

Data Hiding Scheme on Medical Image using Graph Coloring

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2015 J. Phys.: Conf. Ser. 622 012028

(<http://iopscience.iop.org/1742-6596/622/1/012028>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 103.233.100.3

This content was downloaded on 11/09/2015 at 12:06

Please note that [terms and conditions apply](#).

Data Hiding Scheme on Medical Image using Graph Coloring

Widi Astuti¹, Adiwijaya², Untari Novia Wisety³

School of Computing, Telkom University, Bandung

¹astutiwidi@telkomuniversity.ac.id; ²adiwijaya@telkomuniversity.ac.id;

³untarinw@telkomuniversity.ac.id;

Abstract. The utilization of digital medical images is now widely spread[4]. The medical images is supposed to get protection since it has probability to pass through unsecure network. Several watermarking techniques have been developed so that the digital medical images can be guaranteed in terms of its originality. In watermarking, the medical images becomes a protected object. Nevertheless, the medical images can actually be a medium of hiding secret data such as patient medical record. The data hiding is done by inserting data into image - usually called steganography in images. Because the medical images can influence the diagnose change, steganography will only be applied to non-interest region. Vector Quantization (VQ) is one of lossy data compression technique which is sufficiently prominent and frequently used. Generally, the VQ based steganography scheme still has limitation in terms of the data capacity which can be inserted. This research is aimed to make a Vector Quantization-based steganography scheme and graph coloring. The test result shows that the scheme can insert 28768 byte data which equals to 10077 characters for images area of 3696 pixels.

1. Introduction

Nowadays, more digital medical images are transmitted between hospitals which are located differently via computer network [5]. Nevertheless, it is hard to find the network which is secure. Therefore, several methods to protect medical images are developed. The research to protect medical images is mainly classified into two purpose categories : to ensure the images originality or to conceal the patient medical record in medical image[5]. Embedding secret data into an image to conceal the secret data is called steganography. The image will change but the most important thing is that no one can read the secret data.

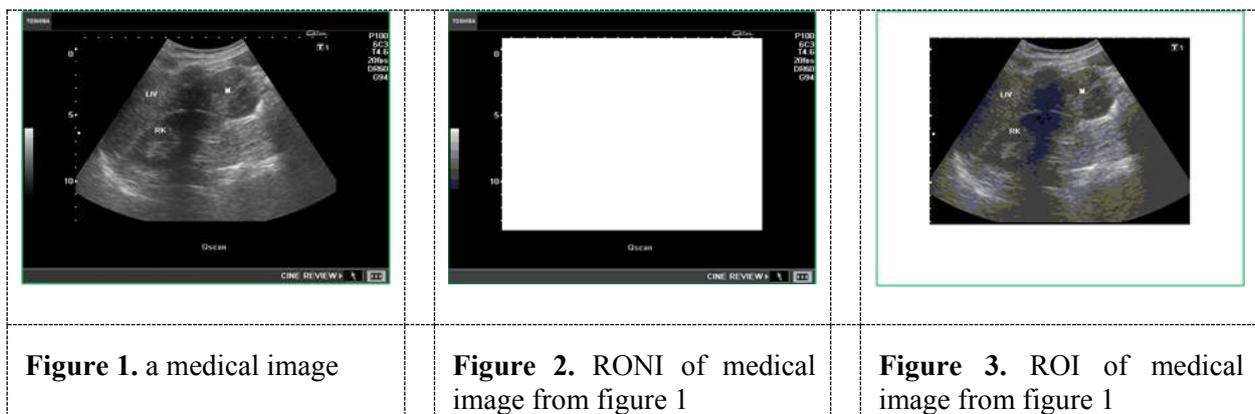
Vector Quantization (VQ) is one of data compression techniques which is lossy, prominent and frequently used. VQ makes use of the fact that many blocks in images have similarities so that an image can be divided into several blocks, and then each block is quantified based on the similarity with a dictionary called codebook. Some data experts attempt to utilize the Vector Quantization in steganography by adopting its principle. VQ and steganography have a similiarity, their process is basically change an image but keep the change visually tolerable. The researches divides the codebook into some clusters then manipulating the quantization process. The clustering process has a very important role since it decided the resulted images quality and data capacity which can be concealed. This clustering process is rarely based on graph coloring. Steganography scheme based on VQ without graph coloring still has limitation in terms of data capacity which can be inserted.



Yue et al. [1] proposed a steganography scheme based on VQ and graph coloring which has good performance with regard to capacity or security. This research attempts to utilize the proposed scheme by Yue et al. for the case of patient medical record data hiding in medical images. Because data hiding in the medical image is very risky due to changes in medical image could lead to changes in diagnosis, this scheme will only be done in its region of non interest, a specific region that is not so sensitive to the change. It is expected that the scheme will be a secure scheme of steganography which will not destroy the important part of medical images but has large capacity of insertion.

2. Data Hiding Scheme in Medical Image

The medical image that will be used in this research is ultrasonography(USG) image since USG is widely used. The medical image can be separated into two region i.e region of interest (ROI) which become the focus in diagnosis making and region of non-interest (RONI) as the figures below shows.



This paper is described generally described in this flowchart.

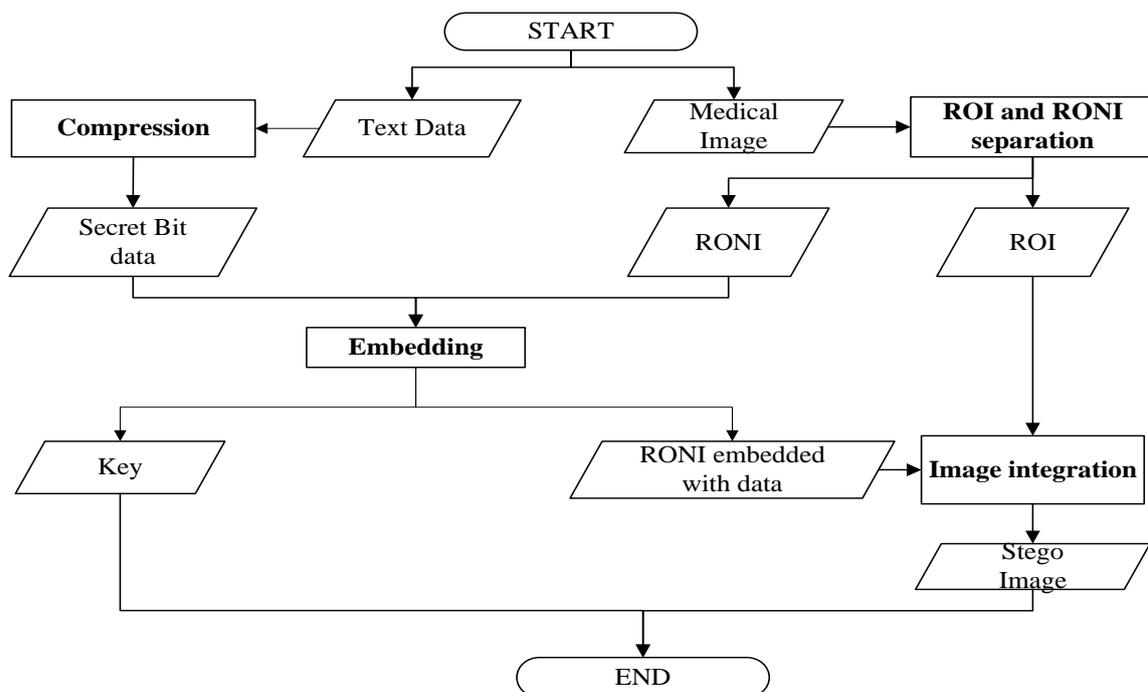


Figure 4. General Scheme

Based on figure 4, secret text data that will be inserted in the image need to be compressed first in order to enlarge the hiding capacity. This research use Huffman Algorithm as the compression method. In the next step, secret data that have been transformed into bit data will be embedded in the image's RONI which had been separated from the ROI, an area that is selected manually. There are two outputs form the embedding process i.e. stego image and key. Stego image is an image that contain secret data. Key is something to protect the secret data. Without the key, the secret data can not be retrieved. Embedding procedure which is modified from [1] is described in figure 5.

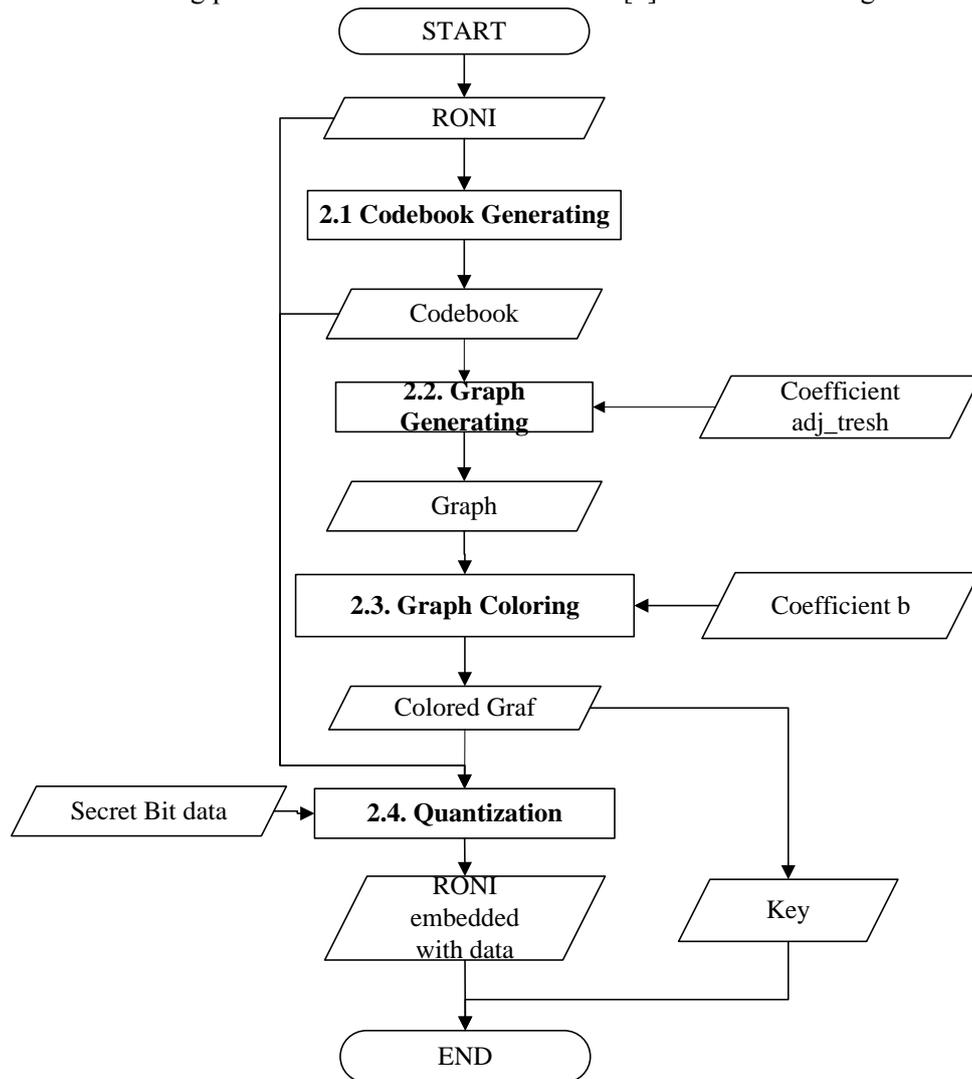


Figure 5. Embedding Procedure

Embedding procedure actually adopt the VQ procedure. First step in VQ is codebook generating, the system construct a codebook that consists of codewords. Each codeword represent a block of pixels which usually the size of one block is 4x4 pixels. Next step in VQ is quantization. In quantization, VQ will process each block by compare it with codebook and find the most similar codeword. The most similar codeword is determined by smallest Euclidian distance between the block and each codeword cw . The equation for the Euclidian distance is shown in Equation (1).

$$D(blok, cw) = \sqrt{\sum_{i=0}^{15} (blok[i] - cw[i])^2} \tag{1}$$

Where,

$D(block, cw)$ = Euclidian Distance between block and codeword cw

block[i] = pixel value in block[i],

cw = pixel value in cw[i],

Then, VQ outputs the index of the most similar codeword as the current compression code in the index table, a table for storing each block compression code. After processing every block in the image, VQ will keep the index table and the codebook.

In this research scheme, before the quantization, scheme do graph coloring in order to clustering the codebook.

2.1. Codebook generating

There are some codebook generating algorithms in VQ. Lu [3] explained that in codebook generating, an image will be divided into some training vectors and the training vectors will undergo a clustering process. This research uses one of the codebook generating algorithm that easily to be implemented i.e. Linde Buzo Gray (LBG) algorithm.

2.2. Graph generating

A graph is constructed by adding a vertex for each codeword in the codebook so that every vertex is representing the codebook. For every vertex $v[i]$ and $v[j]$ in graph representing $cw[i]$ and $cw[j]$, scheme will add an edge between $v[i]$ and $v[j]$ only and if only the values of $cw[i]$ and $cw[j]$ are close enough. The values of $cw[i]$ and $cw[j]$ are considered as close if Euclidian Distance between $cw[i]$ and $cw[j]$ is not exceed adj_tresh , a constant coefficient that determined empirically.

2.3. Graph Coloring

The graph is colored using a number of color called $color_number$. Each color can be represented using an integer so that $color = \{0, 1, 2, 3, \dots, color_number - 1\}$. $Color_number$ value must be satisfied 2^x for an integer x . Each color is corresponding with x bit of secret message. For example, if $color_number = 8$, the colors are: $= \{0, 1, 2, 3, 4, 5, 6, 7\} = \{000, 001, 010, 011, 100, 101, 110, 111\}$. Since graph coloring is a NP complete problem, this research will utilize Particle Swarm Optimization to perform it. Graph Coloring in this scheme is adopting Cui[2] scheme. The difference is Cui's scheme is only used to do graph coloring using 4 colors, but in this scheme, Cui's scheme is modified so that it can be used to do graph coloring using 2^n colors.

2.4. Quantization

The last step on embedding process is a quantization. The quantization is described as follows.

Inputs : RONI, colored graph, codebook, and bit data

Outputs: RONI embedded with data

1. The scheme divides RONI into some blocks
2. The scheme reads the first block
3. The scheme searches codeword c that is most similar with current block using Euclidian Distance
4. The scheme searches vertex v that is corresponding with codeword c .
5. The scheme retrieves x bit from secret bit data where $2^x = color_number$.
6. The scheme searches using Breadth First Search for vertex v' that either directly or indirectly connected with v and satisfied $color(v') = x$; If v' found, go to step 7, If v' can not be found go to step 8
7. Codeword c' that is corresponding with v' will be the output. x bit from secret bit are deleted from secret bit data. Go to step 9.
8. Codeword c will be the output. Go to step 9.
9. The scheme processes next block using step 3-8 until every blocks are processed.

3. Testing and Analysis

In this testing, a same RONI of USG image will be embedded by some secret text using 8 scenarios. Two parameters will be used to measure the scheme performance i.e. hiding capacity and Peak Signal to noise Ratio (PSNR). The scenarios are created by changing three parameters i.e. *codebook_size*, *adj_tresh*, and *color_number*.

3.1. Codebook size

Generally, the size of codebook used in VQ is 256 [1]. But, [1] shows that in a data hiding scheme, codebook size is affecting the scheme performance. Testing in this research changing the codebook size between 64 and 128 because only some region from the image will be embedded by secret data and RONI of medical image tends to be homogenous so there is no need codebook as big as common image.

3.2. Value of adjacency threshold

Adj_tresh is a coefficient that has important role in graph generating phase since the changing of *adj_tresh* will cause the change of graph generated. Test is conducted by changing the value of variable (*adj_tresh* = 60, or *adj_tresh* = 70).

3.3. Color number

Color_number is a parameter that is determining in graph coloring and quantization. Test is conducted by changing the value of variable (*color_number* = 8, or *color_number* = 16).

After conducting the test, the result is shown on the table below.

Scenario	<i>codebook_size</i>	<i>adj_tresh</i>	<i>color_number</i>	PSNR	Capacity (bit)
1	64	8	60	17,35%	22 371
2	64	8	70	16,94%	28 768
3	64	16	60	17,21%	22 650
4	64	16	70	17,78%	94
5	128	8	60	17,46%	22368
6	128	8	70	18,20%	4836
7	128	16	60	16,93%	22668
8	128	16	70	18,10%	1060

The test is aimed at determining the best variable at the scheme. Best result is determined by highest PSNR and hiding capacity. Based on the test results above, it can be seen PSNR will increase as *codebook_size*, *adj_tresh* and *color_number* increase. In term of hiding capacity, as *adj_tresh* and *color_number* increase, the hiding capacity will decrease. It can not be decided yet how codebook size can affect the hiding capacity. It can be conclude that the most optimum parameter is *codebook_size*=64, *color_number* = 16, and *adj_tresh*=60 with byte capacity 28768 equivalent to 10077 characters.

4. Conclusion

Based on the analysis towards the test on the data hiding scheme, it can be concluded that the Vector Quantization and Graph Coloring can be implemented for hiding the data in medical image. The best performance is produced by the parameter *codebook_size*=64, *color_number* = 16, and *adj_tresh*=60 with byte capacity 28768 equivalent to 10077 characters. In term of security, the data which is concealed on medical images using this scheme can not be retrieved without knowing the key, ie the colored graph.

References

- [1] Yue, Shuai *et al.*, 2013, “ An Image Data Hiding Scheme Based On Vector Quantization and Graph Coloring”. Springerlink: Recent Advances in Information Hiding and Applications , 1-17. [1] Cheddad, Abbas (dkk), 2010, “Digital Image Steganography: Survey and analysis of current methods”. Signal Processing , 90 (3), 727-752.
- [2] Cui, Guangzhao, *et al.*, 2007, “Modified PSO algorithm for solving planar graph coloring problem”. Progress in Natural Science 18 (2008), 353-357.
- [3] Lu, Tzu-Chuen, 2010, “A Survey of VQ Codebook Generation”. Journal of Information Hiding and Multimedia Signal Processing , 1 (3), 190-203.
- [4] Memon, Nisar Ahmed, 2010, “ Watermarking of Medical Images for Content”. Phd Thesis, Faculty of Computer Science and Engineering, GIK Institute of Engineering Sciences and Technology, Khyber Pakhtunkhwa
- [5] Adiwijaya, T.A.B. Wirayuda, S.D. Winanjuar, U. Muslimah, “The Multiple Watermarking on Digital Medical Image for Mobility and Authenticity”, Operations Research Proceedings 2012, 457-462