Improved Wavelet-Based Watermarking Through Pixel-Wise Masking

Mauro Barni, *Member, IEEE*, Franco Bartolini, *Member, IEEE*, and Alessandro Piva

***Abstract*—A watermarking algorithm operating in the wavelet domain is presented. Performance improvement with respect to existing algorithms is obtained by means of a new approach to mask the watermark according to the characteristics of the human visual system (HVS). In contrast to conventional methods operating in the wavelet domain, masking is accomplished pixel by pixel by taking into account the texture and the luminance content of all the image subbands. The watermark consists of a pseudorandom sequence which is adaptively added to the largest detail bands. As usual, the watermark is detected by computing the correlation between the watermarked coefficients and the watermarking code, anyway the detection threshold is chosen in such a way that the knowledge of the watermark energy used in the embedding phase is not needed, thus permitting to adapt it to the image at hand. Experimental results and comparisons with other techniques operating in the wavelet domain prove the effectiveness of the new algorithm.**

***Index Terms*—Image watermarking, perceptual noise masking, wavelets.**

I. Introduction

IMAGE watermarking [1], [2] is finding more and more support as a possible solution for the protection of intellectual property rights. To this aim, many techniques have been proposed in the literature over the last few years, and many commercial products are already available. It is possible to state that the most important features a watermarking technique to be used for IPR protection should exhibit are unobtrusiveness and robustness: in practice, it is required that a signal is accurately hidden into image data in such a way to result very difficult to be perceived but also very difficult to be removed. Another important characteristic is blindness [3], i.e., the watermark decoder must not require the original nonwatermarked image for extracting the embedded code.

A research field whose results could undoubtedly help in the design of a watermarking algorithm is image compression. The goal of image compression can be described as one of removing from images all the data which are perceptually unrelevant; similarly the problem of image watermarking is one of adding, to images, data which are perceptually unimportant. To the aim of effectively selecting the image information that can be removed without degrading subjective image quality, theoretical models of the human visual system (HVS) have been deeply studied [4]. Similarly, it is today widely accepted that robust image watermarking techniques should largely exploit the characteristics of the HVS, for more effectively hiding a robust watermark [1], [5], [6].

A widely used technique exhibiting a strong similarity to the way the HVS processes images is the Discrete Wavelet Transform (DWT). As a matter of fact, the next-generation image coding standard JPEG2000 [7], [8] will strongly rely on DWT for obtaining good quality images at low coding rates.

A. Watermarking in the Wavelet Domain

Given its suitability to model the HVS behavior, the DWT has gained interest among watermarking researchers, as it is witnessed by the number of algorithms following this approach that have been proposed over the last few years. In this section some of these algorithms are briefly reviewed, by particularly highlighting the approaches they embraced for considering HVS factors.

Some methods directly take inspiration from the most popular wavelet based compression algorithms. In [9] and [10], for example, the most significant DWT coefficients are selected and modified to carry the watermark. In the first case, some side information (i.e., the location of the modified coefficients) is required to recover the watermark. In the second, an algorithm is proposed for identifying *a posteriori* the modified coefficients. To take into account visual effects, in [9] very large coefficients are left unchanged, while in [10] the watermarking signal is weighted according to a band-dependent value. Xia *et al.* [11] add a pseudorandom sequence to the largest coefficients of the detail bands: perceptual considerations are taken into account by setting the amount of modification proportional to the strength of the coefficient itself. Watermark detection is achieved through comparison with the original un-watermarked image.

Other algorithms hide into images binary logos which are also hierarchically decomposed. Kundur *et al.* [12], for example, first decompose a binary logo through DWT, then repeatedly add it to the subbands of the DWT decomposition of the host image; before being added, the watermark is scaled by a *salience* factor, computed on a block by block basis, related to the local image noise sensibility: visual masking is thus exploited up to a block resolution.

Manuscript received October 26, 1999; revised December 15, 2000. This work was supported in part by the Italian National Research Council (CNR) in the framework of the Project on Cultural Heritage (Progetto Finalizzato Beni Culturali) under Grant 97.00603.PF36, and by the Italian Ministry for University and Scientific and Technological Research (MURST) in the framework of the Project on Monitoring of Quality and Watermarking of Natural and Synthetic Images. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Naohisa Ohta.

M. Barni is with the Department of Information Engineering, University of Siena, 53100 Siena, Italy (e-mail: barni@dii.unisi.it).

F. Bartolini and A. Piva are with the Department of Electronics and Telecommunications, University of Florence, 50139 Firenze, Italy (e-mail: barto@lci.det.unifi.it; piva@lci.det.unifi.it).

Publisher Item Identifier S 1057-7149(01)02477-0.

In [13], the binary logo and the image are hierarchically decomposed (the image through DWT), each detail subband of the logo is then embedded into the corresponding detail subband of the image (more bits of the binary logo are embedded into more active image locations), the original image is required for watermark extraction. Image activity is estimated blockwise through variance computation.

In [14], a binary code is embedded by suitably quantizing some of the coefficients of the detail bands: for watermark recovery the embedded binary code is estimated by analyzing coefficients quantization, once the code has been estimated it is correlated with the watermark and the result compared to a threshold chosen on the basis of a given false positive probability. No particular attention is given to visual masking.

The classical technique proposed in [1] is extended in a straightforward way to the wavelet domain in [15]: each DWT coefficient of highpass bands is modified proportionally to its magnitude. Nicchiotti *et al.* [16] choose to embed the watermark into the low pass band, by imposing a given difference among the mean values of two equally sized, randomly selected, subsets of the low pass image; the original image is not required for watermark detection. No particular care is taken with reference to perceptual masking.

In [17], code bits are hidden within the less significant bits of DWT subband coefficients, the bits to be modified are chosen according to the results of a perceptual analysis step performed on a block by block basis by estimating the local variance; watermark extraction is blind.

An excellent tutorial on the possible approaches to exploit HVS characteristics for watermarking images and videos is presented in [6], where it is proposed to uniformly weigh the strength of the watermark on each DWT band according to a band-dependent factor.

It is also worth mentioning the work of Swanson *et al.* [18], where the DWT transform is applied to a video sequence by decomposing it, along the temporal axis, into stationary and moving components.

B. Novel Algorithm

In this paper, a novel blind watermarking algorithm, which embeds the watermark in the DWT domain by exploiting the characteristics of the HVS, is presented. The main novelty of the algorithm resides in the way the HVS properties are exploited to improve watermark robustness and invisibility. The watermark is an additive one, however watermark strength is modulated according to the local image characteristics. Watermark strength modulation is accomplished through a mask giving a pixel by pixel measure of the sensibility of the human eye to local image perturbations. In contrast to DWT watermarking algorithms proposed so far, which either operate on a subband or a block base, the watermark strength is adjusted pixelwise. In addition, a sophisticated mechanism derived from DWT coding studies is introduced to take into account several HVS phenomena, such as grey-level sensibility, isofrequency and non-isofrequency masking, noise sensibility around edges.

The watermark consists of a binary (± 1) pseudorandom sequence that is added to the DWT coefficients of the three largest detail subbands of the image. As said above, to effectively hide the watermark, each binary value is multiplied, before being added, by a weighing parameter. Stated in another way, a mask is built which, pixel by pixel, gives the maximum amount of modifications that can be applied to the corresponding DWT coefficient in the detail band without compromising watermark invisibility. Mask construction relies on a work by Lewis and Knowles [19], in which the authors propose a method to evaluate the optimum quantization step for each DWT coefficient according to psychovisual considerations. Some modifications to the method by Lewis and Knowles are proposed here, to make it suitable to the computation of the maximum visibly tolerable watermark energy that can be used for each DWT coefficient.

For watermark detection, the correlation between the watermark to be tested for presence, and the marked coefficients is computed. The value of the correlation is compared to a threshold to decide whether the watermark is present or not. An optimum threshold is theoretically set to minimize the probability of missing the watermark subject to a given false detection rate (Neyman–Pearson criterion). An important consequence of the approach we used to set the threshold value is that such a value depends only on the statistics of the watermarked image, and can thus be computed *a posteriori*, without any knowledge about the watermark energy. This permits to adapt the watermark strength to the image at hand, thus making it easier to respect the invisibility constraint without loosing robustness.

Extensive experiments were conducted to prove the validity of the new masking approach and to compare it with conventional wavelet-based embedding strategies. Experiments aimed at assessing the performance of the new system both from the point of view of watermark invisibility and from the point of view of robustness; in particular the system has demonstrated to be resistant to JPEG and wavelet-based compression, median filtering, Gaussian noise addition, multiple marking, cropping plus zero padding, and morphing. The comparison of the new algorithm with other state-of-the-art watermarking systems operating in the wavelet domain confirms the validity of the approach proposed in this work.

II. Perceptual Watermark Embedding

Thanks to its excellent spatio-frequency localization properties, the DWT is very suitable to identify the image areas where a disturb can be more easily hidden. In particular, this property effectively allows to exploit the HVS iso- and near-frequency masking effect: if a DWT coefficient is modified, only the region of the image where the particular frequency corresponding to that coefficient is present will be modified, in contrast to what happens, for example, by using full frame DFT/DCT watermarking [20].

A. Watermark Embedding

The image to be watermarked is first decomposed through DWT in four levels: let us call I_l^θ the subband at resolution level $l = 0, 1, 2, 3$ and with orientation $\theta \in \{0, 1, 2, 3\}$ (see Fig. 1). The watermark, consisting of a pseudorandom binary (± 1) sequence, is inserted by modifying the wavelet coefficients belonging to the three detail bands at level 0, i.e., I_0^0, I_0^1, and I_0^2.

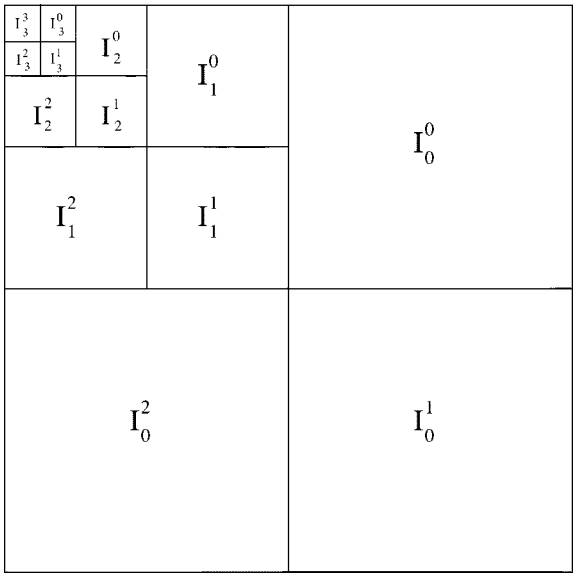


Fig. 1. Sketch of the decomposition of an image in four resolution levels through the DWT.

The choice of embedding the watermark only into the three largest detail subbands was motivated by experimental tests, as the one offering the best compromise between robustness and invisibility. Indeed, inserting the watermark into these subbands could give a lower robustness (e.g., compression or low-pass filtering), but, given the low visibility of disturbs added to these frequencies, a higher level of watermark strength is allowed, thus compensating for such a bigger fragility.

In more detail, a pseudorandom binary sequence $m_h \in \{+1, -1\}$ is arranged in 2-D, in such a way to scan the aforementioned subbands

$$x^{\theta}(i,j) = m_{(\theta MN+iN+j)} \tag{1}$$

where $2M \times 2N$ are the dimensions of the host image and $\theta \in \{0,1,2\}$; subbands coefficients are then modified according to the rule

$$\tilde{I}_0^{\theta}(i,j) = I_0^{\theta}(i,j) + \alpha w^{\theta}(i,j) x^{\theta}(i,j) \tag{2}$$

where α is a global parameter accounting for watermark strength, and $w^{\theta}(i,j)$ is a weighing function considering the local sensitivity of the image to noise. It is this weighing function that allows to exploit the masking characteristics of the HVS.

B. Perceptual Weighing

In order to embed into the images the maximum, but still unperceptible, level of watermark, the weighing function $w^{\theta}(i,j)$ has to consider how the eye perceives disturbs. In their paper [19], Lewis and Knowles tackle the problem of DWT coefficients quantization for compression purposes: they propose to adapt the quantization step of each coefficient according to the local noise sensitivity of the eye. Some modifications of the model proposed in [19] are introduced here, in order to better fit the behavior of the HVS to the watermarking problem. In particular, the following considerations have been taken into account.

- The eye is less sensitive to noise in high resolution bands, and in those bands having orientation of 45° (i.e., $\theta = 1$ bands in our case; see Fig. 1).
- The eye is less sensitive to noise in those areas of the image where brightness is high or low.
- The eye is less sensitive to noise in highly textured areas but, among these, more sensitive near the edges.

Based on these considerations, we computed the quantization step of each coefficient as the weighted product of three terms

$$q_l^{\theta}(i,j) = \Theta(l,\theta)\Lambda(l,i,j)\Xi(l,i,j)^{0.2} \tag{3}$$

where the meaning of each term in this equation is explained below. Note that as described previously, we decided to embed the watermark only at the first decomposition level. Despite this decision, we will analyze the behavior of the weighing function $w_l^{\theta}(i,j)$, even for $l \neq 0$. In a future research, in fact, it could be interesting to further investigate how watermark characteristics change when all the wavelet subbands are marked. Let us start the analysis of $q_l^{\theta}(i,j)$ by the first of the expression in (3). To take into account how sensitivity to noise changes depending on the band (in particular depending on the orientation and on the level of detail), we let

$$\Theta(l,\theta) = \left\{ \begin{array}{ll} \sqrt{2}, & \text{if } \theta = 1 \\ 1, & \text{otherwise} \end{array} \right\} \cdot \left\{ \begin{array}{ll} 1.00, & \text{if } l = 0 \\ 0.32, & \text{if } l = 1 \\ 0.16, & \text{if } l = 2 \\ 0.10, & \text{if } l = 3 \end{array} \right\}. \tag{4}$$

The second term takes into account the local brightness based on the graylevel values of the low pass version of the image. Since Lewis and Knowles assumed that the eye is less sensitive in the regions with high brightness, they proposed to compute this factor in the following way:

$$\Lambda(l,i,j) = 1 + L(l,i,j) \tag{5}$$

where

$$L(l,i,j) = \frac{1}{256} I_3^3 \left(1 + \left\lfloor \frac{i}{2^{3-l}} \right\rfloor, 1 + \left\lfloor \frac{j}{2^{3-l}} \right\rfloor \right). \tag{6}$$

Based on the consideration that the human eye is less sensitive to changes in very dark regions as well, we have modified this factor as follows:

$$L'(l,i,j) = \left\{ \begin{array}{ll} 1 - L(l,i,j), & \text{if } L(l,i,j) < 0.5 \\ L(l,i,j), & \text{otherwise} \end{array} \right\}. \tag{7}$$

Finally, the third term

$$\begin{aligned} \Xi(l,i,j) &= \sum_{k=0}^{3-l} \frac{1}{16^k} \sum_{\theta=0}^{2} \sum_{x=0}^{1} \sum_{y=0}^{1} \left[I_{k+l}^{\theta} \left(y + \frac{i}{2^k}, x + \frac{j}{2^k} \right) \right]^2 \\ &\quad \cdot \operatorname{Var} \left\{ I_3^3 \left(1 + y + \frac{i}{2^{3-l}}, 1 + x + \frac{j}{2^{3-l}} \right) \right\}_{\substack{x=0,1 \\ y=0,1}} \end{aligned} \tag{8}$$

gives a measure of texture activity in the neighborhood of the pixel. In particular, this term is composed by the product of two contributions: the first is the local mean square value of the

DWT coefficients in all detail subbands, while the second is the local variance of the low-pass subband. Both these contributions are computed in a small 2×2 neighborhood corresponding to the location (i, j) of the pixel. Since the first contribution can represent the distance from the edges, whereas the second one the texture, we decided to multiply the two terms, according to our consideration that the eye is less sensitive in textured areas, but more sensitive near edges (in [19], Lewis and Knowles propose to simply add the two contributions).

The fact that $q_l^\theta(i,j)$ is chosen as the quantization step for a DWT coefficient at location (i, j), implies that disturbs having value lower than $q_l^\theta(i,j)/2$ are assumed not to be perceivable. Thus we choose to set the weighing function as

$$w^\theta(i,j) = q_0^\theta(i,j)/2 \tag{9}$$

that is equal to the quantization step of the coefficient to which the watermarking code has to be added. Given that the watermark code is binary, this approach allows to add to each DWT coefficient the maximum unperceivable watermark level.

III. Watermark Detection

Watermark detection is accomplished without referring to the original image. The correlation between the marked DWT coefficients and the watermarking sequence to be tested for presence is computed

$$\rho = \frac{1}{3MN} \sum_{\theta=0}^{2} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \tilde{I}_0^\theta(i,j) x^\theta(i,j) \tag{10}$$

and compared to a threshold T_ρ chosen in such a way to grant a given probability of false positive detection.

To determine T_ρ we adopt the Neyman-Pearson criterion: instead of minimizing the overall error probability, we minimize the probability of missing the watermark subject to a given false detection rate. In this way a twofold aim is pursued: robustness against attacks is augmented and detection is accomplished without any knowledge about the watermark strength α. Given an image $\tilde{I}$ and a watermark X, only three cases are possible.

Case A: image is not watermarked.

Case B: image is watermarked with a code Y other than X.

Case C: image is watermarked with the code X.

To apply statistical decision theory, some assumptions must be made on the random variables forming the observation variable ρ. We know that $x^\theta(i,j)$'s are binary valued, zero mean, independent random variables. Besides, they do not depend on DWT coefficients. As to $\tilde{I}_0^\theta(i,j)$'s, we make the realistic assumption that they are zero mean, independent variables. By exploiting the central limit theorem, we can also assume that ρ is normally distributed. Under these hypotheses, it can be easily demonstrated that the mean values of ρ in cases $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are

$$\begin{aligned}
&\textbf{Case A } \mu_{\rho_A} = 0 \\
&\textbf{Case B } \mu_{\rho_B} = 0 \\
&\textbf{Case C } \mu_{\rho_C} = \frac{\alpha}{3MN} \sum_{\theta=0}^{2} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} E[w^\theta(i,j)].
\end{aligned} \tag{11}$$

We say that the watermark X is present if $\rho > T_\rho$, whereas if $\rho < T_\rho$ the decoder decides for the watermark absence. To estimate the probability of false detection, $P_f = \text{Prob}(\rho > T_\rho \,|\, \textbf{Case A}.\text{OR}.\textbf{Case B})$, we need to compute the variance of the random variable ρ in Case A and Case B. Once again, it can be easily obtained that for Case A

$$\sigma_{\rho_A}^2 = \frac{\sigma_x^2}{(3MN)^2} \sum_{\theta=0}^{2} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} E\left[\left(I_0^\theta(i,j)\right)^2\right] \tag{12}$$

whereas for Case B

$$\begin{aligned}
\sigma_{\rho_B}^2 = \frac{\sigma_x^2}{(3MN)^2} &\sum_{\theta=0}^{2} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} E\left[\left(I_0^\theta(i,j)\right)^2\right] \\
&+ \alpha^2 \sigma_x^2 E\left[\left(w^\theta(i,j)\right)^2\right]
\end{aligned} \tag{13}$$

from which it is readily seen that the latter is the worse case, since the higher the variance the higher the error probability. We can, then, say that

$$P_f \le \frac{1}{2} \text{erfc}\left(\frac{T_\rho}{\sqrt{2\sigma_{\rho_B}^2}}\right). \tag{14}$$

By imposing, for example, that $P_f \le 10^{-8}$, we find

$$T_\rho = 3.97\sqrt{2\sigma_{\rho_B}^2}. \tag{15}$$

To get more insight into the expression of T_ρ, let us consider the mean square value of watermarked coefficients, $E[\tilde{I}_0^\theta(i,j)^2]$. By recalling the embedding rule stated in (2), we can write

$$\begin{aligned}
E\left[\tilde{I}_0^\theta(i,j)^2\right] = E\left[I_0^\theta(i,j)^2\right] &+ \alpha^2 E\left[w_0^\theta(i,j)^2 x_0^\theta(i,j)^2\right] \\
&+ 2\alpha E\left[I_0^\theta(i,j) w_0^\theta(i,j) x_0^\theta(i,j)\right]
\end{aligned} \tag{16}$$

which, by noting that $\sigma_x^2 = \sigma_x^4 = 1, E[I_0^\theta(i,j)] = E[x_0^\theta(i,j)] = 0$, and that $x_0^\theta(i,j)$ does not depend on $I_0^\theta(i,j)$ (and hence on $w_0^\theta(i,j)$), yields

$$\sigma_{\rho_B}^2 = \frac{1}{(3MN)^2} \sum_{\theta=0}^{2} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} E\left[\left(\tilde{I}_0^\theta(i,j)\right)^2\right]. \tag{17}$$

This is a very important result, because it implies that the threshold can be computed *a posteriori* on the watermarked image. Hence, watermark detection can be performed without knowing the value of the watermark strength (α), which can thus be accurately adapted to each image.

In practice the following unbiased estimate of σ_{ρ_B} is used

$$\sigma_{\rho_B}^2 \approx \frac{1}{(3MN)^2} \sum_{\theta=0}^{2} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(\tilde{I}_0^\theta(i,j)\right)^2. \tag{18}$$

IV. Experimental Results

The algorithm has been extensively tested on various standard images and attempting different kinds of attacks: in this section

(a)

(b)

Fig. 2. (a) Original image "Lena" and (b) its watermarked copy.

some of the most significant results will be shown. For the experiments presented in the following, the Daubechies-6 filtering kernel has been used for computing the DWT.

First, watermark invisibility is evaluated: in Fig. 2(a), the original "Lena" image is presented, while in Fig. 2(b), the watermarked copy is shown: the images are evidently undistinguishable, thus proving the effectiveness of DWT watermarking and the masking procedure. In particular, the effectiveness of the weighing function can be appreciated from Fig. 3, where the absolute difference between the original image and the watermarked one, magnified by a factor 8, is shown: it is evident that the watermark is mainly hidden into high activity regions and around edges (see, for example, the high level of watermark at the borders of the hat and the shoulders of the girl, and over the feathers).

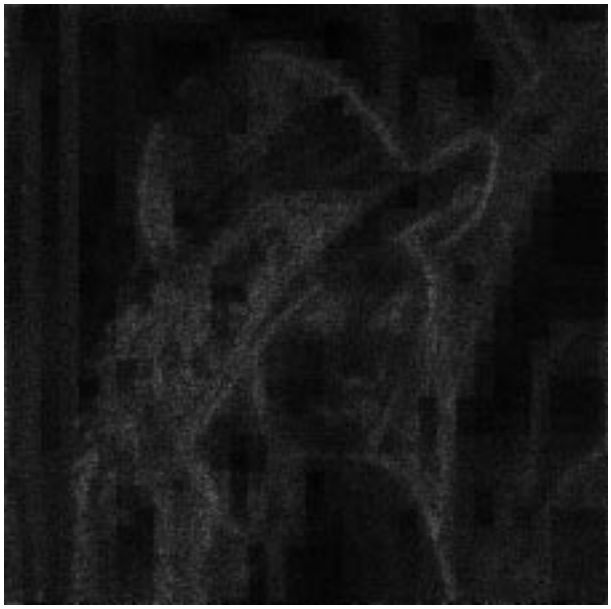


Fig. 3. Absolute difference between the original image and the watermarked one, magnified by a factor 8.

TABLE I
MAXIMUM WATERMARK STRENGTH ACHIEVABLE BY THE THREE VISUAL MASKING METHODS SUBJECT TO THE INVISIBILITY CONSTRAINT. THE VALUES REPORTED IN THE TABLE REFER TO THE AVERAGE α OVER THE WHOLE IMAGE

Image	Our method	Variance based Mask	Constant Mask
Pepper	1.68	1.38	1.64
Boat	2.10	1.15	1.81
Airplane	1.98	0.78	1.74
Lena	2.14	0.96	1.78

We compared the results obtained by our algorithm with those achieved through the use of two standard masking techniques. According to the first one [6], the watermark strength α is constant over each subband. More specifically, as suggested in [6] we chose α according to the psychovisual analysis described in [21]. The second technique is more sophisticated, in that it adapts the watermark strength locally according to the image characteristics. An approach similar to that proposed in [12] is followed. Subbands are divided into nonoverlapping 8×8 blocks, then for each block the local subband variance is calculated and used to insert a higher watermark energy in blocks characterized by a higher activity (i.e., a higher variance), that is we let

$$w_0^{\theta}(i,j) = \frac{1}{K}\mathrm{Var}\left\{I_0^{\theta}(y,x)\right\}_{(y,x)\in B(i,j)} \tag{19}$$

where $B(i,j)$ is the 8×8 block pixel (i,j) belongs to, and K is the maximum value of block variance. Then we inserted a watermark with the maximum admissible strength subject to invisibility. Invisibility was judged by a human observer at a viewing distance equal to six times the image size. The results we obtained on various standard images are reported in Tables I and

TABLE II
MINIMUM PSNR GRANTING WATERMARK INVISIBILITY UNDER DIFFERENT MASKING STRATEGIES

Image	Our method	Variance based Mask	Constant Mask
Pepper	37.98	37.59	38.09
Boat	35.44	38.79	37.22
Airplane	35.87	41.13	37.56
Lena	35.76	41.31	37.39

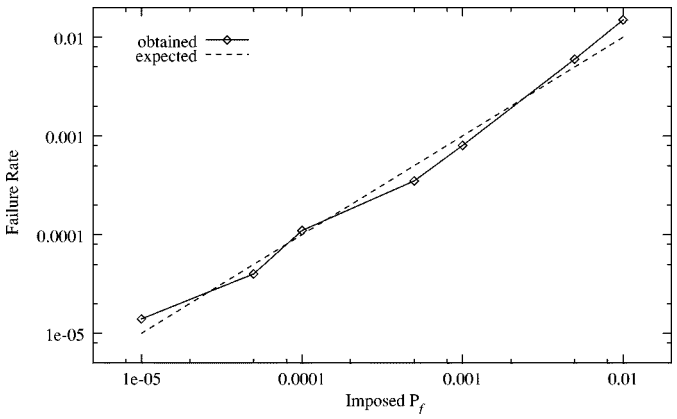


Fig. 4. Experimentally evaluated false detection rate versus the P_f imposed to choose the detection threshold. The theoretical trend (dashed line) is also shown.

II, where the average watermark strength is given along with the PSNR of the watermarked image. As it can be seen, the masking approach introduced in this work outperforms those generally used by conventional wavelet-based watermarking algorithms, in that it permits to hide a watermark with a higher energy, still granting invisibility.

In order to demonstrate the validity of the approach we followed for choosing the detection threshold, the false detection rate was estimated by attempting to detect many different watermarks in the same watermarked image ("Lena"). The threshold T_ρ was set in such a way to grant a given value of P_f. The trial was repeated for values of P_f ranging from 10^{-2} through 10^{-5} (in each case $10 \times P_f^{-1}$ different watermark were tested). The estimated failure rate is plotted in Fig. 4 versus the value of P_f used to set the corresponding thresholds. Experimental results (solid line) are very close to theoretical ones (dashed line), thus supporting the validity of the analysis developed in Section III.

The results presented in the following permit to appreciate the robustness of the new watermarking algorithm with respect to several common signal processing techniques. All the experiments were carried out on the Lena image (see Fig. 2) which was marked with $\alpha = 1.5$.

As a first experiment, JPEG coding with decreasing quality was applied to the watermarked image, and 1000 different watermarks (among them the embedded one) were tested for presence. In Fig. 5, the response of the detector to the embedded watermark is plotted against the compression ratio of the image,

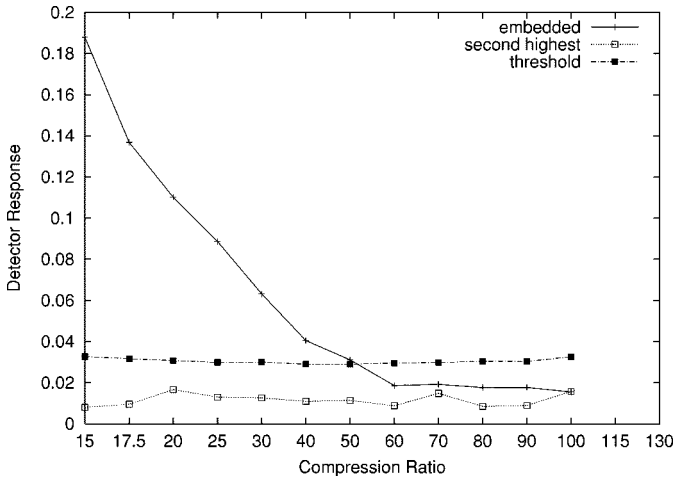


Fig. 5. Plot of the response of the correlation based detector to the embedded watermark (diamond markers), of the threshold (square markers) and of the second highest detector response, corresponding to increasing JPEG compression rates. P_f was set to 10^{-8} and 1000 marks were tested.

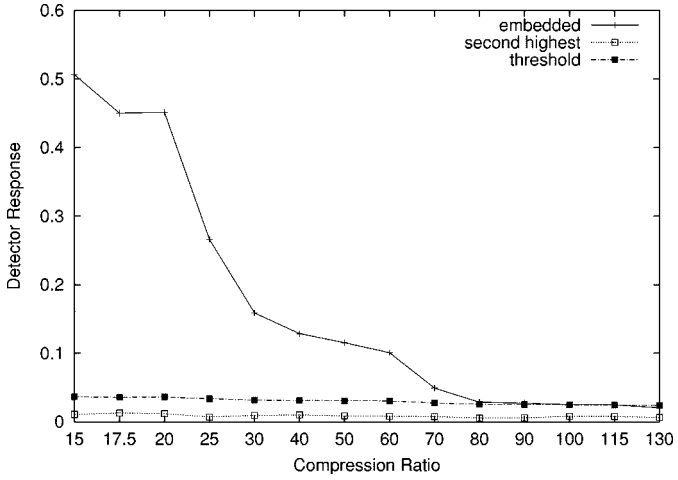


Fig. 6. Plot of the response of the correlation based detector to the embedded watermark (diamond markers), of the threshold (square markers) and of the second highest detector response, corresponding to increasing SPIHT compression rates. P_f was set to 10^{-8} and 1000 marks were tested.

along with the detection threshold, and the second highest response of the detector (i.e., the highest response among those produced by the detector when the 999 fake watermarks are tested). The detector response remains above the threshold until a compression ratio of 50 (corresponding to a quality factor of about 8%) is reached. The second highest detector response always stays below the threshold: the imposed P_f is, in fact, extremely low (10^{-8}) with respect to the number of tested watermarks.

The robustness of the proposed technique against DWT-based compression was also tested by using the SPIHT algorithm [22]. The results are plotted in Fig. 6. In this case the detector behaves significantly better than in the JPEG case, since a compression ratio of about 115 can be tolerated; this could be explained by the fact that the image quality obtained by SPIHT is higher than that obtained by JPEG at the same compression ratio, and, thus, the watermarking signal is better preserved by SPIHT.

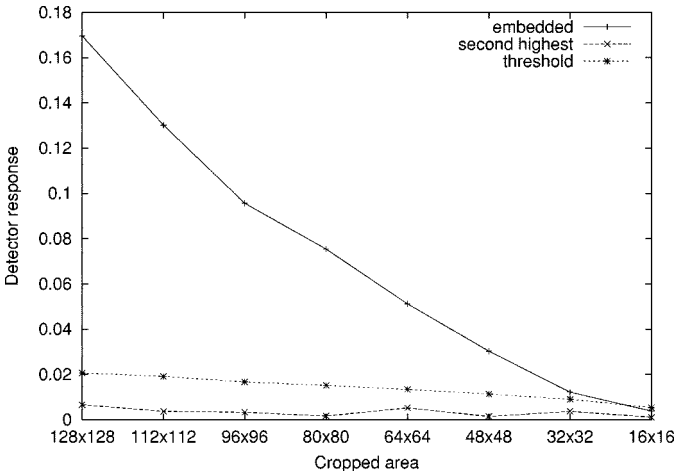


Fig. 7. Plot of the response of the correlation based detector to the embedded watermark (diamond markers), of the threshold (square markers) and of the second highest detector response, corresponding to decreasing cropping dimensions. P_f was set to 10^{-8} and 1000 marks were tested.

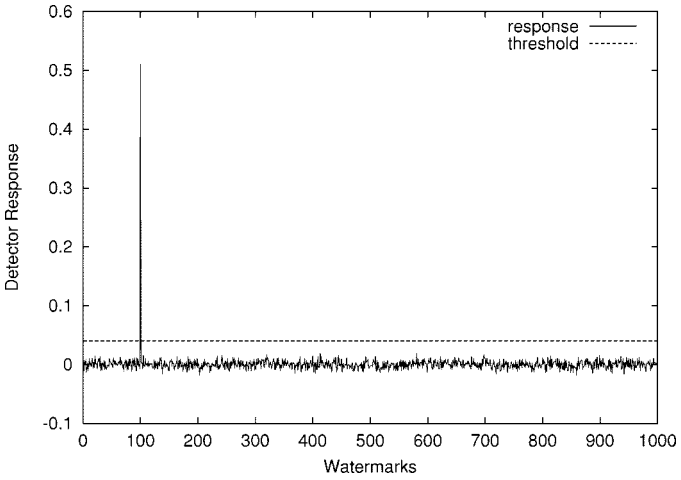
A. Geometric Manipulations

Robustness against geometric manipulations, e.g., resizing, cropping, or rotation, is very important because these manipulations are very common and usually do not degrade too much the quality of the image. The problem of recovering the watermark after the image has been geometrically distorted is often one of loss of synchronization; the watermark is still there but the decoder can not recover it. Among the possible solutions to this problem, the insertion of a synchronization pattern allowing to recover the original image geometry is the most common one [23], [24].

Image cropping deserves a separate discussion, both because of its importance (is one of the most common, usually non malicious, operations that can be applied to an image) and because it presents some peculiarities that make it different from the other geometric distortions. Two different problems must be coped with when considering cropping: loss of synchronization and loss of information. Synchronization problems derive from the fact that, due to cropping, the image space origin is shifted. To cope with this shift, we have to look for the watermark by considering all possible horizontal and vertical shifts. This may be a computationally expensive solution, however, since our detector is based on the correlation between the watermark and the host DWT coefficients, the exhaustive search can be performed efficiently through FFT computation [25]. As to the loss of information, it must be noted that when the image is cropped, part of or all the watermark information is discarded. It is necessary, then, that the watermark is spread all over the image, and that any image subpart contains enough detail to allow the watermark recovery.

Our experiments on cropping mainly aimed at evaluating the resilience of the watermark against the loss of information due to the removal of part of the image, thus witnessing the ability of the decoder to reveal the presence of the watermark even by analyzing a very small subpart of the watermarked image. The results we obtained (see Fig. 7) are quite impressive, since they demonstrate the excellent robustness of the proposed algorithm

(a)



(b)

Fig. 8. (a) Watermarked image "Lena" after an *implode* distortion and (b) the corresponding detector response to 1000 different watermarking codes (among them the embedded one). P_f was set to 10^{-8}.

with respect to cropping. It is, in fact, possible to correctly detect the watermark also when the cropped portion has a size of 32×32 pixels. A possible explanation for these excellent results is that, thanks to the suitability of DWT to model the HVS characteristics, high energy and thus very robust watermarks can be effectively hidden within the images.

The possibility of recovering the watermark even from a very small image subpart has some important consequences. Let us consider, for example, the result reported in Fig. 8, where the effect of a morphing attack is depicted. Very often, such an attack compromises watermark detection, since watermark synchronization is lost; on the contrary, here, as long as at least a small region of the image is left unchanged, the decoder can detect the watermark presence. Even more interestingly, the recovery of the watermark from small subimages can enable the watermarking of separate image objects, to deal, for example, with object-oriented image and video coding standards (e.g., JPEG 2000, MPEG4) [26]. Suppose each object is watermarked separately, and suppose that objects stemming from different documents are put together to compose a new image. If the watermark decoder does not know which objects have been used to build the new image, and if we assume that synchronization problems can be overcome through exhaustive searching in the

TABLE III
EVALUATION OF PERFORMANCES OF THE PROPOSED METHOD AGAINST THE RESULTS OBTAINED BY LETTING $w_0^\theta(i,j)$ VARY ONLY WITH θ, AND BY MAKING IT DEPEND ON THE LOCAL SUBBAND VARIANCE, EQ. (19). IN THE CASE OF JPEG OR SPIHT COMPRESSION, THE MAXIMUM COMPRESSION RATIO THE WATERMARK SURVIVED TO IS SHOWN, WHEREAS FOR CROPPING THE MINIMUM SIZE OF THE CROPPED AREA WHERE THE WATERMARK IS FOUND IS INDICATED

Attack	Our method	Variance based Mask	Constant Mask
JPEG	50:1	30:1	25:1
SPIHT	100:1	70:1	50:1
Despeckle + JPEG	25:1	20:1	10:1
Cropping	32×32	48×48	48×48

FFT domain, the decoder can still recover the watermark of any single object by looking for it in the whole image, due to its ability of detecting a watermark as soon as a small image subpart contains it.

B. Comparisons with Other Wavelet-Based Algorithms

To give a final assessment of the validity of our new watermarking algorithm, some comparisons were performed with other algorithms operating in the wavelet domain. Since the main novelty of our approach resides in the way the HVS characteristics are exploited to hide the watermark, comparisons were carried out by changing the way the watermark sequence is weighted before adding it to DWT coefficients, i.e., by changing the way $w_0^\theta(i,j)$ is calculated. All the other details are left unchanged. In particular, for each method the maximum α still granting invisibility was used, as in Table I. We tested our algorithm against the results obtained by letting $w_0^\theta(i,j)$ vary only with θ, i.e., by using a constant watermark weigh for each subband, and by making it depend on the local subband variance as in (19). Several kinds of attacks were carried out, and for each of them the maximum attack strength the watermark can survive was measured. The results we obtained on the Lena image are given in Table III. As it can be noted, the proposed method outperforms conventional techniques, particularly in the case of lossy compression, thus supporting for the major effectiveness of our approach in adapting the watermarking signal strength to the perceptual image content.

V. CONCLUSION

In this paper, a novel algorithm for image watermarking has been presented. The algorithm embeds the watermark code by modifying the DWT coefficients of the image, and exploits a model derived from image compression techniques for adapting the watermark strength to the characteristics of the HVS.

The performances of the novel algorithm are very good, experimental results, in fact, supported the suitability of DWT watermarking schemes for robustly hiding watermarks into images. In particular, the behavior of the watermark detector with respect to image cropping was surprisingly good. As a matter of fact, DWT schemes do not spread the watermark all over the image (as, for example, the methods based on full-frame DCT [1], [27] do), but, thanks to the similarity of DWT decomposition to the models of HVS, the watermarking energy can be kept so high that even a small portion of the image is sufficient to correctly guess the embedded code.

REFERENCES

[1] I. Cox and M. L. Miller, "A review of watermarking and the importance of perceptual modeling," in *Proc. Electronic Imaging*, Feb. 1997.

[2] M. D. Swanson, M. Kobayashi, and A. H. Tewfik, "Multimedia data-embedding and watermarking technologies," *Proc. IEEE*, vol. 86, pp. 1064–1087, June 1998.

[3] S. Craver, N. Memon, B. Yeo, and M. Yeung, "Resolving rightful ownerships with invisible watermarking techniques: Limitations, attacks, and implications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 573–586, May 1998.

[4] N. J. Jayant, J. Johnston, and R. Safranek, "Signal compression based on models of the human perception," *Proc. IEEE*, vol. 81, pp. 1385–1422, 1993.

[5] A. H. Tewfik and M. Swanson, "Data hiding for multimedia personalization, interaction, and protection," *IEEE Signal Processing Mag.*, vol. 14, pp. 41–44, July 1997.

[6] R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, "Perceptual watermarks for digital images and video," *Proc. IEEE*, vol. 87, pp. 1108–1126, July 1999.

[7] M. Charrier, D. S. Cruz, and M. Larsson, "JPEG2000, the next millenium compression standard for still images," in *Proc. IEEE Int. Conf. Multimedia Computing Systems'99*, Florence, Italy, July 7–11, 1999, pp. 131–132.

[8] B. G. Haskell, P. G. Howard, Y. A. LeCun, A. Puria, J. Ostermann, M. R. Civanlar, L. Rabiner, L. Bottou, and P. Haffner, "Image and video coding—Emerging standards and beyond," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 814–837, Nov. 1998.

[9] H. Inoue, A. Miyazaki, A. Yamamoto, and T. Katsura, "A digital watermark based on the wavelet transform and its robustness on image compression and transformation," *IEICE Trans. Fund. Electron., Commun., Comput. Sci.*, vol. E82-A, pp. 2–10, Jan. 1999.

[10] H.-J. M. Wang, P.-C. Su, and C.-C. J. Kuo, "Wavelet-based digital image watermarking," *Opt. Express*, vol. 3, no. 12, pp. 491–496, Dec. 7, 1998.

[11] X. Xia, C. G. Boncelet, and G. R. Arce, "A multiresolution watermark for digital images," in *Proc. 4th IEEE Int. Conf. Image Processing '97*, Santa Barbara, CA, Oct. 26–29, 1997, pp. 548–551.

[12] D. Kundur and D. Hatzinakos, "A robust digital image watermarking method using wavelet-based fusion," in *Proc. 4th IEEE Int. Conf. Image Processing '97*, Santa Barbara, CA, Oct. 26–29, 1997, pp. 544–547.

[13] C.-T. Hsu and J.-L. Wu, "Multiresolution watermarking for digital images," *IEEE Trans. Circuits Syst. II*, vol. 45, pp. 1097–1101, Aug. 1998.

[14] D. Kundur and D. Hatzinakos, "Digital watermarking using multiresolution wavelet decomposition," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 5, Seattle, WA, May 1998, pp. 2969–2972.

[15] W. Zhu, Z. Xiong, and Y.-Q. Zhang, "Mutliresolution watermarking for images and video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp. 545–550, June 1999.

[16] G. Nicchiotti and E. Ottaviani, "Non-invertible statistical wavelet watermarking," in *Proc. EUSIPCO'98*, vol. 4, Rhodes, Greece, Sept. 8–11, 1998, pp. 2289–2292.

[17] J. Wang and G. Wiederhold, "WaveMark: Digital image watermarking using daubechies' wavelets and error correcting codes," in *Proc. SPIE Int. Symp. Voice, Video, Data Communications*, Boston, MA, Nov. 1998.

[18] M. D. Swanson, B. Zhu, and A. H. Tewfik, "Multiresolution scene based video watermarking using perceptual models," *IEEE J. Select. Areas Commun.*, vol. 16, May 1998.

[19] A. S. Lewis and G. Knowles, "Image compression using the 2-D wavelet transform," *IEEE Trans. Image Processing*, vol. 1, pp. 244–250, Apr. 1992.

[20] A. Piva, M. Barni, F. Bartolini, and V. Cappellini, "Mask building for perceptually hiding frequency embedded watermarks," in *Proc. 5th IEEE Int. Conf. Image Processing '98*, Chicago, IL, Oct. 4–7, 1998, pp. 450–454.

[21] A. Watson, G. Yang, J. Solomon, and J. Villasenor, "Visibility of wavelet quantization noise," *Proc. IEEE*, vol. 6, pp. 1164–1175, Aug. 1997.

[22] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 243–250, June 1996.
[23] S. Pereira and T. Pun, "Fast robust template matching for affine resistant image watermarks," in *Proc. 3rd Information Hiding Workshop*, Dresden, Germany, Sept.–Oct. 1999, pp. 207–218.
[24] A. Piva, M. Barni, F. Bartolini, V. Cappellini, A. De Rosa, and M. Orlandi, "Improving dft watermarking robustness through optimum detection and synchronization," in *GMD Rep. 85, Multimedia Security Workshop ACM Multimedia'99*, Orlando, FL, Oct. 1999, pp. 65–69.
[25] T. Kalker, G. Depovere, J. Haitsma, and M. Maes, "A video watermarking system for broadcast monitoring," in *Proc. SPIE Security Watermarking Multimedia Contents*, vol. 3657, San Jose, CA, Jan. 1999, pp. 103–112.
[26] A. Piva, R. Caldelli, and A. De Rosa, "A DWT-based object watermarking system for mpeg-4 video streams," in *Proc. 7th IEEE Int. Conf. Image Processing '00*, Vancouver, BC, Canada, Sept. 10–13, 2000.
[27] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain system for robust image watermarking," *Signal Process.*, vol. 66, no. 3, pp. 357–372, May 1998.

Mauro Barni (S'88–M'96) was born in Prato, Italy, in 1965. He graduated in electronic engineering from the University of Florence, Florence, Italy, in 1991, and received the Ph.D. degree in informatics and telecommunications in 1995.

From 1995 to 1998, he was a Postdoctoral Researcher with the Department of Electronic Engineering, University of Florence. Since September 1998, he has been with the Department of Information Engineering, University of Siena, Siena, Italy, where he is an Assistant Professor. His main interests are in the fields of digital image processing and computer vision. His research activity is focused on the application of image processing techniques to cultural heritage analysis and preservation, copyright protection of multimedia data (digital watermarking), and transmission of image and video signals in error-prone wireless environments. He has published more than 100 papers on these topics in international journals and conferences. He holds two Italian patents in the field of digital watermarking.

Franco Bartolini (M'96) was born in Rome, Italy, in 1965. In 1991, he graduated (*cum laude*) in electronic engineering from the University of Florence, Florence, Italy. In November 1996, he received the Ph.D. degree in informatics and telecommunications from the University of Florence.

He is now a Postdoctoral Researcher with the University of Florence. His research interests include digital image sequence processing, still and moving image compression, nonlinear filtering techniques, image protection and authentication (watermarking), image processing applications for the cultural heritage field, signal compression by neural networks, and secure communication protocols. He has published more than 80 papers on these topics in international journals and conferences. He holds two Italian patents in the field of digital watermarking.

Dr. Bartolini is a member of IAPR.

Alessandro Piva was born in Florence, Italy, in 1968. In 1995, he graduated (*cum laude*) in electronic engineering from the University of Florence, from which he received the Ph.D. degree in informatics and telecommunications in 1999.

He is now a Postdoctoral Researcher with the University of Florence. His research activity is focused on multimedia systems, digital image sequence processing, image protection and authentication (watermarking), image processing techniques for cultural heritage applications, and secure communication protocols. He has published more than 40 papers on these topics in international journals and conferences. He holds two Italian patents in the field of digital watermarking.