OPTIMUM SIGNAL COMBINING EXTRACTION OF DWT-BASED WATERMARKS

J. Barata, D. Romão, M.P. Queluz and A. Rodrigues

Instituto Superior Técnico – Instituto de Telecomunicações
Av. Rovisco Pais, 1049-001 Lisboa, Portugal

ABSTRACT

The use of watermarking technology in the domain of digital rights management systems is essential due to the necessity of ensuring the protection of digital content. To improve the watermark extraction reliability, signal combining techniques are of great interest. In this paper, a wavelet-based watermarking technique is used as embedding method allowing a scalable multi-channel approach. A new method for estimating the subbands combining weights that does not require the embedding of extra information and can be used for blind and non-blind extraction is proposed. Its performance is assessed in the presence of JPEG/JPEG2000 compression. In the blind case, a new detector structure derived for additive-multiplicative watermarking embedding in Laplacian distributed host features (wavelet coefficients) is also presented and evaluated.

1. INTRODUCTION

In the past few years, several watermarking schemes operating in the wavelet transform domain have been proposed - a comparative study of some of these schemes is provided in [1]. An important advantage of the wavelet domain is to allow a scalable approach for watermarking, which may be particularly useful in the context of content adaptation schemes. In fact, if the same mark is embedded in each wavelet subband, watermark extraction (or detection) may be accomplished using different wavelet decomposition levels (i.e., different image resolutions). Another important aspect in using the wavelet domain is to naturally provide a multi-channel system for watermark transmission. In this sense, watermark extraction can be improved by applying signal combining techniques, common in radio communications with diversity. The benefits of these techniques are enhanced when the combination of the signals is done using weights based on a measure of the reliability for each channel. Some authors have already analyzed a few of these techniques in still image [2,3] and video [4] watermarking. In [2], the reliability factors (or combining weights) are based on the bit error probability of a reference (or pilot) watermark, known by the detector, which is transmitted in each channel. The trade off is that the available "bandwidth" for transmitting the useful watermark is reduced by the presence of the pilot watermark. Another possibility is to estimate the noise variance associated to each channel [3], through a procedure that requires the original host signal to be known. The approach of [4] considers each video frame as an independent channel, and the watermark is extracted considering a group of consecutive frames, simultaneously. The mark extracted for the first group (using a conventional procedure) is used in subsequent groups to compute the reliability factor for each frame.

This paper addresses signal-combining techniques for wavelet-based watermarking of still pictures. Extending the work presented in [4] in terms of signal combining structures, we propose a new method for estimating the subbands combining weights, that does not require embedding of extra information (i.e., a reference watermark), yet can be used in blind detection schemes. Its effectiveness - in terms of resulting bit error rate for the extracted mark - is assessed in the presence of compression (JPEG and JPEG2000), for both blind and non-blind extraction. For the blind case, a new detector structure required for an additive-multiplicative embedding rule in Laplacian distributed host features and originally derived in [8], is also presented and evaluated.

2. WATERMARK INSERTION

As it will be clear in section 3, the method proposed for signal combining can be applied to any (multi-channel) watermarking scheme that performs a soft detection of the embedded symbols. However, in order to quantify its performance, it will be used in a specific (and common) watermarking approach. In a first step, the luminance component of the image to be watermarked is decomposed in L resolution levels, through the digital wavelet transform (DWT). The mark, consisting in N_b binary and antipodal symbols, is embedded using a spread-spectrum (SS) approach. Each one of the N_b sequences that results from the SS modulation is mapped to a sub-set of coefficient positions, of every subband (excluding the lowest resolution one). The mappings are non-overlapping and pseudo-randomly generated, being secret key dependent. After the spatial assignment, and in order to adapt the embedding to the human visual system, the watermark is weighted according to the level and orientation of the subband and scaled by the energy of each wavelet coefficient.

For subband j (j=1 .. $3L$), the embedding process can then be described by

$$y(m,n)=x(m,n) + w(m,n) \quad \text{with}$$
$$w(m,n)=\alpha^j\, |x(m,n)|\, s(m,n)\, bit(m,n)\ , \qquad (1)$$

where $x(m,n)$ and $y(m,n)$ denote, respectively, the original and marked wavelet coefficient in position (m,n), $w(m,n)$ the coefficient modification introduced by the watermarking process, α^j (< 1) is the embedding strength (dependent on the subband level and orientation), $s(m,n)$ ($\in\{+1,-1\}$) denotes the element of the spreading sequence that was mapped to (m,n) and $bit(m,n)$ ($\in\{+1,-1\}$) the watermark bit inserted in position (m,n). DN_j/N_b, where D is the embedding density and N_j is the number of coefficients in the subband, gives the number of modified coefficients per watermarking bit in subband j.

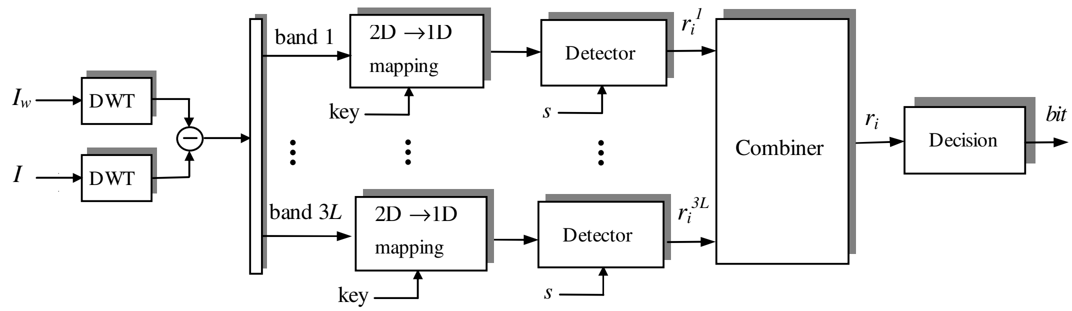


Figure 1 – Watermarking extraction scheme

3. WATERMARK EXTRACTION

3.1 Extraction scheme

Watermarking extraction can be achieved using the system depicted in figure 1, which performs the complementary operations of the embedding system. I_w and I denote, respectively, the watermarked and the original image, the later being available only for the non-blind extraction. The "2D →1D" block converts the watermarked coefficients from a two-dimensional to a one-dimensional space. Each detector, from a total of 3*L* detectors (one per subband) performs a "soft" detection of the watermark bits. Symbol $r_i^{\,j}$ designates the output of detector *j*, for the *i*-th watermark bit. At the "combiner", each detector output is weighted according to its estimated reliability, and summed-up. The last block ("decision") is just a comparator that, based on the combiner output, decides about the correspondent bit value.

3.2 Detectors structure

Optimum watermark extraction requires the knowledge of optimum decoding structures. For the non-blind case, and since the original unmarked coefficients are known (and can be cancelled in (1)) the optimum detector is a simple correlator. The output r_i^j at detector *j*, for the *i*-th watermark bit, can then be expressed as

$$r_i^j = \alpha^j \sum_{(m,n)\in S^j} |\, x(m,n)\,|\, bit_i \qquad (2)$$

where S^j is the set of positions of subband *j*, in which bit *i* was embedded.

For the blind case, the original image is not available - it must then be treated as noise. For an additive embedding rule, optimum detectors where derived in [5] for the case in which the host features follow a generalized Gaussian distribution. For an additive-multiplicative embedding rule, as expressed by eq. (1), the optimum maximum-likelihood detector structures have been investigated for host features that follow a Weibull distribution, as DFT coefficients [6] or a normal distribution [7]. In our case, and since wavelet coefficients are best modeled by a Laplace distribution [8], a Laplacian noise model should be used. For this model, the optimum output r_i^j at detector *j*, for the *i*-th watermark bit, is given by [9]

$$r_i^j = \sum_{(m,n)\in S^j} \left[\lambda^j \,|\, y(m,n)\,|\, \left(\frac{2\alpha^j f(m,n)}{1-\alpha^{j^2}}\right) + \ln\left(\frac{1-\alpha^j f(m,n)}{1+\alpha^j f(m,n)}\right) \right] \qquad (3)$$

where $f(m,n)$= sgn($y(m,n)$).$s(m,n)$, with sgn(.) being the "sign" function and λ^j characterizing the variance of the Laplace distribution of the (unmarked) subband *j*. This parameter can be estimated from the marked image by $\lambda^j = 1/E[y(m,n)]_{(m,n)\in S^j}$ [9], where E[.] denotes expected value.

From (3), the coefficients r_i^j result from the sum of a high number of random variables with the same distribution. Thus, they may be modeled as the output of a additive white gaussian noise (AWGN) channel. Since we are using binary signalling, with antipodal pulses, the bit error probability (P_b^j) associated with the watermark extracted from subband *j*, is

$$P_b^j = Q(\sqrt{\mathrm{SNR}}) \quad \text{with} \quad \mathrm{SNR} = \frac{\mu_j^2}{\sigma_j^2} \qquad (4)$$

where μ_j and σ_j^2 are, respectively, the expected value and the variance of the *j*-th detector output, which can theoretically estimated for each host image [9]. The watermark may be extracted from each subband applying a hard decision to the corresponding detector output by

$$bit_i = \mathrm{sgn}(r_i^j), \quad \forall i \in \{1,\ldots,N_b\} \qquad (5)$$

3.3 Signal combining

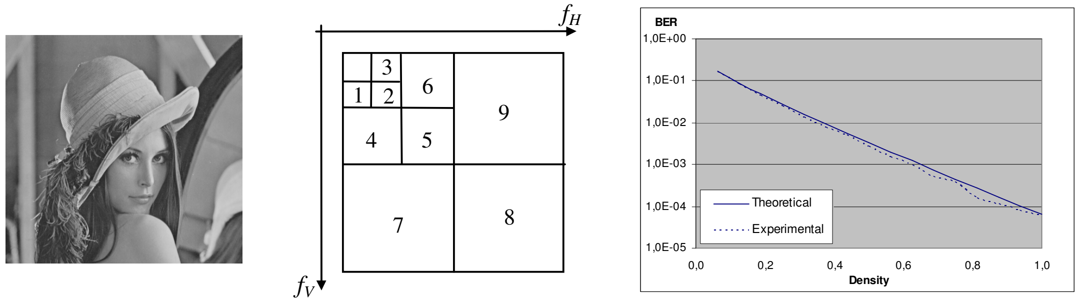
For each embedded bit, the detector's outputs are weighted and summed-up, and the resulting combining value is used for the transmitted symbol decision. The combiner output for the *i*-th bit can be expressed as

$$r_i = \sum_{j=1}^{3L} c_j r_i^j \qquad (6)$$

where c_j is the weight (i.e., reliability measure) associated to channel *j*. For independent and AWGN channels, the optimal weights - in the sense that the SNR at the combiner output ($\mathrm{SNR_{oc}}$) is maximized - are given by [4]

$$c_j = \mu_j \Big/ \sigma_j^2 \qquad (7)$$

resulting in



a) b) c)

Figure 2 – a) “Lena”, 512×512 pixels; b) DWT decomposition, with the subband numbering used in experimental results; c) Theoretical and experimental BER values, as a function of D, for signal combination of bands 1 to 6 (blind extraction).

$$P_b^{oc} = Q(\sqrt{\mathrm{SNR}_{oc}}) \quad \text{with} \quad \mathrm{SNR}_{oc} = \sum_{j=1}^{3L} \left(\frac{\mu_j}{\sigma_j}\right)^2 \tag{8}$$

When *a priori* statistics relative to μ_j and σ_j^2 are not known, they must be estimated from the received data. If in each of the $3L$ channels we perform N_b measurements, considered as independent, the joint probability density function is given by:

$$p(\mathbf{r} \mid \boldsymbol{\mu}, \mathbf{b}, \boldsymbol{\sigma}) = \prod_{j=1}^{3L} \prod_{i=1}^{N_b} (2\pi\sigma_j^2)^{-1/2} e^{-|r_i^j - \mu_j bit_i| / 2\sigma_j^2} \tag{9}$$

where $\mathbf{r}$ is the measurement matrix, of dimension $3L \times N_b$, μ and σ are vectors of dimension $3L$ and $\mathbf{b}$ is a vector of dimension N_b corresponding to the watermark bitstream. Applying the natural logarithm to (9), and maximizing it relatively to μ_i^j and σ_i^j leads to

$$\hat{\mu}_j = \frac{1}{N_b} \sum_{i=1}^{N_b} (r_i^j bit_i); \quad \hat{\sigma}_j = \frac{1}{N_b} \sum_{i=1}^{N_b} (r_i^j bit_i - \hat{\mu}_j)^2; \tag{10}$$

The values $bit_i \in \{-1,1\}$, i=1,…,N_b, required to compute (10) are not known *a priori*. One solution is to use a number of pilot bits (the summations in (10) will be done over these bits), known by the combiner which, as stated in the introduction, decreases the available bandwidth for the useful watermark bits (**method 1**). A non-bandwidth consuming solution is to use the bit values that result from an hard decision applied to each subband detector output, with eq. (5) (**method 2**). In this case, and since all embedded bits are used in (10), if some of them are erroneously extracted (by the subband hard detectors), the estimates of μ_j and σ_j^2 may still be close to the right values. A non-optimal solution is to consider equal and unitary weighting coefficients in (6), i.e. c_j=1, $\forall_j$.

After combining the detector´s outputs, the watermark is extracted applying a hard decision to the combiner output

$$bit_i = \mathrm{sgn}(r_i), \quad \forall i \in \{1, \ldots, N_b\} \tag{11}$$

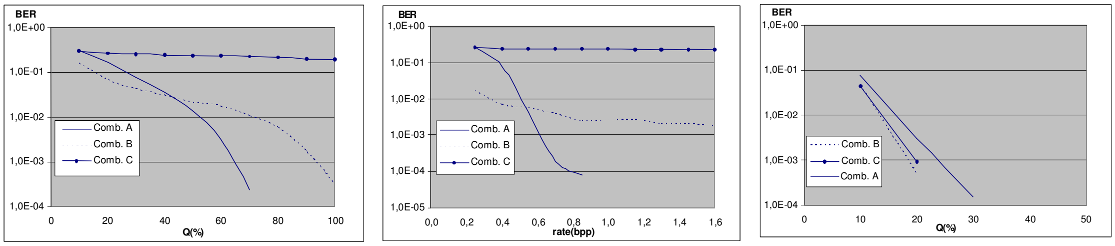
Figure 2-c) presents theoretical and experimental BER (bit error rate) curves obtained for the test image “Lena” (figure 2-a), as a function of the embedding density, D, with N_b=64. The extraction was performed using bands 1 to 6, of a 3 levels (L=3) subband decomposition (figure 2-b), and optimum weighting coefficients (method 2). Bands 7 to 9 were not used in this case, since the resulting BER is too low to be obtained experimentally (theoretically, $P_b \approx 10^{-14}$ for D=0.0625). Similar BER values were obtained for optimal combining with method 1 and for non-optimal combining with equal weights.

Since no other source of distortion has been considered, besides the noise contribution from the host image itself, these curves constitute an upper bound of the performance (lower bound of achievable P_b) that can be attained, when bands 1 to 6 are used. Also, the good match with the experimental BER curves allows the validation of the detector structures and of the overall simulation model.

4. RESULTS

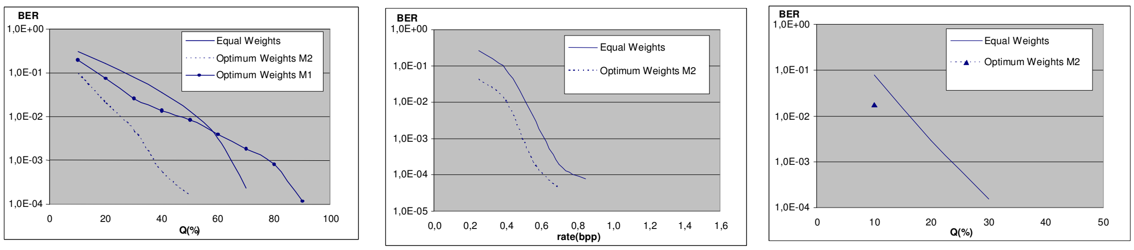
Figures 3 and 4 present experimental BER curves for “Lena”, under JPEG or JPEG2000 compression, and blind or non-blind extraction. For JPEG compression, the curves are plotted as a function of the quality factor (Q); for JPEG2000, the variable is the resulting number of bits per pixel (rate). A 3 levels DWT decomposition, as represented in figure 2-b), was used. The subbands embedding strengths were set to values that guarantee the invisibility of the mark. All tests have been performed using an insertion density of 1 and randomly generated marks with a length of 64 bits.

In figure 3, three different combinations (with equal weights) of subbands were considered: bands 1 to 9 (comb A); bands 1 to 6 (comb B) and bands 1 to 3 (comb C). These curves show that the “best” group of subbands to be considered at the extraction depends on the distortion incurred by the watermarked image, and on the availability (or not) of the original host image. For instance, under non-blind extraction (fig. 3-c) and since the noise contribution from the original image is cancelled, the distortion is mainly due to compression, which penalizes essentially the higher frequencies. Accordingly, watermark extraction should rely mainly on the middle subbands. For blind extraction, the best bands result from two distortion effects - compression (which is dominant for high compression factors) and the noise contribution from the original host image (which is dominant for low to mild compression factors).



a) b) c)

Figure 3 –Experimental BER curves for three (comb A), two (comb B) or one (comb C) DWT levels, and: a) Blind extraction under JPEG compression; b) Blind extraction under JPEG2000 compression; c) Non-blind extraction under JPEG compression.



a) b) c)

Figure 4 –Experimental BER curves for equal weights and optimal weights combining, and: a) Blind extraction under JPEG compression; b) Blind extraction under JPEG2000 compression; c) Non-blind extraction under JPEG compression.

In figure 4, extraction was performed considering subbands 1 to 9, using the different combining strategies analysed in section 3.2: optimal weigths, method 1 (M1); optimal weigths, method 2 (M2); equal weigths. Results concerning method 1 were obtained using 4 pilot bits (64 + 4 embedded bits). Depending on the test image, a higher number of pilot bits may result in lower BER values. However, for all tested cases, method 2 always outperformed method 1. In figure 4-c), no experimental errors were obtained for JPEG quality factors above 10.

Comparing figures 3 and 4, we may conclude that combining the detector´s outputs through the new method proposed in this paper, which does not require extra pilot bits, allows to automatically “tune” the more reliable group of subbands. This conclusion holds for both blind and non-blind extraction.

5. CONCLUSIONS

In this paper we have proposed a new method for estimating the channel reliability weights, which can be used in multi-channel approaches for watermarking that perform a soft-decision of the embedded bits. Its effectiveness was demonstrated in a concrete wavelet-based watermarking system, for distortion scenarios that include JPEG/JPEG2000 compression, and blind or non-blind watermark extraction. A new detector structure required for this particular case was also presented and evaluated.

6. REFERENCES

[1] P. Meerwald and A. Uhl, “A survey of wavelet-domain watermarking algorithms”, in IS&T/SPIE - Proc. of Security and Watermarking of Multimedia Contents – III, Jan. 2001.

[2] D. Kundur and D. Hatzinakos, “Improved robust watermarking through attack characterization”, *Optics Express*, Vol.3, No. 12, Dec. 1998.

[3] D. Kundur and D. Hatzinakos, “Towards Robust Logo Watermarking using Multiresolution Image Fusion Principles”, *IEEE Trans. on Multimedia*, April 2003.

[4] T. Brandão, M.P. Queluz and A. Rodrigues, “Diversity enhancement of coded spread spectrum video watermarking”, *Wireless Personal Communications Journal*, Vol. 23, No.1, pp. 93-104, Oct. 2002.

[5] J. Hernández, M. Amado and F. Pérez-González, “DCT-Domain watermarking techniques for still images: detector performance analysis and a new structure”, *IEEE Trans. Image Processing*, vol. 9, pp. 55-68, Jan. 2000

[6] M. Barni, F. Bartolini, A. Rosa and A. Piva, “Optimum decoding and detection of a multiplicative amplitude encoded watermark”, in IS&T/SPIE Proc. of Security and Watermarking of Multimedia Contents – IV, pp. 409-420, Jan. 2002.

[7] M. Barni, F. Bartolini and A. Rosa, “On the performance of Multiplicative Spread Spectrum Watermarking”, in Proc. of IEEE - MMSP’02, Dec. 2002.

[8] K. Birney and T. Fisher, “On the modeling of DCT and subband image data for compression”, *IEEE Trans. on Image Processing*, Vol.4, no.2, pp. 186-193, Feb. 1992.

[9] J. Barata and D. Romão, “A system for digital watermarking of still images”, Final Year Thesis, Instituto Superior Técnico, Sep. 2003.