A wavelet based method for high frequency subbands watermark embedding

Patrizio Campisi, Alessandro Neri, Marco Visconti

Dipartimento di Ingegneria Elettronica, Università degli Studi di Roma "Roma Tre",
Via della Vasca Navale 84, 00146 Roma, Italy
E-mail: (campisi,neri)@ele.uniroma3.it

ABSTRACT

In this paper a novel wavelet based robust blind watermarking scheme is presented. Taking into account the property of the human visual system, the proposed scheme embeds the watermark into the wavelet coefficients representative of the image's areas highly dense of details. More specifically a two stages unconventional wavelet decomposition is used. In fact, after having performed a 2D one-level wavelet decomposition, the high frequency wavelet subbands are further wavelet decomposed, thus obtaining those coefficients where the mark is finally embedded. Experimental results demonstrate that this embedding strategy allows to greatly increase the local ratio between the mark and the original image energies, thus leading to robustness even against severe image degradations, without transparency loss.

Keywords: Digital watermarking, perceptual models, wavelet transform.

1. INTRODUCTION

The fast diffusion of digital multimedia documents, due to the easiness of their storage and transmission, along with the wide spread of the Internet, has lead to an increasing interest in developing techniques for protection of the ownership right (copyright). A feasible solution consists in embedding a "mark", called "watermark", into the data in such a way that it does not modify them from a perceptual point of view (transparency) while its unauthorized removal should make the data useless (robustness) [1]. In the last years several watermarking schemes have been proposed. With regards to the domain where the watermark embedding occurs, we can distinguish between methods operating in the spatial domain [2], in the DCT domain [3], in the frequency domain [4], and in the wavelet domain [5], [6], [7], [8]. Among all of them, the ones operating in the wavelet domain seem to be very promising because they provide a simultaneous spatial localisation and a frequency spread of the watermark within the host image.

In this work we present a novel robust watermarking method operating in the wavelet domain, which does not require the

Multimedia Systems and Applications III, Andrew G. Tescher, Bhaskaran Vasudev, V. Michael Bove, Jr.,
Editors, Proceedings of SPIE Vol. 4209 (2001) © 2001 SPIE · 0277-786X/01/$15.00

original image to detect the watermark. In order to improve transparency with respect to the existing methods, in the spatial domain we split the host image into a “low frequency” image and a “high frequency” image. This latter, taking into account the image’s details, allows to “hide” the watermark without loss of transparency because of the incapability of the human visual system (HVS) to distinguish between details in areas where their density is high [9]. The watermark embedding is then performed in the wavelet domain. To this aim, contrary to existing methods, after having performed a one-level bidimensional wavelet decomposition, the subbands corresponding to the “high frequency” image’s details (HL, LH, HH) are further wavelet decomposed and the watermark is embedded into the subbands of the second stage of the wavelet decomposition. This allows to increase the energy of the embedded watermark, making the scheme more robust without affecting transparency, due to the aforementioned property of the HVS.

The paper is organized as follows: in Section 2 the proposed watermarking embedding scheme is outlined. The watermarking detection procedure is described in Section 3. Finally the experimental results and the conclusions are drawn in Section 4.

2. WATERMARKING EMBEDDING PROCEDURE

A watermarking embedding scheme, in order to be effective in intellectual property protection (copyright) applications, must be robust against malicious attacks, whose goal is the complete or even partial watermark removal, such as lossy compression, filtering, geometrical distortions, cropping and other kind of processing. Nevertheless the embedded watermark has to be transparent, that is not perceptible, in order not to impair the quality of the host image. These two requirements, that by far represent the main characteristics that a watermarking scheme, devised for copyright assessment, must have, are in contrast each other and a tradeoff must be found.

In the recent literature several watermarking schemes, employing models of the HVS, have been proposed to device a technique to embed a robust and perceptually invisible watermark into the host image [10], [11]. They all rely on the embedding of the watermark into perceptually insignificant parts of an image in order to guarantee transparency, although operating in different domains such as the spatial domain [2], the DCT domain [3], the wavelet domain [6], [7], [9].

In the method here presented a high degree of transparency is achieved into two stages by using visual models both in the spatial and in the wavelet domain in order to understand where and how to embed the watermark by employing the image’s characteristics to come up with an embedding scheme robust and transparent. A preprocessing is carried out in the spatial domain to extract that portion of the host image where the watermark can be embedded without impairing the image’s visual

appearance (Stage 1, see Fig.1). Finally the embedding is performed in the wavelet domain through an unconventional two-level wavelet decomposition whose goal is the extraction of the proper wavelet coefficients that can be modified by adding the watermark coefficients thus leading to robustness and transparency (Stage 2, see Fig.1). The block diagram describing the whole procedure is sketched is Fig.1, and the two stages are detailed described in the next two subsections respectively.

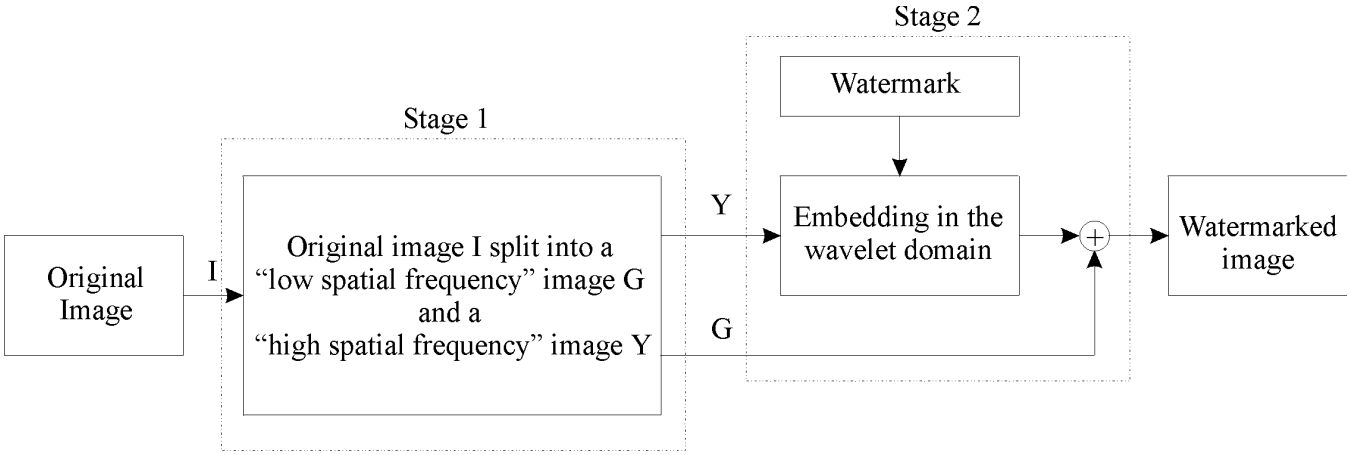

Figure 1: General block scheme of the watermarking embedding procedure.

2.2 Preprocessing in the spatial domain

In summary, given the host image $I[n_1,n_2]$, with $n_1 \in N_1$, $n_2 \in N_2$, its regions characterized by high spatial frequency are extracted thus obtaining an image $Y[n_1,n_2]$, called *layer*. The "residual" image $G[n_1,n_2]=I[n_1,n_2]-Y[n_1,n_2]$, called *ground*, takes into account the regions of the original image characterized by low spatial frequency components. Psychophysical studies have pointed out that the HVS is less sensitive when a noise-like signal, such as a watermark, is introduced in areas with high details density, that is characterized by high local variance. Therefore we resort to embed the watermark into $Y[n_1,n_2]$. The extraction of the *layer* is accomplished by evaluating the local image activity. Specifically, since the host image $I[n_1,n_2]$ can be thought as a sample from a non stationary random field, as indicator of the local image activity we adopt here the local variance $\sigma_I^2[n_1,n_2]$, estimated on a neighborhood $J(n_1,n_2)=\{(p,q)\,|\,n_1-M\le p\le n_1+M,\ \&\ n_2-M\le q\le n_2+M\}$ of size (2*M*+1)x(2*M*+1) of the site (n_1, n_2), i.e.:

$$\sigma_I^2[n_1,n_2]=\frac{1}{(2M+1)^2}\sum_{(p,q)\in J(n_1,n_2)}\left|I[p,q]-m_I[n_1,n_2]\right|^2. \qquad (1.a)$$

where $m_I[n_1,n_2]$ is the local mean estimated on the same neighborhood:

$$m_I[n_1,n_2]=\frac{1}{(2M+1)}\sum_{(p,q)\in J(n_1,n_2)}I[p,q]. \qquad (1.b)$$

Then, if the local variance $\sigma_I^2[n_1,n_2]$ exceeds a threshold γ, the corresponding pixel of the activity mask $M_I[n_1,n_2]$, with $n_1 \in N_1$, $n_2 \in N_2$, is set to 1 otherwise to 0, as expressed by Eq.(2),

$$M_I[n_1,n_2] = \begin{cases} 1 & \sigma_I^2[n_1,n_2] \geq \gamma \\ 0 & \sigma_I^2[n_1,n_2] < \gamma \end{cases}. \tag{2}$$

The portion of the original image $I[n_1,n_2]$ characterized by a high degree of local activity, $Y[n_1,n_2]$, is thus obtained by the product between $I[n_1,n_2]$ and the activity mask, i.e.:

$$Y[n_1,n_2] = I[n_1,n_2] \cdot M_I[n_1,n_2]. \tag{3}$$

The threshold γ can be adapted to the global image activity. Adaptive rules can be based, for instance, on the evaluation of the *N*-th order statistics of the actual local variance distribution, so that the threshold is set to a predefined percentile of $\sigma_I^2[n_1,n_2]$. In other words, the layer $Y[n_1,n_2]$ is constituted by the fraction α of the original image consisting of the most active regions. The remaining pixels are collected into the image $G[n_1,n_2]$. An example of the aforementioned splitting procedure is depicted in Fig.2, where, along with the original image Lena, the two images corresponding to the high spatial frequency image's regions (*layer*) and to the low spatial frequency image's regions (*ground*) are isolated.

Figure 2: From lest to right: original Lena, layer image of Lena, ground image of Lena.

2.2 Embedding in the wavelet domain

The image $Y[n_1,n_2]$, obtained according to the procedure described in the subsection 2.1, becomes the new host image where to embed the watermark. The embedding is performed in the wavelet domain since the wavelet transform can provide both spatial and frequency localization and this well suits with the behavior of the HVS. Therefore, by taking into account the properties of the HVS we can select the subband where to embed the watermark in such a way that the robustness is increased without impairing the perceptual appearance of the host image.

More specifically an unconventional two-level multiresolution-like wavelet decomposition is performed on the *layer* image $Y[n_1, n_2]$. With reference to Fig.3, the first level of the multiresolution decomposition is obtained by performing a discrete wavelet transform (DWT) onto $Y[n_1, n_2]$. In particular an 8 tap Daubachies filter is used.

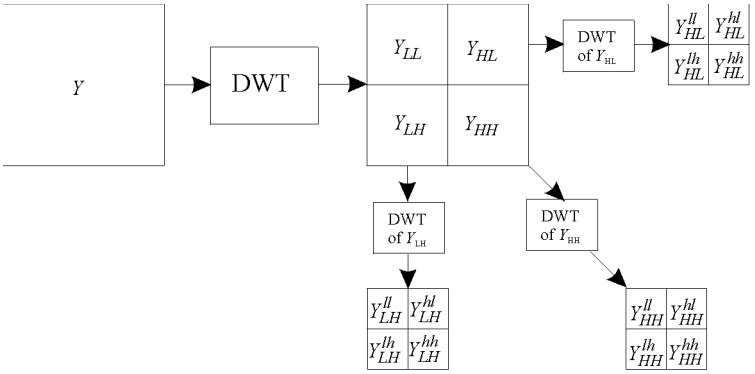

Figure 3: Two level multiresolution-like wavelet decomposition of the layer (Y).

This leads to the subband $Y_{LL}[n_1, n_2]$, which represents the image at a coarser resolution, plus the subbands, $Y_{LH}[n_1, n_2]$, $Y_{HL}[n_1, n_2]$, and $Y_{HH}[n_1, n_2]$, accounting respectively for the "horizontal", "vertical", and "diagonal" image's details also at coarser resolution. Then, the "father" subbands $Y_\eta[n_1, n_2]$, being $\eta \in (LH, HL, HH)$, that is the subbands corresponding to the high frequency components of the image, are further wavelet decomposed, thus obtaining a set of twelve "child" subbands $Y_\eta^\nu[n_1, n_2]$, being $\eta \in (LH, HL, HH)$ and $\nu \in (ll, lh, hl, hh)$ (see Fig. 3). It is worth noting that each "child" subband contains only information related to the high frequency father subbands of the high variance area of the original image $I[n_1, n_2]$. The mark $W[n_1, n_2]$ is then embedded in the active regions of every child subband. Specifically, the local variance of $Y_\eta^\nu[n_1, n_2]$ is computed, based on Eq. (1), and its value is thresholded according to Eq. (2), to form a visibility map $M_{Y_\eta^\nu}[n_1, n_2]$. Then, $Y_\eta^\nu[n_1, n_2]$ is modified according to the following law:

$$\hat{Y}_\eta^\nu[n_1, n_2] = Y_\eta^\nu[n_1, n_2] + kW[n_1, n_2] \cdot M_{Y_\eta^\nu}[n_1, n_2] \tag{3}$$

where $\hat{Y}_\eta^\nu[n_1, n_2]$ represents the marked child subband and *k* establishes the embedding intensity. The mark $W[n_1, n_2]$ is a binary patter having the same size of the child subbands and whose number of "zeros" equals the number of "ones".

After the embedding, the three marked father subbands $\hat{Y}_{LH}[n_1, n_2]$, $\hat{Y}_{HL}[n_1, n_2]$, $\hat{Y}_{HH}[n_1, n_2]$ are obtained by applying the inverse wavelet transform (IDWT) to the marked child subbands. The three marked father subbands $\hat{Y}_{LH}[n_1, n_2]$, $\hat{Y}_{HL}[n_1, n_2]$, and $\hat{Y}_{HH}[n_1, n_2]$, along with the unmarked component $Y_{LL}[n_1, n_2]$, once inverse wavelet transformed, lead to the marked replica of the layer named $\hat{Y}[n_1, n_2]$. The marked original image, $\hat{I}[n_1, n_2]$, is finally obtained as

$$\hat{I}[n_1, n_2] = \hat{Y}[n_1, n_2] + G[n_1, n_2] \tag{4}$$

3. WATERMARKING DETECTION PROCEDURE

To detect the watermark, the original unmarked image is not required. Using the procedure described in the Section 2, we perform the discrete wavelet decomposition on the potentially marked image, say $\tilde{I}[n_1, n_2]$, without splitting the image into *layer* and *ground* since, in case of attack, this operation will give rise to loss of information. For each child subband $\tilde{I}_\eta^\nu[n_1, n_2]$, with $\eta \in (LH, HL, HH)$ and $\nu \in (ll, lh, hl, hh)$, the correlation with the original watermark pattern is evaluated as follows:

$$C_\eta^\nu = \sum_{n_1 \in N_1} \sum_{n_2 \in N_2} W[n_1, n_2] \cdot I_\eta^\nu[n_1, n_2], \quad for\ \eta \in \{LH, HL, HH\}, \nu \in \{ll, lh, hl, hh\}. \tag{5}$$

Then, the value

$$H = \frac{1}{N^2} \prod_{\eta \in I_\eta} \left(\prod_{\nu \in I_\nu} C_\eta^\nu \right), \quad where \quad I_\eta = \{LH, HL, HH\},\ I_\nu = \{ll, lh, hl, hh\} \tag{6}$$

representing a global correlation coefficient, normalized with respect to the images' dimensions, assumed to be without loss of generality *N* x *N*, is computed. Finally *H* is compared with a threshold to authenticate the candidate image $\tilde{I}[n_1, n_2]$. To adapt the detection procedure to the actual image, the threshold evaluation is performed in a similar way. That is, the threshold is set proportional to the value λ obtained by applying Eqs. (5) and (6), to a unitary mark (i.e. $W[n_1, n_2] = 1$) :

$$\lambda = \frac{1}{N^2} \prod_{\eta \in I_\eta} \left(\prod_{\nu \in I_\nu} \tilde{\lambda}_\eta^\nu \right), \quad where \quad I_\eta = \{LH, HL, HH\},\ I_\nu = \{ll, lh, hl, hh\} \tag{7}$$

where

$$\tilde{\lambda}_\eta^\nu = \sum_{n_1 \in N_1} \sum_{n_2 \in N_2} I_\eta^\nu[n_1, n_2], \quad for\ \eta \in \{LH, HL, HH\}, \nu \in \{ll, lh, hl, hh\} \tag{8}$$

In addition, to take into account the structure of the image, $\tilde{I}[n_1, n_2]$, the threshold is normalized with respect to its standard deviation $\sigma_{\tilde{I}}[n_1, n_2]$, thus obtaining the following normalized validation statement

$$\frac{|H|}{\sigma_{\tilde{I}}[n_1, n_2]} \geq |\lambda|. \tag{9}$$

4. EXPERIMENTAL RESULTS AND CONCLUDING REMARKS

The effectiveness of this method has been tested on several standard images using random binary patterns as mark, as well as some structured patterns. In Fig.4 and Fig.5 some of the test images that have been used and the corresponding marked counterparts are shown respectively.

Fig. 4: Original images: Barbara (left), Goldhill (right).

Fig. 5: Watermarked images: Barbara (left), Goldhill (right).

Several distortions have been applied to the watermarked images in order to test the robustness of the proposed algorithm. In particular its resistance to the JPEG compression has been investigated. We embed a watermark into the host image, which is then JPEG compressed using a compression percentage equal to 80%. Then, using as target watermark 499 random generated marks plus the one (no. 251) really embedded into the host image, we apply the detection procedure described in Section 3. In Fig.6 the detector's responses for each of the watermarks are shown. The dotted horizontal line represents the threshold λ, that is the lower bound for the validation statement (see. Eq. (9)). The detector validates only the presence of the watermark no. 251 in the image.

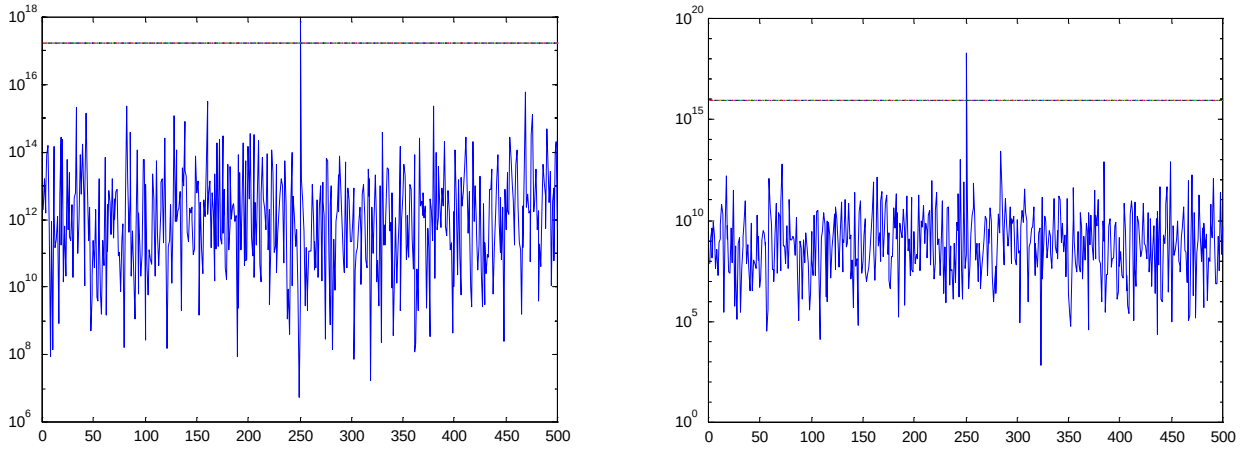

Fig. 6: Detector responses for the jpeg compressed images (compression percentage 80%): Barbara, (left), Goldhill (right). Embedded watermark number 251.

As illustrated by Figs. 7 and 8, the detector's performances are not affect by adding a Gaussian noise, even at very low signal to noise ratios. The method has shown robustness even to cropping as witnessed by Figs. 9-10. Moreover tests have been performed that show the good performances of the algorithm when blurring occurs or elimination of vertical, horizontal, and diagonal lines from the watermarked image, corresponding to the elimination of some spatial frequencies, is done.

Finally, robustness to multiple watermarking has been investigated. Specifically, three different watermarks (no. 201, 301, 401) are embedded into the same host image. The detector is able to reveal the presence of these marks out of other 497 random marks, as shown in Fig. 11.

Fig. 7: Images corrupted by Gaussian noise: Barbara (left), Goldhill (right).

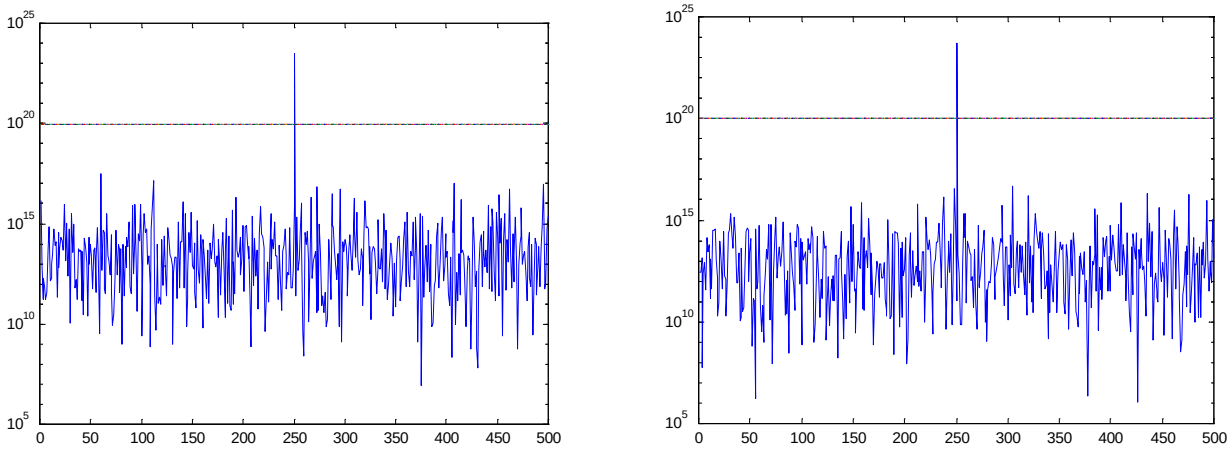

Fig. 8: Detector responses for the noisy images: Barbara (left), Goldhill (right). Embedded watermark number 251.

Fig. 9: Cropped images: the selected parts have been given a higher contrast for presentation purpose (area selected 96x96).

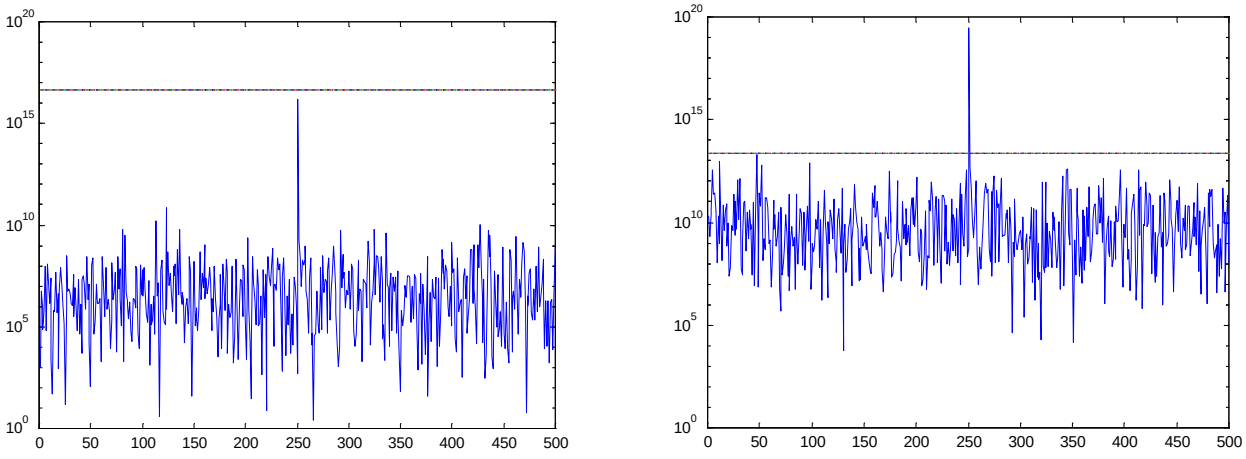

Fig. 10: Detector responses for the cropped images: cropped (96x96) Barbara (left), cropped (96x96) Goldhill (right).

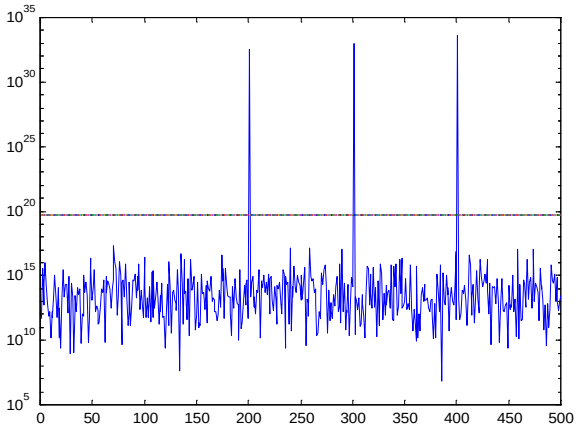

Fig. 11: Detector responses to multiple watermarked Barbara (embedded watermarks number 201, 301, 401).

In summary we have presented an embedding scheme, operating partially in the spatial domain and partially in the wavelet domain, which allows to achieve a high degree of robustness without affecting the transparency. The rational behind our approach lies on the fact that the watermark embedding is transparent when the mark is hidden into that portion of the image that has a high degree of spatial local activity. Moreover the energy of the watermark can be greatly increased when it is embedded into the image's wavelet high frequency subbands, because of the inability of the HVS to catch differences when a noise-like image, such as a watermark, is embedded into the areas of the host image rich of details.

5. REFERENCES

[1] R.M. Wolfgang, C.I. Podilchuk, and E.J. Delp, "*Perceptual Watermarks for Digital Images and Video*", Proceedings of the IEEE, vol. 87, no.7, July 1999.

[2] N. Nikolaidis, and I. Pitas, "*Robust image watermarking in the spatial domain*", Signal Processing, vol. 66, no.3, pp. 385-403, 1998.

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

[4] R.M. Wolfgang, and E.J. Delp, "*A watermark for digital images*", Proc. IEEE Int. Conf. on Image Processing 1996, Lausanne, Switzerland, September 16-19, 1996, pp. 219-222.

[5] R. Dugad, K. Ratakonda and N. Ahuja, "*A new wavelet-based scheme for watermarking images*", Proc. IEEE Int. Conf. on Image Processing 1998, Chicago, IL, October 4-7, 1998, pp. 419-423.

[6] D. Kundur, D. Hatzinakos, "*Digital watermarking using multiresolution wavelet decomposition*", Proc. IEEE Int. Conf. On Acoustic, Speech and Signal Processing, 1998, vol.5, pp. 2969-2972.

[7] H. Inoue, A.Miyazaki, T. Katsura, "*An image watermarking method based on the wavelet transform*", Proc. IEEE Int. Conf. on Image Processing 1999, Kobe, Japan, October 25-28, 1999, pp. 296-300.

[8] H. J. M. Wang, P. C. Su, and C. C. J. Kuo, "*Wavelet-based blind watermark retrieval technique"*, in Proc. SPIE, vol 3528, Conf. On Multimedia Systems and Applications, Boston, MA, November 1998.

[9] M. Barni, F. Bartolini, V. Cappellini, A. Lippi, A. Piva, "*A DWT-based algorithm for spatio-frequency masking of digital signatures"*, Proc.of SPIE vol. 3657, Security and Watermarking of Multimedia Contents, Electronic Imaging '99, San Jose, CA, January 1999.

[10] I. J. Cox and M. L. Miller, "*A review of watermarking and the importance of perceptually modeling*", in Proc. SPIE, Conf. Human Vision and Electronic Imaging-II, San Jose, CA, pp. 92-99, February 1997.

[11] J. F. Delaigie, C. De Vleeschouwer, and B. Macq, "*Watermarking algorithm based on human visual model"*, Signal Processing, vol. 66, no. 3, pp. 319-335, May 1998.