ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 4, June 2012

High Capacity Data Embedding System using Projection Quantization

Navdeep Kaur, Sukhjeet K. Ranade

Department of Computer Science and Engineering, Punjabi University, India.

***Abstract*— In this paper, we propose a high capacity data-embedding system in DCT domain and review the existing high capacity data embedding systems both in spatial domain as well as transform domain. The embedding algorithm is based on the quantized projection embedding method with some enhancement to achieve high embedding rates. The embedding is done on the quantized DCT coefficients using the concept of Hadamard Matrix as base vector. The system has very high capacity as comparable to the existing techniques of DCT domain. The system is highly robust and achieves high visual imperceptibility.**

Keywords: data embedding, high-capacity data hiding, information hiding, hiding capacity, DCT transformation.

I. INTRODUCTION

Image, audio, video, and many other kinds of data are nowadays mostly passed from person to person or from place to place in digital form. The data hiding is the process of hiding the very existence of the data into some cover medium so that the malicious user cannot tamper with it. The embedding of data in digital contents is done for authentication, copyright control, or for secret data hiding. Data-embedding techniques designed to take care of such tasks are commonly classified as watermarking or steganographic techniques in accordance with their functionalities. We know that, with the use of steganographic techniques, it is possible to hide information within digital audio, images and video files which is perceptually and statistically undetectable. The method of embedding secret message (which can be plain text, cipher text, or even images) is usually based on replacing bits of useless or unused data in the source cover (can be audio files, sound, text, Disk space, hidden partition, network packets, digital images, software, or circuitry). There are two common methods of embedding: Spatial embedding in which messages are inserted into the LSBs of image pixels, and Transform embedding in which a message is embedded by modifying frequency coefficients of the cover image (result is called the stego-image).Transform embedding methods are found to be in general more robust than the Spatial embedding methods which are susceptible to image-processing type of attacks. However with respect to steganography the perceptibility (i.e., whether the source cover is distorted by embedding

information to a visually unacceptable level) is a critical property. There is another important issue of steganography, namely, capacity, i.e., how much information can be embedded relative to its perceptibility. We will use digital images as the cover object in this paper in which we embed the hidden information. The challenge of using this data embedding method in cover images is to hide as much data as possible with the least noticeable difference in the output-image and to obtain high robustness.

In this paper we propose a high capacity data embedding system in DCT domain that uses Projection Quantization technique and the Hadamard matrix as base vectors for data embedding.

The rest of the paper is organized as follows: Section (II) describes the existing high capacity data embedding techniques both in spatial domain as well as DCT domain. In section (III) projection quantization technique is described and in section (IV) we describe our proposed system. The conclusion is given in section (V).

II. HIGH CAPACITY DATA EMBEDDING BACKGROUND

Many data embedding methods for hiding data in still images have been proposed [1]-[3]. In order to maintain the secrecy of important data in an image, only a small amount of data can be encoded therein (called the data payload). To obtain higher payloads, image-hiding methods based on least-significant-bit (LSB) substitution have been proposed [6]–[8]. These methods typically utilize some mapping rules to embed the important image in certain LSB planes of the cover image and apply additional pixel-adjustment procedures to reduce the errors introduced in the embedding process. Meanwhile, some studies [5], [9] have considered the characteristics of the human vision system when evaluating the number of bits that can be hidden in an

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 4, June 2012

image. Given that human eyes are most sensitive to edges, these methods usually hide more data in areas with higher spatial variations. The data payload and the imperceptibility are the two most important properties of a steganography system. Intrinsically, these requirements contradict each other, since a high data payload introduces more artifacts into the cover image and hence, increases the perceptibility of the hidden data. So, there is always a trade-off between the capacity and the visual imperceptibility of the data embedding system. Previous attempts [3], [5], [9] to maintain both the imperceptibility and a high data payload have worked from the imperceptibility metric: estimating the degree of alterations that are imperceptible to viewers and then embedding data within this constraint. In military and commercial applications, a large amount of communication by a site tends to expose its position and value. Moreover, due to physical constraints and security concerns, the bandwidth of a communication channel where stego-images are used tends to be low.

In order to improve the capacity of the hidden secret data and to provide an imperceptible stego-image quality, a steganographic method based on least-significant-bit (LSB) replacement and pixel-value differencing (PVD) method [10] is presented. As compared with the PVD method being used alone, the method can hide much larger information and maintains a good visual quality of stego-image. Meanwhile a method based on Discrete Cosine Transform (DCT), Vector Quantization (VQ), and a Pseudo Random Number Generator (PRNG) have been developed that can embed more information than traditional algorithms without compression and provides good imperceptibility and robustness in terms of both JPEG compression and other signal processing attacks.

Another steganography method is that utilizes a two-way block-matching [12] procedure to search for the highest similarity block for each block of the important image. The bases and indexes obtained together with some not-well-matched blocks are recorded in the least significant bits of the cover image using a hop scheme. The method exhibits a high data payload, which reduces the storage and transmission-time requirements and also provides a method that prevents an observer from selectively blocking the transmission of the important image.

There are many scholars pay attention to increase hiding capacity of information hiding algorithm based on DCT transformation in order to spread its application to wider territory. The fast and efficient high-capacity embedding algorithm [13] in DCT domain is developed which is based on Quantized Projection embedding method. It provides good trade-off between robustness, data capacity, and visual quality and achieves very high hiding ratio and a low Bit error rate (BER) under JPEG compression.

Because human vision system is much more sensitive to signal in low frequency than in high frequency, hiding information in low frequency of DCT has better robustness while hiding in median and high frequency has better imperceptibility. In standard quantization matrix commended by JPEG, high frequency has bigger quantization value, thus information embedded in high frequency will be easily filtered by JPEG compression. Then according to that, an information hiding algorithm [16] that can embed information in DCT median and high frequency coefficients is developed which provides strong robustness against lossy compression.

Four-pixel differencing method is implemented as given in [17] and it is based on LSB substitution method. Another method to increase the hiding capacity is using mod-4 embedding method [18] which is based on the image contrast.

III. THE PROJECTION QUANTIZATION TECHNIQUE

Fig. 1 illustrates the geometric interpretation of the projection quantization [18] approach. Let $\tilde{v}$ be the embedded (reconstructed) block vector, $\tilde{v} = [\tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_N]^t$; v be the block vector, $v = [v_1, v_2, \ldots, v_N]^t$. Vector can be expressed as follows:

$$\tilde{v} = v - Pv + B\Delta Q_0(B^H v/\Delta). \quad (1)$$

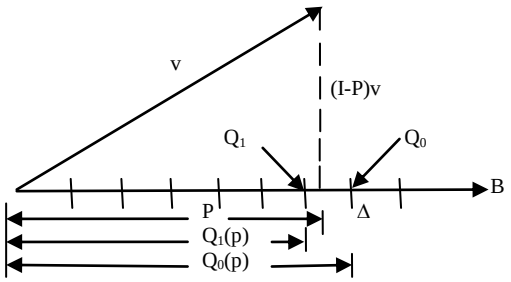

Fig. 1. Quantization Projection

Here, the quantized projection, $Q_0(p) = \Delta Q_0(B^H v/\Delta)$, is equal to $p+\varepsilon_p$. The random variable ε_p is the quantized projection error term. The distribution function of the error term can be modeled as a uniform function within the range ⌊-Δ, Δ⌋ where ⌊ and ⌋ mean exclusive and Δ is the quantization step used. Superscript H indicates the hermitian (or transpose for a real vector) and B denotes the base vector used for projection, with $B = [b_1, b_2, \ldots, b_N]^t$. Vector B is a normalized vector, so its norm is 1. Finally, P is the projection matrix of the subspace of B and is equal to BB^H, and Q_0 and Q_1 are the nearest even and odd number

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 4, June 2012

quantization operators, respectively. Note also that all vectors are of dimension Nx1.
Substituting $p+\varepsilon_p$ in (1) and simplifying the equation, we obtain $\tilde{v}$ expressed in terms of ε_p.

$$\tilde{v} = (I-P)v + B(p+\varepsilon_p) \quad (2)$$
$$\tilde{v} = v + \varepsilon_p B. \quad (3)$$

Now, if the embedded block vector undergoes other quantization attacks such as JPEG compression, we can write down the resulting quantized $\tilde{v}$, $Q_c(\tilde{v})$ as

$$Q_c(\tilde{v}) = v + \varepsilon_p B + E_q \quad (4)$$

Here, E_q is the quantization error random vector with associated quantization steps, δ_is.

IV. THE PROPOSED METHOD

We use the default quantization table (Qtable) of JPEG as our fixed mask for the projection quantization step(s), group DCT coefficients of the host image from low- to high-frequency bands, and embed bits from band to band with specially designed base vectors called Hadamard vectors. The JPEG Qtable is as follows:

Table I
JPEG Standard Quantization Table (Qtable)

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

The procedure of our embedding scheme is described as follows:

1) Calculate the DCT coefficients of 8x8 blocks of the host image.
2) Rearrange the coefficients from the lowest (DC) frequency band to the highest one. Each value in the Qtable serves as the value of the corresponding frequency band. For instance, 16 is the value for the DC band.
3) Use the columns of a Hadamard matrix, as base vector(s) to allow multiple bit embedding.
4) Sort the frequency bands using the Qtable from the smallest value to the largest, except that the DC band is always the first band.
5) To perform multi-bit embedding using base vectors as described in step 3, embed "0" or "1" of data using the "even" or "odd" quantization operators to each block vector of the sorted frequency bands. For example, for a 256x256 image each frequency band size is 32x32. The column vectors of a frequency band are the block vectors. To embed bit "0" we use even number quantization.

$$\tilde{v} = v - Pv + B\Delta Q_0(B^H v/\Delta).$$

To embed bit "1," the odd number quantization is used.

$$\tilde{v} = v - Pv + B\Delta Q_1(B^H v/\Delta).$$

6) To get the reconstructed image after embedding all bits, we reverse the reordering of the DCT coefficients that was performed in the embedding process from band to band.
7) We then apply the inverse discrete cosine transform (IDCT) to the modified and correctly ordered DCT coefficients to produce the embedded image.
8) The embedding process can be stopped at any rate according to the user's preference for visual distortion. The more bits we add, the more distortion we create.

The extraction process is similar to the embedding process except that, in Step 6), we need to calculate the projection ($B^H\tilde{v}$) instead of embedding bits. The computed projection is then rounded with respect to the corresponding Δ. If the result is an even number, then the extracted bit is "0"; otherwise it is "1". For more security the data to be embedded can be encrypted using some encryption mechanism such as RSA or md5. This encrypted data can be decrypted using the decryption mechanism to obtain the original plain data.

V. CONCLUSION

The techniques for embedding data in Spatial domain have larger capacity as compared to that of techniques for data embedding in DCT domain. Our proposed system embeds a large amount of data in DCT domain with minimal distortion. The system is highly robust and secure. The system has low BER under JPEG attacks.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments.

REFERENCES

[1] W. Bender, D. Gruhl, N. Morimoto, and A. Liu, "Techniques for data hiding," *IBM Syst. J.*, vol. 35, pp. 313–336, 1996.

[3] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding—a survey," *Proc. IEEE*, vol. 87, no. 7, pp. 1062–1078, Jul. 1999.

[2] L. M. Marvel, C. G. Boncenet, Jr., and C. T. Retter, "Spread spectrum image steganography," *IEEE Trans. Image Procesing,*, vol. 8, no. 8, pp. 1075–1083, Aug. 1999.

[4] M. Alghoniemy and A. Tewfik, "Progressive quantized projection watermarking scheme," *Proc. ACM Multimedia Conf.*, 1999.

[5] Y. K. Lee and L. H. Cheng, "High capacity image steganographic model," *Proc. Inst. Elect. Eng., Vis., Image, Signal Processing*, vol. 147, no. 3, pp. 288–294, Jun. 2000.

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

[7] C. C. Thien and J. C. Lin, "Asimple and high-hiding capacity method for hiding digit-by-digit data in images based on modulus function," *Pattern Recognition*, vol. 36, pp. 2875–2881, 2003.

[8] C. K. Chan and L. M. Cheng, "Hiding data in images by simple LSB substitution," *Pattern Recogni*tion, vol. 37, pp. 469–474, 2004.

[9] X. Zhang and S. Wang, "Steganography using multiple-base notational system and human vision sensitivity," *IEEE Signal Procesing Let*ters, vol. 12, no. 1, pp. 67–70, Jan. 2005.

[10] H.-C. Wu, N.-I. Wu, C.-S. Tsai and M.-S. Hwang, "Image steganographic scheme based on pixel-value differencing and LSB replacement methods," *IEE Proc.Vis. Image Signal Processing*, Vol. 152, No. 5, pp. 611-615, Oct. 2005.

[11] Y. Y. Chung, "A study of High Capacity Image Steganographic System," *Proc. Tencon 2005 IEEE Region 10*, pp. 1-5, Nov. 2005.

[12]R.-Z. Wang and Y.-S. Chen, "High-Payload Image Steganography Using Two-Way Block Matching," *IEEE Signal Processing Letters*, VOL. 13, NO. 3, pp. 161-164, Mar. 2006.

[13] T.-H. Lan and A. H. Tewfik, "A Novel High-Capacity Data Embedding System," *IEEE Transactions on Image Processing*, Vol. 15, No. 8, pp. 2431-2440, Aug. 2006.

[14] K. B. Raja, Vikas, K. R. Venugopal, and L. M. Patnaik, "High Capacity Lossless Secure Image Steganography using Wavelets," *Proc. Advanced Computing and Communications, ADCOM-2006*, pp. 230-235, Dec. 2006.

[15] X. Jianquan, Y. Chunhua, H. Dazu, "High Capacity Information Hiding Algorithm for DCT Domain of Image," *Proc. Intelligent Hiding and Multimedia Signal Processing, IIHMSP-2008*, Aug. 2008, pp. 269-272

[16] X. Jianquan, X. Qing, H. Dazu, "A Robust High Capacity Information Hiding Algorithm Based on DCT High Frequency Domain," *Proc. Computer Network and Multimedia Technology, CNMT-2009*, pp. 1-4, Jan. 2009.

[17] M.B. O. Medeni, E.M. Souidi, "A Novel Steganographic Method for Gray-Level Images With four-pixel Differencing and LSB Substitution," Proc. IEEE, 2010.

[18] K. Pramitha, L.P.Suresh, K.L. Shunmuganathan, "Image Steganography Using Mod-4 Embedding Algorithm Based On Image Contrast," *Proc. International Conference on Signal Processing, Communication, Computing and Networking Technologies (ICSCCN 2011)*, pp. 364-369, 2011.