A New Fragile Watermarking Scheme Based on Wavelet Edge Feature

Krisda Khankasikam

***Abstract*—An image watermarking is the process of authenticating a digital image by embedding a watermark into it and protecting the image from copyright violation. This paper proposes a new fragile watermarking scheme developed in a wavelet domain based on the discrete wavelet transform and Arnold scrambling algorithm. The original watermark image is transformed into wavelet domain by applying discrete wavelet transform, subsequently the high frequency coefficients which in form of an edge feature image of wavelet transformed image is transformed into binary watermark image by using thresholding method. The binary watermark image is embedded into host image by modifying Arnold scrambling algorithm. The proposed method encompasses three phases including watermark generation phase, watermark embedding phase and tamper detection and localization phase. Experimental results show that the proposed method has satisfactory protection ability and can detect and locate various malicious tampering efficiently. The invisibleness and robustness of the propose method is evaluated by using well known indices including peak signal to noise ratio index and normalized correlation index.**

***Index Terms*—Arnold scrambling, discrete wavelet transform, edge feature image, fragile watermarking scheme, tamper detection.**

I. Introduction

Due to the rapid improvements of modern communication and Internet technology, digital media can be easily transmitted and modified by using image processing tools whether it is malicious or not. Generally, a digital signature scheme, which is adopted in modern cryptography, can be used to detect if an image has been modified. However, this scheme is not able to detect the tamper region of image; moreover a digital signature scheme implies external information additional for each image to be memorized. Digital watermarking solves these issues by applying an authentication key encapsulated directly in the image and identifying directly the tampered zone. Watermarking is regarded as one of effective approaches to resolve images copyright protection and authentication [1]. In general, watermarking technology can be classified as robust watermarking scheme for copyright protection and fragile watermarking scheme for integrity verification [2]. Many researchers have developed fragile watermarking [3]-[8], which focus on image authentication. Fragile watermarking schemes can be typically divided into semi-fragile and completely fragile schemes. The major difference between these two schemes is the integrity criteria [9]. Semi-fragile watermarking [2], [3], [5], [6], [10], [11] called soft authentication [9], uses relatively relaxed integrity criteria. Then, some invisible modifications are allowed such as Joint Photographic Experts Group (JPEG) compression. Semi-fragile watermarking schemes are useful when protected image must be compressed at different rates to satisfy transmission bandwidth. Complete fragile watermarking schemes [4], [7], [8], [12] called hard authentication [2], offer greater protection and integrity than soft authentication. These schemes do not allow any modification or tampering of a protected image. In addition to detecting whether a protected image has been modified, hard authentication schemes must be capable of locating tampered areas.

In view of the above facts, this paper proposes a new fragile watermarking scheme developed in a wavelet domain based on the discrete wavelet transform, defined as DWT, and Arnold scrambling algorithm. The DWT is applied to the original watermark image to obtain a watermark image in wavelet domain, and then, a high frequencies coefficient is adopted. The watermark is embedded into host image by modifying the Arnold scrambling algorithm. The performance of the proposed scheme is evaluated on four test images namely Sail Boat, Lena, Cameraman, and Barbara. Normalized correlation and peak signal to noise ratio are the performance metrics employed for performance evaluation of the proposed method.

The remainder of this paper is organized as follows. In Section II, DWT and Arnold scrambling are briefly described. The proposed fragile watermarking scheme is described in Section III. Section IV gives experimental results to demonstrate the proposed scheme effectively detects and locates a tampered area. Finally conclusions are stated in Section V.

II. Related Works

This section gives the brief descriptions of related works consisting of the DWT, which is used to generate watermark image, and the Arnold scrambling, which is used to embed watermark to original image.

A. Discrete Wavelet Transform

The 2-dimension DWT, which is a linear transform, is commonly used tool in image processing. It decomposes the image into low and high frequency coefficients. The low frequency coefficients give approximation information of image and the high frequency coefficients give detailed

Manuscript received April 13, 2015; revised June 25, 2015. This work was supported in part by the Department of Applied Science, Faculty of Science and Technology, Nakhon Sawan Rajabhat University, Thailand.

K. Khankasikam is with the Department of Applied Science, Faculty of Science and Technology, Nakhon Sawan Rajabhat University, Muang Nakhon Sawan, 60000, Thailand (tel.: +66-81688-0066; e-mail: KrisdaK@gmail.com).

DOI: 10.7763/IJFCC.2015.V4.400

information, especially edge features of image. The edge features provide the structural properties of objects in an image with reduced amount of information [13]. It aids to increase the invisibleness when watermark with less amount of information on the host image. Also these edge features controls the attacks caused by noise, edge strips and acuity.

B. Arnold Scrambling

The Arnold scrambling is usually used in watermarking and encryption techniques. It is used as pre-processing step to embed the watermark, which reduce the spatial relationship between the pixel and makes the image as meaningless one [14], [15]. Let x and y be the coordinates of the original space, x' and y' be the coordinates after iterative computation scrambling and N be the size of image. The 2-dimensional Arnold scrambling can be defined by using following formula.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \bmod N, x, y \in \{0,1,2,...,N-1 \quad (1)$$

III. The Proposed Method

The proposed fragile watermarking method encompasses three phases including watermark generation phase, watermark embedding phase and tamper detection and localization phase. An overview of the proposed method is depicted in Fig. 1. The details of each phase are described following.

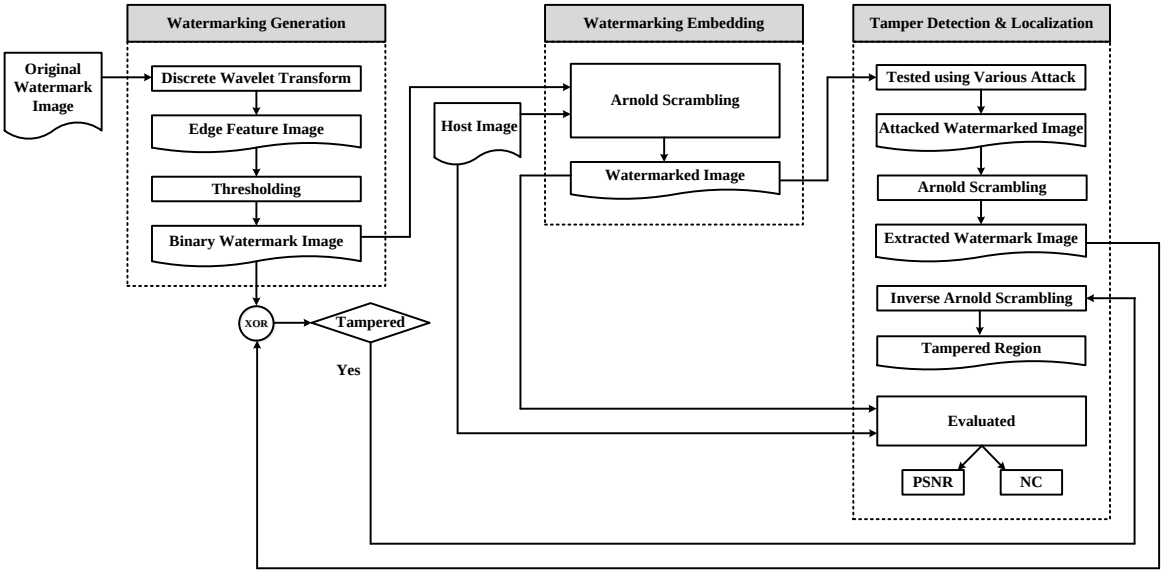


Fig. 1. Overview of the proposed fragile watermarking method.

A. Watermark Generation Phase

The edges are the major features which furnish the information about image content and the wavelet transform is marvelous way to detect the edge features, as, it increases the reliability of edge detection even when it is analyzed at different scales. Hence, the DWT is applied on the original watermark image to acquire the edge image. It decomposes the image into low and high frequency coefficients. These high frequency coefficients contain edge features and it is converted to binary version by using thresholding method to make it feasible to embed on the significant bit. The value of threshold is adaptive one and is computed by summing the mean and standard deviation values of edge features. The edge feature image, defined as F, of watermark image to be embedded on the original host image is derived by using edge image, defined as E, and threshold T which is given by using the following formula.

$$E(i,j) = \begin{cases} 1 \ if \ F(i,j) \geq T \\ 0 \ if \ F(i,j) < T \end{cases} \quad (2)$$

B. Watermark Embedding Phase

The watermarking method begins, only after the generation of watermark edge image was accomplished. Now the resolution of edge image is reduced to twice as the host image. Hence the host image is divided into 2×2 non-overlapping blocks and watermark is embedded on the least significant bit of each block's first pixel element. In fragile watermarking method, watermark must be more sensitive and secure. Therefore, an Arnold scrambling is employed on the host image as a preliminary process. After embedding the edge image, watermarked image is constructed by the inverse Arnold scrambling.

In order to extract the watermark edge image embedded on host image, the Arnold scrambling is applied on watermarked image for number of iterations which is equal to the number of iterations done to embed the watermark edge image. Then it is divided into 2×2 non-overlapping blocks and watermark is extracted from the least significant bit of first element of each block.

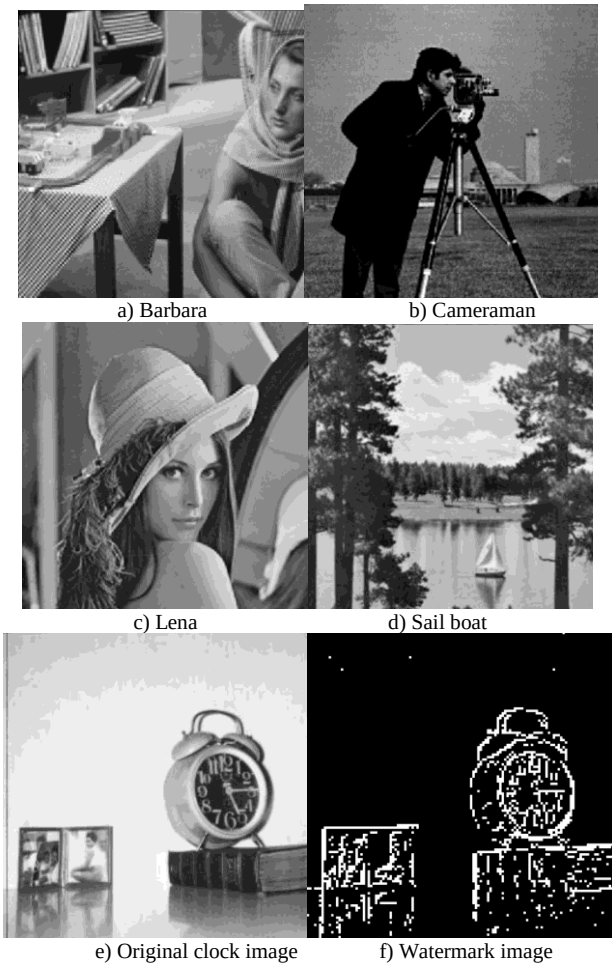
C. Tamper Detection and Localization Phase

The tamper detection process begins after the extraction of watermark edge image. The original and extracted watermark edge images are subjected to XOR operation and it detects the difference among them and the image is decided as tampered or trustworthy based on the difference. Once the image is detected as tampered, the tampered region is localized by

using the inverse Arnold scrambling.

IV. The Experiments

This section is dedicated to the performance evaluation of the proposed watermarking scheme. In order to investigate the effectiveness of the proposed method, the experiments are performed on a computer with Intel Core i5 CPU M480@2.67 GHz and 2G DRAM and with the MATLAB 2011a. Four images, selected from standard image dataset, are used as image test set (Fig. 2). The first four Grayscale image with 512×512 pixels have served as host image while the last two images are used as original watermark and binary watermark image. The experimental results are evaluated by using well known metrics including peak signal to noise ratio and normalized correlation. The description of the experiment is fully described in this section.



a) Barbara b) Cameraman

c) Lena d) Sail boat

e) Original clock image f) Watermark image

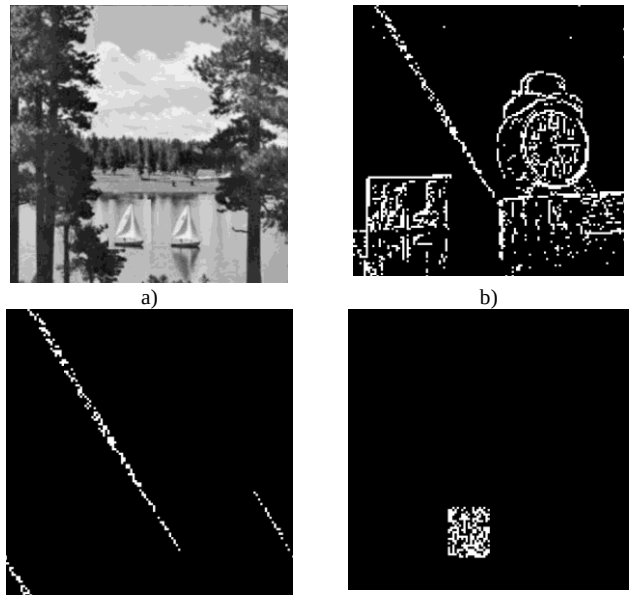
Fig. 2. Tested images and watermark images.

A. Tamper Detection and Localization Results

In this subsection, the watermarked image is examined for several attacks to see how it detects and localizes the tampers through various attacking method including copy and paste attack, text addition attack, image splicing attack and object removal attack. The details of each attack are described in following subsections.

1) Copy and paste attack

In the watermarked image, there is only one sailboat which is shown in the Fig. 2a). In order to do a copy and paste attack experiment, the sail boat is copied and pasted it near to the original sailboat, which is shown in the Fig. 3a). The watermark image, as shown in Fig. 3b), shows some noise. The exclusive-OR (XOR) operation between original watermark image and extracted watermark image is shown in Fig. 3c). This indicates that watermarked image has been subjected to some kind of tampering and Fig. 3d) shows the localized tampered region.



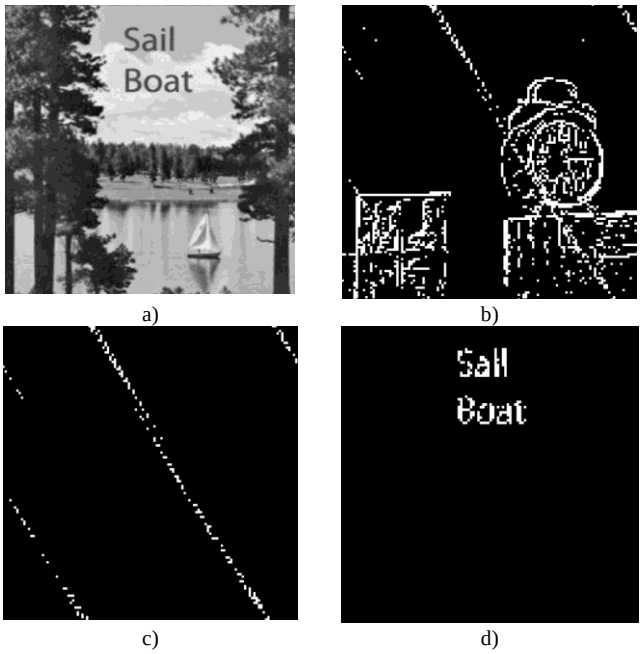
a) b)

c) d)

Fig. 3. Result of copy and paste attack experiment.

2) Text addition attack

To perform this experiment, the text "Sail Boat" is merged to the watermarked image, as shown in Fig. 4a). The watermark image, as shown in Fig. 4b), shows some noise. The XOR operation between original watermark image and extracted watermark image is shown in Fig. 4c). This indicates that watermarked image has been subjected to some kind of tampering and Fig. 4d) shows the localized tampered region.



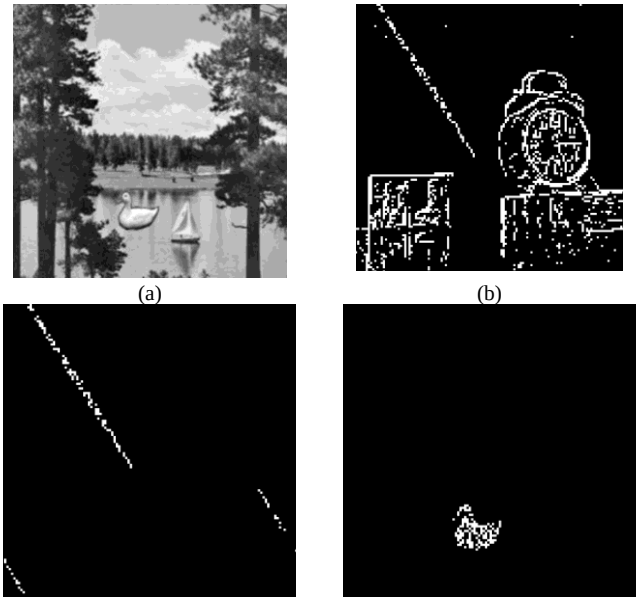
a) b)

c) d)

Fig. 4. Result of text addition attack experiment.

3) Image splicing attack

To test the performance of the proposed method, the object swan is combined in the watermarked image, as shown in Fig. 5a). The watermark image, as shown in Fig. 5b), shows some noise. The XOR operation between original watermark image and extracted watermark image is shown in Fig. 5c). This indicates that watermarked image has been subjected to some kind of tampering and Fig. 5d) shows the localized tampered region.

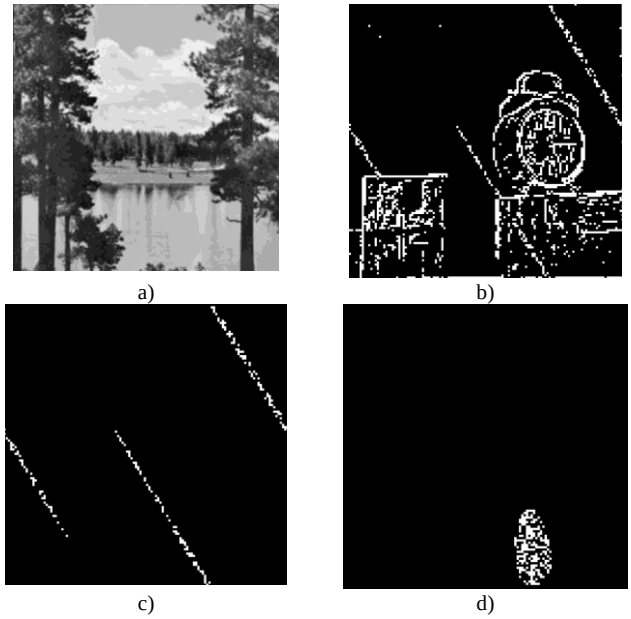


(a) (b) (c) (d)

Fig. 5. Result of image splicing attack experiment.

4) Object removal attack

To carry out this experiment, the object sailboat in the watermarked image is removed, as shown in Fig. 6a). The watermark image, as shown in Fig. 6b), shows some noise. The XOR operation between original watermark image and extracted watermark image is shown in Fig. 6c). This indicates that watermarked image has been subjected to some kind of tampering and Fig. 6d) shows the localized tampered region.



a) b) c) d)

Fig. 6. Result of object removal attack experiment.

B. The Experimental Results

To ensure the information safety, the embedded watermark image should be invisible to human naked eyes; the watermark invisibility is one of the valuable indices to examine the quality of the proposed method. Then the peak signal to noise ratio (*PSNR*) [16] is adopted as a quantitative index to evaluate the effectiveness of the proposed method. The mathematics equation of PSNR is given in (3).

$$PSNR(I,I') = 10\log_{10}\left(\frac{(I_{MAX})^2}{\frac{1}{n\times n}\sum_{i=1}^{n}\sum_{j=1}^{n}(I_{i,j}-I'_{i,j})^2}\right) \quad (3)$$

The more significant index is the watermark robustness which is used to examine the stabilization of a watermark scheme associated with the transform when the watermark is extracted from the watermarked image destroyed by various attack. To measure the stabilization, the normalized correlation (*NC*) [16] is used as the quantification index which and defined as

$$NC(I,I') = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\overline{I_{i,j}\oplus I'_{i,j}}}{n\times n} \quad (4)$$

where *I* and *I*' stand for the original and the processed image, I_{MAX} is the maximum possible intensity value of the image *I*, for an 8-bit per pixel representation I_{MAX} is 255, subscripts *i* and *j* denote the location of the pixel value in the respective image, $\oplus$ denotes the XOR operation and n is the height or width of the square image.

The performance of proposed method also validated by the various images which are taken from standard image processing dataset and its performances are given in Table I.

TABLE I: PERFORMANCE EVALUATION OF THE PROPOSED METHOD

Image	Invisibleness		Robustness (Image is not tampered)	
	PSNR	NC	PSNR	NC
Barbara	57.2131	0.9994	64.2172	0.9989
Cameraman	57.1370	0.9993	64.2026	0.9989
Lena	56.6935	0.9991	64.1892	0.9985
Sail Boat	56.6433	0.9979	64.1851	0.9983
(Average)	(56.9217)	(0.9989)	(64.1985)	(0.9987)

V. CONCLUSION

In this paper, inspired by the methods of DWT and Arnold scrambling, a new fragile watermarking based on wavelet edge feature is proposed. The strength of this scheme against image manipulation attacks is tested on a set of four images in standard dataset and four image manipulation attacks. The experiment is implemented by using MATLAB. Experimental results show that the proposed method retain good watermarked image quality with average PSNR values greater than 56 dB. The obtained results are good in term of accuracy for tamper detection. In future research, more effort will be focused on stereo image authentication to address

issue of copyright protection.

REFERENCES

[1] M. U. Celik, G. Sharmar, and A. M. Tekalp, "Hierarchical watermarking for secure image authentication with localization," *IEEE Transactions on Image Processing*, vol. 11, no. 6, pp. 585–594, 2002.

[2] C. S. Chan and C. C. Chang, "An efficient image authentication method based on hamming code," *Pattern Recognition*, vol. 40, no. 2, pp. 681–690, 2007.

[3] E. C. Chang, M. S. Kankanhalli, X. Guan, Z. Y. Huang, and Y. H. Wu, "Robust image authentication using content based compression," *ACM Multimedia System Journal*, vol. 9, no. 2, pp. 121–130, 2003.

[4] J. Fridrich, M. Goljan, and A. C. Baldoza, "New fragile authentication watermark for images," in *Proc. IEEE International Conference on Image Processing*, 2000, pp. 446–449.

[5] C. Y. Lin, and S. F. Chang, "A robust image authentication method distinguishing JPEG compression from malicious manipulation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 2, pp. 153–168, 2001.

[6] A. H. Paquet, R. K. Ward, and I. Pitas, "Wavelet packets-based digital watermarking for image verification and authentication," *Signal Processing*, vol. 183, no. 10, pp. 2117–2132, 2003.

[7] P. Wong and N. Memon, "Secret and public key authentication watermarking schemes that resist vector quantization attack," in *Proc. SPIE on Security and Watermarking of Multimedia Contents*, 2000, pp. 417–427.

[8] M. Yeung and F. Mintzer, "An invisible watermarking technique for image verification," in *Proc. IEEE International Conference on Image Processing*, 1997, pp. 680–683.

[9] S. H. Liu, H. X. Yao, W. Gao, and Y. L. Liu, "An image fragile watermark scheme based on chaotic image pattern and pixel-pairs," *Applied Mathematics and Computation*, vol. 185, no. 2, pp. 869–882, 2007.

[10] C. Y. Lin and S. F. Chang, "Issues and solutions for authenticating MPEG video," in *Proc. SPIE on Security and Watermarking for Multimedia Contents*, 1999, pp. 54–65.

[11] Z. Peng and W. Liu, "Color image authentication based on spatiotemporal chaos and SVD," *Chaos, Solitons and Fractals*, vol. 36, no. 4, pp. 946–952, 2008.

[12] J. Fridrich, "Security of fragile authentication watermarks with localization," in *Proc. SPIE Conference on Security and Watermarking of Multimedia Contents*, 2002, pp. 691–700.

[13] D. Zhang, Z. Pan, and H. Li, "A contour-based semi-fragile image watermarking algorithm in DWT domain," in *Proc. International Workshop on Education Technology and Computer Science*, 2010, pp. 228-231.

[14] D. Vaishnavi, and T. S. Subashini, "An image watermarking scheme resilient to geometric distortions," *Power Electronics and Renewable Energy Systems*, vol. 326, pp. 1225-1233, 2015.

[15] M. Sui, and J. Li, "The medical volume data watermarking using Arnold scrambling and 3D-DWT," in *Proc. International Conference on Mechatronic Science, Electric Engineering and Computer*, 2013, pp. 1120-1124.

[16] X. Wu, and W. Sun, "Robust copyright protection scheme for digital images using overlapping DCT and SVD," *Applied Soft Computing*, vol. 13, pp. 1170-1182, 2013.

Krisda Khankasikam received the bachelor of engineering degree in computer engineering from Naresuan University, Thailand, in 2002. Later, he received the master of engineering degree in computer engineering from King Mongkut's University of Technology Thonburi, Thailand, in 2005. He received his Ph.D. in knowledge management from Chiang Mai University, Thailand, in 2010.

During the academic years of 2005-2012, he joined the Faculty of Information and Communication Technology at Naresuan University Phayao, where he became an assistant professor in 2010. Currently, he is an assistant professor at the Department of Applied Science, Faculty of Science and Technology, Nakhon Sawan Rajabhat University, Thailand. His research interests include image processing and pattern recognition. In those areas, he has published several papers in refereed journals, and in proceeding of international conferences and symposia.

Dr. Krisda is a senior member of International Association of Computer Science and Information Technology (IACSIT), Singapore. He is also a senior member of Science and Engineering Institute (SCIEI), South Korea.