

# A Psychovisual Approach for Digital Picture Watermarking

J.-F. Delaigle, C. De Vleeschouwer\*, B. Macq

Laboratoire de Télécommunications et Télédétection

Université catholique de Louvain

Bâtiment Stévin - Place du Levant, 2

B-1348 Louvain-la-Neuve

Tel.: +32 10 47.80.72 - Fax: +32 10 47.20.89

E-mail: {delaigle,devlees,macq}@tele.ucl.ac.be

## ABSTRACT

In this paper, we wish to present a process enabling to mark digital pictures with invisible and undetectable secret information. This so-called **watermarking process** is intended to be the basis of a complete copyright protection system. It consists in constructing a band-limited image from binary sequences with good correlation properties and in modulating some randomly selected carriers. The security relies on the secrecy of these carrier frequencies, which are deduced from a unique secret key. Then the amplitude of the modulated images is modified according to a masking criterion based on a model of the Human Visual System. The adding of the modulated images to the original is supposed to be invisible. The resulting image fully identifies the Copyright Owner since he is the only one able to detect and prove the presence of the embedded watermark thanks to his secret key. This paper also contains an analysis of the robustness of the watermark against compression and image processing.

---

\*work supported by a grant F.N.R.S. - Alcatel Bell

# 1 General introduction

The emerging new digital image services and the rapid development of multimedia services rely not only on the increasing availability of digital networks and servers but also on efficient security tools. Among them, there is, nowadays, a strong need for suitable techniques to protect the work authors' and providers' interests<sup>1</sup> .

Indeed, the nature of digital media itself threatens its own viability:

- First, the digital replication of digital works is not only easy but also perfect.
- The ease of transmission and multiple uses is worrying too. Once a single illegal copy has been made, it is instantaneously widely accessible, without any control of the owner of the original picture.
- Finally, the plasticity of digital media is a great menace. Any malevolent user, *or pirate*, can easily modify a digital image at will, which may put many copyright protection methods at risk.

According to these considerations, the design of copyright protection systems is crucial and constitutes a great challenge, because it should be able to cope with these kinds of threats. Without such a system, most authors will not allow their digital works to be widely distributed. This is a major issue since international organizations, like the Mafia, organize financial flows on the use of counterfeit material.<sup>2</sup> Besides, digitization of pictures communication has provoked a parallel growth of technological tools for the creation and broadcasting of fakes. This is why the struggle against counterfeiting begs for ongoing development. For "classical material products", several widely used and very effective technical solutions already exist. Watermarked currency<sup>3</sup> and passports<sup>4</sup> are some examples of applications of these techniques. For other emerging domains, effective protecting methods have still to be conceived. This is the case of all kinds of digital image products. Watermarking is one of the most promising image protection techniques.

The principle of watermarking consists in the robust embedding of copyright information (e.g. time and date, Copyright Identifiers) in the signal to be protected. This signal may be a text<sup>5, 6</sup>, audio contents,<sup>7</sup> but most often watermarking is applied to still or moving images.

Many laboratories and companies have already developed their own watermarking techniques for digital images. Most of them directly work on the luminance with or without any considerations on the quality of the watermarked image.<sup>8-10</sup> Many authors model the image contents as a channel that can convey a certain quantity of information. Some techniques use such an approach on the basis of the spread-spectrum theory to embed a copyright code.<sup>11</sup> Other authors apply the watermarking not to the picture itself but to some of its characteristics, like DCT coefficients,<sup>12</sup> fractal coefficients or motion estimation vectors or the phase of DFT coefficients.<sup>13</sup>

These techniques give different results with diverse degrees of quality, but all have in common that they have to realize a good trade-off between the robustness of the watermark, the quality of the watermarked picture and the computational cost. However, only some of them really use a Human Visual System (HVS) model to analyze the quality of the resulting image. Swanson and al. use a model of the HVS,<sup>14</sup> but apply it to blocks. A method using a model of HVS but which applies on the whole picture is presented in this paper.

This paper introduces an additive watermarking technique for gray scale pictures. It consists in producing a synthetic picture, called *the watermark*, which contains information about the ownership of the original and depends on the picture contents. This watermark is added to the original in such a way that the resulting picture is perceptually identical to the original and the watermark is not detectable.

The most original part of this method is the embedding process i.e. the weighting of each pixel of the watermark before adding it to the original picture. This is based on the masking concept coming from a model of human vision (the perceptual model). From this concept, a method was deduced that reveals itself as genuinely efficient. Another key feature of the proposed method is the availability of two methods for the detection of watermarked pictures without the use of the original. This last point is fundamental for the copyright protection management. This paper concludes with the analysis of the embedding process, the efficiency of the retrieval and the system robustness.

## 2 Masking

### 2.1 Introduction

The aim of a watermarking technique is to provide an invisible embedding of secret information, called the watermark. This watermark must be masked (hidden) by the original picture contents. Masking criteria deduced from physiological and psychophysical studies<sup>15</sup> are widely known. Nevertheless, these criteria were formulated to evaluate picture quality and have to be fashioned efficiently for their use in image watermarking systems.

### 2.2 The Perceptual Model: Approximation of the Working of the Eye

It is common knowledge that the retina splits pictorial information into several components. These components circulate from the eye to the cortex through different tuned channels, each channel being tuned to each component.

The characteristics of one component are:

- the location in the visual field (in the image).
- the spatial frequency (in the Fourier domain: the amplitude in polar coordinates).
- the orientation (in the Fourier domain: the phase in polar coordinates)

In the tuned channels model, one perceptive channel can only be stirred by a stimulus in the corresponding tuned channel. Components that have different characteristics are independent.

### 2.3 The Masking Concept

According to the human vision perceptive model<sup>16</sup> used, signals that have similar (near) components take the same channels from the eye to the cortex. It appears that such signals interact and are submitted to non-linear

effects. Masking is one of those effects.

**Definition:** *the detection threshold* is the minimum level below which a signal can not be seen.

**Definition:** *masking* occurs when the detection threshold is increased because of the presence of another signal.

In other words, there is masking when a signal can not be seen because of the presence of another signal with similar characteristics and at a higher level.

## 2.4 The Masking Model

With the aim of shaping the masking phenomenon, tests have been done on monochromatic signals, called *gratings*. It appears that the eye is sensitive to the contrast of those gratings. This contrast is defined as:

$$C = \frac{2(L_{max} - L_{min})}{L_{max} + L_{min}} \quad (1)$$

where L stands for the luminance.

It is possible to experimentally determine the contrast detection threshold  $C_s$  of a grating signal with regards to the contrast  $C_m$  of the masking grating signal and to the frequency (figure 1.a). Identical curves can also be drawn with regards to orientation.

For simplicity's sake, these experimental curves are modeled as in figure 1.b). Such bi-logarithmic curves are traced for signals of one single frequency and one orientation  $(f_0, \theta_0)$ .

The expression of the detection threshold is as follows:

$$C_s(C_m) = \max[C_0, C_0 \left(\frac{C_m}{C_0}\right)^\epsilon] \quad (2)$$

where  $\epsilon$  (the slope) depends on  $(f_0, \theta_0)$ . Typically,  $0.6 \leq \epsilon \leq 1.1$ .

It is possible to extend that expression to frequency dependence. The general expression of the detection threshold

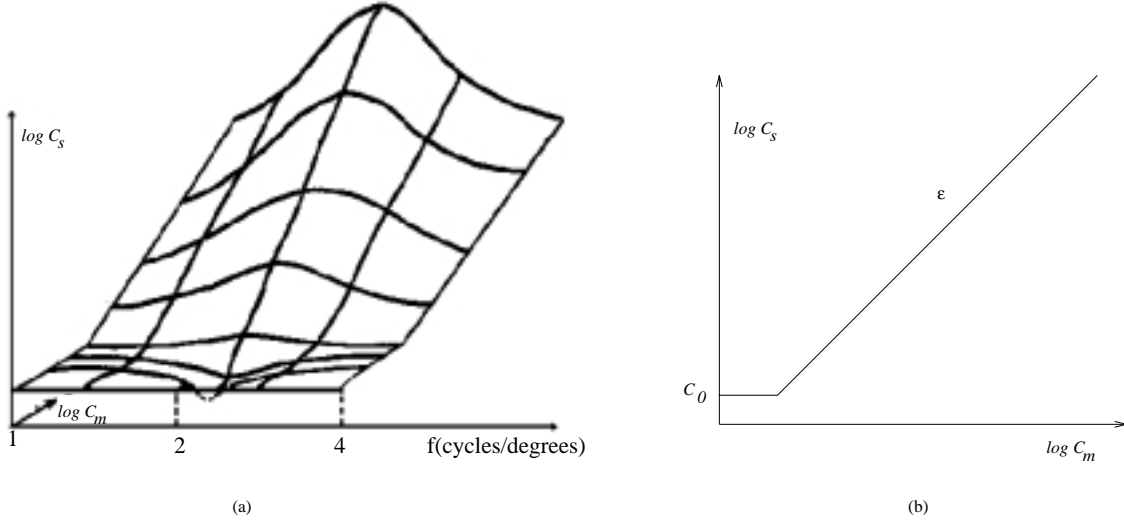


Figure 1: Curves of the masking threshold for a particular orientation: (a) Experimental curve: detection threshold with regard to the masking signal contrast and frequency, (b) Modeled curve for a particular frequency.

becomes:

$$C_s(C_m, f, \theta) = C_0 + k_{(f_0, \theta_0)}(f, \theta)[C_s(C_m) - C_0] \quad (3)$$

where

$$k_{(f_0, \theta_0)}(f, \theta) = \exp \left[ - \left( \frac{\log^2(\frac{f}{f_0})}{F^2(f_0)} + \frac{(\theta - \theta_0)^2}{\Theta^2(f_0)} \right) \right] = G_{(f_0, \theta_0)}(f, \theta). \quad (4)$$

In that expression,  $f_0$  and  $\theta_0$  are relevant to the masking signal,  $f$  and  $\theta$  are relevant to the masked signal,  $F(f_0)$  and  $\Theta(f_0)$  are parameters that represent the spreading of the Gaussian function defined by  $k_{(f_0, \theta_0)}(f, \theta)$  and  $C_0$  is often negligible. The spread of the Gaussian function depends on the frequency  $f_0$ :

- For frequency  $f$ , typical bandwidth at half response is 2,5 octaves at 1 cycle/degree and 1,5 octaves at 16 cycles/degree (c/d) with a linear decrease between both frequencies.<sup>17</sup>
- For orientation  $\theta$ , half bandwidth at half response depends on  $f_0$  and takes typical values like 30 degrees at 1 c/d and 15 degrees at 16 c/d.<sup>18</sup>

These psychovisual results allow us to express  $F(f_0)$  and  $\Theta(f_0)$  in terms of normalized frequency (see appendix A).

According to this expression, the frequency and orientation dependence of the detection threshold has a Gaussian form. Only near frequency signals can interact. When the frequency or the orientation of the masking signal (the mask) are far from the ones of the signal to mask, the detection threshold is almost equal to  $C_0$ .

## 2.5 The Masking Criterion

It is important to notice that those results only concern grating signals. To deduce a masking criterion that can apply to any signal like in real images, the masking condition has to be adapted. It is necessary to generalize the contrast concept in the case of real images. This new concept is called *local energy*.<sup>15</sup>

Local energy is defined on narrow band signals. **The local energy of a narrow band signal corresponds to the square of the amplitude of the signal envelope.**<sup>19</sup> So, in the framework of the analytic representation of signals, signals are described only by positive frequencies and the local energy in a particular signal sample is the square modulus of the local complex value of the signal at this sample position.

A picture is a broadband signal. The contrast related to a particular frequency and orientation results from the analytic filtering of the original picture by the Gabor analytic filter  $G_{(f_0, \theta_0)}(f, \theta)$  (see equation 4). It corresponds<sup>19</sup> to the square of the envelop of the narrow band real signal that would have been obtained through the filtering by a filter having components in both negative and positive frequencies, which is symmetric with regards to the origin of Fourier space (i.e. the zero frequency) and whose components in the positive frequencies are given by  $G_{(f_0, \theta_0)}(f, \theta)$  in equation 4. So, the filter characteristics are tuned to the studied perceptual component. Actually, this filter has been chosen to model human perception. Its effect is close to the filtering effect of the visual cortex cells.

The contrast value in expression (3) is given by this local energy. Since  $C_0$  is negligible, if we assume that the Gabor filters play the same role as  $k_{(f_0, \theta_0)}$  plays in (3), the masking criterion is easily derived.

**The masking criterion:** If the local energy of an additive image is less than the local energy of the masking image, around all the frequencies  $(f_0, \theta_0)$  and for each pixel  $(x, y)$ , then one can say that this additive image is

masked.

Strictly, an additive image (AI) is masked by a masking image (MI) if  $\forall(x, y)$  and  $\forall(f_0, \theta_0)$ ,

$$E_{\text{MI},(f_0, \theta_0)}(x, y) \geq E_{\text{AI},(f_0, \theta_0)}(x, y) \quad (5)$$

For real images, a good approximation of this criterion can be obtained by using a bank of filters whose central frequencies correspond to independent components and which are spread over the entire Fourier space. It is admitted<sup>17,20</sup> that 4 or 5 frequencies and 4 to 9 orientations are sufficient. The standard choice is twenty filters (5 frequencies and 4 orientations).

In our application, the masking image is the original picture and the additive image is the watermark. The latter is narrow band (see section 3.3). The criterion is reduced to the verification of inequality 5 only around the central frequency  $(f_0, \theta_0)$  of the narrow band signal.

## 2.6 Conclusion

This section has led to the expression of an easy to implement masking criterion applicable to any image. This criterion is a simple extension of a theoretical criterion applicable to monochromatic signals. Cases where that criterion does not match are thus possible. Unfortunately, a direct use of this criterion is not possible. The criterion only serves to an a posteriori evaluation of the invisibility of this addition but can not be used to generate an a priori watermark that will be invisible in a given picture. The next sections offer a solution to this problem.

## 3 Secret embedding

### 3.1 Position of the Algorithm in a Copyright Protection Scheme

The main functionality of the watermarking algorithm is to provide robust and reliable authentication of the Copyright Owner (CO). The copyright information carried by the watermark is generated by a secret key owned by the CO, so that it is impossible to detect the watermark without this secret key. The CO, or someone he has fined, can thus use this watermarking scheme to analyze a picture and determine whether it belongs to him or not. Afterwards he can also provide the proof of his ownership in case of conflict.

### 3.2 Embedded Information

The embedded information is a basic picture constructed from *Maximal Length Sequences* (MLS). These binary sequences have good correlation properties as explained in section 6. '0' and '1' symbols are mapped to  $n \times n$  black and white pixel blocks in the picture, as depicted in Figure 2.

### 3.3 Basic Information Modulation: Stamp Definition

In order to benefit from the advantages of human perception and to increase decision reliability, the basic information is modulated redundantly at several frequencies and orientations corresponding to quite independent components. Moreover, care must be taken to filter the initial basic picture with a low pass filter (LPF) (e.g. a Butterworth LPF) so that the resulting signal  $G(x, y)$  is narrow-band. This point is very important because it allows to limit the verification of the masking criterion in the activated perceptive channel.

The positions of the carriers (*inscription keys*) are *secret*. They are deduced from the CO's secret key. In practice, the frequency plan is divided into sectors. Each sector is relevant to one perceptive component and defines a group of pairs  $(f, \theta)$  where the basic picture can be modulated. Only one couple is chosen for each



Figure 2: Example of the basic information used

sector (because couples of a same sector do not stimulate independent components). The picture obtained from the sum of each *modulated grid*  $S_j(x, y)$  is called *the stamp*  $S(x, y)$ .

$$S(x, y) = \sum_{j \in K} S_j(x, y) = \sum_{j \in K} G(x, y) \cdot \cos(f_{x_j} \cdot x + f_{y_j} \cdot y) \quad (6)$$

$K$  represents the set of sectors amongst all possible sectors and  $(f_{x_j}, f_{y_j})$  corresponds to the carrier position chosen in a given sector  $j$  (the choice of the pair is defined by the CO's secret key). It is important to notice that each signal  $S_j(x, y)$  is spread over the whole picture. So, while it fits a given oriented frequency  $(f, \theta)$ , no spatial location of the masking stimuli is taken into account.

## 4 Spatial Weighting of the Stamp

The spread of the stamp leads to a robust watermark but it needs to be spatially weighted to better fit perceptual channels.

### 4.1 Introduction

The purpose of embedding (or inscription) is to adapt the level of each part  $S_j(x, y)$  of the stamp to generate a signal  $W_j(x, y)$  that is invisible once it is added to the picture. The *watermark*  $W(x, y)$ , i.e. the signal added

to the picture, is the sum of all the  $W_j(x, y)$ . So,

$$W(x, y) = \sum_{j \in K} W_j(x, y) \quad (7)$$

In section 4.2, a method is proposed to build a signal  $W_j(x, y)$ , close to  $S_j(x, y)$  but for which the masking criterion is satisfied. Nevertheless, some visibility problems subsist. They are solved in section 4.3.

## 4.2 $S_j$ Embedding Level Adjustment: $W_j$ Building

As mentioned above, each part  $S_j(x, y)$  of the stamp is narrow band. Inscriptions at different frequencies are thus independent and one can deal with the different components of the stamp one at a time. For each frequency  $(f_{x_j}, f_{y_j})$ , the embedding level adjustment is performed through an iterative approach.

### 4.2.1 Iterative Approach

According to the masking criterion and in order to guarantee invisibility, the local energy of the watermark component  $W_j(x, y)$  has to be inferior to the local energy of the picture, this for each pixel and around the inscription frequency. The amplitude of  $S_j(x, y)$  has to be adjusted according to the energy of the original picture in the corresponding component. Nevertheless, to avoid interactions between different parts of the watermark, each  $W_j(x, y)$  has to be kept narrow band. As a conclusion,  $\forall j$ , we have to find a signal  $W_j(x, y)$  so that

- $\forall(x, y) E_{W_j}(x, y) \leq E_{I(f_{x_j}, f_{y_j})}(x, y)$ ,
- $W_j$  is close to a scaled version of the modulated grid  $S_j (= G(x, y) \cdot \cos(f_{x_j} \cdot x + f_{y_j} \cdot y))$ ,
- $W_j$  has a narrow bandwidth

There is an easy way to generate the watermark,  $W_j(x, y)$ , so that it satisfies these three constraints. It consists in multiplying the stamp,  $S_j(x, y)$ , by a  $\alpha_j$  factor. This value is determined by the smallest ratio between the square root of the picture energy and the square root of the energy of  $S_j(x, y)$ . The obvious advantage of this

approach is that it does not deform the stamp to produce the watermark. At the detection stage (see section 5), a correlation is computed between the grid  $G(x, y)$  and a noisy demodulated version of the embedded watermark  $W_j$ . The useful part of this correlation results from the correlation of the grid with the demodulated stamp, i.e. in our case from  $R = \sum_{j \in K} \alpha_j G^2(x, y)$ . The major drawback of this global approach is that  $\alpha_j$  will be close to zero most of the time, because some of the regions of the picture contain very little energy. So,  $R$  will be close to zero.

In order to increase the embedding level of the stamp, this global adjustment is modified to support a local adaptation. If  $\alpha_j$  is greater than the smallest ratio between the square root of the picture energy and the square root of the stamp energy, the energy of some pixels of the watermark will exceed the energy of the original picture. The energy of the watermark in these positions has to be reduced. A way to induce this reduction is to reduce the corresponding watermark pixel values to zero, replacing  $\alpha_j \cdot G(x, y)$  by  $\alpha_j \cdot M_j(x, y) \cdot G(x, y)$ .  $M_j(x, y)$  is a mask equal to 0 at the positions where the watermark is not masked by the picture content. The energy of the watermark will of course decrease for these pixels and their surrounding area. This local approach has the advantage of initially decreasing the energy of the pixels where the energy of the stamp is too high relatively to the picture energy. Its major drawback is that it deforms the stamp to produce the watermark, reducing its correlation with the grid  $G(x, y)$  and increasing the bandwidth of the resulting signal. Concerning the increasing of the bandwidth, a solution is to perform a low pass filtering of  $\alpha_j \cdot M_j(x, y) \cdot G(x, y)$ , resulting in a signal noted  $[\alpha_j \cdot M_j(x, y) \cdot G(x, y)]_{LP}$  in the following.

The optimal choice of  $\alpha_j$  has to be discussed. We propose to choose it in order to maximize the correlation between the large bandwidth signal  $\alpha_j \cdot M_j(x, y) \cdot G(x, y)$  and the grid  $G(x, y)$ . This should provide a high correlation value between  $[\alpha_j \cdot M_j(x, y) \cdot G(x, y)]_{LP}$ , which is actually embedded, and  $G(x, y)$ , which is the useful part of the correlation value  $C$  computed at the detection phase (see section 5).

This optimum can be efficiently computed. First, the square root values of the ratio between the picture energy and the stamp energy are sorted in decreasing order. Each of these values is a possible  $\alpha_j$  value, noted  $Poss_{\alpha_j}$ . By scanning them in decreasing order, one can progressively add non zero values to the mask  $M_j(x, y)$

and take the corresponding points of the grid  $G(x, y)$  into account for the computation of the number  $R(Poss_{\alpha_j}) = Poss_{\alpha_j} \cdot \sum_{(x,y)} M_j(x, y) \cdot G(x, y)$ .  $\alpha_j$  is chosen to maximize  $R(Poss_{\alpha_j})$ .

After this procedure, the energy of the stamp is different from the energy used to compute  $M_j(x, y)$  during the process. Moreover, the low pass filtering of  $\alpha_j \cdot M_j(x, y) \cdot G(x, y)$  may result in some non zero value in areas where the masking criterion had forecast that masking would not work. Some iterations of the process, using  $[\alpha_j \cdot M_j(x, y) \cdot G(x, y)]_{LP}$  in place of  $G(x, y)$ , will allow to solve this problem. Indeed, tests reveal that after a second iteration, the stamp does not change enough to continue the iterations. So, two iterations are sufficient.

#### 4.2.2 First Results Quality Analysis

It appears that the criterion for determining  $\alpha_j$  often leads to impose an optimum with several points equal to zero and a small number of points of high value. The addition of the thus obtained watermark generally entails a degradation of picture quality (figure 3). This emphasizes the weakness of the masking criterion used.

Figure 3 represents the watermark energy and the corresponding picture with regards to pixel location along a horizontal line, in a zone where the watermark is visible after being embedded. Figure 3.a shows that in this problematic zone the watermark energy is higher than the picture energy. The visibility of the watermark could have been correctly predicted by the criterion. Such a situation should not have happened if the inscription process were ideal. Indeed, this artifact is well identified. It is due to the filtering enforced by the constraint of narrow bandwidth of each watermark component. The solution to this problem is to severely limit the embedding level in the surroundings of low energy areas.

Besides, the examination of the second graph of Figure 3, corresponding to another problem zone, reveals a watermark energy inferior to picture energy. The masking criterion thus guarantees the invisibility of the watermark. However, the watermark is not completely masked; it is visible and it decreases the watermarked picture quality.

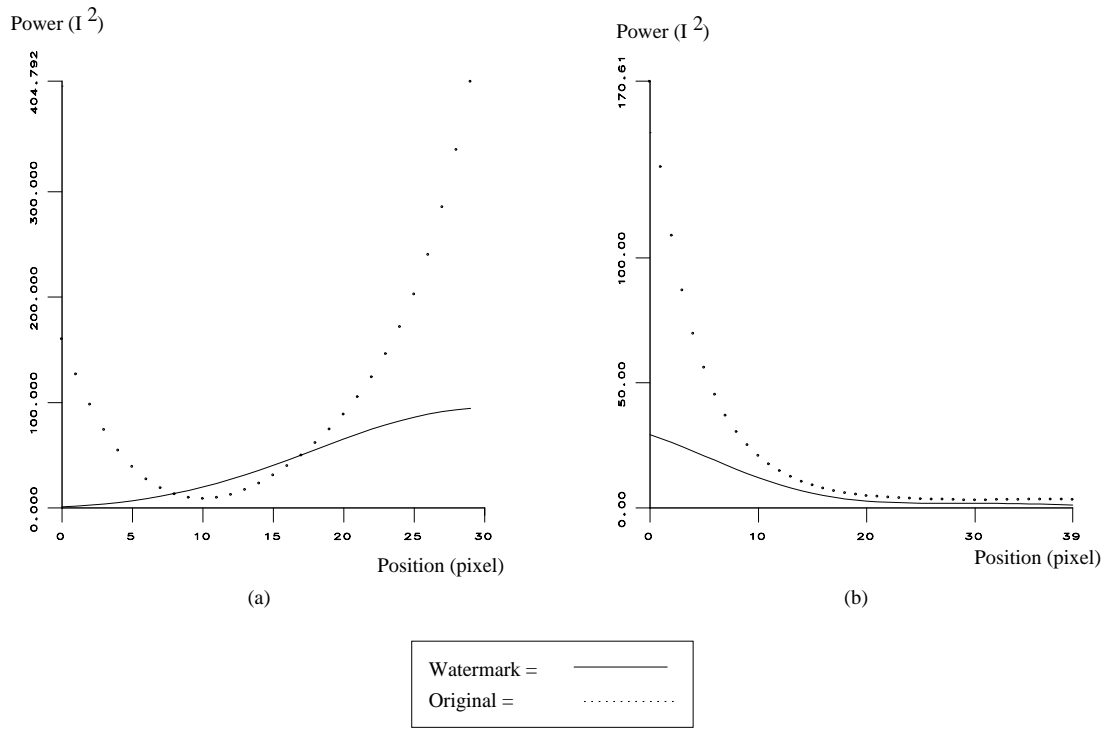


Figure 3: Imperfections of the first implementation. The picture is the watermarked picture of Lena. The graphs show curves of watermark and original picture energy for a particular frequency and orientation with regards to the pixel location along a horizontal line. Graphs (a) and (b) correspond to different areas of the picture where the watermark is visible.

### 4.3 Improvements: Towards a Truly Masked Inscription

The conclusion of the above considerations should be that the used criterion is unsatisfactory. Some improvements have been made after experimental observation. The result of these observations was that the invisibility is only strictly observed in high activity regions, where the high frequencies local energy is important. These regions have to be favored during inscription in the sense that the watermark level will be increased in those regions while it has to be decreased in other regions.

The correction process first isolates high activity regions (figure 4.a). Then, homogenization of the picture is performed by non linear filtering. To this end, morphological tools<sup>21</sup> have been used, i.e. one opening followed by one closing (figure 4.b). This is followed by scaling, which is a division by the mean or mean square value of the homogenized mask, followed by upper thresholding at 1. One obtains a new mask used to multiply the local energy of the picture at all the marking frequencies. This mask gives an advantage to regions of high-frequency energy in comparison with other areas. After that correction, the process is almost identical to the one described previously, except for one point: the choice of  $\alpha_j$  is now limited in order to maintain the number of watermark pixels higher than the number of pixels in which the picture has a high frequency activity. This corresponds to the black region in figure 4.b. The resulting complexity is not increased. Indeed, we first work on the inscription at high frequencies (where there is no quality problem). The value of the high frequency local energy is then used for calculating the correcting mask used for inscription at lower frequencies. The correction scheme is represented in figure 5e.

Another examination of the energy in the problem zones confirms the interest of the correction process (figure 5(a) to (d)). The watermark energy has become inferior to the picture energy (figure 5(a) and (b)). Moreover, figure 5(b) represents a region of the picture where the inscription level was too weak. After correction, the watermark energy follows the picture energy more, the inscription level is increased (figure 5(c) to (d)). These results could have been foreseen due to the high energy level of the picture.



(a)



(b)

Figure 4: Correcting mask for Lena: (a) High frequency areas, (b) Morphological homogenization of the mask (Structurant element is a block whose size is related to picture size by a typical ratio of 32.)

## 5 Detection

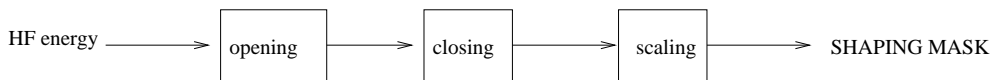
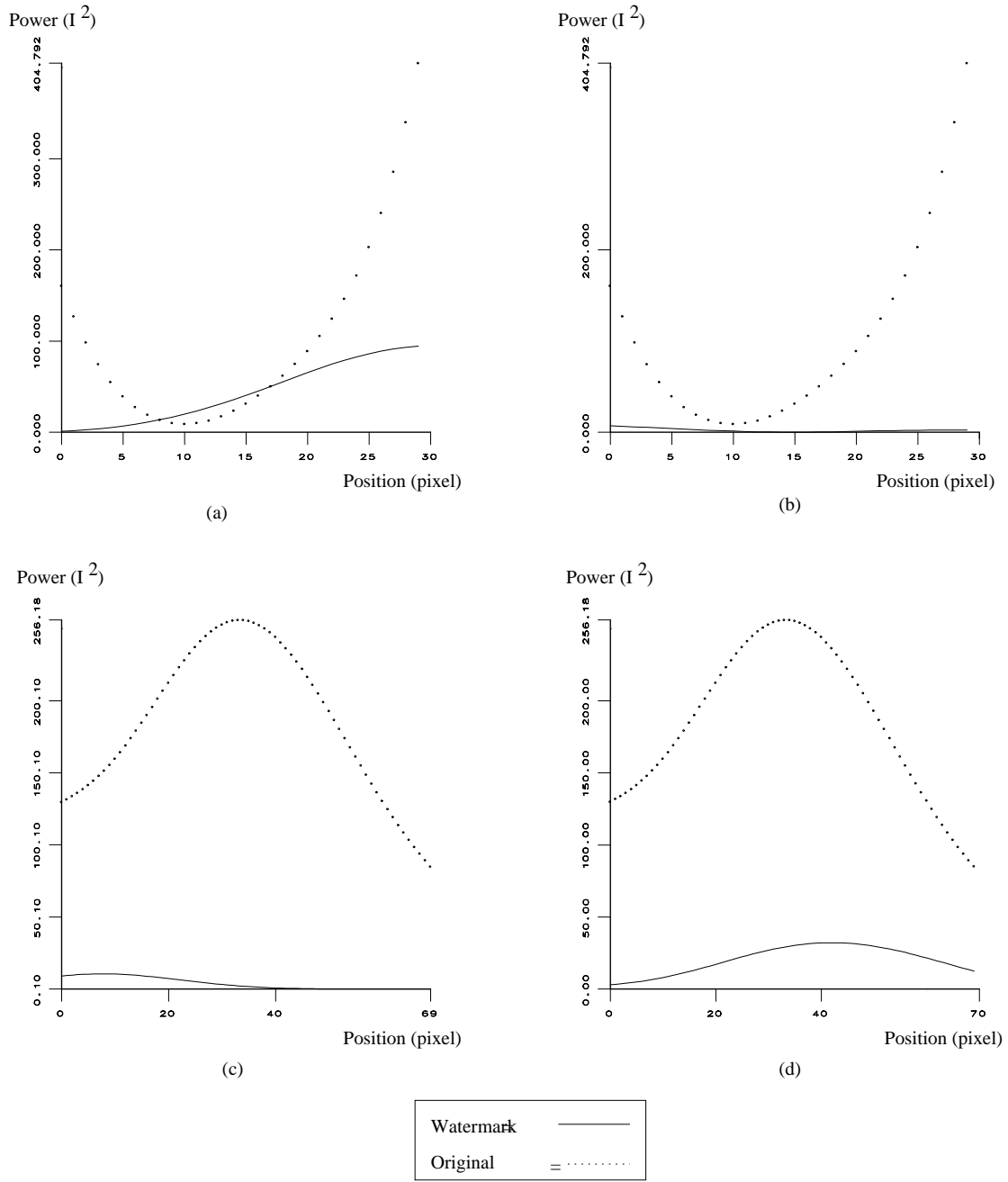
The detection process aims at determining whether a watermark is embedded in a given picture. This can be done with the use of a correlation process. For this purpose, it is necessary to isolate the watermark and then demodulate it in order to reconstruct a signal that is highly correlated with basic information (grid  $G(x, y)$ ).

The formulation of the watermark is:

$$W(x, y) = \sum_j W_j = \sum_{j \in K} [\alpha_j \cdot M_j(x, y) \cdot G(x, y)]_{LP} \cdot \cos(f_{x_j} \cdot x + f_{y_j} \cdot y) \quad (8)$$

In this expression, the level is adjusted in order to make it invisible (see section 4.1) when it modulates the carrier corresponding to a particular perceptive component.

The detection is composed of three steps : demodulation, correlation and decision.



(e)

Figure 5: Improvements provided by the shaping mask. Curves (a) to (d) represent the watermark and original picture energy for a particular frequency and orientation with regard to the pixel location along a horizontal line: (a) problem area, (b) same area when correcting mask is used, (c) area where inscription could be much higher, (d) invisible enforcement of the inscription, (e) way to construct the corrector mask.

- **Demodulation**

Let  $I_W$  be a received picture. For a watermarked picture, it has the form

$$I_W(x, y) = \sum_{j \in K} [\alpha_j \cdot M_j(x, y) \cdot G(x, y)]_{LP} \cdot \cos(f_{x_j} \cdot x + f_{y_j} \cdot y) + I_O(x, y) + N(x, y) \quad (9)$$

where  $I_O(x, y)$  is the original picture and  $N(x, y)$  is an additive channel noise.

The demodulation is achieved independently for each part of the watermark, i.e.  $\forall j \in K$ , the demodulation consists in multiplying  $I_W$  by  $\cos(f_{x_j} \cdot x + f_{y_j} \cdot y)$  and then to filter with a low-pass (LP) filter.

For the  $j_{th}$  part of the watermark, the result is

$$D_j(x, y) = \frac{1}{2} \cdot [\alpha_j \cdot M_j(x, y) \cdot G(x, y)]_{LP} + N_j^*(x, y) \quad (10)$$

$N_j^*(x, y)$  depends on the image and on the additive channel noise around the  $j_{th}$  demodulation frequency.

The other parts of the demodulated signal are of course eliminated by the LP filter.

- **Correlation**

It consists of summing point to point the product of the basic grid  $G(x, y)$  with the sum of all demodulated part of the watermark i.e.  $D(x, y) = \sum_{j \in K} D_j(x, y)$ . This correlation  $C$  is

$$C = \sum_{x, y} \sum_{j \in K} D_j(x, y) \cdot G(x, y) \quad (11)$$

$$= \sum_{x, y} \sum_{j \in K} [G(x, y) \cdot [\alpha_j \cdot M_j(x, y) \cdot G(x, y)]_{LP} + G(x, y) \cdot N_j^*(x, y)] \quad (12)$$

$$\simeq \sum_{x, y} \sum_{j \in K} [G(x, y)^2 \cdot [\alpha_j \cdot M_j(x, y)]_{LP} + G(x, y) \cdot N_j^*(x, y)] \quad (13)$$

$$\simeq \sum_{x, y} \sum_{j \in K} [G(x, y)^2 \cdot [\alpha_j \cdot M_j(x, y)]_{LP} + \sum_{x, y} [G(x, y) \cdot N_j^*(x, y)]] \quad (14)$$

In (14), the first term is even greater than the second, because  $G(x, y)$  and  $N_j^*(x, y) = \sum_{j \in K} N_j^*(x, y)$  are not correlated.

So,  $C$  mainly depends on the watermark value.

- **Decision**

The detection algorithm performs demodulations and correlations at diverse frequencies and with diverse grids. The decision is made after comparison of these correlations. Two methods have been used to determine whether the received picture contains a watermark or not (see section 6).

In the first method, a comparison is performed between  $C$  (see (14)), the correlation resulting from a reception performed with the grid ( $G(x, y)$ ) and the carriers frequencies deduced from the CO's secret keys, i.e. inscription and reception parameters are the same, and  $C^*$  (see (15)), the correlation made with random parameters (grid and frequencies).  $C^*$  is expressed by

$$C^* = \sum_{x,y} G_{random}(x, y) \sum_{j \in K_{rdm}} N_j^{**}(x, y) \quad (15)$$

where  $N_j^{**}(x, y)$  comes from the random filtered parts of the watermarked picture,  $G_{random}(x, y)$  is a random grid and  $K_{rdm}$  is a set of randomly chosen carriers. This correlation has zero mean value. It is supposed to be lower than the one obtained with the correct parameters. This assumption is at the basis of the first decision process.

In the second method, the frequencies used for demodulation are the right ones, but the compared correlations are obtained with grids having decreasing correlations with the embedded grid. If  $G^*(x, y)$  denotes the grid used to perform the correlation, the result obtained has the form

$$C^{**} = \sum_{x,y} \sum_{j \in K} G(x, y) \cdot G^*(x, y) \cdot [\alpha_j \cdot M_j(x, y) \cdot G(x, y)]_{LP} + \sum_{x,y} [G^*(x, y) \cdot N^*(x, y)] \quad (16)$$

$$\simeq \sum_{x,y} \sum_{j \in K} G(x, y) \cdot G^*(x, y) \cdot [\alpha_j \cdot M_j(x, y) \cdot G(x, y)]_{LP} \quad (17)$$

Since  $G^*(x, y)$  and  $N^*(x, y)$  are not correlated, the second term has zero mean value, whatever  $G^*(x, y)$  is. The value of  $C^{**}$  mainly depends on the first term, i.e. on the correlation factor between  $G(x, y)$  and  $G^*$ . This observation is used by the second decision process. It embeds a particular grid  $G(x, y)$  built from a Maximal Length Sequence (MLS) which is a sequence having good autocorrelation properties,<sup>22</sup> i.e. correlations of such a sequence with its shifted versions give low results. The importance of using such sequences has already been pointed out by some authors.<sup>23</sup> As a result of grid construction, shifted versions of the basic grid are also virtually decorrelated with the embedded grid. Due to these good correlation properties, the decision can be made from the observation of the maximal correlation  $C$  compared with all other correlations  $C^{**}$ .

| Image Name              | Optimal correlation | Random correlation 1 | Random correlation 2 | Random correlation3 | Random correlation 4 | Conclusion             |
|-------------------------|---------------------|----------------------|----------------------|---------------------|----------------------|------------------------|
| <b>Lena watermarked</b> | <b>584609</b>       | 92605                | 133920               | 80534               | 143633               | <i>watermarked</i>     |
| <b>Lena original</b>    | 94538               | 98099                | 135492               | 76739               | <b>137120</b>        | <i>Non watermarked</i> |

Figure 6: Results of correlation for Lena and decision.

## 6 Results

The first and probably most important result is the invisibility of the watermark in all images that were tested. Figures 7.a and b compare the original and watermarked picture for Lena. In figure 7.e, one can observe the watermark added to the original picture.

As explained in section 5, two methods are used to determine whether an image is watermarked or not. Figure 6 shows the results obtained by the first method. The main drawback of this method emerges when checking a non watermarked picture. In this case, a correct decision can not be guaranteed, because the correlation made with a given CO's key might be greater than the correlations made with random keys. The spontaneous decision would be to declare the image watermarked with this CO's key, while it is in fact not watermarked. As a result, when the reception declares an image as watermarked, this decision has to be confirmed by many other correlations, in order to obtain a sufficient degree of trustworthiness. This causes an undesirable increase of computational cost.

The second method does not encounter the same drawbacks, since it uses a particular grid  $G(x, y)$  formed from a Maximal-Length Sequence (MLS). Correlations are made with shifted versions of the basic grid. Due to the good correlation properties, the correlation made with the right grid gives a far greater result than the correlations with shifted grids. Results are presented in figure 7(c) and (d). If a picture is watermarked, a peak appears in the center and, even if another peak appears on the correlation graph, the centered one is the highest.

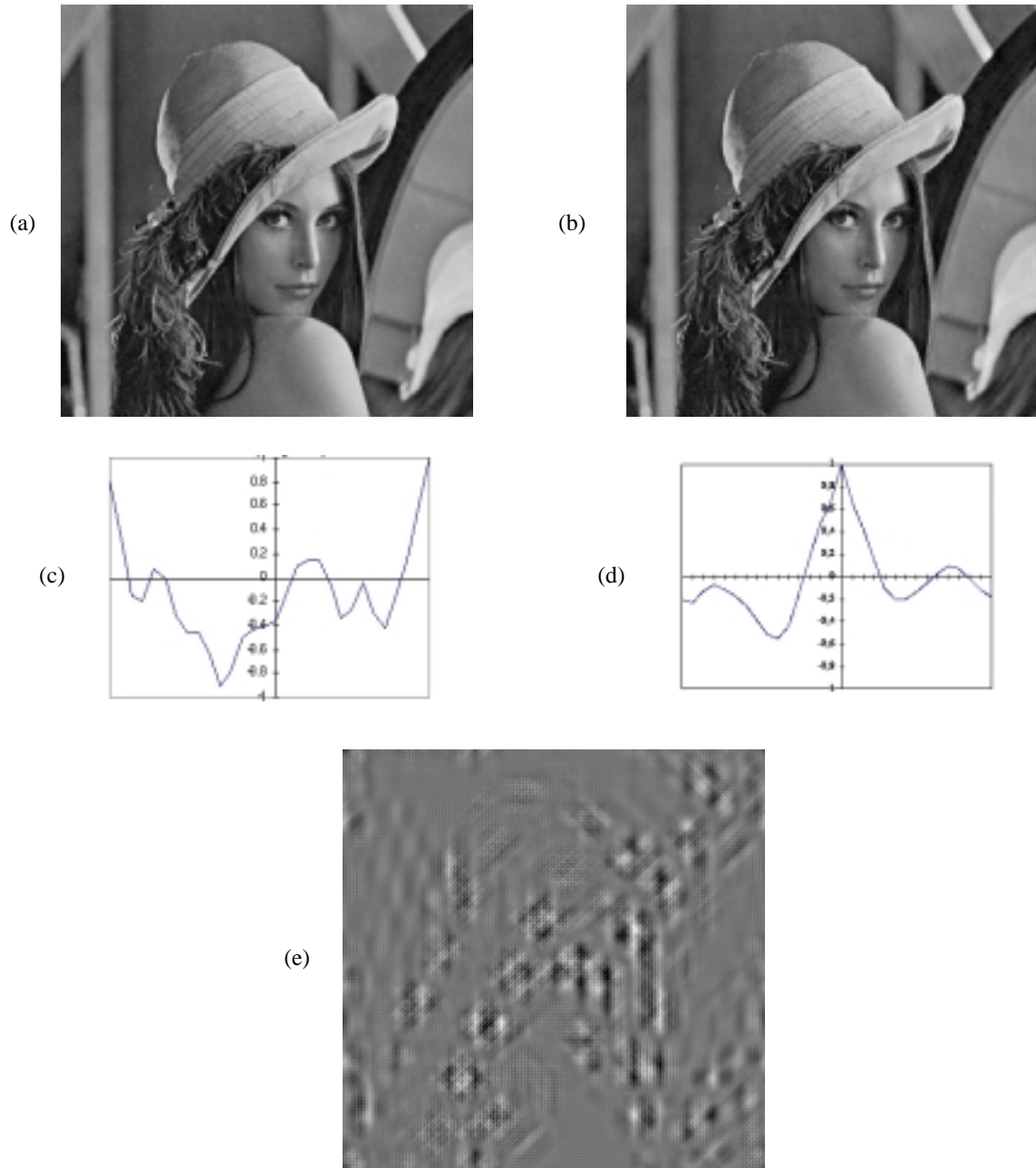


Figure 7: Results for Lena: (a) Original, (b) Watermarked, (c) Correlation graphic for original, (d) Correlation graphic for watermarked, (e) Watermark.

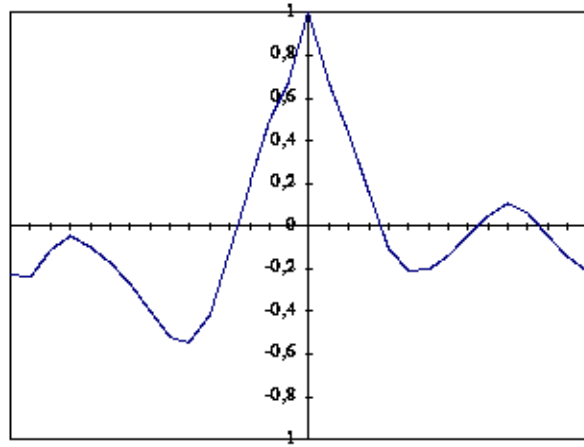


Figure 8: Picture deterioration with a white noise of 100dB variance and corresponding correlation.

## 7 System Robustness

Many tests have been done on usual pictures deterioration in image processing as well as classical pirate attacks.

- Noising

Resistance towards white noise is obvious. This is due to the correlative approach for reception and is illustrated on figure 8. The noise (whose variance is 100dB) has nearly no influence on reception.

- JPEG Coding-Decoding

The aim of a compression algorithm is to remove redundant information e.g., in the case of JPEG, the less significant frequencies with minor importance in the compressed bitstream. The watermark is strongly correlated with the picture. It has a higher level at frequencies with the picture itself at a higher level. So, after compression the most important part of the watermark is still present. Some results are presented in figure 9.a. They show the maximum correlation (always obtained for a null shifting) with regard to the JPEG %. This percentage corresponds to a quality requirement and is directly related to the ratio between

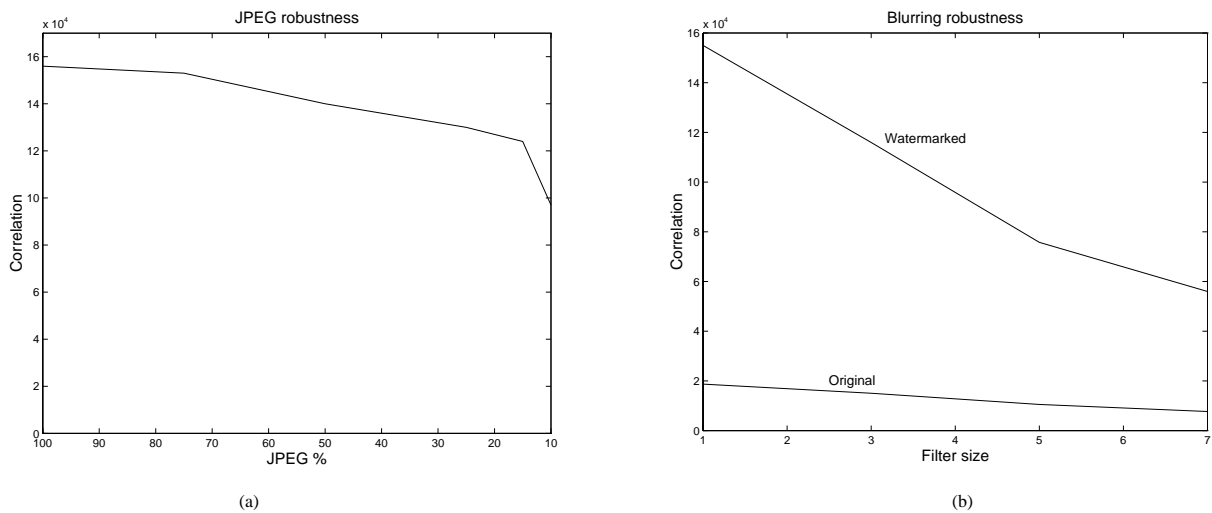


Figure 9: Results for Lena: (a) The correlation decreases with regards to JPEG level of compression, (b) Correlation evolution with regards to the size of the filter used for blurring.

the decoded picture and the original picture bitstream lengths. As this figure demonstrates, the tendency is of course decreasing, due to reduction of information entailed by compression. Nevertheless the decision remains correct even when the picture quality is no longer satisfactory. Figure 10.a shows the reception for Lena coded with JPEG at 15% (corresponding to a compression ratio equal to 12.5). One can still say that the picture was watermarked.

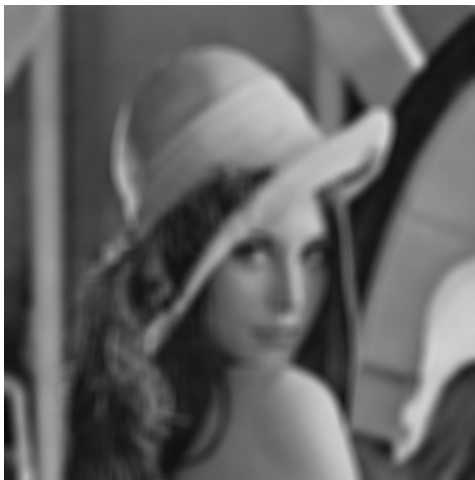
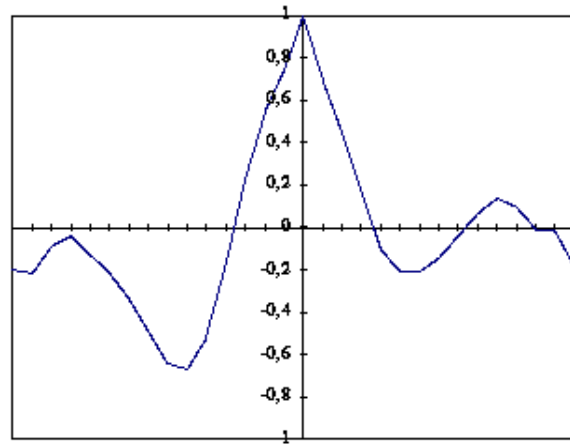
Besides, it is obvious that the information added by the watermark will decrease the compression ratio obtained for a particular quality percentage. However, experiments show that the increase of data is only about 0.6%.

- Scanning of Printed Pictures

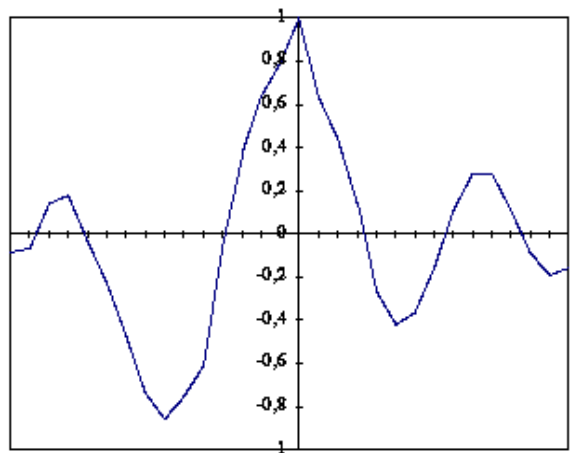
Figure 10.c shows the scanning of a printed watermarked picture of Lena. One notes that, in spite of the bad quality of this particular transmission channel, reception still remains effective. This is an important and encouraging result because it opens the door to a large range of applications: pictures broadcast by photographs or museums will still be protected, even if they have been printed, watermarked identity card pictures printed may be used for authentication, etc.



(a)



(b)



(c)

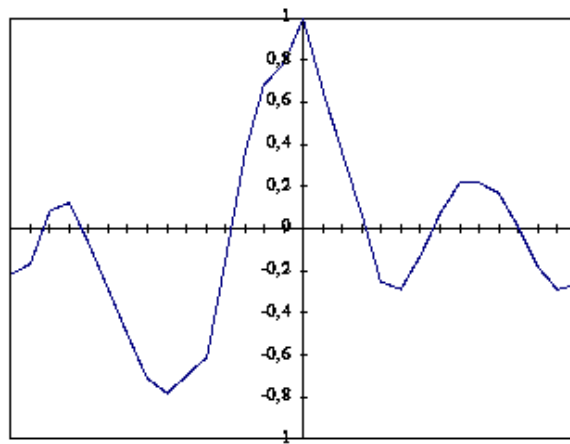


Figure 10: Different deteriorated pictures and corresponding correlations: (a) picture coded and decoded with JPEG 15%,(b) picture blurred with a 7-size filter,(c) Scanned picture.

- Blurring, Smoothing, Simple Frequency Low-Pass Filtering

These treatments result in low-pass filtering. As the watermark also contains low-pass components, it is still present after such filtering, but its level has decreased. Experiments have shown that it is necessary to significantly alter the picture quality so as to make the watermark inefficient. Figure 9.b shows the evolution of the correlation with regard to the size of the blur filter. The influence of the correlation decrease is mitigated by the decreasing trend of the original picture correlation. The decision remains correct up to a 7x7 sized filter. The picture on figure 10.b results from a 7x7 blurring. Despite its bad quality the decision after correlation still remains correct.

- Manipulations

The most common pirate manipulations are: cropping, scaling (also called zoom), rotation, axial symmetry, column and line suppression, analog conversion, format change, etc. After all these manipulations, the watermark is still present without exception, while the picture quality is not too altered. The remaining problem is the location of the watermark. Without extra improvements, the location of the watermark is impossible. These improvements are often obvious, easy to implement and are not time costing at all.

Nevertheless, some more subversive manipulations (i.e. cropping, zoom) are difficult to cope with. The present watermarking algorithm, by its frequential approach, is well conceived to resist such manipulations but in its current implementation, it cannot tackle them because it supposes the picture size has not been modified during transmission. Improvements, that have to be developed in order to solve these problems, are reviewed in the next section.

## 8 Future Research

### 8.1 Robustness

In the preceding section we highlighted some lacks of the algorithm. The current implementation is not sufficient to resist all pirate manipulations, but that is a characteristic of currently existing watermarking methods.

One encouraging point is that the solutions can easily be extended to an object-based approach. After an object-based segmentation of the picture, a watermark could be applied to isolated picture regions. Symmetrically, the watermark detection would be applied after the same segmentation on isolated regions. The difficulty is that this segmentation has to invariably isolate the same regions even after cropping or scaling. Satisfying tools of this kind do not exist. As a result, the robustness against cropping and scaling requires the use of the original in the detection phase.

## 8.2 Detection

An additional improvement would be to allow the system to retrieve the embedded information instead of making a correlation. This would allow the identification of the picture. Indeed, for now the embedded information only aims at identifying the picture ownership; the information is limited to one bit (watermarked or not). This is of course not sufficient for Copyright Owners who want to identify their pictures. The current method only authenticates ownership, but it could be envisaged to extend it in order to embed more than one bit. To this end, it would be possible to vary the embedded MLS, e.g. their phase. The phase could carry information for identification. Of course, the system could become less robust, since the redundancy of the embedding will decrease.

# 9 Conclusion

The watermarking process presented in this paper allows the authentication of the ownership of any picture. The embedding process is based on a perceptual approach. Large-scale tests have demonstrated the effectiveness of the proposed approach both as to the quality of the marking and its robustness. Nonetheless, the current state of the algorithm still needs improvements. The addition of a robust object-based segmentation would be an added value for the resistance against cropping and scaling. It is also possible to extend the current functionalities of the algorithm so as to support an identification of copyrighted pictures.

## 10 REFERENCES

- [1] B. Kahin. The Strategic Environment for Protecting Multimedia. *IMA Intellectual Property Proceedings*, 1:1–8, January 1994.
- [2] Ian M. Lancaster and Lewis T. Kontnik. Progress in Counterfeit Deterrence: The Contribution of Information Exchange. In *Conference 2659 - Optical Security and Counterfeit Deterrence Techniques*, San Jose, January 1996. SPIE Electronic Imaging: science and technology. Accepted for publication.
- [3] Sara E. Church and Thomas A. Ferguson. Evaluation of Security Features for New U.S. Currency. In *Conference 2659 - Optical Security and Counterfeit Deterrence Techniques*, San Jose, January 1996. SPIE Electronic Imaging: science and technology. Accepted for publication.
- [4] Dick van Lingen. The New Dutch Passport. In *Conference 2659 - Optical Security and Counterfeit Deterrence Techniques*, San Jose, January 1996. SPIE Electronic Imaging: science and technology. Accepted for publication.
- [5] K. Matsui and K. Tanaka. Video-Stenography: How to embed a Signature in a Picture. *IMA Intellectual Property Proceedings*, 1(1):187–205, January 1994.
- [6] J.T. Brassil, S. Low, N.F. Maxemchuk, and L. O’Gorman. Electronic Marking and Identification Techniques to Discourage Document Copying. *Proceedings of IEEE INFOCOM’94*, pages 1278–1287, June 1994.
- [7] W. Bender, D. Gruhl, and N. Moromoto. Techniques for Data Hiding. *Proceedings of the SPIE*, 2420(40), February 1995.
- [8] O. Bruyndonckx, J.J. Quisquater, and B. Macq. Spatial Method for Copyright Labelling of Digital Images. *Proceedings of IEEE Workshop on Non-Linear Processing*, pages 456–459, June 1995.
- [9] G. Caronni. Assuring Ownership Rights for Digital Images. *Proceeding of Reliable IT Systems, VIS 95*, June 1995.
- [10] R.B. Wolfgang and E.J. Delp. A Watermark for Digital Images. pages 219–242, Lausanne, Switzerland, September 1996. IEEE International Conference on Image Processing.

- [11] I.J. Cox, J. Kilian, T. Leighton, and T. Shamoan. Spread Spectrum Watermarking for Multimedia. *Proceedings of the SPIE*, 2420:456–459, February 1995.
- [12] E. Koch and J. Zhao. Towards Robust and Hidden Image Copyright Labeling. *Proceedings of IEEE Workshop on Non-Linear Processing*, pages 452–455, June 1995.
- [13] J. J. K. O’Ruanaidh, W. J. Dowling, and F. M. Boland. Phase watermarking of images. pages 239–242, Lausanne, Switzerland, September 1996. IEEE International Conference on Image Processing.
- [14] M.D. Swanson, Bin Zhu, and A.H. Tewfik. Transparent Robust Image Watermarking. pages 211–213, Lausanne, Switzerland, September 1996. IEEE International Conference on Image Processing.
- [15] S. Comes. *Les traitements perceptifs d’images numérisées*. PhD thesis, Université catholique de Louvain, June 1995.
- [16] John Wiley, L.A. Olzak, and J.P. Thomas. *Handbook of Perception and Human Performance. Volume 1: Sensory Processes and Perception. Chapter 7: Seeing Spatial Patterns*. University of California, Los Angeles, California, 1986.
- [17] H.R. Wilson, D.K. McFarlane, and G.C. Phillips. Spatial Frequency Tuning of Orientation Selective Units Estimated by Oblique Masking. *Vision Research*, 23(9):873–847, 1983.
- [18] H.R. Wilson and G.C. Phillips. Orientation Bandwidths of Spatial Mechanisms Measured by Masking. *J. Opt. Soc. Am. A*, 1(2):226–232, February 1984.
- [19] John G. Proakis, editor. *Digital Communications*, chapter 4, pages 152–232. McGRAW-HILL INTERNATIONAL EDITIONS, 1995. ISBN 0-07-113814-5.
- [20] J. Daugman. Uncertainty Relation For Resolution In Space, Spatial Frequency, And Orientation Optimized By Two-dimensional Visual Cortical Filters. *J. Opt. Soc. Am. A*, 2(7):1160–1169, July 1985.
- [21] M. Schmitt and J. Mattioli. *Morphologie Mathématique*. Masson, Paris, 1993.
- [22] D.V. Sarwate and Pursley M.B. Crosscorrelation Properties of Pseudorandom and Related Sequences. *Proceedings of the IEEE*, 68(5):593–617, May 1980.

- [23] C.F. Osborne R.G. van Schyndel, A.Z. Tirkel. A Digital Watermark. pages 86–90, Austin, Texas, November 1994. IEEE International Conference on Image Processing.

## A Definition of the analytic filters parameters

Making the common hypothesis that the observer is located at a distance equal to 6 times the picture height, *cycles/degree* frequencies can be converted into normalized frequencies according to the sampling frequency. Indeed,  $N$  being the number of columns in the picture (i.e. the width of the picture),

$$f_{normalized} \left( \frac{\text{cycles/screen}}{\text{Number of samples on the screen}} \right) = \frac{f(\text{cycles/degree}) \times 9.53(\text{degrees/screen})}{N(\text{Number of samples on the screen})} \quad (18)$$

As the sampling rate is the same in both horizontal and vertical directions, the normalization factor  $\frac{9.53}{N}$  is also valid to convert *cycles/degree* horizontal frequencies into normalized frequencies. So, according to the considerations made in section 2.4 about the bandwidth of the filter defined by equation (4),  $F(f_0)$  and  $\Theta(f_0)$  are defined in term of normalized frequencies  $f_0$  as:

$$F(f_0) = \frac{\frac{-1}{15} \cdot \left( \frac{f_0 \cdot N}{9.53} - 1 \right) + 2.5}{\sqrt{\ln(2)}}$$

$$\Theta(f_0) = \frac{31 - \frac{f_0 \cdot N}{9.53}}{\sqrt{\ln(2)}}$$