DOCUMENT WATERMARKING VIA CHARACTER LUMINANCE MODULATION

Paulo Vinicius Koerich Borges and Joceli Mayer

LPDS: Digital Signal Processing Laboratory
Department of Electrical Engineering
Federal University of Santa Catarina, UFSC, SC, Brazil
{paulo,mayer}@eel.ufsc.br

ABSTRACT

In this paper we propose a new method to embed hidden data in office-like documents, presenting robustness to the print-scan channel. The method slightly modulates the luminance of documents characters and symbols, being this change unnoticeable to the human eye, yet detectable by a scanner. It is based on pulse amplitude modulation communications, using "character shaped" pulses instead of traditional pulses, using text spacing to avoid intersymbol interference. We present an error analysis for the method, along with a new print-scan channel model for text documents. The analyses and applicability of the method are validated by experiments.

1. INTRODUCTION

Two-dimensional bar codes have gained great attention in the past few years, being attractive for many applications [1, 2]. Examples of their utilization include representing encrypted information, serving as an auxiliary authentication channel, or simply acting as a higher capacity version of traditional 1-D bar codes. An important application target for these codes is to serve as an authentication mark. Very often, important paper copies of documents are exchanged between companies and people. The challenge is to define a reliable method for authentication of hardcopy documents. We can use bar codes for this purpose whenever the resulting visual impact and waste of document space is acceptable.

In this work, we propose a system to embed information in printed form documents, presenting low perceptual impact. Besides serving as a source authentication proof, the hidden code can be used to authenticate the information of sensitive parts of a document (names, dates, or values, for example), by hiding this information over the whole document.

Several automated paper authentication techniques have been proposed in literature. When invisibility is an important requirement, printed form authentication techniques fall into the category of *hardcopy watermarking*, which are watermarking techniques able to survive the print-scan (PS) channel.

It is acknowledged that the bulk of documents in the office-like class are composed by black text on white background, being referred to as bi-level documents.

The proposed algorithm embeds information in documents composed by any kind of content, as long as it can be represented by binary text, logos, math symbols, and even line-drawings. In the proposed method, the watermark is embedded by allowing gray level tones to regular text while providing low perceptual impact. A gray level modulation is added to the original text and it is based on Pulse Amplitude Modulation (PAM), but instead of modulating square pulses [1], the characters and symbols themselves are modulated, acting as a random shape transmission pulse. The system is designed to survive the PS operation and visible marks are not required to achieve detection synchronization, a challenging task in hardcopy watermarking.

This paper is organized as follows: in Section 2 we discuss some related methods, pointing their advantages and drawbacks. In Section 3 we discuss and propose improvements for an existing PS model. In Section 4 we propose our technique, and analyze the resulting detection error probability. Experimental results are presented in Section 5, followed by the concluding remarks in Section 6.

2. EXISTING METHODS

Currently, only a relatively small number of papers addressing text watermarking able to survive the PS channel is found in the literature. Brassil et al. [6] developed different watermarking systems which are remarkably robust to the PS channel. One method is called line-shift coding, where a line is moved up or down according to the bit to be embedded. Line centroids can be used as references in blind detection (without the original document). The centroids are expected to be necessarily uniformly spaced, which in practice is not always true. Another method proposed in [6], coined word-shift coding, embeds information by slightly shifting words left or right. In another method, coined feature coding, certain text features like individual characters are moved up or down. These two last methods can embed more bits in comparison to line-shift coding, but the original document must necessarily be available for detection.

Huang and Yan [8] proposed a method based on the average inter-word distance in each line. The distances and the length of each word are adjusted to represent a sine wave with a given frequency and phase. Detection is blind and performed by correlation. Unfortunately, the line and word shifting techniques assume that the text has been correctly segmented. Equations, titles, and variable size logos or symbols complicate the coding process due to non-uniform spacing.

In a different approach, a common way to watermark text is by performing modifications on the characters pixels [9, 11], such as flipping a pixel from black to white (pixel toggling), and vice versa. In the work by Wu and Liu [9], for example, an image is partitioned into blocks and a fixed number of bits are embedded in each block by modifying some pixels in that block. A given pixel is altered depending on a "flipping score," which rates the perceptual impact of changing that pixel, determined dynamically by observing smoothness and connectivity. However, to survive the PS channel, the authors suggest the use of visible registration marks to compensate

This work was supported by CNPq, Proc. No 141078/2004-9.

for the synchronization distortions. Although this class of methods has a large capacity potential, the thresholding error in the binarization process after PS is a major source of error.

Similarly to this work, gray level luminances are added to regular text in [10]. However, only selected sub-character regions can be watermarked, limiting the embedding capacity, as mentioned by the authors. The line centroids need to be determined prior to detection, presenting the same problems of line and word shifting techniques.

3. PROPOSED PRINT AND SCAN CHANNEL

In general, the PS channel is represented by both geometric and pixel value distortions. [3, 4]. We model the pixel value distortions of the PS process by a non-linear gain (brightness/gamma alteration) followed by low-pass filtering and the addition of Gaussian noise:

$$y(m,n) = \{g[s(m,n)] + \eta_1(m,n)\} * h(m,n) + \eta_2(m,n) \quad (1)$$

where $s(m,n)$ and $y(m,n)$ are respectively the original and the print/scanned images, and η_1 and η_2 represent zero-mean white Gaussian noise. The linear system $h(m,n)$ models the system's point-spread-function, combining the printing and scanning operations. Hence, $h(m,n)$ acts as a low-pass filter due to the printing blur [5], to the optics, and to the motion blur caused by the interactions between adjacent CCD arrays elements [1].

We have modeled $h(m,n)$ as a Butterworth low-pass filter, described by

$$H(u,v) = \frac{1}{1 + [D(u,v)/D_0]^{2Q}} \quad (2)$$

where Q is the filter order, D_0 is the cutoff frequency, and $D(u,v)$ is the distance from point u,v to the origin (center) of the frequency spectrum. For this model, we have determined experimentally the optimum filter order Q and cut-off frequency D_0, by adjusting these parameters so that the error $E\{(y-\hat{y})^2\}$ is minimized, where $\hat{y}$ is the estimated y. An error surface is illustrated in Figure 1, and the values $Q = 1$ and $D_0 = 0.17$ are found at the minimum point.

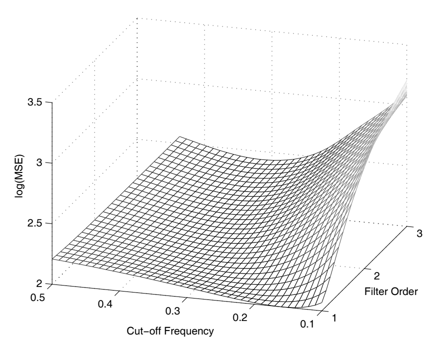

Fig. 1: Error performance surface.

Finally, the non-linear response, $g(\cdot)$, is typically written [1] as $g = \alpha s^\gamma$, where α and γ are constants and depend on the PS system being used. However, to account for an offset (that is, a measurable output when the input is zero), we propose to describe the distortion by

$$g(m,n) = \alpha[s(m,n) + \zeta]^\gamma \quad (3)$$

In the experiments, we have found that typical values for γ and ζ are 0.2 and 30, respectively, when operating in the region $s(m,n) \leq 100$. The parameter α is modeled as constant for a small region (an area corresponding to one fifth of a page, for example), but varies slowly throughout a full page.

Regarding the geometric distortions, it has been observed in [3, 4] and in our own experiments that: (i) When a picture is place on a flatbed scanner, a small misalignment is usually present, producing a rotation distortion. (ii) Not all PS systems work at the same resolution, causing re-scaling modifications. (iii) Prior to high resolution scanning, a cropping region is usually manually selected, causing a cropping distortion specially in text documents, where no evident boundary is present. In addition to the pixel value distortions, the geometric distortions can also interfere with watermark extraction, and the system must be designed to either compensate or be invariant to them.

4. PROPOSED METHOD

In the proposed method, each element in the original digital document is labeled as c_i, with $i = 1, 2, \ldots, K$, where K is the total number of elements. The elements are labeled from left to right, and from top to bottom. A simple luminance gain is added to individual text elements (characters, symbols, lines, etc), according to the bit value b_i, representing a possibly encrypted information to be embedded.

Information is embedded by individually altering the luminance of c_i through an embedding function where each element has its luminance modulated from black to any value in the real-valued discrete alphabet Ω of cardinality S, so that each symbol represents $\log_2 S$ bits. Considering an (m,n) spatial coordinate system for each element, $c_i(m,n)$ is modulated by a gain w_i, $w_i \in \Omega$. Working in the range $c_i(m,n) \in \{0,1\}$ and $w_i \in [0,1]$, consider a modified luminance representation, where the white background level is 0 and the black regions are level 1. The general embedding function is given by:

$$s_i(m,n) = b_i w_i c_i(m,n) \quad (4)$$

The process is illustrated in Figure 2 for $S = 2$, with a very high gain. The proposed system resembles the well-known pulse amplitude modulation (PAM) baseband transmission technique. In particular, when we use only two levels of luminance, the coding scheme is similar to PCM - Unipolar RZ (return-to-zero). The RZ term comes from the fact that between characters we have blank spaces, where no symbol is transmitted, which also avoids ISI (inter-symbolic interference). In contrast to what is illustrated in Figure 2, the luminance gain is chosen to cause low perceptual impact, such that $E\{c(m,n) - s(m,n)\} \leq \epsilon_{max}$. We do not need to use a squared error metric because $c(m,n)$ is known, constant and $c(m,n) \leq s(m,n)$.

$c(m,n)$: S I G N A L

c_1 c_2 $\cdots$ c_K

$\mathbf{b} = [\,0 \;\; 1 \;\; 1 \;\; 0 \;\; 1 \;\; 0\,]$

$\mathbf{w} = [\,0 \;\; A \;\; A \;\; 0 \;\; A \;\; 0\,]$

$s(m,n)$: S I G N A L

Fig. 2: Luminance modulation with visible gain.

The printed document contains a hidden, possibly encrypted, bit string. The decoder scans the document using resolution S_r and the elements are segmented from the background. Again, each element is labeled as $c_1, c_2, \ldots, c_K$, where K is the number of elements in text. When the paper is properly placed in the flatbed scanner, the slight rotation induced by the scanning process does not compromise the labeling process. We use the mean luminance value of each c_i for the bit detection statistic, described in the following.

Unlike 2-D bar codes, which have an embedding capacity expressed in bits/in^2 [2], the capacity of our system depends on the size of the elements, and is expressed in bits/element.

4.1. Bit Error Rate

Consider the 1 bit/element, or $S = 2$, case. By mapping the (m, n) coordinates to an one-dimensional notation, our detection metric d_i for element i is given by:

$$d_i = \frac{1}{N} \sum_{n=1}^{N} y_i(n) \tag{5}$$

where N is the number of pixels in element i and $y(n)$ is described by (1). The detection statistic is given as:

$$\begin{aligned} d_i &= \frac{1}{N} \sum_{n=1}^{N} \{g[s_i(n)] + \eta_1(n)\} * h(n) + \eta_2(n) \\ &= \frac{1}{N} \left\{ \sum_{n=1}^{N} g[s_i(n)] * h(n) + \sum_{n=1}^{N} \eta_1(n) * h(n) + \right. \\ &\left. \sum_{n=1}^{N} \eta_2(n) \right\} \end{aligned} \tag{6}$$

The term $g[s(n)] * h(n)$ in the above equation represents an offset (see (3)) that is modeled as constant for a given element c_i. However, it does vary slightly throughout a full page due to non-constant scanner illumination. We represent this variation by Δ_i, which acts as an offset variable from element to element. Thus we let $g[s(n)] * h(n) = A + \Delta_i$, where A depends on the bit value and Δ is a random variable modeled as $\mathcal{N}(0, \sigma_\Delta^2)$. In the following notation, we express $A = A_1$ for bit 1, and $A = A_0$ for bit 0.

Initially considering that a bit 1 has been transmitted, the expected value of equation (6) is $\mu_d = E\{d\} = A_1$, and the variance is given by:

$$\begin{aligned} \sigma_d^2 &= E\{(d - \mu_d)^2\} \\ &= E\left\{ \left(\frac{1}{N} \sum_{n=1}^{N} (A_1 + \Delta) + \frac{1}{N} \sum_{n=1}^{N} \eta_1(n) * h(n) + \right.\right. \\ &\left.\left. \frac{1}{N} \sum_{n=1}^{N} \eta_2(n) - A_1 \right)^2 \right\} \\ &= E\left\{ \left(\Delta + \frac{1}{N} \sum_{n=1}^{N} \eta_1(n) * h(n) + \frac{1}{N} \sum_{n=1}^{N} \eta_2(n) \right)^2 \right\} \\ &= E\{\Delta^2\} + \frac{1}{N^2} E\left\{ \sum_{n=1}^{N} \eta_1(n) * h(n) \sum_{m=1}^{N} \eta_1(m) * h(m) + \right. \\ &\left. 2 \sum_{n=1}^{N} \eta_1(n) * h(n) \sum_{m=1}^{N} \eta_2(m) + \sum_{n=1}^{N} \eta_2(n) \sum_{m=1}^{N} \eta_2(m) \right\} \end{aligned} \tag{7}$$

Assuming that η_1 and η_2 are zero-mean independent white Gaussian random variables, we find

$$\sigma_d^2 = \sigma_\Delta^2 + \frac{1}{N^2} \sum_{n=1}^{N} \sum_{m=1}^{N} E\left\{[\eta_1(n) * h(n)][\eta_1(m) * h(m)]\right\} + \frac{\sigma_{\eta_2}^2}{N} \tag{8}$$

If we let $z(n) = \eta_1(n) * h(n)$ we have

$$\sigma_d^2 = \frac{1}{N^2} \sum_{n=1}^{N} \sum_{m=1}^{N} E\{z(n)z(m)\} + \frac{\sigma_{\eta_2}^2}{N} + \sigma_\Delta^2 \tag{9}$$

$$= \frac{1}{N^2} \sum_{n=1}^{N} \sum_{m=1}^{N} R_z(m, n) + \frac{\sigma_{\eta_2}^2}{N} + \sigma_\Delta^2 \tag{10}$$

where $R_z(m, n)$ is the autocorrelation function at the blurring filter output. We may write $R_z(m, n) = r_h(m, n) * r_{\eta_1}(m, n)$, by observing the output properties of a linear system with random input [12]. The variables r_{η_2} and r_h represent the autocorrelation functions of η_2 and of the impulse response of h, respectively. Therefore,

$$\sigma_d^2 = \frac{1}{N^2} \sum_{n=1}^{N} \sum_{m=1}^{N} r_h(m, n) * r_{\eta_1}(m, n) + \frac{\sigma_{\eta_2}^2}{N} + \sigma_\Delta^2 \tag{11}$$

Since η_1 is modeled as uncorrelated noise, $r_{\eta_1}(m, n)$ is represented by an impulse at the origin with amplitude $\sigma_{\eta_1}^2$, and equation (11) becomes

$$\sigma_d^2 = \frac{\sigma_{\eta_1}^2}{N^2} \sum_{n=1}^{N} \sum_{m=1}^{N} r_h(m, n) + \frac{\sigma_{\eta_2}^2}{N} + \sigma_\Delta^2 \tag{12}$$

The conditional error probability p_{01} given that bit 1 was transmitted is described by $p_{01} = \Pr(d_i < \lambda | b_i = 1)$. Defining the complementary error function $\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$, we may write $p_{01} = \frac{1}{2}\text{erfc}\left(\frac{A_1 - \lambda}{\sqrt{2\sigma_d^2}}\right)$ where λ is a decision threshold. Similarly, if bit 0 was transmitted, the conditional error probability is given by $p_{10} = \frac{1}{2}\text{erfc}\left(\frac{\lambda - A_0}{\sqrt{2\sigma_d^2}}\right)$. Finally, the average error probability is expressed by $P_e = p_0 p_{10} + p_1 p_{01}$, where p_0 and p_1 are the probabilities of occurrence of bits 0 and 1, respectively. We note that the analysis above is not only valid for the system proposed in this paper, but for any RZ printed 2-D bar code, like the one proposed in [1].

Intersymbol interference (ISI) is another important source of error in the PS channel, due to its low-pass effect. However, in our method, text spacing avoids the smearing of one symbol over another, and the ISI effect can be disregarded.

5. EXPERIMENTS

During the experiments, we have found that typical values for $\sigma_{\eta_1}^2$ and $\sigma_{\eta_2}^2$ are 100 and 60, respectively. Using these noise parameters and the Butterworth filter determined in Section 3, a synthetically generated character represents with fidelity a print-scanned character.

We have conducted experiments with printers HP IJ-855C, HP IJ-870Cxi and HP LJ-1100, and scanners Genius HR6X, HP 2300C and HP SJ-5P.

In our first experiment, a very large sequence of a single character, as in 'aaaa...', is printed, with no luminance alteration. The average luminance value for the scanned characters was $A_0 = 98.72$, which cannot be calibrated. The same sequence is then printed with a luminance gain $w_i = 25$, leading to an average luminance value $A_1 = 112.97$. Clearly A_1 is a controlled variable and depends on w_i. The distortion is expressed as $\epsilon = A_1 - A_0$ and to achieve perceptual transparency in printed paper form, we need $\epsilon \lessapprox 17$. Figure 3 illustrates the obtained PDFs for the detection metric using $w_i = 0$ (black character) and $w_i = 25$ ($\epsilon = 112.97 - 98.72 = 14.25$).

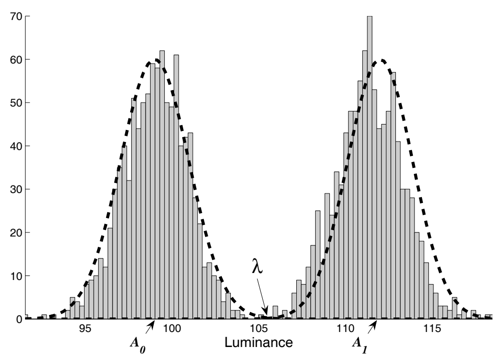

Fig. 3: Detection PDF's.

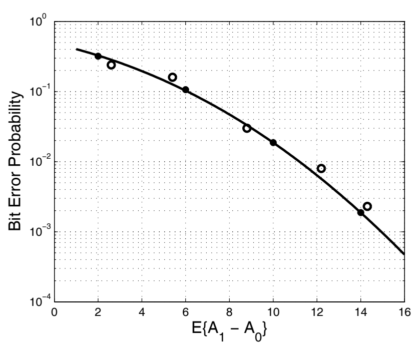

Fig. 4: Bit error probability for different gains.

The plot in Figure 4 shows the similarity between the theoretical (full line) and the experimental (circle marks) bit error rates as a function of the difference $A_1 - A_0$. This example validates the analysis and parameters used. The small dots over the full line represent the error rates from simulations using the PS model of equation (1).

6. CONCLUSIONS

In this paper we have presented and analyzed a simple yet innovative method to transmit hidden information in text documents. The method provides a much lower perceptual impact than 2-D bar codes. It is a document hardcopy watermarking system suitable to any kind of character/symbols/alignments, very robust, and extremely easy to implement. An improved channel model for the print-scan process has been proposed. This model was fully included in the error probability analysis, which is valid not only for our system, but for any 2-D RZ bar code. Experiments validate these analyses and illustrate the direct applicability of the method. Further work includes detection improvement by pre-processing and by using other metrics, in addition to considering the halftoning effects.

7. REFERENCES

[1] N. D. Quintela and F. Pérez-González. "Visible encryption: Using paper as a secure channel." In *Proc. of SPIE*, Santa Clara, USA, January 2003.

[2] R. Víllan, S. Voloshynovskiy, O. Koval, and T. Pun, "Multilevel 2D bar codes: towards high capacity storage modules for multimedia security and management," in *Proc. of SPIE*, Electronic Imaging 2005, San Jose, USA, January 16-20 2005.

[3] K. Solanki, U. Madhow, B.S. Manjunath, S. Chandrasekaran "Modeling the print-scan process for resilient data hiding," in *Proceeding of SPIE*, Electronic Imaging, Vol. 5681, 2005.

[4] Ching-Yung Ling, "Public watermarking surviving general scaling and cropping: an application for print-and-scan process," in *Multimedia and Security Workshop at ACM Multimedia 99*, October 1999.

[5] M. Norris and E. H. B. Smith "Printer modeling for document imaging," in *Proc. of the Int'l Conf. on Imaging Science, Systems and Technology*, CISST '04, June 2004, Las Vegas, USA.

[6] J.T. Brassil, S. Low, N.F. Maxemchuk "Copyright protection for the electronic distribution of text documents," *Proc. of IEEE*, Volume 87, No. 7, pp. 1181-1196, July 1999.

[7] S. Low and N.F. Maxemchuk, "Capacity of text marking channel," in *IEEE Signal Process. Letters*, Volume 7, No. 12, December 2000.

[8] D. Huang and H. Yan, "Interword distance changes represented by sine waves for watermarking text images," in *IEEE Trans. on Circuits and Systems for Video Technology*, Volume 11, No. 12, pp. 1237-1245, Dec. 2001.

[9] Min Wu and Bede Liu, "Data hiding in binary image for authentication and annotation," in *IEEE Trans. on Multimedia*, August 2004.

[10] A. K. Bhattacharjya and H. Ancin, "Data embedding in text for a copier system," *Proc. of IEEE Int'l Conf. on Image Processing*, ICIP 99. Volume 2, 1999.

[11] T. Amano, "A feature calibration method for watermarking of document images," in *IEEE Proc. of the Fifth Int'l Conf. on Document Analysis and Recognition,* ICDAR '99. 20-22 Sept. 1999.

[12] D. Manolakis, V. Ingle, S. Kogon *Statistical and Adaptive Signal Processing*, McGraw-Hill, 2000.