P-3520

A FRAGILE WATERMARKING ALGORITHM BASED ON LOGISTIC SYSTEM AND JPEG COMPRESSION

By

A.KRISHNAVENI

Reg. No. 0920107008

of

KUMARAGURU COLLEGE OF TECHNOLOGY

(An Autonomous Institution affiliated to Anna University of Technology, Coimbatore)

COIMBATORE – 641 049

A PROJECT REPORT

Submitted to the

FACULTY OF ELECTRONICS AND COMMUNICATION ENGINEERING

In partial fulfillment of the requirements

for the award of the degree

of

MASTER OF ENGINEERING

IN

COMMUNICATION SYSTEMS

APRIL 2011

BONAFIDE CERTIFICATE

Certified that this project report entitled "**A FRAGILE WATERMARKING ALGORITHM BASED ON LOGISTIC SYSTEM AND JPEG COMPRESSION**" is the bonafide work of **Ms.A.Krishnaveni** [Reg. no. 0920107008] who carried out the research under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form part of any other project or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

Project Guide

Dr.Rajeswari Mariappan

Head of the Department

Dr. Rajeswari Mariappan

The candidate with university Register no. 0920107008 is examined by us in the project viva-voce examination held on 21.4.11.....

Internal Examiner

External Examiner

ACKNOWLEDGEMENT

First I would like to express my praise and gratitude to the Lord, who has showered his grace and blessing enabling me to complete this project in an excellent manner.

I express my sincere thanks to our beloved Director **Dr.J.Shanmugam**, Kumaraguru College of Technology, for his kind support and for providing necessary facilities to carry out the work.

I express my sincere thanks to our beloved Principal **Dr.S.Ramachandran**, Kumaraguru College of Technology, who encouraged me in each and every steps of the project work.

I would like to thank **Dr.Rajeswari Mariappan** Professor and Head of the department of Electronics and Communication Engineering (ECE), Kumaraguru College of Technology, who rendering us all the time by helps throughout this project.

I would like to express my deep sense of gratitude to our my guide, the ever active **Dr.Rajeswari Mariappan**, Department of Electronics and Communication Engineering, for the valuable suggestions and encouragement which paved way for the successful completion of the project work.

I would also like to express my sincere thanks to **Dr.A.Vasuki**, Professor, Department of ECE for providing me all necessary information and support throughout the project.

In particular, I wish to thank with everlasting gratitude to the project coordinator **Ms.D.Mohanageetha(Ph.D).**, Associate Professor, Department of Electronics and Communication Engineering, for the expert counseling and guidance to make this project to a great deal of success.

I wish to convey my deep sense of gratitude to all the teaching and non-teaching staffs of ECE Department for their help and cooperation.

Finally, I thank my parents and my family members for giving me the moral support and abundant blessings in all of my activities and my dear friends who helped me to endure my difficult times with their unfailing support and warm wishes.

ABSTRACT

To provide authentication and integrity in data transfer, watermarking technique plays a vital role. A fragile watermarking algorithm for image contents integrity verification based on logistic system a Joint Photographic Experts Group is described. Most of them embed watermark data into uncompressed host images. The uncompressed digital image requires high capacity for storage, and high bit rate for communication. In fact, image acquisition equipment often delivers compressed images at its output and the compressed images are more commonly used. Currently, the most common image compression standard is JPEG. From this perspective, the combination of watermarking and JPEG techniques is necessary and significant. A fragile watermarking algorithm for image contents integrity verification based on logistic system and JPEG Joint Photographic Experts Group is proposed. The watermarking process is carried out during JPEG quantization process. Firstly, each 8×8 DCT coefficients block is subdivided into 4 groups based on logistic system; then, in order to obtain the watermarked JPEG image, all of the groups are check coded by the check bits. The watermarked JPEG image remains high quality. The tamper probability is determined by checking the DCT coefficients groups.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

DCT	-------	**Discrete Cosine Transform**
DWT	-------	**Discrete Wavelet Transform**
JPEG	-------	**Joint Photographic Experts Group**
LSB	-------	**Least Significant Bit**
MATLAB	-------	**MATrix LABoratory**
MANET	------	**Mobile Adhoc Network**
PSNR	-------	**Peak Signal to Noise Ratio**

CHAPTER 1
INTRODUCTION

With the rapid revolution in digital multimedia and the ease of generating identical and unauthorized digital data, the requirement to establish reliable methods for copyright protection and authentication has become a matter of concern. The piracy of software, images, video, audio, and text has long been a concern for owners of these digital assets. Data hiding is one efficient solution to face this scenario

1.1 NEED FOR DATA HIDING

The need for data hiding is to provide the carrier document with either of the following

- Covert communication using images (secret message is hidden in a carrier image)
- Ownership of digital images, authentication, copyright
- Data integrity, fraud detection, self-correcting images
- Traitor-tracing (fingerprinting video-tapes)
- Adding captions to images, additional information, such as subtitles, to video, embedding subtitles or audio tracks to video (video-in-video)
- Intelligent browsers, automatic copyright information, viewing a movie in a given rated version
- Copy control (secondary protection for DVD)

1.2 GENERAL SCENARIO OF DATA HIDING

Data hiding is one that embeds data into digital media for the purpose of identification, annotation and copyright. The Figure 1.1 below describes the concept of data hiding.

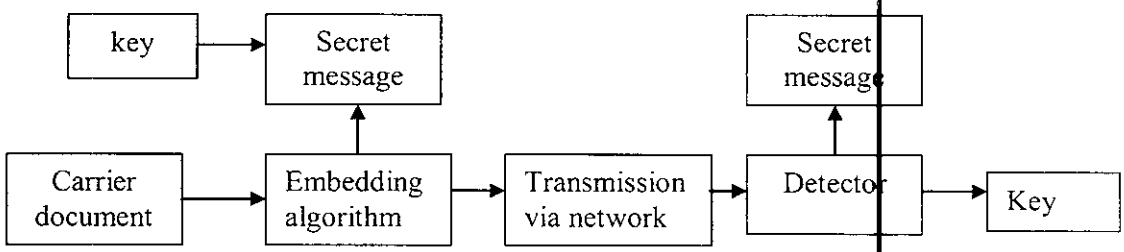

Figure 1.1 Concept of Data Hiding

1.2.1 Disciplines in Data Hiding

Information Hiding is a general term that encompasses many sub-disciplines. Two important sub-disciplines with constitute information hiding are:

- Steganography
    - Hiding: keeping the existence of the information secret
- Watermarking
    - Hiding: making the information imperceptible

1.3 STEGANOGRAPHY

Steganography is the process of embedding information (plain text) within other seemingly harmless information (cover text) in such a way that no one but the intended recipient would try to retrieve it. A comparison between steganography with cryptography is shown below

Table 1.1 Comparison between steganography and cryptography

STEGANOGRAPHY	CRYPTOGRAPHY
Hides the data, without altering	alter, without hiding
Confuses the fact of communication, not the data	Confuses the data, not the fact of communication
preventative - deters attacks	curative - defends attacks

1.3.1 Steganography Techniques

The various techniques in steganography are

1. Substitution methods
    - Bit plane methods
2. Signal Processing methods
    - Transform methods
    - Spread spectrum techniques
3. Coding methods
    - Quantizing, dithering
    - Error correcting codes

1.4 DISADVANTAGES OF STEGANOGRAPHY

- Message is hard to recover if image is subject to attack such as translation and rotation, when the technique is Least Significant Bit encoding
- Significant damage occurs to picture appearance and the message difficult to recover, in Low Frequency Encoding
- The technique is Mid Frequency Encoding is relatively easy to detect.
- Considering High Frequency Domain Encoding, Image gets distorted and message is easily lost if picture subject to compression such as JPEG

1.5 WATERMARKING Vs STEGANOGRAPHY

Watermarking is closely related to steganography but, they differ in some facts. They are

- In watermarking the message is related to the cover
- Steganography typically relates to covert point-to-point communication between two parties .Therefore, steganography requires only limited robustness
- Watermarking is often used whenever the cover is available to parties who know the existence of the hidden data and may have an interest in removing it
- Therefore, watermarking has the additional notion resilience against attempts to remove the hidden data.

1.6 ORGANIZATION OF THE REPORT

The organization of the report is as follows

- **Chapter 2** discusses about the watermarking
- **Chapter 3** deals about the proposed methodology
- **Chapter 4** explains about the results and discussion
- **Chapter 5** discusses the conclusion

CHAPTER 2
WATERMARKING

2.1 INTRODUCTION

Watermarking is an increasingly important technology for protecting intellectual property or ensuring integrity of many forms of digital data. Watermarking means that a digital signature (called a watermark) is added to the protected data – e.g. image - and this signature proves ownership or certifies authenticity of the artifact. Imperceptible watermarks have the advantage that attackers may not even know their presence that makes repudiation very hard. . The significance of watermarking has been rapidly increasing, driven by the dominance of digital media in technology and commerce.

A watermark is a recognizable image or pattern in paper that appears as various shades of lightness/darkness when viewed by transmitted light (or when viewed by reflected light, atop a dark background), caused by thickness or density variations.

Watermarks vary greatly in their visibility; while some are obvious on casual inspection, others require some study to pick out. Various aids have been developed, such as watermark fluid that wets the paper without damaging it. Watermarks are often used as security features of banknotes, passports, postage, and other documents to prevent counterfeiting.

A watermark stored in a data file refers to a method for ensuring data integrity which combines aspects of data hashing and digital watermarking. Both are useful for tamper detection, though each has its own advantages and disadvantages.

2.2 ATTRIBUTES OF WATERMARKING

The main attributes of watermarking algorithm are

- Imperceptibility
- Robustness
- Security
- Capacity

Imperceptibility

A watermark can be embedded into an image as either visible or invisible. The visible watermark is perceptible and is just like a noise. It mostly can be removed by noise removing process. In order to decrease the risk of tracking, cracking, most of the proposed watermarking methods are invisible. On the other hand, the quality of the image is also important. If the watermark embedding process seriously affects the quality of the watermarked image, the watermarked image will draw the attention of attackers or even lose it value. Therefore, the quality between the original image and the watermarked image should not be highly degraded. This property is called Imperceptibility.

Robustness

A watermarking scheme should resist destruction from standard image processing and malicious attacks. The watermarked image may be incurred in several intentional or unintentional attacks to try to remove the embedded watermark. A robust watermarking scheme has to ensure the retrieved watermark is recognized when the image quality does not get harmed. Robustness is most important property and is a requirement of watermarking. The watermark should be able to survive any reasonable processing inflicted on the original image

Security

The watermarked image should not reveal any clues of the presence of the watermark, with respect to un-authorised detection, or (statistical) undetectability or unsuspicious (not the same as Imperceptibility)

Capacity

The capable size of embedding information is defined as the embedding capacity. Due to the reversible watermarking schemes, having to embed the recovery information and the watermark information into the original image, the required embedding capacity of the reversible watermarking schemes is much more than the conventional watermarking schemes. The embedding capacity should be high, but it should not affect the accuracy of the recovered image.

2.3 Digital Watermarking

A digital watermark is a signal or code that is hidden (typically is imperceptible to the user) in a digital signal (such as in the digital audio or a digital image portion) that contains identifying information. The following Figure 2.1 portrays the types of digital watermarking.

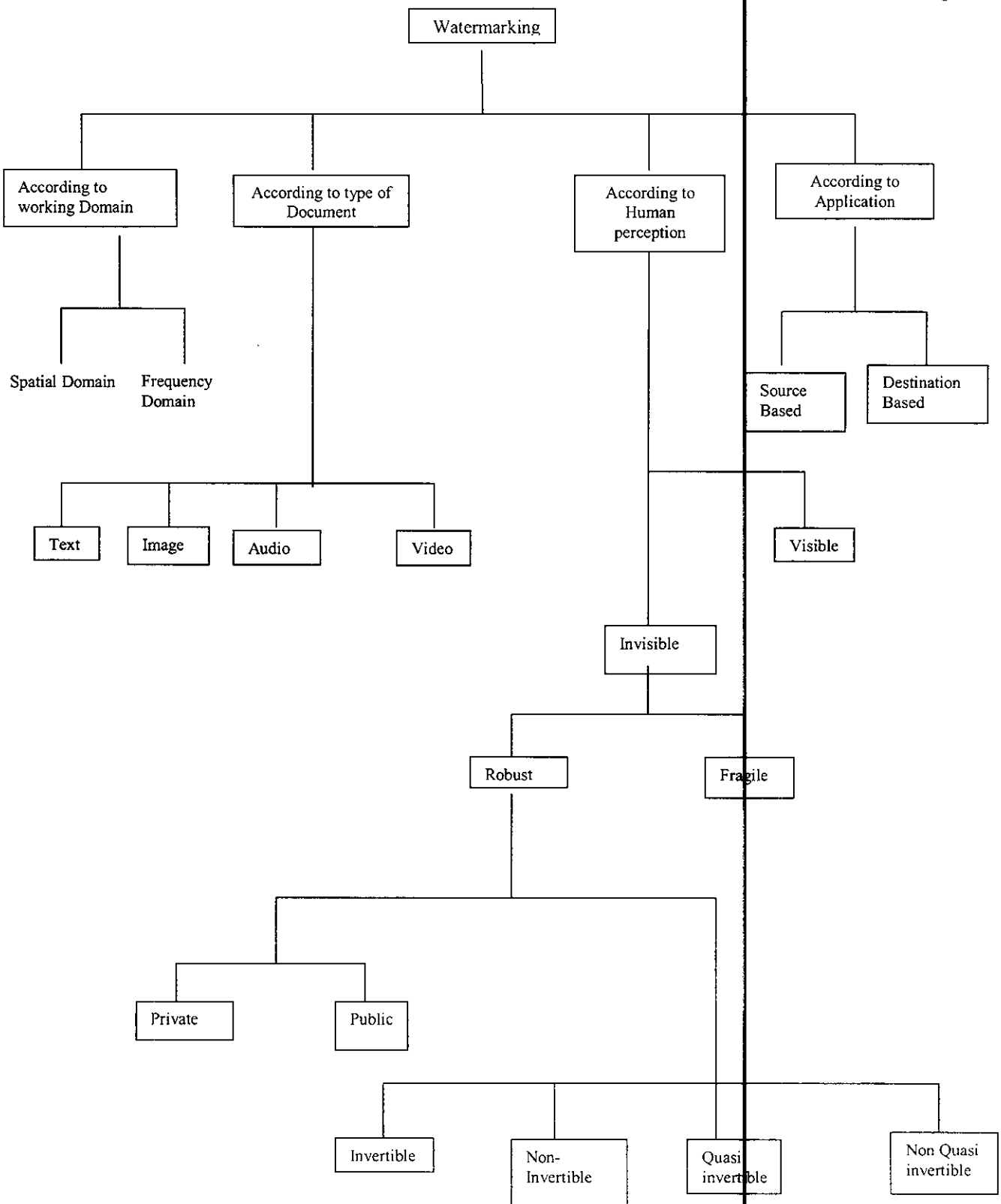

Figure 2.1 Types of watermarks

Ideally a digital watermark would not be destroyed (that is, the signal altered so that the hidden information could no longer be determined) by any imperceptible processing of the overall signal. For example, a digital watermark should not be distorted or lost when the signal is passed through a conversion or compression process.

2.4 Life cycle phases of Digital watermarking

Figure 2.2 shows the lifecycle phases of digital watermarking. The information to be embedded in a signal is called a digital watermark, although in some contexts the phrase digital watermark means the difference between the watermarked signal and the cover signal. The signal where the watermark is to be embedded is called the host signal. A watermarking system is usually divided into three distinct steps, embedding, attack, and detection. In embedding, an algorithm accepts the host and the data to be embedded, and produces a watermarked signal.

Then the watermarked digital signal is transmitted or stored, usually transmitted to another person. If this person makes a modification, this is called an attack. While the modification may not be malicious, the term attack arises from copyright protection application, where pirates attempt to remove the digital watermark through modification. There are many possible modifications, for example, lossy compression of the data (in which resolution is diminished), cropping an image or video or intentionally adding noise.

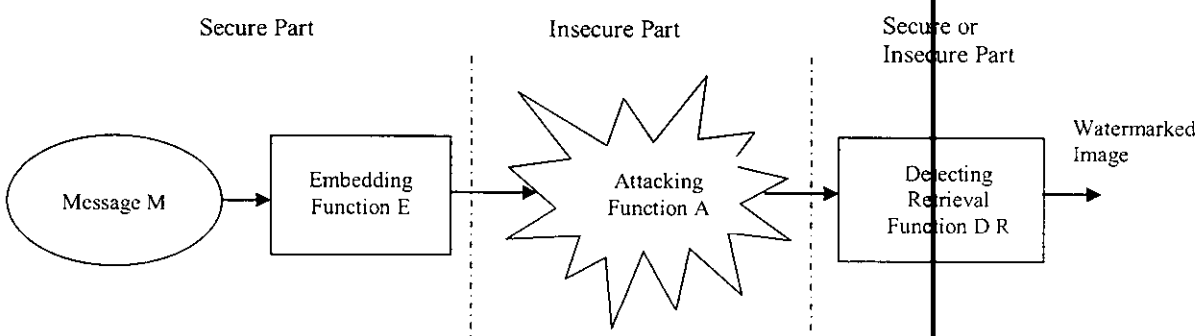

Figure 2.2 Lifecycle phases of digital watermarking

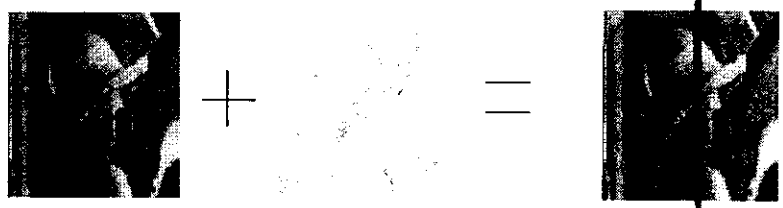

Figure 2.3 Illustration of spatial domain

The drawbacks of spatial domain methods are that in general they are not robust to common geometric distortions and have a low resistance to JPEG compression; moreover, since the watermark casting algorithm can embed only few bits in the image to respect the requirement of unobtrusiveness, they seem to offer a low bit capacity.

2.6.2 Transform domain

In transform domain watermarking techniques, a digital image is processed by means of a specific transform. This process can be applied to the whole image as a single block or to smaller blocks, most commonly of size 8x8 or 16x16. In wavelet based methods, the space of embedding consists of the bands in which the image is decomposed. The advantage of using DCT domain includes the fact that frequency transform is widely used in image and video compression and DCT coefficients affected by compression are well known.

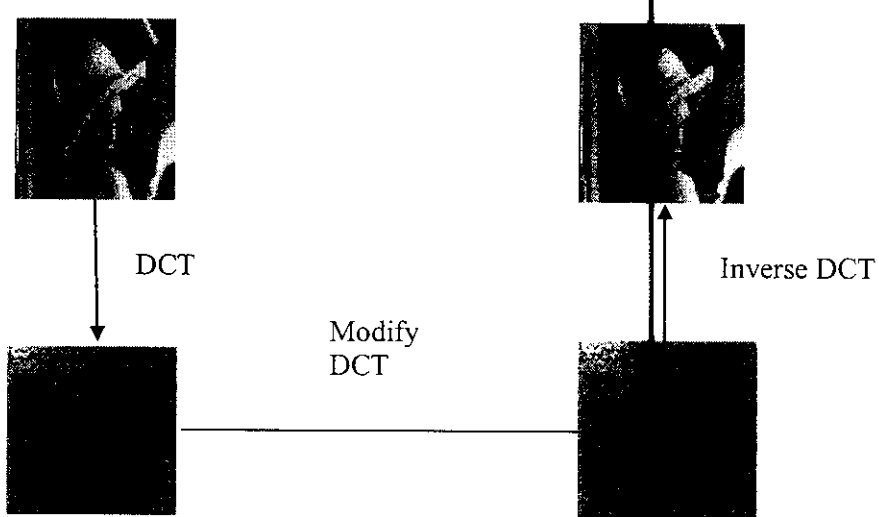

Figure 2.4 Illustration of transform domain

2.7 APPLICATIONS OF DIGITAL WATERMARKING

2.7.1 DIGITAL WATERMARKING TECHNOLOGY FOR RIGHTS MANAGEMENT

One of the traditional applications of the watermark is copyright protection. The primary reason for using watermarks is to identify the owner of the content by an invisible hidden "mark" that is imprinted into the image. In many cases, the watermark is used in addition to the content encryption, where the encryption provides the secure distribution method from the content owners to the receivers, and the watermark offers the content owners the opportunity to trace the contents and detect the unauthorized use or duplications. Without watermarking, there is no way to extend the control of the content owner once the content leaves the protected digital domain and is released to the user. Digital watermark is used to extend the protection and provide the opportunities for the content owners to protect the rights and properties of the electronic distributed contents. The signature of the owner, and usage limitation can be imprinted into the contents, and stay with the contents as far as it travels. This mechanism extends the opportunity of protecting the contents after the release of the contents to the open environment. The major technical requirements for this application are as follows

- The watermark does not incur visible (or audible) artifacts to the ordinary users.
- The watermark is independent of the data format.
- The information carried by the watermark is robust to content manipulations, compression, and so on.
- The watermark can be detected without the unmarked original content.
- The watermark can be identified by some kind of "keys" that are used to identify large number of individual contents uniquely.

The contents may be changed to the other formats, edited or trimmed by the users or compressed for the storage and transmission, and it is desirable to be able to detect the watermark from those processed contents. Usually, the watermark signal embedded into the content does not disappear after the editing of the content, but becomes more and more difficult to detect while the content is distorted. In general, higher robustness can be achieved by increasing the strength of the watermark signal, thus improving the detection capability. In other

words, the robustness of the watermark is a trade-off between the amount of watermark signal that applies to the content and the overhead to the detection. Currently, several commercial products and services using watermarking technology are available. They include applications for watermark embedding/detection and services to search the Internet for the contents with certain designated watermarks. These applications are mainly taking place between the large content owners (e.g. electronic publishers/distributors), and their customers (e.g. the content creators).Because the usage is limited within relatively smaller groups, each group tends to use their own proprietary watermark rather than a common one. Among these groups, the standardization is not an urgent issue until their markets shift to public domain consumers.

2.7.2. DIGITAL WATERMARKING TECHNOLOGY FOR AUTHENTICATION AND TAMPER PROOFING

Another application of digital watermark is contents authentication and tamper proofing. The objective is not to protect the contents from being copied or stolen, but is to provide a method to authenticate the image and assure the integrity of the image. Since low-end digital camera arrived to the consumer market, it rapidly expanded to a number of industrial applications as well, because the use of a digital image is far more cost effective and can also save time and cost for the Developing/Printing/Exposing (DPE) compared to the traditional chemical photos. However, there are some critical issues for applications, where the photos are used as evidence or the material for some kind of business judgment. For instance, automobile insurance companies sometimes use photos of the damaged car sent by the repair shop to estimate the repair cost. A shift to digital photos will save a great amount of time and money for these kinds of processes. However, the digital photos might be altered to exaggerate damage, or even made up from nothing, since the modification of the digital image is getting much easier with some advanced photo-retouching tools be available. This could result in large amounts of extra payment for the insurance company, or more seriously, undermine the credibility of the insurance company itself. A type of digital watermark, called tamper-detect watermark, might resolve this problem, and provide a secure environment for the evidence photos. The way to realize this feature is to embed a layer of the authentication into the subject digital image using a digital watermark. This additional layer of watermark is used as a "sensor" to detect the

alteration. Our recent implementation can even detect the location of the alteration from the altered image itself. The technical requirements for this application are as follows

- ❖ Invisible to the ordinary users
- ❖ Applicable to compressed image format (most digital cameras use JPEG compatible format)
- ❖ Sensitive to content manipulations, compression, and so on

2.7.3. WATERMARKING TECHNOLOGY FOR DVD PLAYBACK AND RECORD CONTROL

In most of the cases, those applications are targeting at a closed environment or exist between limited numbers of member, e.g., between image libraries and content creators, insurance companies and repair shops, and so on. In this section, the focus is on the watermark application that has much more public impact, namely DVD Copy Control.

Watermarking technology can be viewed as a way to provide a secure data channel along with the contents without modifying the installed-base Consumer Electronics (CE) devices. The embedded watermark is transparently passing through the conventional data path, and will only be detected at the digital recorders. When the watermark detection is mandated in these recorders, this watermark can be used to trigger the copy protection mechanism implemented in it. The watermarking data embedded into the video is difficult to remove without damaging the quality of the content because it is carefully "woven" into the visible part of the video data. In this application, the data called Copy Control Information (CCI) is embedded into the video data to indicate that the status of the contents is "Never Copy", "One Copy Allowed" or "Copy Freely". Recording devices will be mandated to facilitate a "watermark detector" to detect the embedded CCI from the incoming and outgoing video data, and responding properly to the recording/playback rules that are defined. The major advantage of the watermark technology is that CCI can be transmitted over the analog video channels. Even advanced digital encryption schemes cannot extend its protection over the analog video channel, but digital watermark could. The embedded CCI will survive even if the video content is transmitted through the analog video channel, recorded to video cassette, and re-digitization. As far as the digital recorders are facilitated to detect the watermark from the video, the copy protection mechanism can be

extended over all the devices. From the implementation viewpoint, because the watermark is completely transparent to the existing system or devices, it does not require that any of the install-base devices to be modified or be made obsolete. The video contents with watermarks looks and works just the same as the contents without watermarks for the devices and channels that have nothing to do with the embedded CCI, thus can be treated transparently by current and future video transmission infrastructures and devices. Although the primary focus of this system is DVD video, the same watermark can also be applied to other forms of video contents, such as videocassettes, laser discs or broadcasted contents. The major functional requirements for this application are

- High robustness
- High image quality
- Low false positive ratio
- Low detection cost
- Real-time embedding/detection capability

The embedded data needs to survive various kinds of video processing, and be detected on the fly with low-cost detection logic to be implemented in consumer electronics devices.

2.8 WATERMARKING ATTACKS

The various attacks in watermarking are

- Robustness attacks: Intended to remove the watermark. JPEG compression, filtering, cropping, histogram equalization additive noise etc.
- Presentation Attacks: Watermark detection failure. Geometric transformation, rotation, scaling, translation, change aspect ratio, line/frame dropping, affine transformation etc.
- Counterfeiting attacks: Render the original image useless, generate fake original, dead lock problem.
- Court of law attacks: take advantage of legal issues.
- Cropping. This is a very common attack since in many cases the attacker is interested in a small portion of the watermarked object, such as parts of a certain picture or frames of a video sequence.

- ❖ Filtering. Low-pass filtering, for instance, does not introduce considerable degradation in watermarked images or audio, but can dramatically affect the performance, since spread-spectrum-like watermarks have non negligible high-frequency spectral contents.

2.9 FRAGILE WATERMARKING

The watermarks of the robust watermarking schemes for copyright protection are expected to survive different types of manipulations to some extent provided that the manipulated media are still valuable in terms of commercial importance or significant in terms of visual quality. Unlike robust schemes, the schemes for the purposes of authentication and content integrity verification are supposed to be fragile, i.e., we expect the watermark to be destroyed when attacks are mounted on its host media so that alarms can be raised when wrong watermark is extracted. Therefore, the emphasis of the fragile watermarking schemes is focused on the sensitivity to attacks or even incidental manipulations in some cases.

To be considered effective, a fragile watermarking scheme must meet the common requirements such as localizing tampering, detecting geometric transformations (e.g., cropping and scaling), signaling removal of original objects, addition of foreign objects, and alerting other image processing operations (e.g., low-pass filtering). In addition, it is more pragmatic to authenticate the media without referring to the original unwatermarked version. This feature is commonly referred to as blind detection.

2.10 LITERATURE SURVEY

The feasibility of coding an "undetectable" digital water mark on a standard 512×512 intensity image with an 8 bit gray scale is discussed in [1] . The watermark is capable of carrying such information as authentication or authorization codes, or a legend essential for image interpretation. This capability is envisaged to application in image tagging, copyright enforcement, counterfeit protection, and controlled access. Two methods of implementation are discussed. The first is based on bit plane manipulation of the LSB, which offers easy and rapid decoding. The second method utilizes linear addition of the water mark to the image data, and is more difficult to decode, offering inherent security. This linearity property also allows some image processing, such as averaging, to take place on the image, without corrupting the water

mark beyond recovery. Either method is potentially compatible with JPEG and MPEG processing.

The drawbacks of the current authentication watermarking schemes for JPEG images, which are inferior localization and the security flaws, are firstly analyzed in this paper. Then, two counterfeiting attacks are conducted on them. To overcome these drawbacks, a new digital authentication watermarking scheme for JPEG images with superior localization and security is proposed. Moreover, the probabilities of tamper detection and false detection are deduced under region tampering and collage attack separately. For each image block, the proposed scheme keeps four middle frequency points fixed to embed the watermark, and utilizes the rest of the DCT coefficients to generate 4 bits of watermark information. During the embedding process, each watermark bit is embedded in another image block that is selected by its corresponding secret key. Since four blocks are randomly selected for the watermark embedding of each block, the non-deterministic dependence among the image blocks is established so as to resist collage attack completely. At the receiver, according to judging of the extracted 4 bits of watermark information and the corresponding 9-neighbourhood system, the proposed scheme could discriminate whether the image block is tampered or not. Owing to the diminishing of false detection and the holding of tamper detection, we improve the accuracy of localization in the authentication process. Theoretic analysis and simulation results have proved that the proposed algorithm not only has superior localization, but also enhances the systematic security obviously [2].

A public key watermarking algorithm for image integrity verification is presented in [3]. This watermark is capable of detecting any change made to an image, including changes in pixel values and image size. This watermark is important for several imaging applications, including trusted camera, legal usage of images, medical archiving of images, news reporting, commercial image transaction, and others. In each of these applications, it is important to verify that the image has not been manipulated and that the image was originated by either a specific camera or a specific user. The verification (the watermark extraction) procedure uses a public key as in public key cryptography, and hence it can be performed by any person without the secure

exchange of a secret key. This is very important in many applications (e.g., trusted camera, new reporting) where the exchange of a secret key is either not possible or undesirable

A novel blind watermarking scheme based on the Back-Propagation Neural Networks (BPNN) for image is presented [4]. First, the convolutional codes encoding is used to refine the watermark for increasing robustness of the scheme. BPNN is developed to memorize the relationships between the wavelet selected samples and a processed chaotic sequence. With wavelet domain of original image being divided into watermarking blocks, then several different BPNN models of selected watermarking blocks are trained simultaneously to form certain relationships, which are employed for embedding the coded watermark bit stream. Compared with conventional watermarking, the proposed scheme based on the trained BPNN models modifies only a small amount of image data such that the distortion on original image is imperceptible.

In [5], a spread-spectrum-like discrete cosine transform domain (DCT domain) watermarking technique for copyright protection of still digital images is analyzed. The DCT is applied in blocks of 8 × 8 pixels as in the JPEG algorithm. The watermark can encode information to track illegal misuses. For flexibility purposes, the original image is not necessary during the ownership verification process, so it must be modeled by noise. Two tests are involved in the ownership verification stage: watermark decoding, in which the message carried by the watermark is extracted, and watermark detection, which decides whether a given image contains a watermark generated with a certain key. A generalized Gaussian distribution to statistically model the DCT coefficients of the original image is applied. The resulting detector structures lead to considerable improvements in performance with respect to the correlation receiver, which has been widely considered in the literature and makes use of the Gaussian noise assumption.

In mobile adhoc networks (MANET) , specific Intrusion Detection systems are needed to safeguard them since traditional intrusion prevention techniques are not sufficient in the protection of MANET. An intrusion detection engine based on neural networks combined with a protection method,which is based on water marking techniques.The advantagesof information visualization and machine learning techniques in order to achieve intrusion detection. The maps

produced by the application of the intelligent techniques using a novel combined watermarking embedded method. The performance of the proposed model is evaluated under different traffic conditions, mobility patterns and visualization metrics, showing its high efficiency[6].

CHAPTER 3

PROPOSED METHODOLOGY

The proposed methodology involves the combination of JPEG compression and watermarking technique. 8×8 block DCT is applied to the input image, every 8×8 DCT coefficients block is subdivided into 4 groups. Then all of the groups are checked by the check bits generated based on logistic system. During the tamper detection procedure, the tamper can be detected by blocks.

3.1 JPEG COMPRESSION

JPEG standard is based on the Discrete Cosine Transform and is adequate for most compression applications and it is most commonly used. For gray scale image, the JPEG compression baseline system is mainly performed in sequential steps, as shown in Figure 3.1.

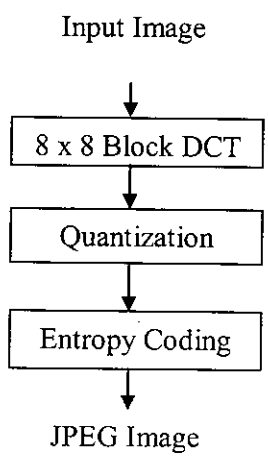

Fig 3.1 Steps of JPEG

In the baseline system, often called the sequential baseline system, the input and output data precision is limited to 8 bits, where as the quantized DCT values are restricted to 11bits.

The compression itself is performed in three sequential steps:

- DCT Computation
- Quantization
- Variable Length Code assignment

3.1.1 DCT Computation

The JPEG image format uses a discrete cosine transform to transform successive 8 × 8 pixel blocks of the image into 64 DCT coefficients each. The DCT coefficients $F(u, v)$ of an 8 × 8 block of image pixels $f(x, y)$ are given by

$$F(u,v) = \frac{1}{4}C(u)C(v)\left[\sum_{x=0}^{7}\sum_{y=0}^{7} f(x,y) * \cos\frac{(2x+1)u\pi}{16}\cos\frac{(2y+1)V\pi}{16}\right] \qquad (3.1)$$

Where $C(x) = 1/\sqrt{2}$ when x equal 0 and $C(x) = 1$ otherwise.

The image is first subdivided into pixel blocks of size 8 × 8, which are processed left to right, top to bottom. As each 8 × 8 block of sub image is encountered, its 64 pixels are level shifted by subtracting the quantity $2n^{-1}$, where 2^n is the maximum number of gray levels. The 2-D discrete cosine transform of the block is then computed, quantized and reordered, using the zigzag pattern to form a 1-D sequence of quantized coefficients.

3.1.2 Quantization

Quantization, involved in image processing, is a lossy compression technique achieved by compressing a range of values to a single quantum value. When the number of discrete symbols in a given stream is reduced, the stream becomes more compressible. For example, reducing the number of colors required to represent a digital image makes it possible to reduce its file size. Specific applications include DCT data quantization in JPEG.

A common quantization matrix is:

$$\begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \end{bmatrix}$$

Typically this process will result in matrices with values primarily in the upper left (low frequency) corner. By using a zigzag ordering to group the non-zero entries and Huffman coding, the quantized matrix can be much more efficiently stored than the non-quantized version.

3.1.3 Entropy coding

Huffman coding is an entropy encoding algorithm used in the proposed methodology. The term refers to the use of a variable-length code table for encoding a source symbol (such as a character in a file) where the variable-length code table has been derived in a particular way based on the estimated probability of occurrence for each possible value of the source symbol.

Huffman coding uses a specific method for choosing the representation for each symbol, resulting in a prefix code (sometimes called "prefix-free codes", that is, the bit string representing some particular symbol is never a prefix of the bit string representing any other symbol) that expresses the most common source symbols using shorter strings of bits than are used for less common source symbols.

Huffman was able to design the most efficient compression method of this type: no other mapping of individual source symbols to unique strings of bits will produce a smaller average output size when the actual symbol frequencies agree with those used to create the code.

Huffman coding today is often used as a "back-end" to some other compression methods. And also it is used in multimedia codec's such as JPEG and MP3 have a front-end model and quantization followed by Huffman coding.

The steps of the Huffman coding are as follows:

1. Sort source outputs in decreasing order of their probabilities
2. Merge the two least-probable outputs into a single output whose probability is the sum of the corresponding probabilities.
3. If the number of remaining outputs is more than 2, then go to step 1.
4. Arbitrarily assign 0 and 1 as code words for the two remaining outputs.
5. If an output is the result of the merger of two outputs in a preceding step, append the current codeword with a 0 and a 1 to obtain the codeword the preceding outputs and repeat step 5. If no output is preceded by another output in a preceding step, then stop.

3.2 EMBEDDING THE WATERMARK IMAGE

The proposed watermarking scheme combines fragile watermarking process with the JPEG compression process. Watermarking technique embeds watermark into the host image, obviously, which will lead to distortion.

As shown in Fig. 4.1, during JPEG compression process, only the quantization step is irreversible and some information will be lost— this means distortion will occur during quantization step.

But, the watermark and quantization of JPEG are both based on psycho visual redundancy. So, the watermarking process parasitic at quantization step of JPEG is appropriate and natural.

The original image is an $M \times N \times 8$ bits grayscale image and [μ, $x1$] is the key of logistic system. The embedding of watermark image is illustrated in Figure 4 2.

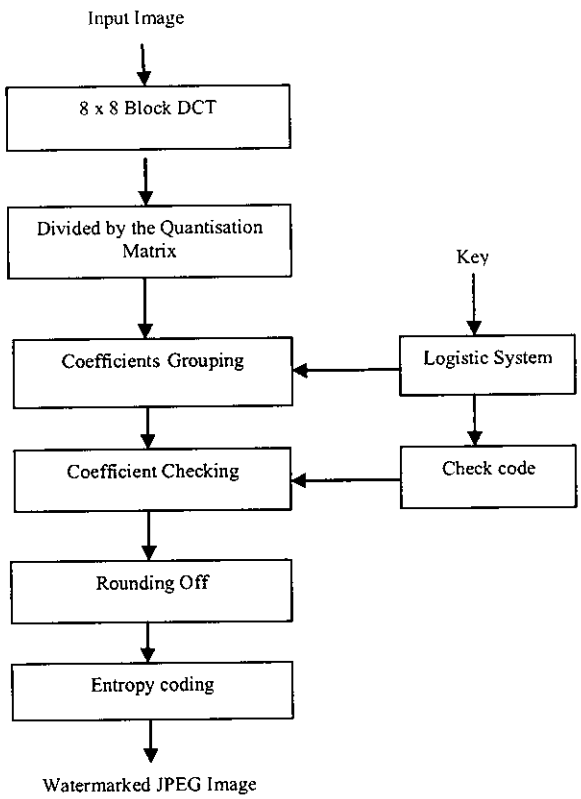

Fig 3.2 Proposed Watermarking Scheme

As shown in Fig. 4.2, after 8×8 block DCT and divided by the quantization matrix, the coefficients of 8×8 blocks are obtained without rounding off. Initially, the coefficients are subdivided in to every 8×8 DCT block into four groups. Then, every group can convey one watermark bit. The process is based on logistic system. The logistic chaotic sequence [6]:

$$x_{n+1} = \mu x_n(1-x_n) \tag{3.2}$$

where n=1, 2,... $0 < x_n < 1$, $3.6 < \mu < 4$

Set the key [μ, x_1] as the seed and generate a logistic chaotic matrix with the size M×N. The logistic chaotic matrix is subdivided into 8×8 blocks.

In this algorithm, all of the coefficients groups are checked by the check bits. The check bits are generated by logistic system. Set the key [μ, x_1] as the seed and generate the check bits.

3.3 TAMPER DETECTION

The tamper detection procedure verifies whether the contents of the watermarked image are tampered or not. The tamper probability is determined by checking the DCT coefficients groups when the tamper location or integrity verification of the image is needed.

The main process of tamper detection system is illustrated in Fig. 4 3

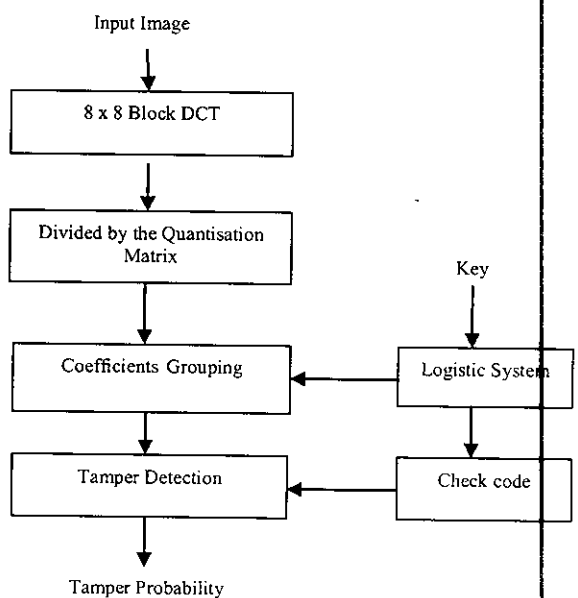

Fig 3.3 Tamper Detection System

Using the key [μ, x_1] as the seed the Tamper Detection is performed just as the watermarking steps.

3.4 PERFORMANCE MEASURE

The performance of the proposed watermarking scheme is evaluated by Peak Signal to Noise Ratio (PSNR).The difference between the input image and the watermarked image is given by the PSNR. The formula for PSNR value is given by Eq 3.1

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{255^2}{\mathrm{MSE}}\right)(\mathrm{dB}) \tag{3.1}$$

where, Mean Square Error (MSE) is,

$$\mathrm{MSE} = \frac{1}{M \times N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(x(i,j) - w(i,j)\right)^2 \tag{3.2}$$

where, x(i,j) is the original image and w(i,j) is the watermarked image.

CHAPTER 4
RESULTS AND DISCUSSION

The simulations are done by using MatlabR2008a.The host image is the well-known test image Lena (512×512×8 bits). The input image for watermarking algorithm is given in Figure 4.1.

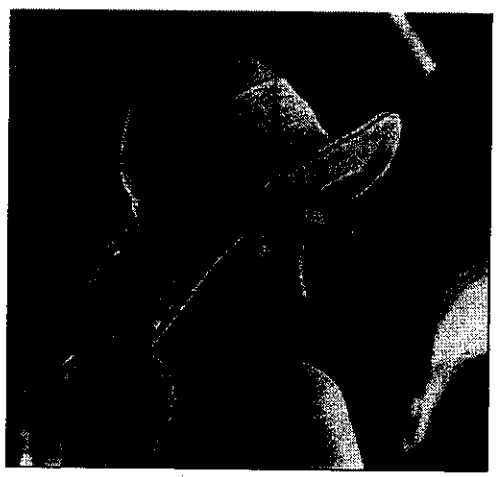

Fig 4.1 Lena as an input image

Image to be hidden inside the Lena image is given in the following Figure 4.2
The key [μ, $x1$] is the seed of logistic system.

Fig 4.2 Image to be Hidden

The watermarking system embeds watermark to host image and compressed the image into JPEG format. The size of the original image is 257KB, the size of the original JPEG is 28.8KB and the watermarked JPEG is 29.3KB. A logo is attached to the watermarked Lena image, as shown in Figure 4.3.

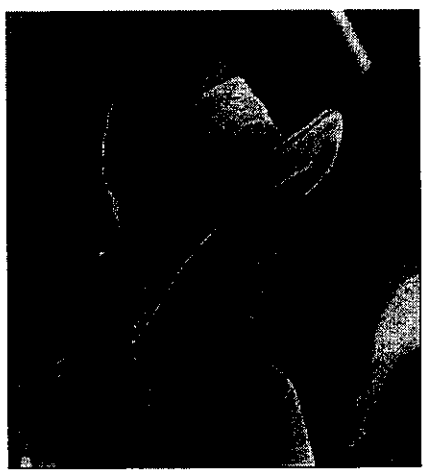

Fig 4.3 Embedded Watermarked image

Peak Signal to Noise Ratio is employed to quantify the degradation of the watermarked images. The PSNR of watermarked image is 37.1214dB. The watermark provides good invisibility. Extracted Hidden Image is shown in the Figure 4.4

Fig 4.4 Extracted Hidden Image

The proposed watermarking algorithm can locate the tamper with 8×8 blocks precision.

The tamper detection result using wrong key is shown in Fig. 4.5.

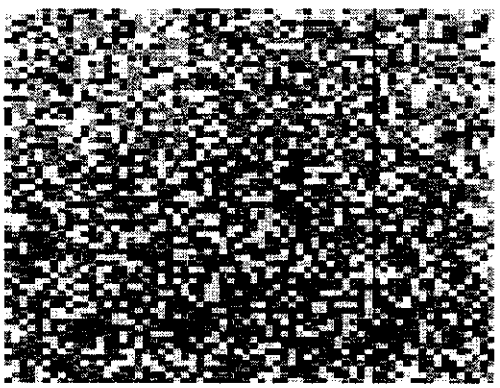

Fig 4.5 Extraction of hiding image using wrong key

When an image without watermark is examined or unauthorized person uses the incorrect key, extracted watermark image is just like the noise pattern. This indicates unauthorized person can't extract embedded watermark image correctly.

The host image is the well-known test image Pepper (512×512×8 bits). The input image for watermarking algorithm is given in Figure 4.6.

Fig 4.6 Pepper as an input image

Image to be hidden inside the pepper image is given in the following Figure 4.7

Fig 4.7 Image to be hidden

The watermarked image for pepper as an input is shown in Figure 4.8

Fig 4.8 Embedded Watermarked Image

The PSNR of watermarked pepper image is 30.33dB. The watermark provides good invisibility. Extracted Hidden Image is shown in the Figure 4.9

Fig 4.9 Extracted Hidden Image

The watermarking system embeds watermark to host image and compressed the image into JPEG format. The size of the original image is 759KB, the size of the original JPEG is 70.8KB and the size of the watermarked JPEG is 72.3KB.

The host image is the well-known test image Gold Hill (512×512×8 bits). The input image for watermarking algorithm is given in Figure 4.10.

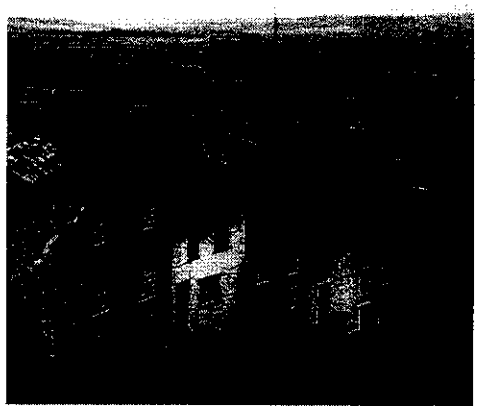

Fig 4.10 Gold Hill as an Input Image

Image to be hidden inside the Gold Hill image is given in the following Figure 4.11

Fig 4.11 Image to be hidden

The embedded watermarked image is shown in Figure 4.12

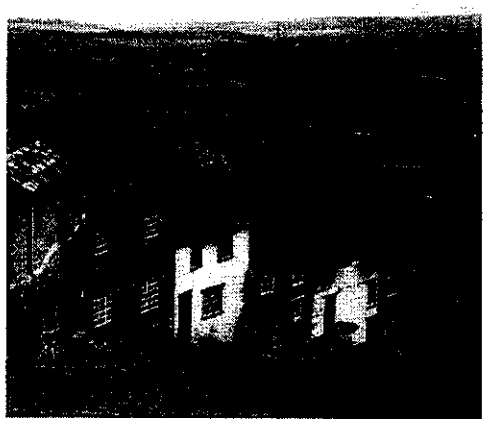

Fig 4.12 Embedded Watermarked Image

The PSNR of watermarked image is 34.73dB. The watermark provides good invisibility. The watermarking system embeds watermark to host image and compressed the image into JPEG format. The size of the original image is 257KB, the size of the original JPEG is 28.2KB and the size of the watermarked JPEG is 29.6KB.

Figure 4.13 shows the extracted hidden image using correct key.

Fig 4.13 Extracted image using correct key

The host image is the well-known test image Hall (512×512×8 bits). The input image for watermarking algorithm is given in Figure 4.14.

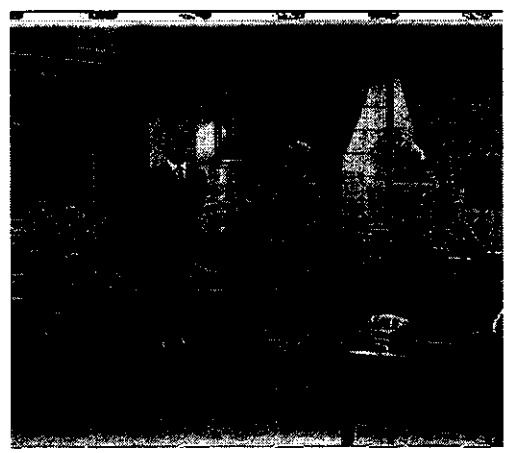

Fig 4.14 Hall as an Input Image

Image to be hidden inside the Hall image is given in the following Figure 4.15

Fig 4.15 Image to be hidden

The embedded watermarked image is shown in Figure 4.16

Fig 4.16 Embedded Watermark Image

The PSNR of watermarked image is 32.1214dB. The watermark provides good invisibility. Figure 4.17 shows the extracted hidden image using correct key.

Table 4.2

Size of original Images, Compressed JPEG Images and Watermarked JPEG Images

Input Images	Size of original Images in Kilobytes	Size of Compressed JPEG Images in Kilobytes	Size of Watermarked JPEG Images in Kilobytes
Lena	257	28.8	29.3
Pepper	759	70.8	72.3
Gold Hill	257	28.2	29.6
Hall	257	26.8	28.3

From the above table, it is observed that size of the original image is compressed by JPEG compression and the size of the water marked image is increased after adding watermarks

CHAPTER 7
CONCLUSION

Fragile watermark is researched in order to authenticate the veracity and integrity of electronic contents. In this project, we focus on the combination of watermarking and JPEG compression; a watermark is embedded into the host image during the procedure of host image JPEG compression. The coefficients of 8×8 DCT blocks are subdivided into 4 groups, then the groups are checked based on logistic system; in the tamper detection steps, the tamper can be detected by blocks. The experiment result shows that the proposed watermarking algorithm provides good invisibility, and can locate the tamper in 8×8 blocks precision.

The future scope of this project is to implement this watermarking algorithm to locate the tamper in 16X16 blocks precision.

Fig 4.17 Extracted Hidden Image

The watermarking system embeds watermark to host image and compression the image into JPEG format. The size of the original image is 257KB, the size of the original JPEG is 26.8KB and the watermarked JPEG is 28.3KB.

Performance measure of the Proposed Method is inferred in the given Table 4.1

Table 4.1
Performance measure of the Proposed Method

Input Image	Image to be hidden	PSNR(dB)
Lena	Dell Emblem	37.12
Pepper	Alphabet A	30.33
Gold Hill	Alphabet jc	34.73
Hall	Eye	32.12

From the above table, it is observed that the PSNR value falls above 30 dB. It is inferred that the watermarking achieves its objective by obtaining the high image quality for the watermarked image. There is no difference between the original image and the watermarked image. The size of original images, compressed JPEG images and watermarked JPEG images are given in the Table 4.2.

KUMARAGURU COLLEGE OF TECHNOLOGY

(An Autonomous Institution Affiliated to Anna University of Technology, Coimbatore)

COIMBATORE - 641049, TAMIL NADU, INDIA.

CERTIFICATE

CITEL 2011

This is to certify that ~~Mr./Ms.~~ A. KRISHNAVENI, ME [COMMUNICATION SYSTEMS] KUMARAGURU COLLEGE OF TECHNOLOGY, COIMBATORE has ~~attended~~ / presented a paper titled A WATERMARKING ALGORITHM FOR JPEG in the 3rd National Conference on **COMMUNICATION, INFORMATION AND TELEMATICS (CITEL 2011)** on 3rd & 4th March 2011, organized by the Department of Electronics and Communication Engineering, Kumaraguru College of Technology, Coimbatore.

Dr. Rajeswari Mariappan
HOD - ECE

Dr. S. Ramachandran
Principal

Dr. J. Shanmugam
Director