2012 International Conference on Information and Knowledge Management (ICIKM 2012)
IPCSIT vol.45 (2012) © (2012) IACSIT Press, Singapore

Applying Yellow Colour Markings as Low-level Security Watermarking of Grayscale Photocopied Documents

Hoong-Cheng Soong [+]

Faculty of Information and Communication Technology (FICT), Universiti Tunku Abdul Rahman (UTAR), Jalan Universiti, Bandar Barat, 31900 Kampar, Perak, Malaysia

Abstract. Digital watermarking is the technique used to provide additional and useful evidences for many application fields especially in copyright infringement detection. The focus of this paper is on Portable Document Format (PDF) watermarking documents and the printed watermarked hardcopy documents. Digital watermarking is indeed suitable to detect unauthorized copies from any authentic sources. Colour theory and colour properties are studied on how to prevent yellow-watermarked stamps being photocopied in grayscale, particularly by using luminance concept on the documents. Please note that the technique applies only for grayscale photocopies, as the detection for coloured copies requires disparity techniques for the coloured detection.

Keywords: Digital Watermarking, Grayscale, Document Management System, Luminance, Relative Luminance, PDF Watermarking

1. Introduction

According to Ross et al. [1], the purpose of hiding information is to make inaccessible certain details that should not affect other parts of a system. As such [16], information hiding techniques have recently been given much attention by researchers and industries. Thus, the interest in information hiding was triggered by concerns over copyright issues as audio, video, documents, and other works are now available in digital formats making it easier to make unauthorized copies. In consequence, watermarking is one of the techniques for information hiding. Apart from that, digital watermarking is defined by Kutter et al. [2] as imperceptible insertion of information into multimedia data. At most [3], Portable Document Format (PDF) digital documents are popular nowadays due to the de facto standard for electronic exchanges of documents and are now, as the standard in the industries for intermediary representation of printed material in electronic prepress systems for conventional printing applications.

2. Related Research

As mentioned before, the focal concentration is in PDF digital watermarking. What is exactly digital water marking? According to van Schyndel et al. [4] and Cox et al. [5], digital watermarking is used as the last resort and as the "last line of defences" as the failure safeguard of the encryption or copy protection occurred and thus the illegal copies are recognized as to against distribution of valuable digital media. Furthermore, [6] stated that a digital watermarking is embedded directly into the system. As for the exemplar, information about copyrights, ownership, timestamps, and the legal recipient possibly will be embedded. However, the implanted digital watermarking by itself is unable to avoid illegally copying, modification and re-distribution of the genuine documents. Nevertheless, watermarking is efficient in tracking and tracing to the rightful owners of detecting unauthorized usage of documents. Thus, punishable actions can be taken to the illegal users provided the watermark can be retrieved and detected from the documents whether it is

[+] Corresponding author. Tel.: + 605–468 8888; fax: +605-466 1672.
E-mail address: *soonghc@utar.edu.my*

invisibly encrypted or visibly stamped as evidences. Exchanges of documents and are now, as the standard in the industries for intermediary representation of printed material in electronic prepress systems for conventional printing applications.

2.1. Governing Equations

Contributions Apart from that, watermarking usually has a set watermarking scheme that has to be done regardless of any type of digital watermarking. The watermarking scheme is as in the equation (1) as given by Dittmann et al. [7] below:

$$Sc = (E, D, R, M, Pe, Pd, Pr) \quad (1)$$

Sc represents the instance of watermarking scheme, E is the embedding method, D is the detecting function, R is the retrieval function, M is the messages, Pe is watermarking parameters, Pd defines the detection parameters, and finally Pr is the retrieval parameters. According to Wolfgang et al. [8], perceptually based watermarks consist of three principles that are robustness, capacity, and transparency. Consequently, digital watermarking can indeed be classified as robustness, capacity, perceptibility/transparency and lastly the embedding methods. Chen and Wornell [9] states that the embedding methods can be classified into three categories are such as spread-spectrum, quantization, and amplitude modulation. In spite of using only perceptible watermark imprinted on the document, illegal grayscale photocopy is easily identified as the yellow-watermarked stamps are impossible to be grayscale photocopied. Hence, colorimetry [10] is the science of measuring color as well as can be translated into the value of luminance to give a new grayscale colour as in equation 2 (luminosity function).

$$F = 683.002lm/W.\int_0^\infty \bar{y}(\lambda)J(\lambda).d\lambda \quad (2)$$

However, [11] by using sRGB colour space it is much easier to denote the relative luminance as compared to luminosity function. Yet, the relative luminance does reflect the luminosity function from the equation (2). From the colour space as denoted in XYZ, the linear Y is derived as the luminance from the colorimetric measurement as in equation (3).

$$Y = 0.2126R + 0.7152G + 0.0722B \quad (3)$$

Based on equation (3), the sRGB constants for R, G and B are not linear and thus another set of equation from linear transformation from RGB colour space is used as shown in equation (4). In addition, the luminance Y (RGB luminance value) from equation (4) is corresponding to intensities of the gray colours from the conversion of RGB images to grey-level images before extraction of the Spatial Grey-Level Dependence Matrices (SGDM) [12].

$$Y = 0.3R + 0.59G + 0.11B \quad (4)$$

3. Methods

As The following subsections are the approaches of the PDF digital watermarking that entail the process flow of the system (system architecture), programming techniques (PDF stamping technique), theoretical watermarking technique (digital watermarking technique and luminance theory), and the prevention of document modification (strategic watermark location). Nevertheless, the focus of the PDF watermarking approach is focussed solely on the hardcopies and the flow of the system instead of the digitally imperceptible watermarking as the watermarked softcopies are deleted as soon the documents are sent to the printers. In spite of that, it is essential to learn in depth the digital watermarking technique for the softcopies as it may be useful for the current research as well as for the future research.

3.1. Governing Proposed system: System Framework/Architecture

By using the system process, certain system flows could increase the watermarking robustness as well as to decrease the chances of perceptibility to the imprinted watermark. Accordingly, the system requires certain level of permission loops before acquiring the watermarked printed documents. For this reason, the digital documents are sent directly to the server for the printing process and the clients are denied access to the softcopy, thus protecting the digital watermarks. As a final point, database records track the digital documents and thus the authenticity can be matched for the detections of the printed watermarked documents. Figure 1 illustrates the process flow of the proposed system known as Document Control System (DocCon).

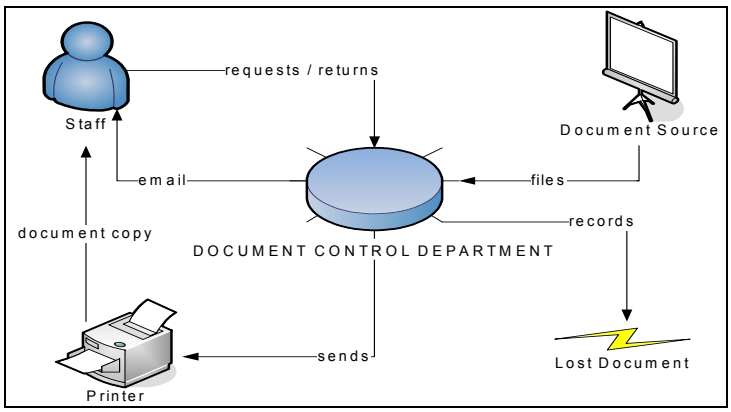

Fig. 1: Original process flow of proposed system.

3.2. PDF Stamping Technique: Appending Digital Watermark

According to Fitzgerald [13], there are steps that must be followed to append the watermark efficiently as illustrated in Figure 2. An extra page must be created prior to be spawned or not spawned to the entire document. As soon as the insertion is completed, the original page and template are deleted to avoid addition of unnecessary pages. The Figure 3 shows those steps are applicable to any programming language, as long as the steps are followed. However, even though the steps are similar for each programming technique, the syntax and sometimes the programming methods may vary.

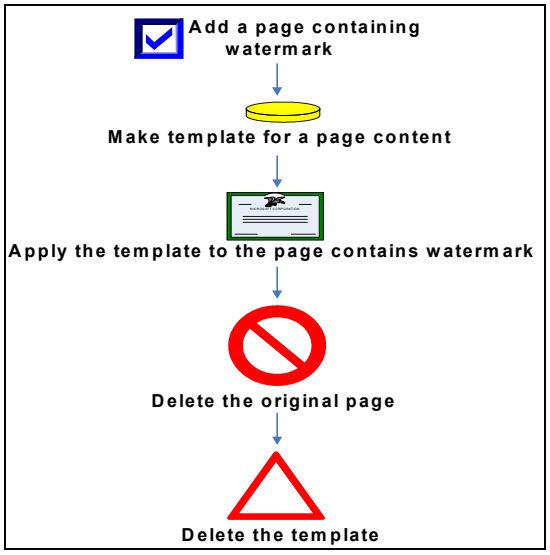

Fig. 2: Steps of appending the watermarks.

```
1  for (var p=0; p<this.numPages; p++)
2  {
3    this.insertPages(p, "/d/jsdocs/wm.pdf");
4    this.createTemplate("onepage", p);
5    op = this.getTemplate("onepage");
6    op.spawn(p+1, true, true);
7    this.deletePages(p);
8    this.removeTemplate("onepage");
9  }
```

Fig. 3: Fragment code of appending the watermarks (JavaScript).

3.3. PDF Stamping Technique: Appending Digital Watermark

Photometry [14] deals with the measurement of visible light as perceived by human eyes. The human eye can only see light in the visible spectrum and has different sensitivities to light of different wavelengths within the spectrum. When adapted for bright conditions (photopic vision), the eye is most sensitive to greenish-yellow light at 555 nm. Hence, the green constant will be the highest, using the linear function of relative luminance: $Y = 0.3R + 0.59G + 0.11B$. Given the yellow intensity [15] is (255, 0, and 0), magenta intensity is (255, 0, and 255) and cyan intensity is (0, 255, and 255) from the RGB value, then map the luminance intensity to the scale of grayscale where the whiter colour will consist of higher luminance since white has the maximum relative luminance as in Figure 4. As a result, yellow will be as near to white because of the higher luminosity, compared to the grayscale generated CMYK colour model from Adobe Photoshop CS3 as in Figure 5. From the results, it is easily detected that the document is grayscale photocopied from the original coloured sources. Therefore, this particular detection is appropriate in the working environment where black and white photocopier is commonly available.

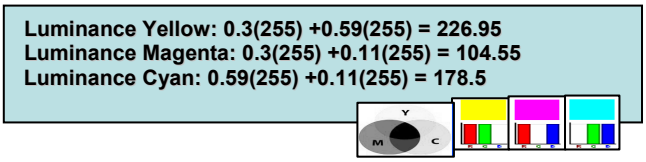

Fig. 4: Luminance values from the colours.

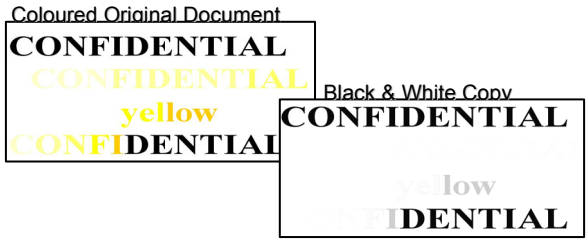

Fig. 5: Left is the original and right is the grayscale copy.

4. Conclusions

From the case study conducted, it is proven that yellow colour cannot be copied using grayscale devices. With the scientific calculation as the proof, it is also found that yellow has the second highest luminance after the colour white. Even though the printed documents cannot be prevented from being copied, the implementation of the yellow watermark can trace or detect unauthorized or illegal copies of the printed documents from any grayscale photocopy environment. Even if the masquerades try to assume ownership of the documents, the details imprinted on the document may prove otherwise. In addition, the manual stamping seal forms a part of the automation system for the Document Control System (DocCon) using PDF stamping technique discussed. As a result, the security for the documents is increased through DocCon as the digital watermarking technique is implemented (PDF stamping) and the softcopy of the files will never be retrieved by the users for any modification as the softcopy files are deleted as soon it is printed from the server side. In conclusion, the research shows that watermarking is one of the information hiding techniques which can be used to control and manage documents. Therefore, information hiding is rapidly becoming the most sought after method to protect industries in music, publications, filming, and other works which may be digitized. Even if the solution is provided for the grayscale photocopies, further enhancements will be done if better solutions are found from further researches. The solution for the colour photocopies is yet to be accomplished further through research which will be more emphasized on the coloured copies of digital watermarking. In addition, the embedding method for digital watermarking techniques should be thoroughly studied to enhance the security of the printed documents. As for the licensing problems encountered for the PDF stamping, various programming PDF stamping techniques are researched for better solution to the Document Control System (DocCon).

5. Acknowledgements

I would like to thank UTAR and FICT for the endless support.

6. References

[1] D. T. Ross, J. B. Goodenough, and C. A. Irvine, (1975). *Software Engineering: Process, Principles, and Goals, IEEE Computer*, **8** (5), May 1975, pp. 17 – 27.

[2] M. Kutter, and F. A. P. Petitcolas (1999). A Fair Benchmark for Image Watermarking Systems. *Electronic Imaging '99. Security and Watermarking of Multimedia Contents*, January 25-27, 1999, 3657, pp. 1–14.

[3] J. Meehan, E. Taft, S. Chernicoff, C. Rose, and K. Ron, (2005). *PDF Reference, Fifth Edition Version 1.6*. California: Peachpit Press.

[4] R. G. van Schyndel, A. Z. Tirkel, and , C. F. Osborne (1994). A Digital Watermark. *Proceedings of the 1994 IEEE International Conference on Image Processing*, 1994, 2, pp. 86–89.

[5] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, (1996). Secure Spread Spectrum Watermarking for Images, Audio, and Video. *Proceedings of the 1996 IEEE International Conference on Image Processing*, 1996, **3**, pp. 243–256.

[6] J. K. Su, F. Hartung, and B. Girod, (1999). Digital Watermarking of Text, Image, and Video Documents. *Preprint submitted to Elsevier Preprint*, August 23, 1999, pp. 1–16.

[7] J. Dittmann, and D. Meg´ıas, A. Lang, and J. Herrera-Joancomart´ı, (2006). A Theoretical Framework for a Practical Evaluation and Comparison of Audio Watermarking Schemes in the Triangle of Robustness, Transparency and Capacity. *Otto-von-Guericke University of Magdeburg, Germany & Universitat Oberta de Catalunya, Spain*, 2006.

[8] R. B. Wolfgang, and C. I. Podilchuk, and E. J. Delp, (1999). Perceptual Watermarks for Digital Images and Video. *AT&T foundation*, 1999 pp. 1-46.

[9] B. Chen, and G. W. Wornell, (2001). Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding. *IEEE Transaction on Information Theory*, May 2001, Vol. 47, No. 4, pp. 1423–1443.

[10] K.. Hirakawa, and T.W. Parks (2005). *Chromatic Adaptation and White-Balance Problem*. IEEE ICIP.

[11] M. Stokes, M. Anderson, S. Chandrasekar and R. Motta (1996). *A Standard Default Color Space for the Internet - sRGB*.

[12] S. N. Ondimu and H. Murase. Thermal properties of living roof greening material by inverse modelling. *Applied Engineering in Agriculture*. 2006, **22**: pp. 435-441.

[13] M. Fitzgerald, (2004). *Using Javascript to Apply Template to a PDF File*. Byte Ryte, The Netherlands.

[14] Y. Ohno, (1999). OSA Handbook of Optics, Volume III Visual Optics and Vision Chapter for Photometry and Radiometry. *Optical Technology Division*, Oct 20, 1999, pp. 1 – 17.

[15] J. Urban. Automatic segmentation of HeLa cell images. *Mid Sweden University in Sundsvall*, 2007, pp. 1-33.

[16] Y. Qin. A New Sample-Based Algorithm for Inpainting Used in Secrete Information Hiding. *Advances in Electronic Commerce, Web Application and Communication; Advances in Intelligent and Soft Computing*. 2012, **149** : 503-509.