High Spatial Capacity Steganographic Algorithms based on Predictive Coding

Thesis submitted in partial fulfilment of the requirements for the award of degree of

Master of Engineering

in

Software Engineering

Submitted By:

Roopam Bamal

(801131021)

Under the supervision of:

Dr. V. P. Singh

Assistant Professor

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT

THAPAR UNIVERSITY

PATIALA – 147004

June 2013

Certificate

I hereby certify that the work which is being presented in the thesis entitled, **"*High Spatial Capacity Steganographic Algorithms based on Predictive Coding*"**, in partial fulfilment of the requirements for the award of degree of Master of Engineering in *Software Engineering* submitted in Computer Science and Engineering Department of Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of Dr. *V. P. Singh* and refers other researcher's work which are duly listed in the reference section.

The matter presented in the thesis has not been submitted for award of any other degree of this or any other University.

(Roopam Bamal)

(801131021)

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

(Dr. V. P. Singh)

Assistant Professor, CSED

Countersigned by

(Dr. Maninder Singh)

Head

Computer Science and Engineering Department

Thapar University

Patiala

(Dr. S. K. Mohapatra)

Dean

Academic Affairs

Thapar University

Patiala

i

Acknowledgement

I would like to express my deepest appreciation to Dr. V. P. Singh, my mentor and thesis supervisor for his constant support and motivation. He had been instrumental in guiding me throughout the thesis with his valuable insights, constructive criticisms and interminable encouragement.

I am also thankful to Dr. Maninder Singh, Head, School of Mathematics and Computer Applications Department for his constant support and encouragement. I would like to thank all the faculty members and staff of the department who were always there at the need of the hour and provided all the help and facilities, which I required, for the completion of this work.

I offer my deepest gratitude to my family for their support and affection and for believing in me always. I also want to thank my colleagues, who have given me moral support and their relentless advice throughout the completion of this work.

At last but not the least I would like to thank ”The Creator of Destinies” for not letting me down at the time of crisis and showing me silver lining in the dark clouds.

Roopam Bamal
(801131021)

Abstract

Information hiding refers to embedding additional digital data in cover objects–e.g. audio, images or video signals – by modifying the cover objects. These techniques are used in a variety of application domains like media forensics, content authentication, copyright protection and covert communications. The main objective of information hiding techniques is to minimize the effect that the embedding process has on the carrier object. Embedding processes introduce distortion to cover images and change the appearance as well as primary statistics of images.

In this dissertation, information hiding algorithms using both colored and grey images have been proposed with three basic goals: improving peak signal to noise ratio, preserving statistical properties and predicting a distortion level. The prediction errors created by the differential and brightness matrix for JPEG quantization used as improving spatial hiding information. The techniques used are lossless predictive coding and adaptive predictive coding. Adaptive predictive coding is an improved hybrid technique concatenating both prediction error expansion method and difference expansion. It is evident from the results that techniques introduced by proposed algorithm are efficient, functional methods for maximum hiding capacity and improved peak signal to noise ratio of the image. Minimization of distortion has been experimentally validated for the improved algorithms.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1. Introduction to Information Hiding

Information hiding refers to the process of embedding additional digital data in host signals, such as image, audio, or video files. The embedded data may vary. It may be identification numbers of media files, copyright information of art works, or information that should be kept secret. The variation in data depends on the purpose of the application. Digital images are the typical choice for the host signals because of their ubiquitous presence in the digital society. In this thesis, discussion is limited to image-based information hiding. Due to mature information and communication technologies and the rapid growth of the Internet, information hiding techniques have recently drawn the attention of the research community as well as industry.

1.1.1. Applications

Many applications have used information hiding techniques including traitor tracing, content authentication, copyright protection, and covert communication.

1.1.1.1.Copyright protection

As accessing the Internet and making digital copies of media have become widespread, copyright protection for digital media is more necessary than ever. Copyright information is embedded in media files before distribution. When needed, it is extracted to prove the ownership for the media. Attacks may occur unintentionally or intentionally. Unintentional attacks include general image processing such as smoothing, sharpening, or compression, while intentional attacks include all attempts to remove the embedded watermark. In order to be useful for copyright protection, the information hiding algorithm ensures that the embedded watermark is detectable after going through the various attacks, so that it could authenticate the copyright of the watermarked media.

1.1.1.2.Traitor tracing

A unique digital ID or signature is called a fingerprint. It is embedded into each piece of distributed digital media. If the distributor detects the unauthorized use of their media, the original user of the unauthorized copies can be traced by the unique fingerprint. In this case, a possible collusion attack between users is a major challenge. For example, multiple users dilute or erase the original fingerprint by averaging multiple copies of the media to avoid detection.

1.1.1.3.Content authentication

Content authentication is another important application [6, 7] of information hiding techniques. The application requires verifying that media files have not been altered since they were marked. This application can be useful for authentication in automatic video surveillance or driver's licenses [5, 8]. The algorithm should be designed in a way that alterations on the media easily destroy the embedded mark, so that the destroyed mark would show that the content has been tampered with and furthermore that where the alteration is located. Fragile watermarking refers to the technique that the mark is designed to be sensitive to any alterations, and semi-fragile watermarking if the mark is sensitive only to malicious alterations.

1.1.1.4. Covert communication

The goal of covert communication is to conceal the existence of the hidden message in the communicated medium. The medium is primary in other applications; the message itself is primary in covert communication. When data hiding technique is used in covert communication, it is called steganography. This word is of Greek origin and means covered or hidden writing [2]. On the other hand, the field of study which is interested in detecting the steganographic communication is called steganalysis. The main requirement for steganographic systems is undetectability: The goal of steganographic methods is that the object containing the concealed message should not be detected as a suspicious object by a steganalysis. The undetectability is challenging because steganalysis methods use visual and statistical inspections to detect the presence of a hidden message.

1.2. Introduction to Steganography

Steganography is an age old art of hiding information. People have hidden secrets throughout ages but the recent growth in computational power and technology has set it to the forefront of current security techniques. The word Steganography has evolved from Greek word, meaning “covered writing”. In simple words “steganography means concealing one piece of data within another, thus hiding the presence of the communicated information”. The first book of steganography was composed by Johannes Trithemius in 1499. The book Steganographia was published later in 1606 and immediately placed on the Index Librorum Prohibitorum. The exploration of this subject in literature may be traced to Simmons, who formulated it as the "Prisoners' Problem" who in 1983 [1]. In this problem, Alice and Bob are in jail, and wish to conceive an escape plan; all their communications pass through the warden, Wendy. If Wendy finds any encrypted messages, he will blow their plan by putting them into solitary confinement. So they need to find some way of hiding their ciphertext in an inoffensive looking covertext. Similarly in cryptography, we assume that the mechanism in use is known to the warden. So the security must depend only on a secret key that Alice and Bob have managed to share somehow.

Steganography is used in various data hiding applications especially in intellectual property, military, personal and diplomatic applications. Various techniques used to perform steganography are available. The most popular ones are invisible inks, digital signatures, character arrangements, microdots, spread spectrum communications and covert channels. These days Steganography has moved towards digital hiding with the advancement in computer and network technologies. It provides high usability communication channels for steganography. A number of digital technologies namely movie images, text files, audio and still images that are used for digital steganography are available. The outcome of steganography is a camouflaged channel of communication that is the system starts by identifying a cover medium’s redundant bits [2]. The secret data is thus transmitted in total secrecy escaping drawing eavesdroppers’ suspicions.

The main terminologies used in Steganography technology are: the secret message, secret key, cover message, and embedding algorithm [3]. Cover message is the carrier of the

message such as image, audio, text, video or some other digital media. The secret message is the data which needs to be concealed in the suitable digital media. The embedding algorithm is the way used to embed the secret information in the cover message. The secret key is used to embed the message depending on the hiding algorithms [4].

1.3. Process of Embedding

The embedding process modifies the original medium in order to embed the message. The modification performs in the spatial domain or transformation. The issue of selecting the location in the domain is also important. For instance, embedding in low frequency coefficients has a different effect from embedding in the high frequency coefficients. It is tough for the human visual system to detect modifications of the high frequency coefficients. So, embedding by modifying the high frequency coefficients can achieve more imperceptible embedding rather than ones in the low frequency coefficients. The location of embedding may be determined by a secret key, for security reasons. The original medium embedded before a message is termed a cover object and the medium embedded after the message is termed a stego object.

The process of embedding includes passing both the secret message and the cover message into encoder. Secret message is embedded into the cover message based on the type of technique. A secret key or stego key is required for embedding process. Stego object should look almost identical to cover message with hidden secret message present inside it.

The decoding process includes interpolation of Stego object into the decoding system. The stego key is used to decode and find the secret message.

So, hiding information into a media requires following elements [13]:

- The cover media(C) that will hold the hidden data
- The secret message (M), may be plain text, cipher text or any type of data
- The stego function (Fe) and its inverse (Fe-1)

- An optional stego-key (K) or password may be used to hide and unhide the message.

The stego function operating over cover media and the message (meant to be hidden) along with a stego-key is used to produce a stego media (S).

The schema of steganographic operation is shown in Fig 1.1 and its projection with images for both encoding and decoding process is explained in Fig 1.2

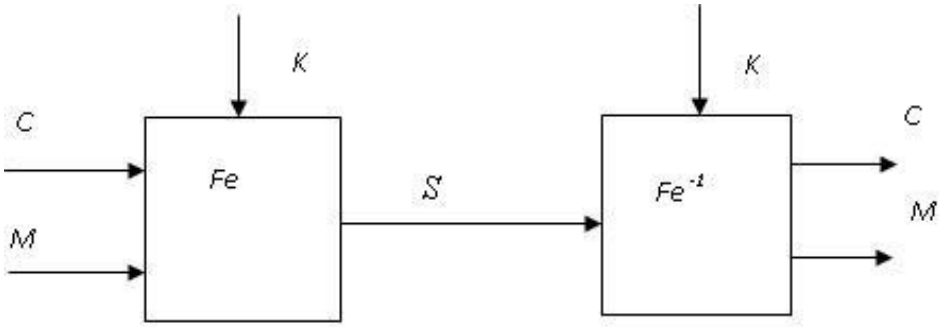

Figure 1.1: The Steganographic operation

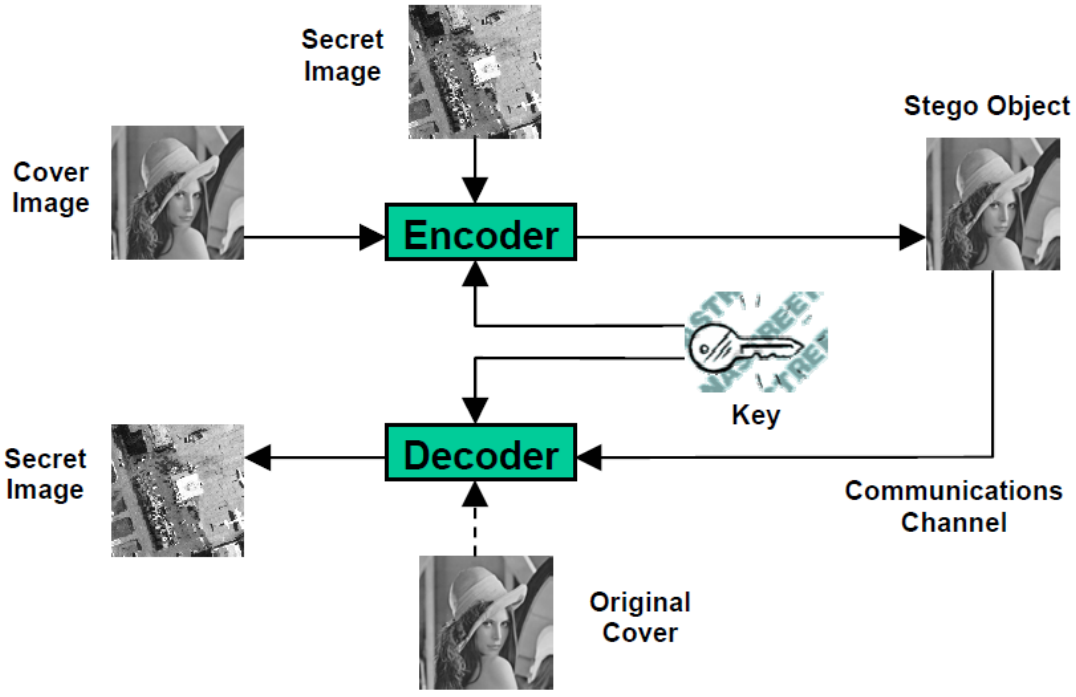

Figure 1.2: Simple Process of Encoding and Decoding [14]

Each embedding process is described based on the following three factors: [15]

- Capacity
- Security
- Robustness

Capacity is defined as the amount of information that can be hidden in the cover medium.

Security is defined as an eavesdropper's inability to detect hidden information.

Robustness is defined as the amount of modification the stego medium can endure before an adversary can destroy hidden information [8].

Steganography strives for high capacity and security, which often requires that the hidden information to be fragile.

1.4. Steganalysis:

The process of detecting Steganography and its use is known as Steganalysis. There are two types of Steganalysis:

Passive Steganalysis: Presence of hidden data is detected.

Active Steganalysis: Attempt is made to retrieve hidden data.

Due to a number of hiding process and algorithms, detection has always been a difficult task and time consuming. So, detection is still an active area of research.

1.5. Predictive Coding:

Predictive coding is information source coding based on data correlation. The basic thought is reducing data dependency on the time domain and space domain. Predictive coding methods employ several heuristics to reduce the prediction errors locally without considering how to minimize the errors globally over the whole image file. According to predictive coding, the values of the adjacent samples are used to predict the current sample value. Then the difference (Prediction Error) between the predicted value and real

sample value is coded and transmitted. The aim is to reduce data dependency. Though the range of the difference values is increased, the majority of the difference values are highly concentrated in a small range near zero owing to the high correlation between neighboring pixels in natural images. Considering the large values in the prediction-error image may be caused by the image content rather than by the data hiding process. A certain threshold is applied to the prediction errors to remove the large values in the prediction error images for steganalysis. Thus it limits the dynamic range of the prediction error image.

There are two approaches usually used to acquire the predictive values:

- Predicting the current pixel value from the adjacent pixel value. It is often happens that the previous pixel value is used to predict the current pixel value.
- Predicting the current pixel value from the adjacent image block. For example, the current pixel value is predicted from the weighted average of the several adjacent pixels.

Main idea of the information hiding algorithm based on the predictive coding is to use the prediction errors for the transmitting of the secret information.

1.5.1. Prediction Error Image:

Generally natural images are continuous, smooth and tend to have a high correlation between neighboring pixels because any object has a certain size. The hidden data may be independent to the cover media. As a consequence, it will reduce the correlation among adjacent pixels, bit planes and image blocks. In steganalysis, the variation caused by data hiding should be increased. The use of neighboring pixels to predict the current pixel forms the basis of predictive coding. The predictions are made in three directions, namely, vertical, horizontal and diagonal since a digital image is actually a 2-D array as shown in Fig 1.3.a to 1.3.c

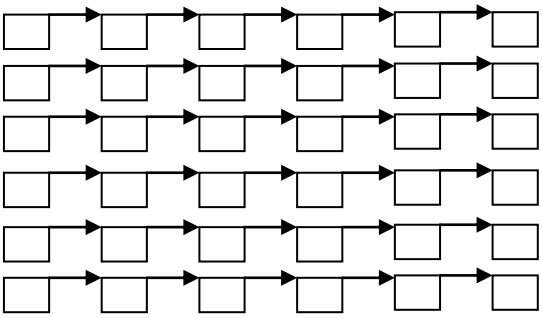

Figure 1.3.a. Prediction Error in Horizontal direction

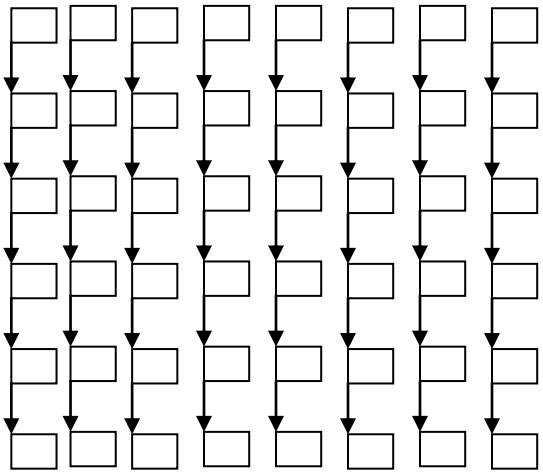

Figure 1.3.b :Prediction Error in Vertical direction

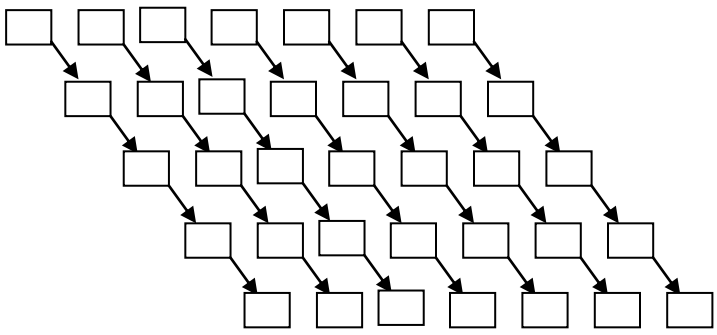

Figure 1.3.c :Prediction Error in Diagonal Direction

The above prediction errors are calculated based on following equations given below as eq (1), eq (2), eq (3):

$$E_h(i,j) = x(i+1, j) - x(i, j) \quad \ldots(1)$$

$$E_v(i,j) = x(i, j+1) - x(i, j) \quad \ldots(2)$$

$$E_d(i,j) = x(i+1, j+1) - x(i, j) \quad \ldots(3)$$

where eh (i, j) indicates the prediction error for pixel (i, j) along horizontal direction while ev (i, j) and ed (i, j) the prediction error for pixel (i, j) on vertical and diagonal directions, respectively.

There are three prediction errors for each pixel of an image,. The prediction errors will form three prediction-error images denoted by Eh, Ev and Ed, respectively. The prediction error can be obtained by subtracting the predicted pixel value from the original pixel value, for each prediction made.

1.5.2. Advantage of Prediction Error:

The utmost importance is given to the distortion in data hiding in steganalysis. This type of distortion may be rather weak and hence covered by other types of noises. With the purpose of enhancing the noise introduced by data hiding, prediction of each pixel value in the original cover image is by made using its neighboring pixels' values. A prediction-error image is obtained by subtracting the predicted image from the test image. It is expected that this prediction-error image removes various information other than that caused by data hiding. Thus it makes the steganalysis more efficient because the hidden data are usually unrelated to the cover media. The prediction-error image is used to erase the image content, in other words.

1.6. High Capacity in Information Hiding

The transform domain algorithm replaces the image's least significant bit (LSB) or all bits of multi unimportant bit plane with the hidden information. LSB algorithm can embed 1 bit information in every pixel. The algorithm has a weaker robustness because it use the most unimportant pixel bit in image, the embedded information is attacked easily

by the operations such as wave filtering, the image quantization, the geometric deformation and the increased noise etc.

To overcome this weakness, the information can be hidden using the Most Significant Bit (MSB) in a 3-color component of each pixel in 24-bit RGB color image, thus giving rise to high capacity in information hiding. This highest bit with the use of multiple plane bits of 3-color components of RGB gives larger information hiding capacity along with better robustness.

1.7. Research Motivation

During the last decade, internet activities became an important part of many people's lives. As these activities number is been increasing, there is a growing in amount of personal information about the users that is now a days stored in electronic form and which is usually transferred using public electronic means. This makes it feasible, and often very easy, to collect, to transfer and to process a huge amount of details about a person.

In view of the data security questions, the information securing technology had already became the hot spot in the research field and one of the important objectives of information hiding research is that secret information is embed as much as possible without the perception of the carrier is affected that is the hidden information does not arise the attacker's attention. This thereby demand for information hiding techniques that can reduce the possibility of being encroached and get rid of the fatal flaw in the data encryption technology.

The need for better algorithm also arises due to the development of the multimedia technology and the widespread application of network communication which raises the information security questions, such as information authentication, copyright protection and secret transmission that have become an increasing concern.

Many scholars have made a lot of research on the information hiding technology. For instance, the improvement on the classical LSB algorithm, the lossless information hiding method based on the space domain and transform domain of the image proposed by

Fridrich, the lossless data hiding method based on prediction errors' expansion proposed by Thodi, the further study on the linear prediction carried on by YumingXie and HongjiPiao and so on so forth. Currently there are many hidden algorithms based on spatial domain and transform domain, but the embedding capacity and the security of information hiding remain to be improved.

The idea of combining high capacity of information hiding with Predictive coding has motivated this research and lead to the formation of new algorithm that not only reduces the distortions but also erases the image content to give higher capacity of information hiding with the use of multi plane bit 3-color components of RGB.

This research is restricted to steganography within photographic images, necessarily implying maximization of payload. The trade-off between perceptibility and robustness, however, depends on the application. In this work, imperceptibility is considered essential while robustness is only necessary so that a change in image format does not affect the hidden data. The chosen requirements result in a system that is eminently suitable for secret communication. Other appropriate uses include those applications where it is expected that the image will not be further processed once the data is embedded (except perhaps for format conversion and lossy compression).

1.8. Organization and Scope of Thesis

This thesis will present information hiding schemes (Predictive Coding, Adaptive Predicting Coding, High Capacity) using JPEG images with three basic goals: minimizing distortion because of embedding, preserving the statistical properties during embedding, and lastly predicting the level of distortion along with the given message length.

Since stego images are created by modifying the carrier images, some level of distortion to the carrier images is unavoidable; but the minimized distortion is a huge requirement in many applications. Especially in high-image-quality-required applications, which includes the photographs for court evidence, fine art works, surveillance videos that may have to be processed again, minimal distortion in the images should be one of the primary requirements.

Another motivation of aiming for minimal distortion is that modification in the image statistics can be minimized by the method of minimizing distortion. Relationship among the distortion and statistical change in information hiding is explored by this thesis. Minimizing distortion due to embedding is the first goal of this thesis.

In addition to producing the distortion, embedding messages also alters statistical properties of the carrier image. When one given algorithm only focuses on minimizing the distortion, it may be disregarding the effect of the statistical alteration, which will make the algorithm more vulnerable to the steganalysis attacks: The altered statistical properties can be used for the steganalysis methods, which are interested in detecting the use of steganographic methods. Preserving statistical properties in images during embedding is the second goal of this thesis. Another goal regarding distortion in this thesis is predicting the distortion level with a given message strength. Distortion always varies depending upon the payload, the length of message. A high payload increases the level of distortion. There has been very little work on how to choose the payload in terms of the tradeoff between payload and distortion. Instead, trial-and-error methods have been used to determine the payload with a given distortion level. In this thesis, the distortion is predicted as a function of message length and one of the cover data statistics. The mathematical analysis will be helpful in determining the payload automatically.

The expected distortion with each algorithm is mathematically analyzed in the chapters. Achieving the three goals is validated by measuring peak signal to noise ratio and distortion. Now, distortion can be measured as a distance between carrier data and stego data. When each embedding algorithm is proposed, the distortion for various embedding rates is presented and the various results are demonstrated in each chapter. The embedded algorithm is continuously improved to decrease distortion and the results show that how low a distortion have been achieved through the improvement. The accuracy in the estimation of distortion due to embedding is shown by plotting the theoretical one and experimental one in the same graph. To show how closely the embedding algorithm preserves the original statistics, differences between the histograms before and after embedding are measured. Lastly, the embedding algorithms have been tested with the well-known steganalysis tests and showed the results.

Chapter 2 – This chapter gives an overview of the literature in the field of information hiding. The literature includes the work done in the past and the improvement required. It forms the basis for the problem statement.

Chapter 3 – This chapter focuses on research gaps and also state problem statement for this thesis.

Chapter 4 – This chapter provides solution to problem stated in previous chapter by proposing new algorithms to improve the capacity of the steganographic process.

Chapter 5 – This chapter demonstrate the proposed solution by implementing it using Java, Matlab and Excel tools. Results of experimental setup are shown.

Chapter 6 – This chapter states the conclusion, contribution of work and future research possible.

Chapter 2
Literature Review

Information hiding is a relatively new research field. Detailed survey of early algorithms and softwares for information hiding techniques are described in the papers by Petitcolas [2] and Swanson et al. [17]. A comprehensive overview of many applications is provided in the books by Cox, Miller, and Bloom[18], Johnson, Duric, and Jajodia [19], Katzenbeisser and Petiticolas [20], and Wayner[21]. An overview focusing on the principles and mathematical methods on information hiding is provided by Moulin in [5].

The first academic conference related to information hiding is International Workshop on Information Hiding (IH), which has been yearly held since 1996. The IH conference proceedings are considered the primary source for the recent information hiding techniques. In addition, two recent peer reviewed journals are IEEE Transactions on Information Forensics and Security and International Journal of Information Security.

The methods proposed in this thesis are applicable to steganography and steganalysis, therefore the review will be limited to those two subfields rather than covering all information hiding techniques that have been developed for the last decade.

2.1. Steganography and Steganalysis:

People have been using steganography to hide information since Greek times. However, digital steganography is a relatively new research field [20]. Since being undetectable is one of the essential requirements for steganography, steganalysis and steganographic applications are evolving in competition with each other.

Fingerprinting and Traitor-Tracing: Steganography is related to traitor tracing and digital fingerprinting. The technique attempts to embed unique identification information into copies of an object in order to trace traitors who release illegal copies. Most of work in fingerprinting deals with cryptographic data only.

Covert Channels: Covert channels have been studied since 1973 in computer security, where one secretly abuses certain parts of a computer system so as to create an anonymous and secret communication channel. In the area of cryptography, covert channels embedded in cryptographic protocols are called subliminal channels, named by Gus Simmons who first studied them in authentication protocols. He introduced the prisoners' problem. One real life example of such a scenario involves the SALT 2 accord between the US and the Soviet Union.

Desmedt generalized the concept of subliminal channels to the more general case of cryptographic protocols. These channels allow an insider to send secret side information through cryptographic protocol messages. It is important to prevent such abuses since they allow insider leakage and electronic espionage through cryptographic communication systems.

Kleptography: Young and Yung further applied this concept to cryptographic hardware and software products. They demonstrated that viruses or malwares with cryptographic capabilities can use subliminal channels in a very damaging and undetectable ways to leak information and to destroy data integrity. This suggests that secure cryptosystems should not be built from externally built blackboxes with computer processing capabilities. Even though subliminal channels and kleptography are special cases of information hiding, their objectives are different. With subliminal channels one hides information into cryptographic protocols, which often contain strings that are uniformly random, whereas uniform random messages are generally considered suspicious and unsuitable for use in modern information hiding. Nowadays secret information is hidden in more natural data such as audio, pictures, or video. This difference leads to entirely different methods used for information hiding.

Anonymous Communications: Chaum introduced the concept of unconditional anonymity and explained the same by dining cryptographers' problem. Unconditional anonymity has many applications in electronic voting and payment systems. In its basic form it involves an end-to-end communication network where no one knows who are the real senders and receivers of each message. Other variants of this concept are MIX

networks, where several layers of servers are used to hide the origin of each messages. Onion routing is a special implementation of MIX networks.

2.2. Modern Day Techniques of Steganography, 'Digital Hiding'

There are a number of digital technologies namely text files, movie images, audio and still images that are used for digital steganography.

2.2.1. **Text Steganography:**

Hiding information in plain text can be done in many different ways. Some techniques consist of relating the characters to hide with the characters of the carrier text, creating a reference dictionary which maps words from the secret text with words from the carrier text. Some include changing the outline of the carrier text such as adding whitespaces or altering the case of certain characters so as to represent secret information. Some of the commonly used text steganography techniques include hiding using whitespace, hiding by selection, hiding in HTML, line and word shifting, abbreviation-based hiding and semantic-based hiding [9, 10].

2.2.1.1. HTML Documents:

Secret text can be easily hidden within HTML documents because HTML tags are case insensitive. For example, the tags <a title="host">, <a TITLE="host">, and <a TitlE ="Left">, are the same and have the same effect on the rendering of the document. Text steganography can be performed by changing the case of the letters that make up the HTML tags in case of HTML documents. In particular, the secret message is represented by the capital version of the tags' letters, or vice versa depending on the algorithm being used. For the recovering process, all the capital letters from the HTML document have to be captured and concatenated together for producing the original camouflaged message [11].

2.2.1.2. Hiding using Whitespace:

A number of extra blank spaces are embedded between consecutive words of cover text. The message to be hidden is first converted into a binary format. After that, every bit whose value is 1 is represented by an extra whitespace between a particular set of two words in the carrier text. Every bit whose value is 0 leaves the original single whitespace between the next particular set of two words. For instance, the statement "the boy went to

school today ” can be deciphered as “101001”. In fact, two spaces exist between “the” and “boy”, between “went” and “to”, and between “today” and the end of the sentence. It results in a bit of value 1 in positions 0, 2, and 5 respectively. In opposite to this, only a single space exists between “boy” and “went”, between “to” and “school”, and between “school” and “today”. The process results in a bit of value 0 in positions 1, 3, and 4 respectively. Basically, the whitespace technique is very suspicious as a normal reader would easily notice the existence of some extra whitespaces in the text. Additionally, this method cannot encode too much information especially in small text [25].

2.2.1.3. Hiding by Selection:

The selection technique selects specific characters in the carrier text to convey the characters of the secret message such as selecting the first character of every word in the carrier text or the second character of every other word. In order to recover the concealed secret message, all first characters of the words of the carrier text are extracted and concatenated together, producing the exact original message. Variation of this technique can be done by selecting the first character from the first word, the second character from the second word, the third character from the third word, and so on, until the characters of the message to hide get exhausted [26].

2.2.1.4. Line and Word Shifting:

In this technique, text lines are shifted vertically and words are shifted horizontally by a fixed space of n inches. This way, distance between lines and words would convey the hidden characters. For example, letter A can be encoded as a 0.01 inch space between two text lines. Similarly, letter B can be encoded as a 0.02 inch space between two other text lines. This technique is more suitable for printed text than for digital text, since printed spaces can be physically measured unlike their digital counterparts [27].

2.2.2. Image Steganography:

Images are considered to be the most popular file formats used in steganography. Image Steganography is beneficial because of the possibility to access any pixel of the image at random. Additionally, the hidden information could remain invisible to the eye.

Compression Methods:

- Lossless compression is known for being preferable if the original data stay in its entirety. In this way, the original image information will never be removed and this makes it possible to reconstruct the original data from the compressed information. This is typical of images in GIF and BMP [28].
- Lossy compression saves storage space by discarding the points the human eyes find difficult to discriminate. In this case the resulting image is expected to be something similar to the original image but not same as the original. JPEG compression uses the technique. A cover image is the image designated to carry the embedded bits or secret information [29].

Most of the Steganographic softwares use lossless 24-bit images such as BMP. Other alternative to this is gray scale image or 256 color image. These kinds of softwares neither support nor recommend using JPEG images. Most common of these found on the Internet are GIF files.

Different ways of information hiding in digital images include:

- Least significant bit (LSB) insertion
- Algorithms and transformations
- Masking and filtering

2.2.2.1. **Least Significant Bit Hiding Technique (LSB):**

Least significant bit insertion is a very common and simple approach to embedding information in a cover file. It conceals the secret message in the RGB image based on it its binary coding. LSB is used to hide information in GIF or BMP images which reconstructs the original message exactly to a JPEG. The LSB algorithm is used to hide the secret messages by using the following algorithm [30, 31].

Algorithm for Least Significant Bit Hiding Technique

Inputs: RGB image, secret message and the password.

Output: Stego image.

Begin

scan the image row by row, encode it in binary.

encode the secret message in binary.

check the size of the image and then the size of the secret message.

start sub-iteration 1:

choose one pixel of the image randomly

divide the image into three parts (Red, Green and Blue parts)

hide two by two bits of secret message in each part of the pixel

in the two least significant bits.

set the image with the new values.

end sub-iteration 1.

set the image with the new values and then save it.

End

LSB manipulation is a fast and easy way to hide information but is vulnerable to small changes resulting from image processing or lossy compression.

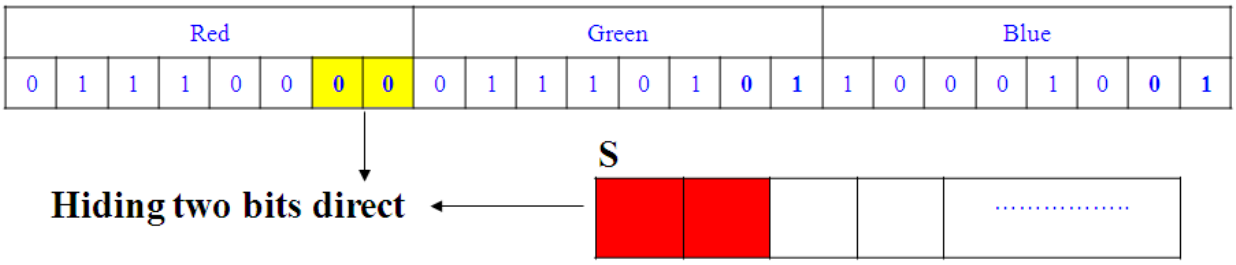

Fig 2.1 : LSB Hiding Technique

In Fig 2.1, the two least significant bits (in this case from Red part) are selected based on the LSB algorithm. These bits are then used to hide the information.

LSB Example:

Lion with Mandrill hidden in it

Fig 2.2.a : Original Image

Fig 2.2.b : 3bits hidden

Fig 2.2.c : 4bits hidden

Fig 2.2.c : 5bits hidden

The LSB technique used to hide the bits is not effective with 5bits as the Mandrill eyes can be seen in the background in Fig 2.2.c while for Fig 2.2.b and Fig 2.2.c Mandrill is perfectly shielded from human eyes and is undetectable. Thus LSB is valid for hiding 3bits and 4bits.

2.2.2.2. Algorithms and transformations:

LSB insertion technique is vulnerable to small changes which occur while converting JPEG images back to GIF or BMP. The information stored could be damaged. For this reason, JPEG images can be used for concealing information.

JPEG images uses **discrete cosine transform (DCT) [32]** coefficients for embedding the data. DCT is a lossy compression transform because the cosine values cannot be exactly calculated and repeated calculations using limited precision numbers introduce rounding errors to the final result. Variations between original data values and restored data values depend on the method used to calculate DCT. With DCT transformation, a signal is transformed from the representation of an image into the frequency domain. This is done by sorting the pixels into (8 × 8) pixel blocks and transforming these blocks into 64-DCT coefficients which are affected by any modification of a single DCT coefficient.

The DCT coefficients F(u, v) of an 8 × 8 block of image pixel f(x, y) is given by

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

Where $c(x) = \frac{1}{\sqrt{2}}$ when x = 0 and c(x) = 1, otherwise

Algorithm for DCT Coefficients:

Input: message, cover image

Output: stego image

while data left to embed do

    get next DCT coefficient from cover image

    if DCT ≠ 0 and DCT ≠ 1 then

        get next LSB from message

        replace DCT LSB with message LSB

    end if

    insert DCT into stego image

end while

Other form of transformation technique is **Fourier transformation and wavelet transformation**[33].

2.2.2.3. Masking and Filtering:

Masking and filtering techniques is performed on gray scale images or 24-bit images. The technique conceals information by marking an image. Digital Watermarking is one of

those methods. Masking is comparatively more robust than LSB insertion with respect to cropping and compression. These techniques embed information in significant areas so that the hidden message is more integral to the cover image than just hiding it in the "noise" level. It makes it more suitable than LSB with, for instance, lossy JPEG images [34].

2.2.3. Audio Steganography:

In audio steganography, secret message is hidden into digitized audio signal which result slight altering of binary sequence of the corresponding audio file. The process of hiding data in audio files can be done in number of ways. Technique includes low bit encoding, Phase coding, Spread Spectrum, MIDI, MP3 and echo data hiding.

2.2.3.1. **Spread spectrum** systems encode data as a binary sequence which sounds like noise but which can be recognized by a receiver with the right key. Spread Spectrum method spreads the secret information over the frequency spectrum of the sound file using a code which is independent of the actual signal. So, the final signal occupies a bandwidth which is more than what is actually required for transmission [22].

2.2.3.2. **Low-bit encoding** substitutes the least significant bit of information in each sampling point with a coded binary string [35].

2.2.3.3. **Phase coding** method replaces the phase of an initial audio segment with a reference phase that represents the information. The phase of consecutive segments is adjusted for preserving the relative phase between segments [36].

2.2.3.4. **MIDI** files are made up of different messages. Some of the messages control the notes you hear when others are silent and make up the file header or change the notes being played.

2.2.3.5. The MP3 format is the most widespread compression format presently used for music files. During the compression process, **MP3Stego** will hide information in MP3 files. The data is then compressed, encrypted and then hidden in the MP3 bit stream. Hiding process takes place at the heart of the Layer III encoding process namely in the inner loop. This inner loop quantizes

the input data and increases the quantizer step size until the quantized data can be coded with the available bits.

2.2.3.6. **Echo Data Hiding** technique inserts data into a host signal by introducing an echo in a discrete signal. The data are hidden by varying three parameters of the echo: decay rate, initial amplitude and offset. The echo blends as the offset between the original and the echo decreases,

2.2.4. Video Steganography:

A video can be viewed as a sequence of still images. The data embedding in videos is very similar to images but with added complexity which arise due to size of the host media and the number of pixels. Video Steganography can be divided into two main classes. One of the classes is embedding data in uncompressed raw video, which is compressed later on. The other class tries to insert data directly in compressed video stream [38].

Since video has both audio and image, a combination of sound and image techniques can be used. Video generally has separate inner files for the video and the sound. Therefore, techniques can be applied in both areas to conceal data. The scope for adding lots of data is too large and hence the chances of hidden data being detected are very low.

2.2.5. Intellectual Property Steganography:

2.2.5.1. Digital Watermarking:

Digital watermarking uses various techniques like music, images, movies which can be imprinted with digital watermarks [39]. A digital watermark is a signal permanently embedded into digital data (audio, images, video, and text) that can be detected or extracted afterwards to confirm the authenticity of the data. The watermark is concealed in the host data in such a way that it cannot be removed without demeaning the host medium [40].

A watermarking system's main goal is to achieve a high level of robustness - that is, it should be impossible to remove a watermark without degrading the quality of data object.

2.2.5.2. Fingerprinting

In Fingerprinting [41], different and separate marks are embedded in the copies of the object that are supplied to different customers. It enables the intellectual property owner

to find customers who break their license agreement by supplying the property to third parties [42].

2.3. Modern Day Technique of Steganalysis

The goal of steganalysis methods is to detect the presence of a hidden message, although the message itself does not need to be decoded. Techniques for image steganalysis are categorized into two: **algorithm-specific steganalysis** and **universal steganalysis**. The first approach of steganalysis methods were designed with respect to the embedding method, and that can be grouped, depending upon the format of the image which uses the embedding algorithm like raw images, JPEG images or palette-based images. For embedding the secret message in raw the images, LSB replacement technique was been used in the pixel domain. There are two described steganalysis techniques to detect the LSB replacement embedding: visual attack [43] and chi-square attack [43].

The visual attack detected changes in the LSB plane of images when the LSBs were replace by a message. Since the LSBs are correlated with the image, embedding of the random message bits changes the correlation between the LSB plane and the image. For example, after embedding with full capacity of the image (i.e., every LSB is replaced with the message bit), the LSB plane of the image looks like random noise.

The chi-square attack detected changes of image statistics by LSB replacement. The LSB replacement results in producing pairs of adjacent coefficients that their values differ by only 1 (e.g., (2,3), (4,5)). The frequencies of those adjacent coefficients have changed depending on the frequencies of embedding message bits (0 or 1). The chi-square attack [43] found that the frequencies of those adjacent coefficients become similar as the amount of the embedded messages becomes large. Since the chi-square attack was proposed, the LSB replacement techniques have been considered as easily detectable by examining histograms [44–47] or structural analysis [48]. LSB steganalysis approaches that also provided message length estimation were proposed for gray scale images [49, 50] and color images [51].

The Hide algorithm increments or decrements the sample value in order to match its LSB to the message bit, which cannot detected by simple LSB steganalysis. This ±1 embedding technique used in the Hide algorithm was detected by looking at the neighborhood colors for each color value as proposed in [52]. The next group of steganalysis is designed with respect to embedding methods forpalette-based images. Many steganography methods [53] for palette-based images (example GIF images) have been proposed. The early proposed steganalysis method [54] for palette-based images used visual inspection or simple statistics for detection. Later, EzStego-specific steganalysis methods [55] explored color pairs and obtained once the sorting of palette is completed. The phenomenon leads to the changes in their patterns after the embedding.

The last group of steganalysis is designed with respect to embedding methods for JPEG images. Due to emergence of many steganographic methods [56, 57] employing JPEG images, many JPEG steganalysis methods have been proposed. Two steganalysis methods [50,51] have proposed to detect the F5 and Outguess algorithms using DCT coefficients histogram and blockiness. The steganalysis technique in [58] proposed a method of estimating the cover image by cropping the stego JPEG image by 4 columns and then encoding the cropped image using the same quantization table. Stego image is compared with the histogram of the estimated cover image, and then the length of message was estimated.

The steganalysis technique to detect Outguess was proposed in [59]. Since Outguess can preserve the original histogram of DCT coefficients, the technique of estimating the original histogram is not effective. Blockiness was used instead of Outguess in steganalysis. Blockiness refers the spatial discontinuities along the 8×8 DCT blocks. Blockiness increases with an increase in the length of the embedding message.

While the first group of steganalysis methods was designed for a specific embedding method, the second type of steganalysis methods are designed to detect any embedding algorithms, called Universal Steganalysis or Blind Steganalysis Technique. Universal steganalysis basically designs a classifier that can differentiate the cover images and stego images. It is very important to find the features that can capture the statistical changes introduced in the images after embedding.

A universal steganalysis method proposed in [60,61] decomposed an image using quadrature mirror filters and then high order statistics were calculated from the high frequency bands to be used as features for steganalysis. In [62], the authors used the low-low wavelet subbands to calculate the statistical moments of characteristic functions for their steganalysis. Another steganalysis method estimated a cover image of the JPEG format and the statistics of the estimated cover image was used as the features to detect data embedded in JPEG images.

The methods in [63,64] detected spread spectrum data hiding method using the interpixel dependencies and a Markov chain model. The Morkov features were also applied in [65, 66] attacking JPEG steganography methods.

2.4. Image Based Information Hiding System

As the thesis focuses on designing information hiding algorithms using JPEG image, overviews of image based information hiding system is given below:

Generally, information hiding methods modify a cover object to embed a message and create a stego object.

Figure 2.3 shows an overview of the information hiding system. As the thesis designed image based information hiding methods, X and $\hat{X}$ are images. A cover image

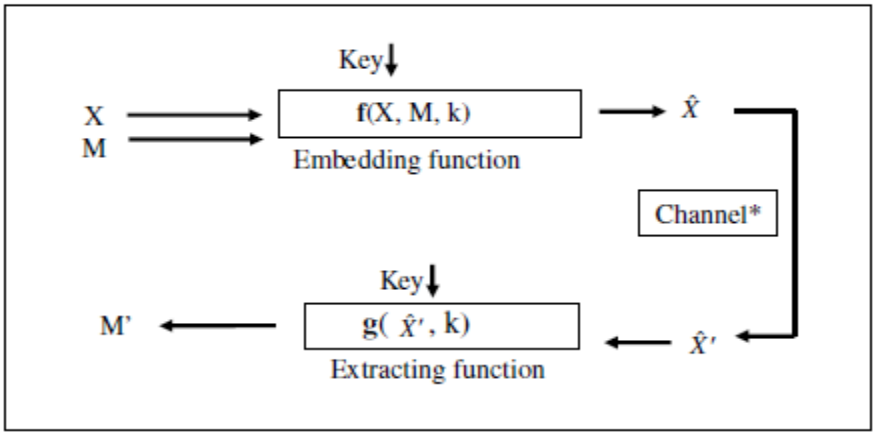

Figure 2.3 : An overview of information hiding system

(X) and a message (M) is sent to the embedding function with a secret key. Typically, the message is assumed to be encrypted for security reasons, and the secret key and the embedding function are assumed to be known to the receiver. An information hiding algorithm has a pair of functions f and g such that

$$\hat{X} = f\ (X, M),\ M = g\ (\hat{X}) \qquad \text{,…(5)}$$

The typical embedding function finds the insignificant part of the cover image for the human visual system and modifies them to embed a message. The output of the embedding function is the stego image and ($\hat{X}$). It is not easy to tell the stego image from the cover image by the human visual system. Channel is used to send the stego image to the receiver. The channel may modify the stego image by adding noises. Stego image is sent to the extracting function with the key. The extracting function retrieves a message from the stego image. The message may or may not be the same to the message that the sender has sent. The usage of the extracted message varies depending on the applications.

JPEG image formatting removes some image details to obtain considerable saving of storage space without much loss of image quality. The savings are based on the fact that humans are less sensitive to changes of the high frequency components than of the low frequency components. Information hiding techniques using JPEG files operate in the frequency domain. One of the benefits of embedding in the frequency domain is that of not being affected by visual attacks. When some stego images are examined by the visual attack technique, embedded messages can be seen on the low-bit plane of the image; this usually happens when messages are embedded in BMP images using early information hiding techniques.

Chapter 3

Problem Statement and Objective

During literature review, some research gaps have been encountered. A problem has been formulated in an attempt to overcome some of those gaps.

3.1. Research Gaps

The research gaps observed during literature review are as follows:

- There are many techniques of steganography but the basic process and output description of the hypothesis can be explained in a simpler manner as

Table1: Information hiding applications

Capacity	**Distortion**	**Robustness**	**Imperceptibility**
Medium	Medium/High	Low	Medium/High

Most of the algorithms used for information hiding for both grey images and colored images works and provide results as shown in table1. It is clear that with so much diversity in the requirements, the design problem is very different for information hiding techniques which are intended for steganography.

- Due to distortion most of time steganography is easily been detected through steganalysis tools without even ample amount of time.

3.2. Problem Formulation

Due to recent advances in network technology, steganalysis tools had become more advanced to find vulnerability. The growth of detective steganalysis tools provide added complicities for the steganographic techniques. Steganography has very wide range of applications with high possibilities for improvements. There are two different ways for implementing steganography. First is, either to develop an application-specific solution

that utilizes specific features of the application. Other solution that utilizes the communication channel for the steganography and thus might be used on wider range of applications.

So this dissertation basically deals with how much data can be imperceptibly embedded into an image, assumption is taken that the sender and receiver have access to a shared image and that there is no modification of the communication during its transmission. These information hiding technologies also includes low distortion level with the improvement in peak signal to noise ratio.

3.3. Objectives of Thesis

To fulfill the above aim, the following goals are set for the thesis work:

- To study the existing technology of information hiding methods.
- Develop the algorithms for adaptive predictive coding while embedding secret information.
- To implement the proposed algorithm in Java, MatLab and verify results with the conventional technology.

3.4. Methodology used in Proposed Solution

Following are the steps performed to implement the above objectives:

- Algorithms and techniques are done for both the processes encoding and decoding correctly resulting improved minimum hiding ratio.
- Image used as secret information can be of both colored and grey also, it can be in any format type like jpeg, png, gif, and tiff.
- Hybrid prediction error expansion method is attained for grey images.
- Implementation is done using java coding and MatLab for the desired results for grey, color images with secret data in any form, format of image or document.
- Algorithm using high capacity information hiding with improved predictive coding has been formulated.
- Differential matrix is introduced along with the brightness matrix for each carrier image during encoding.
- At the time of decoding camouflaged brightness matrix is introduced.

- Presumptive secret keys table through the high capacity number generator attains improved prediction error.

Chapter 4
Proposed Solution

In this chapter, new methodologies along with algorithms, are presented for encoding and decoding.

This algorithm is proposed by the use of the predictive coding, a high capacity algorithm is taken into consideration to carry on the information hiding in color images. To restrain the error diffusion which usually appears during the information hiding process, an improved high capacity algorithm with predictive coding been is put forward. Through the experiments, the performance of the basic algorithm is tested, resulting into the proof of the thread correctness. As the multimedia technology is fast developing and network communication applications are widespread, the information security is concatenated with many questions, such as secret transmission, information authentication and copyright protection, became matter of concern. Therefore, these information hiding technologies arise fast and develop rapidly. Many scholars and researchers had done a lot of research on the topic of information hiding. As instance, the improvement done on the classical LSB algorithm [67], the lossless information hiding method [68, 69] based on the spatial domain and transform domain of the image proposed by Fridrich, the lossless method of data hiding [70] which is based on prediction errors expansion proposed by Thodi, the further study done on the topics like linear prediction [71, 72] carried on by Yuming Xie and Hongji Piao and so on with new improved techniques. In this research, the solution proposed will firstly interpret the predictive coding concept, and then the high capacity algorithm is used for embedding the information and finally the extraction is done taking predictive coding as its base.

4.1. PREDICTIVE CODING

Predictive coding is said as a kind of information source coding which is based on data correlation, and the reliable thought is reducing data dependency on both the domains time and space. That is explained with the values of the adjacent samples which are used to predict the value of current pixel value. After that the difference, which is represented as prediction error, between the predicted value and real sample value is been coded and transmitted. In case of the image, usually there are two approaches which are used to acquire the predictive values. One is through the prediction of the current pixel value from the value of adjacent pixel as shown in figure 4.1.a. Here the value of current pixel comes from the subtraction of previous pixels It often happens that the previous pixel value is used to predict the value of current pixel. The other method is through the prediction of the current pixel value from the adjacent block of image as presented in figure 4.1.b.

$$\hat{x} = \begin{Bmatrix} \max(a,b) & c \leq \min(a,b) \\ \min(a,b) & c \geq \max(a,b) \\ a+b-c & otherwise \end{Bmatrix} \quad \text{,…(6)}$$

Where a,b,c helps in determining pixel x and $\hat{x}$ is the prediction value of x.

a	x	B

Figure 4.1.a : Prediction context based on two adjacent bits

x	b
a	c

Figure 4.1.b : Prediction context based on three adjacent bits

4.2. THE ALGORITHM FOR COLORED IMAGES

The basic idea of the information hiding algorithm through predictive coding is to fully utilize the prediction errors for the transmission of the secret information. Here, the bits of the corresponding difference signal or the prediction error which will be obtained from the presumptive secret key table are compared with the bits of secret information i.e. the information which will be embedded to carry on the information hiding. Here, the presumptive secret key table will be established in advance and it assigns a bit for each difference possible difference.

4.2.1. Improved Algorithm

Here, the color image is used as the carrier image. Figure 4.2 illustrates the algorithm for data hiding capacity in color images.

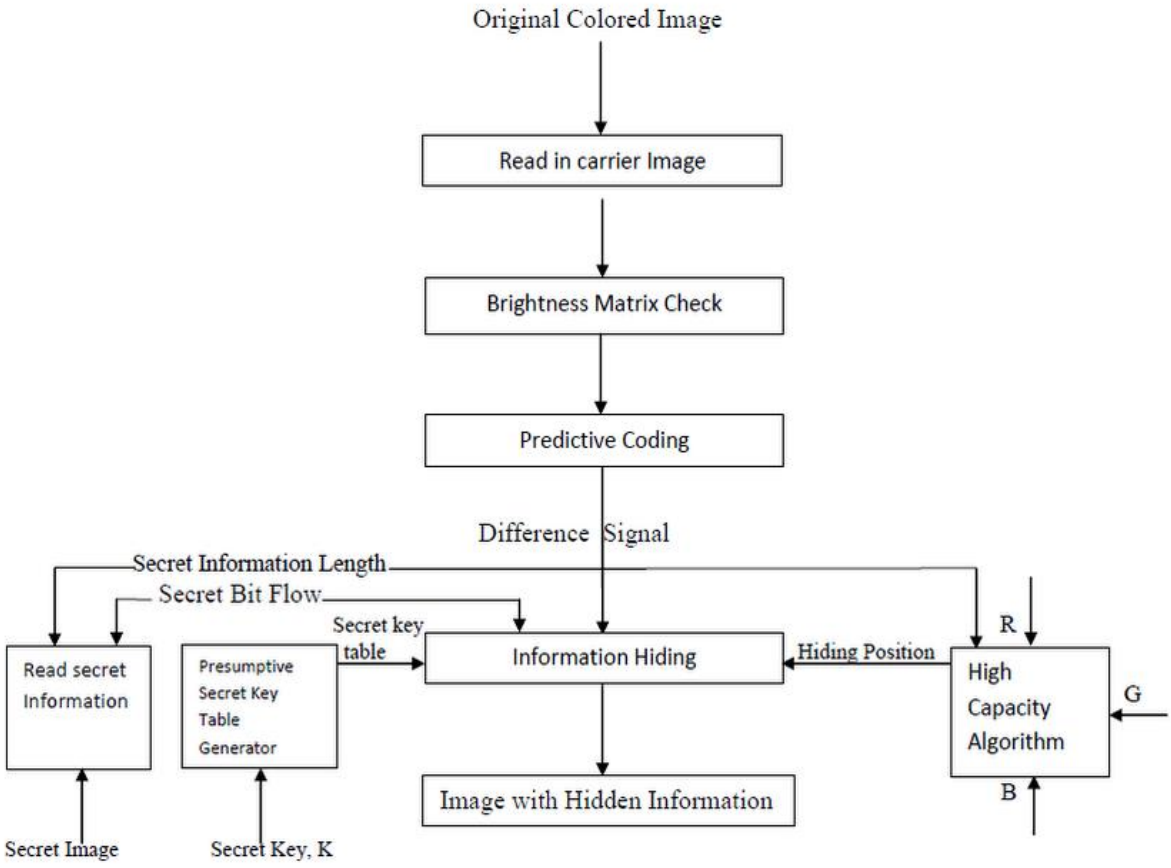

Figure 4.2 : Algorithm for high data hiding capacity in color image

1) Step 1: Read the original colored carrier image to acquire the brightness matrix.

2) Step 2: According to the results of the brightness matrix, carry on the predictive coding to gain the difference signal. The previous pixels value is used to predict the value of current pixels, thus the signal value in the previous position is subtracted from the signal value in the current position and then it gives the difference signal. Here, the matrix made from the difference signals is known as difference matrix.

$$x = \begin{Bmatrix} \left[log_2\left(\sum_{i=0}^{7} r_i \times 2^i\right)\right] \sum_{i=0}^{7} r_i \times 2^i > 0 \\ \sum_{i=0}^{7} r_i \times 2^i = 0 \end{Bmatrix} \quad ,\ldots(7)$$

$$r_i^* = \begin{Bmatrix} w_j & z \le i \le x - y \\ r_i & other \end{Bmatrix} \quad ,\ldots(8)$$

3) Step 3: Determine the position where the secret information will be embedded.

It is implemented through the high capacity algorithm which is based on the phenomenon that whenever the information is hidden through image as the carrier, the hidden information is considered as a weak signal which is superimposed on the original image acting as strong background. The visual system of human eye cannot feel the existence of the signal as long as the superimposed signal is lower than the contrastive threshold. Thus, this threshold is influenced by illumination of background, complexity of the background's texture and frequency of signal. The threshold increase with the illumination of background and the complexity of texture (or edge is rich), which is called luminance masking and texture masking [69]. The main idea of the algorithm is: the 3-color component of each pixel in 24-bit RGB color image (R_i, G_i, B_i) respectively are taken, Among them, i={7,6,5…,0}, this algorithm embed information separately into each color of every pixel 's RGB 3-color in image. At First the red component is chosen and every bit is checked in (R_7, R_6, R_5, R_4, R_3, R_2, R_1, R_0) from the high to the low. When the No. x bit is not zero, the algorithm embed information from the No. a-b bit (b>=1) until the No. c. bit. If c>a-b, this component of this pixel is not the embedded information, that is, the value after the No. i bit in (8), R（i∈{0,1,…,7}）is the pixel's red component is embedded and the bit sequence that is ready to embed is w_j （j

$\in\{1,2,\ldots,L\}$, here the value of j would be the number of bits which has been embedded. The value of j will add up by 1 whenever every 1 bit is embedded. The embedded strength is decided by the value of b. The course between imperceptibility and the embedded capability is been steered in this manner. The imperceptibility would be better when b will be having bigger value, but the hidden information should be less. The required robustness against filtering decides the value of c. If the value of c is zero, only than the LSB bit will be included. This algorithm deals with the blue and green components of this pixel in the same way as the red component. According to the characteristic of human visual system the value of b may be the different value to the blue components and green components,

4) Step 4: Input the secret key k, and get the presumptive secret keys table by the high capacity number generator. The range of the difference signals determines the length of the presumptive secret keys table.

5) Step 5: The secret information in binary bits are read in, and form the flow of secret bits which are to be hidden.

6) Step 6: Finally embed the secret information into the carrier image.

For the No.i secret bit, the corresponding Δ_i is found from the embedding position determined by high capacity algorithm from Step3. Now, do Compare the Δ_i with the corresponding bit found in the presumptive secret keys table. If it is the same bit as the secret bit, than it does not need any revision. And if it different from the secret bit, do find Δ_i which is very near to Δ_i and same as the secret bit in the presumptive secret keys table. The embedding process is explained through the following example: Let the embedded 8th secret bit is 0. Here in the hiding position information, the position obtained is row 6 and column 12. And the difference in this position is 46. The corresponding presumptive secret keys table is shown in Table 2. The bit according to 46 is 1 in the table. It is clearly different from the secret bit i.e. 0. In this type, the difference which is nearest to 46 and whose value is equal as of secret bit 0. The result is 47. So the difference in Row 6 and Column 12 is modified as 47 to hide the 8th secret bit. Repeat the same for every case till the embedding process is finished for each secret bit and information.

TABLE 2: PRESUMPTIVE SECRET KEYS TABLE

Δ	40	41	42	43	44	45	46	47	48
M	0	1	0	0	1	1	1	0	0

7) Step 7: Lastly finish hiding the secret information and save the image with secret information.

4.2.2. Decoding Algorithm

The basic process of the decoding algorithm:

1) Step1:Read the image with hidden information, and obtain the camouflaged brightness matrix.
2) Step 2: Predictive coding for the camouflaged brightness matrix is done to obtain the difference signal.
3) Step3:Use the same method of information hiding to determine the position of the secret information which is embedded.
4) Step4: Presumptive secret key table will be generated. The method used is same as that of the generation of the presumptive secret key table when hiding the secret data.
5) Step5: Extract the secret information which was embedded. For the No.i bit, examine the information (m,n) in the camouflaged position, and obtain the corresponding difference Δ_i from the difference signal. As per the presumptive secret key table, secret information bit is obtained.
6) Step 6: Secret information can be finally saved as a secret document and the process of extraction is completed.

4.3. High data hiding capacity Algorithm for grey images

4.3.1. Information Embedding Algorithm

The proposed method combines the two new techniques adaptive embedding as well as pixel selection. First according to the optimal adaptive embedded threshold divide the image pixels into two regions.

a) Flat region

b) Rough region.

Embed 2 bits into the selected flat region pixels &embed only and only one bit into the

rough region pixels .Now instead of embedding 1 bit into a pixel having large prediction error, it thought to be better to embed additionally 1bit into an already embedded pixel which has original prediction-error is sufficiently too small. This fact is been used to exploit the given pixels with relatively very small prediction errors. Here further proceedings are based on this observation, unlike to the conventional PEE which uniformly embeds only 1 bit into each pixel which is expandable, it can also embed more than one bit into the expandable pixels located into the flat regions. This process has been named as adaptive embedding.

Taking the gray-scale images as carriers, the information embedding process of the predictive coding based on image separation is described in figure 4.3.

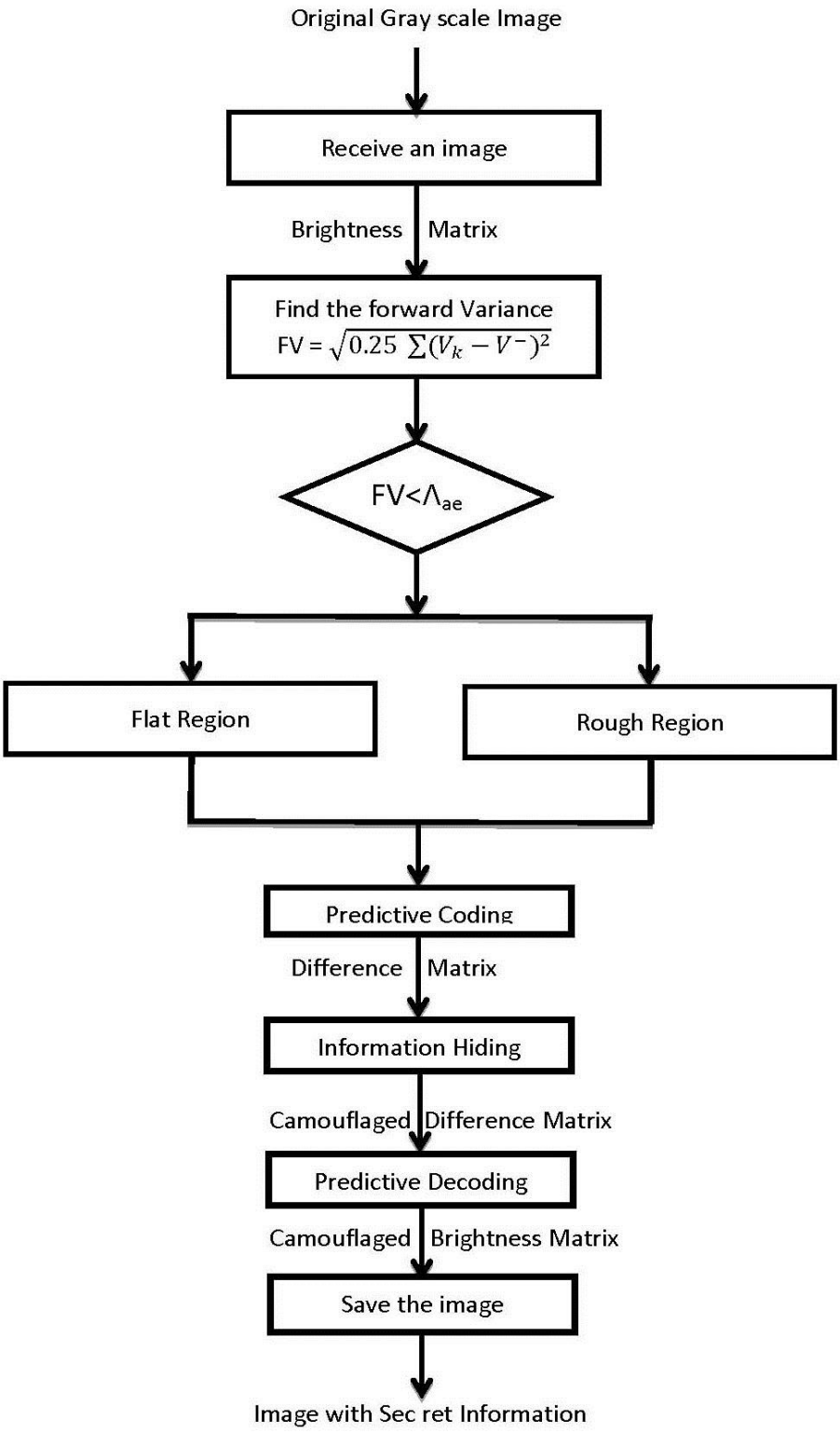

Figure 4.3 : High data hiding capacity algorithm for grey image.

1) Step 1: Accurate Carrier image is selected and undergoes through the read in process to acquire the brightness matrix.

$$\hat{x} = \begin{Bmatrix} \max(a,b) & c \leq \min(a,b) \\ \min(a,b) & c \geq \max(a,b) \\ a+b-c & otherwise \end{Bmatrix} \quad \text{,…(9)}$$

2) Step 2: Find the Variance and gradient adjusted pixel values of the neighboring bits to implement the predictive coding:

$$\text{FV} = \sqrt{0.25 \sum (V_k - V^-)^2} \quad \text{,…(10)}$$

$$d_h = \left|I_{i-1,j} - I_{i-2,j}\right| + \left|I_{i,j-1} - I_{i-1,j-1}\right| + \left|I_{i,j-1} - I_{i+1,j-1}\right| \quad \text{,…(11.1)}$$

$$d_v = \left|I_{i-1,j} - I_{i-1,j-1}\right| + \left|I_{i,j-1} - I_{i,j-2}\right| + \left|I_{i+1,j-1} - I_{i+1,j-2}\right| \quad \text{,…(11.2)}$$

Where, d_h& d_v are vertical & horizontal gradients

Value is predicted based on the bit position shown in figure 4.4:

	j-1	J	j+1	j+2
i-1	u1	u2	u3	
i	u4	Ii,j	v1	x
i+1	v2	v3	v4	
i+2	y	Z		

Figure 4.4: Adjacent bits to the real sample bit Ii,j

Value of Ii,j is calculated based on equation 12.

$$\widehat{I_{i,j}} = \begin{cases} v_1 & if\ d_v - d_h > 80 \\ \frac{(v_1 + I'_{i,j})}{2} & if\ d_v - d_h \in (32, 80) \\ \frac{(v_1 + 3I'_{i,j})}{4} & if\ d_v - d_h \in (8, 32) \\ I'_{i,j} & if\ d_v - d_h \in (8, 32) \\ \frac{(v_3 + 3I'_{i,j})}{4} & if\ d_v - d_h \in (-32, -8) \\ \frac{(v_3 + I'_{i,j})}{2} & if\ d_v - d_h \in (-80, -32) \\ v_3 & if\ d_v - d_h < -80 \end{cases} \quad ,...(12)$$

Prediction Error is then calculated as:

$$P_{i,j} = I_{i,j} - \widehat{I_{i,j}} \quad ,\ldots(13)$$

3) Step 3: Image is now been separated into two parts i.e information embedding part E and part NE (non-information embedding part).

4) Step 4: According to brightness matrix of the carrier image, carry on adaptive predictive coding for the pixels in part E, and obtain the difference matrix DE. The hybrid predictive coding is used which is attained by embedding two types of predictive techniques i.e. PEE predictive error expansion which became very much popular due to its potential to well exploit the spatial redundancy in natural images. Prediction Error expansion said as the improved version of the existing Tian's Difference Expansion PEE algorithm was originally developed by Thodi & Rodriguez [73] in 2007. Prediction-error expansion combines all the advantages of given expansion embedding along with the correlating abilities of a predictor, which results in a higher data-embedding capacity than with DE. After that no. of PEE methods developed using difference prediction algorithm In the Conventional PEE, first each pixel is predicted the earlier prediction method

$$\hat{x} = \begin{cases} \max(a, b) & c \le \min(a, b) \\ \min(a, b) & c \ge \max(a, b) \\ a + b - c & otherwise \end{cases} \quad ,\ldots(14)$$

Where a, b, and c are taken as the bits which are the right, lower, and diagonal neighbors of x.

Accordingly, the PE is calculated by e=x−x^. The PE histogram which is a Gaussian-like distribution centered at 0.

a)IF the PE belongs to the inner region, i.e., e ∈[−t, t], then take the pixel value as

x' = x^+ e'= x+ e + b,

Where e'= 2e+b

Is the expanded PE, and b∈{0, 1} is a prediction bit. As a result, one bit is embedded into this pixel.

b)If the PE belongs to the outer region then formulae will be not same as above

To minimize the distortion in prediction error expansion method, the capacity parameter is considered as the smallest integer such that the inner region can provide sufficient expandable pixels to embed the payload. To prevent the overflow/underflow (overflow, inabbreviation), some of the pixels cannot be expanded or shifted which meant that the values watermarked should be restricted into range of [0,255].

5) Step 5: Determine the secret information embedding position. It is implemented by the stochastic permutation algorithm based on the MD5 algorithm. The main idea of the algorithm is: assume each input to be i, and i is an integer less than the sum of the units embedded into the carrier. The number of Ji can be used to express the embedding required position of No.i bit in the given secret information. The step used for generating Ji is:

$$v = [i/X]; \quad \ldots(15)$$

$$u = i \bmod X; \quad \ldots(16)$$

$$v = (v+MD5(u,k1)) \bmod Y; \quad \ldots(17)$$

$$u = (u+MD5(v,k2)) \bmod X; \quad \ldots(18)$$

$$v = (v+MD5(u,k3)) \bmod Y; \quad \ldots(19)$$

$$Ji = vX+u. \quad \ldots(20)$$

Here, X and Y represent the row and column respectively.k1, k2 and k3 are three secret keys. The binary bits of secret information are read in, and form the secret bits flow waiting for being hidden.

6) Step 6: Embed the secret information in the difference matrix abbreviated as DE. It is embodied as given below:

a) Create the presumptive secret key table: Input the secret key k, and get the presumptive secret keys table by the stochastic number generator. The length of the presumptive secret keys table is determined by the range of the difference signals. Here, the brightness signals belong to type of unit 8. It ranges [0,255], and the difference ranges [-255,+255], thus the length of the presumptive secret keys table is 511. So according to the first digit of the stochastic series, the difference is -255. And according to the second digit of the stochastic series, the difference is -254. The rest may be deduced by analogy.

b) Embed the secret information: Here the information is embedded by the relation of the value of bit value in the presumptive secret table and in secret information when the difference signals are the same. For the No.i secret bit, find the corresponding Δi from the embedding position in difference matrix DE. Compare the Δi with the corresponding bit in the presumptive secret keys table. If it is the same as the secret bit, does not do any revises. But if it differs from the secret bit, find $\Delta i'$ which is nearest to in the presumptive secret keys table and replace Δi in this position in DE with $\Delta i'$ to complete the secret information hiding.

4.3.2. Algorithm for Information Extraction in Grey image

The algorithm for Information extraction in Grey images is explained in figure 4.5.

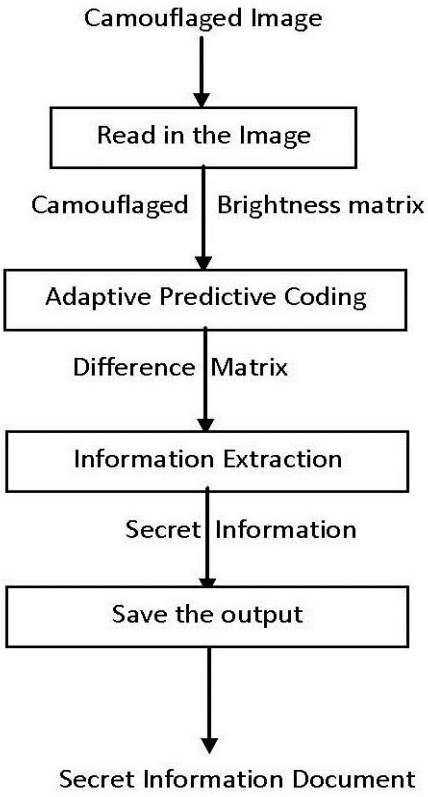

Figure 4.5: High Capacity data hiding algorithm for camouflaged images

The algorithm that follows is:

1) Step 1: Read in the image containing secret information, and then camouflaged brightness matrix is obtained.

2) Step 2: Carry on the adaptive predictive coding for the camouflaged brightness matrix, obtain the difference signal.

3) Step 3: Secret keys k1, k2 and k3, uses the same method as in information embedding process for determining the embedding position of the secret data.

4)Step 4: As per the secret key k4, use the earlier method as the generating of the presumptive secret key table, when to hide the information which gives the presumptive secret key table.

5)Step 5: Secret information is extracted . For the No.i bit, examine the information (m,n) in the camouflaged position, and obtain the corresponding difference Δi from the

camouflaged difference signal. According to presumptive secret key table, secret information bit is obtained.

6)Step 6: Hidden information is saved as a secret document to complete the extraction process.

Chapter 5

Implementation and Experimental Results

The proposed solution gives the following results:

5.1. Stego Images

Stego images are formed using the proposed solutions.

Four types of Stego images are formulated using the methods of high capacity and adaptive predictive coding. The types namely are:

a) Hiding Text into color image,

b) Hiding Text into grey image,

c) Hiding color image into color image and

d) Hiding grey image into color image

These images are then analysed in 5.2 and later.

5.1.1. Hiding text into color image

In this, a colorstego image is formed by taking a colored carrier image (Standard lena image at -34.2 db) and text is hidden into the same.

The following code is used to achieve the above given information:

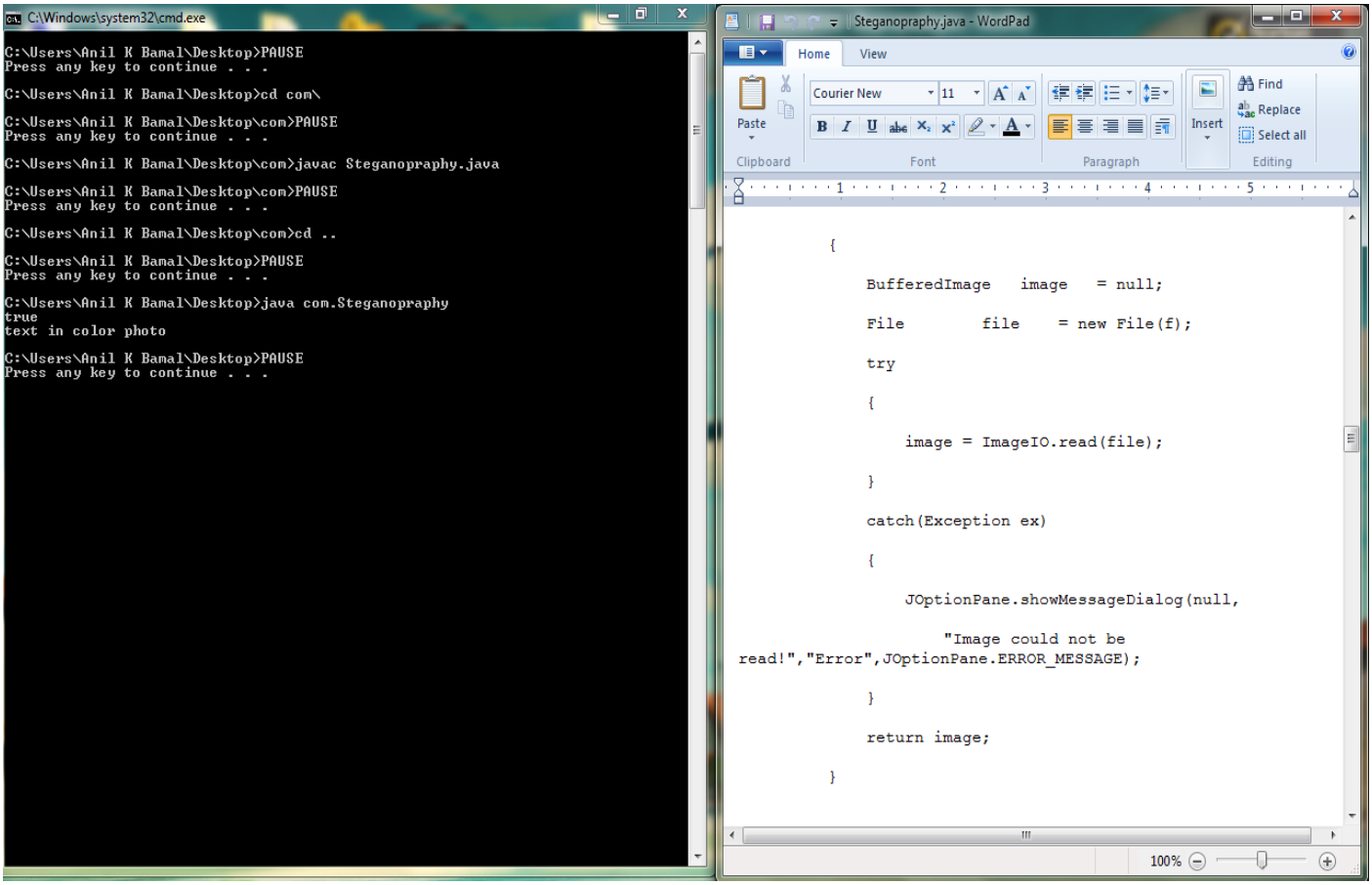

Figure 5.1: Code to hide text into colour image

Following is the implemented Lena image:

Figure 5.2.aOriginal

Figure 5.2.b Stego-Image

The resulted stego-image of standard image Lena that has been used at the embedding strength of -34.2 db is compared with the standard original image at the given value. There has some distortions in the stego image but that are mostly undetectable to human eye.

5.1.2. Hiding Text into Grey Image:

In this a stego image is formed by embedding text into a grey image.

The following code is used to implement the above:

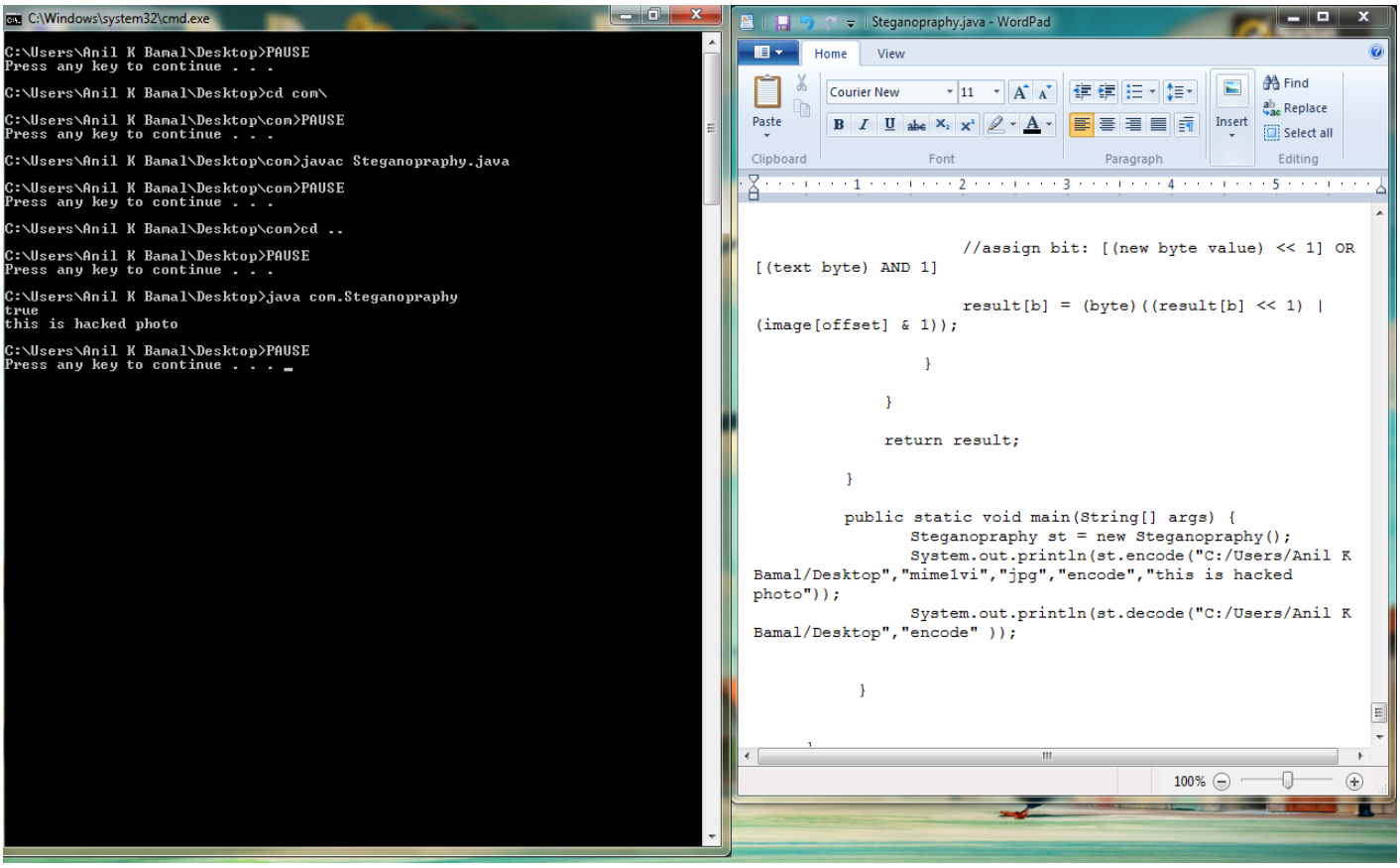

Figure 5.3: Code to hide text into grey image

Following images are used to implement the code:

Figure 5.4.a: Carrier Image

Figure 5.4.b: Stego Image

The resulted stego image has a text hidden into it. The text that is hidden is “This is hacked photo”.

5.1.3. Hiding Color Image into color image:

In this, a color image is used as secret message and is embedded into a cover message with the use of the following code:

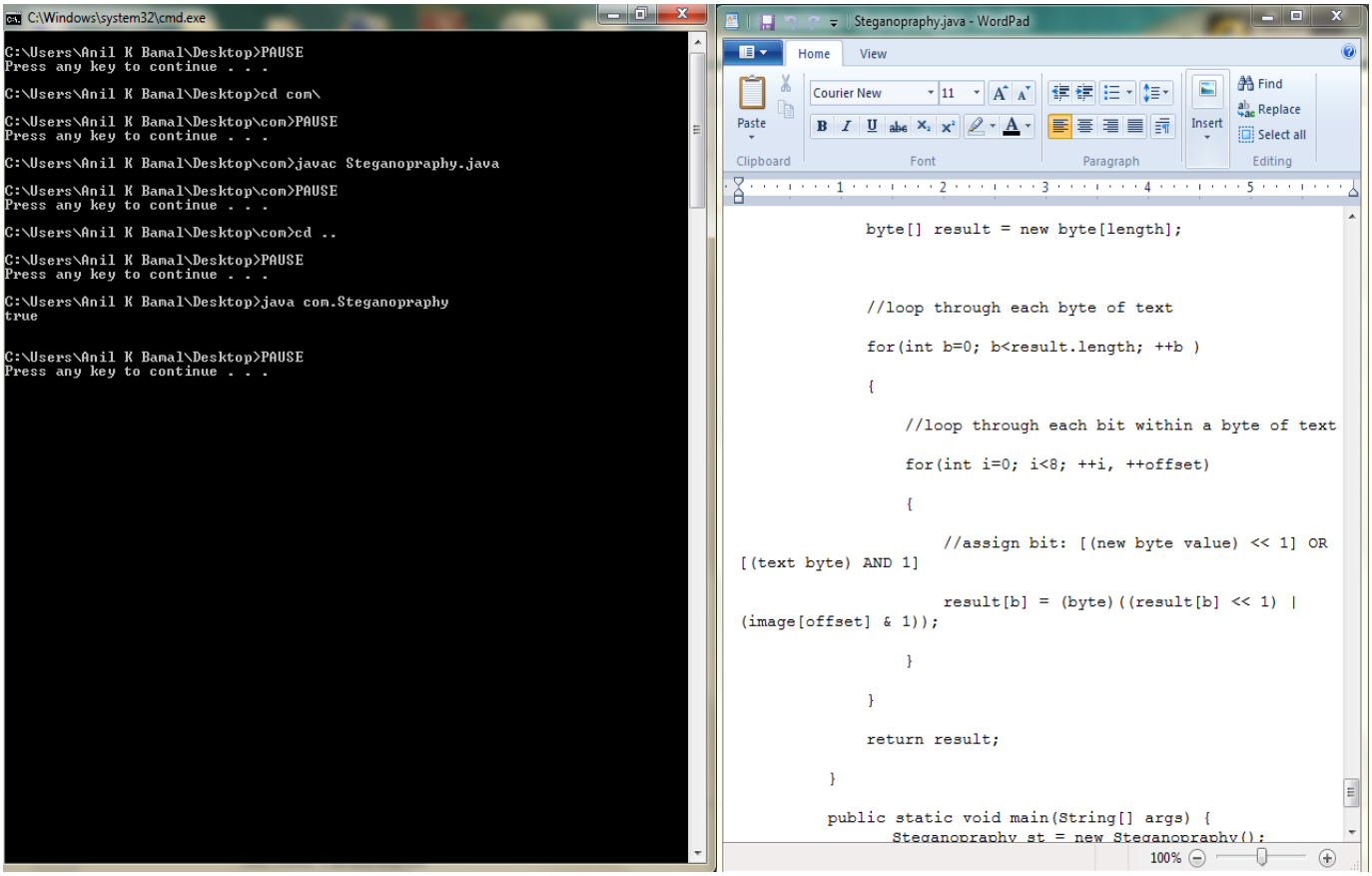

Figure 5.5: Code to hide color image into a color image

Following images are used:



Fig 5.6.a: Secret Image Fig 5.6.b: Cover Image Fig 5.6.c: Stego Image

The secret image (standard image) Peppers is hidden into the cover image Lena. The Stego Image of Lena is formed in the process which is represented in Fig 5.6.c

5.1.4. Hiding Color Image into Grey Image

In this, a color image is hidden into a grey image with the use of the following code:

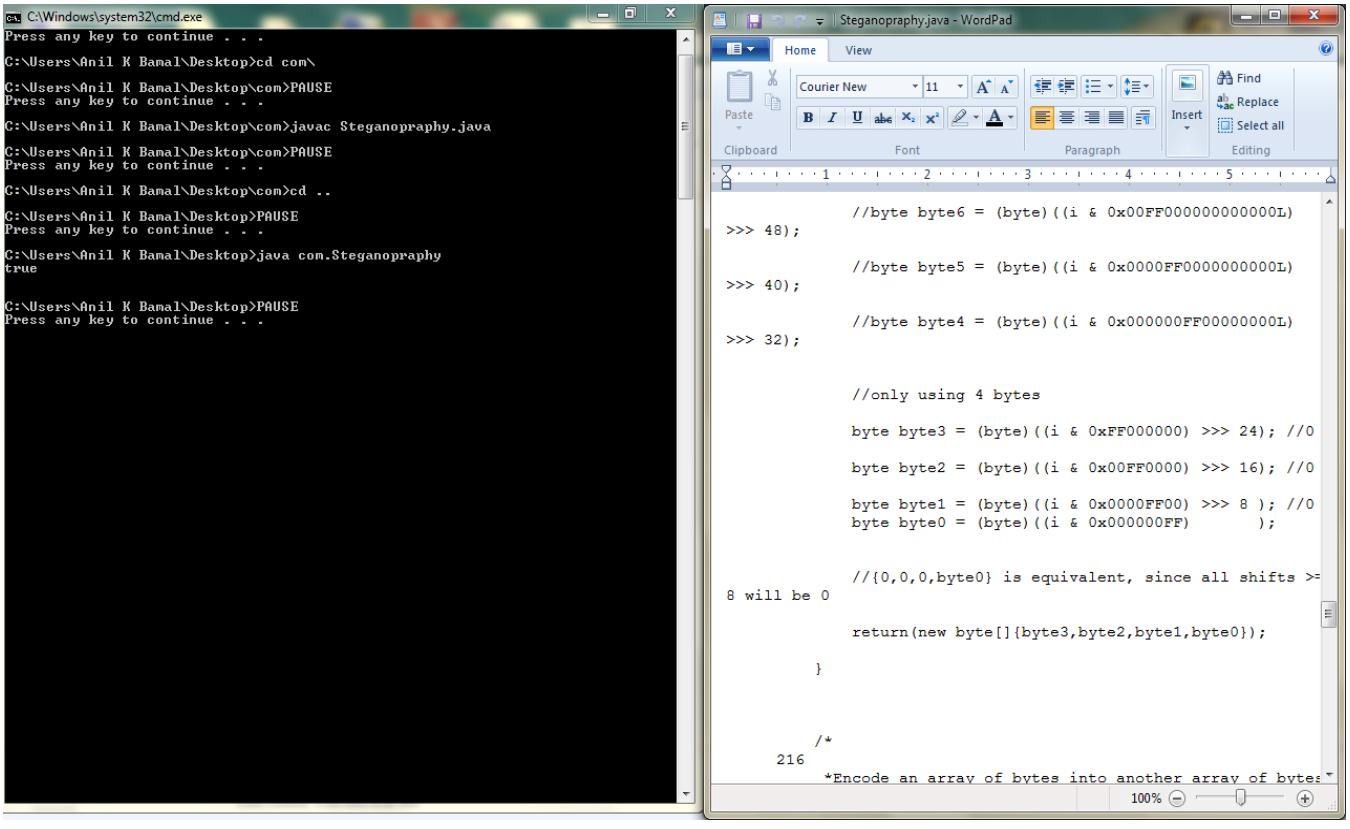

Figure 5.7: Code to hide color image into a grey image

Following images are used to implement the above code:

Fig 5.8.a Secret Image Fig 5.8.b: Cover Image Fig 5.8.c: Stego Image

The standard peppers image of smaller size is hidden into the grey image.

5.2. Prediction Error for Capacity Parameter T

Value of Prediction error at T=2.4 as compared with the occurrence of number of bits in the standard image Lena used.

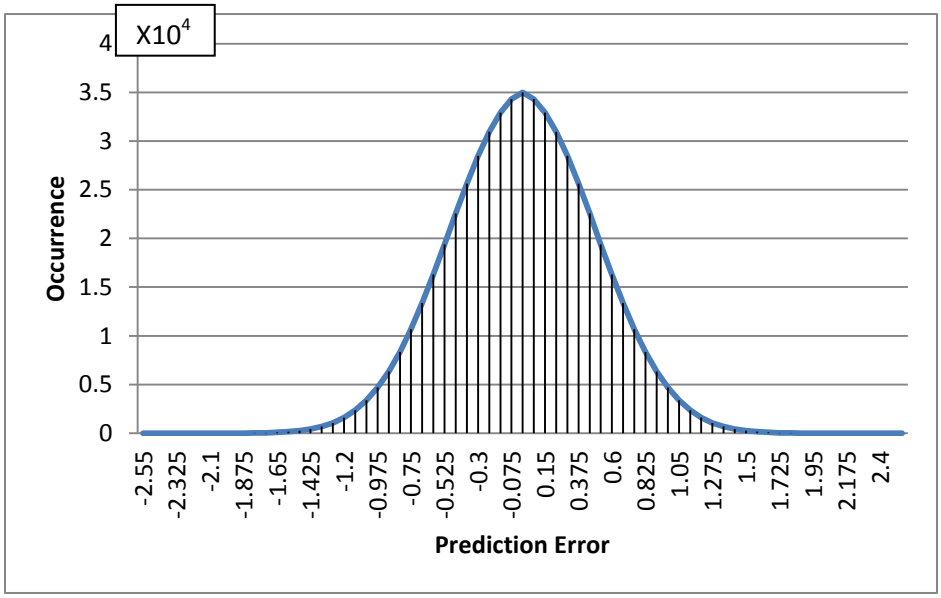

Figure 5.9: Capacity Parameter based on Prediction error with T=2.4 (Prediction Capacity)

The above image shows that maximum occurrence of bits i.e. the maximum capacity of the standard image lies between a standard T values of prediction error. Beyond the T value, the occurrence of bits is significantly low. The T region forms the inner region

which will carry the 1bit and the region outside T forms the outer region where the pixels are shifted to eliminate ambiguity.

5.3. Prediction Error for Capacity Parameter 2T

The Prediction error is further compared with the occurrence of bits in an image while using 2T=4.8

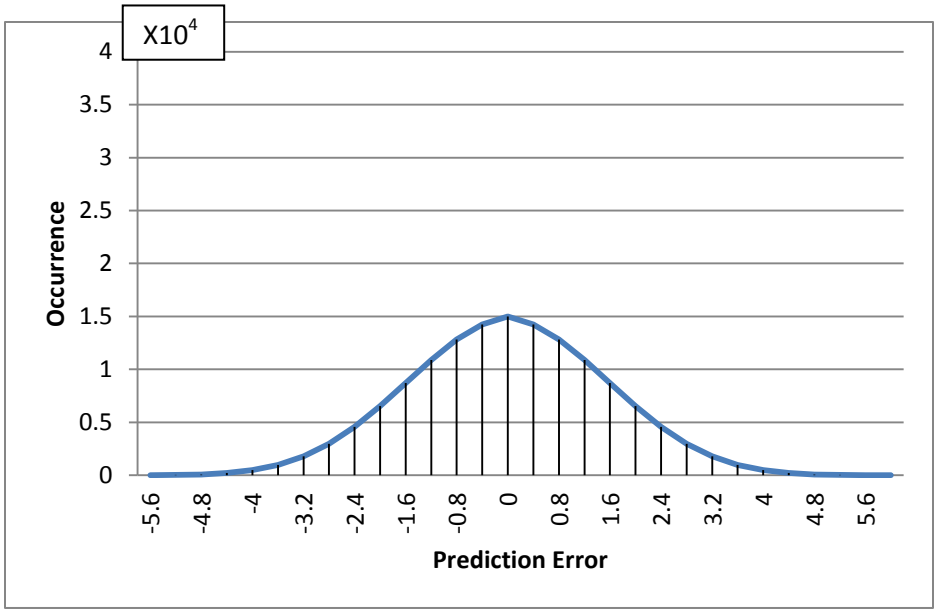

Figure 5.10: Capacity Parameter based on Prediction error with 2T (T=2.4 i.e. the prediction capacity)

The above graph is formed by doubling the standard T value region i.e. 2T which now forms an expanded inner region. This shows that expansion of inner region causes a shift in the outer region. This shift in pixel value (both expansion and shifting) leads to expansion or shift of prediction error in the prediction error histogram.

5.4. PSNR Value Calculation

The PSNR value is calculated using the MatLab software. All the stego images that are formed after embedding the secret message or image are then analyzed using Matlab to calculate their PSNR value. This value is then compared with the PSNR value of cover image that is image before embedded with hidden information. The following code is used for the same:

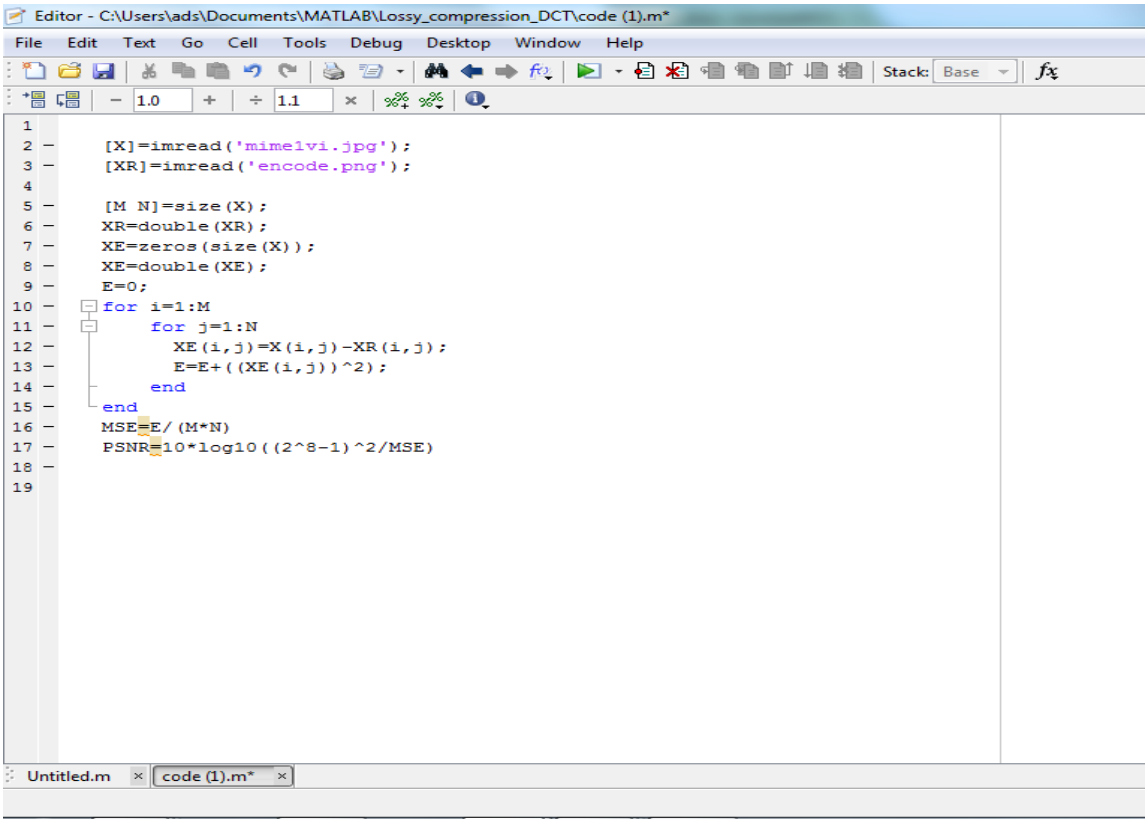

Figure 5.11: Code to calculate PSNR value for the Images

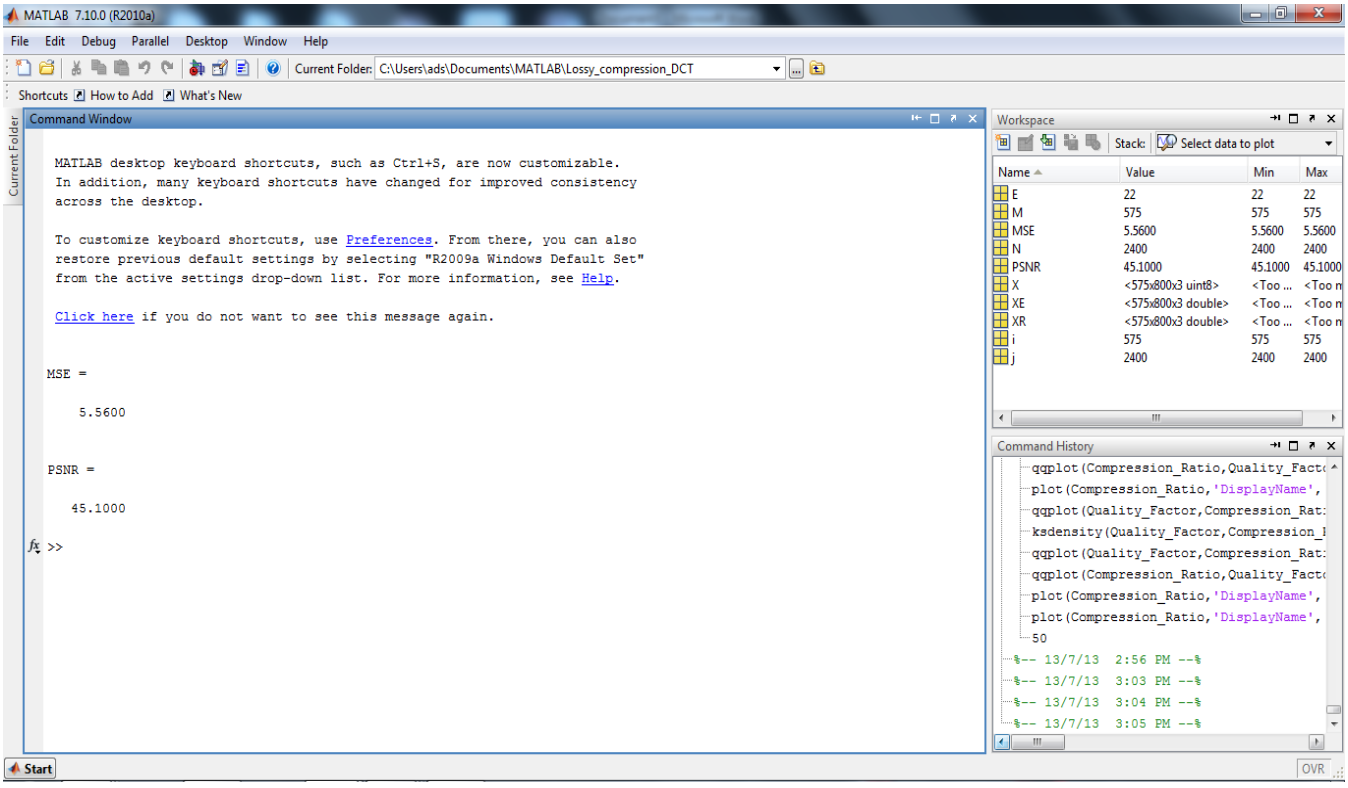

Figure 5.12: Sample PSNR and MSE values

The PSNR value is then compared between the conventional method and the proposed improved method using the standard images of Lena and Peppers.

The PSNR (Peak signal to noise ratio) values are compared for the conventional method and the improved method in order to analyze the efficiency of the improved method. The PSNR value is calculated using the MatLab software for standard images of Lena and Peppers for 10kB, 50kB and 100kB images. The following graph (figure 5.13) and table 3 gives an overview of the implemented method and values arrived.

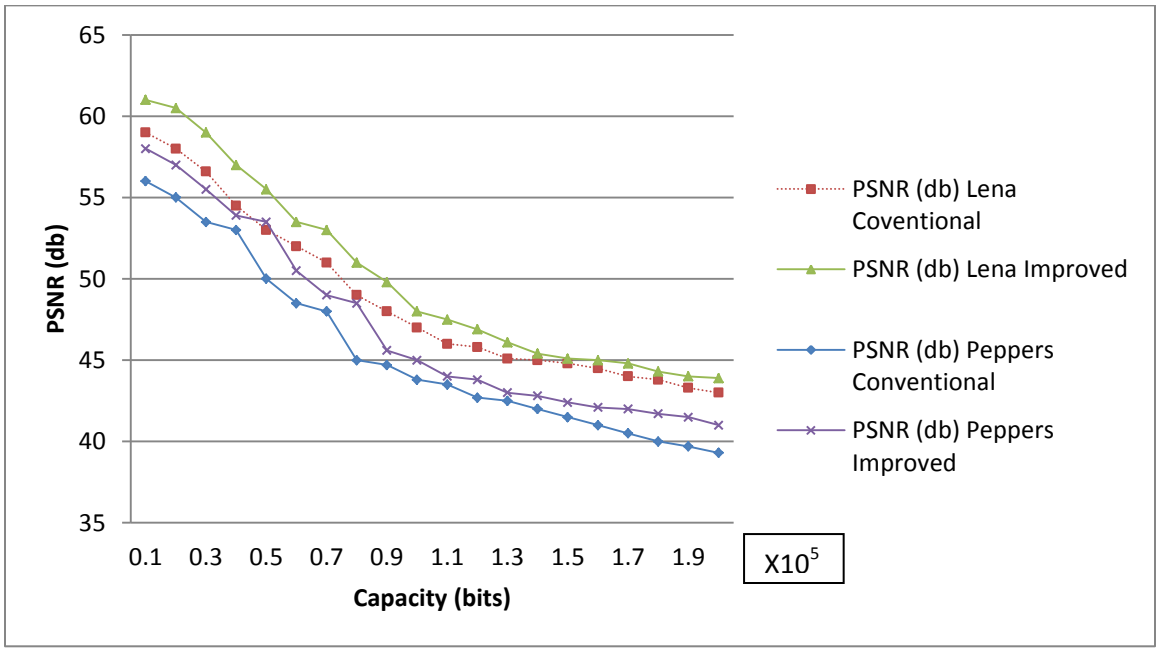

Figure 5.13: PSNR value comparison between the Conventional and Improved method

Table 3: PSNR values using conventional and improved method for bits per image

Capacity (bits)	PSNR (db) Lena Coventional	PSNR (db) Lena Improved	PSNR (db) Peppers Conventional	PSNR (db) Peppers Improved
0.1	59	61	56	58
0.2	58	60.5	55	57
0.3	56.6	59	53.5	55.5
0.4	54.5	57	53	53.9
0.5	53	55.5	50	53.5
0.6	52	53.5	48.5	50.5
0.7	51	53	48	49
0.8	49	51	45	48.5
0.9	48	49.8	44.7	45.6
1	47	48	43.8	45
1.1	46	47.5	43.5	44
1.2	45.8	46.9	42.7	43.8
1.3	45.1	46.1	42.5	43
1.4	45	45.4	42	42.8
1.5	44.8	45.1	41.5	42.4
1.6	44.5	45	41	42.1
1.7	44	44.8	40.5	42
1.8	43.8	44.3	40	41.7
1.9	43.3	44	39.7	41.5
2	43	43.9	39.3	41

For 10^5number of bits capacity image, the PSNR value is calculated using conventional and improved methods. The comparison shows that the improved method gives higher PSNR value than the conventional method thus proving improvement in the algorithm.
The improved algorithm thus has higher capacity to embed more information. Also, the improved algorithm provides prediction error which is used to calculate the difference matrix and thus giving better position information and improved hiding capacity.

5.5.The working of high capacity formula and how the value of x is calculated

$$x = \begin{Bmatrix} \left[log_2\left(\sum_{i=0}^{7} r_i \times 2^i\right)\right] \sum_{i=0}^{7} r_i \times 2^i > 0 \\ \sum_{i=0}^{7} r_i \times 2^i = 0 \end{Bmatrix} \quad (7)$$

The above eq (7) has been implemented in the following way to calculate the x value. For the bits r0 to r7 and i (0 to 7), the bit value for each pixel (R,G,B) is calculated separately to give the resultant x values which is then further used to calculate the embedding position of high capacity region.
For example, for a given set of pixels calculated from prediction error method explained in the proposed solution, value of x for R, G, B is calculated. The RGB value which is maximum among the pixels is thus selected as the high capacity embedding pixel in a colored image.

i		2	(Power of 2 based on		bit value		r*2i
0		1		r0	0		0
1		2		r1	0		0
2		4		r2	1		4
3		8		r3	1		8
4		16		r4	0		0
5		32		r5	1		32
6		64		r6	1		64
7		128		r7	0		0
						Summation	108
						Log value	6.754888
					x value is	**6.754887502**	

Figure 5.14: Display the usage of formula to calculate high capacity of a pixel (R,G,B)

5.6. Results

Following are the results of the thesis based on proposed solution:

Table 4: Result based on the Proposed Solution

Capacity	**Distortion**	**Robustness**	**Imperceptibility**
High	Medium/Low	Medium	Medium/High

CHAPTER 6

CONCLUSION

This Chapter concludes the work explained in this thesis. Some future directions have been proposed at the end of this chapter, which can be considered to implement using other compression standards.

6.1 Conclusion

Steganography through predictive coding is possible in both colored and grey images. But there is a problem of high distortion and low hiding capacity in earlier approaches. For colored images high capacity algorithm for increasing the minimum hiding ratio (MHR) is introduced. Similarly for grey images, hiding capacity has been increased through the implementation of adaptive predictive coding. The proposed algorithm improves predictive coding for the processes which takes every bit into account from neighboring bits to diagonal bits. Hence imperceptibility in also improved along with the peak signals to noise ratio(PSNR). An information hiding method based on adaptive predictive coding includes the hybrid properties of predictive error expansion and differential error. In order to restrain the error diffusion which possibly appears during the inverse predictive coding in the information hiding process, an improved adaptive predictive coding algorithm is put forward. The analysis of invisibility and hiding capacity verifies the validity of the algorithm. It also confirms that improved predictive coding algorithm is superior in performance to the basic predictive coding algorithm.

6.2 Future Scope

Some improvements still need to be made in this algorithm. Some of which are listed below:

1. In the present work, only still images are analyzed. This technique can be further extended for moving pictures and sound.
2. There is more scope of improvement in peak signal to noise ratio and minimum hiding ratio.
3. Predictive coding algorithms can be improved to reduce distortion and noise ratio for different case

BIBLIOGRAPHY

[1] R. J.Anderson and F. A. Petitcolas, “On the limits of steganography,” IEEE Journal on Selected Areas in Communications, vol. 16, no. 4, pp. 474–481, May 1998.

[2] F. A. P. Petitcolas, R. J.Anderson, and M. G.Kuhn, “Information hiding–A survey,” Proceedings of the IEEE, vol. 87, no. 7, pp. 1062–1078, July 1999.

[3] M. Barni, “What is the future for watermarking? (part 1),” IEEE Signal Processing Magazine, pp. 55–59, September 2003.

[4] F.Johnson, “What is the future for watermarking? (part 2),” IEEE Signal Processing Magazine, pp. 53–57, November 2003.

[5] P. Moulin and R. Koetter, “Data-hiding codes,” Proceedings of the IEEE, vol. 93, no. 12, pp. 2083–2126, December 2005.

[6] F. Bartolini, A. Tefas, M. Barni, and I. Pitas, “Image authentication techniques for surveillance applications,” Proceedings of the IEEE, vol. 89, no. 10, pp. 1403–1418, December 2001.

[7] E. T. Lin and E. J. Delp, “A review of fragile image watermarks,” in Proceedings of the Multimedia and Security Workshop Multimedia Contents, pp. 25–29, October 1999

[8] B. C. Emin Martinian and G.Wornell, “On authentication with distortion constraints,” in International Symposium on Information Theory, January 2001.

[9] GJ Simmons, "The Prisoners' Problem and the Subliminal Channel", in Proceedings of CRYPTO '83, Plenum Press, pp 51-67, March 1984

[10] R.J. Anderson and F.A.P. Petitcolas, “On the Limits of Steganography,” J. Selected Areas in Comm., vol. 16, no. 4, pp. 474–481, May 1998

[11] M. D. Swanson, B. Zhu and A. H. Tewfik, “Robust Data Hiding for Images”, IEEE Digital Signal Processing Workshop, University of Minnesota, pp 37-40, September 1996.

[12] W, Peter, “Disappearing Cryptography”, Information Hiding: Steganography & Watermarking (second edition), Vol 3, pp 192-213, November 1992

[13] Bret Dunbar, “Steganographic Techniques and their use in an Open-Systems Environment”, The Information Security Reading Room, SANS Institute 2002, Available Online, URL :http://www.sans.org/reading-room/whitepapers/covert/677.php

[14] Jonathan Cummins, Patrick Diskin, Samuel Lau and Robert Parlett, "Steganography And Digital Watermarking", School of Computer Science, The University of Birmingham, April 2004

[15] B. Chen and G.W. Wornell, "Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding," IEEE Trans. Information Theory, vol. 47, no. 4, pp. 1423–1443, June 2001.

[16] NIELS PROVOS AND PETER HONEYMAN, "Hide and Seek: An Introduction to Steganography", University of Michigan, IEEE Security and Privacy, August 2003

[17] M. Swanson, M. Kobayashi, and A. Tewfik, "Multimedia data-embedding and watermarking technologies," in Proceedings of the IEEE, vol. 86, pp. 1064–1087, June 1998.

[18] I. Cox, M. L. Miller, and J. A. Bloom, "Digital watermarking", San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., December 2002.

[19] N. Johnson, Z. Duric, and S. Jajodia, "Information Hiding: Steganography and Watermarking — Attacks and Countermeasures", Boston: Kluwer Academic Publishers, February 2000.

[20] S. Katzenbeisser and F. Petitcolas, "Information Hiding: Techniques for Steganography and Digital Watermarking", Norwood, MA: Artech House Books, May 2000.

[21] P. Wayner, "Disappearing Cryptography" San Francisco: Morgan Kaufmann, December 2002.

[22] Mohammad Shahreza, "Text Steganography by Changing Words Spelling", ICACT, May 2008

[23] Youssef Bassil, "A Text Steganography Method Using Pangram and Image Mediums", International Journal of Scientific & Engineering Research (IJSER), ISSN: 2229-5518, Vol. 3, No. 12, December 2012

[24] SudeepGhosh, "StegHTML: A message hiding mechanism in HTML tags", March 2007

[25] B. Pfitzmann, "Information hiding terminology", First International Workshop in Information Hiding, vol. 1174, pp. 347–350, Springer, October 1996

[26] Adnan Gutub and ManalFattani, "A Novel Arabic Text Steganography Method Using Letter Points and Extensions", World Academy of Science, Engineering and Technology, Vol. 27, September 2007

[27] S. Low, N. Maxemchuk, J. Brassil, L. O'Gorman, "Document marking and identification using both line and word shifting", Proceedings of the 14th Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 95, May 1995

[28] N.F. Johnson and S. Jajodia, "Exploring steganography: seeing the unseen." IEEE Computer Journal. Vol 31(2), pp. 26-34, February 1992.

[29] A. Cheddad, J. Condell, K. Curran and P.M. Kevitt, "Digital image steganography: survey and analysis of current methods." Signal Processing Journal. Vol 90(3), pp. 727-752, October 2010

[30] N. Johnson, "Digital Watermarking and Steganography: Fundamentals and Techniques", The Computer Journal, July 2009.

[31] I. Cox et al., "A Secure, Robust Watermark for Multimedia," Proc. First Int'l Workshop Information Hiding, Lecture Notes in Computer Science No. 1, 174, Springer-Verlag, Berlin, pp. 185-206, April 1996.

[32] E. Koch, J. Rindfrey, and J. Zhao, "Copyright Protection for Multimedia Data," Proc. Int'l Conf. Digital Media and Electronic Publishing, Leeds, UK, December 1994

[33] X-G. Xia, C.G. Boncelet, and G.R. Arce, "A Multiresolution Watermark for Digital Images," IEEE Int'l Conf. Image Processing, IEEE Press, Piscataway, N.J., September 1997

[34] Neil F. Johnson, Sushil Jajodia, "Exploring Steganography: Seeing the Unseen", George Mason University, April 1999

[35] MengyuQiao, Andrew H. Sung ,Qingzhong Liu "Feature Mining and Intelligent Computing for MP3 Steganalysis" International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing, March 2009

[36] Jossi, F., "Hiding in plain sight", Available online,
URL :http://www.wired.com/wired/archive/9.06/mustread.html?pg=9

[37] Shankaranarayanan, M., "Audio file steganography", Available Online,
URL :http://www.cise.ufl.edu/~smanamal/steganography.htm

[38] Bin Liu., Fenlin Liu., Chunfang Yang and Yifeng Sun, "Secure Steganography in Compressed Video Bit streams", The Third International Conference on Availability, Reliability and Security, April 2008

[39] L. Chun-Shien, "Multimedia security: steganography and digital watermarking techniques for protection of intellectual property", USA: Idea Group Publishing, pp. 1-253, October 2005

[40] Lin, Eugene T. and Delp, Edward J, “A Review of Fragile Watermarks”, Available online,
URL:
http://citeseer.nj.nec.com/cache/papers/cs/14065/ftp:zSzzSzskynet.ecn.purdue.eduzSzpubzSzdistzSzdelpzSzacm99zSzpaper.pdf/lin99review.pdf

[41] R.J. Anderson and F.A.P. Petitcolas, “On the limits of steganography”, IEEE Journal of Selected Area in Communications [On line]. 16(4), pp. 474-481, May 1998
URL: http://www.cl.cam.ac.uk/~rja14/Papers/jsac98-limsteg.pdf

[42] R Anderson, C Manifavas , "Chameleon - A New Kind of Stream Cipher", to appear in Proceedings of the 4th Workshop on Fast Software Encryption, June 1997

[43] A. Westfeld and A. Pfitzmann, “Attacks on steganographic systems,” in Proceedings of International Workshop on Information Hiding, ser. LNCS, no. 1768. Springer-Verlag, pp. 61–75, April 1999.

[44] K. Lee, A.Westfeld, and S. Lee, “Category attack for LSB embedding of JPEG images,” in Proceedings of the International Workshop on Digital Watermarking, ser. LNCS, vol. 4283. Springer-Verlag, pp. 35–48, October 2006.

[45] R.F. Young, “Generalized category attack - improving histogram-based attack on JPEG LSB embedding,” in Proceedings of the International Workshop on Information Hiding, pp. 378–391, September 2007

[46] X. Yu, Y. Wang, and T. Tan, “On estimation of secret message length in jsteg-like steganography,” in Proceedings of International Conference on Pattern Recognition, vol. 4. Washington, DC, USA: IEEE Computer Society, pp. 673–676, November 2004

[47] T. Zhang and X. Ping, “A fast and effective steganalytic technique against jsteg-like algorithms,” in SAC ’03: Proceedings of the 2003 ACM symposium on Applied computing. New York, NY, USA: ACM, pp. 307–311, April 2003.

[48] A. Ker, “A general framework for structural analysis of LSB replacement,” in Proceedings of 7th International Workshop on Information Hiding, ser. LNCS, no. 3727. Springer-Verlag, pp. 296–311, December 2005.

[49] J. Fridrich, M. Goljan, and R. Du, “Detecting LSB steganography in color and grayscale images,” IEEE Multimedia Magazine, pp. 22–28, October 2001.

[50] S. Dumitrescu, X. Wu, and N. D. Memon, “On steganalysis of random LSB embedding in continuous-tone images,” in Proceedings of IEEE International Conference on Image Processing, pp. 641–644, November 2002.

[51] J. Fridrich, R. Du, and M. Long, "Steganalysis of LSB encoding in color images," in Proceedings of IEEE International Conference on Multimedia and Expo, vol. 3, pp. 1279–1282, May 2000.

[52] J. Fridrich, M. Goljan, and R. Du, "Reliable detection of LSB steganography in color and grayscale images," in Proceedings of ACM Workshop on Multimedia and Security, pp. 27–30, October 2001.

[53] C. Maroney, "Hide and seek." Available Online,
URL: http://www.rugeley.demon.co.uk

[54] N. F. Johnson and S. Jajodia, "Steganalysis of images created using current steganography software," in Proceedings of the Second International Workshop on Information Hiding. London, UK: Springer-Verlag, pp. 273–289, December 1998

[55] J. J. Fridrich, M. Goljan, D. Hogea, and D. Soukal, "Quantitative steganalysis of digital images: estimating the secret message length," Multimedia Syst., vol. 9, no. 3, pp. 288–302, April 2003.

[56] D. Upham, "Jsteg." Available Online URL: http://zooid.org/ paul/crypto/jsteg/

[57] Y. Kim, Z. Duric, and D. Richards, "Towards lower bounds on embedding distortion in information hiding," in Proceedings of International Workshop on Digital Watermarking, ser. LNCS, no. 4283. Springer-Verlag, pp. 362–376, November 2006.

[58] J. Fridrich, M. Goljan, and D. Hogea, "Steganalysis of JPEG images: Breaking the F5 algorithm," in Proceedings of International Workshop on Information Hiding, ser. LNCS, no. 2578. Springer-Verlag, pp. 310–323, September 2002.

[59] J.W.Wright, "Attacking the OutGuess," in Proceedings of ACM Workshop on Multimedia and Security. ACM Press, June 2002.

[60] H. Farid, "Detecting hidden messages using higher-order statistical models," in International Conference on Image Processing, Rochester, NY, September 2002. Available Online. URL: www.cs.dartmouth.edu/farid/publications/icip02.html

[61] S. Lyu and H. Farid, "Detecting hidden messages using higher-order statistics and support vector machines," in 5th International Workshop on Information Hiding, Noordwijkerhout, The Netherlands, January 2002. [Online]. Available: www.cs.dartmouth.edu/farid/publications/ih02.html

[62] Y. Q. Shi, G. Xuan, D. Zou, J. Gao, C. Yang, Z. Zhang, P. Chai, W. Chen, and C. Chen, "Image steganalysis based on moments of characteristic functions using wavelet decomposition, prediction-error image, and neural network," in ICME, pp. 269–272, March 2005.

[63] K. Sullivan, U. Madhow, S. Chandrasekaran, and B. S. Manjunath, "Steganalysis of spread spectrum data hiding exploiting cover memory," in IS&T/SPIE's 17th Annual Symposium on Electronic Imaging Science and Technology, January 2005.

[64] K. Sullivan, U. Madhow, S. Chandrasekaran, and B. Manjunath, "Steganalysis for Markov cover data with applications to images," IEEE Transactions on Information Forensics and Security, vol. 1, no. 2, pp. 275–287, June 2006.

[65] Y. Q.Shi, C. Chen, and W. Chen, "A Markov process based approach to effective attacking jpeg steganography," in Proceedings of International Workshop on Information Hiding, May 2006.

[66] T. Pevn´y and J. Fridrich, "Merging Markov and DCT features for multi-class jpeg steganalysis," in Proceedings of SPIE Electronic Imaging, January 2007.

[67] Bing Fu, "A method for raising the hiding of LSB information", Journal of Yangtze University (Natural Science Edition), ,vol.3, pp. 73-75, December 2006

[68]Fridrich J, Goldjan M, Du R., "Invertible authentication", Proceedings of SPIE. San Jos, California, vol. 4314, pp.197-208, January 2001.

[69] Goldjan M, Fridrich J, Du R., "Distortion-free data embed-ding", Proceedings of the 4th Information Hiding Work-shop, Pittsburg, PA,USA, pp.27-41, April 2001,

[70]Thodi D M, Rodrfguez J J., "Reversible watermarking by prediction- error expansion", *Proceedings of the 6th IEEE* Southwest Symposium on Image Analysis and Interpretation, Lake Tahoe, Nevada, vol.6, pp.21-25, March 2004.

[71] Yuming Xie, Yimin Cheng, and Yixiao Wang, "Lossless Data Hiding Method in Image Based on Linear Prediction", Journal of Computer- Aided Design & Computer Graphics, China Computer Federation, Beijing, vol.18, pp. 585-591, April 2006

[72] Hongji Piao, Pin Zheng, Xiong Tian, "Data Hiding Algorithm Based on Linear-Prediction and Bit-Operation", Computer Technology and Development, China Computer Federation Shanxi Province branch, Xi'an, Shanxi, vol.18, pp.185-187, January 2008

[73] D.M. Thodi and J.J .Rodriguez, "Expansion embedding techniques for reversible watermarking," IEEE Trans. Image Process, vol. 16, no.3, March 2007

List of Publications

[1] Roopam Bamal, Dr. V. P. Singh, "Steganography : A modern day Art and Science for Data Hiding", accepted by International Journal of Latest Research in Science And Technology, July 2013.

[2] Roopam Bamal, Dr. V.P. Singh, "High Capacity Steganographic Algorithm with Predictive Coding", communicated to International Journal of Computer Applications, July 2013

[3] Roopam Bamal, Dr. V.P. Singh, "Steganographic Algorithm based on Adaptive Predictive Coding", communicated to International Journal of Computer Science and Software Engineering, July 2013