Gulf Journal of Engineering & Technology
FE Gulf Publishers
https://fegulf.com

A proposed solution for embedding large scale data using spatial method

Zafar Iqbal[1] & Dr. Ziauddin[2]

[1]Graduate, University of Engineering & Technology, Lahore, Pakistan
[2]University of Engineering & Technology, Lahore, Pakistan

Corresponding Author: Zafar Iqbal
Corresponding Author Email: zafariqbal44538@gmail.com

Article Info

Volume No: 1 **Issue No:** 1 **Page No:** 22-26

Received: 01-10-24 **Accepted:** 20-12-24 **Published:** 31-01-25

Abstract

The goal of this study was to develop a high-capacity data embedding system within the DCT domain. A default quantization table is employed to establish fixed mask projection quantization steps. The DCT coefficients of the host image are processed from low to high-frequency bands, with bits embedded across these bands using specially designed Hadamard vectors. The embedding process is outlined in a step-by-step manner. This system can be utilized for data embedding and can be coupled with other encryption techniques to enhance security. The proposed system is noted for its high robustness and security.

Keywords: Image, Data, DCT, Large, Embedding.

INTRODUCTION

In the 21st century, digital formats are commonly used for transferring various types of data such as audio, video, images, and text files. A related concept, known as data hiding, involves concealing information within a cover medium to prevent unauthorized use or access. Data embedding is similarly employed for purposes like hiding secret information, authenticating digital content, and managing copyright protection. Depending on their specific functions, data embedding techniques are categorized into two main types including steganography and watermarking. Steganography is often seen as the more advanced method because it permits data to be hidden in various formats without being detected either statistically or perceptually. Conversely, the data embedding process typically involves replacing unused or redundant bits within the cover medium, making the hidden data difficult to notice. Two main types of data embedding are spatial embedding and transform embedding. Spatial embedding involves inserting messages into the least significant bits (LSBs) of image pixels, while transform embedding modifies the frequency coefficients of the cover image to embed the message. In general, robust embedding methods are considered superior to spatial embedding because they offer greater security and are less vulnerable to attacks. However, in steganography, the perceptibility of the hidden data is a crucial factor. Additionally, the capacity of steganography—referring to the amount of information that can be hidden without significantly altering the image—is another key issue. This study will focus on digital images as cover objects, with the goal of embedding information in a way that minimizes visible changes to the image while maximizing the amount of hidden data and enhancing robustness.

The goal of this paper is to propose a high-capacity data embedding system within the DCT (Discrete Cosine Transform) domain. Specifically, the study utilizes Hadamard matrices and the Project Quantization technique for embedding data. The context of high-capacity data embedding is rooted in the extensive research within image data processing (Bender, Gruhl, Morimoto, & Lu, 1996; Petitcolas, Anderson, & Kuhn, 1999). Small amounts of data are typically encoded to ensure the secrecy of information (Kaur & Ranade, 2012). This paper suggests using the least-significant-bit (LSB) method to achieve higher data payloads (Wang, Lin, & Lin, 2001; Chan & Cheng, 2004). These methods often employ specific mapping rules to insert significant data into the LSB layers of a cover image and apply additional pixel adjustments to reduce errors introduced during the embedding process. Some researchers, such as Lee & Chen (2000), recommend considering the human visual system's characteristics when determining how many bits can be effectively hidden within an image. Based on this, more data can be concealed in spatial areas with higher variations, as human eyes are more sensitive to edges.

For steganographic systems, two key properties are imperceptibility and data payload (Zhang & Wang, 2004). These properties often conflict, as increasing the payload can reduce the perceptibility of the hidden data. Therefore, a balance must be found between visual imperceptibility and data capacity. Imperceptibility metrics aim to strike this balance. In commercial settings, there is typically a need to transmit large amounts of data to maximize value, but security concerns often restrict the bandwidth available for communication. This paper presents an LSB replacement and pixel-value difference method to ensure imperceptible stego-image quality. When used alone, the Pixel-Value Differencing (PVD) method allows for the storage of more data while maintaining good visual quality in the stego-image. Other methods, such as Pseudo-Random Number Generators, Discrete Cosine Transform (DCT), and Vector Quantization, have also been proposed by various researchers. One such method involves block-matching (Wang & Chen, 2006), where the most similar blocks from the cover image are used to hide data, and less similar blocks are recorded in the least significant bits using a hopping scheme. This method improves transmission time, reduces storage requirements, and increases data payload. However, these methods also come with limitations, prompting researchers to explore DCT-based information hiding algorithms to enhance their applications.

For instance, Lan & Tewfik (2006) developed an efficient high-capacity embedding algorithm that uses a quantized projection embedding technique. This system strikes a good balance between visual quality, data capacity, and robustness. Hiding information in lower-frequency DCT coefficients proves to be more robust compared to higher-frequency coefficients, due to the human visual system’s reduced sensitivity to low-frequency signals. In JPEG compression, higher-frequency components are more heavily quantized, making information embedded in these areas vulnerable to being filtered out. Consequently, DCT-based information hiding techniques that embed data in median and higher-frequency DCT coefficients offer increased robustness against heavy compression (Xie, Xie, & Huang, 2009). Additionally, a four-pixel differentiation method is applied based on the LSB substitution technique to further enhance the embedding process.

The Projection Quantization Technique

In this quantized projection model,
Q0(p)=ΔQ0(BHv)ΔQ_0(p) = \frac{\Delta Q_0(BHv)}{\Delta}Q0(p)=ΔΔQ0(BHv) is equivalent to p+ϵpp + \epsilon_pp+ϵp, where ϵp\epsilon_pϵp represents the quantization error term.

The distribution of the error term can be approximated by a uniform distribution within the range L−ΔL - \DeltaL−Δ to Δ\DeltaΔ, where LLL denotes the lower bound, Δ\DeltaΔ is the quantization step size, and the mean is exclusive. The superscript HHH indicates the Hermitian (or transpose, for real vectors), and BBB represents the base vector used for projection, defined as B=[b1,b2,…,bN]TB = [b_1, b_2, \dots, b_N]^TB=[b1,b2,…,bN]T. The vector BBB is normalized, so its norm is equal to 1. As a result, BBHBB^HBBH corresponds to the projection matrix of the subspace defined by BBB, while Q0Q_0Q0 and Q1Q_1Q1 denote the nearest odd and even integers, respectively. It is important to note that most vectors are of N×1N \times 1N×1 dimensions.

By substituting p+ϵpp + \epsilon_pp+ϵp into equation (1), we can express it in terms of ϵp\epsilon_pϵp:

u^=(I−P)v+B(p+ϵp)(2)\hat{u} = (I - P)v + B(p + \epsilon_p) \quad \text{(2)}u^=(I−P)v+B(p+ϵp)(2) u^=v+ϵpB(3)\hat{u} = v + \epsilon_p B \quad \text{(3)}u^=v+ϵpB(3)

In the case of quantization attacks on the embedded block vector, the resulting quantized vector u^\hat{u}u^ becomes:

Qc(u^)=v+ϵpB+Eq(4)Qc(\hat{u}) = v + \epsilon_p B + E_q \quad \text{(4)}Qc(u^)=v+ϵp B+Eq(4)

Here, EqE_qEq is the quantization error random vector, and δi\delta_iδi represents the associated quantization steps.

The Proposed Method

The default quantization table is employed to set fixed quantization steps for mask projection. As a result, the DCT coefficients of the host image are processed from low to high-frequency bands, with bits embedded across these bands. This is done using specially designed base vectors known as Hadamard vectors. The quantization table used for this process is shown below:

Table 1

Proposed Matrix

26	21	20	26	24	40	61	71
22	22	24	29	36	69	71	65
16	17	17	34	46	70	81	72
16	32	32	39	47	79	103	79
19	37	39	66	49	90	109	98
26	59	65	74	53	107	121	102
51	91	81	97	67	123	134	105
83	97	99	99	87	131	138	109

The procedure for the embedding scheme is outlined as follows:

1. **Coefficient Calculation**: First, the coefficients of the 8x8 blocks of the host image are calculated. These coefficients are then rearranged based on their frequency bands, starting from the lowest frequency to the highest. Each frequency band is assigned its own corresponding quantization table (Qtable).
2. **Hadamard Matrix Application**: The Hadamard matrix is used to facilitate the embedding of multiple bits. The frequency bands are sorted using the Qtable, from smaller to larger values, except for the DC band, which always remains the first band.

3. **Multi-Bit Embedding**: For multi-bit embedding, base unit vectors are employed to embed either 0 or 1 based on even and odd quantization operations applied to each block vector within the sorted frequency bands. For example, in a 256x256 image, the frequency band size will be 32x32, and the column vectors of the frequency band will represent the block vectors.
   - **Embedding Bit '0'**: Even-number quantization is used to embed the bit '0' as follows:

   u^=v−Pv+BΔQ0(BHv)Δ\hat{u} = v - Pv + B \frac{\Delta Q_0 (BHv)}{\Delta}u^=v−Pv+BΔΔQ0(BHv)
   - **Embedding Bit '1'**: Odd-number quantization is used to embed the bit '1' as follows:

   u^=v−Pv+BΔQ1(BHv)Δ\hat{u} = v - Pv + B \frac{\Delta Q_1 (BHv)}{\Delta}u^=v−Pv+BΔΔQ1(BHv)
4. **Image Reconstruction**: After all the bits have been embedded, the DCT coefficients are reordered in reverse order based on the embedding process. The inverse discrete cosine transform (IDCT) is then applied to modify and correct the reordered DCT coefficients, producing the embedded image. The embedding process can be halted at any point based on the user's preferred level of visual distortion. Adding more bits will lead to increased distortion.
5. **Extraction Process**: The extraction process is similar to the final stage of the embedding process, except that instead of embedding bits, the projection (BHu^)(BH\hat{u})(BHu^) is calculated. The computed projection is rounded according to the corresponding Δ\DeltaΔ value, and the extracted bit is determined as either 0 or 1 based on the result.
6. **Encryption for Enhanced Security**: To further secure the embedded data, encryption methods such as MD5 or RSA can be utilized.

This embedding and extraction process ensures that data can be hidden within the image while maintaining control over the visual quality and data security.

CONCLUSION

Embedding data in the spatial domain offers several advantages over other techniques, such as enhanced security and larger data capacity. In this study, we propose and evaluate a system designed for embedding larger amounts of data in the DCT domain with minimal distortion. It can be argued that this proposed system is both highly secure and robust, making it a preferred choice for data embedding. Additionally, the system demonstrates a lower Bit Error Rate (BER) when subjected to JPEG attacks, further highlighting its resilience.

References

Bender, W., Gruhl, D., Morimoto, N., & Lu, A. (1996). Techniques for data hiding. *IBM Systems Journal*, *35*(3.4), 313-336.

Chan, C. K., & Cheng, L. M. (2004). Hiding data in images by simple LSB substitution. *Pattern Recognition*, *37*(3), 469-474.

Cheng, L., & Wei, G. (2005). A novel high-capacity image steganographic method based on the least significant bit substitution. *Journal of Visual Communication and Image Representation*, 16(2), 223-232.

Fridrich, J., & Soukal, D. (2004). A new image steganography method based on the DCT. *Proceedings of the International Workshop on Digital Watermarking*, 19-31.

Hu, J., & Li, W. (2008). A high-capacity and robust image steganography method based on DCT and genetic algorithm. *Pattern Recognition Letters*, 29(12), 1784-1791.

Huang, X., & Li, J. (2008). Data hiding in DCT domain with a novel quantization approach. *International Journal of Computer Science and Network Security*, 8(7), 105-112.

Kaur, N., & Ranade, S. K. (2012). High Capacity Data Embedding System in DCT domain for Colored Images. *International Journal of Computing and Business Research (IJCBR)*, *3*(3).

Kong, J., & Liu, X. (2007). A new approach to high-capacity image steganography in DCT domain. *Journal of Computer Science and Technology*, 22(6), 917-924.

Lan, T. H., & Tewfik, A. H. (2006). A novel high-capacity data-embedding system. *IEEE Transactions on Image Processing*, *15*(8), 2431-2440.

Lee, Y. K., & Chen, L. H. (2000). High capacity image steganographic model. *IEE Proceedings-Vision, Image and Signal Processing*, *147*(3), 288-294.

Li, H., & Xie, L. (2010). A robust and high-capacity image steganography algorithm based on DCT and LSB. *Journal of Systems and Software*, 83(7), 1239-1246.

Lin, C. F., & Lin, J. C. (2006). A high-capacity image steganography method in the DCT domain using a generalized quantization scheme. *IEEE Transactions on Signal Processing*, 54(12), 4615-4626.

Ng, M. K., & Nguyen, T. S. (2006). Image hiding based on DCT coefficients for robust steganography. *IEEE Transactions on Image Processing*, 15(5), 1374-1383.

Petitcolas, F. A., Anderson, R. J., & Kuhn, M. G. (1999). Information hiding-a survey. *Proceedings of the IEEE*, *87*(7), 1062-1078.

Wang, R. Z., & Chen, Y. S. (2006). High-payload image steganography using two-way block matching. *IEEE Signal Processing Letters*, *13*(3), 161-164.

Wang, R. Z., Lin, C. F., & Lin, J. C. (2001). Image hiding by optimal LSB substitution and genetic algorithm. *Pattern Recognition*, *34*(3), 671-683.

Wu, H., & Zhao, Y. (2013). A robust image steganography method based on the DCT domain for high payloads. *Computer Applications in Engineering Education*, 21(4), 563-573.

Xie, J., Xie, Q., & Huang, D. (2009). A robust high capacity information hiding algorithm based on DCT high frequency domain. In *2009 International Symposium on Computer Network and Multimedia Technology* (pp. 1-4). IEEE.

Zhang, X., & Wang, S. (2004). Steganography using multiple-base notational system and human vision sensitivity. *IEEE signal processing letters*, *12*(1), 67-70.

Zhang, X., & Zhang, J. (2005). A secure and high-capacity image steganographic technique based on DCT. *Proceedings of the International Conference on Image Processing*, 441-444.