Data Hiding Method Replacing LSB of Hidden Portion for Secret Image with Run-Length Coded Image

Kohei Arai[1]
[1]Graduate School of Science and Engineering
Saga University
Saga City, Japan

***Abstract*—Data hiding method based on steganography improving secret image invisibility through replacing Least Significant Bit: LSB with run-length coded Image is proposed. Although, the proposed method is based on Discrete Wavelet Transformation: DWT, run-length coded secrete image is embedded in the LSB of the high frequency component before reconstruction of the image for opening to the public for improvement of the secrete image invisibility. Before embedding the coded secret image, the coded secret image bits are rearranged the order by using random number. Therefore, secret image invisibility is much improved. Through experiments, it is confirmed that secret images are almost invisible in distribute images. Data hiding performance in terms of invisibility of secret images which are embedded in distribute images is evaluated with the Peak Signal to Noise Ratio: PSNR and the Root Mean Square: RMS difference between the original secret image and extracted one from the distribute images. Meanwhile the conventional Multi-Resolution Analysis: MRA based data hiding is attempted with a variety of parameters, level of MRA and the frequency component location of which secret image is replaced to it and is compared to the proposed method. It is found that the proposed data hiding method is superior to the conventional method.**

Keywords—Wavelet; DWT; Steganography; Random number based Permutation; Data hiding; Data compression

I. Introduction

Wavelet analysis applications are getting more popular in time series analysis, image analysis, information security area, etc.[1],[2]. Data hiding is important for information contents security, in particular, for protection of copy right. Using data hiding methods, some important information such as copyright, signature, etc. can be embedded. Data hiding based on wavelet analysis, in particular, Multi-Resolution Analysis: MRA is widely used. One of the problems on data hiding methods is visibility of the embedded information on the available distribute images [3]-[7]. The other problem is robustness against image processing which is applied to the distribute images including data compressions. It sometime occurs that small amount of information on the embedded image appears on the distribute images slightly due to the embedding mechanism of the data hiding.

In order to improve invisibility of the secret images in the distribute images, run-length coded binarized secret images are used. The locations of the codes after the data compression in one of the frequency component images after the dyadic wavelet transformation [8] are determined with random numbers generated by Mersenne Twister of random number generator. After all, reconstructed image (inverse dyadic wavelet transformation) is used for distribute images. The original secret images are almost invisible in the distribute images. Although the method for data hiding based on Legall 5/2 (Cohen-Daubechies-Feauveau: CDF 5/3) wavelet with data compression and random scanning of secret imagery data is proposed [9], invisibility of the secret image is not good enough. Therefore, improvement of the invisibility is still required.

This paper deals with the current problems on the widely used MRA based data hiding method (Conventional data hiding method). One of the problems is visibility of secret image in the distribute images followed by robustness against distribute image manipulations including image deformation, geometric conversion, data compression, etc. In order to overcome the aforementioned problems, a method for data hiding based on lifting dyadic wavelet transformation with run-length coding of data compression which is applied to secret image together with pixel order exchange is proposed. This is well reported in the previously published paper [10]. The method proposed is to improve the invisibility of the secret images by using the process of which the run-length coded secret image is replaced to several bits from the Least Significant Bit: LSB of one of the decomposed images. It is called as “Steganography” that the coded secret image is embedded in the LSB. LSB is not so visible. Therefore, embedding the coded secret image in the LSB implies improvement of secret image invisibility. Before replace the secret image to one of the decomposed components of the original image, secret image is encoded by run-length coding with Modified Huffman code. Then the coded image is scanned with the random numbers generated with Mersenne Twister and is replaced to the one decomposed components with the information of compressed image size, initial value of the random number and the parameters of wavelet transformation based on lifting dyadic wavelet. Thus the secret image becomes invisible.

First, the aforementioned problems of the conventional MRA based data hiding method are described followed by the proposed method. Then it is enhanced that the proposed method achieved significantly improvement of invisibility of the secret images through experiments with the standard image database for data compression performance evaluations.

II. PROPOSED METHOD

A. Conventional Data Hiding Method Based on Discrete Wavelet Transform

Wavelet utilized MRA allows decompose the image with wavelet coefficients (high and low frequency components) and also the original image can be reconstructed with the decomposed wavelet coefficients perfectly. If the high frequency component is replaced with secret image to be hidden, and if the reconstructed image is distributed to public, then secret image can be extracted from the distributed image if receiver knows which component is replaced with secret image. In this case, secret image has to be invisible in the distributed image. Also even if image processing including geometric conversion (linear transformation) and data compression (nonlinear transformation) is applied to the distributed image, secret image has to be extracted. The aforementioned "invisibility" and "robustness against image processing" are very important for data hiding.

The proposed method for data hiding is based on dyadic wavelet transformation. Dyadic wavelet allows to separate frequency components keeping image size with that of original image. Dyadic wavelet is called as a binary wavelet and has high pass and low pass filter components, $\{h[k], g[\mathrm{k}]\}$ and reconstruction filter $\{h[k], g[\mathrm{k}]\}$. Low and high frequency components, c_n and d_n are expressed as follows,

$$c_n\,[\mathrm{i}]=\Sigma_{\mathrm{k}}\, h[k\,]C_{n-1}\,[i + k_{2n-1}\,] \tag{1}$$
$$d_n\,[\mathrm{i}]=\Sigma_{\mathrm{k}}\, g\,[k\,]C_{n-1}\,[i + k_{2n-1}\,] \tag{2}$$

Then original image is also reconstructed with the low and high frequency components as follows,

$$c_{n-1}\,[i\,]=1/2\Sigma_{\mathrm{k}}\mathrm{h}[k\,]c_n[i{-}k2^{n-1}]+\Sigma_k g[k\,]\mathrm{d}_{\mathrm{n}}\,[i{-}k2^{n-1}\,] \tag{3}$$

If a new parameter s[m] is employed, then lifting dyadic wavelet is defined as follows,

$$h^{\mathrm{new}}[k\,]=h^{\mathrm{old}}\,[\mathrm{k}\,] \tag{4}$$
$$h^{\mathrm{new}}[k\,]=h^{\mathrm{old}}\,[\,k\,] + \Sigma_{\mathrm{m}}\, s[-m]\, g^{\,\mathrm{old}}\,[k{-}m] \tag{5}$$
$$g^{\mathrm{new}}[k\,]=g^{\mathrm{old}}\,[\,k\,] - \Sigma_{\mathrm{m}}\, s[m]\, h^{\mathrm{old}}\,[\,k{-}m] \tag{6}$$
$$g^{\mathrm{new}}[k\,]=g^{\,\mathrm{old}}\,[k\,] \tag{7}$$

In Fig.1, the secret binary image of Mandrill [11] with a certain threshold is embedded in the HL component of the dyadic wavelet transformed images derived from the original image of Lenna [11] with dyadic wavelet transformation. At the left bottom corner of Fig.1, a reconstructed image (image for circulation) is shown. The secret image can be recognizable in the circulation image, unfortunately. In these cases, Daubechies wavelet base function (support length=2) [12] is used. On the other hand, Fig.1 (b),(c),(d) shows reconstructed images of Mandrill of secret image (a) embedded Lenna of original images of which the secret image is embedded at the LH, HH, and HL of frequency components, respectively. Image size is not changed for original and dyadic wavelet transformed images.

B. Problem Description

Fig.2 shows a schematic process flow of the proposed data hiding based on lifting dyadic wavelet transformation. It is possible to hide the embedded image at the certain location of wavelet transformation images then distribute images containing the embedded image can be reconstructed through inverse wavelet transformation. In this case, although binary secret images are assumed, half tone, colored images are also available for secret images.

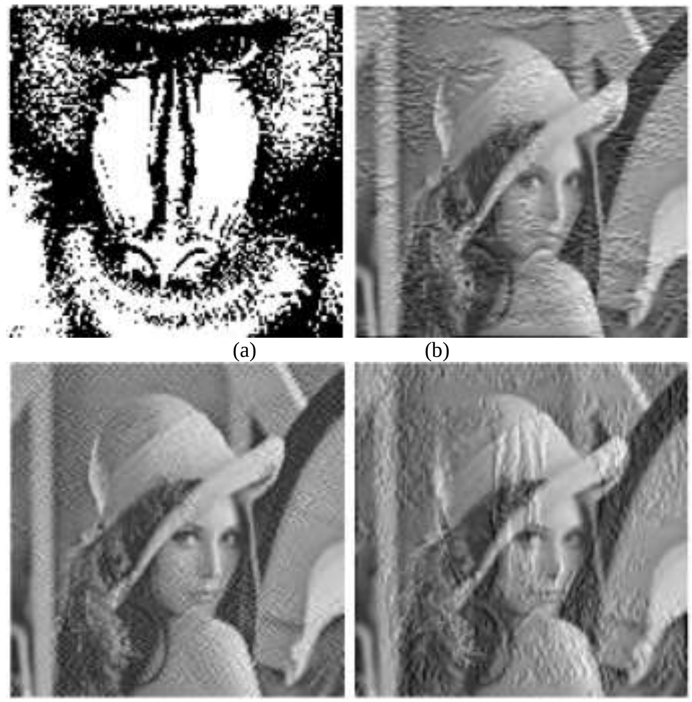


Fig. 1. (a) The binalized secret image of Mandrill, (b) Reconstructed image of Mandrill image embedded (LH) Lenna image for circulation, (c) Reconstructed image of Mandrill image embedded (HH) Lenna image for circulation, (d) Reconstructed image of Mandrill image embedded (HL) Lenna image for distribution (Image size of the secret and the reconstructed images are 128 by 128 pixels and quantization bit.is 8 for both images)

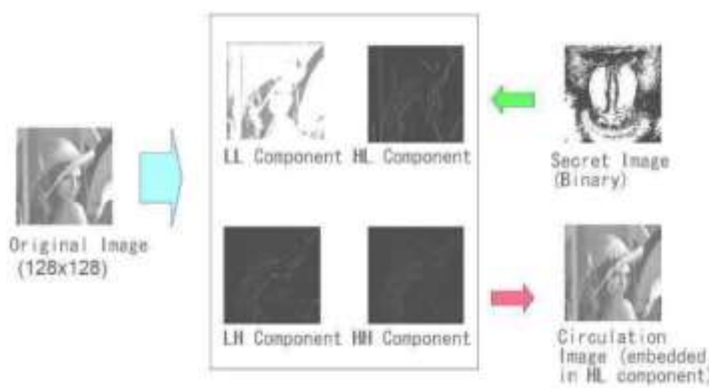
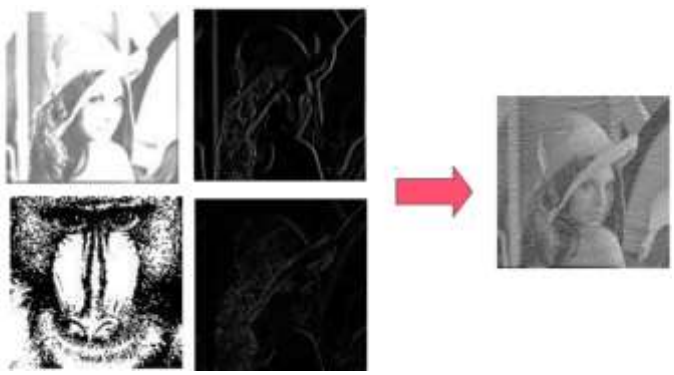
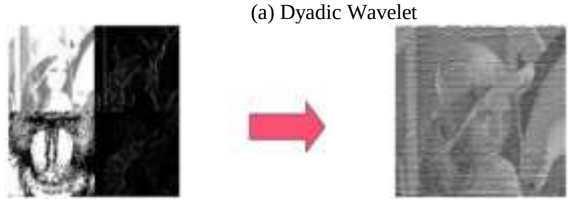


Fig. 2. Schematic process flow of the proposed data hiding

Root Mean Square: RMS difference between original and the reconstructed images is used for the evaluation of data hiding performance. The parameters of the wavelet transformation are (1) Basis function, (2) Frequency component in which the secret image is replaced, (3) Level of the wavelet transformation (wavelet transformation can be applied recursively. The number of recursive corresponds to filter bank frequency) and (4) Downsizing parameter (Dyadic wavelet transformation makes downsizing parameter is one while downsizing parameter of the ordinal wavelet transformation is used to be 1/2). Fig.3 shows RMS difference of a variety of parameters of wavelet transformation based data hiding. Fig.3 (a) shows one of examples of reconstructed image of the data hiding method with dyadic wavelet transformation while Fig.3 (c) shows another example of the reconstructed image of the data hiding method with the conventional Daubechies wavelet transformation. Meanwhile, RMS difference between the original image and the reconstructed images for the dyadic wavelet transformation based data hiding is shown in Table 1 while that for the conventional Daubechies wavelet transformation is shown in Table 2. In Fig.3, Lv1 denotes level 1 while Dbx denote Daubechies base function with support length of x.



(a) Dyadic Wavelet



(b) Daubechies wavelet

Fig. 3. RMS difference of a variety of parameters of wavelet transformation based data hiding

TABLE I. RMS DIFFERENCE FOR DYADIC WAVELET BASED DATA HIDING

Lv1	Db1	Db2	Db4
LH	22.93	22.54	22.32
HH	20.5	20.38	20.3
HL	23.66	22.86	22.56

TABLE II. RMS DIFFERENCE FOR DAUBECIES WAVELET BASED DATA HIDING

Lv1	Db1	Db2	Db4
LH	80.2	79.37	79.59
HH	80.22	79.51	79.74
HL	80.89	79.87	80.12

It is quite obvious that the dyadic wavelet transformation based data hiding method is superior to the conventional Daubechies wavelet based data hiding method.

C. Previously Proposed Data Hiding Method

First, secret image is binarized. It is usual that secret images are used to be authorization images for protection of original image content (copyright). Before binarized secret images are replaced to one of the high frequency component images, run-length coding is applied to secret images in order to improve an invisibility of the secret images in the distribute images. Fig.4 shows schematic process flow of the run-length coding method. The number of pixels in the original binary image is 27 while the number of pixels in the compressed image is just 6 (quantization level is variable).

XXXXX YYY XXX YYYYY X YYYYYYYYYY 27

↓

5 3 3 5 1 10 6 of run-length

Fig. 4. Schematic process flow of run-length coding

Then run-length coded data are replaced to one of the high frequency components with the pixel order exchanges based on generated random numbers which are generated by Mersenne Twister. Only if the receiver who knows the initial value of random number of Mersenne Twister and how to decode run-length coding, then such the receiver can extract the secret images. Thus the copyright holders can assert their copy right through extraction of secret images. It, however, is not enough in terms of invisibility of the secret image.

D. Proposed Data Hiding Method

Fig.5 shows the process flow of the proposed data hiding (hide the secret image into the original image and create distribute image embedded the secret image then extract the secret image from the distribute image)

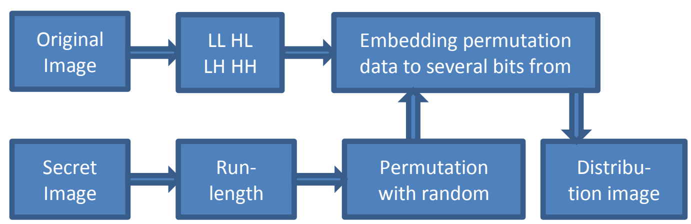


Fig. 5. Process flow of the proposed method

The method proposed is to improve the invisibility of the secret images by using the process of which the run-length coded secret image is replaced to several bits from the Least Significant Bit: LSB of one of the decomposed images derived from DWT. Several bits from LSB is almost invisible in the one of the high frequency components in the decomposed images based on DWT of wavelet transformation. Before replace the secret image to one of the decomposed components of the original image, secret image is encoded by run-length coding with Modified Huffman code which is shown in Fig.6.

Modified Huffman code consists of terminating code and make-up code and End of Line: EOL. An example is shown in the bottom of Fig.6.

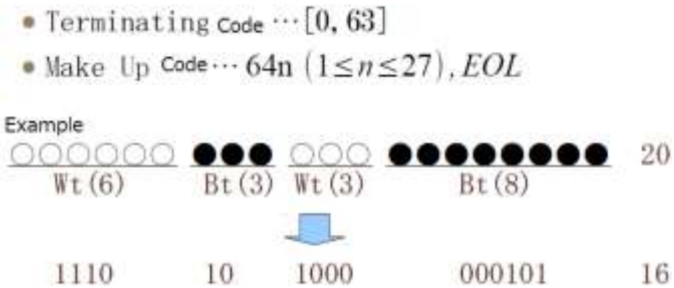


Fig. 6. Modified Huffman coding

Then the coded image is scanned with the random numbers generated with Mersenne Twister and is replaced to the one decomposed components with the information of compressed image size, initial value of the random number and the parameters of wavelet transformation based on lifting dyadic wavelet. All the required information for reconstruction of the secret image, such as wavelet transformation parameters, compressed image data size and the initial value of the random number generator are stored in several bits from the LSB of the one of the decomposed image of high frequency component of the secret image. Thus the secret image becomes much invisible in comparison to the previously proposed method.

III. EXPERIMENTS

A. Data Used

“Lenna” in the SIDBA image database is used for the original image while “Shuttle cock” and “Mandrill” in the SIDBA image database is used for the secret images. Fig.7 shows the original and the secret images. The 8-bits secret image is represented as 8 bit-plane (binary data). Each bit-plane can be encoded with run-length coding. The coded data is embedded in several bits of HH component image derived from DWT. In this process, pixel order is permutated with random number derived from Merssene Twister.

Fig. 7. Original (Top) and secret images (Bottom) used for the experiments

Not only RMS difference, but also the following Peak Signal to Noise Ratio: PSNR in unit of dB is evaluated in the experiments.

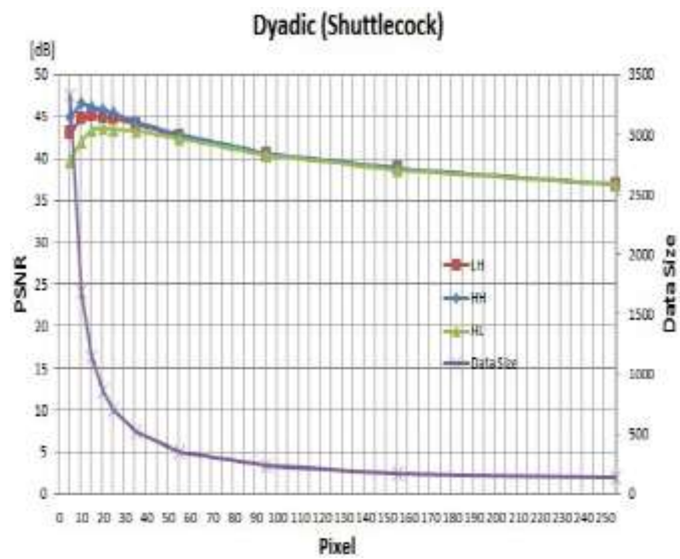
$$PSNR=20\log_{10}(Max/RMSE) \quad (8)$$

where Max=255 (8-bit quantization) and RMSE denote RMS difference between the original and the reconstructed images.

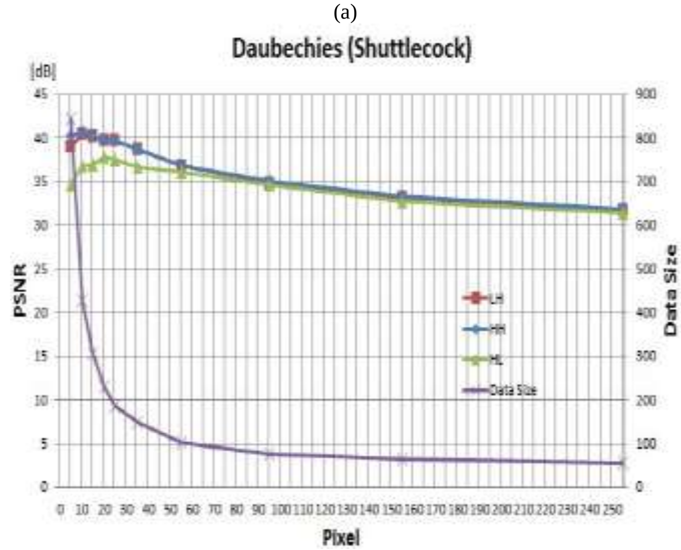
B. Preliminary Experimental Results

Level 1 of wavelet transformation is applied to the original image with dyadic and the conventional Daubechies based wavelet transformations. The secret images are replaced to the high frequency components, LH, HL, and HH after the wavelet transformation. Then the image is reconstructed for evaluation of PSNR. Fig.8 shows the evaluation results of PSNR and data size as a function of the number of pixels of the secret image size ranged from 4 to 256.

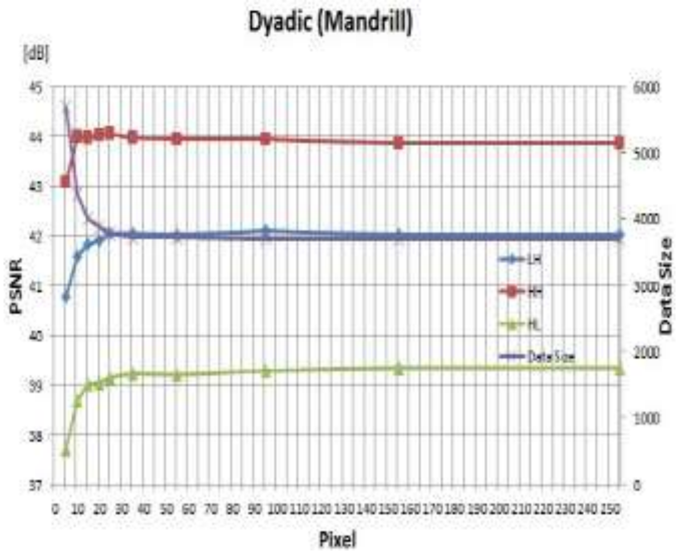
PSNR shows a peak at the secret image size of 8. Then the evaluated PSNR decreases in accordance with the increasing of the secret image size. Also it is found that HH component is most appropriate component in which the encoded secret image inserted to LSB of the HH component. Furthermore, it is found that dyadic wavelet transformation based data hiding is superior to the conventional Daubechies wavelet transformation based data hiding.



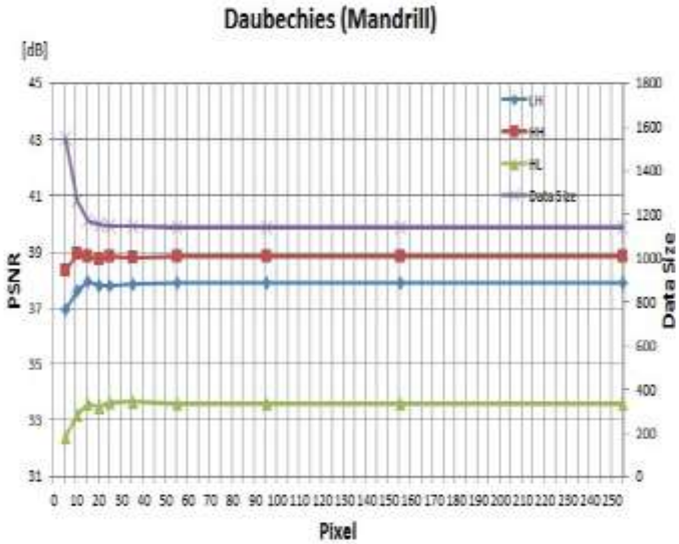
(a)



(b)



(c)



(d)

Fig. 8. Evaluation results of PSNR and data size as a function of the number of pixels ranged from 4 to 256

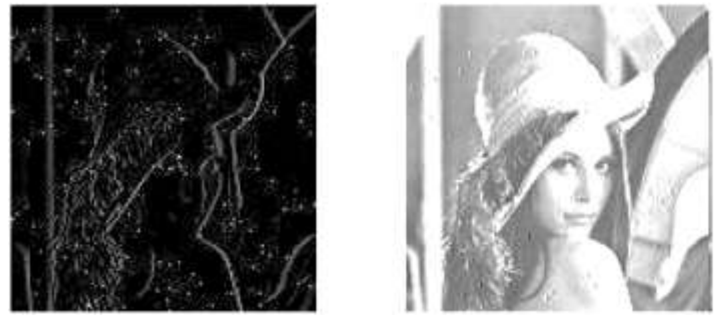
C. Influence Due to Threshold

The secret image of "Mandrill" is 8-bit quantized image. Therefore, data hiding performance, in terms of invisibility of the secret image depends on the threshold used for generation of the secret image.

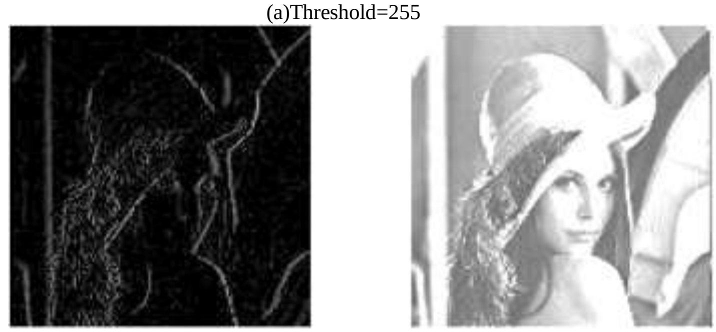
Fig.9 shows the HH component image (on the left) and the reconstructed image (on the right). PSNR for the threshold= 255 is 36.90 dB while that for the threshold=15 is 45.16 dB. Meanwhile, the number of data for the threshold=255 is 133 while that for the threshold=15 is 1140. Therefore, it is better to select the most appropriate threshold before applying data hiding.

D. Comparison Between with and without LSB Insertion

There are two different proposed data hiding methods, with and without LSB insertion. Namely, the proposed method seems to be worked without LSB insertion. The data hiding method without LSB insertion is called standard method (STD) while that with LSB insertion is called LSB insertion method (LSB). "Shuttle cock" and "Mandrill" of the secret images are quantized with quantization bit ranged from 4 to 32.



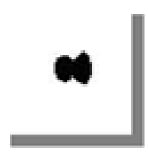
(a)Threshold=255



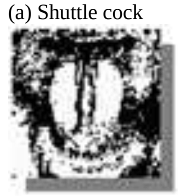
(b)Threshold=15

Fig. 9. Influence due to the threshold

Then the quantized secret images are encoded with run-length coding. After that the proposed data hiding methods with and without LSB insertion is applied. PSNR is evaluated for the proposed methods, with and without LSB insertion. Fig.10 shows the secret images. Table 3, 4 shows the evaluated PSNR for the secret image of Shuttle cock and Mandrill, respectively. In accordance with increasing of the number quantization bit, the number of hidden data is decreased obviously. Also, PSNR decreases in accordance with increasing of the quantization bit for the proposed data hiding method with LSB insertion while PSNR does not depend on the data hiding method without LSB insertion. Furthermore, PSNR of the data hiding method with LSB insertion is superior to the data hiding method without LSB insertion. Therefore, it is better to select the data hiding method with LSB insertion with a small quantization bit.



(a) Shuttle cock



(b) Mandrill

Fig. 10. Secret images

TABLE III. PSNR FOR THE PROPOSED DATA HIDING METHODS, WITH AND WITHOUT LSB INSERTION (SHUTTLE COCK)

Type#1	Hidden_Data	STD	LSB
4bir	1395	38.59	52.12
8bit	608	41.16	48.19
16bit	306	40.7	44.4
32bit	158	38.82	40.65

TABLE IV. PSNR FOR THE PROPOSED DATA HIDING METHODS, WITH AND WITHOUT LSB INSERTION (MANDRILL)

Type#2	Hidden_Data	STD	LSB
4bit	2085	36.85	52.14
8bit	1378	38.43	49.0
16bit	1169	38.71	45.33
32bit	1147	38.63	42.84

E. Evaluation of the Proposed Method

Another dataset is prepared as shown in Fig.11. The original images, “Lenna” and “Bridge” are 128 by 128 pixels of 8-bit quantized greyscale images while the secret images, “Barbra” and “SIDBA Title” are 32 by 32 pixels of 8-bit quantized greyscale images.

Fig. 11. Original and secret images used for the experiments

In accordance with the procedure of the proposed data hiding method, the secret images, Barbra and SIDBA-Title are embedded in the original image, Lenna and Bridge. In case of Barbra is embedded in the original images, data compression ratio is 71.76% (Fig.12 (a)) while SIDBA-Title is embedded in the original images, data compression ratio is 65.92% (Fig.12 (b)).

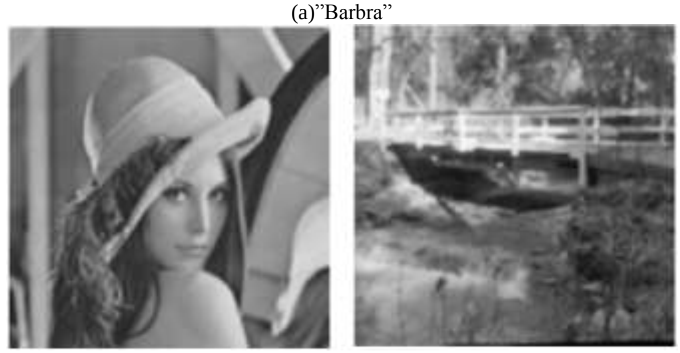
PSNR for these cases are shown in Table 5.

TABLE V. PSNR FOR THE PROPOSED DATA HIDING METHOD

PSNR	Lenna	Bridge
Barbra	46.83	47.21
SIDBA-Title	46.90	47.40

PSNR for all the cases are over 46 dB and are beyond the preliminary case studies. Therefore, it may conclude that the secret images are almost invisible.

(a)”Barbra”



(b)”Sidba-Title

Fig. 12. Distribution images for the cases that Barbra and SIDBA-Title is embedded in the original images

IV. CONCLUSION

Data hiding method based on steganography improving secret image invisibility through replacing Least Significant Bit: LSB with run-length coded Image is proposed. Lifting dyadic wavelet is used for fundamental data hiding. Although, the proposed method is based on Discrete Wavelet Transformation: DWT, run-length coded secrete image is embedded in the LSB of the high frequency component before reconstruction of the image for opening to the public for improvement of the secrete image invisibility. Before embedding the coded secret image, the coded secret image bits are rearranged the order by using random number. Therefore, secret image invisibility is much improved.

Through experiments, it is confirmed that secret images are almost invisible in distribute images. Data hiding performance in terms of invisibility of secret images which are embedded in distribute images is evaluated with the Peak Signal to Noise Ratio: PSNR and the Root Mean Square: RMS difference between the original secret image and extracted one from the distribute images. Meanwhile the conventional Multi-Resolution Analysis (MRA) based data hiding is attempted with a variety of parameters, level of MRA and the frequency component location of which secret image is replaced to it and is compared to the proposed method.

It is found that the PSNR shows a peak at the secret image size of 8. Then the evaluated PSNR decreases in accordance with the increasing of the secret image size. Also it is found that HH component is most appropriate component in which the encoded secret image inserted to LSB of the HH component. Furthermore, it is found that dyadic wavelet transformation based data hiding is superior to the conventional Daubechies wavelet transformation based data hiding. PSNR for all the cases of the experiment conducted for the proposed data hiding method are over 46 dB and are beyond the preliminary case studies. Therefore, it may conclude that the secret images are almost invisible.

Further investigations are required for utilization of relational information, shape information, and so on. Also, it is desirable to conduct a research on image retrieval based on deep learning. On the evaluation of the image quality, not only PSNR but also other measures have to be created in an realistic manner.

ACKNOWLEDGEMENTS

Author would like to thank Mr.Yuji Yamada, former Saga University student for his great effort to conduct the experiments.

REFERENCES

[1] Kohei Arai, Fundamental Theory on Wavelet Analysis, Morikita Shuppan Publishing Co. Ltd., 2000.

[2] Kohei Arai, Self Learning on Wavelet Analysis, Kindai-Kagakusha publishing Co. Ltd.. 2006.

[3] Kohei Arai and Kaname Seto, Data hiding method based on Multi-Resolution Analysis: MRA, Visualization Society of Japan, 22, Suppl.No.1, 229-232, 2002.

[4] Kohei Arai and Kaname Seto, Data hiding method with coordinate conversion in feature space, Visualization Society of Japan, 25, Suppl.No.1, 55-58, 2005.

[5] Kohei Arai and Kaname Seto, Improvement of invisibility of secret images embedded in circulate images based on MRA with coordinate conversion and Principal Component Analysis: PCA, Journal of Image and Electronics Society of Japan, 36, 5, 665-673, 2007.

[6] Kohei Arai and Kaname Seto, Improvement of invisibility of secret images embedded in circulate images based on MRA with scanning scheme conversion, Visualization Society of Japan, 29, Suppl.No.1, 167-170, 2009.

[7] Kohei Arai, Improvement of security and invisibility of secret images embedded in circulate images and based on MRA, Report of RIMS-Research Institute for Mathematical Sciences Kyoto University, ISSN188-2818,1684,93-113,2010.

[8] S.Mallat and S.Zhong, ``Characterization of signals from multiscale edges," IEEE Trans. Pattern Anal. Machine Intell., 14, pp.710-732, 1992.

[9] Kohei Arai, Method for data hiding based on Legall 5/2 (Cohen-Daubechies-Feauveau: CDF 5/3) wavelet with data compression and random scanning of secret imagery data, International Journal of Wavelets Multi Solution and Information Processing, 11, 4, 1-18, B60006 World Scientific Publishing Company, DOI: I01142/SO219691313600060, 1360006-1, 2013

[10] Kohei Arai, Yuji Yamada, Improvement of Secret Image Invisibility in Distribute Image with Dyadic Wavelet Based Data Hiding with Run-Length Coded Secret Images of Which Location of Codes are Determined with Random Number, IJARAI, Special Issue on Artificial Intelligence, 33-40, 2012.

[11] http://vision.kuee.kyoto-u.ac.jp/IUE/IMAGE_DATABASE/STD_IMAGES/ (Accessed on March 11 2011).

[12] Kohei Arai and Leland Jameson, Earth observation satellite data analysis based on wavelet analysis, Morikita-Shuppan Publishing Co., Ltd., 2001.

AUTHORS PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 and also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post-Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a counselor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission-A of ICSU/COSPAR since 2008. He received Science and Engineering Award of the year 2014 from the minister of the ministry of Science Education of Japan and also received the Bset Paper Award of the year 2012 of IJACSA from Science and Information Organization: SAI. In 2016, he also received Vikram Sarabhai Medal of ICSU/COSPAR and also received 20 awards. He wrote 34 books and published 520 journal papers. He is Editor-in-Chief of International Journal of Advanced Computer Science and Applications as well as International Journal of Intelligent Systsems and Applications. http://teagis.ip.is.saga-u.ac.jp/