Wavelet Multi Resolution Analysis based Data Hiding with Scanned Secrete Images

Kohei Arai
Faculty of Science and Engineering
Saga University, Saga City, Japan

***Abstract*—Wavelet Multi Resolution Analysis (MRA) based data hiding with scanned secrete images is proposed for improvement of invisibility of the secrete images. Daubechies (biorthogonal basis was adopted as the wavelet, but it was demonstrated that the key image (or secret image data) information can be restored with the biorthogonal wavelet. Also, the information of what to adopt as the biorthogonal wavelet is hidden. Key image information can also be protected by doing so, that the horizontal biorthogonal wavelet of the image does not have to be the same as the vertical biorthogonal wavelet, and the insertion position of the secret image data can be freely selected. It is also possible to divide the bit string of the secret image data and insert it into an arbitrary high frequency component, that the information hiding capability changes depending on the number of bit strings (information amount) of the secret image data, and the secret image in the public image data. Random scanning is effective for improving the visibility of data, selection of scanning method type, random number initial value It was shown that sharing only among parties is useful for improving confidentiality, resistance to noise, resistance to data compression, and resistance to tampering with data.**

Keywords—Multi-Dimensional wavelet transformation; multi resolution analysis: MRA; image data hiding; scanned secrete image; Daubechies basis function; invisibility

I. Introduction

Although personal works are often represented by digital format files, etc., the current situation is that the method of claiming the copyright of the digital contents works is unbearable. In other words, copyright cannot be protected even if it is plagiarized without knowing how to claim the copyright. The importance of digital forensics is being emphasized. That is, evidence is getting more important. How should the proof of copyright infringement be left behind? That is the question of the research. For this purpose, the digital content itself is hidden and hidden only between the recipient and the third party.

There is a method to send and receive so that it does not exist. Data hiding technology. Data hiding is a general term for steganography and digital watermarking. When the information to be embedded is important and its existence is not known, steganography, and when the content itself in which the secret information is embedded is important, it is generally referred to as a digital watermark [1].

In steganography, there is a trade-off between the quality of multimedia content and the amount of information that can be embedded. In digital watermarking, there is a trade-off between resistance to attacks and the amount of information that can be embedded [2]. To efficiently perform a cryptographic protocol, such as a digital fingerprint system, a method that suppresses the amount of calculation and communication may be an excellent method. In addition, it is important to have a digital watermark technology that is resistant to attacks such as falsification and deletion of embedded digital fingerprints [3].

The data hiding introduced in this paper allows digital contents to keep secret keys of authors in circulation so that copyright can be claimed. This makes it possible for an author who can know the secret key to claim the copyright by taking out the distribution content from the distribution content.

The secret key must not be visible to the distributed contents, and this invisibility is important. It is also important to improve the confidentiality by devising a method to keep the secret key in the distribution contents. One of the methods is to hide the secret key in the decomposition factor in wavelet multiresolution analysis. Especially, if it is hidden in high wavelet frequency components, the visibility is generally high [4], [5], [6], [7]. The wavelet-based data hiding method includes reversible data hiding by the histogram gap method based on the integer wavelet, in addition to the method based on this multiresolution analysis [8].

Wavelets allows time-frequency analysis. Wavelet Multi Resolution Analysis: MRA based on biorthogonal basis function of Daubechies is applicable for a variety of application fields [9], [10], [11]. One of the application fields is data hiding.

If the frequency component in which the secret key is embedded is searched by the brute force method or the like, the secret key may be stolen or tampered with. Therefore, it is extremely dangerous to simply perform data hiding using multi-resolution analysis. Therefore, in this paper, data hiding by multi-resolution analysis is preprocessed, and the parameters of the preprocessing that only authors who can do it also need to know together with the information about the frequency component to be embedded. The author devised it so that the author could not find the key. To improve the invisibility of the secret key image in the distribution image, the secret key image is rescanned in accordance with the Hilbert scan algorithm or random scanning algorithm as a preprocessing of the MRA-based data hiding.

Section II outlines data hiding based on multi-resolution analysis, and Section III proposes a method for performing sequence order conversion and permutation conversion processing by random scanning on the bit array of the secret key. The data hiding process in which the image data in the database is selected as the original image is exemplified, and the confidentiality and the visibility difficulty of the secret content in the distribution content are evaluated in Section 4. Sections 5 and 6 gives conclusions and future work, respectively.

II. Outline of Data Hiding based on Multi-Resolution Analysis

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 [12]. Improvement of secret image invisibility in circulation image with Dyadic wavelet-based data hiding with run-length coding is also proposed [13]. Meanwhile, noble method for data hiding using Steganography Discrete Wavelet Transformation: DWT and Cryptography Triple Data Encryption Standard: DES is proposed and well reported [14].

In this paper, MRA based data hiding method with random scanning of the insert secrete image is proposed.

A. Wavelet Multi-Resolution Analysis

The wavelet transforms of a given discrete scalar signal $f = (f_1, f_2, ..., f_n)^T$ is described as $C_n f$ by a square matrix C_n composed of a sequence $\{p_k\}$ and a sequence $\{q_k\}$. pi is for low-frequency components, coefficient q_i is for high-frequency components, C_n divides f into low-frequency components and high-frequency components, and is composed of sequences $\{a_k\}$ and sequences $\{b_k\}$ The square matrix H_n is expressed as follows,

$$H_n C_n = I_n \tag{1}$$

where l_n is an identity matrix. And then, the following equation is defined.

$$H_n = C_n^T \tag{2}$$

When Eq. (2), the biorthogonal wavelet transform is an orthogonal wavelet transform, that is, the orthogonal wavelet transform is a kind of biorthogonal wavelet transform.

B. ２D(Two Dimensional) Discrete Wavelet Transformation

For 2D image signals, this process is performed horizontally and vertically one level at a time. Fig. 1 shows the band components when two-dimensional DWT is performed twice. In the figure, L indicates a low frequency component, and H indicates a high frequency component. The image is decomposed into four bands (LL, LH, HL, HH) by the first two-dimensional DWT, and the lowest band component (LL) is further divided into four bands (LLLL, LLLH, LLHL, LLHH).

Following are the related research works: Data hiding method replacing LSB of hidden portion for secrete image with Run-Length coded image is proposed [15]. Meanwhile, Data hiding method with Principal Component Analysis: PCA and image coordinate conversion is proposed for improvement of invisibility of the secret key image [16].

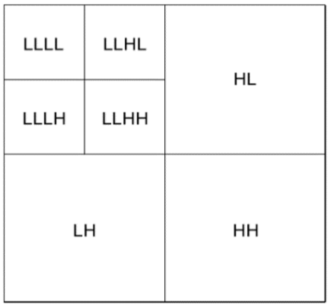


Fig. 1. Band Components after the 2D DWT.

III. Proposed Method

When the DWT is applied to n time series data in one stage, it can be decomposed into *n* / 2 high frequency components and *n* / 2 low frequency components. By further subjecting the *n* / 2 low frequency components to a one-stage DWT, the *n* / 4 low frequency components and the *n* / 4 high frequency components can be decomposed. By repeating this, the number of data becomes 1 or 2. This is shown in Fig. 2. This is called the Laplacian Pyramid.

In this case, the size of the image in each stage is halved both vertically and horizontally by the DWT. A Dyadic Wavelet that does not downsize is also proposed. Also, in this case, a certain low-frequency component image is decomposed into four, but a Multi Wavelet that decomposes this into 16 images is also proposed. It is reported that they are effective for noise removal and data compression, respectively. These Wavelets and many others are published in the Special Issue on Visualization Information Society of Reference [9], so please refer to them.

The original time series data can be completely restored by applying the inverse wavelet transform (Inverse DWT: IDWT) for the number of transform stages using the high frequency components and the low frequency components of each stage generated by this decomposition. Of the decomposed frequency component data, the fact that "the human eye has a low resolution of high frequency components" is used to embed the secret data into one of the high frequency components and reconstruct it with the secret data embedded. When attempting to restore to the original image level, the secret data is embedded in the high frequency component, so that data like the original image is reconstructed in a state where it is difficult to see.

The data generated in this way is called distribution data (content). The distribution contents are contents that are open to the public and can be obtained by anyone. Therefore, they are exposed to the risk of plagiarism. This distributed content is almost the same as the original content but differs from the original content in that the secret data is embedded in the high frequency component. Even if the distributed content is stolen, the copyright holder can claim the copyright by extracting and showing the secret content (e.g., copyright) embedded in the high frequency component. Fig. 3 shows a series of processing flow from embedding secret data for asserting copyright in such copyrighted content, generating a distribution image, and restoring the secret data and the original content from the distribution image.

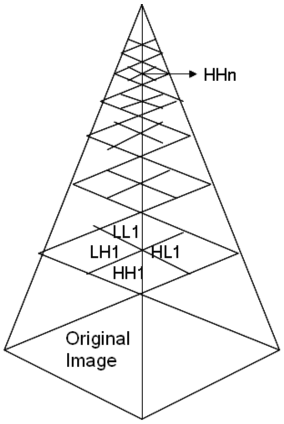


Fig. 2. Laplacian Pyramid.

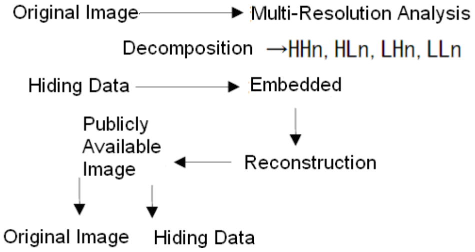
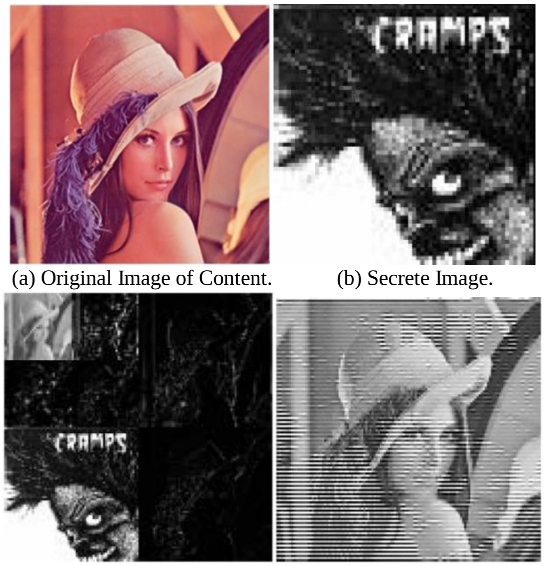


Fig. 3. Process Flow of Data Hiding based on Multi Resolution Analysis (MRA).

The visibility of the secret content in the distributed content is greatly influenced by the confidential data to be embedded (frequency component of the content) and the place to be embedded (frequency component). Therefore, as shown in Fig. 4, the location where the secret content is embedded is a very important factor in considering the visibility of the secret image content in the distribution image content. In this case, LH1 of MRA (Fig. 4 (c)) of the original image (Fig. 4 (a)) is replaced to the secrete image of “CRAMPS” (Fig. 4 (b)). Then the reconstructed image is derived by Inverse DWT.


(a) Original Image of Content. (b) Secrete Image.

(c) Data Hidden Image. (d) Reconstructed Image.

Fig. 4. LH1 of MRA is replaced to the Hiding Image Content of “CRAMPS”.

The scanning method of the secret image data can be changed from the normal line sequential scanning to the random scanning to improve the visibility of the secret image in the distribution image. The author proposes a method to obtain a distribution image in random scan (rand) by using the support length (dbn) of Daubechies basis function used in MRA and the initial value (rand50 / 5000) of uniform random numbers used in random scan as parameters.

IV. Experiment

A. Preliminary Results

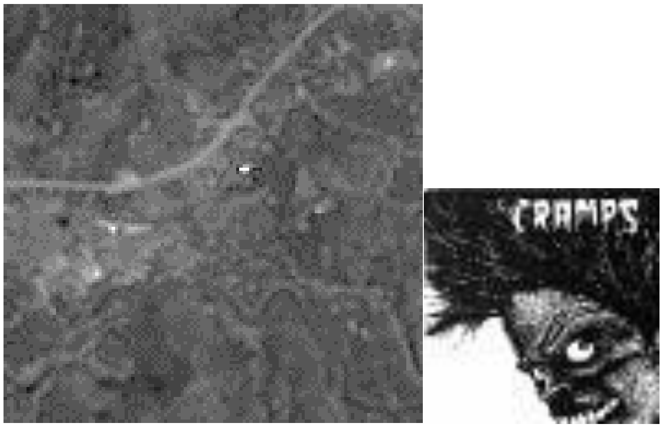
A method has also been proposed to improve the visibility of the secret data in the distribution image by scanning it again before embedding the secret data. It is premised that the rules are shared. In contrast to normal image data that is line-sequential scanning, a secret image is converted to random scanning that determines the scanning order by, for example, generating a random number. It converts and stores 2D spatial data into a dictionary array (1D data).

The conversion of this scanning method can be performed by a permutation conversion matrix. At this time, if the random number generation rule information is shared between the sending and receiving parties, the inverse matrix of the permutation conversion matrix is applied after extracting secret data in random scanning. By doing so, the reverse conversion of the scanning method becomes possible and the secret data in the line sequential scanning can be reproduced. Since it is difficult for a third party to obtain the information of the scanning method when embedding the secret data, it is difficult to obtain the secret data. The confidentiality of information is also improved.

The experimental results are as follows: The used data is the original image shown in Fig. 5 (a) (The band 3 red area in which the Thematic Mapper: TM sensor mounted on the Landsat satellite observed near the Yamato interchange of Nagasaki Highway near Saga city) The original image is composed of 128x128 pixels, the secret data is composed of 64x64 pixels, and the quantized bits are 8 in each case. Landsat / TM is a 30m spatial resolution multi-spectral scanner with spectral bands of five bands from blue to near infrared and one band in thermal infrared.

The wavelet division was applied to the original image by one stage, and the secret data was embedded in HH1. At that time, the secret data was embedded in HH1 with line sequential scanning and the random number generation method of Merthenne Twister was used. Compared with the method of generating uniform random numbers, scanning again based on that, and embedding in HH1, the distribution image obtained by reconstructing using the embedded image is almost the same as the original image.

When the author tries to reconstruct the HH1 using this method, the secret data can be restored as shown in Fig. 7 (a), (b) and (c) for line sequential scanning (raster scan), random scanning and Hilbert scan, respectively. An example of Hilbert scan is shown,



(a) Original. (b) Hiding Data.

Fig. 5. Original Image of Landsat-5/TM Band 3 Data of Saga City and Hiding Data.

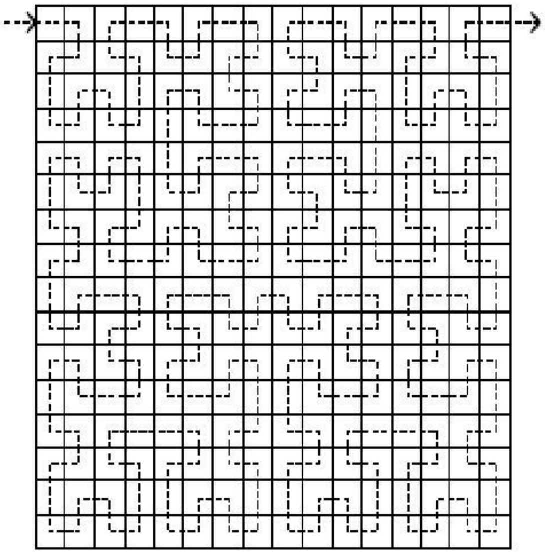


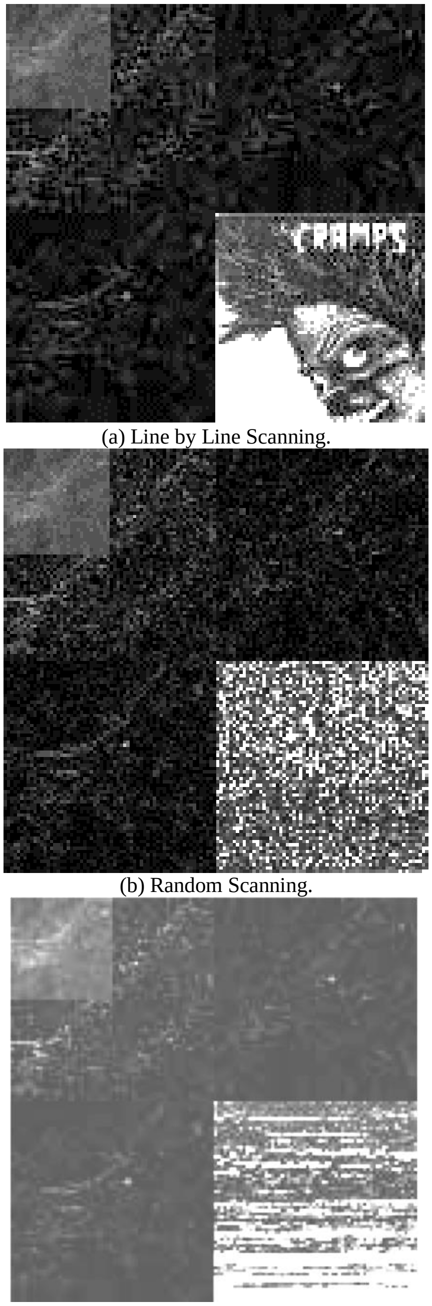
Fig. 6. Hilbert Scan.

The Hilbert curve is one of the simplest curves which pass through all points in a space (Fig. 6). Many researchers have worked on this curve from the engineering point of view, such as for an expression of two-dimensional patterns, for data compression in an image or in color space, for pseudo color image displays, etc.

If the secret data is embedded in the original image as it is, the secret data itself can be restored in HH1, but in the case of random scanning, the secret data cannot be restored without knowing the control parameters for random number generation.

On the other hand, Fig. 8 (a) shows the secrete image of "CRAMPS" derived from the Hilbert scanning. Also, Fig. 8 (b) and (c) are the randomly scanned secrete image and the raster scanned secrete image, respectively. Not only random scan, but also Hilbert scan can be used for improvement of invisibility of the hidden secrete image from the reconstructed image. Fig. 8 (d) shows the reconstructed image derived from the decomposed image embedding the secrete image with Hilbert scanning. Also, Figs. 8 (e) and (f) show the reconstructed images derived from the decomposed image embedding the secrete image with raster scanning and random scanning, respectively.

These are images obtained by scanning the key image by raster, Hilbert, and random scanning, replacing the original image with HH1 after MRA, and performing wavelet transform on the reconstructed image. As is apparent from these, in the case of raster scanning, the key image itself appears and there is no confidentiality. On the other hand, in Hilbert scanning, only horizontal stripe noise appears, and in random scanning, only random noise appears, so it is difficult to visually recognize the key image.

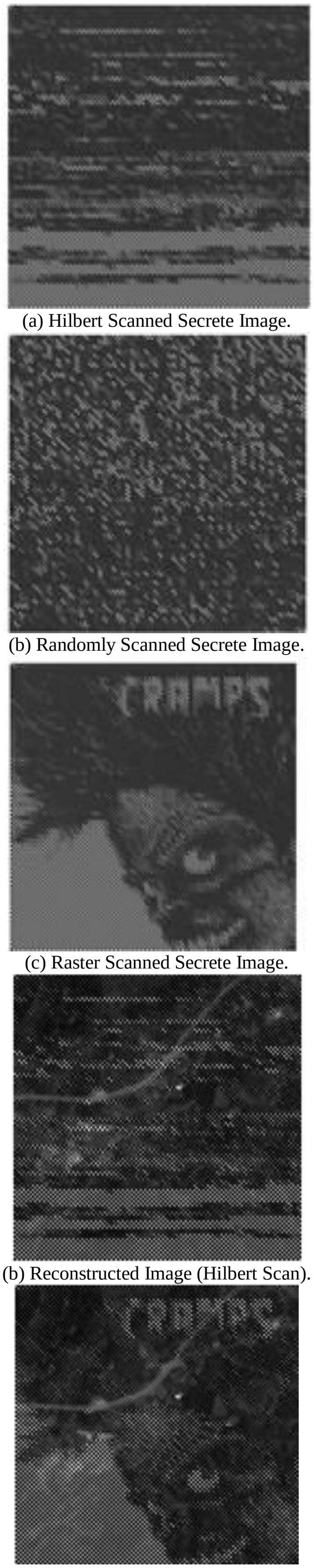


(a) Line by Line Scanning.

(b) Random Scanning.

(c) Hilbert Scan.

Fig. 7. Reconstructed Hiding Data from Publicly Available Image Content Derived from the MRA based Methods with the Different Scanning Schemes.



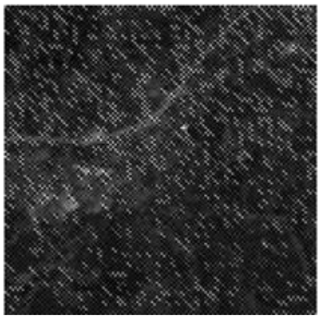
(a) Hilbert Scanned Secrete Image.

(b) Randomly Scanned Secrete Image.

(c) Raster Scanned Secrete Image.

(b) Reconstructed Image (Hilbert Scan).

(c) Reconstructed Image Raster Scan.



(d) Reconstructed Image (Random Scan).

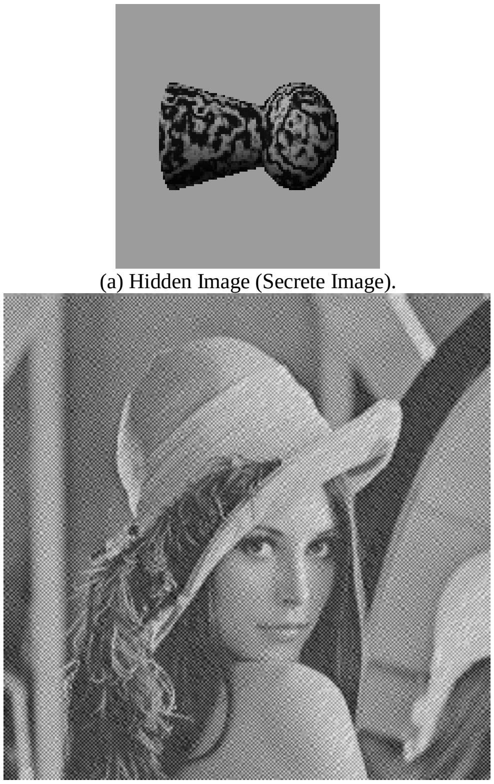
Fig. 8. Secrete Image of "CRAMPS" Derived from the Hilbert Scanning and the Reconstructed Image Extracted from the Decomposed Image Derived from the Decomposed Image Embedding the Secrete Image with Hilbert, Random and Raster Scanning.

Furthermore, by understanding the parameters related to the scanning order generation method in random scanning and Hilbert scanning only between the sending and receiving parties, only the parties can know the key image, and the confidentiality and confidentiality can be improved.

B. Experimental Results

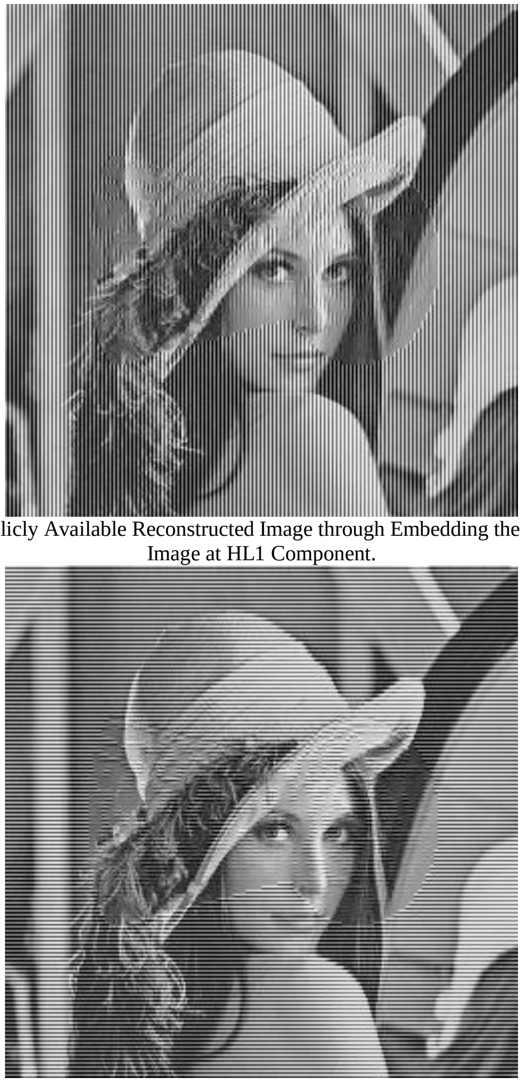
Fig. 9(a) is an example of a secret image. Fig. 9(b), (c), and (d) show the distribution images when this is inserted into each of HH1, HL1, and LH1 of the above-mentioned original image (Lena).

As is clear from Fig. 9, the secret image data can be visually recognized on the distribution image. As shown in Fig. 10, this secret image data is changed from line sequential scanning to random scanning to improve visibility.



(a) Hidden Image (Secrete Image).

(b) Publicly Available Reconstructed Image through Embedding the Hiding Image at HH1 Component.



(c) Publicly Available Reconstructed Image through Embedding the Hiding Image at HL1 Component.

(d) Publicly Available Reconstructed Image through Embedding the Hiding Image at LH1 Component.

Fig. 9. Hidden Image and Publicly Available Reconstructed Images through Embedding the Hiding Image at HH1, HL1 and LH1 Components.

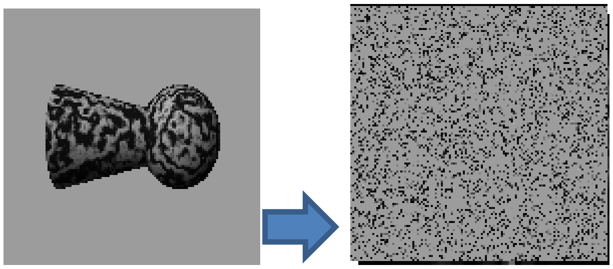
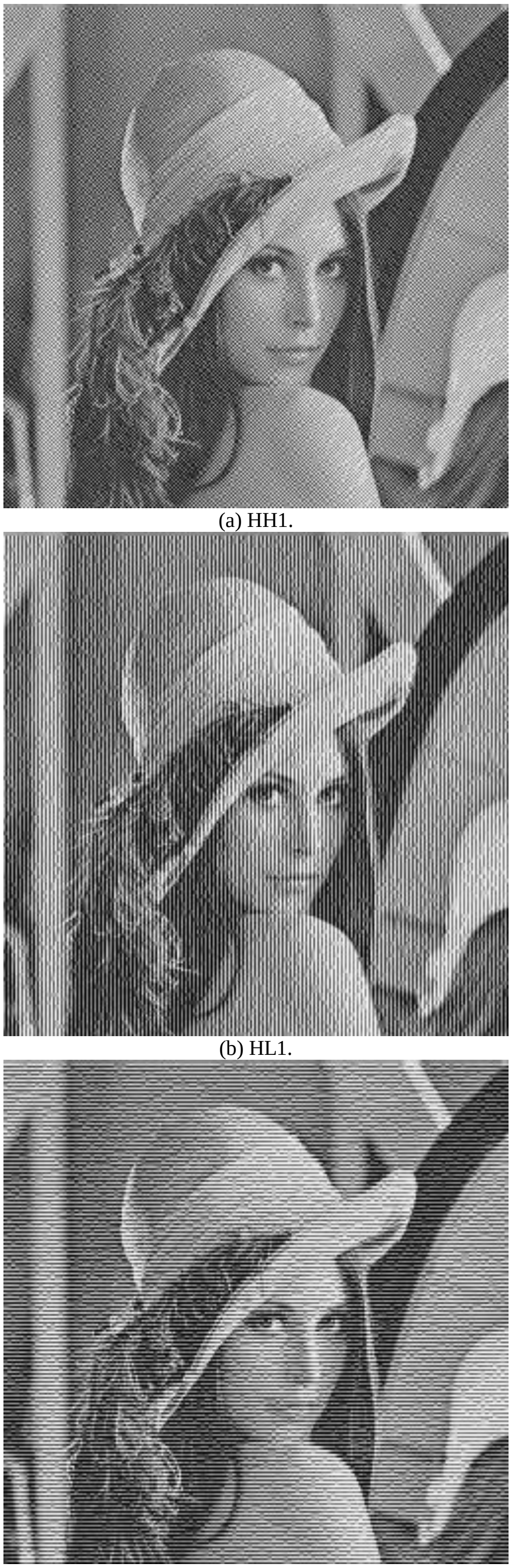


Fig. 10. Scanning Scheme Conversion from the Line-by-line to Random.

Fig. 11(a), (b), and (c) show the results of improving the visibility of the secret image data on the distribution image due to the change of the scanning method. Fig. 10(a), (b), and (c) show the circulation images when they are inserted into HH1, HL1, and LH1 of the original image (Lena).



(a) HH1.

(b) HL1.

(c) LH1.

Fig. 11. Hidden Image and Publicly Available Reconstructed Images through Embedding the Hiding Image at HH1, HL1 and LH1 Components after the Scanning Scheme Conversion for Hidden Image from Line-by-line to Random.

Distribution image and original image in line sequential scanning (Normal) and random scanning (rand) using the support length (dbn) of the Daubechies basis function used for MRA and the initial value of uniform random numbers (rand50 / 5000) used for random scanning as parameters Table I shows a comparison of the Root Mean Square Difference (RMSD) between the original and the publicly available reconstructed images and the results show that random scanning is better than line-sequential scanning, and that longer support length is better than short support length. It can be seen that the initial value of the random number is not so affected.

Since the visibility of the secret image data on the distribution image does not depend on the initial value of the random number used, if this initial value is hidden by steganography between the sending and receiving parties, only the party who knows this initial value will have the secret value. Image data can then be restored.

TABLE I. COMPARISONS OF ROOT MEAN SQUARE DIFFERENCE (RMSD) BETWEEN THE ORIGINAL AND THE PUBLICLY AVAILABLE RECONSTRUCTED IMAGES THROUGH DATA HIDING BASED ON MRA WITH EMBEDDING THE HIDING IMAGE TO HL1, HH1 AND LH1 AND WITH SCANNING SCHEME CONVERSION FROM LINE-BY-LINE TO RANDOM

Scanning Method	HL1	HH1	LH1
Normal(db2)	69.594	69.137	69.183
Normal(db4)	69.397	69.089	69.058
Normal(db8)	69.518	69.069	69.056
rand50(db2)	68.790	68.297	68.340
rand50(db4)	68.609	68.215	68.247
rand50(db8)	68.568	68.135	68.123
rand5000(db2)	68.856	68.357	68.427
rand5000(db4)	68.665	68.291	68.316
rand5000(db8)	68.633	68.182	68.202

V. CONCLUSION

The author has introduced a method that improves the confidentiality by applying principal component transformation and oblique coordinate transformation as preprocessing for data hiding based on wavelet multiresolution analysis. The author investigated the confidentiality when a third party attempts to extract secret data from only the data for distribution.

The method introduced in this paper allows only the author who knows the characteristics of the original multispectral image to recover the secret data, i.e., when the information of the original image needs to be protected. The author also showed how to convert the scanning method of the secret data from line-sequential to random scanning, which leads to the improvement of the confidentiality of the secret data and the visibility difficulty in the distribution image. By sharing the equation parameters only between the sending and receiving parties, more confidential data hiding can be realized.

In this paper, the Daubechies basis function is adopted as the wavelet, but the secret data can be restored by using the biorthogonal wavelet, and the secret data can be protected by hiding what is adopted as the biorthogonal wavelet.

VI. FUTURE RESEARCH WORKS

In the future, the author will compare the proposed method with conventional data hiding methods such as steganography method.

ACKNOWLEDGMENT

The author would like to thank Dr. Kaname Seto of former student of Saga University and Dr. Leland M. Jameson of Naval Research Laboratory for their contribution in this study. The author, also, would like to thank Professor Dr. Hiroshi Okumura and Professor Dr. Osamu Fukuda for their valuable discussions.

REFERENCES

[1] Tirkel, A., et al., "Electronic Water Mark" Proceedings DICTA 1993, 666-672, 1993.

[2] Fabien A. P. Peticolas, Ross J. Anderson, Markus G. Kuhn : "Attacks on copyright marking systems", David Aucsmith(ed), Information Hiding, Proceedings, LNCS 1525, Springer-Verlag, ISBN 3-540-65386-4, pp.219-239. 1998.

[3] H. Keith Melton : "The Ultimate Spybook", Dorling Kindersley Limited, London, 1996.

[4] Kohei Arai, Kaname Seto, Data Hiding Based on Wavelet Multiresolution Analysis, Journal of Visual Information Society, Vol.22, Suppl.No.1, 229-232, 2002.

[5] Kohei Arai, Kaname Seto, Data Hiding Based on Multiresolution Analysis Using Information Bias by Eigenvalue Expansion, Journal of Visual Information Society, Vol.23, No.8, pp.72-79, 2003.

[6] Kohei Arai, Patent Application No .: 2004-29933, Digital Watermark Insertion / Extraction Device and Method.

[7] Kohei Arai, PCT application number: PCT / JP2005 / 13512, coordinate transformation method, data compression and data hiding method using the same, and their devices, 2005.

[8] Yao Qiuming, Xuan Guorong, Yang Chengyun, Shi Yunquin, Lossless Data Hiding Using Histogram Shifting Method Based on Integer Wavelets, Proceedings of the IWDW 2006: Digital Watermarking pp 323-332 | Cite as, 2006.

[9] Kohei Arai, Basic Theory of Wavelet Analysis, Morikita Publishing (November 2000).

[10] Kohei Arai, Leland Jameson, How to use earth observation satellite data by wavelet analysis, Morikita Publishing (July 2001).

[11] Kohei Arai, Self-study wavelet analysis, published by Modern Science Co., Ltd. (June 2006).

[12] 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.

[13] Kohei Arai and Yuji Yamada, Improvement of secret image invisibility in circulation image with Dyadic wavelet based data hiding with run-length coding, International Journal of Advanced Computer Science and Applications, 2, 7, 33-40, 2011.

[14] Cahya Rahmed Kohei Arai, Arief Prasetyo, Noriza Arigki, Noble Method for Data Hiding using Steganography Discrete Wavelet Transformation and Cryptography Triple Data Encryption Standard: DES, IJACSA, 9, 11, 261-266, 2018.

[15] Kohei Arai, Data Hiding Method Replacing LSB of Hidden Portion for Secrete Image with Run-Length Coded Image, International Journal of Advanced Research on Artificial Intelligence, 5, 12, 8-16, 2016.

[16] Kohei Arai, Data Hiding Method with Principal Component Analysis and Image Coordinate Conversion, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 12, No. 8, 25-30, 2021.

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 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 councilor 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 a Science Council of Japan Special Member since 2012. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Science Commission "A" of ICSU/COSPAR since 2008 then he is now award committee member of ICSU/COSPAR. He wrote 55 books and published 620 journal papers as well as 450 conference papers. He received 66 of awards including ICSU/COSPAR Vikram Sarabhai Medal in 2016, and Science award of Ministry of Mister of Education of Japan in 2015. He is now Editor-in-Chief of IJACSA and IJISA. http://teagis.ip.is.saga-u.ac.jp/index.html.