A NEW SPREAD SPECTRUM WATERMARKING METHOD WITH SELF-SYNCHRONIZATION CAPABILITIES

I. Mora - Jiménez, Student Member, IEEE, and A. Navia - Vázquez, Member, IEEE

Department of Communication Technologies, Universidad Carlos III de Madrid
Avda de la Universidad 30, 28911 Leganés-Madrid, Spain
inmoji@tsc.uc3m.es, navia@tsc.uc3m.es

ABSTRACT

Among the many techniques available for information concealment, those based on spread spectrum modulations have proven to yield improved results when robustness against attack is at a premium. In this paper, we propose a new spread spectrum-based watermarking procedure that combines space and frequency marks to provide good robustness properties against both spatial (affine) and transform-based compression attacks, without needing the original image as a reference (blind detection). It provides a mechanism for N-bit concealment and also improves the detection-of-presence process by gathering all watermark energy into a single value (sufficient statistic for detection). The recovery of every single bit is also improved by taking into account the so-called "watermark-print" or "waterprint" instead of looking at a single correlation value. It additionally provides the means to recover synchronization under affine transformations in the blind detection scenario. These characteristics will be analyzed by means of several practical examples.

1. INTRODUCTION

Many watermarking algorithms relying on Spread Spectrum (SS) have been proposed in the literature, adding the watermark in either spatial [1][2][3] or transformed domains [7][4][5][6], among many other. None of them combines both techniques, even knowing that synchronization (recovery from geometric affine transformations [1]) can be more easily achieved in the spatial domain, while frequency methods are usually more resilient against strong image compression and smoothing (usually attained retaining only a few coefficients of a transformed image). Nevertheless, it seems interesting to diversify the watermark in both spatial and frequency domains such that the maximum watermark power is retained under a worst attack scenario. This diversification has to be done carefully, in order not to lose detection capability. In order to maximize the probability of detection, all this watermark energy should be converted back into a single sufficient statistic ρ. At the same time, it is interesting to deal with N-bit watermarks, not just the binary presence/abscense case. Unfortunately, many of the proposed schemes for N-bit concealment decode every bit separately (using either N different watermarks [8] or splitting the image into N different regions [1]), both reducing the power available for every bit. In this paper we present an SS modulation scheme which provides N-bit information concealment as well as an unified sufficient statistic for detection. It also provides means for self-synchronization in the absence of the original image. The outline of the paper is as follows: in Section 2 we describe the proposed scheme for both detection-of-presence and bit recovery. In Section 3 we present the self-synchronization mechanism, in Section 4 we conduct several experiments and we close the paper in Section 5 with some conclusions.

2. THE AMPLITUDE SS MODULATION AND THE DETECTION MECHANISM

The general SS watermarking insertion procedure, interpreted as a communication process, is presented in Figure 1,

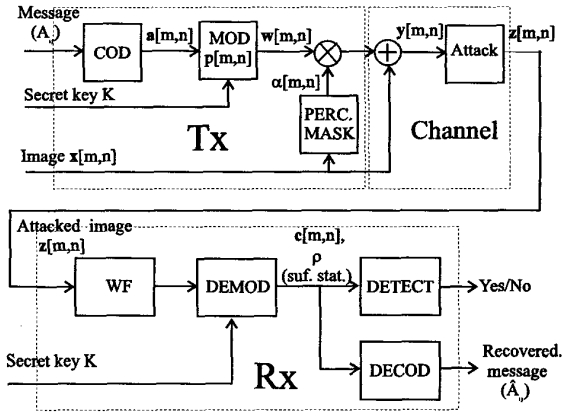


Fig. 1. Watermark insertion and recovery scheme.

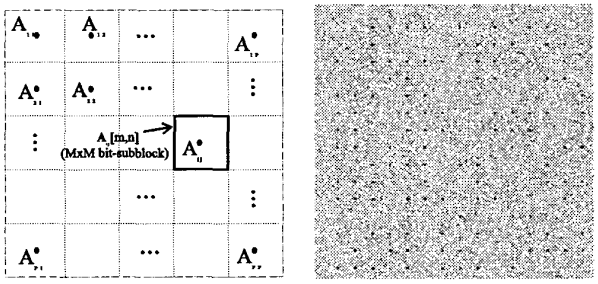
0-7803-6297-7/00/$10.00 © 2000 IEEE

It can be observed that the watermark $\mathbf{w}[m,n]$ is generated using a key-dependent pseudorandom sequence, weighted by a perceptual mask (to attain imperceptibility) and finally added to the image. After applying whitening filter (WF), implemented as an Adaptive Wiener Filter (AWF), the watermarked image is correlated with the key-dependent demodulating signal (frequency de-spreading). Based upon this correlation image $\mathbf{c}[m,n]$ the receiver performs two tasks: watermark detection and bit recovery.

Some other mechanisms using SS amplitude modulation have been proposed in the literature, but the usual approach is to use non-overlapping regions of the image to hide information concerning every different bit [1][2][3]. We believe that proceeding this way, the whole scheme becomes much more sensitive to synchronization loss and blurring, because information concerning different bits may be mixed or lost (the so called patchwork-like methods are also very sensitive to this effect [9][10][11]). We propose to hide every bit over the entire image ("holographic" property, following [12]), taking advantage of the orthogonality properties between a pseudorandom sequence $\mathbf{p}[n]$ and any shifted version of it $\mathbf{p}[n-kN]$. Information bits $A_{j,k}\{0,1\}$ are firstly represented as equally-spaced pulses, and then modulated using $p[n]$. The base watermark is, therefore, constructed using

$$\mathbf{w}[m,n] = \mathbf{p}[m,n] \otimes \sum_{j,k} A_{j,k}\delta[m-jM, n-kM] \quad (1)$$

where $\otimes$ denotes circular convolution. The proposed modulation scheme is as shown in Figure 2(a), and the output of the correlation receiver for the 256-bit case is depicted in (b).

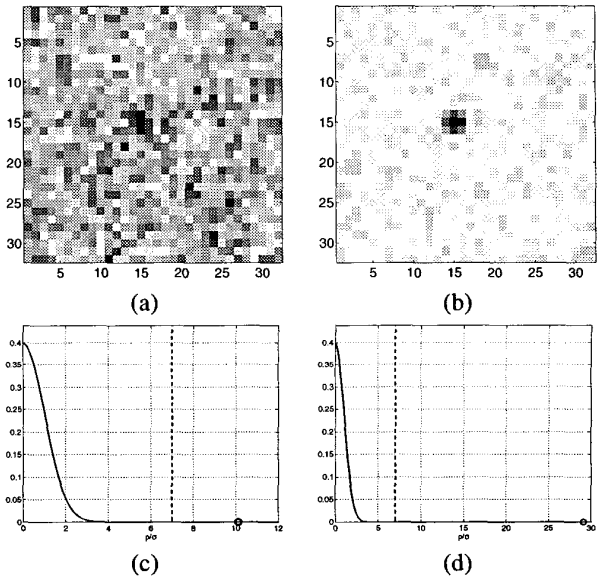


(a) (b)

Fig. 2. Proposed modulation scheme: location of bits before transmitting (a) and correlation image $\mathbf{c}[m,n]$ in the receiver (b).

Note that every pulse $A_{j,k}$ is centered in a different $M \times M$ image subblock, as shown in figure 2(a), the modulating function $\mathbf{p}[m,n]$ spreads every one of them over the entire image, such that the watermark will look like random noise. This basic signal $\mathbf{w}[m,n]$ will be inserted in both spatial (multiplied by the corresponding perceptual mask) and transformed domains (main coefficients of the Discrete Cosine Transform (DCT)). In the receiver, after applying a whitening filter (Adaptive Wiener Filter) and correlating the watermarked image with a reproduction of the watermark (using key K), we obtain a correlation image $\mathbf{c}[m,n]$ with peaks located at points where $A_{j,k}$=1, as shown in figure 2(b) for the particular case of $N = 256$ bits. The bit recovery problem is to spot these peaks in the noisy correlation image.

One of the advantages of this scheme is that, for detection of watermark presence purposes, the correlation image can be blockwise ($M \times M$) averaged, such that peaks associated to bits reinforce each other while incorrelated information (noise variance is reduced). In figure 3(a) we show one of these subblocks, and in (b) we see the result of block averaging.



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

Fig. 3. Detection scenario: comparison between direct (a)(c) and averaged decision (b)(d).

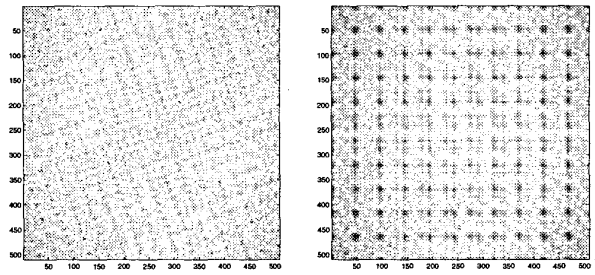
It can be observed that the background noise power (σ_n^2) in (a) is larger than in (b). When testing for the presence of watermark, the maximum peak value has to be compared with σ_n, and if it exceeds some threshold (here chosen to be $7\sigma_n$), we decide that the watermark is present. The decision scenario shown in figure 3(d) ($\rho/\sigma_n \approx 28$, marked as 'o') is, therefore, superior to that in (c) ($\rho/\sigma_n \approx 10$). The averaging mechanism improves the robustness of the watermark detection test and, therefore, reduces the risk of both false positives and detection misses.

Many of the attacks tend to spread the watermark en-

ergy over the correlation image such that, in the receiver, we obtain what we call a "waterprint" (blurred peak, so to speak) instead of a sharp peak; in those cases the message may be lost if a single value of $\mathbf{c}[m,n]$ is used for the decision. We can use the average waterprint (see figure 3(b) as an example) to infer where the watermark energy is concentrated, information that can finally be used to improve the bit recovery process, by converting the watermark energy back to a single value (integration of energy). The 'out-of-mask' noise is used to estimate σ_n such that an appropriate threshold can be selected. This is the procedure used in the experiments, and its performance will be analyzed in the experimental section.

3. SELF-SYNCHRONIZATION MECHANISM

One of the common attacks (either intentional or not) consists in geometric (affine) manipulations suffered by the image (synchronization loss) which, unless corrected, may lead to a total failure in detection even for small distortions. Due to the symmetry properties of the proposed modulation scheme, it is possible to infer the nature of this type of attack, by simply observing the autocorrelation function of the watermarked image (after whitening). Peaks associated to bits appear arranged in a grid which reveals the affine transformation suffered by the image. See, for instance, Figure 4 where the autocorrelation function of the marked image under two situations is shown: (a) anti-clockwise rotation of 20 degrees and (b) resizing by 2/3. Once the affine manipulation has been blindly estimated (without needing the original image), it is possible to adequately revert the process before proceeding as in Section 2.



(a) (b)

Fig. 4. Sync recovery signals: 20 degree rotation (a) and 2/3 re-scaling (b).

Unfortunately, this sync information may not survive some of the common attacks, but this process can still be used in controlled scenarios where an intentional hard attack is not expected.

4. EXPERIMENTAL RESULTS

We present in this section some preliminary results concerning robustness of the proposed method. We have embedded the watermark using our scheme into "lena" image (512×512) using a perceptual mask and such that the image quality is about 4.5 (following the ITU-R Rec. 500, where level 5 means excellent/imperceptible, and level 4 is good/perceptible, not annoying). In table 1 we detail the results after applying most common attacks (with parameters such that the image quality is not severely damaged).

Table 1. Attacks performed to test robustness, 64 bits encoded. Cases (15)-(20) represent combined attacks.

No	Attack	Watermark Detected?	No err. bits
1	AWGN	yes	0
2	Median(3×3)	yes	0
3	Median(5×5)	yes	0
4	2^4 Quantiz. levels	yes	0
5	2^3 Quantiz. levels	yes	0
6	AWF(5×5)	yes	0
7	AWF(15×15)	yes	0
8	JPEG 50%	yes	0
9	JPEG 80%	yes	0
10	Cropping 90%	yes	0
11	Dithering	yes	0
12	Decim.($\downarrow$2)+interp.($\uparrow$2)	yes	0
13	Smooth(5×5)	yes	0
14	Smooth(15×15)	yes	8
15	"(3)+(6)"	yes	2
16	"(2)+(8)"	yes	0
17	"(4)+(10)"	yes	1
18	"(12)+(2)"	yes	0
19	"(2)+(4)+(8)"	no	-
20	"(12)+(2)+(4)"	yes	0

It can be observed how the watermark survives most attacks, even when some of them are combined (cases (15)-(20)). In order to further test robustness, we apply some severe attacks to the image, and the results have been collected in Figure 5 where, in every upper left corner, is also shown the associated waterprint.

Finally, the capacity of lena's image (grey-scale) using this scheme image has been estimated to be of 4096 bits, this is roughly the number of bits we have been able to conceal in it (under no attack scenario). This number can serve as a guide for those applications where robustness is not the main goal, but transmitting as many bits as possible

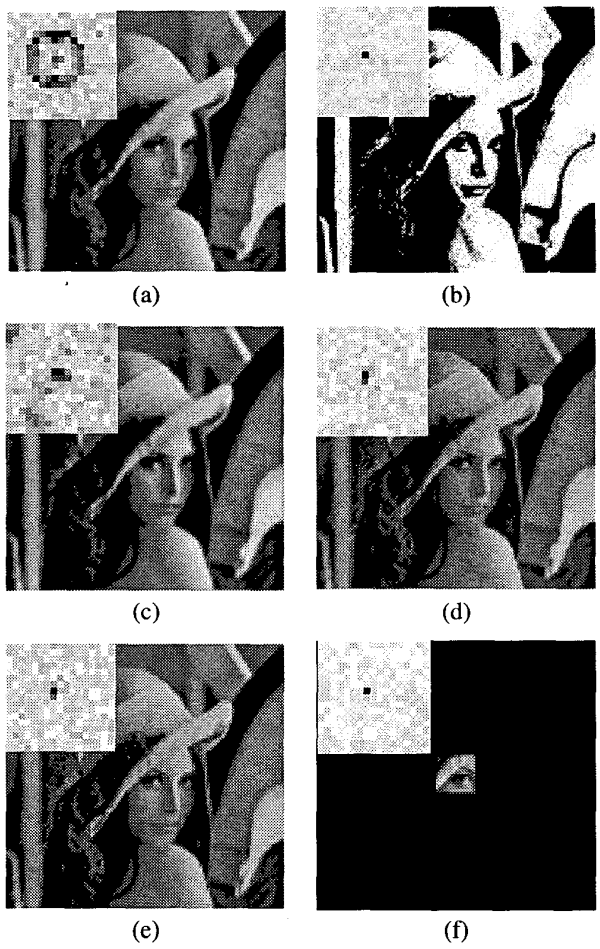


Fig. 5. Illustration of some extreme attacks and their effect on the watermark-print (upper-left corner): (a) Median 7×7, (b) binarization, (c) smoothing 15 × 15, (d) dithering, (e) Jpeg 10% quality, and (f) cropping 90%.

5. CONCLUSIONS

In this paper we have proposed a new spread spectrum watermarking scheme with three main benefits. Firstly, the watermark is distributed in both spatial and frequency domains, such that resistance against attacks is maximized. Secondly, a mechanism is proposed, such that all the diversified watermark energy can be used to build a single presence-of-watermark test, while allowing the storage of many bits, and its improved recovery by means of the waterprint. Thirdly, the particular geometry of the modulating signal provides a method for self-synchronization when the original image is not available. Robustness of the proposed method is tested using standard attacks, and seems to perform quite well under rather severe conditions.

6. REFERENCES

[1] J. R. Hernández and F. Pérez-González, "Statistical analysis of watermarking schemes for copyright protection of images", in *Proceedings of the IEEE*, vol.87, pp. 1142-1166, 1999.

[2] M. Kutter, F. Jordan, and F. Bossen, "Digital signature of color images using amplitude modulation", in *Proc. SPIE-EI97*, pp. 518-526, 1997.

[3] C. Langelaar, J. C. A. ven der Lubbe, and R. L. Lagendijk, "Robust labeling methods for copy protection of images", in *Proc. Electronic Imaging*, San Jose, CA, vol. 3022, pp. 298-309, 1997. [Online]. Available at http://www-it.et.tudelft.nl/~gerhard/home.html.

[4] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain system for robust image watermarking", in *Signal Processing*, vol. 66, no. 3, pp. 357-372, 1998.

[5] A. Piva, M. Barni, F. Bartolini, and V. Cappellini, "DCT-based watermark recovering without resorting to the uncorrupted original image", in *Proc. Intl. Conf. on Image Processing*, vol. 1, pp. 520-523, 1997.

[6] Chiou-Ting-Hsu, and Ja-Ling-Wu, "Hidden digital watermarks in images", in *IEEE Transactions on Image Processing*, vol.8, no.1, pp. 58-68, 1999.

[7] R.B. Wolfgang, C.I. Podilchuk, and, E. J. Delp, "Perceptual Watermarks for digital images and video", *Proc. of the IEEE*, vol.87, no. 7, pp. 1108-1126, 1999.

[8] G.B. Rhoads, "Stenography methods employing embedded calibration data", United States Patent # 5.636.292, 1997.

[9] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding", *IBM Syst. J.*, vol.35, nos. 3 and 4, pp. 313-336, 1996.

[10] C. Dautzenberg and F. Boland, "Watermarking images", Dept. of Electrical Engineering, Trinity College, Dublin, Tech. Rep., 1994.

[11] O. Bruyndockx, J. J. Quisquater, and B. Macq, "Spatial method for copyright labeling of digital images", in *Proc. IEEE Nonlinear Signal Processing Workshop*, 1995, pp. 456-459.

[12] A.M. Bruckstein, and T.J. Richardson, "A holographic transform domain image watermarking method", *Circuits, Systems, and Signal Processing*, vol.17, no.3, pp.361-389, 1998.