Information Systems International Conference (ISICO), 2 – 4 December 2013

Document Authentication Using Print-Scan Image Watermarking Based on DCT (Discrete Cosine Transform) Algorithm

Nazori Agani*, M. Iman Wahyudi, Riyanto****
* Department of Electrical Engineering And Computer Science, Universitas Budi Luhur Jakarta Indonesia
** Department of Computer Science, Universitas Budi Luhur Jakarta Indonesia

Keywords:

Document authentication
Print-scan watermarking
DCT algorithm

ABSTRACT

The piracy of copyright issue has become very serious problem to be dealt. Important printed documents such as certificates, insurance policy, and other important documents can be manipulated by someone according to the technology development.

In this study, we propose a technique on document authentication using watermarking. The image which has a copyright sign (copyright image) is embedded to the cover image such as logo or any other image to be watermarked image, then the watermarked image will be embedded on document, the watermarked image will be extracted when the document will be authenticated. The method is based on DCT (Discrete Cosine Transform) algorithm.

The copyright image which has embedded on the cover image either on 24-bit grayscale images or 24-bit color images are detected through print-scan process, although grayscale 24-bit cover image is better than color 24-bit cover image. It shows that the technique is efficient for document authentication.

A printer which is used for this technique to be very influential on the quality of the printed document and it will be influenced on extraction process, the laser printer is recommended than inkjet printer. The next future work is expected that the 24-bit color image is able to be extracted as well as the 24-bit grayscale image.

Corresponding Author:

Nazori Agani,
Department of Electrical Engineering And Computer Scince,
Universitas Budi Luhur,
Jl. Raya Ciledug, Jakarta Selatan, Indonesia (12260)Telp. (021)5853753, Fax.(021)5853752.
Email: nazori.agani@gmail.com

1. INTRODUCTION

The developments of information and communication technologies has given rise the piracy in attempt to manipulate and to copy without permission the creation technology such as picture video, audio and digital document. The problem of copyright issues in this field is not just about copying and distribution alone, but also on the label ownership. Important printed documents such as certificates, insurance policy, and other important documents can be manipulated by someone according to the technology development. The important document must be protected by any technique that will not be manipulated by other, because the manipulation of this document will cause the loss of the owner.

Watermarking is an implementation of steganography which focused on signing in digital documents, these documents can be a proof of ownership of an organization who has patented its copyright even in digital image data, voice data, or video data form. The technique of watermarking [1] is the process of adding a permanent identification code into digital data. The identification code can be text, images, sounds, or video. In addition it will not damage the data of digital products that will be protected, the inserted code should have a resistance (robustness) of various advanced processing such as conversion, transformation geometry, compression, and encryption.

Digital watermarking is a method to hide some information that is being integrated with multimedia objects. Such objects may include images, sound, video, or writing. Watermarking itself is used in various applications, such as ownership evidence, fingerprinting, authentication and integrity verification, and content labeling and protection.

Digital watermark must satisfy the following criteria [2] according to the document authentication:

a. Imperceptibility; the watermark should not imperceptible by human observe
b. Secure and reliable; the embedded watermark cannot be deleted and retrieved from the host image
c. Robust; the watermark can survive from various attack like compression, cropping, and resizing
d. Unambiguous; the watermark logo must unambiguously identifying the owner

The image watermarking technique [3] generally consists of two stages: 1) Embedding watermark; watermark image embedded to cover image as shown on figure 1, and 2) The detection or extraction of watermark as shown on figure 2, the technique has two ways; blind and nonblind watermarking. Blind watermarking system does not require the original image to extract the watermark, whereas the nonblind watermarking need to be able to extract the original image watermark.

Watermarked image and then distribute for example, published in the web or sold to customers or embedding on important document. During transmission and distribution, watermarked image is distorted due to the common image processing, such as compression, improved contrast, resize, re-sampling, gamma correction, and so on. All the distortions imposed on watermarked image is viewed as an attack. Any attack contributes noise (n) on the image and can interfere with the detection process. Good watermarking methods should be robust against attacks that can damage or destroy the watermark in the image.

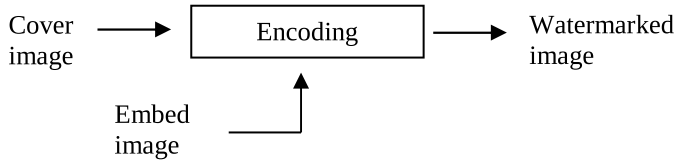


Figure 1. Embedding watermark image process

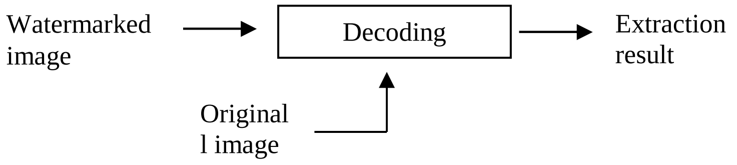


Figure 2. Extraction watermarked image process

The Research and Implementation of watermarking has been widely applied, both watermarking technique based on LSB (Least Significant Bit), DCT (Discrete Cosine Transform), (DFT) Discrete Fourier Transform, Discrete Wavelet Transform (DWT) and Discrete Laguerre Transform (DLT).

Watermarking itself has implemented in both in digital data or data that has been printed. Watermarking implementation on the print media known as Print-Scan Watermarking.

There have been many algorithms and techniques offered to address deficiencies in the implementation of the Print-Scan watermarking. Anja Keskinarkaus [4] proposed some of the techniques in printing images watermarking; (a). method for adding resilience watermarked image with perceptual weighting (b). Synchronizing multi bits messages resistance to extraction. Fr´ed´eric Lef`ebvre et.al. [5] Proposed a method which combines an additive watermarking algorithm in the spatial domain, the additive watermark in the spatial domain is based on an original generalized 2-D cyclic pattern for secret message embedding and exhaustive search and a synchronization template in the Fourier domain Hironori Takimoto [6] proposed template techniques to image two-dimensional template form which embedded in the watermarked image. Qibin Dajun He and Sun [7] gave the solution to the watermark image attachment strategies by classifying blocks and pasting distinction based on the block. Yen chiu chung et.al [8] and Tan Yi-zhouming et.al [9] proposed print-scan watermarking and Coding Synchronization of Peak Locations in Frequency Domain algorithm based on DFT algorithm. Anu Pramila [10] proposed print-scan watermarking using multiple domain both spatial and frequency domain in DWT. Solanki et.al.[11] present methods for hiding information into images in a manner that is robust to printing and scanning, the proposed methods are blind, the original image is not required at the decoder to recover the embedded data, two methods for hiding information resilient to print-scan operation are proposed. The first technique, called selective embedding in low frequencies (SELF), hides data in the magnitude of dynamically selected low-frequency DFT coefficients.

In this paper propose DCT algorithm through print-scan watermarking in purpose to validate important document which has embedded watermark image to be extraction.

2. DCT (Discrete Cosine Transform)

Discrete Cosine Transform (DCT) is used to convert a signal into frequency components basically. DCT was first introduced by Ahmed, Natarajan and Rao[12].

DCT has two main properties for image and video compression; for 2D digital image, DCT 2D throught matrix I M x N is formulated [13] as follows:

$$C(p,q) = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m,n) \cos \frac{\pi 2(m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}$$

....(1)

Values C(p,q) are called by DCT coeficient from tranformation of image I, and its inverse of DCT (IDCT) is formulated as follow:

$$I(m,n) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \alpha_p \alpha_q C(p,q) \cos \frac{\pi 2(m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}$$

Where

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}} & , p = 0 \\ \sqrt{\frac{2}{M}} & , 1 \le p \le M-1 \end{cases} \qquad \alpha q = \begin{cases} \frac{1}{\sqrt{N}} & , q = 0 \\ \sqrt{\frac{2}{N}} & , 1 \le Q \le N-1 \end{cases}$$

.....(2)

DCT domain divides the image into three sub-band frequencies (low, middle, and high) it is shown on figure 3. Embedding watermark image into low frequency can destroy the image, and embedding watermark into high sub-band will erase easily by quantitation operation easily, like lossy compression. Therefore the best way is embedding watermark image into middle frequency.

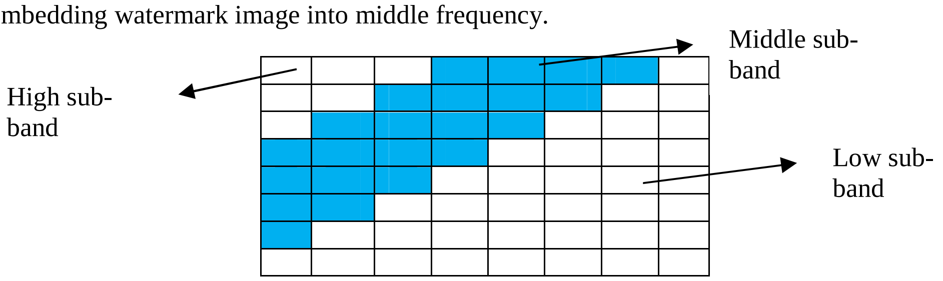


Figure 3. three sub-band frequencies in DCT

Discrete Cosine Transform represents an image of the sum of the magnitude and frequency sinusoid changing. The nature of the DCT is changing significantly the image information is concentrated on just a few DCT coefficients. Therefore DCT is often used for image compression as the JPEG.

3. RESEARCH METHOD

This study was conducted to create a technique that can be used to validate reliably an important document as a reference that the document was issued by the issuing party. Watermarking technique is used as a problem-solving to create a document in which integrated a sign embedded either logo or other image-owned by the owner. After successfully embedded the document then print with printers media. The document will be validated with scanning on the documents when the extraction process, when the embedded image is detected it show that the document is valid. In this study propose the watermarking techniques and extraction based on DCT (Discrete Cosine Transform).

The watermarking information by using the DCT algorithm can be described as follows:

For two-dimensional signals such as digital images, two-dimensional DCT on the size of the matrix I M x N is defined as follows: [13] describe on (1). For instance we call DCT coefficient stored in array v. embed watermark image (w) into v using:

$$V(i) = v(i) + \alpha \, |v(i)| w(i) \quad ....(3)$$

In this case α is the factor to indicate the strength of the watermark ($0 < \alpha < 1$).

The values of C (p, q) are called the DCT coefficients of the image I. Inverse DCT transformation is expressed by the equation describe on (2).

DCT watermark embedding methods known to be very powerful for robustness quality, because some research on watermarking method attacked the watermarked image by a variety of attacks such as compression, cropping, gamma correction [10] but the embedded image still detected, we are focused on the 256 x 256 and 512 x 512 pixel image either greyscale or colour on one logo image in table 1 and 2.

In this study the framework are described as follows:

1. Embed the image with the DCT algorithm, with condition that the image will be embedded should have a smaller resolution than the cover image
2. Embed the watermarked image on the important document
3. Print the document
4. Scan the document with 300 dpi setting
5. Normalization of the scanned document by:
   a. Cropping the watermarked image
   b. Convert the watermarked image to 24 bits and bmp extension
6. Eextract the watermarked image using DCT algorithm

The implementation algorithm in this paper is using DCT algorithm which has proposed by Crish Shoemaker [14] who propose only for watermarking and extraction process using DCT algorithm, in this paper this algorithm develop to be print-scan watermarking process.

The embedding watermark algorithm in this system describe as follows:

1. Convert cover image of 24-bit jpg extensions to bmp extension.
2. Perform DCT transformation the cover image
3. Embed watermark image of 8-bit bmp extension with smaller resolution than the cover image
4. Restore the image of the 24-bit as watermarked image.

While the algorithm for image extraction are as follows:

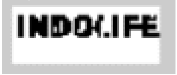
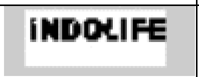
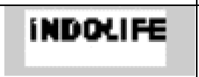
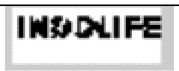
1. Convert 24 bit of jpg extension image after print-scan process to bmp extension
2. Perform DCT transform for extraction to watermarked image

In this study in extraction process using blind method which need original image for extraction process in order that the embedded image is original from the owner of the document.

4. RESULT AND ANALYSIS

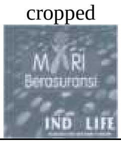
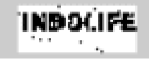
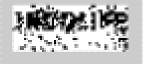
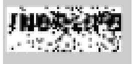
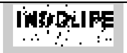
In the explanation of the previous chapter the study of image watermarking using DCT algorithm has the result listed as follow;

Tabel 1. The result without Print-scan image watermarking

Nu	Original image	Embedded image	Watermarked image	Extraction result	Extraction time	Note
1	grayscale 24 bit.jpg 512 x 512	grayscale 8 bit.bmp 50 x 20	grayscale 24 bit.bmp 512 x 512		3.3852	Detected clearly
2	color 24 bit.jpg 512 x 512	grayscale 8 bit.bmp 50 x 20	color 24 bit.bmp 512 x 512		3.3072	Detected clearly
3	color 24 bit.jpg 256 x 256	grayscale 8 bit.bmp 50 x 20	color 24 bit.bmp 256 x 256		1.0296	Detected clearly
4.	grayscale 24 bit.jpg 256 x 256	grayscale 8 bit.bmp 50 x 20	grayscale 24 bit.bmp 512 x 512		1.0608	Detected clearly

This study is focused on 24 bit and jpg extension with resolution 256 x 256 pixel, and 512 x 512 pixel either grayscale or color image, and the printer which used in this study is inkjet canon MP and got the result as follow:

Tabel 2. The result with Print-scan process image watermarking

Nu	Original image	Embedded image	Watermarked image	Print-scan process	Extraction result	Extraction time	Note
1	grayscale 24 bit.jpg 512 x 512	grayscale 8 bit.bmp 50 x 20	grayscale 24 bit.bmp 512 x 512	grayscale 24 bit.bmp 512 x 512 cropped 		2.7768	Detected clearly
2	color 24 bit.jpg 512 x 512	grayscale 8 bit.bmp 50 x 20	color 24 bit.bmp 512 x 512	color 24 bit.bmp 512 x 512 cropped		3.3852	Detected unclearly
3	color 24 bit.jpg 256 x 256	grayscale 8 bit.bmp 50 x 20	color 24 bit.bmp 256 x 256	color 24 bit.bmp 256 x 256 cropped		0.8736	Detected unclearly
4	grayscale 24 bit.jpg 256 x 256	grayscale 8 bit.bmp 50 x 20	grayscale 24 bit.bmp 256 x 256	grayscale 24 bit.bmp 256 x 256 cropped		0.8736	Detected clearly

The Table 2 shows no 2, and 3 shows that the result are not good according to the image attack degradation of the color of the image after print-scan process, it is the same case as Mei Jiansheng [15] which attack the image with the various technique. Although it can be conclude that print-scan method can be implemented on image with DCT algorithm for document authentication on important document, and the grayscale image is better for this technique.

5. CONCLUSION

From the study, it can be concluded that watermarking techniques using DCT can be implemented efficiently for the print-scan watermarking. Implementation of DCT algorithms proposed with little development can be done for the print-scan watermarking, Results from the testing of samples produced that the image grayscale better than color image for print-scan technique, the shortcomings of this technique is the extraction time embedding and relatively long, when the image has a large resolution, and works only on the same amount of pixels between the cover image and the image to be extracted.

In this study due to limitations of the printer which used was very simple, in the future studies are expected to use a better printer to obtain a better result for print-scan watermarking so that the criteria for image watermarking as robustness, safe, imperceptibility can be obtained, and also this technique must be obtain good result on color image in extraction as the grayscale image.

REFERENCE

[1] Ir. R. Munir, M.T., “Steganografi dan Watermarking,” Departemen Teknik Informatika Institut Teknologi Bandung. 2004.

[2] W.Y. Chen, *et.al.*, “Digital Watermarking Using DCT Transformation,” Department of Electronic Engineering, National Chin-Yi Institute of Technology, pp.173-184.

[3] R. Munir, “Image Watermarking untuk Citra Berwarna dengan Metode Berbasis Korelasi dalam Ranah DCT,” Program Studi Teknik Informatika ITB Sekolah Teknik Elektro dan Informatika ITB.

[4] A. Keskinarkaus, “Digital Watermarking Techniques For Printed Images,” *Thesis*, *University Of Oulu Graduate School;* University Of Oulu, Faculty Of Technology, Department Of Computer Science And Engineering. 2012.

[5] F.E. Lef`ebvre, *et.al.*, “A Print And Scan Optimized Watermarking Scheme,” Laboratoire de T´el´ecommunications et T´el´ed´etection Universit´e catholique de Louvain – Belgium.

[6] H. Takimoto, *et.al.*, “Invisible Print-Type Calibration Pattern Based On Human Visual Perception,” *IEEE 17th International Conference on Image Processing Hong Kong* 2010.

[7] D. He *et.al.*, “A Practical Print-Scan Resilient Watermarking Scheme Institute” for Infocomm Research, 2005.

[8] Y.C. Chiu, *et.al.*, “Copyright Protection against Print-and-Scan Operations by Watermarking for Color Images Using Coding and Synchronization of Peak Locations in Frequency Domain,”*Journal Of Information Science And Engineering* 22, 2006. pp. 483-496

[9] T. Yi-zhou, *et.al.*, “An Optical Watermarking Solution for Color Personal Identification Pictures College of Mechatronics Engineering and Automation,” College of Science, National University of Defense Technology, Changsha, Hunan, China..

[10] A. Pramila, *et.al.*, “Multiple domain watermarking for print-scan and JPEG resilient data hiding,” MediaTeam Oulu. Department of Electrical and Information Engineering. University of Oulu.

[11] K. Solanki , *et.al.*, “Print and Scan’ Resilient Data Hiding in Images Member,” *Ieee Transactions On Information Forensics And Security, Vol. 1, No. 4, December.* 2006, pp. 464-478

[12] N. Ahmed.**;** Natarajan, T.; Rao, K. R., “Discrete Cosine Transform,” *IEEE Transactions on Computers* C–23 (1), 1974, pp.90–93.

[14] C. Shoemaker, “Hidden Bits: A Survey of Techniques for Digital Watermarking,” Independent Study, Research Report, 2002.

[15] M. Jiansheng, L. Sukang, and T. Xiaomei, “A Digital Watermarking Algorithm Based On DCT and DWT,” *Proceedings of the 2009 International Symposium on Web Information Systems and Applications (WISA’09) Nanchang*, P. R. China, May 22-24, 2009, pp. 104-107

BIBLIOGRAPHY OF AUTHORS

	Nazori, AZ, has received the PhD degree in Electrical Engineering from University Teknologi Malaysia in 2007. Now, he is staying in Universitas Budi Luhur as a lecture, and head of magister of Computer Science. His Research interest are; Image Processing, Computer Vision and Computational techniques.
	M. Iman Wahyudi, has received his Bachelor of English Education from IAIN Banten. He is currently studying his Magister of Computer Science program in Universitas Budi Luhur Jakarta Indonesia. He is currently working as a teacher of ICT at Assa’adah Islamic Boarding School Banten Indonesia.
	Riyanto, has received his Bachelor Degree in 2008 in Information and Technology Faculty in Universitas Budi Luhur, and continued magister of Computer Science in Universitas Budi Luhur until now, currently he is working at an insurance company; Indolife Pensiontama as supervisor of information technology.