Features Analysis and Extraction Techniques for the Image Steganography

Abdul Yabar Rafiqi[a], Archana Singh[b]

[a, b] Amity School of Engineering and Technology, Amity University Uttar Pradesh
[a]rafiqiyabar@gmail.com, [b]asingh27@amity.edu

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 20 April 2021

Abstract: The image stenography is utilized to provide security to sensitive data. Several techniques are planned so far for creating an effective stegno image. The existing technique makes the utilization of discrete wavelet transform for detecting the edges from the images and the final stegno image is produced by inserting text on the detected edges. This research work deploys the Grey Scale Co-occurrence Matrix (GLCM) to create the stegno image so that the attributes can be extracted. The text is inserted into the image using PCA (Principal Component Analysis). MATLAB is applied to test the performance of suggested model. The testing depicts that the suggested model provides efficient performance concerning PSNR (Peak signal-to-noise ratio) and MSE (mean squared error).

Keywords: GLCM, PCA, PSNR, MSE, Entropy

1. Introduction

The major concern of information technology and communication is to secure the information as the internet is developing rapidly. Cryptography is put forward to keep the communication confidential. A number of methods are constructed to encrypt and decrypt the messages for their safety. But, these techniques are inefficient for securing the contents of a message. Thus, the existence of message is kept as a secret in new technologies. Steganography is a scheme using which invisible communication can be provided to the users [1]. Other information conceals the information which leads to keep the existence of communicated information secret with the help of this technique. The image steganography technique is implemented to conceal the information in images. The steganography is different from the cryptography as the contents of message are concealed using the cryptography approach while the steganography approach keeps the existence of message secret. Even though, these approaches focus on protecting the information from unauthorized users, they can be compromised somehow due to the emergence of technology in recent years [2]. The main intend of steganography faces failure even during the detection of presence of hidden information. The applicability of this approach can be enhanced by integrating the steganography with cryptography. The technologies which are similar to steganography are Watermarking and fingerprinting. The requirements of these techniques are different in contrast to steganography because these techniques assist in tackling the issue related to the protection of intellectual properties of users. All the components of an object are marked in the watermarking technique using a same technique. In particular, a signature is useful to signify the ownership of information so that copyright protection can be ensured in watermarking [3].The intellectual property owner is capable of identifying the clients who supply the property to third parties and break their licensing agreement. In watermarking and fingerprinting, the files hide the information which can be seen in some cases. But it is essential to imperceptible the information in steganography. An attack is launched in the steganogrphic system, in case the attacker comes to know that some information is concealed in the file. A successful attack is occurred on a watermarking system when the mark is eliminated in place of detecting it. A number of image steganography methods contain several domains. The spatial domain technique is the major domain among them [4]. Numerous modifications of spatial steganography are suggested so that some bits are changed in the image pixel values in the hiding data. The steganography based on LSB (least significant bit) is a technique in which secret message is concealed in the LSBs of pixel values without any insertion of huge perceptible distortions. This technique is very simple. The modifications performed in the value of LSB cannot be seen by human eyes. The transform domain is another complex approach which assists in hiding the data in an image. The information is concealed in this domain with the deployment of several algorithms and transformation. A variety of algos introduced to embed the transform domain are called domain of embedding methods. Unlike the frequency domain of a signal, the procedure to embed the data is more efficient in the embedding domain principles which are planned on the basis of time domain [5].The transform domains carry out operations of most of the strong steganographic systems that provides more advantages than the spatial domain. The reason is the concealing of information in areas that have minimal exposure to compression, cropping and image processing in an image. In the distortion methods, knowledge related to the original cover image is required through the decoding process. The decoder functions are utilized to determine the differences amid the original and distorted cover image so that the secret message can be restored. The encoder adds the sequence of changes for the cover image [6]. At last, the signal distortion is useful for storing the information. The masking and filtering is a technique in which an image is marked similar to the paper watermarks for hiding the information. These methods focus on embedding the information in place of hiding it into noise level. The cover image considers the hidden message as integral. For covering the image, the hidden message is known to be integral.

2. Literature Review

Muhammad Arslan Usman et.al (2018) discussed the significance of image steganography in image security [7]. The use of this approach had been carried out as an alternate technique to secure healthcare data from cybercrimes prevalent in healthcare sector. In this work, a fresh strategy of image steganography was presented for securing clinical data. This work applied loss-free compression and manifold encryption to the payload using the Swapped Huffman tree coding before inserting into the cover image. Apart from this, this work embedded the undisclosed data by using merely the edge areas of the cover image. This phenomenon increases the scale of imperceptibility. In the results, the introduced technique not only ensured confidentiality and privacy of patient data but also maintained imperceptibility.

Mingming Jiang, et.al(2017) presented a FLD Ensemble classifier models and content-adaptive quadtree algorithm-based color image steganography strategy [8]. The new approach initially made use of the quad-tree algorithm to divide the color cover into many overlapped sub-blocks. The priority was assigned to every block according to the complexity of texture. Next, the statistical features of blocks based on high priority were extracted from three channels one by one. The extracted features were fed as input to the Ensemble classifier models. Finally, the output of the Ensemble classifier models was used to devise the additive distortion function of color images and measure the pixel deformity. Secret message embedded with STC were used into R, B, G channels correspondingly. The tested outcomes showed the effectiveness of the presented technique over other existing steganography methods.

Sreekutty S Kumar, et.al (2017) proposed a new image steganography technique which improved both image quality as well image security [9]. The presented framework separated the edges and smooth areas by using the canny edge detection approach. In contrast to non-edge areas, Adaptive LSB substitution scheme added more bits to the edge areas. Improved LSB substitution approach preserved edges along with embedding the message bits even when the embedding task completed. The embedding was carried out just in the areas of high entropy. A key altering the cover image earlier than embedding made the system more secure. In the results, the proposed framework outperformed the other approaches by obtaining greater PSNR value.

Tarun Jain, et.al(2017)proposed an image steganography method based on an Adaptive pattern [10]. The new approach addressed the main issues such as adaptivity in data embedding and used mask encryption algorithm to improve the system security. Both of these approaches together provided suitable outcomes. The tested outcomes revealed that the introduced method was highly efficient with low visual distortions. This work also made comparison of the new algorithm with several old algorithms based on the certain features.

Anjana Rodrigues et.al(2017) presented a data hiding approach that replaced the LSB of certain cover image pixels with the undisclosed message bits [11]. Hamming code (7, 4) was thebase of the selection. However, this work presented a basic modified approach that improved the randomness of embedding with embedding efficiency. This algorithmic approach was suitable for use intelemedicine due to the overall error-free extraction of the cover image. This algorithm could also be used for preventing the piracy of digital images. The cover image requiring security can be inserted with some noise as secret message. Also, original keys were the only means to get the genuine denoised image.

Hetal N. Patel,et.al (2017) presented a new image steganography approach based on color palette for the iterated Julia- set fractal image [12]. First of all, a color palette was generated and arranged according to the brightness level of the pixels. The scanning of the palette was carried out for embedding the secret bit. Also, the secret value of the bot was used to modify the palette index. This work recommended and efficiently applied a random pixel selection method to improve the algorithm strength. In the results, the presented algorithm obtained highest PSNR (Peak Signal to Noise Ratio) of 61 and 0.999 NC (Normalized Correlation) for a sum of 50 pictures.

3. Proposed Method

This work is carried on the basis of image encryption and the base paper method is deployed on enciphering application to transmit the image on unsecured channels. The image is encrypted while transmitting on unsecured channels. For this, the image is split into blocks. After the partition of an image, the rearrangement of these divided blocks is done for encrypting the image. The key, which is assisted in encryption, decides a pattern at which the blocks are shuffled. The key is extracted on the basis of association among the pixels of an image. The suggested approach provides better performance and promising outcomes are generated from this while dealing with various attacks. The future work will focus on key generation stage for driving a key on the basis of textural attributes of the image to reduce the pixel loss during the decryption process. Different stages are executed to implement the suggested algorithm: -

1. Pre-processing Phase: This stage utilizes two images as input. The initial image is image for which encryption is required and the second image assist in generating the key.

2. Feature extracted: The second stage includes the implementation of Grey Scale Co-occurrence Matrix (GLCM) for extracting the textual attributes from the first image. This algorithm will extract the attributes such as energy, entropy etc. image.

GLCM algorithm

1. To count all the number of pixels in the matrix that saves the data.

2. For storing the counted pixels in matrix P[I,j].

3. To check similarity among pixels in the matrix with the help of histogram technique.

4. To compute the contrast factor from the matrix:

$$g = \exp[\frac{mean(I) - minimum(I)}{maximum(I) - mean(I)}]$$

5. The pixels are divided to normalize the elements of g.

$$g = \begin{bmatrix} 0.8\, if g < 0.8 \\ 1.2\, if g > 1.2 \\ g otherwise \end{bmatrix}$$

3. **Apply PCA algorithm**: The third stage employs the PCA (Principal Component Analysis) for choosing the extracted attributes from the initial image. The attributes can be extracted using different methods such as LDA (Linear Discriminant Analysis), Independent Component Analysis and PCA. The PCA is best algorithm among them to perform the image formation. This algorithm is capable of recognizing the data illustrations, similarities and differences among them. This algorithm is adaptable to mitigate the dimension for whicha strategic distance is maintained from redundant information without much loss. The statistics and a portion of the mathematical methods such as Eigen esteems and Eigen vectors are utilized to analyze the PCA in detail. This algorithm is valuable statistical and common technique which is employed to discover the application in fields, for instance, to compress and recognize an image. PCA is a mathematical methodology in which LT (Linear Transformations) are executed for mapping the data from high dimensional space to low dimensional space.Eigen vectors of the covariance matrix controls the low dimensional space.

The stages involved in Principal Component Analysis are:

- The mean valve S of the given data set "S" is discovered

- The mean value is subtracted from S. These valves provide a novel matrix denoted with "A"

- Covariance is obtained from the matrix which implies C = AAT Eigen values are offered through the covariance matrixes which are V1V2V3V4…VN,

- At last,the calculation of Eigen vectors is done for covariance matrix C

- Any vector S or $S - \bar{S}$or can be written as linear combination of eigen vectors which is expressed in the Equation given below.

- Because covariance matrix is symmetric it form basis $V_1V_2V_3V_4 \ldots V_N$,

$$V_N S - \bar{S} = b_1u_1 + b_2u_2 + b_3u_3 + \cdots + b_Nu_N$$

- Only Largest eigen values are kept for generating lower dimension data set:

$$\hat{S} - \bar{S} = \sum_{I=0}^{1} b_1u_1\,; 1 < N$$

The components in lower dimension space are known as principal components whose freedom is ensured in case information set is mutually appropriated. PCA has susceptibility against scaling of the original variables. On the basis of field of application, this algorithm is recognized as discrete KLT (Karhunen-Loève Transform)or the Hotelling transform.

4. **Encryption of second Image**: -The subsequent step is concerned with generation of keys from the second image. This step divides the first image into blocks and encrypts every block separately for making the ultimate encrypted image. There are basicaly two phases involved in the chaos-based image cryptosystem. The plain image is fed at input into this system. The confusion stage corresponds to the pixel permutation where the location of the pixels is scrambled over the complte image without disturbing the pixels' values and the obatined image is simply unrecognizable. A chaotic system perfroms the pixel permutation. The initial conditions and control parameters determined from the 16-character key control the chaotic behavior. The next phase of the encryption process improves the security by changing the value of each pixel in the overall image. This isa important tool that protects an image from from attackers.

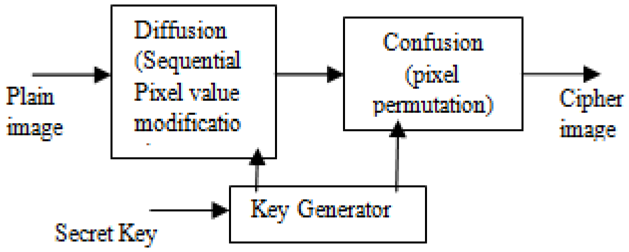
Plain image
Diffusion (Sequential Pixel value modificatio
Confusion (pixel permutation)
Cipher image
Key Generator
Secret Key

Figure 1. Architecture of Chaos-based image cryptosystem

The confusion stage refers to the pixel permutation. This stage scatters the location of the pixels over the overall image without troubling the value of the pixels and the obtained image is completely unidentifiable. Hence, the initial conditions and control parameters are used as the secret key. The permutation stage alone is not sufficiently secure to be applied as an attack can easily break its security wall. The next phase of the encryption process aiming at altering the value of all image pixels is implemented for improving the system security.

The process of diffusion is equivalent to an approach applied by the chaotic map. This process for the most part relies upon the initial conditions and control parameters. The diffusion stage modifies the pixel values sequentially using the sequence created by one of the three chaotic systems selected using the outside key. The complete confusion-diffusion round is repeated numerous times to obtain a strong level of security. The haphazardness feature in chaotic maps increases its reasonability for image encryption.

4. Evaluation of the proposed method

The proposed algorithm is an image stenography algorithm. This algorithm uses GLCM,PCA algorithm for the feature analysis and feature selection respectively. This algorithm performs image encryption using chaos-based strategy. This work uses MATLAB software for implementing the new algorithm and analyze results in the context of MSE and PSNR values.

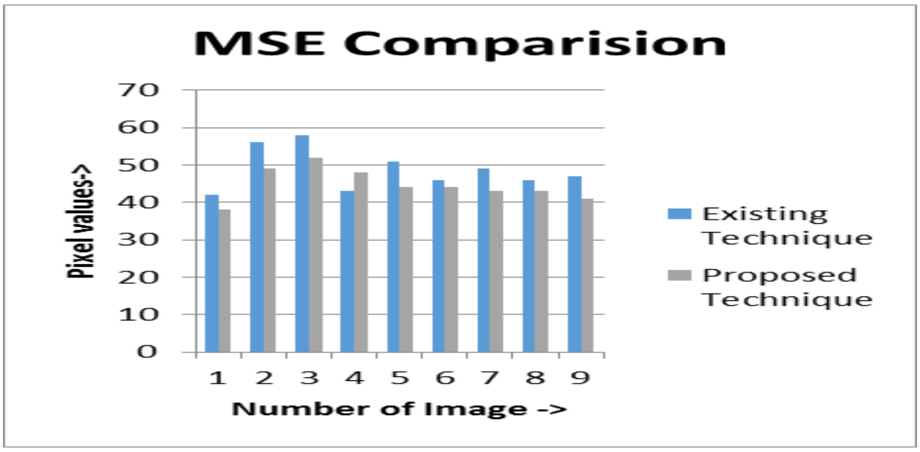

Fig 2: MSE Comparison

Figure 2 depicts the comparison of the new and old method in the context of MSE value. The analytic results reveal that the use of GLCM and PCA algorithms increases the MSE value of the new technique.

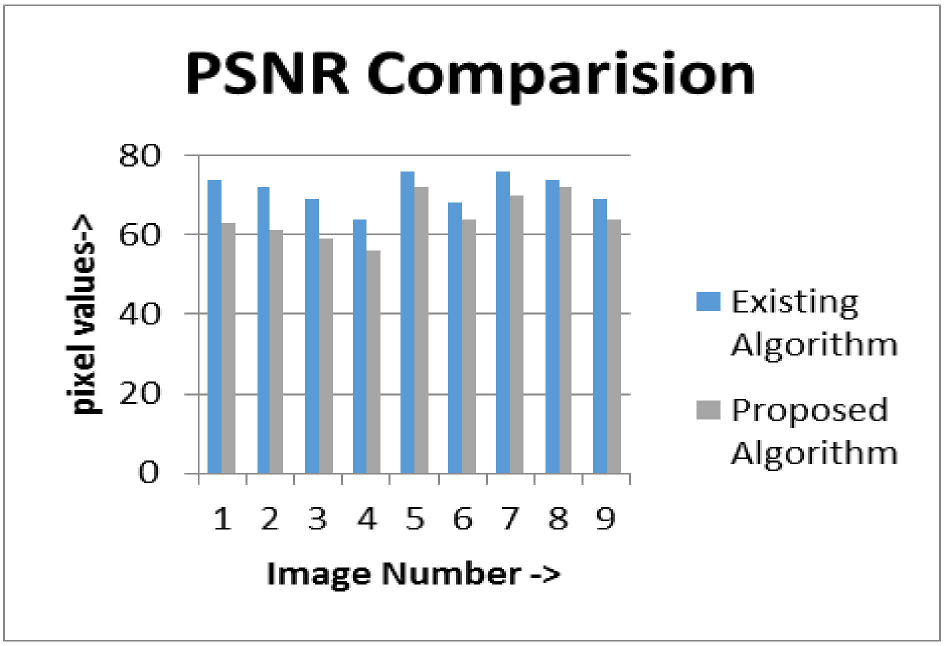

Fig 3: PSNR Comparison

Figure 2 depicts the comparison of the new and old method in the context of PSNR value. The analytic results reveal that the use of GLCM algorithms increases the PSNR value of the new technique.

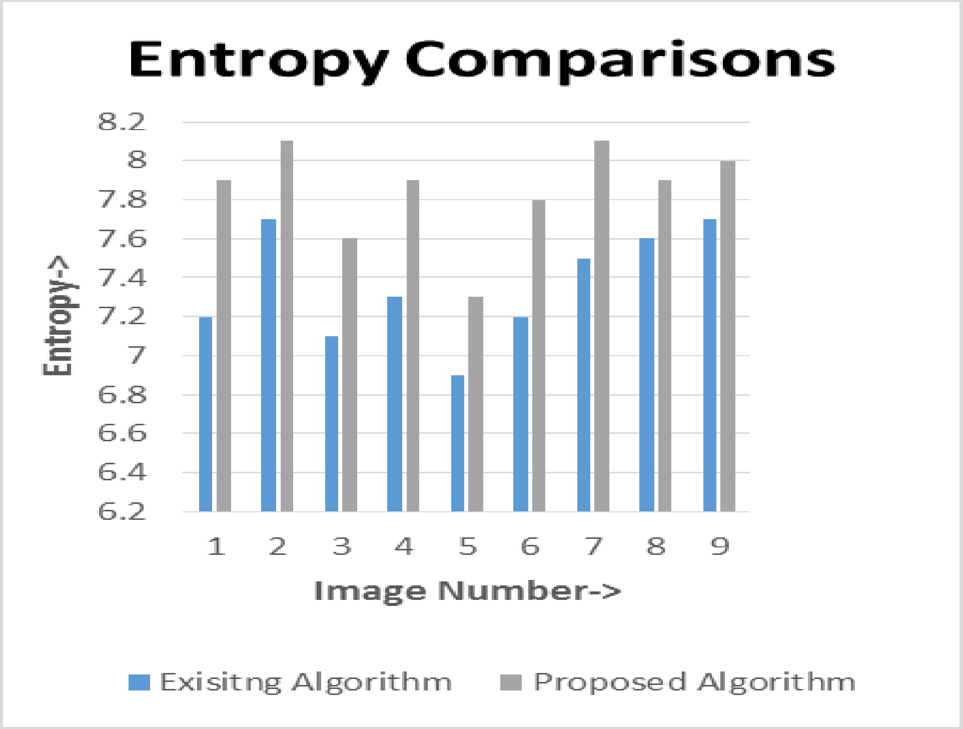

Fig 4: Entropy Comparison

Figure 4 depicts the comparison of the new and old method in the context of Entropy value. The analytic results reveal that the use of GLCM and PCA algorithms increases the Entropy value of the new technique.

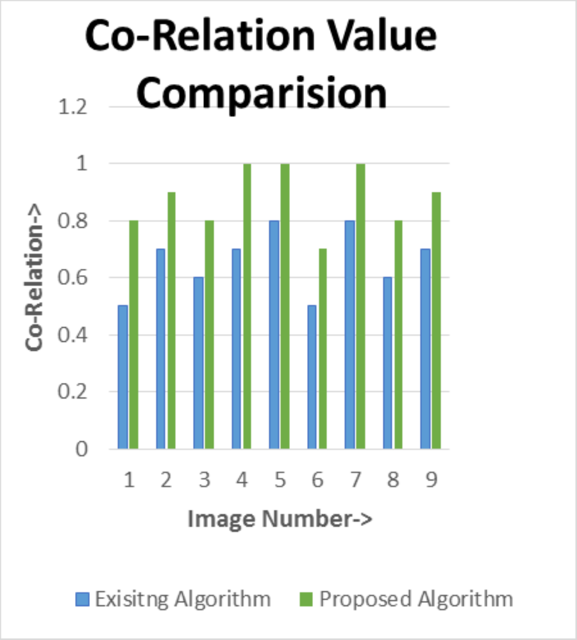

Fig 5: Co-Relation Comparison

Figure 5 depicts the comparison of the new and old method in the context of Co-relation between adjacent pixels. The analytic results reveal that the use of proposed algorithms increases the Co-relation between the adjacent pixels.

5. Conclusion

This work concludes that image stenography is a powerful approach for hiding sensitive data. Up till now, researchers have presented a number of methods for generating ultimate stegno image. The available framework of image steganography makes use of genetic algorithm for the edge detection and inserts text on the edge in the image. This work introduces a new strategy for image steganography. The new approach uses GLCM and PCA algorithms. The simulation outcomes of the proposed strategy show an increase in the PSNR value of the system and decrease in MSE value, as compared to the existing techniques which are considered as benchmarks in this work

References

1. M. A. Usman, M. R. Usman, & S. Y. Shin, "Quality assessment forwireless capsule endoscopy videos compressed via HEVC: Fromdiagnostic quality to visual perception." Computers in Biology andMedicine Vol. 91, pp: 112-134, 2017.
2. L. Q. Kuang, Y. Zhang, & X. Han, “Watermarking image authenticationin hospital information system,” Information Engineering and ComputerScience, ICIECS, pp. 1–4, 2009.
3. C. K. Chan, & L. M. Chen, “Hiding data in images by simple LSBsubstitution,” Pattern Recognition, vol. 37, no. 3, pp. 469–474, 2004.
4. M.S. Nambakhsh, A. Ahmadian, & H. Zaidi, “A contextual based doublewatermarking of PET images by patient ID and ECG signal,” in Comput.Methods Progr. Biomed, vol. 104, no. 3, pp: 418–425, 2011.
5. H. Golpira, & H. Danyali, “Reversible blind watermarking for medicalimages based on wavelet histogram shifting,” in Proceedings of IEEEInternational Symposium on Signal Processing and InformationTechnology, pp. 31–36, 2010.
6. K. K. Tseng, J. M. Jiang, J. S. Pan, L. L. Tang, C. Y. Hsu, & C. C. Chen, “Enhanced Huffman coding with encryption for wireless data broadcasting system”, in Computer, Consumer and Control (IS3C), 2012International Symposium on, pp. 622-625, Jun. 2012.

7. Muhammad Arslan Usman and Muhammad Rehan Usman, “Using Image Steganography for Providing EnhancedMedical Data security”, 2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)
8. Mingming Jiang, Guangming Tang, Yi Sun,Shunxiang Yang, “Color Image Steganography Scheme Based on FLD Ensemble Classifiers”, 2017 3rd IEEE International Conference on Computer and Communications
9. Sreekutty S Kumar, Sylish S V, “IMAGE STEGANOGRAPHY IN HIGHENTROPY REGIONS USING A KEY &MODIFIED LSB FOR IMPROVED SECURITY”, Proceedings of the IEEE 2017 International Conference on Computing Methodologies and Communication(ICCMC)
10. Tarun Jain, Anita Shrotriya, Vivek Kumar Verma, Horesh Kumar, “Mask Encryption Based Highly Secure ImageSteganography”, 2017 International Conference on Intelligent Communication and Computational Techniques (ICCT)
11. Anjana Rodrigues and Archana Bhise, “Reversible Image Steganography Using CyclicCodes and Dynamic Cover Pixel Selection”, IEEE WiSPNET 2017
12. Hetal N. Patel,Dipanjali R. Khant and Darshana Prajapati, “Design of a Color Palette Based ImageSteganography Algorithm forFractal Images”, IEEE WiSPNET 2017
13. Z. Xiao-Ping, T. Li-Sheng, P. Ying-Ning, The design of a kind of chirp-like mother wavelet by neural network, IEEE, in: Signal Processing, 1996., 3rdInternational Conference on, vol. 2, 1996, pp. 1381–1384
14. Signal and image processing institue, university of sourtherncalifornia, [Online]. Available: http://sipi.usc.edu/database/database.php?volume=misc&image=2#top.