An Efficient Steganography Method for Hiding Patient Confidential Information

Hayat Al-Dmour, IEEE Student Member, Ahmed Al-Ani and Hung Nguyen

Abstract— **This paper deals with the important issue of security and confidentiality of patient information when exchanging or storing medical images. Steganography has recently been viewed as an alternative or complement to cryptography, as existing cryptographic systems are not perfect due to their vulnerability to certain types of attack. We propose in this paper a new steganography algorithm for hiding patient confidential information. It utilizes Pixel Value Differencing (PVD) to identify contrast regions in the image and a Hamming code that embeds 3 secret message bits into 4 bits of the cover image. In order to preserve the content of the region of interest (ROI), the embedding is only performed using the Region of Non-Interest (RONI).**

Index Terms—**Steganography, PVD, hamming code, ROI.**

I. INTRODUCTION

There has been an increase reliance on the Internet for the transmission of information between various organization and administrative systems. This transmission requires a secure network that is highly resistant to fraud and data theft.In particular, dispersing and maintaining medical records and data has become quite important for a number of medical applications. For example, during an orthopaedic surgery, the surgeon may need to send data messages that contain patient's details and medical images to a pathologist and/or radiologist. However, it is essential during this process to maintain the security and privacy of the patient's information. In fact, the US Department of Health and Human Services (DHHS) foisted a law for data security and privacy under the Health Insurance Portability and Accountability Act (HIPAA) 1996 [8].

A number of hospitals use cryptography for protecting patients' information. Cryptography is the process of hiding the meaning of message by encrypting the message [3], [5]. However, the security of cryptographic systems is not guaranteed as each cryptography algorithm is vulnerable to certain type(s) of attack, which would enable intruders to gain access to patients' information. Accordingly, steganography has been introduced to overcome some of the limitations of cryptography [5]. Cox et. al. [4] defined steganography as the little and much younger sister of cryptography, as it represents an alternative tool for privacy and security. Unlike cryptography that encrypts the message, steganography hides the message in innocuous-looking objects, such that its presence is concealed. Digital media such as text, image, audio or video is used as the cover object. Steganography has started to play a vital role in hiding data in digital media. According to [1], steganography can also be used in the implementation of multi-level access control in medical image transmission.

Least Significant Bit (LSB) substitution is a well-known steganoraphy technique, which embeds a secret message of L bits in L pixels of the cover image. This simple technique however causes noticeable distortion when the number of embedded bits per pixel exceeds three [3]. Wu and Tasi proposed the Pixel Value Differencing (PVD) technique to upgrade the embedding rate without introducing noticeable visual artefacts to the stego image. It aims to embed more bits in sharp contrast areas of the image, while keeping modifications to smooth areas at a minimal level. PVD determines the number of secret bits to be embedded by calculating the difference between each two adjacent pixels [9].

In this paper we introduce a new biomedical image steganography algorithm based on PVD to identify the pixel locations of sharp contrast regions for embedding, and a (7,4) hamming code to conceal the secret message. Our scheme not only embeds the secret message data into the cover image but also maintain the quality of an image and its security. In addition, the region of interest (ROI) is used to preserve the original medical information.

The rest of this paper is organized as follows.The related work is described in section 2. Details of the proposed method are presented in section 3. Section 4 presents the experimental results, and the conclusion is given in section 5.

II. RELATED WORK

Jain et. al. proposed a steganography method to conceal biometric data in cover images [6]. This could be convenient in disseminated systems where the patient information may be transmitted over non-secure network. Also, hiding biometric data in cover images might help keep intruder from getting sensitive data. Prabakaran et. al. proposed a method that gives proficient and storage security mechanism to protect digital medical images. Integer Wavelet Transform (IWT) was utilized to secure an MRI medical image into a single holder. The holder image was taken and flip left to obtain a dummy holder. At that point the patient's medical diagnosis image was taken as mystery image and Arnold transform was performed and scrambled secret image was acquired. The scrambled secret image was hidden into the dummy holder and Inverse IWT was applied to get the secret image [7].

H. Al-Dmour, A. Al-Ani and H. Nguyen are with the Faculty of Engineering and Information Technology, University of Technology, Sydney, NSW 2007, Australia (e-mail: HayatShahir.T.Al-Dmour@student.uts.edu.au, Ahmed.Al-Ani@uts.edu.au, Hung.Nguyen@uts.edu.au).

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

Zhou et al. introduced a strategy that attaches digital signature and electronic patient record (EPR) into a medical image. Their strategy utilizes LSB substitution method to hide the signature [10]. Chao et al. suggested a protected information concealing procedure the is dependent upon the bipolar mutiple-base transformation to permit a mixed of EPR information to be concealed inside the same mark image. The mark image could be the sign of a clinic used to recognize the origin of an EPR. Their strategy permits dividing and rebuilding of concealed information by authorized clients [2]. The Pixel Value Differencing (PVD) steganography technique has attracted the attention of many researchers due to its high embedding capacity and ability to maintain the visual artefacts at a minimal level. It divides the cover image into non-overlapping blocks of two consecutive pixels (p_i and p_{i+1}). The absolute difference value, $d = |p_i - p_{i+1}|$ is categorized into one of pre-defined non-overlapping ranges. The number of secret bits that can be embedded in the two pixels (p_i and p_{i+1}) depends on the range that d belongs to [9].

The next section presents our proposed method, which takes into consideration preserving the ROI from any modifications and is based on PVD and a (7,4) hamming code.

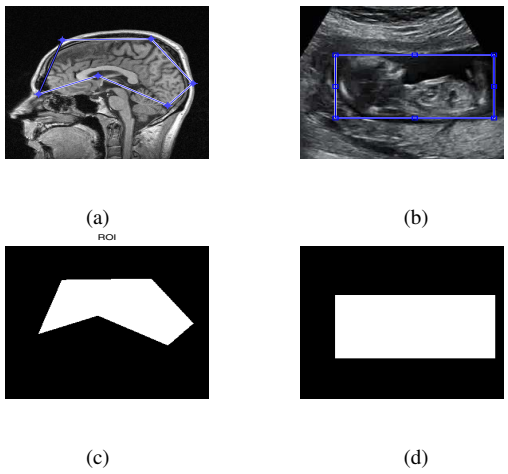


Fig. 1. (a) and (b): ROI of MRI and ultrasound cover images. (c) and (d): Corresponding outline of ROI

TABLE I
PVD RANGE TABLE

SR1	SR2	SR3	SR4	SR5	SR6
[0-7]	[8-15]	[16-31]	[32-63]	[64-127]	[128-255]
Smooth Region		Sharp Region			

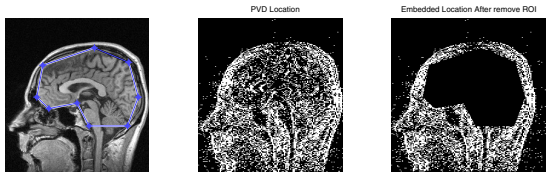
III. THE PROPOSED METHOD

It is well-known that when observing an image, the human visual system (HVS) is less sensitive to changes in sharp contrast areas compared to uniform areas. Thus, it is logical to mainly consider image regions that have strong differences between adjacent pixels. Because scans of different parts of the body usually have such property, we decided to choose PVD as basis for our proposed steganography method. In order to increase the embedding efficiency and security, the proposed method also utilizes a Hamming code. Below is a description of the message embedding and extraction processes.

A. The Embedding Process

The embedding process is implemented as described in the following steps.

1) *Identification of the ROI:* In order to sustain the most valuable part of the image, the user is asked to identify the ROI. The secret message will only be embedded into the region of non-interest (RONI), while keeping the ROI not affected. A binary image is generated by this process, where pixels that have value of '1' belong to ROI, while those that have a value of 0 belong to RONI, as shown in fig.4. Only pixels of RONI will be utilized in embedding the message, while those of ROI will not be modified. The coordinates of ROI are stored in the last row of the stego image to allow the extraction process to exactly identify the ROI pixels.
2) *PVD:* In the second stage, the cover image is divided into blocks of two consecutive non overlapping pixels. The grey level difference between the two pixels of each block is calculated and then classified into six sub ranges as shown in table I. Blocks formed using ROI pixels are excluded in order to protect the important region of the image from being modified, and hence, only blocks that belong to RONI are considered for embedding.
This method does not use every block in the RONI to hide the secret message, for that we defined threshold value to improve embedding efficiency and security. The threshold variable is used to select blocks with high difference to hide the message, because as mentioned earlier hiding data in sharp regions is hard to detect by the human eye compared to smooth regions. The difference of each block is compared to the initial threshold value of 128. If the numbers of blocks that are greater than the threshold (can be used for embedding) are enough to embed the entire message, then the process halts. Else the threshold is lowered to the first value of the next smaller sub-region (64 for SR5). If it is still not possible to embed the entire message based on this new threshold, then the threshold is lowered to 32. This process continues until there are enough blocks to embed all bits of the message (the two LSBs of each of the selected pixels are used for embedding). The threshold value is stored in a pre-defined pixel of the cover image.
Fig. 2(a) shows the cover image with an outline of the ROI. Fig. 2(b) shows the resulting PVD image, while Fig. 2(c) highlights the potential embedding locations after excluding the ROI.
3) *Embedding using a (7,4) Hamming Code:* In this method, 2 LSBs from each selected pixel is used in embedding to enhance the embedding capacity. To reduce the number of modified pixels we use a (7, 4)



(a) (b) (c)

Fig. 2. (a) ROI (b) PVD Location (c) Final Location for Embedding

Hamming code. In this stage, hamming code is used to hide each 3 bits of the secret data (m_1, m_2 and m_3) into 4 cover bits (p_1, p_2, p_3 and p_4) that come from two embedding pixels (selected as described in the previous step). A codeword, C, is formed by arranging the seven bits in the following order: $C = [m_1, m_2, p_1, m_3, p_2, p_3, p_4]$. The parity check matrix and the codeword are multiplied to determined which of the cover pixels need to be modified, as shown in Eq. 1.

$$S = H * C^T \quad (1)$$

$$\text{where } H = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

where S indicates the bits that need to be modified based on the codeword C. If S is equal to the 3rd, 5th, 6th or 7th column of H, then one bit is changed from the cover pixels. When $S = [0, 0, 1]$ then p_1 is changed. If S is equal to the 1st, 2nd or 4th column of H, then two bits of cover pixels are changed. For example, if $S = [1, 0, 0]$ then p_3 and p_4 are changed. If $S = [0, 1, 0]$ then p_2 and p_4 are changed. Finally, if $S = [1, 1, 0]$ then p_1 and p_4 are changed.

4) *Correction of pixel values after embedding:* It is necessary to check the new difference of each block after operation embedding. If the new difference of the two corresponding pixels remains within the same sub-range, then there is no need for modifications. However, if the new difference belongs to a different sub-range, then then 3rd LSB of one of the pixels should be changed so that the new and the old differences belong to the same sub-range.

B. The Extraction Process

The extraction process begins by retrieving the coordinates of ROI from the last row of the stego image. The threshold value is also retrieved. After that, the stego image is partitioned into blocks of two pixels to determine the blocks used to embed the secret message. If a given block does not belong to the ROI and the difference value of its two pixels is greater than or equal to the threshold, then three message bits will be extracted from it, otherwise it will not be used for extraction. Finally, the three secret bits m_1, m_2 and m_3 will be extracted from the block using the following XOR operations:

$$m_1 = p_1 \oplus p_2 \oplus p_4$$
$$m_2 = p_1 \oplus p_3 \oplus p_4$$
$$m_3 = p_2 \oplus p_3 \oplus p_4$$

IV. Experimental Results

In this section, the proposed method is evaluated using 5 Ultrasound and 5 MRI cover images (all of them are gray level of size 256×256). The length of the secret message (data capacity) is used as one of the evaluation criteria, which is defined as the amount of bits that can be embedded into the cover image. The embedding capacity is computed using Eq. 2.

$$E = \frac{K}{WH}(bpp) \quad (2)$$

where K is the number of the data message bits, while W and H are the width and height of the cover image respectively (both cover and stego images are of the same size. For the considered images, $W = H = 256$). There is no unique method to measure imperceptibility of steganography methods. One of the commonly used measures of imperceptibility is the Peak Signal-to-Noise Ratio (PSNR) between the cover and stego images, which is calculated as shown in Eq. 3.

$$PSNR = 20 \log_{10} \left(\frac{255}{MSE} \right) (dB) \quad (3)$$

where MSE is the mean square error between cover and stego images, which is defined as:

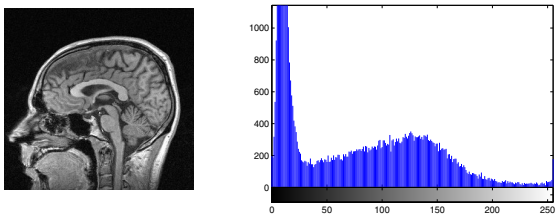
$$MSE = \frac{1}{WH} \sum_{i=1}^{W} \sum_{j=1}^{H} (c_{ij} - s_{ij})^2 \quad (4)$$

where c_{ij} and s_{ij} are the gray values of pixel (i, j) of the cover and stego images respectively. The weighted Peak signal-to-Noise Ratio ($wPSNR$) is an alternate measurement imperceptibility. It utilizes an extra parameter called Noise Visibility function (NVF). $wPSNR$ is roughly equivalent to PSNR for flat areas because NVF is close to one in smooth regions. However, for regions with sharp contrasts, $wPSNR$ is higher than $PSNR$, because NVF is close to zero for complex regions. Hence, $wPSNR$ attempts to reflect how the HVS perceives images.

$$wPSNR = 10 \log_{10} \left(\frac{\max(C)^2}{\|NVF(S - C)\|^2} \right) (dB) \quad (5)$$

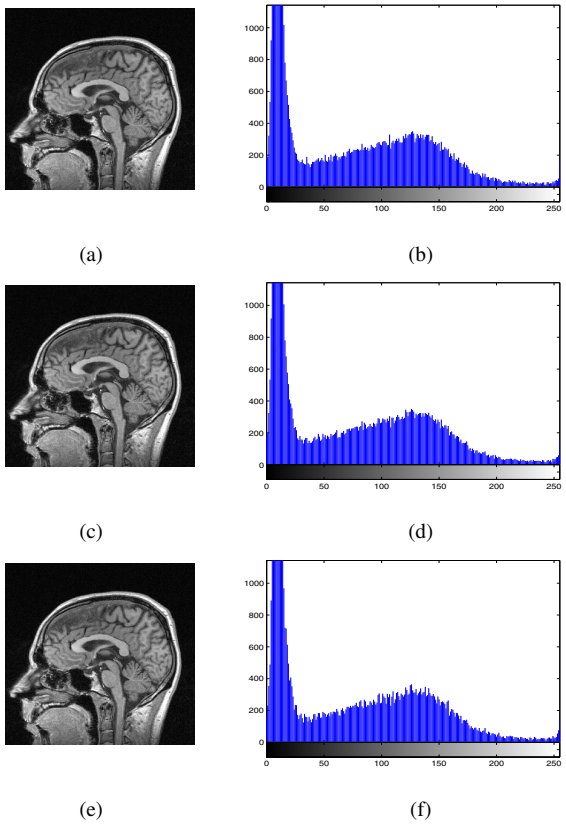
Figs. 3(a) and 3(b) show one of the cover images used in the experiment and its histogram.

Figs. 4(a), 4(c) and 4(e) show the stego images produced by our method using different threshold values, i.e., different message lengths. As explained in the previous section, the stego image is produced by embedding message bits in blocks that have difference values greater than or equal to the threshold. The three figures indicate that it is almost impossible to visually differentiate between any of the stego images and the cover image. The histograms of the sego images shown in Figs. 4(b), 4(d) and 4(f) represent a high degree of similarity between the four images.



(a) (b)

Fig. 3. (a) Cover image 256×256 (b) cover image histogram



(a) (b)

(c) (d)

(e) (f)

Fig. 4. (a), (c) and (e) Stego Image using Th =32 , 16, and 8 resectively (b), (d) and (f) Stego Image Histogram using Th =32 , 16, and 8 resectively

Table II indicate the quality of the stego images generated by our proposed method. We can see that even when applying a threshold value of 8, the $PSNR$ is more than 52 dB and the $wPSNR$ is more than 67 dB, which indicate a very high degree of similarity with the cover image. As one would expect, the degree of similarity is even higher when consider a higher threshold, however this comes with the price of compromising the payload. The threshold value of 8 enables the hiding of a large amount of patient record data (up to 32300 bits in total, considering that 2 bits are utilized from each pixel). This represents a very acceptable payload, given the fact that the ROI was not used in embedding the secret message.

TABLE II

COMPARISON USING DIFFERENT THRESHOLD VALUES

	Threshold	64	32	16	8
MRI image	Data Length (bits)	2323	8775	19800	32300
	Embedding Rate	3.54%	13.38%	30.21%	49.28%
	MSE	0.028	0.119	0.263	0.408
	PSNR	62.33	57.39	53.93	52.03
	wPSNR	76.20	72.70	69.15	67.08
	SSIM	1.00	0.999	0.997	0.999
	Average Difference	-0.0004	-0.0013	0.0006	-0.002
ultrasound	Data Length (bits)	117	1323	7980	22428
	Embedding Rate	0.178%	2.01%	12.17%	34.22%
	MSE	0.0022	0.0217	0.1310	0.2994
	PSNR	74.73	64.76	56.96	53.37
	wPSNR	88.53	79.11	71.23	68.32
	SSIM	1.00	1.00	0.9998	0.9997
	Average Difference	-0.0002	-0.0004	-0.001	-0.003

V. CONCLUSIONS

This paper has introduced a steganography method for biomedical images in order to protect patient's information. This method achieves both high payload and good quality of stego image with $PSNR$ values of more than 50 dB. The highly efficient performance of the proposed method is achieved by using PVD to select sharp regions for embedding, which are less sensitive to changes by the HVS. In addition, the distortion is further reduced by using a hamming code for concealing secret data, which also adds security to the embedded message.

REFERENCES

[1] J Nafeesa Begum, Krishnan Kumar, and V Sumathy. Design and implementation of multilevel access control in medical image transmission using symmetric polynomial based audio steganography. *arXiv preprint arXiv:1004.1682*, 2010.

[2] Hui-Mei Chao, Chin-Ming Hsu, and Shaou-Gang Miaou. A data-hiding technique with authentication, integration, and confidentiality for electronic patient records. *Information Technology in Biomedicine, IEEE Transactions on*, 6(1):46–53, 2002.

[3] Abbas Cheddad, Joan Condell, Kevin Curran, and Paul Mc Kevitt. Digital image steganography: Survey and analysis of current methods. *Signal Processing*, 90(3):727–752, 2010.

[4] Ingemar Cox, Matthew Miller, Jeffrey Bloom, Jessica Fridrich, and Ton Kalker. *Digital watermarking and steganography*. Morgan Kaufmann, 2008.

[5] Nagham Hamid, Abid Yahya, R Badlishah Ahmad, and Osamah M Al-Qershi. Image steganography techniques: an overview. *International Journal of Computer Science and Security (IJCSS)*, 6(3):168–187, 2012.

[6] Anil K Jain, Arun Ross, and Umut Uludag. Biometric template security: Challenges and solutions. In *Proceedings of European Signal Processing Conference (EUSIPCO)*, pages 469–472. Citeseer, 2005.

[7] G Prabakaran, R Bhavani, and PS Rajeswari. Multi secure and robustness for medical image based steganography scheme. In *Circuits, Power and Computing Technologies (ICCPCT), 2013 International Conference on*, pages 1188–1193. IEEE, 2013.

[8] Yeshwanth Srinivasan, Brian Nutter, Sunanda Mitra, Benny Phillips, and Daron Ferris. Secure transmission of medical records using high capacity steganography. In *Computer-Based Medical Systems, 2004. CBMS 2004. Proceedings. 17th IEEE Symposium on*, pages 122–127. IEEE, 2004.

[9] Da-Chun Wu and Wen-Hsiang Tsai. A steganographic method for images by pixel-value differencing. *Pattern Recognition Letters*, 24(9):1613–1626, 2003.

[10] XQ Zhou, HK Huang, and Shieh-Liang Lou. Authenticity and integrity of digital mammography images. *Medical Imaging, IEEE Transactions on*, 20(8):784–791, 2001.