) displayed Fourier Mellin transforms, which may be used for digital watermarking. Fourier Mellin transform is alike to Fourier Transform in log-polar coordinate system for an image. This scheme is robust against geometrical attacks like rotation, scaling, etc [33].

De Rosa et al. (1999) proposed a scheme to embed a watermark by direct modification in the middle frequency bands of the DFT coefficients [34]. Ramkumar et al. (1999) also suggested a DFT based data concealing scheme, where they modified the magnitude of the DFT coefficients. The suggested scheme was displayed to be resilient to JPEG compression [41].

Lin et al. (2001) presented a geometrically resilient watermarking algorithm in which the watermark is implanted in the magnitude coefficients of the Fourier transform which was sampled by using log-polar mapping. This method is however susceptible to attacks such as cropping and JPEG compression [36].

Pitas and Solachidis (2001) suggested a new watermarking method. They implanted a “circularly symmetric watermark” in the magnitude coefficients of the Fourier transform [31]. Because the watermark was rounded in form with its centre coinciding with the image centre, this technique has been proved robust against geometric transformations. The watermark is positioned around the mid-band frequency region of the DFT magnitude coefficients. Masking of neighbouring pixel variance is utilised to decrease any visible artefacts.

A semi-blind watermarking method was proposed by Eskicioglu and Ganic (2004). They implanted spherical watermarks with one in the higher frequency while the other one in the lower frequency. Their work was encouraged by [35]. They followed the same reasoning as that recommended by inserting watermarks in the lower frequency components, which was proved to be robust against one set of attacks, while inserting watermarks in the higher frequency components was robust to another set of attacks.

2.3.6 TEMPLATE BASED EMBEDDING

Pun and Pereira (2000) suggested a watermarking algorithm resilient to affine transformations. They familiarized the concept of the template. A template is inserted in the DFT domain to evaluate the transformation factor. Once the image experiences a transformation this template is examined to resynchronize the image, and then the detector is used to obtain the implanted spread spectrum watermark [30].

2.4 FFT AND DHT DOMAIN WATERMARKING

Pereira et al. (1999) suggested an FFT based watermarking algorithm which is robust against geometrical transformations and JPEG compression attacks. Its embedding algorithm is template based, which is alike to the one reviewed in the previous section. In addition to the template, a logo is entrenched to prove proprietorship. If the image experiences a geometric alteration the template is overturned back to its previous location and then extraction of the watermark is carried out. They make use of the idea of log-polar maps and log-log maps to retrieve the hidden template. This method has proved to be robust against print and scan attack, cropping; however, it is tough to employ [37].

Lim and Falkowski (2000) suggested non-blind watermarking algorithm based on Discrete Hadamard transform and multi-resolution transform [39]. Firstly, the multi-resolution Hadamard transform was employed in order to divide the image into its constituent frequency bands like LL, LH, HL and HH. The lowest frequency band that is LL is then partitioned into 8x8 sub-blocks and 2D Discrete Hadamard transform is utilised. Watermark is implanted in this domain by modifying the phase component of the most substantial image component. The watermark is implanted in the phase component rather than in the magnitude component because phase modulation is evidently more robust to noise as compared with amplitude modulation. The proposed scheme proved to be robust against various attacks like image scaling, JPEG compression, cropping, dithering, and successive watermarking.

In another watermarking procedure suggested by Skodras and Gilani (2001), the watermark was implanted by adjusting the high frequency Hadamard coefficients. The image underwent dual frequency transform, firstly, by Haar Wavelet Transform and subsequently by Hadamard Transform. This gave rise to the multiresolution Hadamard Frequency domain. Most of the energy is concentrated in the upper left corner, therefore, it is nominated to implant watermark information. The writers claimed that the higher frequency bands of the Hadamard transform are robust against noise and can consequently withstand JPEG compression attacks at a lower quality factor [38].

CHAPTER 3

THE WATERMARKING APPROACH

Fig. 3.1 demonstrates the flowchart of the watermarking approach. Firstly, the image is applied with an interest point extraction scheme such as Scale-Invariant Feature Transform (SIFT). The information obtained from these interest points, which include the orientations, locations, and scales shall be used for determination of the invariant areas for watermark implanting. In the pre-processing step, various unsuitable interest points shall be removed and the invariant region for watermark embedding will be established around the remaining points. Essentially, the image is segmented into regions related with numerous interest points and every invariant region shall be enlarged in order to form a larger area. The watermark image comprising the necessary hidden information is entrenched in the DCT coefficients. The rationale behind choosing DCT is that in DCT the block transform can be computed effectively and perceptual models on DCT are readily available. The implanted watermark can then be weighted in accordance with the perceptual model in order to ensure good quality watermarked images. Along with the watermark, the pilot signal will also be embedded into the DCT coefficients. In watermark extraction, the interest points are obtained and then local searching is carried out about the extracted points to ascertain possible areas carrying the watermark. If the detection efficiency isn’t a major concern, the pre-processing step might be omitted.

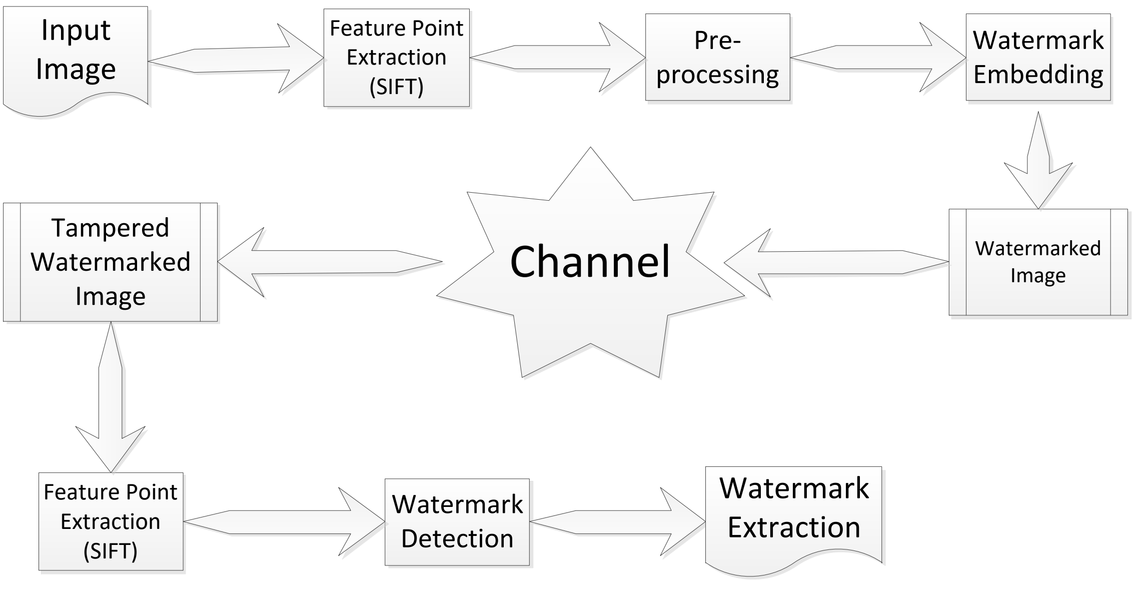


Figure 3.1: The Watermarking approach

Detection of interest points might be influenced by many factors, such as geometrical attacks, JPEG compression, and even the watermark embedding itself. Hence, the locations of the interest points so extracted might not be precise enough for reliable watermark recognition. Local searching will be employed to accommodate probable deviations of extracted interest points. When compared with the pre-existing watermarking methodologies, this algorithm demonstrated better performance against geometrical attacks. Next, each step used in the design will be discussed in detail.

3.1 INVARIANT AREA DETERMINATION

The first step in the algorithm is to ascertain probable areas for watermark embedding/ detection in order to attain synchronization. Hereafter, we use image "Lena" (Fig 2.1(a)) shall be used to elucidate the watermarking procedure. Fig. 2.1(b) demonstrates the interest points so extracted by employing SIFT and they are marked with white dots. The invariant regions, which are nothing but squares with the interest points being located at their centre. The rationale behind selecting the square shape is that watermark shall be embedded after a block transform. Moreover, these squares aid to comprise a broader section of the image for watermark embedding or extraction by extending. The width of a side of this square is established by multiplication of the representative scale of the corresponding SIFT interest point, λ, and a pre-defined positive value, τ , which is selected so as to develop adequate invariant squares with sizes comparable to the one used in the watermarking algorithm. The alignment of the invariant region is also determined by gradient information of the SIFT interest point

SIFT typically produces a large quantity of interest points and various invariant squares overlap each other. Hence, selection of desirable interest points for the signal embedding is essential. Initially, we employ JPEG compression with Quality Factor as 50, followed subsequently by Gaussian filtering, in order to select only those interest points which are common between that of the original image and the attacked one. Thereafter, the representative scales of interest points are inspected. Since the extracted invariant area would be embedded by a two-dimensional image of fixed size, scaling of either the image content or the hidden signal pattern is essential. If the two sizes are considerably dissimilar from one another, the scaling itself shall affect the implanted signal significantly. Thus, if the image size used for watermark embedding or subsequent detection is 32×32, we shall select those invariant regions having their sizes ranging between 30 × 30 to 34 × 34. Ordering of the interest points based on their sizes is also essential. The invariant region with its size closest to 32×32 shall be selected on highest priority basis whilst the selection of invariant areas smaller or larger than this shall be deferred. Such a scheme intends to speed up the detection procedure because the detection algorithm shall follow a similar order. Additionally, the points so selected should be separated from one other by an acceptable distance. If the distance of the interest point in question from its nearest chosen point is smaller than a threshold distance, T_{dist}, then the interest point in question shall be ignored.

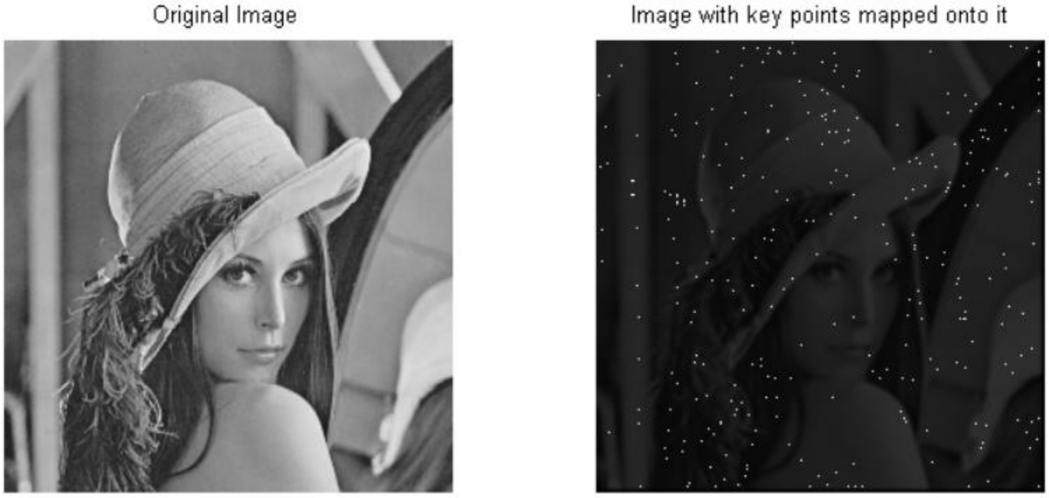


Figure 3.2 : Image on the right displaying the result after Feature Point Extraction using SIFT algorithm of the image on the left

Certain prevailing schemes might employ only these regions for the watermark embedding or subsequent detection but, if the area is small, the payload or detection certainty shall be affected. Conversely, if a larger size is employed, the implanted signal will be influenced by the local distortions easily. In this work, the invariant regions or grids shall be elongated to cover a larger area for signal embedding or subsequent detection. The inclusion or omission of interest points based on the threshold distance as explained before is also dependent on the locations of these white dots. The rationale behind employing such a design is to assist with the detection process.

Next, the pixels covered by this extended grid shall be recognized and implanted with the watermark image. This extension is carried out to guarantee that the embedder and detector follow similar algorithms. Similar procedure is repeated for each and every extended grid and the expansion carried out from one initial grid shall be limited within. A few grids may be inserted at the boundaries provided that the inserted grids shall not overlap others.

Most all these grids shall be implanted with signals, excluding those covering interest points. The reason for excluding the grids consisting the interest points is to prevent the implanted signal from altering the descriptors or even making the interest points imperceptible. Even though the interest point extraction algorithms are becoming increasingly robust, we still can’t rule out the plausibility of affecting the descriptors by the implantation of the watermark. One might repeat the interest point extraction procedure on the selected regions immediately after embedding the watermark to verify whether these interest points chosen were reliable. However, this scheme makes the embedding process much more complicated than it already is. Since the pixels in the grid containing the interest point are unaffected, the first extended grid of each interest point should always be implanted with signals. This scheme will make the detection procedure easier since now, only a single grid related with an examined interest point will be verified first in the detection phase. Hence, along with the constraints of robustness, threshold distance and scale, interest points selection also takes the relative positions of the grids into consideration that is, an interest point in question which might overlap with the extended regions of the already selected points shall be ignored.

3.2 WATERMARK EMBEDDING

In order to implant the watermark, a conventional spread spectrum scheme DCT has been employed. Other watermarking techniques, such as quantization index modulation, may also be employed due to the reason that synchronization issue will be dealt with while embedding or detection of the pilot signal. Now, the watermark image is scaled and implanted into particular DCT coefficients. Watson’s perceptual model [33] is employed to compute the Just Noticeable Difference (JND) for particular DCT coefficients as the weights in accordance with the watermark image. Even though

Watson's model could be used in variable sized blocks, a question might arise that the model only safeguards the inconspicuousness of the noises inside the blocks. Hence, we select a smaller block size to alleviate such a concern. A square grid with side equal to $\lambda \times \tau$ will be standardized by scaling and subsequent rotation into a 32 × 32 pixel block, which shall be further sub-divided into 8 × 8 sub-blocks, upon which 8 × 8 DCT is computed. The watermark image shall be implanted into the low-medium frequency components because these components are much more robust against geometrical attacks and JPEG compression.

Given a grid, **G**, as mentioned earlier, we shall morph it into square blocks having 32×32 size and then the watermark signal shall be inserted into 8 × 8 Discrete Cosine Transform coefficients $c'_{i,j,h}$ given by equation 3.1

$$c'_{i,j,h} = \begin{cases} c_{i,j,h} + w_{i,j,h} \times m_{i,j,h}\,, \ if\ sgn(c'_{i,j,h}) = sgn(c_{i,j,h}) \\ \qquad\qquad 0,\ otherwise \end{cases} \tag{3.1}$$

Wherein $c_{i,j,h}$ denotes the DCT coefficients which were watermarked. In view of the fact that the watermark sequence would be long and the quantity of coefficients considered for watermarking are limited in number, we will be inserting the sequence into numerous grids. An interest point serves as the main point of interest and the watermark shall be inserted into conforming location.

Insertion of pilot signals is similar to watermark embedding. Even though this signal shall be perceived in the spatial domain, we shall be embedding it in frequency domain, *i.e.*, in the DCT coefficients, in order to guarantee that the signal be made indiscernible by means of the visual model. Firstly, a 32 × 32 pilot signal with values ±1 is created and divided further into 8 × 8 sub-blocks. In order to avoid generating periodic patterns, various pilot signals are allocated to extended grids in a Voronoi sub-region of an interest point. The resulting 32×32 pilot signal is denoted as $\mathbf{T}_f$ where f signifies the frequency. The implanted coefficients, $c_{i,j,h}$ can then be calculated by Eqn. (3.1). After implantation of both the signals, Inverse Discrete Cosine Transform (IDCT) has been applied in order to develop the embedded block **G**. The suitable pixels to be selected on the grid in the watermarked image shall be computed by interpolation based on grid **G**. Orientation of the interest point plays an essential role during interpolation. Meticulous execution will certainly lead to enhanced image quality and further improves robustness. The selected pixels on the grid shall be further examined to determine whether or not they're lying within the 32 × 32 block boundary by reverse mapping.

3.3 SIGNAL DETECTION

Given the watermarked image, similar interest point extraction algorithm shall be employed in order to ascertain the feasible regions, from where we shall compute correlation between the tested signal and the signal so retrieved. Hence, we would be able to determine whether a hidden signal is present or not. Now, if a signal is detected, the watermark sequence would be extracted and an expanding process will be carried out in order to determine more such areas for signal detection. After the detection of interest points, invariant grids having size similar to 32×32 grid shall be scrutinized first, as was the case in signal embedding. This approach helps us to fasten the detection procedure significantly. Elimination of interest points based on lower correlation values may not be viable in the signal detection stage in order to avoid wrong omittion of various interest points, especially in the scenario in which the image had been distorted by the channel. However, in this work, the procedure of pilot signal detection is time consuming. Hence, various interest points having enormously large or small characteristic scales may be omitted altogether. A range of detection scale may be chose for instance 22 × 22 to 40 × 40 in order to ensure that the hidden signal will be detected even if the image experiences about 25% up-sampling or 25% down-sampling, while the range of embedding was chosen to be around 28 × 28 to 36 × 36 pixels.

A distorted watermark might be present in a certain grid of a watermarked image and the detector might falsely eliminate this interest point since it is to compute correlation with a known watermark signal. A suitable alternative is to employ an exhaustive search but it would increase the detection time significantly. Hence, the interest points so detected offer a reference for local searching. The scale and the characteristic orientation of an interest point are taken into consideration in order to establish a plausible grid. It must be noted however, that slight distortion of the grid is allowed by various attack in order to preserve the image quality.

CHAPTER 4
CLASSIFICATION OF VARIOUS ATTACKS ON WATERMARKED IMAGE

A fair and automated assessment of watermarking algorithms is essential for chosen application areas. Here, we analyse the performance of the algorithm by subjecting the watermarked image to various attacks and then evaluating the performance by measuring various parameters like normalised correlation, PSNR, and Tamper Assessment Function (TAF) etc. A classification of extensive watermarking attacks for assessing the robustness is presented in [6], where the attacks are characterised into geometrical, protocol, security, and removal attacks. Furthermore, in [46] the imperceptibility of different watermarking methods has been evaluated; this thesis illustrates a few such conventional attacks in the following sections.

4.1 REMOVAL ATTACKS

These attacks are aimed at removal of the watermark from the host image. This categorization includes image de-noising, lossy compression, remodulation, averaging and collusion attacks and quantization.

4.1.1 IMAGE DE-NOISING AND LOSSY COMPRESSION ATTACKS

These types of attacks are comparatively broad in nature and comprises various conventional image processing operators such as image de-noising, lossy compression, and quantization. Image de-noising is a procedure by which we may reconstruct a signal from its noisy counterpart. Various de-noising filters exist in literature such as the mean filter, weighted mean filter, median filter, etc. Compression is another prevalent technique for attacking watermarked images. Two most commonly employed image compression techniques are: lossy compression, such as JPEG compression, and Vector Quantization (VQ) compression. In order to remove a hidden watermark, an attacker may condense the watermarked image with another VQ code and then later decode the VQ indices. The VQ compression technique has proved to be very effective against some pre-existing algorithms.

4.1.2 REMODULATION ATTACKS

Since image de-noising and lossy compression have been so extensively appeared in literature, with various applications including image enhancement and low bit-rate coding respectively; it isn’t surprising that these attacks are also quite famous in the watermarking community. In contrast, the remodulation attacks are a rather fresh addition to the watermarking community. An efficient remodulation attack was first demonstrated by Langeelar et al. in [47]. In this algorithm, the authors tried to forecast the watermark by subtracting the median filtered version of the steganographic image from the host steganographic image. The forecasted watermark this obtained was truncated and later high-pass filtered and then and subtraction is carried out from the steganographic image. Median filtering primarily eliminates noise in the high-frequency band. An equivalent attack based on weighted mean forecasting was demonstrated by Holliman et al. [48]. In this work, the authors were able to successfully remove watermark. Furthermore, in [49], a Wiener attack has been proposed. The proposed attack consists of the following three steps: predictive Weiner filter-based forecasting of the watermark, subsequent subtraction from the source steganographic image along with some strength parameters to the forecasted watermark and lastly adding stationary Gaussian noise.

4.1.3 AVERAGING AND COLLUSION ATTACKS

With regard to the collusion attacks, we have various instances of the same data set. Now, we create an attacked data set is created by acquiring only a little portion of each data set and subsequent reconstruction a novel attacked data set from those selected portions. Statistical averaging illustrates an attack wherein various samples of a pre-conditioned data set are available, computed every time by means of a different secret key. Further, these are averaged in order to compute the attacked data. For instance, while watermarking a video, each frame may be interleaved by means a different watermark computed via a different key. Provided that the data set is adequately large, the inserted watermark can’t be detected anymore since it shall output zero mean on average. Another attack that deteorates the detection and decoding of the watermark is mosaic attack [50].

This attack was devised in the structure of the automatic copyright protection frameworks that inspect the web and downloads images for verifying the existence of the watermarked images on pirated websites. The mosaic attack is not intended to remove or alter the watermark by means of some image processing techniques, instead it hampers watermark detection via deviding the image into smaller fragments.

4.2 GEOMETRICAL ATTACKS

When compared with removal attacks, geometrical attacks don’t aim at removal of the embedded watermark but they aim at alteration of the watermark via spatial transformations. Various attacks included in this category are rotation, scaling, translation, alteration of aspect ratio, etc. Almost all the current watermarking techniques are robust against these categories of attacks due to the presence of various synchronization techniques. This thesis report shall also tend to such attacks on the watermarked images on order to to test the efficiencies of various algorithms.

4.3 CRYPTOGRAPHIC ATTACKS

Cryptographic attacks in watermarked images are quite similar in nature to the those applied in the art of cryptography. These are the forced attacks which aim at unearthing sensitive information by employing exhaustive search algorithms. Since various watermarking techniques employ a secret key, it is essential to use keys with a safe length.

4.4 IMAGE SHIFTING AND LINE DELETION

The attackers might even alter the watermarked image by means of shifting the image either horizontally or vertically, or even eliminate a complete line of pixels, in order to completely or partially deform the embedded watermark. For watermark embedding in DCT domains, image shifting might result in loss of synchronisation while extracting the watermark from the watermarked images.

Even though the above-mentioned categorization makes it quite simple to have an understanding of the various types of attacks that an attacker may employ, still the type of attack that the attacker may employ is quite uncertain since an attacker typically employs a combination of various such attacks at the moment.

CHAPTER 5

PERFORMANCE EVALUATION PARAMETERS

In order to estimate the performance of the algorithm, various geometrical attacks were carried out over the watermarked image in order to deteriorate its performance and compute the robustness of the watermark. The following three parameters were used to quantitatively analyse the performance of the watermarking algorithm:

5.1 PEAK SIGNAL TO NOISE RATIO (PSNR)

Peak signal to noise ratio, usually abbreviated as **PSNR**, is a ratio of the maximum power of a signal and the power of corrupting noise that affects the reliability of its representation. PSNR is typically expressed in logarithmic decibel scale.

$$MSE = \frac{1}{m * n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [X(i,j) - Y(i,j)]^2$$

PSNR in decibels(dB) is represented as

$$PSNR = 10 * \log_{10}\left(\frac{MAX_I^2}{MSE}\right)$$

$$= 20 * \log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right)$$

$$= 20 * \log_{10}(MAX_I) - 10 * \log_{10} MSE$$

Wherein m and n are the dimensions of the images X and Y respectively. MAX_I is the maximum possible pixel value of the image. This depends of the number of bits used to encode a particular pixel for instance if the pixels are being represented using 8 bits per sample, this value is equal to 255. Greater the value of PSNR better is the concealment of the watermark i.e. the watermark is more imperceptible which is so desired.

5.2 NORMALISED CORRELATION (NC)

This is another such performance evaluation criterion wherein we compute the correlation been the embedded watermark and the extracted watermark.

$$NC = \frac{\sum(x - x\prime)(y - y\prime)}{\sqrt{\sum(x - x\prime)^2(y - y\prime)^2}}$$

Where x is the original watermark image and y is the extracted watermark. Its value lies between 0 and 1 where 0 stands for no correlation between the images and 1 stands for complete correlation between images. Hence, a higher value of Normalised Correlation between the original watermark and extracted watermark is desirable in our case.

5.3 TAMPER ASSESSMENT FUNCTION (TAF)

In this performance evaluation criterion, we can determine quantitatively the extent to which the extracted watermark has been tampered with by comparison with the original watermark. Firstly, we convert our greyscale image into a binary image. Then, an exor operation is computed pixel by pixel in both the original watermark image and extracted watermark image and later, all these individual values are summed and subsequently normalised. Hence, a lower the value of the tamper assessment function denotes that the similarity index between both images is very high which further indicates that our watermarked image has not been tampered with by attackers which is desirable. A higher value of Tamper Assessment Function denotes that the image been tampered with which is undesirable.

.

CHAPTER 6

EXPERIMENTAL RESULTS AND DISCUSSIONS

In order to test the performance of the algorithm, we rotated, scaled, and shifted the grid, as displayed in Fig. 6.1. Basically, we resize square grids from the watermarked image with slightly different positions, sizes, and orientations, in order to form normalised 32 × 32 sized grids for easier signal detection. Here, small square grids are employed since they won't be distorted gravely by a geometrical attack and this scheme shall also help us in reducing the detection time significantly.

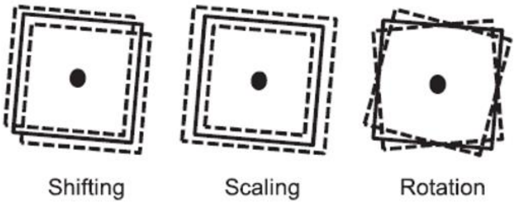


Fig 6.1: Geometrical Transformation of image

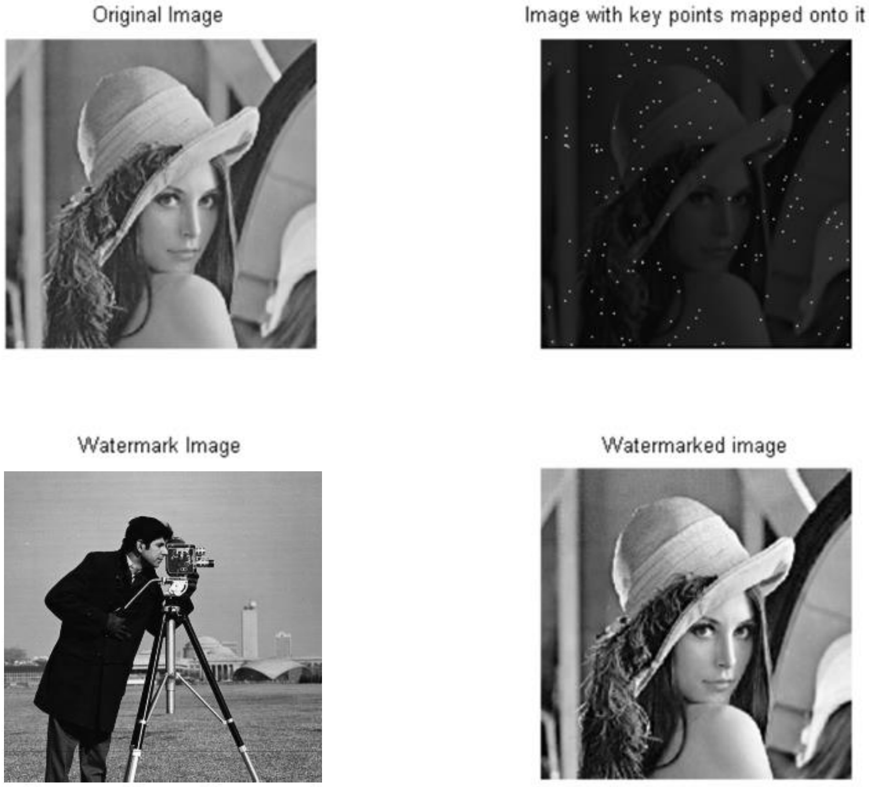


Fig 6.2: Watermark embedding

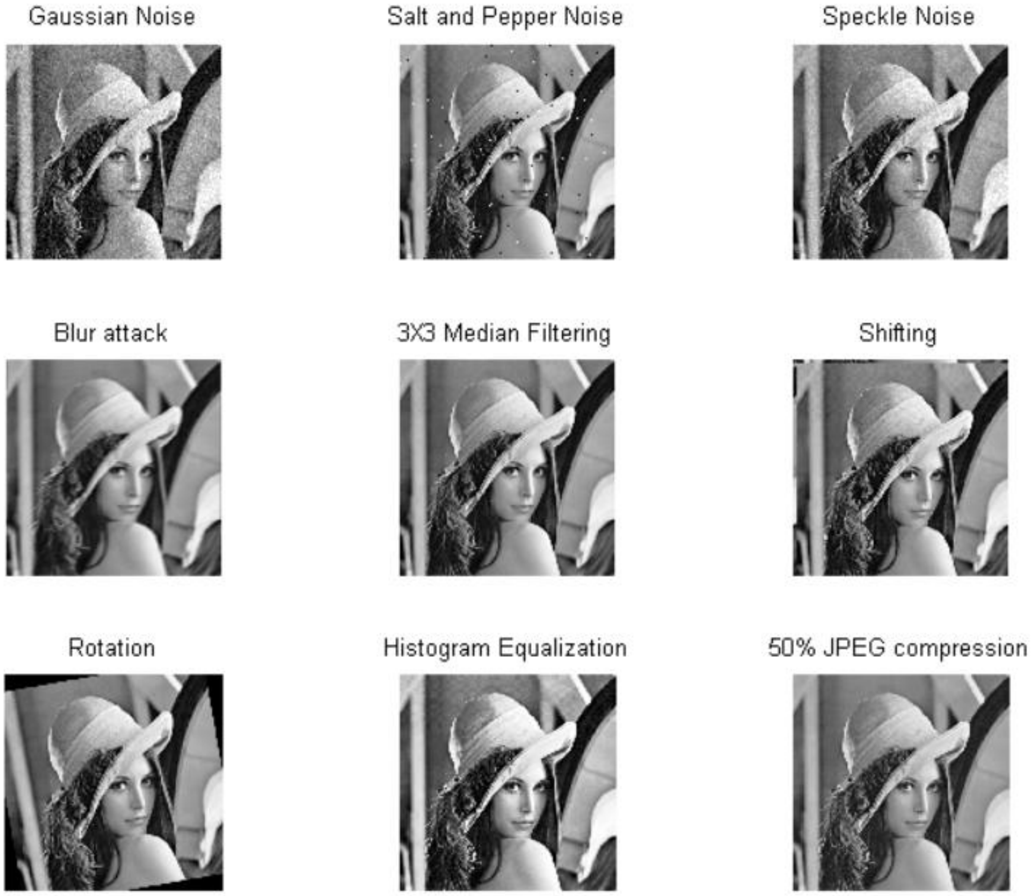


Fig 6.3: Watermarked image corrupted with various types of noise.

The above figure shows various instances of watermarked images in presence of attacks namely

(a) Gaussian Noise with variance 0.005

(b) Salt and Pepper Noise with noise density 0.005

(c) Speckle Noise with noise density 0.005

(d) Blurring attack

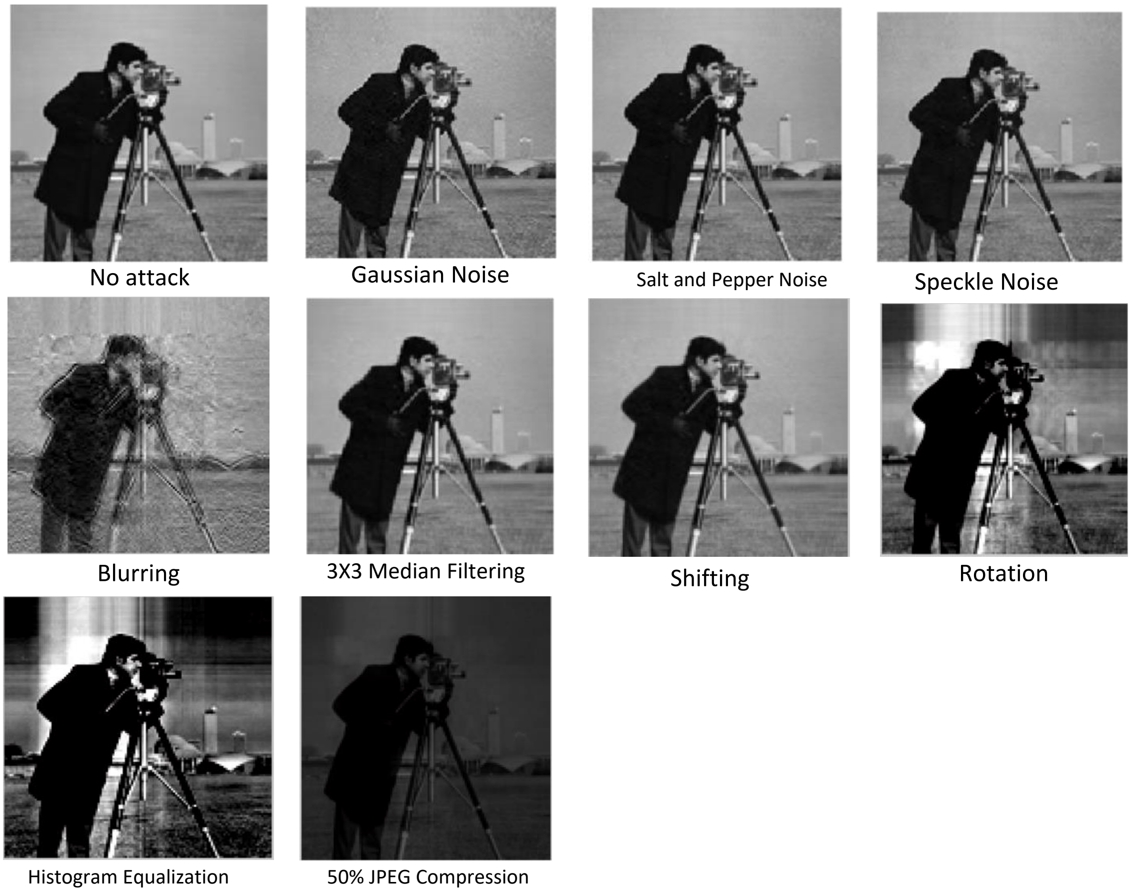
(e) 3X3 window Median Filtering attack

(f) 5% Circular Shifting

(g) Rotation (10^0)

(h) Histogram Equalization

(i) 50% JPEG compression



No attack
Gaussian Noise
Salt and Pepper Noise
Speckle Noise
Blurring
3X3 Median Filtering
Shifting
Rotation
Histogram Equalization
50% JPEG Compression

Figure 6.3: Extracted watermark in presence of various attacks

The above figure shows various instances of extracted watermark images from presence of attacks namely

(a) Gaussian Noise with variance 0.005

(b) Salt and Pepper Noise with noise density 0.005

(c) Speckle Noise with noise density 0.005

(d) Blurring attack

(e) 3X3 window Median Filtering attack

(f) 5% Circular Shifting

(g) Rotation (10^0)

(h) Histogram Equalization

(i) 50% JPEG compression

6.1 PERFORMANCE REVIEW

Type of attack	DCT			DWT			DWT-SVD			Presented work		
	PSNR	CC	TAF	PSNR	CC	TAF	PSNR	CC	TAF	PSNR	CC	TAF
No attack	41.41	1	0	35.81	1	0	37.55	1	0.21	21.05	1	0
Circular Shift	12.53	0.47	26.37	14.75	0.03	47.73	14.75	1	0.21	14.2	0.99	2.74
Rotation (10°)	10.49	0.21	51.56	12.18	0	50.26	12.15	0.63	45.9	12.2	0.88	31.81
Median Filter [3X3]	33.67	0.91	4.59	32.54	0.62	18.7	33.21	0.75	21.13	20.87	0.99	1.85
Blurring	24.08	0.92	4	25.74	0.16	37.31	25.47	0	51.13	20.03	0.89	15.31
Gaussian noise (Var=0.005)	23.17	0.95	2.24	22.54	0.2	43.54	22.87	0.53	27.55	18.55	0.99	3.42
Salt and Pepper(Density=0.005)	27.80	0.97	1.36	27.56	0.36	0.89	27.88	0.77	18.26	19.62	0.99	1.50
Speckle noise(Density=0.005)	28.28	0.99	0.39	27.56	0.36	35.54	27.95	0.8	16.93	19.64	0.99	2.97
Histogram Equalisation	24.47	1	0	18.84	0.27	42.82	19.17	0.84	18.19	19.09	0.84	41.4
JPEG Compression 50%	29.95	0	53.71	30.2	0.27	57.91	33.14	0	58.85	28.97	0.97	58.31

Table 6.1: Quantitative Performance review of various watermarking techniques subjected to various attacks

6.2 CONCLUSION & FUTURE SCOPE

Feature point extraction based digital image watermarking with watermark embedding based on modification of DCT coefficients in the area of interest obtained via SIFT is implemented and is shown to be robust against various geometrical attacks when compared with other conventional algorithms such as DCT, DWT, DWT-SVD. This algorithm is, no doubt, computationally expensive due to the inclusion of extraction of features points using SIFT in the algorithm, but it is also more robust to geometrical attacks.

APPENDIX 1

HUMAN VISUAL SYSTEM (HVS)

In various applications related to image processing operations, the constraints of the human visual system (HVS) can be utilized to improve performance of the algorithm from a visual quality point of view. Such HVS-model based approaches are only slowly replacing classical schemes, in which the quality metric consists of a simple pixel-based difference measure, like the mean squared error (MSE). The improvement achieved by employing an HVS-based approach is quite significant and applies to diverse image processing applications. For instance, quality assessment tools try to predict subjective ratings, image compression schemes reduce the visibility of introduced artifacts, and watermarking schemes hide more robustly information in images.

The HVS model is based on mimicking the behavior of our eyes to various stimulus presented to it. They can be understood as a complicated camera continually in motion, allowing accommodation to different light levels and to objects at various distances. Our eyes have certain optical defects such as chromatic aberration and optical blur, but usually these do not affect the rest of the processing chain. Retina lies at the back of the eyes, wherein a dense layer of interconnected neurons sample and process the visual information. The role of the retina is preponderant, because the processing that the retina performs governs the rest of the visual chain.” The retina encodes visual information based on the image formed on it and then transmits it to the optical nerve. The ratio between the number of receptors i.e. Rods and cones in the retina and the number of fibers in the optical nerve is about a 100:1 which implies that a compression of visual information takes place at this stage. This compression is achieved by a replacement of the photographic image with spatial, temporal and chromatic characteristics such as contours, color and motion. The primary function of the retina is the sampling of the optical signal by photoreceptors. There are two kinds of photoreceptors, rods and cones. Rods are sensitive to low levels of luminosity and saturate in photopic conditions, under which images are usually viewed. In most image processing operations, we exclude the contribution of rods because of the non-existence of rods in the center of the visual field. In accordance to their sensitivity to various wavelengths cones are classified as L-, M- and S-cones according to their sensitivity to long, medium and short wavelengths, respectively. The cones do not provide detailed spectral information, but a weighted summation over the different sensitivity spectra. This means that three values should be sufficient to reproduce human color distinction capabilities, which leads to the description of color by tri-stimulus values.

Human color perception is not directly related to the cone responses, but rather to their differences. These are represented by an achromatic channel and two opponent-color channels, which code red-green and blue-yellow color differences. This coding decreases the redundancy between the signals of the three cone types, because it follows the principal components of natural scenes. This efficient coding takes place in the retina, and Derrington et al. proposed a color space based on the null response of color-opponent retinal neurons which respond to color differences. In image processing this coding is exploited in several color spaces such as YC_BC_R, where Y is the luminance channel and C_B, C_R the color-difference channels.

APPENDIX 2
WATSON'S PERCEPTUAL MODEL

Here, we quickly illustrate the procedure of establishing Just Noticeable Difference (JND) of DCT coefficients in Watson's Perceptual model. The model typically takes two masking effects into consideration: the contrast masking and the luminance masking. The contrast masking designates that the threshold value for a visual model would be lowered by the existence of other patterns whereas the luminance masking indicates to the reliance of the visual threshold value on the average luminance of the local image region. The evaluation of the visual threshold value begins with a perceptual model insinuated by Peterson and Ahumada which was independent of the type of image used. The visual threshold value $v_{i,j}$ related to each DCT coefficient with frequency indices *(i, j)*, $0 \leq i, j < 8$, is computed. $V_{i,j}$ is a function of the global display and perceptual parameters, such as display luminance, display resolution, and the viewing distance. The threshold value after adjusting luminance is computed by

$$a_{i,j,h} = v_{i,j} \left(\frac{c_{0,0,h}}{c'_{0,0}} \right)^{a_T}$$

Where $c_{0,0,h}$ gives the DC value for DCT of the sub-block *h*, $c'_{0,0}$ is the mean value of the DC coefficients for the pixels included in the sub-block, and a_T is the luminance-masking exponent having a conventional value of 0.65. Hence, luminance masking coefficient, $a_{i,j,h}$, is governed by the DC term and the location of the pixel denoted by *(i, j)* in a particular sub-block.

REFERENCES

[1] M. Barni, “Effectiveness of exhaustive search and template matching against watermark desynchronization,” *IEEE Signal Process. Lett.*, vol. 12, no. 2, pp. 158–161, Feb. 2005.

[2] J. Lichtenauer, I. Setyawan, T. Kalker, and R. Lagendijk, “Exhaustive geometrical search and the false positive watermark detection probability,” *Proc. SPIE*, vol. 5020, Security and Watermarking of Multimedia Contents V, pp. 203–214, Jan. 2003.

[3] J. Ruanaidh and T. Pun, “Rotation, scale and translation invariant spread spectrum digital image watermarking,” *Signal Process.*, vol. 66, no. 3, pp. 303–317, May 1998.

[4] C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, M. L. Miller, and Y. M. Lui, “Rotation, scale, and translation resilient watermarking for images,” *IEEE Trans. Image Process.*, vol. 10, no. 5, pp. 767–782, May 2001.

[5] S. Pereira and T. Pun, “Robust template matching for affine resistant image watermarks,” *IEEE Trans. Image Process.*, vol. 9, no. 6, pp. 1123–1129, Jun. 2000.

[6] M. Kutter, “Watermarking resisting to translation, rotation and scaling,” *Proc. SPIE*, vol. 3528, Multimedia Systems and Applications, Boston, MA, USA, pp. 423–431, Nov. 1998.

[7] M. Kutter, S. Bhattacharjee, and T. Ebrahimi, “Towards second generation watermarking schemes,” in *Proc. IEEE ICIP*, vol. 1. Oct. 1999, pp. 320–323.

[8] P. Bas, J. Chassery, and B. Macq, “Geometrically invariant watermarking using feature points,” *IEEE Trans. Image Process.*, vol. 11, no. 9, pp. 1014–1028, Sep. 2002.

[9] J. Seo and C. Yoo, “Image watermarking based on invariant regions of scale-space representation,” *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1537–1549, Apr. 2006.

[10] X. Gao, C. Deng, X. Li, and D. Tao, “Geometric distortion insensitive image watermarking in affine covariant regions,” *IEEE Trans. Syst., Man Cybern. C, Appl. Rev.*, vol. 40, no. 3, pp. 278–286, May 2010.

[11] X. Wang, J. Wu, and P. Niu, “A new digital image watermarking algorithm resilient to desynchronization attacks,” *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 4, pp. 655–663, Dec. 2007.

[12] C. Deng, X. Gao, and D. T. Xuelong Li, “A local Tchebichef moments based robust image watermarking,” *Signal Process.*, vol. 89, no. 8, pp. 1531–1539, Aug. 2009.

[13] D. Zheng, S. Wang, and J. Zhao, “RST invariant image watermarking algorithm with mathematical modeling and analysis of the watermarking processes,” *IEEE Trans. Image Process.*, vol. 18, no. 5, pp. 1055–1068, May 2009.

[14] C. Harris and M. Stephens, “A combined corner and edge detector,” in *Proc. Alvey Vis. Conf.*, vol. 15. 1988, pp. 1–6.

[15] D. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[16] S. Voloshynovskiy, F. Deguillaume, and T. Pun, “Multibit digital watermarking robust against local nonlinear geometrical distortions,” in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2001, pp. 999–1002.

[17] Cox, LI, Miller, ML & Bloom, JA 2002, Digital Watermarking, Morgan Kaufmann Publisher, San Francisco, CA, USA.

[18] I.J Cox, J. Kilian, F.T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia" in IEEE Transactions on Image Processing, vol. 6, no. 12, Dec.1997, pp:1673 -1687

[19] Guan-Ming Su, "An Overview of Transparent and Robust Digital Image Watermarking".

[20] Lu, C-S., Liao, H-Y., M., Huang, S-K., Sze, C-J., "'combined Watermarking for Images Authentication and Protection", in Ist IEEE International Conference on Multimedia and Expo, vol. 3, 30 July-2 Aug. 2000 pp. 1415 - 1418

[21] Kaewkamnerd, N., Rao, K.R., "Multiresolution based image adaptive watermarking scheme", in EUSIPCO, Tampere, Finland, Sept. 2000.

[22] Kaewkamnerd, N., Rao, K.R., "Wavelet based image adaptive watermarking scheme" in IEE Electronics Letters, vol.36, pp.3 12-313, 17 Feb.2000

[23] Voyatzis, G., Pitas, I., "Digital Image Watermarking using Mixing Systems", in Computer Graphics, Elsevier, vol. 22, no. 4,pp. 405-416, August 1998

[24] Xiao, W., Ji, Z., Zhang, J., Wu, W., "A watermarking algorithm based on chaotic encryption", in Proceedings of IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering TENCON, vol. 1, pp. 545-548, 28-31 Oct. 2002

[25] Zhao, Y., Campisi, P., Kundur, D., "Dual Domain Watermarking for Authentication and Compression of Cultural Heritage Images", in IEEE Transactions on linage Processing, vol. 13, no. 3, pp. 430-448, March 2004.

[26] Raval, M.S., Rege, P.P., "Discrete wavelet transforn based multiple watermarking scheme", Conference on Convergent Technologies for Asia-Pacific Region, TENCON 2003, vol. 3, pp. 935 - 938, 15-17 Oct. 2003

[27] Kundur. D., Hatzinakos, D., 'Towards Robust Logo Watermarking using Multiresolution Image Fusion," IEEE Transactions on Multimnedia, vol. 6, no. 1, pp. 185-198, February 2004

[28] Tao, P & Eskicioglu, AM 2004, 'A Robust Multiple Watermarking Scheme in the Discrete Wavelet Transform Domain', in Symposium on Internet Multimedia Management Systems V, Philadelphia, PA.

[29] Ganic, E., Eskicioglu, A. M., "Robust digital watermarking: Robust DWT-SVD domain image waternarking: embedding data in all frequencies", Proceedings of the 2004 multimedia and security workshop on Multimedia and Security, September 2004, pp. 166 - 174.

[30] Pereira, S., Pun, T., "Robust Template Matching for Affine Resistant Image Watermarks," in IEEE Transactions on Inage Processing, vol. 9, no. 6, pp. 1123-1129, June 2000

[31] Solachidis, V & Pitas, I 2001, 'Circularly Symmetric Watermark Embedding in 2-D DFT Domain', in IEEE Transactions on Image Processing, vol. 10, no. 11, pp. 1741-1753.

[32] Ruanaidh, J. J. K. O., Dowling, W. J., Borland, F. M. "Phase watermarking of digital images," in Proc. IEEE Int. Conf: Image Processing, pp. 239-242, Sept. 16-19, 1996.

[33] Ruanaidh, J. J. K. O., Pun, T., "Rotation, scale and translation invariant spread spectrum digital image watermarking," Signal Process, vol. 66, no. 3, pp. 303-317, 1998.

[34] De Rosa, A., Bami, M., Bartolini, F., Cappellini, V., Piva, A. "Optimum Decoding of Non-additive Full Frame DFT Watermarks", in Procedings of the 3rd Workshop of Information Hiding, 1999, pp. 159-171.

[35] Ramkumar, M., Akansu, A.N., Alatan, A.A., "A Robust Data Hiding Scheme For Digital Images Using DFT', in IEEE ICIP, vol 2, pp 211-215, October 99.

[36] Lin, C-Y, Wu, M, Bloom, JA, Cox, U, Miller, ML & Lui, YM 2001, 'Rotation, Scale and Translation Resilient Watermarking for Images', IEEE Transactions on Image Processing, vol. 10, no. 5, pp. 767-782.

[37] Pereira, S, ORuanaidh, JJK, Deguillaume, F, Csurka, G & Pun, T 1999, 'Template Based Recovery of Fourier-Based Watermarks using Log-polar and Log-log Maps', in Proc. IEEE Int. Conf. Multimedia Computing and Systems, vol. 1, 1999, pp. 870--874. Florence, Italy.

[38] Gilani, A.M., Skodras, A.N., "Watermarking by Multi-resolution Hadamard Transform," in Proceedings Electronic Imaging & Visual Arts (EVA 2001), pp. 73-77, Florence, Italy, March 26-30, 2001.

[39] Falkowski, B.J., Lim, L.S., 'image Watermarking Using Hadamard Transforms', in IEE Electronics Letters, United Kingdom, vol. 36, no. 3, pp. 211-213, February 2000.

[40] Lu, C. S., Huang, S.-K., Sze, C.-J., Liao, H.-Y., "A new watermarking technique for multimedia protection," in Multimedia Image and Video Processing, L. Guan, S.-Y. Kung, and J. Larsen, Eds. Boca Raton, FL: CRC, 2001, pp. 507--530.

[41] Choi, Y., Aizawa, K., "Digital Watermarking Technique using Block Correlation of DCT Coefficients" in Electronics and Comununications, Japan, Part 2, vol. 85, no. 9, 2002

[42] Huang, J, Shi, YQ & Shi, Y 2000, 'Embedding Image Watermarks in DC Components', IEEE Transactions on Circuits and System for Video Technology, vol. 10, no. 6, pp. 974-979.

[43] Tao, P., Eskicioglu, A.M., "A Robust Multiple Watermarking Scheme in the Discrete Wavelet Transform Domain", in Symposium on Internet Multimedia Management Systems, Philadelphia, PA. October 25-28, 2004.

[44] Lee, C., Lee, H., "Geometric attack resistant watermarking in wavelet transform domain," in Optics Express vol. 13, no. 4, pp. 1307-1321 2005

[45] Feig, E., "A fast scaled DCT algorithm", in Proc. SPIE Image Processing Algorithms and Techniques, vol. 1224, pp. 2-13, Feb. 1990.

[46] Lee H.Y., Kim H., Lee H.K. Robust image watermarking using local invariant features, Opt. Eng. 45 (3), pp. 1-11, 2006

[47] Langelaar G.C., Lagendijk R.L., Biemond J. Removing spatial spread spectrum watermarks by nonlinear filtering. Proceedings of the European Signal Processing Conference, 1998

[48] Holliman M., Memon N., Yeung M. Watermark estimation through local pixel correlation. IS&T/SPIE Electronic Imaging'99, Session: Security and Watermarking of Multimedia Contents, pp. 134–146, 1999.

[49] Su J., and Girod B.. Power-spectrun condition for energy-efficient watermarking, IEEE ICIP-99, 1999

[50] Petitcolas F.A.P., Anderson R.J., and Kuhn M.G. Information hiding—A survey. Proceedings of the IEEE, special issue on protection of multimedia content, vol. 87(7), pp. 1062-1078, 1999.

[51] Naveen, Chegguju. (2013, October). SIFT (Scale Invariant Feature Transform) Algorithm. Retrieved from https://in.mathworks.com/matlabcentral/fileexchange/43723-sift--scale-invariant-feature-transform--algorithm

[52] Su, Po-Chyi, Yu-Chuan Chang, and Ching-Yu Wu. "Geometrically Resilient Digital Image Watermarking by Using Interest Point Extraction and Extended Pilot Signals", IEEE Transactions on Information Forensics and Security, 2013.