Australian Journal of Basic and Applied Sciences, 7(14) December 2013, Pages: 195-203

AENSI Journals

Australian Journal of Basic and Applied Sciences

Journal home page: www.ajbasweb.com

A New Mechanism to Control Marked Image Quality of Reversible Data Hiding Based on Histogram Shifting

[1]Hamida Mohamed Almangush, [2]Mohd Khanapi Abd Ghani, [3]Ahmed Bashir Abugharsa

[1]Biomedical Computing and Engineering Technologies (BIOCORE Malaysia) Research Group .Faculty of Information and Communication Technology Universiti Teknikal Malaysia Melaka Hang Tuah Jaya 76100 Durian Tunggal, Melaka Malaysia.
[2]Biomedical Computing and Engineering Technologies (BIOCORE Malaysia) Research Group. Faculty of Information and Communication Technology Universiti Teknikal Malaysia Melaka Hang Tuah Jaya 76100 Durian Tunggal, Melaka Malaysia..
[3]Faculty of Information and Communication Technology Universiti Teknikal Malaysia Melaka Hang Tuah Jaya 76100 Durian Tunggal, Melaka Malaysia.

ARTICLE INFO

Article history:
Received 20 Jun 2013
Received in revised form 20 December 2013
Accepted 23 December 2013
Available online 1 February 2014

Key words:
Computational Shifted Histogram, Closest Pair Of Points, Histogram Modification, Reversible Data Hiding.

ABSTRACT

Background: The common histogram-based reversible data hiding technique is based on the notion of selecting the peak and the minimum points. This technique suffers from some problems because it is considered to be capacity–based. The first problem is that this technique does not consider the size of the payload, even if it is a small size, the peak point is chosen to represent the data, therefore distortion in the marked image is almost equally the same for large or small payloads. The second problem occurs when a cover image histogram includes more than a single peak point and a minimum point using a distance measure to select the closest pair of points which does not ensure a reduced number of modified pixels. **Objective:** To solve these problems, a new mechanism is proposed that considers the size of the payload and the number of the modified pixels to determine the optimal pairs of alternatives peaks (container colors) and minimum points. Therefore, no extra amount of pixels will be modified. **Results:** Compared to the traditional histogram-shifting methods. The experimental results show that the proposed mechanism improves the quality of the marked image of histogram shifting technique under different embedding rates. **Conclusion:** applying reversible data hiding based on a computational shifted histogram technique the marked image quality is improved. Further, the proposed mechanism for selecting the best pair considers the factors that affect the distortion in the image quality of a marked image, namely the number of the modified pixels between the peak and minimum points and the payload size. In comparison to the traditional histogram-shifting method, the proposed method significantly improves the marked image quality and the proposed work can be considered to be both capacity and distortion-based.

To Cite This Article: Hamida Mohamed Almangush, Mohd Khanapi Abd Ghani, Ahmed Bashir Abugharsa., A New Mechanism to Control Marked Image Quality of Reversible Data Hiding Based on Histogram Shifting. ***Aust. J. Basic & Appl. Sci.,*** 7(14): 195-203, 2013

INTRODUCTION

The main goal of steganography is to communicate securely in a completely undetectable manner and to avoid drawing attention to the transmission of hidden data. It is not to keep others from knowing the hidden information, but it is to keep others from thinking that the information even exists. Steganography can be said to protect both messages and communicating parties. If a steganography method causes someone to suspect the carrier medium, then this method is defeated. A steganographic system has two main aspects: steganographic capacity and imperceptibility. These two characteristics are at odds with each other. Furthermore, it is quite difficult to increase the steganographic capacity and simultaneously maintain the imperceptibility of a steganographic system.

Digital image-based steganography, as a method of secret communication, aims to convey an amount of secret data, called the payload, between communicating parties. Additionally, it aims to avoid the suspicion of non-communicating parties towards this kind of communication. In this paper, some characteristics and properties of digital images have been employed to enhance the marked image quality (imperceptibility). The term ‘quality’ refers to the ability to detect differences between the cover image and the marked image. Thus, if no difference can be detected between these two images then this steganography method is imperceptible. In other words the closer the marked image is to the cover image, the higher the security which is measured in terms of Peak Signal to Noise Ratio (PSNR). A high PSNR value indicates high security because it indicates

Corresponding Author: Hamida Mohamed almangush. Faculty of Information and Communication Technology,Universiti Teknikal Malaysia MelakaHang Tuah Jaya 76100Durian Tunggal, Melaka Malaysia
E-mail: hamida.almangush@yahoo.com

minimum difference between the original and the marked values. So no one can suspect the hidden information exists (Hemalatha *et al*. 2013).

The rest of this paper is organized as follows: Related works are described in Section 2. The characteristics of the proposed algorithm and its details are described in Section 3. Experimental results are presented in Section 4, and lastly conclusions are made in Section 5.

Related Works:

In the histogram shifting technique for data embedding by Ni *et al*., all pixel values between the peak and minimum points have to be modified by one unit to leave a gap for data hiding (Ni, Shi *et al*. 2006). However, this method does not utilise a distortion control mechanism because the distortions are almost equally the same for a large or small payload. The authors do not mention the method of selection of the peak and minimum points.

(Yang,*et al*. 2009) presented an algorithm which stands on the waveform of the histogram, and selects the maximum point from the wave crests and minimum point from the wave troughs. They considered the ratio value $\frac{\text{Capacity}}{\log(\text{the amount of pixels moved})}$ when finding the best pair of a wave crest and a trough with many wave crests and troughs in the histogram. The capacity is computed by h(a) - size of the overhead information. The first drawback is that this technique is time consuming because the selection of the wave crests depends on their next and previous neighbour frequencies and not on the size of the payload. The second drawback is that the equation does not find the best pairs with the least amount of pixels moved. For example, assume the two pairs are (c1, t1) and (c2, t2) and that they have the same value of overhead information, h(c1)=6000. The amount of pixels moved between c1 and t1 is 4000, h(c2) =5000, the amount of pixels moved between c2 and t2 is 2000. After applying the equation R1= 6000/log(4000)=1665.7135 R2=5000/log(2000)=1524.6788, according to the algorithm given by the authors, the first pair will be selected. However, in fact the first pair includes a greater amount of pixels moved compared to the second pair.

(Hong, Chen *et al*. 2010) proposed a modified version of the Ni *et al*. (2006) method to solve the major drawback of that technique. Instead of shifting all the pixels between the peak and minimum point before embedding, they combined the shifting and embedding processes together so that just the necessary amount of pixels would be modified for a given payload. Therefore, no extra amount of pixels would be modified compared to the Ni *et al*.'s method. The advantage of this technique is that the quality of the marked image is improved. The drawback of this improved technique is that the peak point is selected to represent the secret data even if the payload is a small size.

Proposed Method:

The main idea in the modified-histogram data hiding method is to find a pair of peak and minimum points in the image pixel intensity histogram and then shift the intensity of those pixels within the peak and minimum frequency range by one level, towards the minimum frequency level. Therefore the distortion in the marked image is almost equally the same for a large or small payload. Using a distance measure to select the closest a minimum frequency colour to a peak is not the ideal way to reduce the distortion in a marked image. Figure 1 presents an example of the number of modified pixels and the distance between the peak point 'P' and the two minimum frequency colours 'P-3' and 'P+2'. Note that pair (P, P-3) is better than pair (P, P+2) because it has less number of pixels between P and P-3, although it is farther to P than P+2.

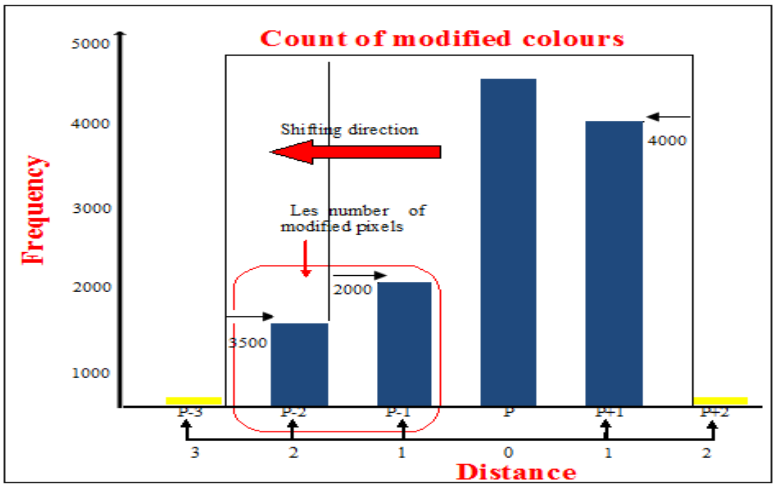


Fig. 1: A comparison between modified pixels and distance measures.

Problem: Best pair of colours:
Given:

- Set G={0,1,2,3,…………..,255}
- Set of Peak Point Colours P, it satisfies:
  -It is a proper subset of G, $P \subset G$;
  -Frequency (p) ≥ payload, $p \in P$
- Set of minimum frequency colours M, it satisfies
  -It is a proper subset of G, $M \subset G$;
  -Frequency (m) = minimum{ Frequency (0),…….., Frequency (255)}, $m \in M$
- Intersection of the sets P and M , $P \cap M = \{\}$
- Distance(D) measure
- For w, x, y, z ∈ G, the Distance(D) measure satisfies:
  - D(x, x)=0
  - 0 < D (x, y) ≤ 255, where x≠ y
  - D(x, y)=D(y, x)
  - D (x, y)=| x-y |
  - D(w, x) < D(w, y) < D(w, z) where w < x < y < z or w > x > y > z
- Number of modified pixels (NMP) measure, it satisfies:
- For w, x, y, z ∈ G, the number of modified pixels (NMP) measure satisfies
  - NMP (x, y) = 0 ,where |x-y | ≤ 1
  - NMP (x, y) > 0 , where |x-y | > 1
  - NMP (x, y) = NMP (y,x)
  - NMP (w, x) ≤ NMP (w, y) ≤ NMP (w, z) where w < x < y < z or w > x > y > z
  - NMP (x,y)= Frequency (x+1)+ Frequency (x+2)……+ Frequency (y-1) where x < y ;
    Frequency (y+1)+ Frequency (y+2)……+ Frequency (x-1) where x > y

Find:

- Pair (p, m) p ∈ P and m € M such that NMP (p, m) is minimum

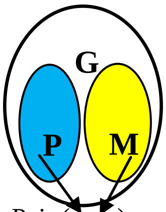


Fig. 2: Pair of colours.

A naive solution to this problem is to compute the number of modified pixels between p and m in each pair of points in Pair set, then pick the pair that has a lesser amount of modified pixels, where Pair set is obtained from the Cartesian product of peak point colours P set and minimum frequency colours M set, Pair=P × M={ (p, m) | p ∈ P and m ∈ M}.

To solve the problems described above, a new mechanism is proposed to control the distortion in a marked image. The new mechanism comprises three proposed techniques to select the best pair in terms of the least number of modified pixels. The first technique is named Minimum and Container Colours Filtering Technique. The second technique is named the Closest Colour Searching. The third technique is named Best Pair Filtering Technique. The details of these techniques are described in this section. In this study the technique of Hong *et al*. is applied to embed the payload.

Minimum and Container Colours Filtering Technique:

This procedure selects the containers colours and the minimum frequency colours .The containers colours or the alternative of the peak points which their frequencies are equal to or bigger than the payload size. The payload includes the secret data and the original values of the least significant bits of the pixels that are modified to represent the bits of the overhead information 'OI'. The overhead information involves the primary overhead information 'POI' and the map location 'ML'.

Payload size= length of the secret data + Overhead information (1)

Overhead information = POI+ ML (2)

The primary overhead information POI includes the values of the container colour, the minimum frequency colour and the minimum frequency and the last position (x,y) of the embedded bit. The values of the container colour and the minimum colour $\in \{0,\ldots 2^8\}$, each colour is represented by 8 bits .The size of minimum frequency value is presented by the reserved number of bits RB, the number of bits that required to represent the last position is NB bits

$$POI = 2 \times 8 + RB + NB \quad (3)$$

Where

NB denotes the number of bits that is required to represent the coordinates x and y within the cover image.

$$NB = 2 + \lfloor Log(W-1)/ Log(2) \rfloor + \lfloor Log(H-1)/ Log(2) \rfloor \quad (4)$$

W and H are the width and the height of the cover image respectively.

The map location ML includes the position of the pixels that their values equal to the minimum frequency colour

$$ML = NB \times MF \quad (5)$$

MF refers to the minimum frequency in the cover image histogram. The maximum value of MF is 2^{RB}-1

Figure 3 shows the result after applying the filtering technique for a payload size of 2500 bits. In part of the cover image histogram from colour C_1 to colour C_n, the container colours are highlighted by a blue colour and above the red line which indicates the number of bits that can represent the payload . Algorithm 1 presents the steps of the Minimum and Container Colours Filtering Technique.

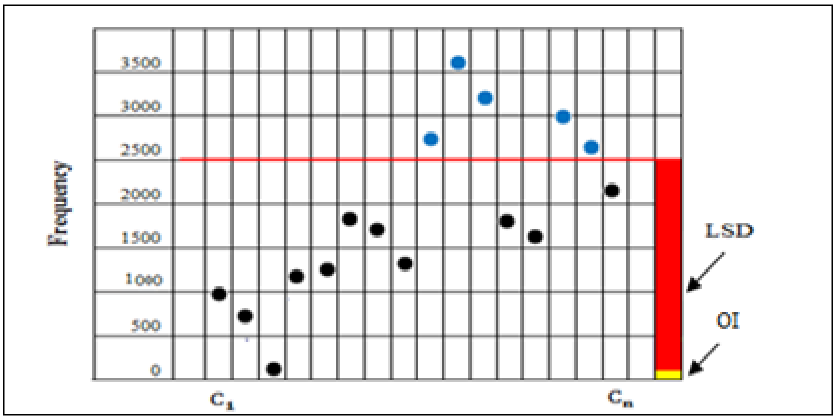


Fig. 3: Container colours.

Algorithm 1: Minimum and Container Colours Filtering:
Input : cover image, RB , Length of the Secret Data (LSD) in bit stream form.
Output: Container Colours array, Minimum Colours array, cover image histogram h(256)
Steps:
Step 1: Create the histogram of the cover image h(256)
Step 2: Scan the histogram to find the minimum frequency.
Step 3: W =Cover-Image.Width
Step 4: H =Cover-Image.Height
Step 5: $NB = 2 + floor(Log(W-1) / Log(2)) + floor(Log(H-1) / Log(2))$
Step 6: $POI = 2 \times 8 + RB + NB$
Step 7: ML= 2 × NB × Minimum Frequency
Step 7: PZ = LSD + POI + ML
Step 8: Scan the histogram again in a sequential order from 0 to 255. When a colour is found with a frequency equal to the minimum frequency then add the colour to the Minimum Colours array or when a colour is found with a frequency greater than or equal to PZ then add the colour to the Container Colours array.
Step 9: End

Closest Colour searching:

The searching technique provides a mechanism to control the marked image quality. It is designed to find pairs of container colours and their closest minimum frequency colours. Algorithm 2 includes the steps of determining the pairs of the closest neighbour colours.

In this section an example is given to present the closest pairs assuming the minimum frequency colour values are 1, 2, 3, 4, 5, 6, 7, 8, 150, 184, 227, 228, 239, 241, 247, 251and 253 and the container colour values are 12, 13, 14, 163, 235 and 243. In Figure 6(a) the minimum frequency colours are highlighted by a yellow colour and the container colour values are highlighted by a blue colour. Figure 4(b) presents a histogram showing the distribution of these colours. Figure 5 presents the closest neighbour container colour of each minimum frequency colour. Note that the minimum frequency colours with values of 1, 2, 3, 4, 5, 6, 7, 227, 251, and 253 are not listed because they are not the closest neighbour minimum frequency colours of any container colour. Note also that the container colours with value of 13 is not listed because it is not the closest neighbour container colours of any minimum frequency colour .The set of pairs (p) is {(12, 8) (14, 150) (163, 150) (163, 184) (235, 228) (235, 239) (243, 241) (243, 247)}.The number of the pairs is 8 instead of the 102 that could be obtained by a Cartesian product (17× 6 =102).

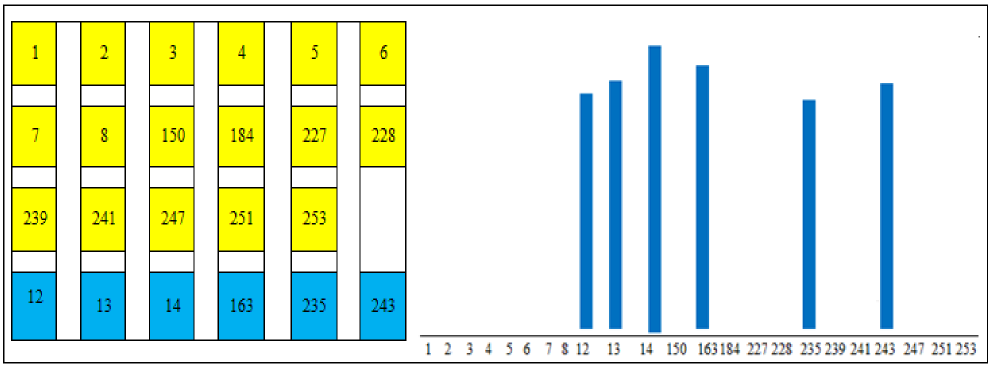


Fig. 4: The distribution of the minimum frequency and the container colours.

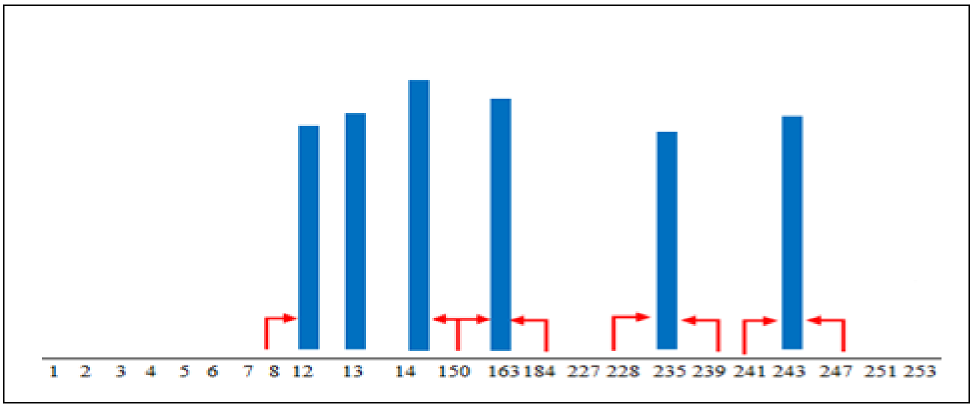


Fig. 5: Pairs of the minimum frequency colours and the container colours.

Algorithm 2: Closest Colour Searching:

Input : sorted Container Colours array C(n), sorted Minimum Colours array M(k)
Output : Pairs array
Steps
Step 1: i=1
Step 2: j=1
Step 3: A= c(i)
Step 4: B= m(j)
Step 5: If c(i) < m(j) Then go to step12
Step 6: While m(j) < c(i)
Step 7: B = m(j)
Step 8: j = j + 1

Step 9: If j = k+1 Then go to step 19
Step 10: End While
Step 11: Add pair (A ,B) to the Pairs array
Step 12: While c(i) < m(j)
Step 13: A= c(i)
Step 14: i = i + 1
Step 15: If i = n+1 Then go to step 19
Step 16: End While
Step 17: Add pair (A , B) to the Pairs array
Step 18: Go to step 6
Step 19: Add pair (A, B) to the Pairs array
Step 20: End

Best Pair Filtering Technique:

This technique computes the number of modified pixels that are between the container and the minimum frequency colours in each pair, and then returns the pair that has the minimum number of modified pixels. Using the number of modified pixels as a measure instead of the distance measure reduces the distortion within a marked image.

Algorithm 3: Best Pair Filtering:

Input :Cover image histogram h(256) array , Pairs(z) array
Output : The best pair of colours
Steps:
Step 1: For i=1 to z
Step 2: If Pair(i).C > Pair(i). M then
Step 3: Pair(i).NMP = h(Pair(i).C -1) +…………..+ h(Pair(i).M +1)
Step 4 : Else
Step 5 : Pair(i).NMP = h(Pair(i).C +1) +………….. + h(Pair(i).M -1)
Step 6: End If
Step 7 : Next
Step 8: Best Pair=Minimum(Pair(1).NMP, Pair(2).NMP, ……….., Pair(z).NMP)
Step 9 :End

Experimental Results:

To evaluate the proposed scheme, six medical images with 256 gray scales were used. The hiding capacity of the modified histogram based on the reversible data hiding method is primarily determined by the peak height of the histogram of the corresponding cover image (Shih, 2012), thus the maximum payload-number of bits that can be hidden into the image is equal to the maximum frequency of the original cover image histogram. In this study the embedding rate is defined as follows:

$$E_R = \frac{\text{Number of Embedded bits}}{\text{Maximum payload}} \times 100 \qquad (6)$$

In our experiments, the quality of the marked image is measured by the peak signal-to-noise ratio (PSNR) which is the most popular criterion to measure the distortion between the cover image and the marked image. The PSNR is often expressed as a figure on a logarithmic scale in decibels (dB). A higher PSNR indicates a better reconstructed or marked image (Krishna, Rahim *et al*., 2010).
The PSNR is represented by the following equation:

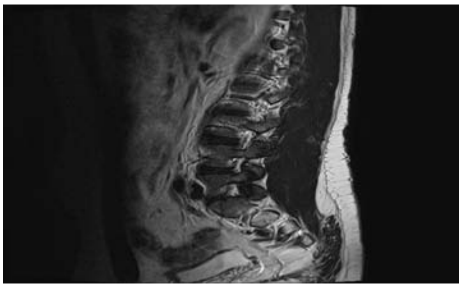
$$\text{PSNR} = 10\text{Log}_{10}\left(\frac{255^2}{\text{MSE}}\right) \qquad (7)$$

where MSE denotes the Mean Square Error which is given by:

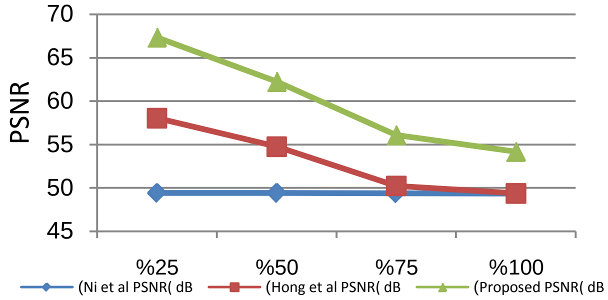
$$\text{MSE} = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}\left(S_{Xy} - C_{Xy}\right)^2 \qquad (8)$$

where x and y are the image coordinates, M and N are the dimensions of the image, Sxy is the generated marked image and Cxy is the cover image (Jain and Ahirwal 2010).

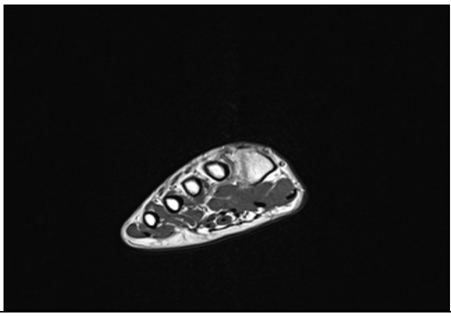
Table 1 lists the number of the minimum and container colours within five medical images. In the table, "m" refers to the minimum frequency colours, "c" refers to the number of the container colours, (m ×c) refers to the number of pairs that are obtained by the Cartesian Products and "P" refers to the number of pairs that are obtained by applying the proposed closest pair searching technique under different embedding rates. Table 2 and Figure 6 (f-j) show a comparison of the proposed scheme with the methods of Niet *et al*. and Hong *et al*. in terms of the marked image quality under different embedding rates.



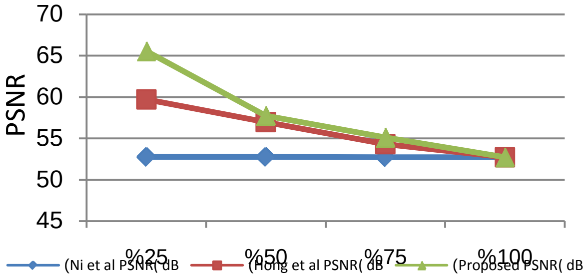
(a)



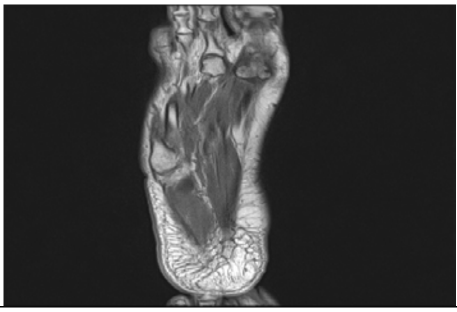
(f)



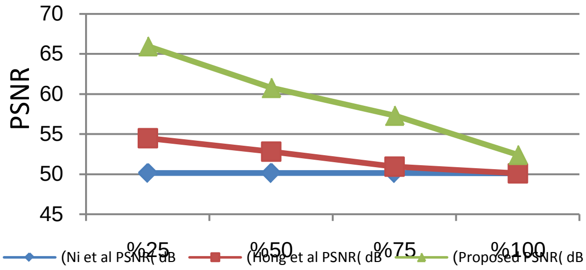
(b)



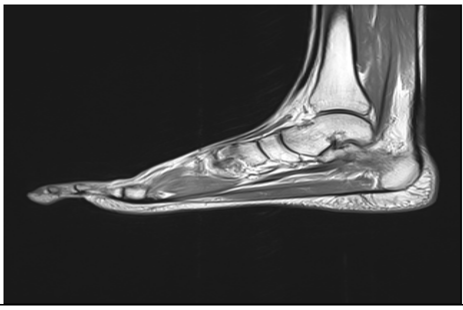
(g)



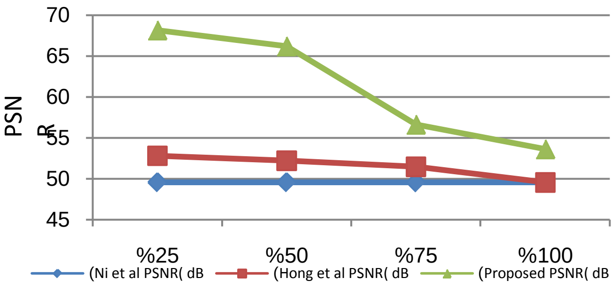
(c)



(h)

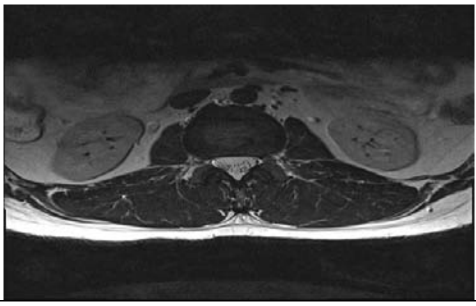


(d)

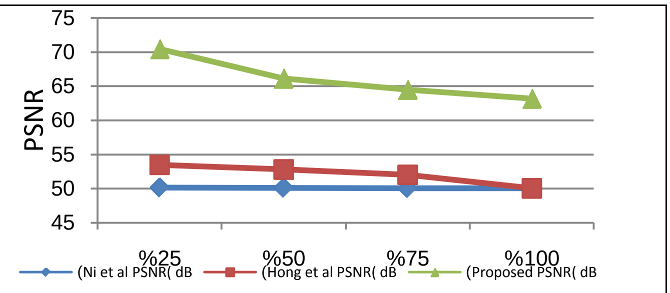


(i)

Australian Journal of Basic and Applied Sciences, 7(14) December 2013, Pages: 195-203



(e)



(j)

Fig. 6: (a-e) medical images, (f-j) Comparison of the PSNR- Embedding Rate between Ni *et al*., Hong *et al*. and the proposed methods.

Table 1: The numbers of the minimums and containers colors.

Image number	M	Embedding Rate											
		25%			50%			75%			100%		
		C	m ×c	p	C	m ×c	P	C	m ×c	p	C	m ×c	P
1	39	34	1326	2	25	975	2	11	429	2	1	39	2
2	2	154	308	2	67	134	2	12	24	2	1	2	2
3	39	136	5304	2	89	3204	2	25	975	2	1	39	2
4	25	110	2750	2	42	1050	2	12	300	2	1	25	2
5	108	18	1944	24	1	108	2	1	108	2	1	108	2

Table 2: Comparison of the proposed scheme with the methods of Niet *et al*. and Hong *et al.*

Image number	E_R	SPNR (dB)		
		Ni *et al*	Hong *et al*	Proposed
1	25%	49.43645	58.05008	67.36883
	50%	49.41646	54.76347	62.24742
	75%	49.39609	50.24379	56.09599
	100%	49.37684	49.38265	54.18689
2	25%	52.7636	59.7052	65.53961
	50%	52.7473	56.9826	57.73571
	75%	52.7302	54.2969	55.09399
	100%	52.7136	52.7243	52.7243
3	25%	50.13949	54.48047	65.97996
	50%	50.13017	52.80748	60.77403
	75%	50.12065	50.95091	57.3002
	100%	50.11242	50.11812	52.41825
4	25%	49.58357	52.81964	68.13615
	50%	49.57352	52.21995	64.90429
	75%	49.56317	51.46663	56.60864
	100%	49.55363	49.55381	53.62696
5	25%	50.14822	53.47608	70.48079
	50%	50.10906	52.80254	66.12922
	75%	50.07026	52.04534	64.50687
	100%	50.03096	50.03962	63.18605

Conclusions:

In this paper, the researchers have shown that by applying reversible data hiding based on a computational shifted histogram technique the marked image quality is improved. Further, the proposed mechanism for selecting the best pair considers the factors that affect the distortion in the image quality of a marked image, namely the number of the modified pixels between the peak and minimum points and the payload size. In comparison to the traditional histogram-shifting method, the proposed method significantly improves the marked image quality and the proposed work can be considered to be both capacity and distortion-based.

ACKNOWLEDGMENT

This paper is part of the PhD work in the Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM).

Australian Journal of Basic and Applied Sciences, 7(14) December 2013, Pages: 195-203

REFERENCES

Hemalatha, S., U.D. Acharya, *et al*., 2013. "A Secure and High Capacity Image Steganography Technique." Signal.

Hong, W., T.S. Chen, *et al*., 2010. "A modified histogram shifting based reversible data hiding scheme for high quality images." Inform. Technol. J., 9(1): 179-183.

Jain, Y.K. and R. Ahirwal, 2010. "A Novel Image Steganography Method with Adaptive Number of Least Significant Bits Modification Based on Private Stego Keys." International Journal of Computer Science and Security, 4(1): 40-49.

Krishna, S., B.A. Rahim, *et al*., 2010. Lossless embedding using pixel differences and histogram shifting technique. Recent Advances in Space Technology Services and Climate Change (RSTSCC), 2010, IEEE.

Ni, Z., Y.Q. Shi, *et al*., 2006. "Reversible data hiding." Circuits and Systems for Video Technology, IEEE Transactions on 16(3): 354-362.

Shih, F.Y., 2012. Multimedia Security: Watermarking, Steganography, and Forensics, CRC Press.

Yang, H.W., I.E. Liao, *et al*., 2009. Reversible data hiding based on grayscale value histogram with optimal multiple pairs of crests and troughs. Information Networking, 2009. ICOIN 2009. International Conference on, IEEE.