Reversible Image Data Hiding Using Predictive Coding Technique Based on Steganograpic Scheme

Ananthi S[1], Anjanadevi A[2]
[1]*Assistant Prof. /Dept of IT, M.A.M. College of Engineering*
[2]*Assistant Prof. / Dept of ECE, M.A.M. College of Engineering*

Abstract: *- Steganography is the science that involves communicating secret data in an appropriate multimedia carrier, e.g., image, audio, and video files. This scheme provides the ability to hide the data into a host image and then recover the host image without losing any information when the secret data is extracted. The reversible image steganographic scheme based on predictive coding is proposed by embedding secret data into the error values during the lossless image compression. This technique embeds secret data into the error values of that image. During the predictive coding stage, this scheme embeds secret data into error values by using Median Edge Detective (MED) predictor. In decoding stage, the secret data can be extracted by referring to the Error values and the host image can be recovered during the predictive decoding stage. The host image can be reconstructed without losing any information when the secret data is extracted. In Steganograpy data hiding provides protection against detection and also provides protection against removal. The existence of the message is hidden secret so that it can avoid the removal of content from the unauthorized users. The most frequent hacking techniques can be avoided by using this predictive coding technique*

Keywords: *- Cryptography, Steganography, Steg-analysis, Data hiding, Reversible Image, Irreversible Image, Least Significant Bit, Predictive Coding, Attacks.*

I. INTRODUCTION

Cryptography is the art of changing the plain text (Original message) to Cipher text by using the key value. It hides the contents of a secret message from a malicious people. In cryptography the content of the message alone is kept secret. The existence of the message is not kept secret. The structure of a message is scrambled to make it meaningless. While seeing this meaningless information hackers came to know the existence of the message. The system is broken when the attacker can read the secret message [1]. Data hiding conceals the existence of secret messages while cryptography protects the content of message.

The word steganography comes from the Greek *Steganos*, which mean covered or secret and *graphy* mean writing or drawing ie., the art of hiding information in ways that prevent detection[2],[3]. Steganography is the process of hiding a secret message within another message. Steganography can be an invaluable tool in maintaining confidentiality, which is based on the computer security, integrity and availability [4]. Steganography can be useful when the use of cryptography is forbidden: where cryptography and strong encryption are outlawed, steganography can circumvent such policies to pass message covertly. However, steganography and cryptography differ in the way they are evaluated: The disciplines that study techniques for deciphering cipher messages and detecting hide messages are called *cryptanalysis* and *steganalysis* [3],[4].

Steganography technique is not a new technique. They are some older practices in message hiding such as invisible ink, tiny pin punctures on selected characters and pencil mark on typewritten characters. In term of the key management, steganography is more secure than cryptography. If The goal of steganography is to hide secret message into the cover image and it generates the stego-image[5].

The goal of steganography is to avoid drawing suspicion to the existence of a hidden message. This approach of information hiding technique has recently become important in a number of application areas. Digital audio, video, and pictures are increasingly furnished with distinguishing but imperceptible marks, which may contain a hiding copyright notice or serial number or even help to prevent unauthorized copying directly[5].

Reversible image has the ability to reconstruct the secret information from the cover image without losing of any information to both cover images as well as to the secret data. A basic steganographic model is shown in Fig 1.1. The message ‘M’ is the secret data that the Sender wishes to hide without any suspicion. The secret data can be audio, video, image, text. The cover ‘X’ is the original image, audio file, video file, in which the secret message ‘M’ is to be embedded. The cover ‘X’ is also called as “Message Wrapper”. It is not necessary that the cover ‘X’ and the message ‘M’ should have homogeneous structure. For example, text message or an audio file can also be hidden into video or image. The cover ‘X’ and Message ‘M’ are images.

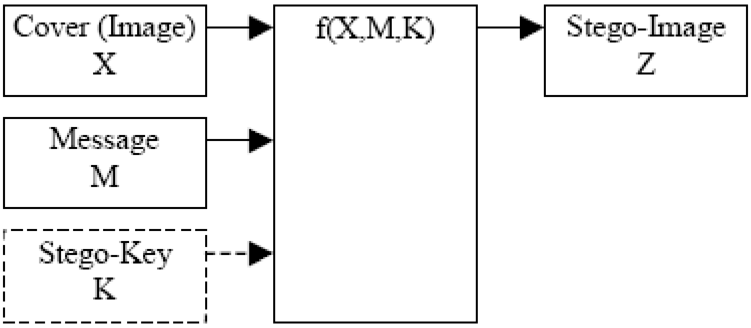


Figure 1.1 basic digital steganography encoder

Stego – Image 'Z' is basically the image in which the secret image 'M' is embedded. It should be ensured that at any point, the stego-image should resemble the cover image else it will cause suspicion. Stego-key 'Z' is provided to the receiver so that only he can be able to extract the secret image from the cover image[6].

Image steganography can simply be divided into two types. The first type is an irreversible image steganographic scheme which embeds the secret data into the host image in order to form the stego-image. The host image encounters a slight destruction because of the hidden secret data. Consequently, the stego-image cannot recover the host image without losing information when the secret data is extracted from the stego-image. There are many irreversible image steganographic schemes which have been proposed The advantage of an irreversible image steganographic scheme is that it provides a high capacity for embedding secret data.[5],[6]. For example, the least significant bit (LSB) substitution scheme can provide the embedding capacity in 1 bit per pixel (hereinafter referred to as bpp)–4 bpp for each host image [7]. Least Significant Bit (LSB) technique provides the ability to embed the data into the least bit [8]. So the removal of all least significant bit will destroy the information hidden over the cover image.

The second type is the reversible steganographic scheme which involves embedding secret data into the host image to create the stego-image which can produce a lossless recovery of the host image when the secret data is extracted. There are many reversible steganographic schemes that have been proposed [9]. A reversible scheme which compresses one of the least significant bit planes of the host image, and then appends the secret data to compress the bit planes and encrypt it. Finally, this scheme replaces the host image bit plane with the encrypted bit plane [9], [10]. However, the capacity of this scheme is low. Then it is improved by proposing a reversible scheme, and also utilizing the concept of quantization and lossless compression in order to provide low-distortion and high capacity. A reversible scheme by using difference expansion for a pair of pixels [9]. This scheme can embed several bits in each vector. Normally, the host image can simply be divided into two formats: the raw format and the compressed format. There are many image steganographic schemes which hide secret data into raw format, but they have a large image size. It does not suit transmission on the Internet because the bandwidth of data transmission is limited. Most images will use lossy-compression or lossless compression in order to generate a compressed form for data transference [10].

The image steganographic schemes that have been proposed to hide the secret data into compression codes. The benefit of embedding secret data into compression codes provides a small image size and saves data transmission time, it proposes a high capacity image steganographic scheme based on the concept of lossless and near-lossless image compression. This scheme embeds the secret data into the error values during the predictive coding process [11].

In this work, a reversible image steganographic scheme has been proposed which improves upon the vulnerability of the slight destruction of the host image encountered with the scheme of Yu et al. The proposed scheme embeds the secret data into the error values when the host image is in the predictive coding stage [12]. The proposed scheme also refers to an error values in order to extract the secret data and provide lossless recovery of the host image in the decompression stage [12].

II. EXISTING SYSTEM

Least Significant Bit (LSB) insertion is a common simple approach to embedding information in a cover file. Unfortunately it is vulnerable to even a slight image manipulation. Converting an image from a format like GIF or BMP which reconstructs the original message exactly. But while using format like JPEG, which does not reconstruct back the original message, and could destroy the information hidden in the LSB's [12].

LSB requires that only half the bits in an image be changed. It can hide data in the Least significant bits (i.e., it divides the bytes into 8 bits each and data is hidden the Least significant bit) and still the human eye would not be able to discern it [5], [8], [10].

Even though LSB provide security it is not robust. Sensitive to any kind of filtering. Hackers can destruct the messages by removing or zeroing the LSB. While removing the LSB it will not affects the quality and it is not identified by the end user. So the content could be destroyed [12].

In Irreversible image data hiding, while reconstructing the secret image from the Stego-image there is some loss in the Cover image [9].

III. PROPOSED SYSTEM

In this work, we introduce the Median Edge Detector (MED) predictor. The MED predictor is commonly used for lossless and near-lossless image compression techniques based on Reversible Image Steganographic scheme.

3.1 Embedding Procedure

The MED predictor is used for the LOw-COmplexity LOssless COmpression for Image (LOCO-I) scheme. LOCO-I is the core algorithm for the lossless and near-lossless image compression of ISO/ITU standard. The MED predictor is used in the predictive coding stage. For each image pixel, the predictive values are generated by a MED predictor, which is called the predictive image.

By finding the difference between the original image and the predictive image, the error values are generated, and then coded in the entropy coding stage. The predictive template shown in fig 1 uses neighboring pixels to generate the predictive value. *x* is the current pixel designated to be predictive, and *a*, *b* and *c* are neighboring pixels of *x*. The MED predictor uses past data, *a*, *b* and *c*, in order to detect vertical or horizontal edges in the predictive template.

When a vertical edge appears on the left side of *x*, the MED predictor will use *b* as the predictive value.

When a horizontal edge is on the upper side of *x*, the MED predictor will use *a* as the predictive value. If no edge appears in the predictive template, the MED predictor will use $a + b - c$ for the predictive value.

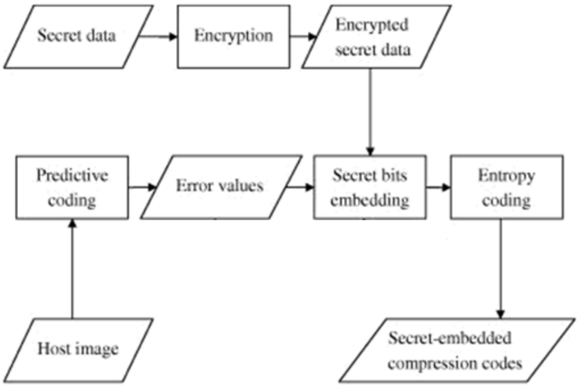


Figure 3.1.1 embedding procedure

C	***B***
A	***X***

Figure 3.1.2 host image template

If x' is the predictive value of *x*, the predictive rule can be represented as

C	***B***
A	***x'***

Figure 3.1.3 predictive image template

Predictive images are calculated by using the below formula for N X N pixels in the selected block.

$$: x' = \begin{cases} \mathbf{Min\ (a,b)} & \mathbf{if\ \ c \geq Max\ (\ a,\ b)} \\ \mathbf{Max\ (a,b)} & \mathbf{if\ \ c \leq Min\ (a,\ b)} \\ \mathbf{a + b - c} & \mathbf{otherwise} \end{cases} \qquad \mathbf{(1)}$$

Where max (*a*, *b*) and min (*a*, *b*) are the functions that find the maximum and minimum values of *a* and *b*, respectively.

In the recovery stage, the predictive image also can be generated by the same MED predictor with the error values of entropy decoding. The original image can be recovered by adding the predictive image and error values[12].

An image-hiding scheme via predictive coding that is based on the concept of lossless and near-lossless image compression. They embed the secret data by modification of error values during the predictive coding stage. In their scheme, each error value can embed three secret bits at most [11].

Table 3.1.1 Example of HI and PI

164	159	155	149
209	127	119	118
166	169	170	175
207	205	200	198

164	-5	-4	-6
45	-77	-8	-1
-43	42	9	6
41	-2	-5	-2

Table 3.1.2 Example of EV

0	164	159	155
164	204	127	119
209	127	161	169
166	207	205	200

Table 3.1.3 Example of AEV

164	5	4	6
45	77	8	1
43	42	9	6
41	2	5	2

Steps for Embedding:

1. Choose the Host or Cover image in which the secret data has to be embedded.

   Let *HI* be the gray-scale host image with *N* × *N* pixels represented as,

   $$\mathbf{: HI = \{h_{ij}\ \vert\ 0 \leq i < N,\ 0 \leq j < N,\ h_{ij}\ \text{€}\ \{0,1,\ldots.225\}} \qquad \mathbf{(2)}$$

2. Calculate the Predictive value by using eqn. (1)

   Let *PI* be the predictive image with *N* × *N* pixels represented as

   $$\mathbf{: PI = \{p_{ij}\ \vert\ 0 \leq i < N,\ 0 \leq j < N,\ p_{ij}\ \text{€}\ \{0,\ 1,\ldots.,\ 225\}} \qquad \mathbf{(3)}$$

3. The encrypted secret data is embedd into the error value. Error values are the differences between HI and

PI.

Let *EV* be the Error value set for N X N error value represented as,

$$: \mathbf{EV = \{e_{ij} \mid 0 \leq i < N, 0 \leq j < N, e_{ij} € \{-255,-254, ..., 225\}} \quad \mathbf{(4)}$$

Where,

e_{ij} - error value.

$$\mathbf{e_{ij} = h_{ij} - p_{ij}.}$$

4. The round of value of the Error value is calculated by AEV

The range of each e_{ij} is [−255, 255]. AEV be the Absolute Error Value,

$$: \mathbf{AEV = \{a_{ij} \mid 0 \leq i < N, 0 \leq j < N, a_{ij} € \{0, 1, ..., 225\}} \quad \mathbf{(4)}$$

Where,

$$\mathbf{a_{ij} = \mid a_{ij} \mid}$$

5. In these Error value the secret data is embedding.

There are two basic ways to manipulate GIS videos images for hiding data. The first class of approaches changes low-level features such as flipping a black pixel to white or vice versa. The second class of approaches changes high-level features such as modifying the thickness of strokes, curvature, spacing, and relative positions. Since the number of parameters that can be changed by the second class of approaches is limited, especially under the requirements of invisibility and blind detection (i.e., without using the original image in detection), the amount of data that can be hidden is usually limited except for special types of images [12].

3.1.1 Embedding of Secret Message

Directly encoding the hidden information in flippable pixels (e.g., set to black if to embed a "0" and to white if to embed a "1") may not allow the extraction of embedded data without the original image. The reason is that the embedding process may change a flippable pixel in the original image to a pixel that may no longer be considered flippable. As a simple example, suppose only black pixels that are immediately adjacent to white pixels are considered as "flippable"[4].

One such flippable pixel, marked by thick boundary in Fig. 1, is changed to white to carry a "1", as shown in Fig. 1. It can be seen that after embedding, this pixel is no longer considered flippable if applying the same rule. This simple example shows the difficulty for the detector to correctly identify which pixel carries hidden information without using the original image [7].

3.2 Extraction Procedure

The hidden data in the Image or Video or Audio is extracted by referring the Absolute Error Value (AEV). During the predictive decoding the hidden data's are extracted and MED Predictor is applied to regenerate the original information. While extracting the hidden data there is no loss in the original information by using the Reversible Image steganography.

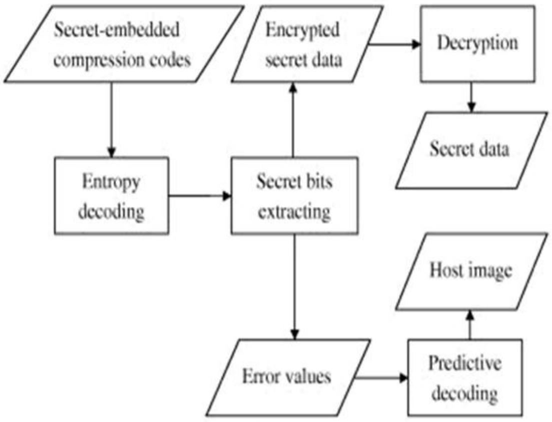
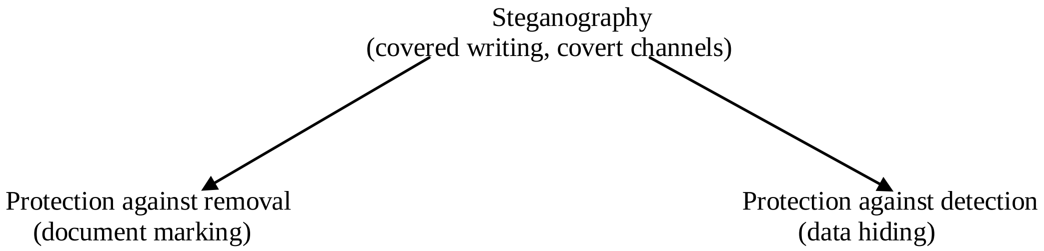


Figure 3.2.1 extraction procedure

IV. Detection and Removal against Protection

Steganography provides both protection against removal as well as protection against detection [13]. In predictive coding technique the protection against detection of the secret message from the stego-image is achieved, because of the existence of message is hidden secret in steganography. In LSB substitution technique the hackers can easily removing or zeroing the least bit will destroy or remove the secret information hidden over the stego-image [8]. But in this Predictive Coding technique it prevents the information against detection as well as removal of secret data.


Steganography
(covered writing, covert channels)
Protection against removal
(document marking)
Protection against detection
(data hiding)

V. Types of attacks

5.1 Stego-only attack

Only the stego-object is available for analysis. For example, only the stego-carrier and hidden information are available.

5.2 Known cover attack

The original cover-object is compared with the stego-object and pattern differences are detected. For example, the original image and the image containing the hidden information are available and can be compared.

5.3 Known message attack

A known message attack is the analysis of known patterns that correspond to hidden information, which may help against attacks in the future. Even with the message, this may be very difficult and may be considered the same as a stego-only attack.

5.3 Chosen stego attack

The steganography tool (algorithm) and stego-object are known. For example, the software and the stego-carrier and hidden information are known.

5.4 Chosen message attack

The steganalyst generates a stego-object from some steganography tool or algorithm of a chosen message. The goal in this attack is to determine corresponding patterns in the stego-object that may point to the use of specific steganography tools or algorithms.

5.5 Known stego attack

The stegonography tool (algorithm) is known and both the original and stego-object are available.

5.6 Passive Attack

A passive attack involves detecting the use of steganography and is a prelude to actually deciphering the hidden message[14]. Methods of steganalysis include:

- Viewing the file
- Listening to the file
- Performing comparisons on a file (if you have the original)
- Statistical Attack – this involves detecting changes in the patterns of
- Pixels of Least Significant Bits.
- Signature

Obviously the first two methods of analysis will not often yield accurate results. The purpose of steganography is that the changes are well hidden. Therefore simply viewing or listening to the file is not meant to reveal anything other than what the file appears to be. The first four methods in fact involve performing comparisons against the original file. While this can often indicate that one file is a carrier and therefore be successful, if steganography is used 'The Right Way' then an attacker will rarely have access to the original unmodified file. If a person is intending to use steganography to hide a secret message, it would be foolish to use a well known or readily available image or sound file to conceal the message in. Sensibility suggests that one would use a file preferably never before seen by anyone else or at the very least, chosen from an inconspicuous location on the Internet [14].

Hence, the embedding of data takes place only by referring the Absolute Error Value. So unauthorized users can't identify where the actual content is embedded. So all these attacks can be predicted in this predictive coding technique.

VI. CONCLUSION

The data has been compressed, encrypted, hidden without any visible increase in file size. The data containing cover image also reconstructed without losing any information. The reversible data hiding in the file is thus implemented objectively. The data retrieval after successful decryption and decompression is also achieved. This work proposes a reversible image steganographic scheme which is based on predictive coding. The secret data is encrypted by a secret key and is embedded in modified or absolute error values. By referring to the absolute error value, the embedded secret data can be extracted from modified error values, and the host image also can be lossless recovered. On the one hand, compared with the existing reversible schemes in our experimental results, our scheme has the highest embedding capacity. By referring to the error value, scheme can provide a large embedding capacity and can recover the host image without losing any information when the secret data is extracted. The unauthorized users can't identify where the actual content is embedded. Prevention of data or image against detection and removal is achieved. Stego attacks can be avoided by using this technique.

REFERENCES

[1] B. Schneier, "Applied Cryptography," Wiley, New York, *1996.*

[2] N. Provos, P. Honeyman, "Hide and seek: an introduction to steganography", *IEEE Security and Privacy Magazine 1 (2003) 32–44.*

[3] Ki-Hyun Jung, Kee-Young Yoo, "Data Hiding method using Image Interpolation", Republic of Korea.

[4] Armin Bahramshahry, Heasm Ghasemi, Design of Data Hiding Application using Steganograhy, *April 2007.*

[5] Shaifizat Mansor, Roshidi Din & Azman Samsudin, "Analysis of Natural Language Steganography", *International Journal of Computer Science and Security (IJCSS), Volume (3) : Issue (2).*

[6] Shrikant Khaire, sanjay l. nalbalwar, "Review: Steganography – Bit Plane Complexity Segmentation (BPCS) Technique", *International Journal of Engineering Science and Technology Vol. 2(9), 2010, 4860-4868*

[7] R.Z. Wang, C.F. Lin, J.C. Lin, " Image hiding by optimal LSB substitution and genetic algorithm ", *Pattern Recognition 34 (2001) 671–683.*

[8] J. Mielikainen, " LSB matching revisited ", *IEEE Signal Processing Letters 13 (2006) 285–287.*

[9] M.U. Celik, G. Sharma, A.M. Tekalp, E. Saber, "Reversible data hiding", *Proceedings of the IEEE International Conference on Image Processing (2002) 157–160.*

[10] A.M. Alattar, "Reversible watermark using the difference expansion of a generalized integer transform", *IEEE Transactions on Image Processing 13 (2004) 1147–1156.*

[11] C.C. Thien, J.C. Lin, "A simple and high-hiding capacity method for hiding digit by- digit data in images based on modulus function", *Pattern Recognition 36 (2003) 2875–2881.*

[12] Y.H. Yu, C.C. Chang, Y.C. Hu, "Hiding secret data in images via predictive coding", *Pattern Recognition 38 (2005) 691–705*

[13] Jonathan Cummins, Patrick Diskin, Samuel Lau and Robert Parlett,"Steganography and Digital watermarking",School of Computer Science, The University of Birmingham.

[14] Lachlan McGill, "Stegnograpy Steganography: The Right Way", SANS Institute Reading Room site

Asst.Prof S. Ananthi received her B.E Degree in Computer Science from Dhanalakshmi Srinivasan Engineering College, Perambalur and M.E Degree in Computer Science from Annamalai University, Chidambaram. Currently, She is working as an Assistant Professor in M.A.M College of Engineering, Siruganur, Trichy. She had presented a paper about Steganography in National Conference on Multimedia Signal Processing, NCMSP' 2011 in Annamalai University on 2011. Her Paper got Selected in the National Conference on Information Security (NCIS-2012) held at SASTRA University, Thanjavur, during June 14-15, 2012. She participated in the National Conference on Confluence of Multidisciplinary in Engineering Fields NCCME'12. She had published a paper in International Journal of Engineering and Innovative Technology (IJEIT).

Asst. Prof A. Anjana devi received her B.E Degree in Electronics and communication engineering from Paavai Engineering College, Namakkal and M.E Degree in Applied Electronics from Sathyabama University, Chennai. Currently, She is working as a Assistant Professor in M.A.M College of Engineering, Siruganur, Trichy. She had published a paper in International Journal of Engineering and Innovative Technology (IJEIT).