RESEARCH ARTICLE OPEN ACCESS

Digital Image Steganography using a Novel Uniform Distortion Function with all Possible DCT Magnitudes

N.Jeyakumar[1], M..Thangasivagamaselvi (Assistant Professor) [2]

Dept of CSE, PSN Engineering College, Tirunelveli.

---*****************************---

Abstract:

Secure steganographic algorithms hide confidential messages within other, more extensive data (carrier media). An attacker should not be able to find out, that something is embedded in the steganogram (i. e., a steganographically modified carrier medium). The existing distortion functions are minimal distortion function is used for both side-informed and non-side-informed secure JPEG steganography. UED tries to spread the embedding modification uniformly to quantized discrete cosine transform (DCT) coefficients of all possible magnitudes and achieves statistical detectability. The inappropriate use of DC and zero AC coefficients in JPEG steganography may lead to additional block artifacts in stego image and decrease in the efficiency of JPEG compression. That is why the most existing JPEG steganography schemes use only non-zero AC coefficients as possible cover elements to make the embedding naturally content-adaptive. In our proposed system, certain DC and zero AC coefficients in the texture regions could indeed be incorporated to further data embedding without decreasing, even increasing the security performance. It is proposed to embed the payload while minimizing a suitably defined distortion function. The embedding distortion is computed as a sum of relative changes of coefficients in a cover image.

Keywords: ***Uniform Embedding Distortion (UED), JPEG Steganography***

---*****************************---

I. INTRODUCTION

Since the rise of the Internet one of the most important factors of information technology and communication has been security of message transmission. Steganography is the art and science of Invisible communication. Steganography means is not alter the structure of the secret message, but hides it inside a cover object (original image). After hiding process, cover object and stego object (carrying hidden information object) are same. Steganography differs from cryptography in the sense that where cryptography focuses on keeping the contents of a message secret, steganography focuses on keeping the existence of a message secret.

Steganography mainly focus on two conflicting objectives, i.e., Undetectability and embedding payload should be carefully considered when devising a steganographic scheme.

Steganography uses almost all digital file formats, but the file formats that are more suitable with a high degree of redundancy. Redundancy can be defined as the bits of an object that provide more accuracy far greater than necessary for the object's use and display. Image and audio files especially comply with these requirements. Especially on the internet, large amount of redundant bits present in the digital representation of an image. Images are the most popular cover object for digital steganography. This paper will focus on hiding information in digital image.

Digital color images are typically stored in 24bit files and use the RGB color model. When working with large images of larger images of greater bit depth, the image tend to become too large to transmit over an internet. So reduce the file size with image compression techniques. In images there are two types of compression: lossy and lossless. Lossy compression creates smaller file size by discarding the excess image data from the original image, resulting in close approximations of the original image, although not an exact duplicate. JPEG (Joint Photographic Experts Group) uses this compression technique. JPEG file format is the most popular image file format on the internet, because of the small size of the image.

The quantization block the DCT transforms an 8x8 brightness value into 8x8 frequency coefficient. A modification of a single DCT coefficient affects all 64 image pixel value in the block. After the DCT, the quantization suitably rounds the frequency values into integer values. Once the lossy quantization is over, the Hamming coding ensures the redundancy free encoding for reducing the size.

Distortion techniques need knowledge of the original cover image during the decoding process, where the decoder function to check for differences between the original cover image and discovered cover image in order to reproduce secret data hidden in the cover image. Uniform distortion function used to spread the modification uniformly in the DCT coefficient. The message is encoded at pseudo-randomly chosen pixels. The encoder adds a sequence of changes to the cover image. Using this technique, a stego object is created. After encoding the message, the DCT coefficient LSB pixel value is changed, if the stego image is different from the cover image at the given message pixel, the message bit is a “1” otherwise the message bit is a “0”. In this approach, it is hard to distinguish stego media secret message from the stego media.

II. RELATED WORK

Many steganographic systems are weak against visual and statistical attacks. Systems without these weaknesses offer only a relatively small capacity for steganographic messages. The newly developed algorithm F5 withstands visual and statistical attacks [2], yet it still offers a large steganographic capacity. F5 implements matrix encoding to improve the efficiency of embedding. Thus it reduces the number of necessary changes. F5 employs permutative straddling to uniformly spread out the changes over the whole steganogram.

J.Fridrich proposed to determine the steganographic capacity of JPEG images (the largest payload that can be undetectably embedded) with respect to current best steganalytic methods [3]. Additionally, by testing selected steganographic algorithms we evaluate the influence of specific design elements and principles, such as the choice of the JPEG compressor, matrix embedding, adaptive content-dependent selection channels, and minimal distortion steganography using side information at the sender. they conclude that the average steganographic capacity of grayscale JPEG images with quality factor 70 is approximately 0.05 bits per non-zero AC DCT coefficient.

Steganography is the art of covert communication [7]. This paper presents an efficient JPEG steganography scheme based on the block entropy of DCT coefficients and syndrome trellis coding (STC). The proposed cost function explores both the block complexity and distortion effects due to flipping and rounding errors. The STC provides multiple solutions to embed messages to a block of coefficients. The proposed scheme determines the best one with minimal distortion effect. In this way, the total distortions are significantly reduced, which corresponds to less detect ability of steganalysis. Compared with similar schemes, experiment results demonstrate the superior performance of the proposed scheme in terms of secure embedding capacity against steganalysis.

Calibration was first introduced in 2002 as a new concept [8] for attacking the F5 algorithm. Since then, it became an essential part of many feature-based blind and targeted steganalyzers in JPEG as well as spatial domain. The purpose of this method is to shed more light on how, why, and when calibration works. In particular, this paper challenges the thesis that the purpose of calibration is to estimate cover image features from the stego image. We classify calibration according to its internal mechanism into several canonical examples, including the case when calibration hurts the detection performance. All examples are demonstrated on specific steganographic schemes and steganalysis features. Furthermore, a modified calibration procedure that improves practical steganalysis

A modern direction in steganography calls for embedding while minimizing a distortion function defined in a sufficiently complex model space [9]. In this proposed work show that, quite surprisingly, even a high-dimensional cover model does not automatically guarantee immunity to simple attacks. Moreover, the security can be compromised if the distortion is optimized to an incomplete cover model. Demonstrate these pitfalls with two recently proposed steganographic schemes and support our arguments experimentally.

Currently, the most secure practical steganographic schemes for empirical cover sources embed their payload while minimizing a distortion function designed to capture statistical detectability [10]. Since there exists a general framework for this embedding paradigm with established payload–distortion bounds as well as near-optimal practical coding schemes, building an embedding scheme has been essentially reduced to the distortion design. This is not an easy task as relating distortion to statistical detectability is a hard and open problem. In this article, proposed an innovative idea to measure the embedding distortion in one fixed domain independently of the domain where the embedding changes (and coding) are carried out. The proposed universal distortion is additive and evaluates the cost of changing an image element (e.g., pixel or DCT coefficient) from directional residuals obtained using a Daubechies wavelet filter bank. The intuition is to limit the embedding changes only to those parts of the cover that are difficult to model in multiple directions while avoiding smooth regions and clean edges. The utility of the universal distortion is demonstrated by constructing steganographic schemes in the spatial; JPEG, and side-informed JPEG domains, and comparing their security to current state of the-art methods using classifiers trained with rich media models.

J.Kodovsy and J.Fridrich proposed a rich model of DCT coefficients in a JPEG file [11] for the purpose of detecting steganographic embedding changes. The model is built systematically as a union of smaller sub models formed as joint distributions of DCT coefficients from their frequency and spatial neighborhoods covering a wide range of statistical dependencies. Due to its high dimensionality, we combine the rich model with ensemble classifiers and construct detectors for six modern JPEG domain steganographic schemes: nsF5, model-based steganography, YASS, and schemes that use side information at the embedded in the form of the uncompressed image: MME, BCH, and BCHopt. The resulting performance is contrasted with previously proposed feature sets of both low and high dimensionality. We also investigate the performance of individual sub models when grouped by their type as well as the effect of Cartesian calibration. The proposed rich model delivers superior performance across all tested algorithms and payloads.

Today, the most accurate steganalysis methods for digital media are built as supervised classifiers on feature vectors extracted from the media. The tool of choice for the machine learning seems to be the support vector machine (SVM). J.Kodovsy and J.Fridrich proposed [12] an alternative and well- known machine learning tool ensemble classifiers implemented as random forests and argue that they are ideally suited for steganalysis. Ensemble classifiers scale much more favorably the number of training examples and the feature dimensionality with performance comparable to the much more complex SVMs. The significantly lower training complexity opens up the possibility for the steganalyst to work with rich (high-dimensional) cover models and train on larger training sets two key elements that appear necessary to reliably detect modern steganographic algorithms. Ensemble classification is portrayed as a powerful developer tool that allows fast construction of steganography detectors with markedly improved detection accuracy across a wide range of embedding methods. The power of the proposed framework is demonstrated on three steganographic methods that hide messages in JPEG images.

The main purpose of steganography is to hide the occurrence of communication [15]. While most methods in use today are invisible to an observer's senses, mathematical analysis may reveal statistical anomalies in the stego medium. These discrepancies expose the fact that hidden communication is happening. This paper presents improved methods for information hiding. One method uses probabilistic embedding to minimize modifications to the cover medium. Another method employs error-correcting codes, which allow the embedding process to choose which bits to modify in a way that decreases the likelihood of being detected. In addition, hide multiple data sets in the same cover medium to provide plausible deniability. To prevent detection by statistical tests, preserve the statistical properties of the cover medium. After applying a correcting transform to an image, statistical steganalysis is no longer able to detect the presence of steganography. N.Provos present an a priori estimate to determine the amount of data that can be hidden in the image while still being able to maintain frequency count based statistics. These ways quickly choose an image in which a message of a given size can be hidden safely. To evaluate the effectiveness of proposed approach, present statistical tests for the JPEG image format and explain how this method defeats them.

LIBSVM is a library for Support Vector Machines (SVMs) [16]. This package actively developed since the year 2000. The goal is to help users to easily apply SVM to their applications. LIBSVM has gained wide popularity in machine learning and many other areas. LIBSVM support various optimization problems, theoretical convergence, multi-class classification, probability estimates, and parameter selection are discussed in detail.

III. PROPOSED SYSTEM

The objective of proposed joint coefficient uniform distortion embedding frame work is to improve the security performance of steganography.

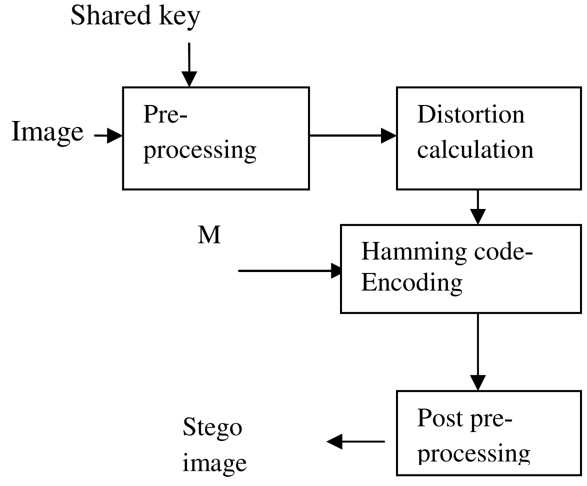

Fig1. Data Embedding

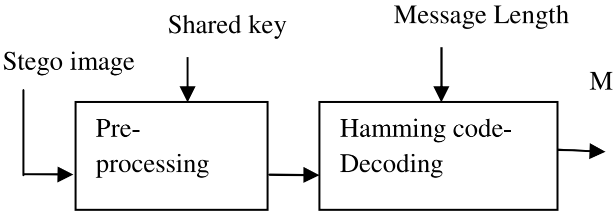

Fig2. Data Extraction

Data Embedding:

Preprocessing:

The preprocessing is adopted to generate the cover image. Let I (x, y) denote cover image with x=0,1,2,…M, and y=0,1,2,…N. this M x N cover image is divided into 8 x 8 blocks and DCT is performed on each blocks.

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

For u=0, 1, 2…7 and v=0, 1 …7.

Distortion Calculation:

Compute the embedding distortion for each non zero coefficient.

$$\rho_{ij} = \sum_{d_{ia} \in N_{ia}} (|C_{ij}| + |d_{ia}| + \alpha_{ia})^{-1} + \sum_{d_{ir} \in N_{ir}} (|C_{ij}| + |d_{ir}| + \alpha_{ir})^{-1}$$

Where $N_{ia} = \{C_{i+1,j}, C_{i-1,j}, C_{i,j+1}, C_{i,j-1}\}$

$N_{ir} = \{C_{i+8,j}, C_{i-8,j}, C_{i,j+8}, C_{i,j-8}$ are the inter block and intra block neighborhood of coefficient of $C_{i,j}$ respectively. α_{ia} and α_{ir} are adjustment parameters and determined experimentally as 1.3 and 1 respectively in [13].

Hamming code:

Hamming code is applied to embed the secret message M into the cover image. The entropy encoding is applied to embed the message bit in the cover image.

Post preprocessing:

Once the message is embedded to the cover image, then apply the Enhanced least significant bit algorithm to change the LSB value of data embedded bits. After encoding the message, the DCT coefficient LSB pixel value is changed, if the stego image is different from the cover image at the given message pixel, the message bit is a “1” otherwise the message bit is a “0”. In this approach, it is hard to distinguish stego media secret message from the cover media.

Data Extraction:

Preprocessing:

The quantized DCT coefficients are reproduced from the stego image by entropy decoding with the shared key.

Hamming code:

Finally apply the hamming code to decode the secret message from the stego image.

IV. PERFORMANCE ANALYSIS:

A class of new distortion function known as Joint Coefficients uniform embedding distortion functions for JPEG steganography. The proposed UED embed the message only to the non-zero AC coefficients fall into the realm of conventional JPEG steganography. We concentrate on minimizing the number of coefficient modified during the embedding operation. If reduce the number of change in a quantized DCT coefficient in data embedding process, it provide less coding loss in the data extraction process. Compared to the existing work, the proposed work reduces the number of coefficient modified in the data embedding process.

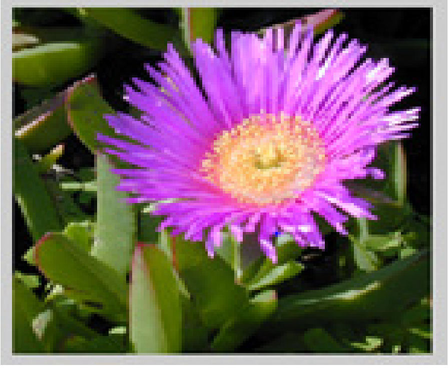

Fig 2(a) cover image

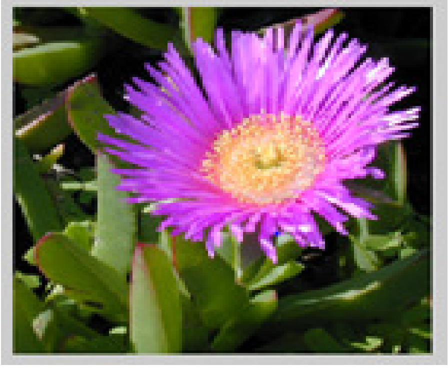

Fig 2(b) stego image

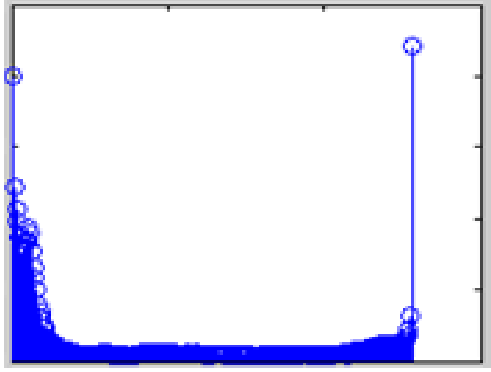

Fig 3. Shows the number of coefficient modified in the cover image

The fig 3 shows, number of coefficient modified during 38kb text file embedded in the cover image. Compared to the existing work, the proposed work reduces the number of change in the DCT coefficient of cover image and produces a stego image. So it is hard to distinguish stego media secret message from the stego media.

CONCLUSION

Many steganographic algorithms offer a high capacity for hidden messages, but they are weak against visual and statistical attacks. Uniform-distortion embedding framework is a practical approach to implement steganography with high embedding efficiency by uniformly "spreading" the embedding modifications to quantized DCT coefficients of all possible magnitudes. The average changes of the coefficient are possibly minimized, especially in the small coefficient, which leads to less statistical detectability, and hence, more secure steganography.

REFERENCES:

[1] T. Filler, J. Judas, and J.Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes" IEEE Trans. Inf. Forensics Security, vol. 6, no. 3, sep. 2011.

[2] A. Westfeld, "F5—A steganographic Algorithm" in Proc 4th inf. Hiding conf. vol. 2137. 2001.

[3] J. Fridrich, T. pevny and J. Kodoyshy, "Statistically undetectable JPEG steganography: Dead ends challenges, and opportunities" in Proc. 9th ACM Workshop Multimedia Security, Dallas, TX, USA, Sep. 2007.

[4] Y. Kim, Z. Duric, and D. Richards, "Modified matrix encoding technique for minimal distortion steganography" in Proc.8th Inf. Hiding Conf, vol. 4437, 2006.

[5] V. Sachney, H. J. Kim, and R. Zhang "Less detectable JPEG steganography method based on heuristic optimization and BCH syndrome coding" in Proc. 11th ACM Workshop Multimedia Security SEP.2009.

[6] T. Filler and J. Fridrich, "Design of adaptive steganographic schemes for digital images" Proc. SPIE, vol. 7880, p.78800F, Jan. 2011.

[7] C. Wang and J. Ni, "An efficient JPEG steganographic scheme based on the block entropy of DCT coefficient," in Proc. IEEE ICASSP, Kyoto, Mar 2012.

[8] J. Kodoysky and J. Fridrich, "Calibration revisited" in proc. 11th ACM Workshop Multimedia Security, New York, NY, USA, Sep. 2009.

[9] J. Kodoysky, J. Fridrich and V. Holub, "On dangers of overtraining steganography to incomplete cover model" in Proc. 13th ACM Workshop Multimedia Security, New York, NY, USA, pp. 69-7. 2011.

[10] V. Holub and J. Fridrich, "Digital image steganography using universal distortion" in Proc. 1st ACM Workshop Inf. Hiding Multimedia Security, 2013.

[11] J. Kodoysky and J. Fridrich, "Steganalysis of JPEG images using rich models," Proc. SPIE, vol. p.83030A, Jan. 2012.

[12] J.Kodoysky, J.Fridrich and V. Holub, "Ensemble classifiers for steganalysis of digital media" IEEE Tran. Inf. Forensics Security, vol. 7, pp. 432-444,Apr 2012.

[13] L. Guo, J. Ni, and Y.Q.Shi "An efficient JPEG steganographic scheme using uniform embedding," in Proc. 4th IEEE Int. Workshop Inf. Forensics Security, Tenerife, Spain, pp. 169-174, Dec.2012.

[14] P. Bas, T. Filler, and T. Pevny, "Break our steganographic system –The ins and outs of organizing boss," in Proc. 13th Inf. Hiding Conf., pp. 59-70, 2011.

[15] N. Provos, "Defending against statistical steganalysis" in proc. 10th USENIX Security symp. Washington, DC, USA, 2001.

[16] C.-C. Chang and C.-J. Lin, "LIBSVM:A library for support vector machines," ACM Trans. Intell. Syst. Technol., vol. 2, no. 3, pp. 27:27, 2011.