Data Hiding and Retrieval using Integer Wavelet Transform

K. Jayasakthi velmurugan[1*], S. Hemavathi[2]

[1]Faculty, Computer Science and Engineering, Jeppiaar Engineering College, Chennai, Tamil Nadu, India
[2]Faculty, Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai, Tamil Nadu, India

**Corresponding author*

Abstract

A reversible data hiding (RDH) scheme for encrypted digital images using integer wavelet transform, histogram shifting and orthogonal decomposition is presented. This scheme takes advantage of the Laplacian-like distribution of integer wavelet high-frequency coefficients in high frequency sub-bands and the independence of orthogonal coefficients to facilitate data hiding operation in encrypted domain, and to keep the reversibility. Experimental results has demonstrated that this scheme outperforms all of other existing RDH schemes in encrypted domain in terms of higher PSNR at the same amount of payload. Compared with the state-of-the-arts, the proposed scheme can be applied to all natural images with higher embedding rate.

Keywords: *Integer wavelet transform, Reversible data hiding, Encrypted images, Histogram shifting.*

1. Introduction

Reversible data hiding (RDH) technique is able to erase completely the distortion caused by data embedding after the hidden data have been extracted out. This important technique is widely used in law forensics, medical imagery, remote sensing imagery, where no distortion of the original cover is desired (Ni et al. 2006). Many RDH techniques have been proposed in recent years, which are mainly based on the following three strategies (Zhang et al. 2014): lossless compression (Fridrich and Goljan 2002), difference expansion (DE) (Tian 2003) and histogram shift (HS) approaches (Ni et al. 2006). Almost all RDH algorithms comprise two steps. The first step generates a host sequence with small entropy, e.g., the host has a sharp histogram which usually can be realized by using prediction errors (PE) with the sorting technique (Sachnev et al. 2009) or the pixel selection (Li et al. 2011). The second step reversibly embeds the additional message bits into the host sequence. With the increasing demand of privacy protection, signal processing in encrypted domain has attracted considerable research interests.With regard to providing confidentiality for multimedia content, encryption is an effective and practical approach since it can protect multimedia information from illegal access by transforming the original information into encrypted content during the processes of transmission, storage, etc. As is well-known, partial encryption is an approach to reduce the computational resources for huge volumes of multimedia data in low power network (Puech and Rodrigues 2005;Korshunov and Ebrahimi 2014). Therefore, current image encryption mainly focus on partial encryption. However, in some scenarios, a content owner may not trust the ocessing service provider, and does not want the service provider to access the content of original multimedia, such as Cloud-based storage service. The content owner will encrypt the image before submitting (Xiong et al. 2015; Xia et al. 2016; Fu et al. 2016). The service provider would embed some additional messages into the encrypted image for authentication, notation and copyright protection. Therefore,
many RDH schemes in encrypted domain have been proposed recently. Zhang proposed ateast significant bits (LSBs) modification scheme to achieve the RDH for encrypted image (Zhang 2011). Hong et al. (2012) improved Zhang's method via using side match technique. And this scheme chooses a better metric to measure the block smoothness as preprocessing. In order to extract data, these two methods need to rely on decrypted images. However, the decrypted images may not be provided or be unknown in order to maintain the confidentiality of images. To separate data extraction from image encryption, Zhang introduced a separable method, in which a receiver having the data hiding key can extract the additional data and a receiver having the encrypted key can decrypt received data to obtain an image similar to the original one (Zhang 2012). However, these methods can only achieve low embedding capacity or generate marked image with poor quality in the case of high embedding capacity. Zhang et al. proposed a reversibility improved method. In this method, embedding space is vacated firstly by shifting the histogram of estimating errors of some pixels (Ia) chosen by only when neighborhood of these pixels includes at least T pixels in E (E is a set of estimating values, T is a threshold (1 <= T <= 8)) (Zhang et al. 2014). Therefore, the method is also defined as re-

serving room before encryption (RRBE). This kind of encoding method generally comprises three steps. The first step generates a host sequence with reserving room. The second step encrypts the host sequence. The third step reversibly embeds the additional message bits into the encrypted sequence. In the decoding phase, the data hider can extract the additional bits by a hidden key. The data decryption operator can decrypt the encrypted-marked data by a decrypted key. However, this method is not applicable to some images where no enough pixels can be selected to form vacating room, i.e., the data embedding rate is rather limited. Additionally, this method would cause high computation complexity due to the specific arrangements for embedding space. Moreover, in reversible data hiding (RDH) schemes for encrypted images, image encryption should be arranged to the content owner. Meanwhile, data embedding is supposed to be accomplished by the service provider (Wu and Sun 2014). Therefore, RDH scheme in encrypted domain needs to keep independence between the extraction and decryption steps. Based on the above analysis, this paper proposes an integer wavelet transform based scheme for reversible data hiding in encrypted images, which can boost the image quality under the same payload, and can increase the capability of data extraction and image recovery.

2. Related Work

[1] The efficient processing and analysis of the large imaging data sets that have accompanied the rise in popularity of medical imaging have presented significant challenges that have yet to be successfully overcome. Medical image analysis and volume visualization are topics that attract much attention from the image-processing community. [2]MIGS-GPU's computations are performed on the GPU by means of the compute unified device architecture (CUDA) in order to achieve fast performance and increase the utilization of available system resources. MIGS-GPU can be an advantageous and useful tool for biomedical laboratories, offering a user-friendly interface that requires minimum input in order to run. [3] The evaluation of this SVA Bayesian estimator is then relaxed into a problem that can be computed efficiently by iteratively solving a convex total-variation denoising problem and a least-squares clustering (K-means) problem, both of which can be solved straightforwardly, even in high-dimensions, and with parallel computing techniques. [4] Analysis of leukocytes provides valuable information to medical specialists, helping them in diagnosing different important hematic diseases, such as AIDS and blood cancer (Leukaemia). However, this task is prone to errors and can be time-consuming. the detected WBCs are categorized into five classes: basophil, eosinophil, neutrophil, lymphocyte, and monocyte. [5] The detection of a brain tumor and its classification from modern imaging modalities is a primary concern, but a time-consuming and tedious work was performed by radiologists or clinical supervisors. The accuracy of detection and classification of tumor stages performed by radiologists is depended on their experience only, so the computer-aided technology is very important to aid with the diagnosis accuracy. [6] The retinal images provide vital information about the health of the sensory part of the visual system. Retinal diseases, such as glaucoma, diabetic retinopathy, age-related macular degeneration, Stargardt's disease, and retinopathy of prematurity, can lead to blindness manifest as artifacts in the retinal image. [7] Automated melanoma recognition in dermoscopy images is a very challenging task due to the low contrast of skin lesions, the huge intraclass variation of melanomas, the high degree of visual similarity between melanoma and non-melanoma lesions, and the existence of many artifacts in the image. In order to meet these challenges, we propose a novel method for melanoma recognition by leveraging very deep convolutional neural networks (CNNs).

3. Proposed System and Methodology

Iinteger wavelet transform (IWT) and Encryption method used in this paper. Then we introduce data hiding algorithm to embed additional data into the encrypted image. In this way, the additional data is embedded in encrypted domain..

3.1. Architecture Diagram

Firstly, an image is processed using the integer wavelet transform (IWT). Secondly, all frequency coefficients are encrypted. Thirdly, the encrypted frequency coefficients are embedded with additional data. Fourthly, the encrypted coefficients containing embedded data can be processed by Integer Wavelet Reverse Transform to obtain an encrypted image containing embedded data. In the decryption phase, we can obtain the decrypted image containing embedded data. In the data extraction phase, the embedded data can be directly extracted from encrypted high frequency coefficients containing embedded data and we can also obtain encrypted image without embedded data. Meanwhile the embedded data can also be extracted from decrypted image containing embedded data. Finally, we can obtain a recovered image

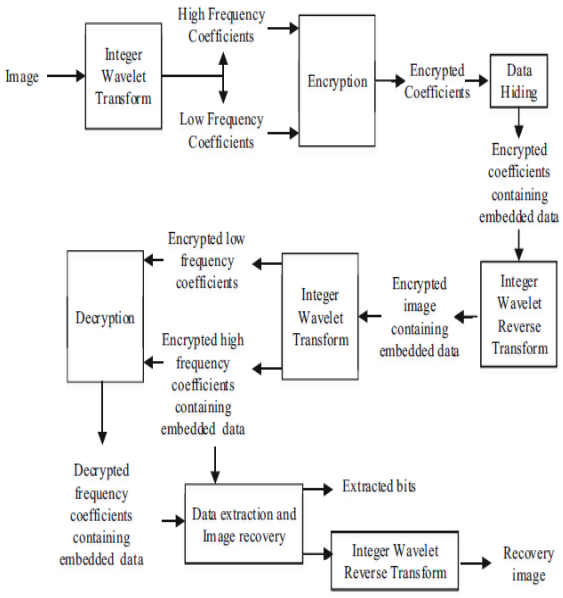

Fig.3.1.Architecture Diagram

3.2 Integer Wavelet Transform

We use the integer lifting scheme based on wavelet transform in this framework as it is required to reconstruct the original image. We adopt the CDF(2,2) integer wavelet transform in the experiment as the same in Xuan et al. (2002, 2006). After the IWT, it has four sub-bands. Denote the high frequency coefficients as X, and the other coefficients as $\acute{X}$.

Using orthogonal decomposition based on B, X can be represented as $X = B \cdot Y$ or $Y = B{-}1 \cdot X$, where Y is a set of the orthogonal coefficients. The matrix B can be divided into two sub-matrixes, i.e., $B = (R, S)$, thus R and S can be expressed as $R = (b1, b2, \ldots, bm)$ and $S = (bm{+}1, bm{+}2, \ldots, bn)$. The vector Y is also divided into two sub-vectors, i.e., $Y = (Y1, Y2)T$. Therefore, R and S can be used to control $Y1$ and $Y2$ separately.

3.3. Encryption

In this scheme, as a data encryption operator, R and X are known. Firstly, we compute $Y1$ by Eq. (3), then $Y1$ is encrypted with Eq. (5)

$Y1e = Enc\ (Y1, K) = (Enc\ (RT \cdot R){-}1 \cdot RT \cdot X, K)$ -(5)

Where $Y1e$ is the encrypted $Y1$, $Enc(\cdot, \cdot)$ represents an encryption algorithm and K is an encryption key. Therefore, we obtain encrypted X according to Eqs. (2)–(4) as follows. $Xe = R \cdot Y1e{+}S \cdot Y2 = R \cdot Enc\_(RT \cdot R){-}1 \cdot RT \cdot X, K\_{+}X - R(RT \cdot R){-}1 \cdot RT \cdot X$ (6) where Xe is the encrypted X. Similar to Eq. (3), $Y1e = (RT \cdot R){-}1 \cdot RT \cdot Xe$, Eq. (6) takes as input K, R and X, and outputs Xe. And other frequency coefficients $\acute{X}$ are also encrypted with $Enc(\cdot, \cdot)$ as follows. $\acute{X}e = Enc(\acute{X}, K)$

Therefore, we can obtain an encrypted whole data. Define $\tilde{X}e$ as the encrypted whole data. $\tilde{X}e = \acute{X}e||Xe$ where _ represents the combination of four sub-bands.

4. Data Embedding Process

Set a threshold T>0, to let the number of the coefficients in [-T, T] is greater than M. And set the Peak = T. Step 2. In the histogram of coefficients, move the histogram (the value is greater than Peak) to right-hand side by one unit to leave a zero-point at the value Peak+1. Then embed data in this point. Step 3. If there are to-be-embedded data remaining, let Peak = (-Peak), and move the histogram (less than Peak) to the left-hand side by 1 unit to leave a zero-point at the value (-Peak-1). And embed data in this point. Step 4. If all the data are embedded, then stop here and record the Peak value as stop peak value, S. Otherwise, Peak = (-Peak-1), go back to (2) to continue to embed the remaining to-be-embedded data.

5. Data Extraction Process

Step 1. Set Peak = S. Step 2. Decode with the stopping value Peak. Extract all the data until Peak+1 becomes a zero-point. Move all the histogram (greater than Peak+1) to the left-hand by one unit to cover the zero-point .Step 3. If the extracted data is less than M. Set Peak = -Peak-1. Continue to extract data until it becomes a zero-point in the position (Peak-1). Then move histogram (less than Peak-1) to the right-hand side by one unit to cover the zero-point. Step 4. If all the hidden bits have extracted, stop. Otherwise, set Peak = -Peak, go back to (2) to continue to extract the data.

6. Decryption and Data Extraction

Firstly Separating Xeh from $\acute{X}e||Xeh$ by wavelets transform. And then, the data hider extracts the embedded data from Xeh. At the same time, the data decryption operator decrypts Xeh and $\acute{X}e$ with a decrypted key. There are two scenarios considered in Zhang et al. (2014): data extraction before image decryption and image decryption before data extraction.

Data extraction before image decryption In this scheme, for the data hider, S and Xeh are known. According to Sect. 2, the data hider can get $Y2h$ from the encrypted high frequency coefficients containing embedded data, Xeh. The procedures are as follows.

$$Y2h = (ST \cdot S){-}1 \cdot ST \cdot Xeh$$

Here, we obtain $Y2h$, and then the embedded data can be extracted from $Y2h$ with extraction procedures described in Section II.C. Similarly, $Y2h$ can be also recovered to $Y2$.

Image decryption before data extraction

At the same time, for the data decryption operator, R, Xeh and $\acute{X}e$ are known. The procedures are as follows.

$Y1e = (RT \cdot R){-}1 \cdot RT \cdot Xeh$, $Y1 = Dec\ (Y1e, K) = Dec\_(RT \cdot R){-}1 \cdot RT \cdot Xeh$, $K\_$, and $\acute{X} = Dec(\acute{X}e, K)$where $Dec(\cdot, \cdot)$ is a decryption function. In this way, we can obtain a decrypted data containing embedded data, $\acute{X}||Xh$, by the following equation. $Xh = R \cdot Dec\_(RT \cdot R){-}1 \cdot RT \cdot Xeh$,

$$K\_{+}\ Xeh - R \cdot (RT \cdot R){-}1 \cdot RT \cdot Xeh \quad (10)$$

Based on the above analysis, the R and S can be used as the control codes of encryption and data hiding respectively, which keep the independence between data extraction and image decryption

6.1 Integer Wavelet Transformation

We first introduce integer wavelet transform (IWT) and encryption method used in the paper. Then we introduce data hiding algorithm to embed additional data into the encrypted image. In this way, the additional data is embedded in encrypted domain

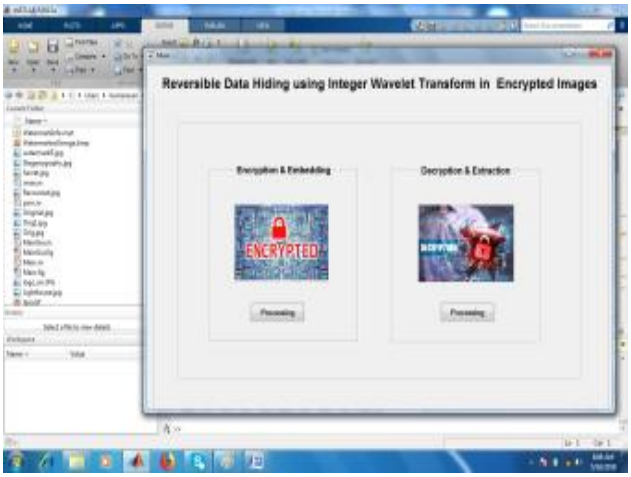

Fig.3.2 Data Hiding using Integer Wavelet Transformation

Then we introduce data hiding algorithm to embed additional data into the encrypted image.

In this way, the additional data is embedded in encrypted domain. Firstly, an image is processed using the integer wavelet transform (IWT).

6.2 Encryption and Data Hiding Process

Secondly, all frequency coefficients are encrypted. Thirdly, the encrypted frequency coefficients are embedded with additional data.

6.3 Decryption and Data Extraction Process

In the decryption phase, we can obtain the decrypted image containing embedded data. In the data extraction phase, the embedded data can be directly extracted from encrypted high frequency coefficients containing embedded data and we can also obtain encrypted image without embedded data. Meanwhile the embedded data can also be extracted from decrypted image containing embedded data. Finally, we can obtain a recovered image.

6.4 Reversible Integer Wavelet Transformation

The encrypted coefficients containing embedded data can be processed by Integer Wavelet Reverse Transform to obtain an encrypted image containing embedded data. In the decryption phase, we can obtain the decrypted image containing embedded data.

7. Implementation

We first introduce integer wavelet transform (IWT) and encryption method used in the paper. Then we introduce data hiding algorithm to embed additional data into the encrypted image. In this way, the additional data is embedded in encrypted domain. Firstly, an image is processed using the integer wavelet transform (IWT). Secondly, all frequency coefficients are encrypted. Thirdly, the encrypted frequency coefficients are embedded with additional data. Fourthly, the encrypted coefficients containing embedded data can be processed by Integer Wavelet Reverse Transform to obtain an encrypted image containing embedded data. In the decryption phase, we can obtain the decrypted image containing embedded data. In the data extraction phase, the embedded data can be directly extracted from encrypted high frequency coefficients containing embedded data and we can also obtain encrypted image without embedded data. Meanwhile the embedded data can also be extracted from decrypted image containing embedded data. Finally, we can obtain a recovered image.

Result and Discussion

The proposed scheme can embed much more data into the Baboon and Barbara images than the scheme reported in Zhangit has a wider applicability. Additionally, the other data embedding methods can be also applied in this scheme

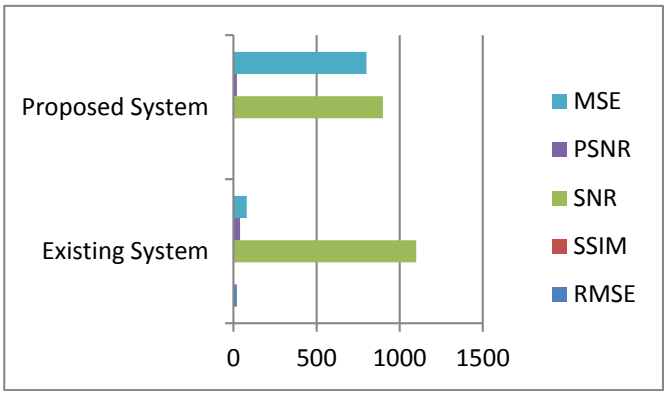

Fig.3.3.Comparison Graph

Conclusion

A novel scheme for reversible data hiding in encrypted images has been reported. It is based on Integer Wavelet Transformation (IWT), histogram shifting and orthogonal decomposition. In the proposed scheme, integer wavelet high-frequency coefficients have a Laplacian-like distribution and the histogram shifting technique can be well carried out in these coefficients. The independence of orthogonal coefficients (Y) can facilitate data hiding operation in encrypted domain and keep the reversibility. The experimental results has shown that its superior performance over the current state of- the-arts in terms of higher PSNR at the same amount of payload. The proposed scheme can embed much more data into the Baboon and Barbara images than the scheme reported in Zhangit has a wider applicability. Additionally, the other data embedding methods can be also applied in this scheme. The paper aims to propose a novel scheme for reversible data hiding in encrypted images. As for the other data embedding and encryption methods, we will discuss these methods and explore a better performance method in future work.

REFERENCES

1. 1] Fridrich, J., & Goljan, M. (2002). Lossless data em embedding for all image formats. In SPIE proceedings of photonics west, electronic imaging, secure ty and watermarking of multimedia contents (Vol. 4675, pp. 572–583). San Jose.
2. 2] Fu, Z., Ren, K., Shu, J., Sun, X., & Huang, F. (2016). Enabling personalized search over encrypted outsourced data with efficiency improvement. IEEE Transactions on Parallel and Distributed Systems, 27(9), 2546– 2559.
3. 3] Hong, W., Chen, T. S., & Wu, H. Y. (2012). An improved reversible data hiding in encrypted images using side match. IEEE Signal Processing Letters, 19(4), 199–202.
4. 4] Korshunov, P., & Ebrahimi, T. (2014). Scrambling-based tool for secure protection of JPEG images. In 2014 IEEE international conference on image processing (ICIP) (pp. 3423–3425).
5.

5] Li, X. L., Yang, B., & Zeng, T. Y. (2011). Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection. IEEE Transactions on Image Processing, 20(12), 3524–3533.

6] http://decsai.ugr.es/cvg/dbimagenes/g512.php. Ni, Z. C., Shi, Y. Q., Ansari, N., & Su, W. (2006). Reversible data hiding. IEEE Transactions on Circuits and Systems for Video Technology, 16(3), 354–362.

7] Puech,W.,&Rodrigues, J. M. (2005). Crypto-compression of medical images by selective encryption of DCT. In 2005 13th European signal processing conference (pp. 1–4). IEEE.

8] Sachnev, V., Kim, H. J., Nam, J., Suresh, S., & Shi, Y. Q. (2009). Reversible watermarking algorithm using sorting and prediction. IEEE Transactions on Circuits and Systems for Video Technology, 19(7), 989–999.

9] Tian, J. (2003). Reversible data embedding using a difference expansion. IEEE Transactions on Circuits and Systems for Video Technology, 13(8), 890–896.

10] Wu, X.,&Sun,W. (2014).High-capacity reversible data hiding in encrypted images by prediction error. Signal Processing, 104, 387–400.

11] Xia, Z., Wang, X., Zhang, L., Qin, Z., Sun, X., & Ren, K. (2016). A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing. IEEE Transactions on Information Forensics and Security,. doi:10.1109/TIFS.2016.2590944.

12] Xiong, L., Xu, Z., & Xu, Y. (2015). A secure re-encryption scheme for data services in a cloud computing environment. Concurrency and Computation: Practice and Experience, 27(12), 4573–4585.

13] Xuan, G.,Yao, Q.,Yang, C.,Gao, J., Chai, P., Shi,Y. Q.,&Ni, Z. C. (2006).Lossless data hiding using histogram shifting method based on integer wavelets. In 2016 international workshop on digital watermarking (IWDW), lecture notes in computer science (Vol. 4283, pp. 323–332). Berlin, Heidelberg: Springer.

14] Xuan, G., Zhu, J., Chen, J., Shi, Y. Q., Ni, Z., & Su, W. (2002). Distortionless data hiding based on integer wavelet transform. Electronics Letters, 38(25), 1646–1648.

15] Zhang, X. P. (2011). Reversible data hiding in encrypted image. IEEE Signal Processing Letters, 18(4), 255– 258.

16] Zhang, X. P. (2012). Separable reversible data hiding in encrypted image. IEEE Transactions on Information Forensics and Security, 7(2), 826–832.

17] Zhang,W. M.,Ma, K. D., & Yu, N. H. (2014). Reversibility improved data hiding in encrypted images. Signal Processing, 94, 118–127.