*Journal of Information Hiding and Multimedia Signal Processing* ©2016 ISSN 2073-4212
**Ubiquitous International** Volume **7**, Number **1**, January **2016**

# Hyper Edge Detection with Clustering for Data Hiding

Raniyah Wazirali and Zenon Chaczko

Centre for Real-time Information Networks (CRIN)
Faculty of Engineering and Information Technology (FEIT)
University of Technology, Sydney(UTS)
Australia, NSW, Sydney
Raniyah.a.wazirali@student.uts.edu.au; Zenon.chaczko@uts.edu.au

Received April, 2015; revised August, 2015

Abstract. *Hiding of information is a great significance in information security. For information security and intellectual property, effective methods are required. Steganography is the art of hiding information in such a way that it cannot produce distrust about its presence. In this way only the sender and recipient are aware of hidden information. There are two critical characteristics of steganography which are as stego image imperceptibility and payload capacity. Recently, many edge detection based steganogaphy approaches have been developed to provide high imperceptibility of the stego image. However, the capacity of these approaches can be limited. This paper presents an effective edge detection method based on gradient, 9x9 mask of LOG and zero crossing in combination with clustering to classify the image to edge and non-edge regions. Different amount of secret message will be hidden in each cluster after the detection of the edge regions. It declares the stability between stego image imperceptibility and high payload capacity .*
**Keywords:** Hyper Edge detection; Data hiding; Clustering.

1. **Introduction.** Huge quantities of material over public networks are spread with a wide area of internet facilities which give users an ease of accessibility of information. Some of this materials contain secret information and involve a certain level of security. Cryptography is one of the common method which is being used for many years and its represent useful techniques for securing secret message by transfer its content to an unreachable and unreadable format or text called cipher text thus only the authorized person will have the right to extract the secret message by entering a key. On the other hand, cryptography also has some weaknesses as a result of the presence the cipher text. Therefore, steganography has been one of the interest topics in the recent year as it involves hiding the existence the secret information rather than hiding its content.

Steganographic system has two important criteria which is payload capacity and stego image quality. However, steganographic payload capacity and image imperceptibility are the case of trad off situation. For example, hiding more subtle elements in secure pictures (higher capacity) exhibits more relics into the carrier file and then it increases the distortion of the hidden information and thus decrease the quality of the stego image. Moreover, it is impractical to simultaneously expand the security and capability of a steganographic system. This is a direct result of the tradeoff between the amount of relics introduced to the carrier (quality), safety against stego file modification and amount of embedded information.

In order to obtain a high payload capacity with minimum changes, the secret message should be embedded carefully in some specific region based on the characteristics of the cover image (carrier). Although edge detection approaches are relatively simple techniques, its provide robust hiding techniques against attacks by preserving the abrupt modification in image intensities. Wu and Tasi (2003) proposed pixel-value differencing (PVD) scheme that applies the difference value between two pair pixels as the size of the embedding capacity [7]. The main concept of PVD based on exploiting the differences between pixels values as a priority of insert secret message. The value of the differences indicates the secret message and allow to keep the degree of the differences unmodified after embedding the information. In PVD the difference values are categorized and allocated to select grades, then higher grades specify a larger dissimilarity value and the lower grades show minimum differences. Fundamentally, a higher differences between two consecutive pixels signify of two pixels are more applicable for modification for embedding some data. If two consecutive pixels has a slight difference value, then only small amount of data can be embedded in this pair because it allocated in a smooth region and therefore any changes can be degrade the visual quality. In 2010, Singh et al. proposed a novel technique for hide secret message in digital color image [4]. The method based on modifying a few pixel color values that used some selected pixel to represent characters instead of color value which then leads to high visual quality of stego object.

Many adaptive steganographic techniques take into consideration human vision sensitivity as well as local texture characteristics component [2, 7]. In 2010, Chen used hybrid edge detector that combines fuzzy edge detection with canny edge detector [7]. Many other studies used canny edge detector to capture edge regions and then insert the data using $k$ LSBs [3, 7]. However, edge detection scheme based on 3x3 window and then secret message embedded in edge region using first component modification method provides high imperceptibility [1]. An algorithm that uses three parameters which is the size of the of Gaussian mask, low threshold value and finally high threshold value can lead to various output of the same carrier and secret message [7]. Wazirali et. al used score matrix with Genetic Algorithm to finds best mapping between secret message and the cover image [6].

This paper introduce a new data hiding approach in edge area based on based on gradient approach, $9x9$ mask of LOG and zero crossing to detect the area of embedding data. The organization of this paper as follow; section 2 provides literature review. Section 3, describes the proposed method and section 4 evaluates the experimental results.

## 2. Related Works.

### 2.1. Classical LSB Steganography.

Least Significant bits substitution is one of the most common hiding approach due to its simplicity. The secret message is considered as bit stream and will be concealed into cover image by changing the LSBs of the cover image with the boot stream of the secret file. For example, by assuming that cover image is 8 bits gray-scale value, the original values of first three pixels are $\{Pixel_1$, $Pixel_2$and $Pixel_3\}$ which have the following values respectively [1 0 0 1 0 1 1 0 ],[1 0 1 1 1 0 0 1] and [1 1 1 1 0 1 1 1]. In order to hide the secret message of {1,0,1}, the work involves replacing the last bit of each pixel. Therefore, the values and the corresponding pixels in the stego image will be [1 0 0 1 0 1 1 1 ],[1 0 1 1 1 0 0 0] and [1 1 1 1 0 1 1 1]. With only 1 bit substitution, the alteration between the cover and stego image is hardly noticeable by human visual system. However, when the size of secret message is high, the

(a) k=1 (b) k=2 (c) k=3

(d) k=4 (e) k=5 (f) k=6

FIGURE 1. Stego image of hiding Airplane image in lenna cover image with $k$ LSBs from 1 to 6

embedding capacity will increase. Therefore, more bits will be used to conceal the large secret message and consequently, more degradation may introduce to the stego image and hence effects its imperceptibility. Figure 1 demonstrates the stego image quality with embedding $k$ LSBs from 1 to 6 in $512x512$ lena image. Obviously, the deterioration with $k = 6$ is noticeable to human visual system and therefore can not be acceptable as a good approach for hiding large secret message.

2.2. **Edge Detector.** Edge detection is a tool utilized in image processing for feature recognition and abstraction, whose purpose is to classify points in an image where intensity of image deviations suddenly. The aim of edge detection is meaningfully decreasing the quantity of data in an image and reserves the important possessions for additional image processing. In a gray scale image the edge is an indigenous feature that, with in a neighborhood splits areas areas in each of which the gray level is more or less abrupt change on the two sides of the edge or texture representative the end of one area in the image and the commencement of another.

The main idea behind edge detection embedding is to identify areas of tone change or object boundaries of the original image in order to find carrier pixels best suitable for data insertion. Most of existing edge based embedding use Laplacian, Canny or Fuzzy edge detection operators to separate usable edge pixels from uniform tone areas which ensure minimum changes to the cover image and therefore hardly detected any alteration by human visual system. However, most of the existing approaches are used in gray-scale image while there is limited edge detection in color images which reflect lower embedding capacity in gray-scale images than the color images. Therefore, effective techniques to extract the edge region in color images are needed to be used in the embedding process. As this work pretends to develop a new LOG of edge detector, the following subsection will describe the basic idea of LOG.

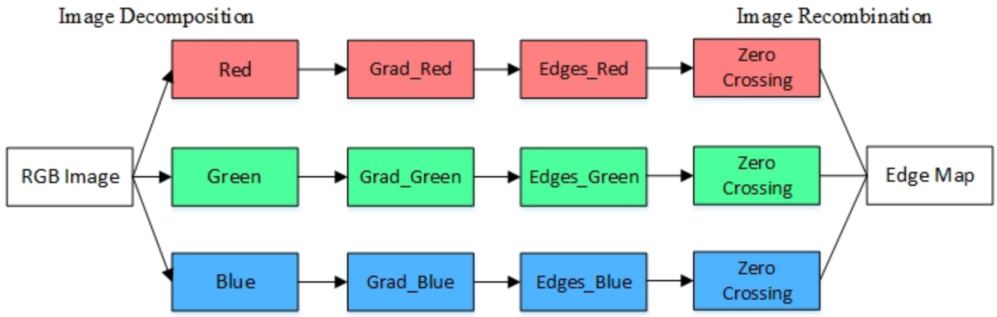


FIGURE 2. Flowchart of the gradient LOG scheme

2.2.1. *Laplacian of Gaussian method.* Laplacian of Gaussian (LOG) mainly based on smoothing the host image before detecting the edges. The noise removes through complication between LOG operator and Gaussian shaped kernel as given in equation 1 followed by Laplacian Operator which is shown in equation 2.

$$G(x, y) = e^{\frac{-x^2+y^2}{2\sigma^2}} \tag{1}$$

Where $\sigma$ represents the standard deviation which if coil with an image, will make it unclear. The amount of blurring is compute by the value of $\sigma$.

$$\nabla^2 G(x, y) = \frac{\partial^2 G(x, y)}{\partial x^2} + \frac{\partial^2 G(x, y)}{\partial y^2} = \left[\frac{x^2 + y^2 - 2\sigma^2}{\sigma^4}\right] e^{\frac{-x^2+y^2}{2\sigma^2}} \tag{2}$$

The laplacian $L(x, y)$ with intensity of $I(x, y)$ of a specific image can be calculated as follows

...

$$L(x, y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \tag{3}$$

The standard $5x5$ LOG mask is giving in equation 4 as follows

$$\begin{matrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 16 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{matrix} \tag{4}$$

3. **Proposed Method.** In order to produce balance between high embedding capacity while keeping high quality of the stego image, a new approach that ensure greater accuracy of edge detection in color images are needed. The proposed approach based on gradient, $9x9$ mask of LOG and zero crossing.

3.1. **The Edge Detection Phase.** The work involves image decomposition of R, G and B colors at the begging and finishes with image recombination at the end of edge detector phase as shown in figure The edges are the regions with powerful intensity contraindication and moving in intensity from pixel to the neighbor pixels can involve major variation in image imperceptibility. By applying first and second derivations such intensity contrast can be detected. The first derivation in image processing is considered as gradient. The gradient function $f(x, y)$ is defined in the following equations

$$\nabla f = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \tag{5}$$

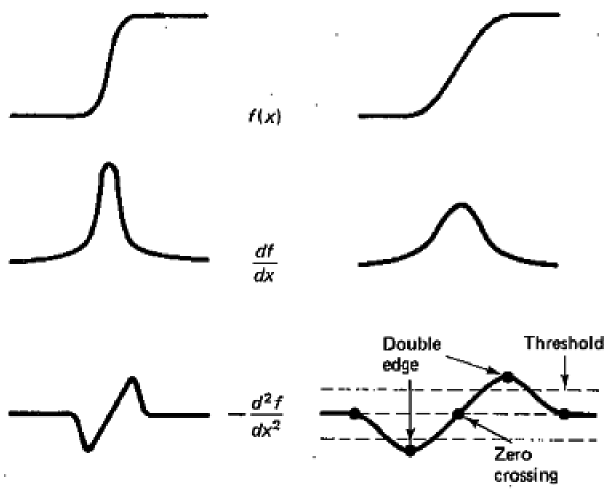


FIGURE 3. First and second derivatives for edge detection

$$\nabla f = mag(\nabla f = [g_x^2 + g_y^2]^{\frac{1}{2}} = \left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\right]^{\frac{1}{2}} \tag{6}$$

Sometimes, the quality can be approximated by omitting the square root operation as shown in equation 7or by using the absolute value as shown in equation 8.

$$\nabla f \approx g_x^2 + g_y^2 \tag{7}$$

$$\nabla f \approx |g_x^2| + |g_y^2| \tag{8}$$

This study uses new improved Laplacian of Gaussian edge detector that based on $9x9$ mask which will produce better accuracy for edge detection using LOG. The values of the new mask are given in the following equation 9.

$$\begin{array}{ccccccccc} 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 4 & 4 & 4 & 3 & 1 & 0 \\ 0 & 3 & 4 & 2 & 0 & 2 & 4 & 3 & 0 \\ 1 & 4 & 2 & -10 & -15 & -10 & 2 & 4 & 1 \\ 1 & 4 & 0 & -10 & -50 & -10 & 0 & 4 & 1 \\ 1 & 4 & 2 & -10 & -15 & -10 & 2 & 4 & 1 \\ 0 & 3 & 4 & 2 & 0 & 2 & 4 & 3 & 0 \\ 0 & 1 & 3 & 4 & 4 & 4 & 3 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{array} \tag{9}$$

As a result of applying second derivations, the resulted gradient more sensitive to noise and the threshold magnitude may generates double edges as can be seen in figure 3. For these reasons, LOG with $9x9$ mask will combine with zero-crossing to detect edge location. Zero crossing detector is defined as given in equation 10 where $c$ normalizes the addition (sum) of elements of a $9x9$ mask.

$$h(m,n) = c\left[1 - \frac{x^2+y^2}{\sigma^2}\right] exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{10}$$

3.2. **Region Classification.** In order to classify the region to be either edge region or non-edge region, K-mean clustering will be used. One of the main important factors in k-mean clustering is to define the number of cluster N. In one hand, when N=2, the k-mean will distinguish whether edges or non-edges region. However, the non-edges pixels govem the total of pixel population. Therefore, losing some edges region as some points in spectral space be inclined to fit in to non-edge clusters instead of edges region. Consequently, N should be greater than two. In this situation, the cluster with bidder population is considered as non edges cluster and all the remaining N-1 cluster is combine in one cluster and considered as edge cluster. Experiments show that N=[4, 8] is the best rang of cluster as edge detection where outcome does not change any more.

3.3. **Data Hiding Phase.** Based on the cluster region, different percentages of LSBs will be applied to embed the secret data. Therefore, the sharp edge region will hide higher percentage, while the smooth region will hide less percentage.

As green color has more visual perception of intensely of the Human Visual System while the blue color has the lowest perception contribution to the color image. Therefore, the last byte of each pixel is the most suitable byte to hide more data.

In smooth regions 2-1-3 embedding technique will be used. 2:1:3 of R:G:B LSBs embedding technique will be used to hide the data in RGB color image respectively. That means, 2 bits in the least significant bit of the red component, 1 bit in the last LSBs of green component and 3 bits in 3 LSBs of blue component. Therefore, 6 bits will be hidden in each pixel of the smooth region.

In sharp edge regions 4-2-6 LSBs embedding technique will be used to hide data which means this region will hide more bits than the smooth region. 4 bits in the least significant bit of the red component, 2 bits in the last 2 component of green and 6 bits in 6 LSBs of blue component will be utilized for data embedding. Therefore, 12 bits will be hidden out of 24 bits of each pixel of the complex region.

Therefore, more and more data will be hidden in edge region which has minimum effect to Human Visual System and ensure minimum degradation of the stego image.

Table 1. The performance of $PSNR$ for various embedded capacity in Kilobyte (kb) for the proposed method and other embedding techniques

| Embedding Rate | Proposed | GA Score-Matrix [6] | LSB | PVD |
|---|---|---|---|---|
| 5% | 58.8165 | 58.0452 | 56.9199 | 57.5445 |
| 10% | 58.7355 | 58.4485 | 55.7563 | 57.2557 |
| 15% | 57.7804 | 57.4974 | 55.1875 | 56.8258 |
| 20% | 57.5530 | 57.0525 | 54.8538 | 56.3584 |
| 25% | 57.0431 | 56.3597 | 54.2597 | 55.9059 |
| 30% | 56.8443 | 55.9545 | 53.5787 | 54.9526 |
| 35% | 56.1788 | 55.5987 | 53.1589 | 54.2598 |
| 40% | 55.5968 | 54.2658 | 52.2575 | 53.3598 |
| 45% | 54.3496 | 53.3597 | 51.4554 | 53.0674 |
| 50% | 53.4864 | 52.985 | 50.5898 | 52.6584 |

4. **Experimental Results.** The results achieved throughout the research work indicate that the combination of graphical user and Matlab increases the ability to attain the probable results. Obviously, the suggested method is better than the classical LSB method

TABLE 2. The performance of $wPSNR$ for various embedded capacity in Kilobyte (kb) for the proposed method and other embedding techniques

| Embedding Rate | Proposed | GA Score Matrix [6] | LSB | PVD |
|---|---|---|---|---|
| 5% | 72.0897 | 71.9878 | 69.5984 | 70.9874 |
| 10% | 71.8599 | 71.5984 | 68.5959 | 70.3698 |
| 15% | 71.5287 | 71.1587 | 68.0259 | 69.5875 |
| 20% | 71.2547 | 70.9842 | 67.6468 | 69.0258 |
| 25% | 70.8574 | 70.2598 | 67.2599 | 68.5874 |
| 30% | 69.5747 | 69.7845 | 66.8469 | 67.4859 |
| 35% | 69.0588 | 69.2548 | 66.4878 | 67.1589 |
| 40% | 68.9584 | 68.4489 | 65.4847 | 66.3684 |
| 45% | 67.5587 | 67.1879 | 64.5489 | 65.9548 |
| 50% | 66.5557 | 66.0258 | 63.4481 | 65.5789 |

(a) k=10 (b) k=20 (c) k=30 (d) k=40

(e) k=50 (f) k=60 (g) k=70 (h) k=80

FIGURE 4. The cover image and the resulted stego image of the proposed method with different embedding rate

in enhancing the imperceptibility of stego images though retain high payload capacity. In order to get a comprehensive comparison, the evaluation of the proposed method will compare with two classical work that include LSB and PVD and one of region based work [6]. The suggested method turns out to be more effective than the commonly used LSB method and PVD and others for optimizing the capacity of stego images with minimum effecting the quality of the stego image.

4.1. **Evaluate the Imperceptibility.** Peak Signal to Noise Ratio (PSNR) measures the quality of the resulted stego image based on the measurement of the Mean Square Error (MSE) with respect of the cover image. 11 and 12 show the related equations to calculate the PSNR and MSE.

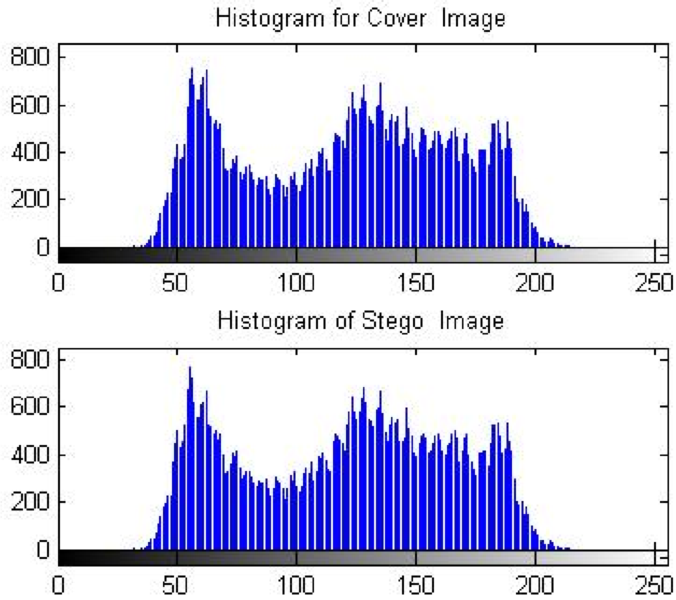


FIGURE 5. Histogram Comparison with 50% embedding rate for Lenna

$$PSNR = 10 \times \log_{10} \frac{Max^2}{MSE} \tag{11}$$

$$MSE = \frac{1}{m \times n} \sum_{i-1}^{m} \sum_{j-1}^{n} (A_{ij} - B_{ij})^2 \tag{12}$$

PSNR and MSE are only represent the total difference between the cover and stego image such as pixels error. Therefore, this paper will also evaluate the result based on weighted Peak Signal to Noise Ratio ($wPSNR$) which takes into account HVS properties by using the principle of less sensitivity of human eye toward changes in texture areas than in smooth areas. $wPSNR$ is an alternative quality metric developed by Voloshynovskiy et al. based on computation of Noise Visibility Function (NVF) as a correction factor that classified the image property based on textured and edge region as embedding data in this region can hide more data effectively with less distortion [5]. For smooth region, $NVF$ is almost1 while for edge regions is close to 0 which indicates that for flat region. $wPSNR$ provides more accurate results than $PSNR$ for methods that based on edge detection approaches.

$$wPSNR = 10 \log_{10} \frac{MAX^2}{||(C_{ij} - S_{ij})||^2_{NVF}} \tag{13}$$

$$wPSNR = 10 \log_{10} \frac{MAX^2}{||(C_{ij} - S_{ij}).NVF||^2}$$

$$NVF(i,j) = \frac{1}{1 + \theta \sigma_x^2(i,j)} \tag{14}$$

Where $\sigma_x^2(i,j)$ represents the local alteration of the image in a window sprinted on the pixel with coordinates $(i,j)$ and $\theta$ is a modification parameter consistent to the specific image which can be calculated as given in 1516.

$$\overline{x}(i,j) = \frac{1}{(2L+1)^2} \sum_{k=-L}^{L} \sum_{l=-L}^{L} x(i+k, j+L) \tag{15}$$

$$\sigma_x^2(i,j) = \frac{1}{(2L+1)^2} \sum_{k=-L}^{L} \sum_{l=-L}^{L} (x(i+k, j+L))^2 \tag{16}$$

The image based on modification parameter is given as $\theta = \frac{D}{\sigma^2_{x(max)}}$ Where $\sigma_x^2$ is the maximum local variance for a particular image and $D \in [50, 100]$ is a determined value.

Figure 1 represents evaluation of the suggested method with other methods in term of the embedded capacity with$PSNR$ Value. Table 1 shows the PSNR value of the suggested system with its associated capacity in$512x512$ images for Lena image with comparison of other steganography method. Figure 4 shows the cover image and resulted stego image for embedding various size of secret message. The improvement of the proposed method over other techniques is noticeable. It is obvious from table 1 that the method can embed high payload capacity while retaining acceptable visual quality.

4.2. **Evaluate the Deductibility.** Histogram Error (HE) measures the absolute error value between the cover image and the stego image. Histogram Error can be calculated as given in 17 where $h_c$and $h_s$indicate the resulted histogram of the cover image and the histogram of the stego image respectively.

$$HE = \frac{\sum_{j=1}^{512}(h_c - h_s)}{512} \tag{17}$$

Table3 shows the result of embedding different sizes of secret message to $512x512$ caver image. Although the increase of HE value with the increase of the size of the secret message, all the resulted HE values are acceptable and good. Figure 5 shows the histogram comparison between the original host image and the resulted stego image of embedding the 85KB of secret message in Lena image. The differences between the histogram graphs is very tiny and cannot be observe by the human visual system. Therefore, the proposed technique can be considered as a secure method against the statistical and visual attack.

Table 3. Histogram Value of the proposed approach with different sizes of the secret message

| Embedding Size | Histogram Error |
|---|---|
| 5% | 2.15 |
| 10% | 2.36 |
| 15% | 3.17 |
| 20% | 3.63 |
| 25% | 3.75 |
| 30% | 4.25 |
| 35% | 4.69 |
| 40% | 5.03 |
| 45% | 5.16 |
| 50% | 5.30 |

5. **Conclusion.** In this paper a new color image stegnography approach is proposed in order to improve the capacity in high visual imperceptibility, and it is used to improve Laplacian of Gaussian edge detector which is based on combination of zero crossing and $9x9$ mask which follow by clustering the resulted blocks to make a difference whether the block in sooth region or edge region. A balanced substitute between security of edge detection based embedding and payload capacity of standard LSB can be achieved on the work based on a non-uniform data distribution algorithm with small computational requirements and improved data carrying surface. It is also added that non-linear data distribution which will also help to secure and provide reasonable resistance to statistical detection of hidden data. It has been approved from the experimental results that not only high imperceptibility could achieve by this proposed method but also make certain high embedding rate.

## REFERENCES

[1] S. Arora and S. Anand, A Proposed Method for Image Steganography Using Edge Detection*International Journal of Emerging Technology and Advanced Engineering,*, 2013

[2] H. C. Huang, J. S. Pan, Y. H. Huang, F. H. Wang, and K. C. Huang, Progressive watermarking techniques using genetic algorithms, *Circuits, Systems, and Signal Processing*, vol.26, no.5, pp.671–687, 2007.

[3] J. S. Pan, H. Luo, and Z. M. Lu, Look-up table based reversible data hiding for error diffused halftone images, *INFORMATICA*, vol.18, no. 4, pp.615–628, 2007.

[4] H. C. Huang, J. S. Pan, and C. M. Chu, Optimized copyright protection systems with genetic-based robust watermarking, *Proc. of the 3rd Int'l Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Kaohsiung, Taiwan, pp.II-465–II-468, 2007.

[5] C. C. Lai, H. C. Huang and C. C. Tsai, Image watermarking scheme using singular value decomposition and micro-genetic algorithm, *Proc. of the 4th Int'l Conf. on Intelligent Information Hiding and Multimedia Signal Processing*, Harbin, China, pp.239–242, 2008.

[6] Y. H. Chen, *A Study on Digital Watermarking and Its Application on Network Multimedia*, Ph.D. Thesis, National Chiao-Tung University, Hsinchu, Taiwan, 2008.

[7] F. C. Chang, H. C. Huang, and H. M. Hang, Combined encryption and watermarking approaches for scalable multimedia coding, in *5th Pacific Rim Conference on Multimedia*, K. Aizawa, Y. Nakamura, and S. Satoh (eds.), Berlin/Heidelberg/New York, Springer-Verlag, LNCS 3333, pp.356–363, 2004.