International Journal of Biotechnology and Biomedical Sciences
p-ISSN 2454-4582, e-ISSN 2454-7808, Volume 2, Issue 3; July-December, 2016 pp. 229-233

http://www.krishisanskriti.org

Reversible Data Hiding Using Integer to Integer Wavelet Transform for Medical Images

Shruthi N[1], Savitha C[2] and M Z Kurian[3]

[1]M.Tech Student, Digital Electronics, SSIT, Tumkur, India, 572105
[2,3]Dept. of ECE, SSIT, Tumkur, India, 572105
E-mail: [1]shruthi16dec@gmail.com, [2]savitha_c21@yahoo.com, [3]mzkurian@yahoo.com

Abstract—*A digital watermark is a kind of marker embedded in a cover media such as image data. Watermarking is one of the promising solutions for protection of hidden digital content. However, watermarking may cause damage to the sensitive information present in the cover image. Therefore, at the receiving end, the exact recovery of cover image may not be possible. Additionally, there exist certain applications that may not tolerate even small distortion in the cover image. In such applications, instead of conventional watermarking, reversible data hiding is employed. Reversible data hiding (RDH) is a technique which enables images to be authenticated and then restored to their original form by removing the digital watermark and replacing the image data that had been overwritten. Histogram shifting (HS) is a useful technique of reversible data hiding. Invertible integer to integer wavelet transform based on lifting is issued in the proposed scheme. It maps integers to integers and does not cause any loss of information through forward and inverse transforms. In medical images, generally, there is very large background area, which contains no information and this leads to the sequence of 0's. This sequence of 0's in medical images can be used for data embedding. In this paper, a reversible data hiding scheme for medical images is proposed which utilizes integer to integer wavelet transform and histogram modification for watermark embedding.*

Index Terms: *Reversible data hiding, histogram shifting, entropy, integer to integer wavelet transform, medical images.*

1. INTRODUCTION

Data hiding is a process of hiding data into a cover media. Media can be anything like audio, image and video. Hiding is done by modifying the contents of the digital media. If the cover media is an image, then data hiding is done by modifying some pixels of the image. The hiding process is done in such a way that modification of pixel values should be undetectable to the viewers. Image after data hiding is known as watermarked image.

The data hiding process links two sets of data, a set of the embedded data and another set of the cover media data. The relationship between these two sets of data characterizes different applications. For instance, in covert communications, the hidden data may often be irrelevant to the cover media. In authentication, however, the embedded data are closely related to the cover media. In these two types of applications, invisibility of hidden data is an important requirement. In most cases of data hiding, the cover media will experience some distortion due to data hiding and cannot be inverted back to the original media. That is, some permanent distortion has occurred to the cover media even after the hidden data have been extracted out. In some applications, such as medical diagnosis and law enforcement, it is critical to reverse the marked media back to the original cover media after the hidden data are retrieved. The marking techniques satisfying this requirement are referred to as reversible, lossless, distortion-free, or invertible data hiding techniques. Reversible data hiding (RDH) facilitates immense possibility of applications to link two sets of data in such a way that the cover media can be recovered without loss after the hidden data have been extracted out, thus providing an additional avenue of handling two different sets of data.

Histogram shifting (HS) is one among various reversible data hiding techniques. HS method utilizes the histogram of the cover image. Based on the peak and zero points of the histogram, pixels values are shifted for embedding the secret data.

Conventional wavelet transform is not applicable to the reversible watermarking scheme since it does not guarantee the reversibility. For example, an image block consisting of integer-valued pixels is transformed into a wavelet domain using a floating-point wavelet transform. If the values of the wavelet coefficients are changed during watermark embedding, the corresponding watermarked image block is no longer guaranteed to have integer values. Any truncation of the floating-point values of the pixels may result in a loss of information and may ultimately lead to the failure of the reversible watermarking systems, that is, the original image cannot be reconstructed from the watermarked image. Furthermore, the conventional wavelet transform is, in practice, implemented as a floating-point transform followed by a truncation or rounding since it is impossible to represent transform coefficients in their full accuracy. To avoid this problem, an invertible integer-to-integer wavelet transform based on lifting is issued in the proposed scheme. It maps integers to integers and does not cause any loss of information through forward and inverse transforms.

2. LITERATURE REVIEW

In [1] Ni et al. proposed the first histogram shifting (HS) based RDH method. This method utilizes peak (maxima) and minimum points of the pixel-intensity-histogram to embed the data. HS adds gray values to some pixels to shift a range of classes of the image histogram to create a gap near the histogram maxima. Pixels which belong to the class of the histogram maxima (carrier class) are then shifted to the gap or kept unchanged to encode one bit of the message i.e, '0' or '1'. The other pixels (non-carriers) are simply shifted.

Basic idea is to shift each pixel value at most by 1, thus a good visual quality of marked image can be obtained. The PSNR of the marked image versus the original image is guaranteed to be higher than 48 dB. However, its embedding capacity is quite low and this method does not work well if the cover image has a flat histogram.

Tian [2] proposed differential expansion (DE) algorithm for a pair of pixels to devise a high-capacity and low-distortion reversible watermark. In DE algorithm, the host image is divided into pixel pairs, and the difference value of two pixels in each pair (that are not expected to cause an overflow or underflow) is expanded to carry one bit data. This method can provide an embedding capacity up to 0.5 bits per pixel (BPP) in a single pass. Tian employed a location map to record all expandable locations, and afterwards, the technique of location map is widely adopted by most RDH algorithms. The location map that indicates the modified pairs is compressed and included in the payload.

In [3] Thodi et al. introduced Expansion Embedding (EE) technique which is a generalization of Difference Expansion proposed by Tian et al. which expands the difference between two adjacent pixels by shifting to the left its binary representation, thus creating a new virtual least significant bit (LSB) that can be used for data insertion. Since then, EE has been applied in some transformed domain such as the wavelet domain or to prediction-errors. EE is usually associated with LSB substitution applied to samples that cannot be expanded due to the signal dynamic limits or to preserve the image quality.

The authors achieved a significant improvement by incorporating DE with HS. In addition, instead of the difference value, they suggested utilizing the prediction-error for expansion embedding since this can better exploit local correlations within neighboring pixels. The prediction-error expansion (PE) algorithm is essentially a form of difference expansion. Thodi et al. employed a pixel's three-neighbor context to predict the pixel value and used the expansion of prediction-error between the original pixel value and the estimated one to embed message. The PE algorithm achieved a maximal embedding rate of 1 bit per pixel (bpp).

In [4] Sachnev et al. presents a reversible watermarking algorithm for images without using a location map in most cases. The algorithm employs prediction errors to embed data into an image. A sorting technique is used to record the prediction errors based on magnitude of its local variance. Sachnev et al. points out that RDH algorithm with no location maps, or smaller in some cases, are very desirable.

Celik et al. [5] devised a low-distortion, reversible watermark that can embed as high as 0.7 bits/pixel. The algorithm first quantizes each pixel by a quantizer of step size L, compresses the quantization noise and appends a payload to it, then adds an L-ary representation of the result to the quantized image.

Kamstra et al. [6] utilized low-pass image to predict expandable locations so that the location map can be remarkably compressed. The authors improved Tian's methods by sorting least significant bits (LSBs) or pairs of pixels to be watermarked with respect to the obtained values heuristically. The sorting improves the coding efficiency of the lossless compression, so that the overall performance is improved.

3. PRELIMINARIES

3.1 Histogram Shifting

- Embedding Algorithm:

In the histogram, there are two points namely, a zero point and a peak point. A zero point corresponds to the grayscale value which has no pixel in the given image. A peak point corresponds to the grayscale value which has the maximum number of pixels. Note that zero point defined above may not exist for some image histograms. The concept of minimum point is hence more general. By minimum point, it means a grayscale value 'b', that has a minimum number of pixels, assume this value be h(b), is minimum. Accordingly, the peak point 'a' is referred to as maximum point.

For an $M \times N$ image, each pixel grayscale value $x \in [0, 255]$.

a) Generate its histogram H(x).
b) In the histogram H(x), find the maximum point h(a), $a \in [0, 255]$ and the minimum point zero h(b), $b \in [0, 255]$.
c) If the minimum point $h(b) > 0$, recode the coordinate (i, j) of those pixels and the pixel grayscale value 'b' as overhead bookkeeping information (referred to as overhead information for short). Then set h(b)=0.
d) Without loss of generality, assume a<b. Move the whole part of the histogram H(x) with $x \in [a, b]$ to the right by 1 unit. This means that all the pixel grayscale values (satisfying $x \in [a, b]$) are added by 1.
e) Scan the image, once meet the pixel (whose grayscale value is 'a'), check the to-be-embedded bit. If the to-be

embedded bit is '1', the pixel grayscale value is changed to 'a+1'. If the bit is '0', the pixel value remains 'a'.

- Extraction Algorithm:

For the sake of brevity, only the simple case of one pair of minimum point and maximum point is described here because, as shown above, the general cases of multiple pairs of maximum and minimum points can be decomposed as a few one pair cases. That is, the multiple pair case can be treated as the multiple repetition of the data extraction for one pair case. Assume the grayscale value of the maximum point and the minimum points are 'a' and 'b', respectively. Without loss of generality, assume $a<b$. The marked image is of size $M \times N$, each pixel grayscale value $x \in [0, 255]$.

a) Scan the marked image in the same sequential order as that used in the embedding procedure. If a pixel with its grayscale value 'a+1' is encountered, a bit '1' is extracted. If a pixel with its value 'a' is encountered, a bit '0' is extracted.

b) Scan the image again, for any pixel whose grayscale value $x \in [a, b]$, the pixel value 'x' is subtracted by 1.

c) If there is overhead bookkeeping information found in the extracted data, set the pixel grayscale value (whose coordinate (i, j) is saved in the overhead) as 'b'.

In this way, the original image can be recovered without any distortion.

3.2 Integer to Integer Wavelet Transform

Conventional wavelet transform is not very useful in reversible data hiding as it is not fully reversible. For example, if an image is decomposed using conventional wavelet transform, it may not be guaranteed that the wavelet coefficient values after watermark embedding would be in integer. Information may be lost if any floating-point value is truncated and original cover image cannot be recovered in its entirety. Therefore, to deal with this issue, an invertible integer-to-integer wavelet transform based on lifting scheme is utilized in the proposed scheme. In integer-to-integer wavelet transform, integers are mapped to integers and thus, do not cause any information loss.

3.3 Entropy

In information theory, entropy can be defined as the measure of the uncertainty of a random variable. Entropy of a random variable X with probability mass function p(x) can be defined as:

$$H(X) = -\Sigma\, p(x) \log p(x)$$

Here, the log is to the base 2 and entropy is measured in bits. The entropy of an image can be calculated approximately by the histogram of image. Entropy demonstrates the average information and therefore, it can be used to find the average information present in an image.

4. PROPOSED METHOD

A reversible data hiding scheme is proposed which utilizes integer-to-integer wavelet transform and histogram modification for watermark embedding. The proposed embedding and extraction procedure is shown in Fig. 1 and Fig. 2 respectively.

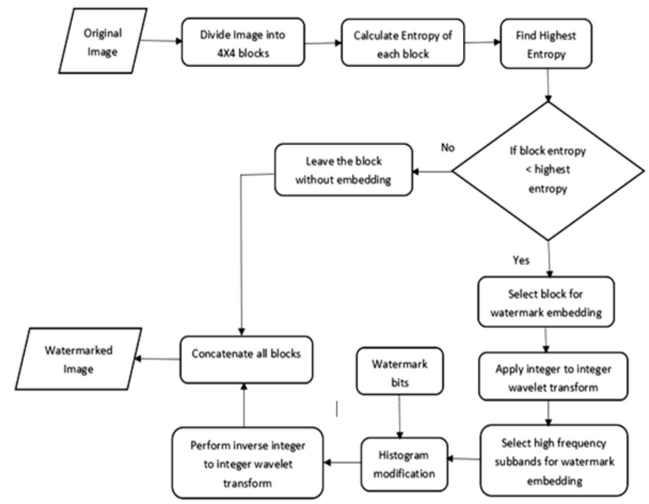

Fig. 1: Embedding Procedure

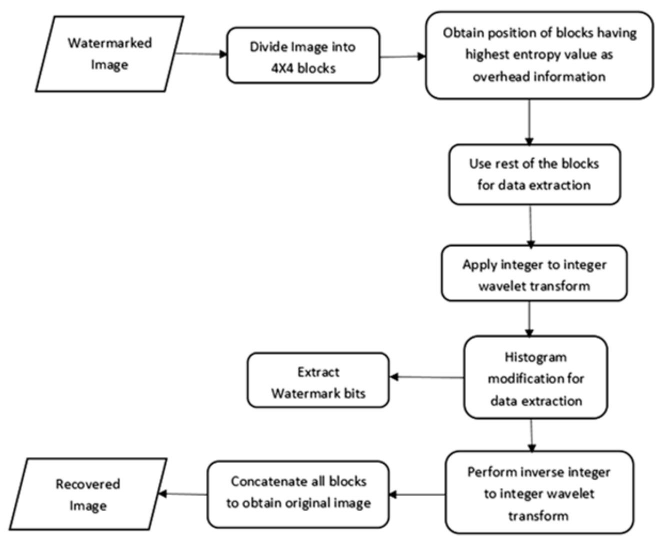

Fig. 2: Extraction Procedure

A. Embedding Algorithm

1. Divide the original medical image I into 4X4 sized non-overlapping blocks.
2. Calculate the entropy of each block. Find out the block having highest entropy.
3. If the entropy of the block is less than the highest entropy, use the block for embedding.

4. Apply integer-to-integer wavelet transform over all the selected blocks and obtain the low frequency sub band LL and high frequency sub bands HL, LH and HH.
5. Apply the basic histogram-bin-shifting based technique on all high frequency sub bands i.e. LH, HL and HH.
6. Perform inverse integer-to-integer wavelet transform on all modified blocks.
7. Concatenate all 4X4 blocks to obtain the watermarked image I_w.

B. Extraction Algorithm

1. Consider the watermarked image I_w. Divide I_w into 4X4 sized non-overlapping blocks.
2. Based on the overhead information, obtain the position of block having highest entropy.
3. Except the block having highest entropy, decompose all other blocks (watermarked blocks) into low and high frequency sub bands using integer-to-integer wavelet transform.
4. Apply extraction and recovery algorithm using basic histogram shifting to extract the watermark bits and original content of each watermarked block.
5. Apply inverse integer-to-integer wavelet transform over all blocks.
6. Concatenate all the blocks to obtain the original cover image I^1.

5. SYSTEM IMPLEMENTATION

5.1 Embedding Procedure

The proposed embedding algorithms have been applied on a medical image of size 256 × 256 as shown in Fig. 3a. This medical image is obtained from public access database.

(https://visualsonline.cancer.gov/closeup2016)

Input: A medical cover image I (Fig. 3a) and text message data to-be-embedded (Fig. 3b).

Output: A watermarked-image I_w with text message data embedded in it (Fig. 3c).

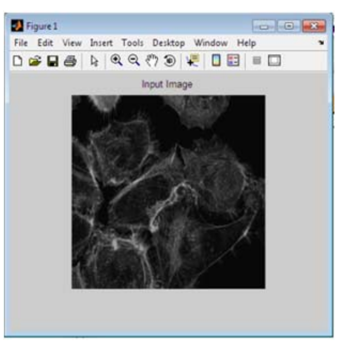

Fig. 3a: Input cover image

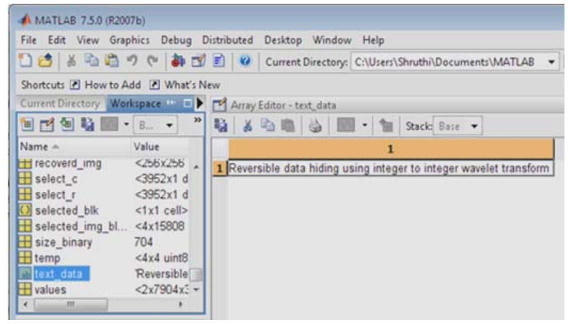

Fig. 3b: Text data to-be-embedded

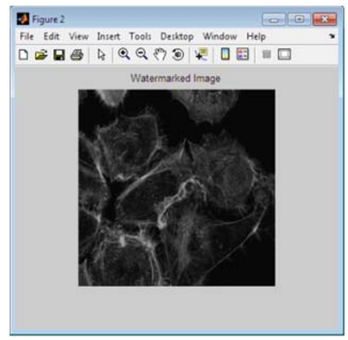

Fig. 3c: Watermarked image

Steps:

1. Read an original medical image I and convert it into grayscale image of size 256 X 256. Divide the resized image into 4 X 4 non-overlapping blocks.
2. Get the text data to be embedded and convert it into corresponding binary values.
3. Calculate the entropy of all 4 X 4 blocks.
4. Find out the highest entropy value.
5. Now select the blocks having entropy values less than highest entropy. Use these blocks for watermark embedding.
6. Leave the blocks having entropy equal to the highest entropy value.
7. Apply integer-to-integer wavelet transform over all the selected blocks and obtain the low frequency sub band LL and high frequency sub bands LH, HL, HH.
8. Now embed the watermarked bits into high frequency sub bands i.e. LH, HL, HH by using basic histogram shifting technique.
9. Perform inverse integer-to-integer wavelet transform on all modified blocks.
10. Concatenate all the modified blocks (obtained from step 9) and the remaining blocks (from step 6) to obtain the watermarked image I_w.

6. CONCLUSION

In this paper, some of reversible data hiding techniques for digital images were investigated in the literature review. As a preliminary to the proposed system, some concepts such as histogram shifting, entropy and integer to integer wavelet transform were discussed in section 3. Finally, a system for reversible data hiding using integer to integer wavelet transform for medical images has been designed and it is implemented up to the embedding part. The implementation of the extraction procedure will be done in the future work.

REFERENCES

[1] Zhicheng Ni, Yun-Qing Shi, Nirwan Ansari, and Wei Su, "Reversible Data Hiding", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 16, No. 3, March 2006.

[2] Jun Tian, "Reversible Data Embedding Using a Difference Expansion", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 13, No. 8, August 2003.

[3] Diljith M. Thodi and Jeffrey J. Rodríguez, "Expansion Embedding Techniques for Reversible Watermarking", IEEE Transactions on Image Processing, Vol. 16, No. 3, March 2007.

[4] Vasiliy Sachnev, Hyoung Joong Kim, Jeho Nam, Sundaram Suresh, and Yun Qing Shi, "Reversible Watermarking Algorithm Using Sorting and Prediction", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 19, No. 7, July 2009.

[5] Mehmet Utku Celik, Gaurav Sharma, Ahmet Murat Tekalp and Eli Saber, "Lossless Generalized-LSB Data Embedding", IEEE Transactions on Image Processing, Vol. 14, No. 2, February 2005.

[6] Lute Kamstra and Henk J. A. M. Heijmans, "Reversible Data Embedding into Images Using Wavelet Techniques and Sorting", IEEE Transactions on Image Processing, Vol. 14, No. 12, December 2005.

[7] Rajkumar Ramaswamy and Vasuki Arumugam, "Lossless Data Hiding Based on Histogram Modification", The International Arab Journal of Information Technology, Vol. 9, No. 5, September 2012.