Annual ADFSL Conference on Digital Forensics, Security and Law 2013

Jun 11th, 1:35 PM

An Image Forensic Scheme with Robust and Fragile Watermarking for Business Documents

Sai Ho Kwok

Department of Information Systems, Business Statistics and Operations Management, The Hong Kong, University of Science and Technology, jkwok@ust.hk

AN IMAGE FORENSIC SCHEME WITH ROBUST AND FRAGILE WATERMARKING FOR BUSINESS DOCUMENTS

(**Briefing Paper/Presentation**)

Sai Ho KWOK
Department of Information Systems
Business Statistics and Operations Management
The Hong Kong University of Science and Technology
Clear Water Bay, N.T.
Hong Kong SAR
Phone: (852) 2358 7652
Fax: (852) 2358 2421
Email: jkwok@ust.hk

ABSTRACT

This paper proposes an image forensic scheme with both robust and fragile watermarking techniques for business documents. Through a dual watermarking approach, the proposed scheme can achieve image forensics objectives of (a) identification of source; (b) authentication of documents; and (c) locating the tempered areas of documents due to attacks. An example is presented to prove the concepts of the proposed scheme.

Keywords: Image Forensics, Fragile and Robust Watermarking, Business Document.

1. INTRODUCTION

In business, digital documents include legal documents, official reports, contracts, agreements and so on. Currently, most of business documents are in MS Word and PDF formats and they contain both text and graphs. In this paper, business documents are represented as digital images technically. Digital forensic becomes crucial when these business documents are treated as evidence of a crime or attack. In the private sector, digital forensic is required during internal corporate investigations or intrusion investigation.

In general, digital forensic can be used to attribute evidence to specific suspects, confirm alibis or statements, determine intent, identify sources, or authenticate documents (Wikipedia, 2013). Zhou et al. (Zhou et al., 2011) point out that digital image forensics can be used to (a) judge whether an image is from a particular digital camera; (b) determine whether an image was produced by the same type of cameras or other equipment or software; (c) to determine whether an image has been processed (or attacked); and (d) to determine whether an image is the original one or not.

Digital watermarking is a key process in active image forensics (Xi et al., 2011; Zhou et al., 2011). It is a process of embedding relevant information (such as a logo, fingerprint, and serial number) into a digital content (Kung et al., 2003; Lin et al., 2001; Xi et al., 2011). The embedded data can be viewed as digital watermarks. There are two types of digital watermarks and they are visible (or fragile) and invisible (or robust) watermarks.

A visible or fragile watermark is the translucent logos that often appear at the corner of images, in an attempt to prevent copyright infringement. Such watermarking process operates in the spatial domain, where the corresponding pixel values are modified directly or indirectly (Kutter et al., 1998; Voyatzis et al., 1997). However, these visible watermarks can be targeted and removed rather simply by cropping the image, or overwriting the logos (Xi et al., 2011). A visible or fragile watermark is used for content authentication applications to verify or authenticate the integrity of digital documents.

An invisible or robust watermark can often resist intentional and unintentional attacks to the images. Common attacks are often referred to image processing, such as compression, rotation, filtering, zooming and so on (Kung et al., 2009). The robust watermarking usually operates in the transform domain (Briassouli et al., 2004; Lin et al., 2008; Lu et al., 2009; Tsui et al., 2008; Zou et al., 2008). In other words, robust watermarking techniques can embed data in the transform domain, such as frequency domain through Discrete Fourier Transform (DFT) (Tsui et al., 2008; Zou et al., 2008), Discrete Cosine Transform (DCT), and Discrete Wavelet Transform (DWT) (Lin et al., 2008; Lu et al., 2009). A robust watermark can be used to determine the source of the image and copyright protection (Zhou et al., 2011).

The objectives of digital image forensics for business documents include (a) identification of source; (b) authentication of documents; and (c) locating the tempered areas of documents due to attacks. Due to the nature of business documents, which are different from ordinary images, image forensics techniques for business documents could be quite different from those techniques in the literature of digital image forensics. This paper intends to fill this gap and proposes an image forensics scheme for business documents specifically. The proposed scheme falls into the category of active image forensics as digital watermarking is applied after the documents have been produced. The proposed scheme adopts a dual watermarking process, which is a rather innovative and promising approach in tackling image forensics problems (Zhou et al., 2011).

2. PROPOSED IMAGE FORENSICS SCHEME

The following are the terms used in the proposed image forensics scheme.

- A business document, $\boldsymbol{D}$ is treated as an 8-bit gray-scaled image, where its pixel values ranging from 0 to 255.
- A fragile watermark, $\boldsymbol{W}$ is a pre-defined watermark specifically for an organization, for example a company logo or trademark. This is also an 8-bit gray-scaled image.
- An authentication data, $\boldsymbol{AD}$ is data for identifying a particular person or computer in an organization. It can be a staff ID, an IP address of a computer, and so on. In digital watermarking, it is commonly known as the ***Key*** of a watermarking process.
- The embedded watermark, $\boldsymbol{W_f}$ is the watermark to be embedded into $\boldsymbol{D}$. It is a fragile watermark and is produced by a robust watermarking process, whose inputs are $\boldsymbol{W}$ and $\boldsymbol{AD}$.
- A watermarked document, $\boldsymbol{D_w}$ is an 8-bit gray-scaled image. It is resulted from a fragile watermark insertion process, whose inputs are $\boldsymbol{D}$ and $\boldsymbol{W_f}$. A simple implementation can be $\boldsymbol{D_w = D + W_f}$.
- An evidence of business document, also known as a modified watermarked document is denoted as $\boldsymbol{D_w'}$, which a watermarked document picked up by the police force or internal/external audit for examination. A $\boldsymbol{D_w'}$ can be (i) exactly the same as $\boldsymbol{D_w}$. so the content of the document is identical to the original document, or (ii) a modified version of the original document, where the content of the document has been modified by attacks.

2.1 Watermark Insertion

In the proposed scheme, the watermark insertion process as depicted in Figure 1 is responsible for inserting a digital watermark, $\boldsymbol{W_f}$ into a business document, $\boldsymbol{D}$. The watermark, $\boldsymbol{W_f}$ is applied to all pages of the document. The process consists of two different watermarking techniques: a robust watermarking technique and a fragile watermarking technique. It is also known as dual watermarking technique. The primary objective of the robust watermarking process is to incorporate authentication data, $\boldsymbol{AD}$ into a fragile watermark, $\boldsymbol{W}$. It results in another fragile watermark, $\boldsymbol{W_f}$ for inserting to the business document, $\boldsymbol{D}$.

The fragile watermark insertion process operates at the spatial domain. A simple implementation of the process can be an additional operator, which can be represented as follows.

$$D_w = D + W_f$$

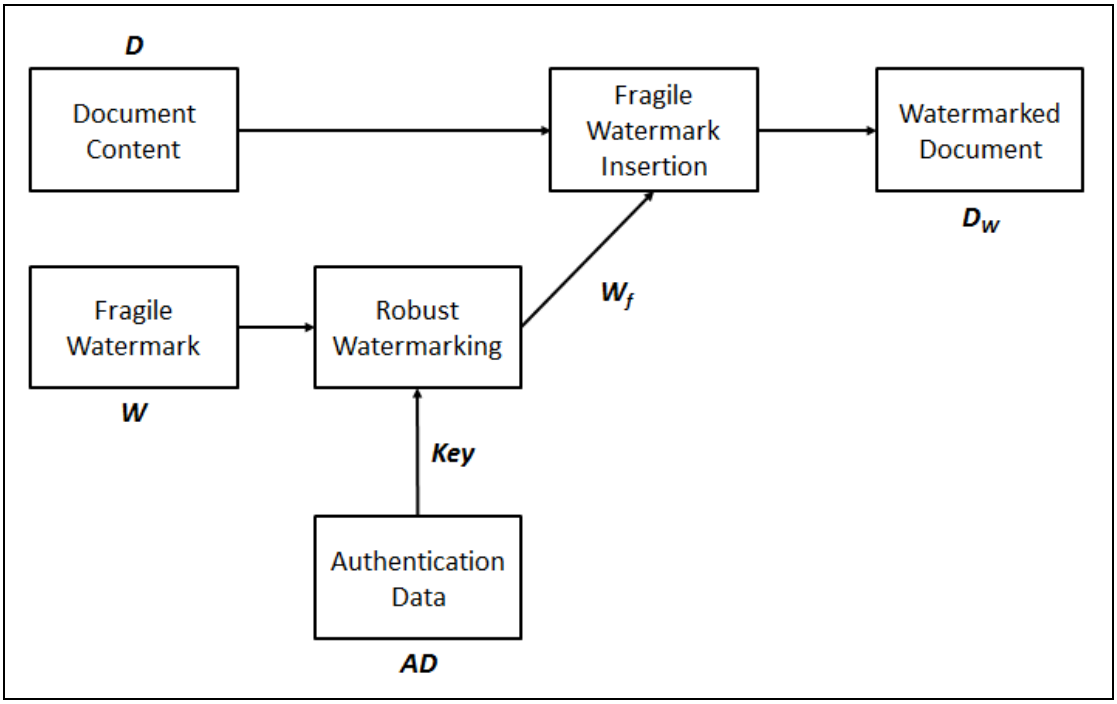

Figure 1 Watermark insertion process

2.2 Watermark Verification

The proposed watermark verification process is depicted in Figure 2. The objective of the watermark verification process is to verify whether the embedded watermark, W_f has been modified or not. If W_f is modified, the process is able to locate the tempered area. As fraudulent documents are usually very much similar to the original documents, except several major parts of the document, it is expected that D_w' should be very much similar to D_w. Based on this assumption, authentication can be achieved through the robust watermarking process.

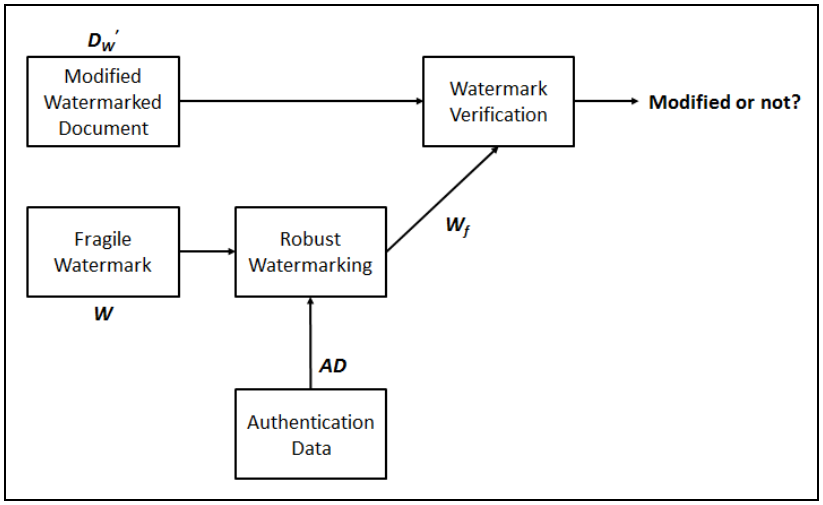

Figure 2 Watermark verification process

2.3 Proof or Concepts

Assume the original business document, ***D*** contains only a letter "H". Figure 3 shows all involving images during the watermark insertion process for the business document. The resulting watermarked document, $\boldsymbol{D_w}$ is produced. Then the watermark verification process is performed as $\mathbf{D_W}' - \mathbf{W_f}$.

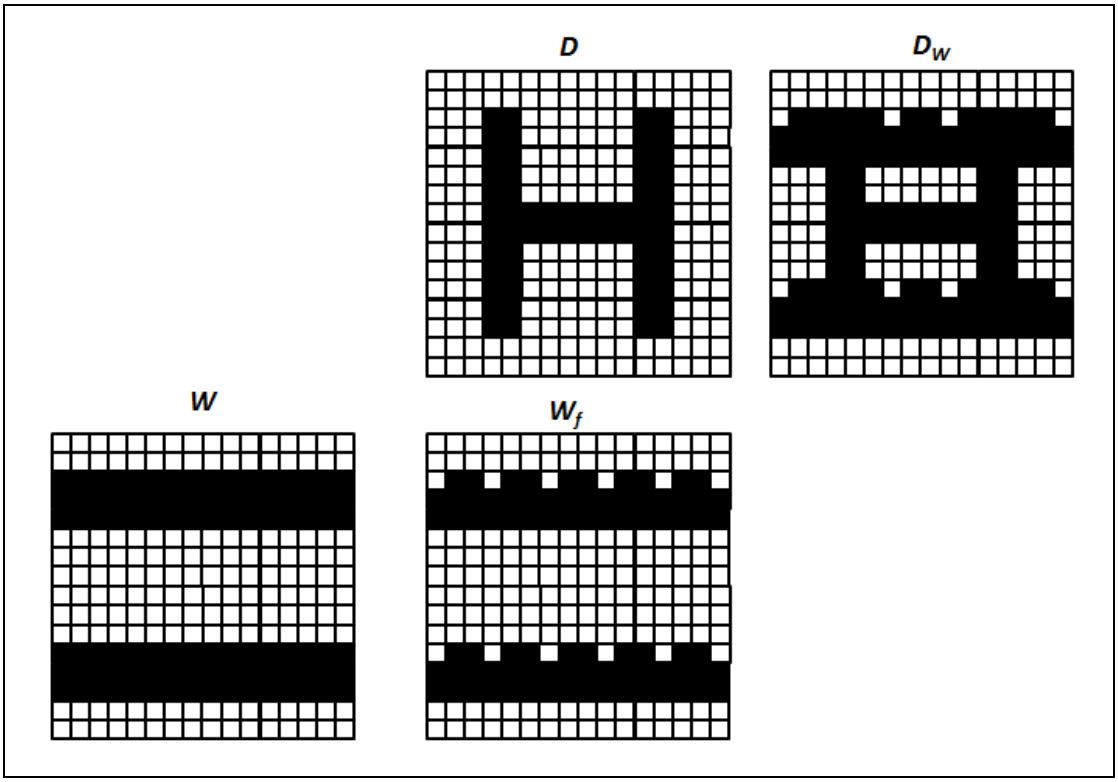

Figure 3 The watermark insertion process with an input of the letter "H"

Assume a digital evidence $\boldsymbol{D_w}'$ is presented for examination. The content of the business document has been changed from the letter "H" to the letter "I". Figure 4 illustrates the watermark verification process with $\boldsymbol{D_w}'$. Figure 5 presents the result of the watermark verification when the $\boldsymbol{D_w}'$ is derived from a business document with the letter "I" without any changes.

By comparing the results of Figures 4 and 5, it is shown that the watermark verification process can distinguish between a changed document and an unchanged document.

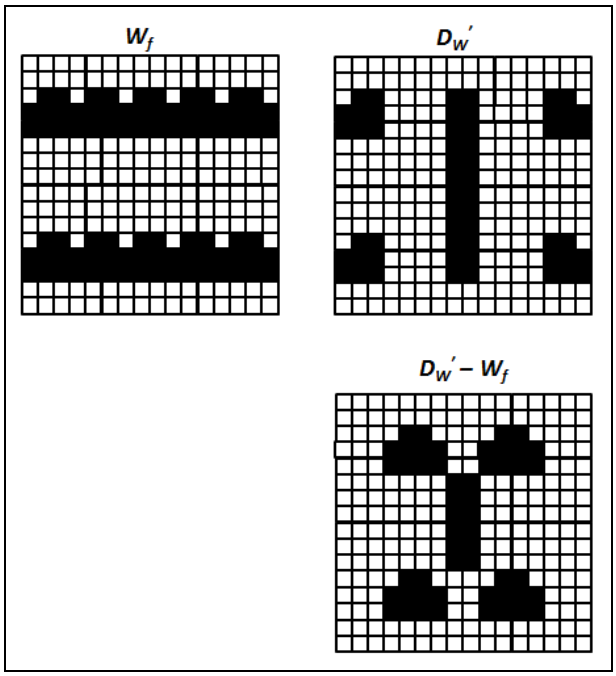

Figure 4 The watermark verification process with $\boldsymbol{D_w}'$ of the letter "I"

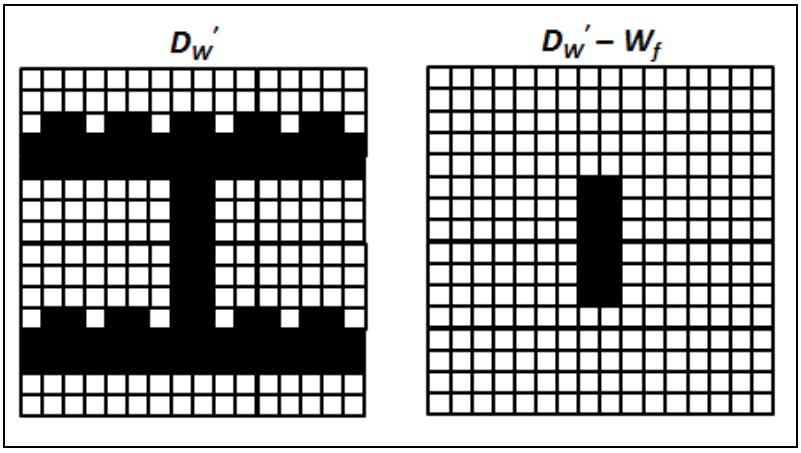

Figure 5 The result of the watermark verification process with the $\boldsymbol{D_w}'$ derived from the letter "I"

3. CONCLUSIONS

The proposed image forensics scheme can achieve the following major objectives of image forensics.

a) Identifying the source of business documents as a fragile watermark, ***W*** of company logo is used in the scheme;
b) Authenticating the documents as authentication data, ***AD*** are embedded into the documents through watermarking; and
c) Locating the tempered areas through the watermark verification process.

The following are the remaining problems with the proposed scheme.

- The choice of robust and fragile watermarking techniques: The objectives of the watermarking techniques in the proposed scheme have been specified, but which particular watermarking techniques can provide the best performance are under reviews.

- The choice of the file format of watermarked business documents: The file format is likely to be PDF as this has been wide used by many organizations but further research is needed.
- The design of fragile watermark: In the example, a black-and-white fragile watermark was used but it is likely to be a gray-scaled image in practice. Further investigations are required for the design of fragile watermarks, which is likely to be adopted by organizations.

The proposed scheme is still under development and further tests are required to evaluate the performance and applicability of the scheme in business.

REFERENCES

Briassouli, A, Strintzis, M. G. (2004). Locally Optimum Nonlinearities for DCT Watermark Detection. *IEEE Transactions on Image Processing, 13*(12), 1604-1617.

Kung Chih-Ming, Chao Shu-Tsung, Chao, Tu Yen-Chen, Yan Yu-Hua, and Kung Chih-Hsien (2009). A Robust Watermarking and Image Authentication Scheme used for Digital Content Application *Journal of Multimedia*, *4*(3), 112-119.

Kung C. M., Jeng, J. H., and Kung C. H. (2003). Watermarking Base on Block Property. *16th IPPR Conference on Computer Vision, Graphics and Image Processing*, 2003.

Kutter M., Jordan F., and Bossen F. (1998). Digital Watermarking of Color Images using Amplitude Modulation. *Journal of Electronic Imaging*, 326-332.

Lin, C. Y., and Chang, S. F. (2001). A Robust Image Authentication Method Distinguishing JPEG Compression from Malicious Manipulation. *IEEE Transaction on Circuits and Systems of Video Technology*, 11(2), 153-168.

Lin W H, Horng Shi-Jinn, Kao Tzong-Wann, Fan Ping-Zhi, Lee Cheng-Ling, and Pan Yi. (2008). An Efficient Watermarking Method based on Significant Difference of Wavelet Coefficient Quantization. *IEEE Transactions on Multimedia*, *10*(5), 746-757.

Lu Wei, Sun Wei, and Lu Hongtao. (2009). Robust Watermarking based on DWT and Nonnegative Matrix Factorization. *Computers & Electrical Engineering*, *35*(1), 183-188.

Tsui T. K., Zhang, X. P., and Androutsos D. (2008). Color Image Watermarking using Multidimensional Fourier Transforms. *IEEE Transactions on Information Forensics & Security*, *3*(1), 16-26.

Voyatzis, G., and Pitas, I. (1997). Embedding robust watermarks by chaotic mixing. *13th International Conference on Digital Signal Processing (DSP'97)*, 1997.

Wikipedia. (2013). Digital forensics – Wikipedia. Retrieved from http://en.wikipedia.org/wiki/Digital_forensics on March 14 2013.

Xi Ahao, Philip Bateman, and Anthony T. S. Ho. (2011). Image Authentication using Active Watermarking and Passive Forensics Techniques, in *Multimedia Analysis, Processing and Communications Studies in Computational Intelligence*, (eds) Weisi Lin, Dacheng Tao, Janusz Kacprzyk, Zhu Li, Ebroul Izquierdo, and Haohong Wang, Springer.

Zhou Guojuan, and Lv Dianji. (2011). An Overview of Digital Watermarking in Image Forensics. *Fourth International Joint Conference on Computational Sciences and Optimization*, 2011, China.

Zou Lu-Juan, Wang Bo, Feng Jiu-Chao (2008). A Digital Watermarking Algorithm based on Chaos and Fractional Fourier Transformation. *ACTA PHYSICA SINICA*, *57*(5), 2750-2754.