Wavelet packet based digital watermarking

Jacques Lévy Véhel, Anne Manoury

Wavelet packet based digital watermarking

J. Lévy Véhel
Projet Fractales, INRIA Rocquencourt,
78153 Le Chesnay Cedex, France
and Ircyn, BP 92101, 1 rue de la Noé,
44321 Nantes Cedex 3, France
jlv@bora.inria.fr

A. Manoury
Irccyn, BP 92101, 1 rue de la Noé,
44321 Nantes Cedex 3, France
Anne.Manoury@irccyn.ec-nantes.fr

Abstract

We present a method for digital image watermarking based on the modification of certain subsets of the wavelet packet decomposition. These subsets are determined both from a secret key and an image dependent procedure that chooses a best basis from an energy criterion. The mark is set by imposing a parity constraint at each level of the decomposition. We elaborate on the choice of some of the parameters of the model, showing how they can be tuned so as to obtain good resistance to attacks. Examples are displayed to assess the validity of our approach.

1 Introduction

Digital image watermarking has attracted a lot of interest in recent years, due in particular to the development of Internet and the World Wide Web. The aim is to protect ownership by including in the image a copyright information. This information, or mark, has to be set in such a way that it is invisible: indeed, it must not alter the viewing content and, in addition, it should not be easy to remove. Furthermore, the mark must be resistant to attacks directed at erasing it. Such attacks are of two kinds, whether one uses cryptographic or image processing methods. We shall be concerned in this paper only with the second type of attacks.

A number of methods have been proposed to insert robust and invisible watermarks. Some operate directly in pixel space [1], other in a transform domain, such as Fourier [2] or DCT [3]. We propose here to study a wavelet packet based watermarking procedure. Working in the wavelet domain yields a number of advantages: First, it allows to control in a precise way the location both in space and scale of the mark. Second, wavelet coefficients give a structured way of representing the information: as is well known for instance from studies in image compression, in most cases, only a few coefficients are large, indicating where the information lies in a given image in terms of scale and space. This is useful both for invisibility and robustness concerns. Using wavelet packets adds another degree of freedom because it allows to select frequency independently of scale. Finally, wavelet based algorithms are fast and allow to reconstruct the image.
Section 2 briefly recalls some basic facts about wavelet packets decomposition. Section 3 presents our watermarking algorithm. Some details on the choice of various parameters are discussed in section 4. Finally, section 5 presents numerical experiments.

2 Recalls on Wavelet Packets

Wavelet packet decomposition (WPD) is a generalization of the dyadic wavelet transform (DWT) where the low pass parts are further analyzed. Figure 1 shows the decomposition tree of an image: the coefficients of each packet are obtained by successive filtering and decimation along lines and columns. As is apparent from figure 1, this representation is redundant, and it is possible to extract a basis by selecting coefficients in an appropriate way. Figure 2 shows three levels of decomposition of an image, along with one of the bases corresponding to a particular tiling of the space/frequency domain. Usually, one defines a "best basis" as a basis that optimizes a certain criterion. A popular choice is to minimize the entropy of the representation [4].
A word on notation: $C_{p,i,j}$ denotes the packet at resolution p in the frequency region indexed by i, j. Individual values of the coefficients in this packet are denoted $C_{p,i,j}(k)$, where k codes for the spatial translation.

3 Watermarking based on the parity of certain subsets associated to a best basis

Our method consists in inserting a mark by modifying the best basis associated to an image so that it respects some parity constraints. More precisely, the mark will be a sequence of 0-s and 1-s ; we first compute the best basis using an energy criterion detailed below (the classical entropy criterion is not fitted to our needs because it is not robust to compression). We then extract certain subsets from this

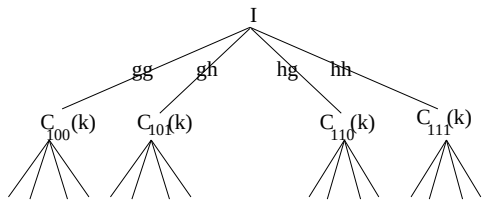

Figure 1. Quaternary wavelet packet decomposition tree, obtained by successive filtering and decimation along the lines and the columns. h and g denotes respectively the high pass and low pass filters of the DWT.

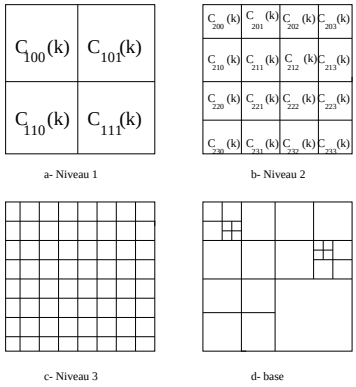

Figure 2. (a), (b), (c) : first three levels of the space/frequency decomposition of an image ; (d) : example of an extracted basis

basis, using a private information coded in a secret key. The image is modified so that, at each level p of the decomposition, the number of packets in the best basis which are selected by the secret key is odd or even according to whether the p^{th} element in the mark is 0 or 1. This procedure inserts in some sense a "virtual" mark, since it is set on the *structure* of the best basis. We now make the different steps precise (see figure 3).

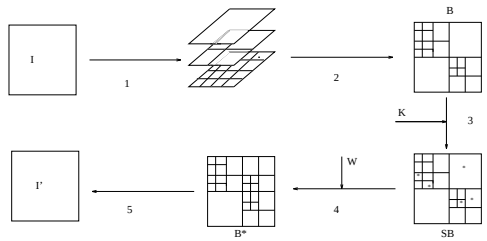

Figure 3. Different steps in the watermarking process.

- step 1: Wavelet packet decomposition.
- step 2: Best basis selection.
  Since the mark is set on the best basis, the basis should be both robust to attacks and contain enough vectors so that sufficiently rich marks may be implemented. For instance, we do not wish to select only packets at the highest or lowest resolution levels because they would not fulfill these constraints. The following criterion for best basis selection helps to attain our objectives. In short, we shall elect a packet to be in the basis B if it has sufficient energy and if its offsprings do not share this property. Formally: $C_{p,i,j} \in B$ iff $C_{p,i,j} \in \mathcal{F}_\lambda$ and $C_{p+1,2i,2j} \notin \mathcal{F}_\lambda$, where the set $\mathcal{F}_\lambda$ is defined in a recursive way:

$$\mathcal{F}_\lambda = \{C_{0,0,0}\} \bigcup \{C_{d,x,y} \ / \ \forall(\delta_1,\delta_2) \in \{0,1\}^2, \sum_k |C_{d,x+\delta_1 r_1, y+\delta_2 r_2}(k)|^2 \geq \lambda, r_1 = 1 - 2 \times (x[2]), r_2 = 1 - 2 \times (y[2]), C_{d-1,E(x/2),E(y/2)} \in \mathcal{F}_\lambda\}$$

  where $x[2]$ means x modulo 2. The best basis is thus of maximal depth and contains only packets which have "sufficient" energy: a packet is in the best basis if it belongs to $\mathcal{F}_\lambda$ but its offsprings do not. $\mathcal{F}_\lambda$ is the set of packets which have energy larger than λ and such that their brothers, father and uncles also have energy larger than λ. The important question of the optimal choice of the threshold λ is addressed in the next section.
- step 3: Extraction of the sub-bases.
  Let $\mathcal{B}$ be the set of all admissible bases corresponding to the image. We define an operator E :

$$\begin{array}{ccc} \mathcal{B} & \longmapsto & \mathcal{P}(\mathcal{B}) \\ B & \longmapsto & SB = B_1, B_2, ..., B_m \end{array}$$

  where $B_1, B_2, ..., B_m$ are disjoint subsets of B on which the mark will be implemented. SB is the secret key associated to the process and we have : $SB \subset B$. Note that, of course, the extraction operator E should not easily be invertible.

  To each sub-base B_i, we then associate a scalar e_i : e_i is 0 if the number of elements in B_i is even and is 1 otherwise (see figure 4).

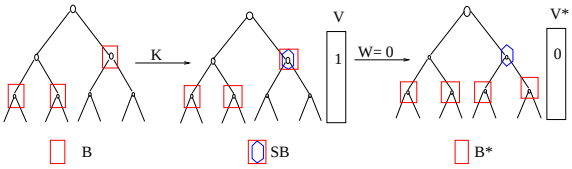

Figure 4. Encoding of the mark on a sub-base.

- step 4: Encoding of the mark.
  We first modify SB according to the watermark that identifies the owner of the image. This is modeled by an operator :

$$\begin{array}{ccc} \mathcal{P}(\mathcal{B}) \times \mathcal{W} & \longmapsto & \mathcal{P}(\mathcal{B}) \\ (SB, W) & \longmapsto & SB^* = B_1^*, B_2^*, ..., B_m^* \end{array}$$

  where the set of watermarks $\mathcal{W}$ is simply $\{0,1\}^m$. SB^* is computed in the following way:

1. if $w_i = e_i$, then the elements of SB^* at are equal to those of SB
2. otherwise, we suppress the element of SB that has highest energy and add at scale below its four offsprings to fill B. SB is filled by the offsprings selected by E.
3. the same procedure is performed iteratively for all the subset.

Once SB^* is completely defined, we modify the image to get the corresponding best basis B^*. To this end, the coefficients of the offsprings of the elements suppressed from SB are increased in absolute value so that the energy in their packet becomes larger than λ. This procedure is iterated for each subset starting from the lowest resolution (see figure 4). Finally, to make the process more resistant to attacks, we slightly modify the whole image in the following way : we multiply those coefficients which are "close" to the thresthold by a value that move them away from the thresthold.

Finally, the decoding process simply follows the same steps 1,2,3 as above, and then replaces step 4 by the computation of the parity of the number of vectors at each level of the m secret sub-bases. The presence of the mark is then classicaly assessed through computation of the Hamming distance between the observed parity vector and the theoretical one. Note that the decoding process does not require the knowledge of the original image.

4 Fine tuning of the parameters

Our method requires that a number of choices be made regarding the various parameters. In general, it is a delicate matter to derive values which are universally optimal. This would imply the knowledge of all possible images and all possible attacks, which is clearly impossible. A way out is to restrict the considered range of images and/or attacks. For some applications, for instance photographic reproduction of artworks, it is quite natural to focus on a given type of images. In this paper, we will rather specify a set of possible attacks and keep all freedom on the kind of marked images.

We have chosen to use as a model for the set of all possible attacks the *Stirmark* software, which has served as a benchmark for testing the robustness of watermarks by various authors [5]. Our aim in this section is thus to try and set some of the parameters of our method so as to maximize its resistance to the attacks implemented in *Stirmark*. More precisely, we will consider the choice of the energy threshold in the best basis selection. To simplify notation, we denote from now on c_i the *energy* of the packet at "location" i, where i an index that subsumes the information of scale and frequency. Obvious requirements concerning the energy threshold λ are the following : on the one hand, λ should be set small enough so that sufficiently many packets have energy larger than λ, and that a best basis may be found. On the other hand, a large λ is desirable, because selection of low energy packets should be forbidden (they are not robust to attacks such as compression or filtering). If we consider specifically the problem of robustness to attacks, something more precise may be said. Let I be the set of the indices of all wavelet packets coefficients, J be the subset of all indices in the marked image such that the corresponding coefficients are eligible to be in the best basis, i.e. $i \in J$ if $c_i > \lambda$, and K be the complement of J in I. Since the secret key is chosen independently of the best basis, an ideal choice for λ (as far as robustness is concerned) would be one that makes both the sets J and K stable under the action of any attack. In other words, an ideal situation would be one where $c_i > \lambda \Rightarrow d_i > \lambda$ and $c_i < \lambda \Rightarrow d_i < \lambda$, where c_i denotes the energy of a packet in the marked image and d_i the corresponding energy in the image after an attack. Indeed, such a choice would insure that the same coefficients are considered whether the image has been attacked or not, yielding in turn the same parity at each level, and making the mark insensitive to attacks.

While it is not possible to find such a λ in general, we may try to be as close as possible to this ideal case. In this view, a reasonable approach is to use an asymmetric procedure and to look for two reals (λ, λ') so that the number of coefficients for which $c_i - \lambda$ and $d_i - \lambda'$ have opposite signs is minimum in some sense. Denoting $\overline{e}$ and $\underline{e}$ the maximum and minimum of the energy of all packets, we consider the following stochastic modeling of our problem.

For a fixed image D, we consider the set C of all its wavelet packets energy coefficients c_j. The set Ω of elementary events is composed of "all identified attacks", for instance all attacks implemented in *Stirmark*. Our probability triplet is thus $(\Omega, \mathcal{B}, \mathcal{P})$, where $\mathcal{B}$ is the natural algebra associated with Ω and $\mathcal{P}$ is the uniform probability. For each index $i \in I$, the wavelet packet energy coefficient d_i is a random variable whose probability distribution F_i is determined by the result of the action of all possible attacks on the particular coefficient c_i.

We consider the following random function :

$$S(\lambda, \lambda') = \sum_{i \in I} \phi(c_i - \lambda) sgn(d_i - \lambda')$$

where $sgn(x) = \frac{x}{|x|}, sgn(0) = 0$, and ϕ is a C^∞ function that approximates the sgn function, such as the arctangent (this will simplify the optimization step). We seek to maximize $B(\lambda, \lambda') = \mathsf{E}(S(\lambda, \lambda'))$ w.r.t. $(\lambda, \lambda') \in \Lambda = [\underline{e}, \overline{e}]$, where E denotes expectation and the bounds set on (λ, λ'), are a weak version of the requirements described at the beginning of this subsection, namely that enough coefficients should be eligible, but packets with too low energy should not be included. Such restrictions also allow our problem

to be well posed, otherwise $(\lambda,\ \lambda')$ = $(0,0)$ and $(\lambda,\ \lambda')$ = (∞,∞) would be degenerate solutions independent of the input image. It is easy to check that:

$$B(\lambda,\lambda') = \sum_{i\in I} \phi(c_i-\lambda)(1-2F_i(\lambda')) \qquad (1)$$

If the F_i are C^2, so is $B(\lambda,\lambda')$[1], and its maximum may be found by classical gradient descent once the F_i are known. To learn the F_i-s, we adopted the following strategy: for each image in a learning set, we computed for each c_i the empirical distribution corresponding to all attacks. We checked that, for a given position i, the empirical distributions $F_i^{\mathcal{I}}$ were roughly similar for all images $\mathcal{I}$. We then approximated the common F_i using a shifted Gamma distribution with density of the form:

$$f_i(x) = \gamma(x-a-c_i)^{b-1}e^{-c(x-a-c_i)}H(x-a-c_i)$$

where $\gamma = \frac{c^b}{\Gamma(b)}$ and H is the Heaviside function. a,b,c are free parameters to be estimated from empirical distribution. The choice of the Gamma distribution and the form of the shift were made on the basis of the observed empirical distributions and on "semi-heuristic"considerations. In particular, if we assume that one can model the action of attacks as the addition of Gaussian white noise to the wavelet coefficients, then the energies should be distributed as a chi-square, which is a particular case of Gamma distribution. In addition, the fact Gamma distributions are stable by linear combinations is also useful for our purpose. Figure 5 shows examples of empirical distributions of the wavelet packets energy which have been submitted to the 89 attacks implemented in Stirmak. Figure 6 shows a fitted Gamma distribution using a maximum likelihood estimate of a,b,c. Once the parameters and thus the F_i are known, it is an easy task to compute (λ,λ') that maximize $B(\lambda,\lambda')$.

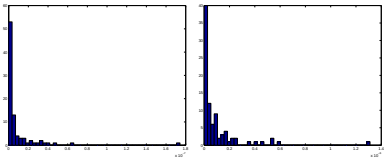

Figure 5. Empirical distributions of the attacked wavelet packet energies, for packet $C_{8,203,201}$ in the Lena (left) and Barbara (right) images.

While the above modeling works fine for attacks such as compression or small angle rotations, other ones based on cropping or low-pass filtering are not well controlled by this approach. These attacks can however be taken care of efficiently through adequate restrictions on the secret key. This will be presented elsewhere.

[1]While there are a finite number of attacks in Stirmark, we may still assume that F_i has a density by considering that the various parameters (e.g. the compression rate in JPEG) can be tuned in a continuous way.

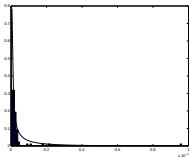

Figure 6. Empirical distributions of the attacked wavelet packet energies and the fitted version.

5 Numerical Experiments

We show an experiment on a boat image of size 256×256 pixels coded on 8 bits. Figure 7 shows the original signal along with a marked image containing a 32 bits watermark with redundancy equal to 71 . The PSNR between the two images is 49,27 dB. Figure 8 shows the watermark detector response to 1000 randomly generated watermarks after an attack consisting of JPEG compression with 30 % quality. This and other experiments made on several images with various attacks show that our method is an efficient and robust watermarking technique.

Figure 7. Original (left) and watermarked image (right).

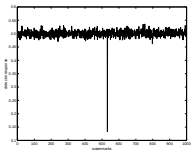

Figure 8. Watermark detector response to 1000 randomly generated watermarks after JPEG compression.

References

[1] N. Nikolaidis and I. Pitas. Copyright protection of images using robust digital signatures. *IEEE International Conference on Acoustic Speech and Signal Processing*, 4:2168–2171, May 1996.

[2] F.M. Bordland J.J.K.O. Ruanaidh, W.J. Dowling. Phase watermarking of digital images. *IEEE International Conference on Image Processing*, 2, September 1996.

[3] E. Koch and J. Zhao. Towards robust and hidden image copyright labeling. In *IEEE Workshop on Nonlinear Signal and Image Processing*, pages 452–455, Neos Marmaras, Greece, June 1995.

[4] R. Coifman and V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE trans. Info. Theorie*, 38:713–718, 1992.

[5] M. Kutter and F.A.P. Petitcolas. A fair benchmark for image watermarking systems. *Electronic Imaging: Security and Watermarking of multimedia Contents*, 3657:226–239, January 1999.