IMPROVED EMBEDDING OF MULTIPLICATIVE WATERMARKS VIA SPACE-TIME BLOCK CODING

Irene G. Karybali and Kostas Berberidis

Dept. of Computer Engineering and Informatics and CTI/R&D
University of Patras, 26500 Rio-Patras, Greece
phone: +30 2610960425, fax: +30 2610960374, emails: {karybali, berberid}@ceid.upatras.gr

ABSTRACT

In this paper, a new scheme for image watermarking in spatial domain based on space-time block coding is proposed. Specifically, a 4×4 real orthogonal design is employed for embedding a multiplicative watermark in the image. The image is divided into four blocks, which, after some simple operations, can be viewed as four different flat fading channels. At the receiver's end, a low cost maximum likelihood decoding is performed based only on linear processing. This scheme turns out to perform much better than repetitive watermarking, taking advantage of the well-known merits of space-time block coding, i.e., it achieves full diversity offering at the same time the maximum possible transmission rate for real constellations.

1. INTRODUCTION

Copyright protection and authentication of digital data via watermarking is an issue of intensive research worldwide, in recent years. There are some contradictory requirements in the watermarking process. Watermark robustness, invisibility and sufficient informative capacity are simultaneously required.

Watermark repetition as an encoding process along with perceptual masking at embedding are known to substantially increase robustness. However, there are some disadvantages from the encoding efficiency viewpoint and a compromise has to be found between watermark redundancy and coding efficiency. Possible solutions to this compromise are diversity techniques or sophisticated encoding approaches such as trellis-coded modulation [1].

In [2], diversity through watermark repetition is used. A non-stationary parallel binary symmetric channel model is introduced and channel estimation is achieved through a reference watermark.

The authors of [1] consider diversity with an interperiod optimal signal encoding based on iterative codes such as turbo codes or low-density parity check codes, with the decoder performing watermark channel state estimation based on a reference pilot watermark. Binary phase shift keying modulation is used at embedding, while a generalized watermarking channel is considered that includes geometrical attacks, fading and additive non-Gaussian noise.

Li et al. in [3] present a watermarking technique in spatial domain based on a particular form of transmit diversity. More specifically, the RGB components of color images are considered as three independent slowly fading channels. The same watermark in different forms, modulated by different pseudorandom sequences and interleavers, is put into the three color channels simultaneously. The extraction is based on comparing the energy strengths of the three channels or accumulating the energy of the three channels. The corresponding process in DCT domain is presented in [4].

In this paper we present a novel watermarking technique motivated by the so-called space-time block coding (STBC) which has been relatively recently proposed in wireless communications as a transmit diversity technique, [5], [6]. STBC generalizes the transmission scheme proposed by Alamouti [7] to an arbitrary number of transmit antennas and is able to achieve the maximum (full) diversity that can be provided by a specific configuration of transmit and receive antennas. This scheme is much less complex than space-time trellis coding and it does not require any channel state information feedback from the receiver to the transmitter. Moreover, STBC schemes employ a very simple maximum likelihood decoding algorithm based only on linear processing at the receiver. Furthermore, for real signal constellations (such as PAM), they provide the maximum possible transmission rate allowed by space-time coding theory [8].

In the proposed watermarking scheme the case of four transmitters and one receiver is investigated, although extension to other configurations is possible following a similar procedure. The image is divided into four blocks that after some simple manipulation in spatial domain can be interpreted as four different flat fading channels. The watermark is multiplicatively embedded in the image according to a 4×4 real orthogonal design [5]. At the receiver, maximum likelihood decoding is achieved in a simple way. A proper matched filter is imposed in the received signal that decouples the transmitted signals. A linear correlation is subsequently used for the watermark detection. It is proved that the proposed watermarking scheme has superior performance as compared to the repetitive watermarking and is robust to many different attacks as additive white gaussian noise, filtering, JPEG compression, etc.

In Section 2, the problem is formulated and the channel model is presented. The proposed watermarking scheme (embedding and detection) is described in Section 3, and its improved performance is justified. Experimental results are provided in Section 4 and finally, in Section 5, the work is concluded and further research directions are discussed.

2. PROBLEM FORMULATION AND THE CHANNEL MODEL

The image to be watermarked is divided into four blocks and each one corresponds, after some operations, to a different

This work was supported by the European Social Fund (ESF), Operational Program for Educational and Vocational Training II (EPEAEK II / Program IRAKLEITOS), and the Research Academic Computer Technology Institute of Patras.

flat fading channel. The embedding is multiplicative and the watermarked image is given by

$$y_i = x_i + x_i w_i a = x_i(1 + w_i a) \quad (1)$$

where a is a scalar that determines the watermark's strength and $i = 1,2,3,4$ denotes the image block index. If repetitive watermark embedding with four blocks was considered, then the four watermarks w_i would be equal to a single watermark generated by a zero-mean pseudorandom generator of $\{-1,1\}$ using a secret key. As will be shown in next section, in the proposed scheme the four watermarks are different but properly related in accordance to the STBC concept.

The cover image x can be written as a sum of its local mean value and the corresponding error as

$$x_i = m_i + e_i \quad (2)$$

where e_i is a zero mean sequence. Using (2) in (1) we obtain

$$y_i = m_i(1 + w_i a) + e_i(1 + w_i a) \quad (3)$$

with m_i being a constant for each block. As will be seen next, the mean value of each block can be viewed as the attenuation parameter of the corresponding flat fading channel. The error sequence e_i can be considered as noise. We also assume that white gaussian noise, n, with zero mean is added.

3. THE PROPOSED ALGORITHM

The concept of diversity has been extensively studied in wireless communications and proved to be a very powerful tool for improving wireless systems' performance. Since watermarking can be actually considered as a communications problem, we seek here a means to improve watermark detection by applying transmit diversity based on orthogonal codes. For simplicity, we adopt a configuration with four transmitters and one receiver. Thus, the following 4×4 real orthogonal design [5] is properly used

$$\begin{pmatrix} w(4k+1) & w(4k+2) & w(4k+3) & w(4k+4) \\ -w(4k+2) & w(4k+1) & -w(4k+4) & w(4k+3) \\ -w(4k+3) & w(4k+4) & w(4k+1) & -w(4k+2) \\ -w(4k+4) & -w(4k+3) & w(4k+2) & w(4k+1) \end{pmatrix} \quad (4)$$

where for a given k the i-th column contains four consecutive data symbols transmitted by the corresponding i-th transmitter. Note that the same four symbols (with different ordering and signing) are transmitted by all transmitters.

3.1 Watermark Embedding

Using the orthogonal design of (4), the original watermark is "transformed" into four different watermarks that are embedded in the image blocks as shown in Fig. 1. The data to be transmitted via the four different channels (blocks) are $w(4k+1)$, $w(4k+2)$, $w(4k+3)$ and $w(4k+4)$, where $k = 0,1,...,\frac{N^2}{16} - 1$ and N^2 is the number of image pixels. Each image block contains $N^2/4$ pixels taken row-wise. Obviously, the number of watermark bits to be embedded is also equal to $N^2/4$. Thus, the four different watermarks are properly related in accordance to the STBC concept.

By adding the four image blocks, the received signals corresponding to indices $4k+1$, $4k+2$, $4k+3$ and $4k+4$ are

$$r(4k+1) = a[m_1 w(4k+1) + m_2 w(4k+2) + m_3 w(4k+3) + m_4 w(4k+4)] + \varepsilon(4k+1) + (m_1+m_2+m_3+m_4) + n(4k+1) \quad (5)$$

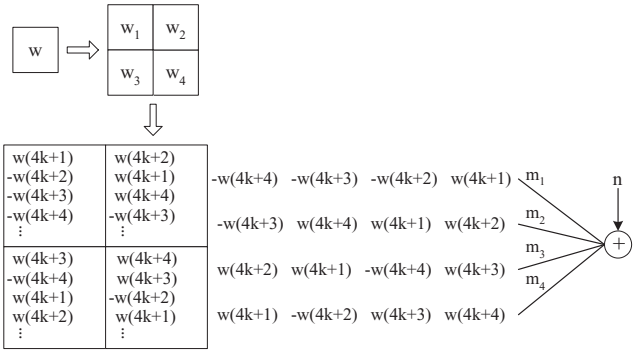


Figure 1: Proposed watermark embedding.

where $\varepsilon(4k+1) = e(4k+1)[1+w(4k+1)a] + e(4k+2)[1+w(4k+2)a] + e(4k+3)[1+w(4k+3)a] + e(4k+4)[1+w(4k+4)a]$,

$$r(4k+2) = a[m_2 w(4k+1) - m_1 w(4k+2) + m_4 w(4k+3) - m_3 w(4k+4)] + \varepsilon(4k+2) + (m_1+m_2+m_3+m_4) + n(4k+2) \quad (6)$$

where $\varepsilon(4k+2) = e(4k+2)[1+w(4k+1)a] + e(4k+1)[1-w(4k+2)a] + e(4k+4)[1+w(4k+3)a] + e(4k+3)[1-w(4k+4)a]$,

$$r(4k+3) = a[m_3 w(4k+1) - m_4 w(4k+2) - m_1 w(4k+3) + m_2 w(4k+4)] + \varepsilon(4k+3) + (m_1+m_2+m_3+m_4) + n(4k+3) \quad (7)$$

where $\varepsilon(4k+3) = e(4k+3)[1+w(4k+1)a] + e(4k+4)[1-w(4k+2)a] + e(4k+1)[1-w(4k+3)a] + e(4k+2)[1+w(4k+4)a]$, and

$$r(4k+4) = a[m_4 w(4k+1) + m_3 w(4k+2) - m_2 w(4k+3) - m_1 w(4k+4)] + \varepsilon(4k+4) + (m_1+m_2+m_3+m_4) + n(4k+4) \quad (8)$$

where $\varepsilon(4k+4) = e(4k+4)[1+w(4k+1)a] + e(4k+3)[1+w(4k+2)a] + e(4k+2)[1-w(4k+3)a] + e(4k+1)[1-w(4k+4)a]$, respectively.

If we introduce the vectors

$$\mathbf{r} = \begin{bmatrix} r(4k+1) \\ r(4k+2) \\ r(4k+3) \\ r(4k+4) \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} w(4k+1) \\ w(4k+2) \\ w(4k+3) \\ w(4k+4) \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon(4k+1) \\ \varepsilon(4k+2) \\ \varepsilon(4k+3) \\ \varepsilon(4k+4) \end{bmatrix},$$

$$\mathbf{n} = \begin{bmatrix} n(4k+1) \\ n(4k+2) \\ n(4k+3) \\ n(4k+4) \end{bmatrix} \text{ and } \mathbf{1}_{4x1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},$$

and the matrix

$$H = \begin{bmatrix} m_1 & m_2 & m_3 & m_4 \\ m_2 & -m_1 & m_4 & -m_3 \\ m_3 & -m_4 & -m_1 & m_2 \\ m_4 & m_3 & -m_2 & -m_1 \end{bmatrix}$$

the received sequences can be expressed as

$$\mathbf{r} = H\mathbf{w}a + \boldsymbol{\varepsilon} + \mathbf{n} + (m_1+m_2+m_3+m_4)\mathbf{1}_{4x1}. \quad (9)$$

Note that the "channel matrix" H is orthogonal such that

$$H^T H = (m_1^2 + m_2^2 + m_3^2 + m_4^2) I. \quad (10)$$

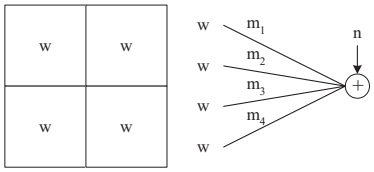


Figure 2: Repetitive watermark embedding.

3.2 Watermark Detection

Assuming that the image blocks' mean values remain approximately unchanged after watermark embedding, the "channel matrix" H can be estimated from the watermarked image, which means that the detection is blind. Multiplying $\mathbf{r}$ with H^T in the receiver, we obtain

$$\begin{aligned}\mathbf{z} = H^T\mathbf{r} &= H^T H\mathbf{w}a + H^T\boldsymbol{\varepsilon} + H^T\mathbf{n} \\ &+ (m_1+m_2+m_3+m_4)H^T\mathbf{1}_{4x1} \\ &= a(m_1^2+m_2^2+m_3^2+m_4^2)\mathbf{w} + \mathbf{v} + \mathbf{c} \end{aligned} \quad (11)$$

where $\mathbf{v} = H^T\boldsymbol{\varepsilon} + H^T\mathbf{n}$ and $\mathbf{c} = (m_1+m_2+m_3+m_4)H^T\mathbf{1}_{4x1}$. The vector $\mathbf{c}$ consists of constant elements. More specifically

$$\mathbf{c} = \begin{bmatrix} (m_1+m_2+m_3+m_4)^2 \\ (m_2+m_3)^2 - (m_1+m_4)^2 \\ (m_3+m_4)^2 - (m_1+m_2)^2 \\ (m_2+m_4)^2 - (m_1+m_3)^2 \end{bmatrix}$$

and can also be estimated from the watermarked image. Note that H^T is in fact the matched filter and $\mathbf{z}$ is the matched filter output.

Then we subtract $\mathbf{c}$ from $\mathbf{z}$ to obtain

$$\mathbf{z}' = a(m_1^2+m_2^2+m_3^2+m_4^2)\mathbf{w} + H^T\boldsymbol{\varepsilon} + H^T\mathbf{n}. \quad (12)$$

This sequence will be used for the watermark detection procedure. Note that the sequences $H^T\boldsymbol{\varepsilon} + H^T\mathbf{n}$ contain error variables and noise. If these terms were known, $\mathbf{w}$ could be estimated after proper scaling and slicing of $\mathbf{z}'$.

We follow here a commonly used strategy for watermark detection, i.e., a linear correlation based detector. We compute the linear correlation between $\mathbf{z}'$ and $\mathbf{w}$, and since the involved sequences (watermark, error and noise) are uncorrelated, we have that

$$E[\mathbf{z}'\mathbf{w}] = a(m_1^2+m_2^2+m_3^2+m_4^2)\sigma_w^2 \quad (13)$$

where E is the expectation operator.

For comparison reasons let us see the case of repetitive watermark embedding shown in Fig. 2. In this case, the linear correlation based detector's output is

$$E[y'w] = a(m_1+m_2+m_3+m_4)\sigma_w^2 \quad (14)$$

where $y' = y_1+y_2+y_3+y_4+n$.

The performance of the watermark detection can be measured in terms of the false alarm probability (P_F) and the probability of detection (P_D). P_F is the proportion of keys for which we decide that the desired watermark is present, while it is absent. P_D is the proportion of keys for which we correctly decide that the desired watermark is present.

(a) (b)

Figure 3: (a) Cover image, (b) Watermarked image

A way of depicting the performance of a detector is to plot P_D versus P_F. The false alarm probability is defined as

$$P_F = Q\left(\frac{\eta+\mu}{\sigma}\right) \quad (15)$$

and the probability of detection is defined in terms of P_F as:

$$P_D = Q\left[Q^{-1}(P_F) - \sqrt{SNR}\right] \quad (16)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{+\infty} e^{-\left(\frac{t^2}{2}\right)}dt$, η is a threshold and $SNR = \frac{m^2}{\sigma^2}$, with m and σ^2 being the mean and variance of the detector's output (under the hypothesis that the watermark is present). Each point of the plot corresponds to a value (P_F,P_D) for a given threshold η. As η increases, P_F decreases and so does P_D (and vice-versa). This type of performance depiction is called Receiver Operating Characteristic (ROC) [9]. Hence, as shown in (16), the ROC of the watermark detector depends exclusively on the value of *SNR*. The larger the value of *SNR*, the larger the P_D associated with a certain P_F and, as a consequence, the better the performance of the detector.

Here, it can be shown that $SNR_{E[\mathbf{z}'\mathbf{w}]} > SNR_{E[y'w]}$, where $SNR_{E[\mathbf{z}'\mathbf{w}]} = \frac{m^2_{E[\mathbf{z}'\mathbf{w}]}}{\sigma^2_{E[\mathbf{z}'\mathbf{w}]}}$ and $SNR_{E[y'w]} = \frac{m^2_{E[y'w]}}{\sigma^2_{E[y'w]}}$. This has also been verified by extensive experimental results. Thus, the proposed scheme performs considerably better than repetitive watermarking.

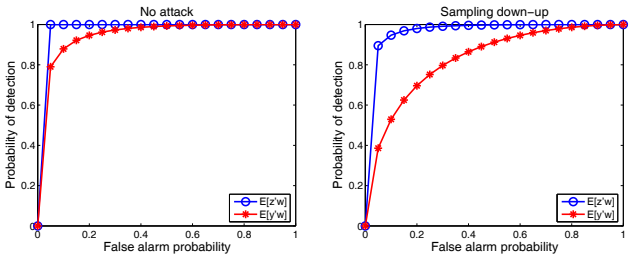
4. EXPERIMENTAL RESULTS

Extensive experiments have been conducted in order to test the detector's performance for the proposed scheme, as compared to that of the repetitive watermarking. Due to the limited space, we present here results only for the image of Lenna (256×256). In all cases, the watermark embedding and detection processes have been applied to 1000 different keys taken at random. The embedding was such that the PSNR (Peak Signal to Noise Ratio) between the cover and the watermarked image was $40dB$. The cover and the watermarked image are shown in Fig. 3. For each examined case the SNR was computed and the ROCs curves were plotted for standard values of the false alarm probability.

We examined attacks such as noise, JPEG compression, filtering (linear and non-linear) and sampling the image down (to the $3/4$ of its original dimensions) and up (back to its

Table 1: SNR(dB) computations for different attacks

Attacks	$SNR_{E[\mathbf{z'w}]}$	$SNR_{E[\mathbf{y'w}]}$	SNR_{gain}
No attack	15.5012	7.7878	7.7133
AWGN			
15dB	11.1565	7.7547	3.4017
10dB	10.2439	7.5988	2.6451
5dB	8.2833	7.3628	0.9205
JPEG Comp.			
QF:75	7.4756	0.5278	6.9478
QF:50	3.8729	-3.7032	7.5761
Wiener (3x3)	1.0322	-6.4998	7.5320
Median (3x3)	2.2647	-5.4501	7.7148
Sample down-up	9.2423	2.6375	6.6048



(a) (b)

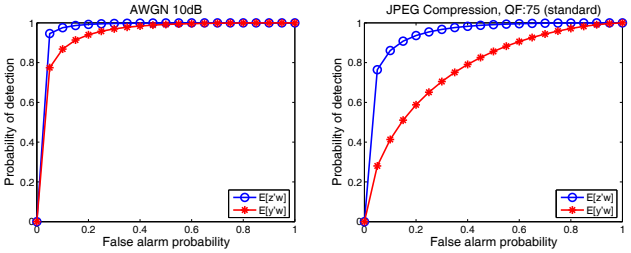
Figure 4: ROCs for (a) no attack and (b) sampling down-up

original dimensions). The $SNR_{E[\mathbf{z'w}]}$ and $SNR_{E[y'w]}$ were computed for the two compared techniques, as well as the $SNR_{gain} = SNR_{E[\mathbf{z'w}]} - SNR_{E[y'w]}$. The results are shown in Table 1. It is probable that the detection is notably improved for the proposed scheme, even in severe attacks' cases. Then, some characteristic ROCs curves are shown verifying the results of Table 1. In Fig. 4 we can see the curves for the cases of no attack and sampling down and up. The curves for the noise ($10dB$) and JPEG compression (QF:75) and the curves for the Wiener and median filtering are shown in Fig. 5 and in Fig. 6, respectively.

5. CONCLUSION AND FUTURE WORK

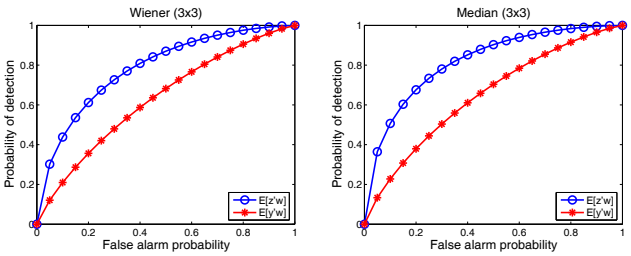
A new scheme for image watermarking has been proposed that is based on space-time block codes, and specifically a 4×4 real orthogonal design. Such a coding scheme provides full diversity, the maximum possible transmission rate, and low complexity. The performance of the proposed scheme (as far as detection is concerned) compares very favorably with the standard repetitive watermarking method.

Following similar derivation lines, the proposed scheme can be generalized for M transmitters. Moreover, it can be properly combined with standard repetitive watermarking in order to be robust to more attacks, as for example cropping. Finally, it can be used along with an appropriate perceptual mask that substantially increases the watermark strength [10] yielding a further performance improvement.



(a) (b)

Figure 5: ROCs for (a) AWGN and (b) JPEG compression



(a) (b)

Figure 6: ROC's for (a) Wiener and (b) median filtering

REFERENCES

[1] S. Voloshynovskiy, F. Deguillaume, S. Pereira and T. Pun, "Optimal Adaptive Diversity Watermarking with Channel State Estimation," in *Proc. SPIE: Security and Watermarking of Multimedia Content III*, vol. 4314, San Jose, CA, USA, Jan. 22–25, 2001.

[2] D. Kundur and D. Hatzinakos, "Diversity and Attach Characterization for Improved Robust Watermarking," *IEEE Transactions on Signal Processing*, vol. 49, no. 10, pp. 2383–2396, Oct. 2001.

[3] X. Li, Z. Du, W. Li, P. Lu and X. Xue, "Multi-channel Data Hiding Scheme for Color Images," in *Proc. IEEE ITCC 2003*, Las Vegas, NV, USA, Apr. 28-30, 2003, pp. 569–573.

[4] X. Li and X. Xue, "Improved Robust Watermarking in DCT Domain for Color Images," in *Proc. IEEE AINA 2004*, vol. 1, Fukuoka, Japan, Mar. 29-31, 2004, pp. 53–57.

[5] V. Tarokh, H. Jafarkhani and A. R. Calderbank, "Space-Time Block Codes from Orthogonal Designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[6] V. Tarokh, H. Jafarkhani and A. R. Calderbank, "The Application of Orthogonal Designs t Wireless Communication," in *Proc. IEEE Information Theory Workshop*, Killarney, Ireland, Jun. 1998, pp. 46–47.

[7] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," *IEEE Journal on Select Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[8] V. Tarokh, S. Seshadri and A. R. Calderbank,"Space-Time Codes for High Data Rate Wireless Communication: Performance analysis and Code construction," *IEEE Transactions on Information Theory*, vol. 44, pp. 744–765, Mar. 1998.

[9] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume II, Detection Theory*. Prentice-Hall, 1998.

[10] I. G. Karybali and K. Berberidis, "Efficient Spatial Image Watermarking via New Perceptual Masking and Blind Detection Schemes," submitted to *IEEE Transactions on Image Processing*.