

## Cover Page

1) Title of the paper:

**A robust image watermarking technique based on quantization noise  
visibility thresholds**

2) authors' affiliation and address:

**IRCCyN-IVC, (UMR CNRS 6597), Polytech' Nantes  
Rue Christian Pauc, La Chantrerie, 44306 NANTES, France.  
Tel : 02.40.68.32.47  
Fax : 02.40.68.30.66**

3) e\_mail address:

**Florent.Autrusseau@polytech.univ-nantes.fr**

4) Journal & Publisher information:

**Elsevier Signal Processing  
<http://www.elsevier.com/locate/sigpro>**

5) bibtex entry:

```
@article{Autrusseau2007sigpro,  
  Author = {F. Autrusseau and P. Le Callet},  
  Journal = {Elsevier Signal Processing},  
  Number = {6},  
  Pages = {1363-1383},  
  Title = {A robust image watermarking technique based on quantization  
noise visibility thresholds},  
  Volume = {87},  
  Year = {2007}}
```

doi:10.1016/j.sigpro.2006.11.009

# Robust image watermarking technique based on quantization noise visibility thresholds

Florent Autrusseau, Patrick Le Callet

*Ecole polytechnique de l'Université de Nantes,  
Rue, Ch. Pauc, La Chantrerie,  
44306 Nantes Cedex 3, FRANCE*

---

## Abstract

A tremendous amount of digital multimedia data is broadcasted daily over internet. Digital data being very quickly and easily duplicable, intellectual property rights protection techniques, appeared about fifty years ago (see (14) for an extended review). Digital watermarking was born. Since its birth, many watermarking techniques appeared, in all possible transformed spaces. However, an important lack in watermarking literature concerns the human visual systems models. Several Human Visual System (HVS) model based watermarking techniques have been designed in the late 1990's. Due to the weak robustness results, especially concerning geometrical distortions, the interest in such studies have been reduced. In this paper, we intend to take benefit from the last advances in HVS models and watermarking techniques to revisit this issue. We hereby demonstrate that it is possible to resist to many attacks, including geometrical distortions, in HVS based watermarking algorithms. The used perceptual model takes into account very advanced features of the HVS, fully identified from psychophysics experiments conducted in our lab. This model have been successfully applied in quality assessment or image coding schemes. In this paper, the human visual system model is used to create a perceptual mask, in order to optimize the watermark's strength. The obtained optimal watermark ensures both invisibility and robustness requirements. Contrary to most watermarking schemes using advanced perceptual masks, in order to best thwart the de-synchronization problem induced by geometrical distortions, we propose here a Fourier domain embedding and detection technique optimizing the watermark's amplitude. Finally, the obtained scheme's robustness is assessed against all attacks provided by the Stirmark benchmark. This work proposes a new digital rights management technique using an advanced human visual system model and able to resist to various kind of attacks including many geometrical distortions.

*Key words:* image watermarking, perceptual model, quantization noise, visibility thresholds, geometrical distortions

## 1 Introduction

### 1.1 Invisibility versus robustness: HVS and spread spectrum

Data hiding have been used for several purposes such as steganography, indexing, authentication, fingerprinting copyright protection or even copy control. The requirements strongly differs for each of these applications (17). Concerning the digital rights management, the main requirements are the copyright invisibility, the embedded data capacity and the robustness against most attacks the image could undergo. The first watermarking schemes performed slight luminance modifications (7), (8) or less significant bits substitutions (31). Since these techniques easily ensure the watermark invisibility and a high embedding capacity, the robustness requirements are not fulfilled. Ensuring the best invisibility versus robustness tradeoff is not obvious, for instance it is well-known that a watermark embedded into perceptually non-significant data components would easily be removed by an appropriate perceptual lossy compression (12). This observation brought the Spread Spectrum theory into watermarking techniques. In spread spectrum theory, the media (images, videos, ...) are considered as a communication channel and the embedded watermark is viewed as the signal to be transmitted through this channel. The goal is then to spread the watermark data over as many frequencies as possible. This ensures a good invisibility versus robustness tradeoff. Since most watermarking techniques are actually based on ideas from spread spectrum communications (2), we will not go here through all the details of this theory, but rather recommend readers to refer to these pioneering works using spread spectrum in watermarking context (13), (12), (2). In fact, such techniques don't guarantee the optimal invisibility, which could only be provided by using a HVS model. The interoperability between spread spectrum techniques and HVS models remains a complex issue.

### 1.2 Visibility and watermarking

Concerning the vision aspects, the watermark's invisibility is usually either empirically assumed or only tested with simplified quality metrics such as Peak Signal to Noise Ratio (PSNR) or Root Mean Square Error (RMSE). Most watermarking approaches aiming in the optimization of the robustness versus invisibility tradeoff are inspired on well-known perceptual properties from a qualitative point of view rather than on advanced quantitative visual models. This is all the more surprising since several image processing applications, such as quality assessment (9), or compression (37) made the implementation of complex perceptual models possible. In watermarking applications, a few

studies were conducted on perceptual masks creation in watermarking context. A perceptual mask is supposed to provide for every image site the maximum amount one can add or subtract without producing any visible difference. A typical example of qualitative HVS properties exploitation is addressed in (19), where the authors made the following heuristic assumptions: the noise sensitivity is weak on the image edges, smooth areas are very sensitive to variations and textured areas have a high noise sensitivity level. An edges-texture classification is then used in order to create Just Noticeable Difference (JND) Masks. The so obtained content based watermarking technique was found to be resistant against several attacks, such as JPEG compression, cropping, or Gaussian noise addition. However, although such heuristic properties can be exploited to implement simple JND masks, some very useful HVS features are not taken into account. For instance, using an advanced HVS model could allow to fully exploit the masking effects, and thus, to optimize both the invisibility and the mark's robustness. Hence, several works exploiting advanced HVS models have been conducted. The next sub section is devoted to a brief presentation of the most significant works dealing with the use of advanced HVS models in watermarking applications.

### *1.3 Advanced HVS models in watermarking*

An interesting study was conducted by Bartolini et al. (46) on perceptual masks. Here the authors built several JND masks. They used a multiple channels HVS model, which was designed to predict the visibility thresholds for simple sine wave gratings. The authors tested the watermark's robustness against JPEG compression, circular cropping, de-speckle filtering and dithering. They claimed that the masks based on heuristic considerations presented better detection results than the proposed HVS based mask. However, the exploitation of such JND masks in a spatial domain watermarking technique could probably not resist to any geometrical distortions as a de-synchronization problem would inevitably occur. Furthermore, the assumption that the watermark may be assimilated to a simple signal is erroneous. Therefore, the HVS model is not adequate in this context.

Delaigle et al. (15) have performed an interesting watermarking scheme exploiting visual discrimination of edge and texture. In this work, the authors computes the image's local energy by using analytic filters, and consider the watermark as masked if its energy remains below the computed mask's energy. In fact, an important key-point in this work was the use of watermarks presenting good auto-correlation properties. Besides these two ideas, the watermark location was not fully described and its spreading over the spectrum was not clearly detailed. Their watermarking algorithm have been tested against three attacks: noise addition, JPEG compression and low-pass filtering.

Podilchuk and Zeng (28) proposed an image adaptive watermarking algorithm for both DCT and DWT domains. On one hand, DCT domain JND masks are issued from quantization matrices established by Watson in (40). On the other hand, JND masks applying to the DWT domain are computed from visual thresholds for wavelet quantization error (35). For both embedding domains, the watermark's robustness have been tested against JPEG compression, Rescaling and Cropping. It is important to notice in these cases, that the tested scaling rates were quite small, and an interpolation was applied prior using the detection algorithm.

Kutter and Winkler presented in (3) an original watermarking technique mixing both a vision model and the spread-spectrum theory. This work takes into account advanced HVS properties, such as masking effects or band-limited local contrast, but it is based on simple signals detection (sine wave gratings). The authors clearly mentioned that their goal was to introduce optimal weighting functions, rather than to provide a full digital watermarking technique. Thus, they only tested the scheme's robustness against JPEG compression. They discussed on a few possible extensions of this work in order to resist to geometrical distortions, such as using a reference watermark for spatial synchronization, or the concept of self-reference (multiple watermark embedding).

While these techniques usually ensure a good invisibility versus robustness tradeoff, their robustness against geometrical transforms is not well established. No HVS based schemes were ever tested against all Stirmark attacks. Such studies highlight the interoperability problem between HVS models and watermarking techniques in classical transformed spaces (DCT and DWT). In the previously cited works, this implies a HVS model simplification, and thus, a loss of the model's accuracy.

#### *1.4 The proposed approach*

Besides the usual weaknesses of HVS based watermarking techniques one can cite :

- The used HVS models are based on simple signals detection thresholds. Such signals are not realistic regarding to the watermark properties.
- The interoperability between the HVS model and the watermark embedding technique may not be optimum. To improve the mark's robustness, many authors chose to embed the watermark in DCT or DWT transformed spaces, whereas these domains do not allow the implementation of a suitable HVS model.

As in (28), the work presented here also uses a model of quantization noise visibility thresholds (on complex signals rather than simple sine waves gratings) to determine JND values. These latter are evidently represented in the spatial domain, but the extraction technique operates in the Fourier domain in order to best resist to most attacks. Previous psychophysics experiments conducted in our lab led to define band-limited local contrast and associated optimal quantization laws with remarkable properties. The obtained HVS model, mainly used for quality assessment or image coding purpose have also been exploited for data hiding frameworks, limited to JND mask implementation (5), and spatial watermark embedding (6). Besides the invisibility requirement, a very important issue in watermarking schemes implementation is the definition of an optimal embedding space regarding to robustness requirements. Contrary to previous schemes exploiting perceptual masks (19), (46), (6), an important aim of this contribution is to provide a geometrical distortions resilient watermarking algorithm. Although the spatial domain is well known to provide an important embedding capacity while still ensuring the mark's invisibility, it presents very weak robustness features. The major interest of DCT or DWT based watermarking techniques lies in their possible adaptation into compression standards. However, such methods might not allow an efficient watermark detection after geometrical distortions as the ones introduced in the Stirmark benchmark (20), (27), (26). Furthermore, as explained in section 2, contrary to Fourier domain, which is well suited to model the Human Visual System behavior, both DCT and DWT domains present serious incompatibilities with the use of advanced Human Visual System models. Hence, among all possible transformed spaces, we opted for a Fourier space watermarking technique as it allows good robustness properties to many kind of distortions. Furthermore, unlike most presented HVS based watermarking schemes, this work proposes a perceptual model composed of strictly defined overlapping visual sub-bands. Such decomposition ensures that masking effects are restricted into the visual sub-band, and thus, the watermarks are completely included into the sub-bands.

Our study presents several differences with the previously cited works. The first difference lies in the weighting coefficient computation, which, in our case, is based on quantization noise visibility thresholds established during psychophysics experiments with several observers on complex images. Contrary to Kutter and Winkler (3) we wish here to build a full watermarking scheme, and thus, to provide an efficient detection algorithm. Hence, we hereby propose a spatial perceptual mask creation, combined with a Fourier domain detection technique able to resist to various kind of attacks, including several geometrical distortions. Finally, another important strength of the presented schemes lies in the clear definition of the perceptual sub-bands (unlike Delaigle et al. (15)), allowing a very good delimitation of the masking effects in the Fourier spectrum. The HVS model ensures each frequency watermark to be maintained within the visual sub-band, and the interactions with other

channels are strictly avoided. An important key-point of this work lies in the use of a Fourier domain watermarking embedding technique with an adapted weighting scheme using a spatial perceptual mask.

The next section is devoted to a detailed presentation of the used HVS model and the quantization noise visibility thresholds computation is given (equation 4). The watermarking technique presented in section 3 takes benefit of the HVS model's spectral decomposition to optimize the tradeoff between invisibility and robustness. This section details the theoretical watermark perceptual stretching to best match the perceptual mask (given in a spatial representation). Finally, experimental data on the optimum watermark strength and spectral content are given and the technique's robustness is assessed in section 4 against all Stirmark attacks.

## 2 Human visual system model

### 2.1 *visibility*

In the proposed watermarking scheme, we plan to exploit the HVS properties in order to control the watermark's invisibility. We need to compute "Just Noticeable Difference" masks (JND) as they should give the exact amount of data one can add or subtract to each image pixel without producing visible artifacts. Therefore, we should address low level parts of the HVS related to visibility thresholds mechanisms. Unfortunately, even the low level parts are not easy to model and several decades of psychophysics have been necessary to provide elements of visibility threshold prediction. Experiments on sine wave gratings have driven the emergence of the contrast sensitivity function (CSF) concept, providing the just noticeable contrast threshold at a given spatial frequency. In this case, the HVS is considered as a mono channel behavior whose MTF is the CSF. Several CSF models have been proposed in the literature (47), (48). For luminance component (grey levels images), it is well admitted that the CSF is band-pass although some studies suggested that this is not true for suprathreshold conditions leading to the concept of contrast constancy. In watermarking applications, we are concerned with visibility around threshold so usual CSFs could be exploited. Nevertheless, CSFs are not adapted to predict visibility for complex signals, essentially because they are not able to emulate masking effects. Masking effect happens when the visibility of a signal is affected by other signals. One can see masking effect as a modification of the CSF and some attempts to define the parameters that affects the CSF shape can be found in the literature. Since CSFs are related to simple signals experiments, they are not appropriate in real image context. In fact, masking effect modeling needs to consider HVS as a multichannel rather than a mono channel. Several physiological studies showed that most

cells in the HVS are tuned to specific visual information, such as color, orientation or frequency. Psychophysics experiments (10) (29) have pointed out the HVS multi-resolution structure. The HVS behavior can be modeled by spatio-frequency visual channels. The visual channels are modeled by a filter bank separating each perceptual channel also called perceptual sub-band. The major asset of such HVS decomposition would lie in the clear definition of visibility thresholds for each single visual channel. Furthermore, multi-channel HVS decomposition is perfectly suited to model masking effects. Masking effect occurs when the combination of two signals, which could be visible taken independently, produces an overall imperceptible signal. Masking effects have been the subject of many studies and several models have been proposed (21) (16) (18) (34). Most authors agree on the general shape of the HVS decomposition, this latter is usually represented as a polar representation of separable channels. Nevertheless, although the decompositions are usually based upon angular and radial selectivities, a wide variation is found on the filters parameters (30), (39), (11),(22). The complexity of the masking phenomenon led many researchers to study these effects on simple sine-waves gratings. Several studies were conducted on the masking effects of oriented sine-waves gratings (16), for different spatial frequencies (21), or even on chromatic and achromatic signals masking (24). Besides these models presenting good properties for simple sine-wave gratings, several psychophysics studies have been conducted to identify the visibility of more complex and realistic signals such as quantization noise. The results have shown some important differences comparing with simple signal visibility, but modeling in such context remains difficult. One attempt have been done by Watson (38) in order to predict the visibility of quantization noise in DWT domain. An interesting study pointing out the serious incompatibilities of the wavelet domain with HVS models have recently been conducted (41). Effectively, the Fourier spectrum properties, namely, the conjugate symmetry property are incompatible with the wavelet sub-bands. For instance, on figure 1, the  $(III, 2)$  and  $(III, 6)$  sub-bands are completely independent, each one has its own visibility threshold (see equation 4), whereas for the same spatial frequency locations, the two obtained wavelet sub-bands represent the same image content, in fact, these two sub-bands are the same in the wavelet decomposition (see (25) for a comparison between these two decompositions). We have previously developed in our lab a model of quantization noise visibility with a coherent approach according to HVS multichannel decomposition. This allows to use a more realistic visibility model than those based on simple signals. The next sections describes this visibility model, it introduces a channel decomposition and its associated masking effect model.



## 2.2 Perceptual Channel decomposition

Based on psychophysics experiments conducted in our lab, we have derived a Perceptual Channel Decomposition (PCD). The PCD's filters are similar to the cortex filters developed by Watson (36) (39). However, they have been adapted to the frequency splitting of figure 1, which is not dyadic according to radial frequencies. Moreover, in this decomposition the angular selectivity is not constant. The PCD presented above in figure 1 uses a set of three band-pass radial frequency channels (crown *III*, *IV*, *V*) each being decomposed into angular sectors with an oriented selectivity of  $45^\circ$ ,  $30^\circ$  and  $30^\circ$  respectively. Channel number *II* has been merged with the low-pass channel (crown *I*), which is non-directional, and gives rise to a simple low-pass radial frequency channel (this latter is denoted here as LF). The used Cortex filters are defined as the product between Dom filters which characterize the radial selectivity and Fan filters providing the angular selectivity. Interested readers may consult

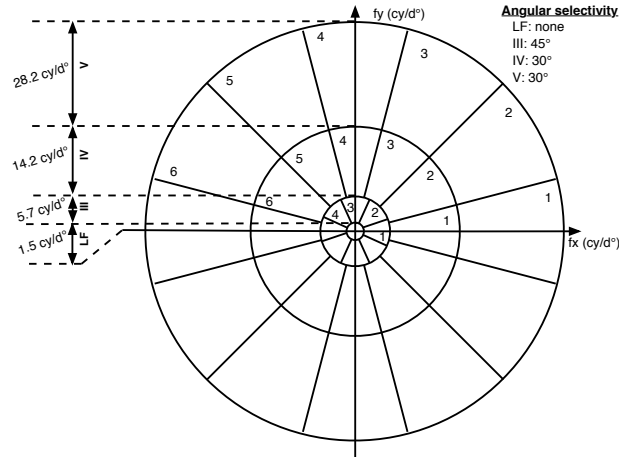


Fig. 1. Perceptual sub-band decomposition performed onto the Fourier Spectrum (32) for more details on the filters creation.

## 2.3 Masking effects through quantization noise visibility

The local band limited contrast introduced by E. Peli (23) takes into account the important fact that the perception of a detail depends on its local neighborhood. The computation of local contrast goes through a decomposition of the image into perceptual sub-bands. This local contrast for a given  $(m, n)$  image location and for a given  $i$  sub-band is defined as the ratio between a sub-band luminance and the mean luminance of the considered channel which

means the sum of all luminances located under the considered sub-band.

$$C_i(m, n) = \frac{L_i(m, n)}{\sum_{k=0}^{i-1} L_k(m, n)} \quad (1)$$

where  $i$  represents the  $i^{th}$  radial channel and  $\sum_{k=0}^{i-1} L_k(m, n)$  is the low frequency signal corresponding to the  $i^{th}$  sub-band. For the PCD presented in section 2, this local band limited contrast is slightly modified to take into account the angular selectivity :

$$C_{i,j}(m, n) = \frac{L_{i,j}(m, n)}{\sum_{k=0}^{i-1} \sum_{l=0}^{Card(l)} L_{k,l}^i(m, n)} \quad (2)$$

where  $L_{i,j}(m, n)$  and  $C_{i,j}(m, n)$  respectively represents the luminance and contrast at the  $(m, n)$  position of  $i^{th}$  radial channel and  $j^{th}$  angular sector.  $Card(l)$  is the number of angular sectors in the  $k^{th}$  radial channel. The contrast can be rewritten as :

$$C_{i,j}(m, n) = \frac{L_{i,j}(m, n)}{\bar{L}_i(m, n)} \quad (3)$$

where  $\bar{L}_i(m, n)$  is the local mean luminance at the  $(m, n)$  position (i.e. the spatial representation of all Fourier frequencies below the considered visual sub-band). The local contrast definition is chosen to determine the allowable watermark strength. Previous studies, conducted on the perceptual decomposition given in figure 1, determined invisible quantization thresholds (33). In this study, each perceptual sub-band went through a quantization process and the image overall quality was assessed by a set of observers. The threshold contrast notion ( $\Delta C$ ) has been introduced in order to provide the maximum quantization step, which do not visually affects the image. From psychophysics experiments (33), this threshold contrast is given by equation 4.

$$\Delta C_{i,j} = \frac{E_{i,j}}{L_0} \left( \frac{\Delta f_i}{f_{0,i}} \right)^{\lambda_{i,j}} \quad (4)$$

where  $\Delta C_{i,j}$  is the threshold contrast, and  $E_{i,j}$  is the  $(i, j)$  sub-band's power.  $L_0$  is the screen luminance,  $\Delta f_i$  is the  $i^{th}$  radial bandwidth,  $f_{0,i}$  is the central frequency of sub-band  $i$ , and  $\lambda_{i,j}$  is a constant depending on sub-band  $(i, j)$ . The  $\lambda_{i,j}$  values for each crown are given in the table 1.

<i>Crown</i>	II	III	IV
$\lambda_{i,j}$	-1.52	0.094	-0.28

Table 1

Parameters for intra channel masking for each DPC crown (33)

#### 2.4 Perceptual mask

As previously seen in section 2.3, we are able to define the visibility of complex signals. These thresholds will now be used in a watermarking context for the strength determination process. Since this model operates in a psychovisual space, we must transform digital images in luminance. This very important statement, which have never been taken into consideration in any watermarking algorithm, implies that a watermark might be invisible on a specific screen but would present quite annoying artifacts while viewed on another monitor. An ideal perceptual watermark should not only be image-dependent, but also display-dependent. The screen's "Gamma function" obtained from the monitor's calibration is usually used to transform the digital grey level image values into perceived luminance.

We saw in section 2.2 that the Perceptual Channel Decomposition may be used to extract any visual sub-band. The watermark can be embedded in any of the so-obtained sub-bands. The proposed masking effect model suggest that we can control the visibility on each spatial site of each sub band. So the image's most adequate sites can be easily defined by simply extracting one or several spectrum sub-band, this selection may be content based, i.e. one could select for each crown, the sub-band having the biggest energy. Derived from equation 3, the maximum variation  $\Delta L_{i,j}(m, n)$  (maximum watermark strength) allowable for each  $(i, j)$  sub-band and for each  $(m, n)$  pixel position without providing visible artifacts is given by

$$\Delta L_{i,j}(m, n) = \Delta C_{i,j} \times \bar{L}_{i,j}(m, n) \quad (5)$$

In the following, the watermark will be assimilated to a quantization noise. According to equation 5, it clearly appears that the detection threshold depends on both the local mean luminance ( $\bar{L}_{i,j}(m, n)$ ), and the quantization step value ( $\Delta C_{i,j}$ ). It is important to notice that, for each orientation of a particular crown, the local mean luminance is the same (represented by all the lowest frequencies), this means that the only difference between two masks provided by two sectors of a same crown will come from the  $\Delta C_{i,j}$  value of this particular crown. It is also interesting to note that for the seven tested images, within a frequency crown the computed  $\Delta C_{i,j}$  values for each sub-band showed only a weak variation.

### 3 Watermarking process

In (5), we have already exploited this HVS model for a JND mask creation technique. The obtained JND mask has been tested in different spaces (DCT, DWT and spatial embedding). According to the obtained results the mask implementation steps have been slightly modified (6). The perceptual mask was first created according to the content of three visual sub-bands (e.g. (III,1) (IV,2) (V,4)) and located in the image edges. Detection reliability tests showed a higher watermark robustness in the uniform areas rather than on the image edges. Although a watermark embedding in the uniform areas offers a very good robustness, the use of a powerful HVS model is crucial as the modifications in such regions are very easily perceivable. The use of the HVS model in these studies ensured the watermark invisibility for every context and optimized the robustness against many attacks. These techniques, operating a spatial domain extraction process, showed very interesting robustness properties against several distortions. However, these methods suffered of weak robustness results against most geometrical attacks. Due to its well-known adaptation to the HVS behavior, the Fourier domain presents interesting features for image watermarking techniques. Our goal in the next section is to propose a Fourier domain embedding and extraction scheme while ensuring the best watermark amplitude below the given perceptual mask.

#### 3.1 JND adaptive watermark

Once the previously detailed HVS model implemented and the sub-band dependent JND mask obtained, the watermark embedding technique itself is rather simple, the watermark is weighted to best fit within the perceptual mask. As previously detailed (section 2.4), the obtained perceptual mask is spatially defined, which means it provides a spatial JND threshold for each image pixel. As previously emphasized, for the robustness requirements, the watermark embedding and extraction should be performed in the frequency domain. Thus, regarding to the detection process, this will allow to store only a small frequency watermark patch, instead of its spatial representation, spread over the whole image's size. A frequency watermark is created and modulated onto a frequency carrier within a perceptual sub-band, its spatial representation is computed and compared to the spatial perceptual mask, the smallest difference between these two images provides the weighting coefficient. The watermark is then weighted by this weighting parameter and finally embedded into the original data. Due to the Fourier transform linearity property, the perceptual weighting coefficient can be applied either in the frequency or spatial domain. In the presented algorithm, the watermark is a square shaped zero-mean Gaussian random variable. The perceptual weighting coefficient (

$K_{i,j}$ ) is given according to equation 6.

$$K_{i,j} = \operatorname{argmin}_{m,n} \left( \left| \frac{\Delta L_{i,j}(m,n)}{W_S(m,n)} \right| \right) \quad (6)$$

where  $\Delta L_{i,j}(m,n)$  represents the previously defined visual mask and  $W_S(m,n)$  depicts the watermark's spatial representation. It is important to notice that to avoid phase reversals, only the real part of the spectrum is modified, while the imaginary part is kept unchanged. Evidently, to respect the spectrum's symmetry, the watermark's symmetry must also be respected, as shown in figure 3(a). Once weighted, the obtained watermark is finally added to the original image. This addition may be performed either spatially into the image or, after a Fourier transform, into the original spectrum coefficients. Figure 2 shows

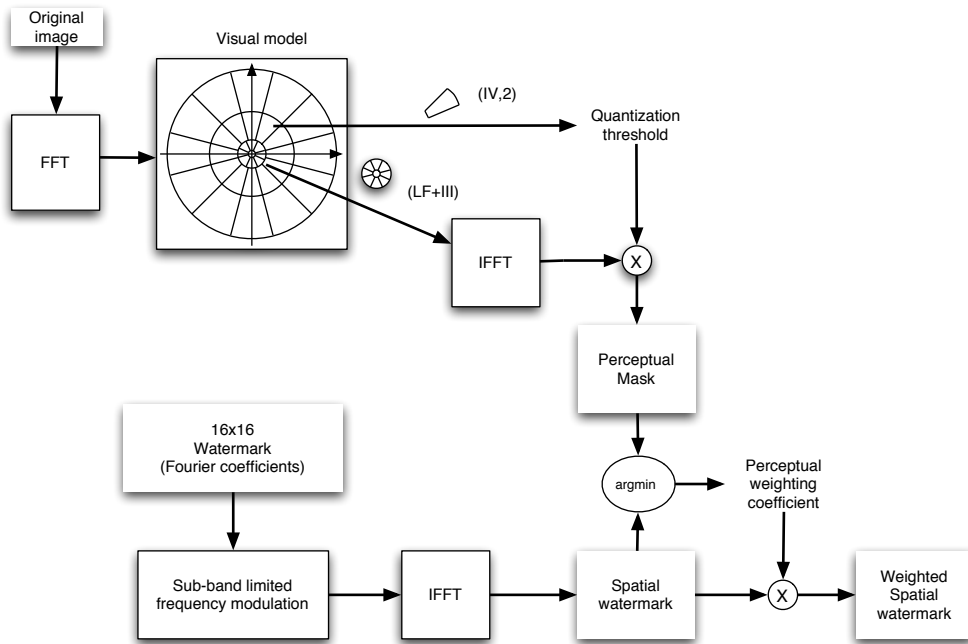
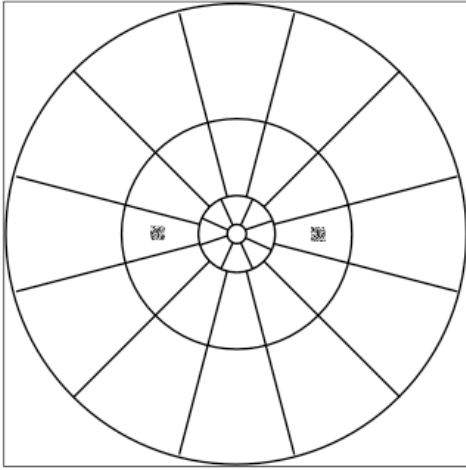
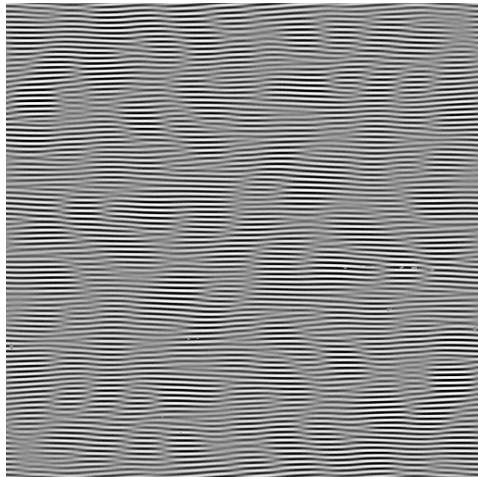


Fig. 2. Weighted watermark computation

the different steps of the weighted watermark computation. The upper branch in figure 2 represents the mask creation steps. The Fourier transform is first computed on the original image, and the Perceptual Channel Decomposition is applied on the obtained spectrum. The perceptual mask is then obtained from equation 5, and the perceptual weighting coefficient computed according to equation 6. The watermark is finally weighted with this coefficient before being embedded into the image (lower branch in figure 2). Figure 3 shows an example of frequency watermark modulated onto a frequency carrier placed in the middle of a visual sub-band (represented by the two square noise patterns, superimposed to the PCD in figure 3(a)), the corresponding watermark's spatial representation (figure 3(b)) and the obtained marked image (figure 3(c)).



(a) Frequency watermark



(b) Spatial watermark



(c) Marked image

Fig. 3. Frequency watermark in sub band (IV,1) (3(a)), spatial representation of a frequency watermark (3(b)), and the corresponding marked image (3(c))

Here, the watermark is designed as a square patch of random variables ( $16 \times 16$  coefficients), located in the visual sub-band (IV,1) (see figure 1).

### 3.2 *Detection process*

For each tested image, the cross-correlation was computed between the stored original watermark and the portion of the image's spectrum supposedly containing the mark. The only needed data during the retrieval procedure is the

original frequency watermark, and the frequency carrier where the mark is embedded (2-D coordinates). The previously computed weighting coefficient could also be stored along with the watermark frequency values in order to allow a watermark extraction, providing a reversible watermarking scheme. The cross-correlation function between two data sets  $x$  and  $y$  is given by  $r(d)$  in equation 7, where  $\bar{x}$  and  $\bar{y}$  represents the sequences' mean value.

$$r(d) = \frac{\sum_i [(x_i - \bar{x}) * (y_{i-d} - \bar{y})]}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_{i-d} - \bar{y})^2}} \quad (7)$$

The most intuitive behavior we could expect from such technique would be to obtain good detection features for soft geometrical distortions, but this schemes might not present very competitive detection rates for stronger geometrical attacks. However, it is important to notice that this technique might present very interesting detection features after an approximative geometrical distortion compensation. The detection process was reinforced by combining this one-dimensional cross-correlation function with a two-dimensional version, very useful for a re-synchronization process. The two-dimensional cross-correlation function between  $x$  and  $y$  is given according to the Fourier theorem by

$$x \star y = \mathcal{F} [\overline{X(\nu)} \cdot Y(\nu)] \quad (8)$$

where  $\star$  is the correlation operator,  $X(\nu)$  and  $Y(\nu)$  respectively represent the Fourier spectrums of  $x$  and  $y$ , and  $\overline{X}$  is  $X$ 's complex conjugate. The main advantage of such correlation is to help locating a correlation peak shift in two dimensions. This peak shift information may be very useful to compensate a possible geometrical distortion. Effectively, by comparing the new location of the correlation peak to the watermark's original location (the stored frequency carrier), we can easily determine the possible distortions the image underwent. Although we tested here the watermarking scheme's detector response, we did not focus on the threshold selection or the false alarm rate (43), (44), (45). Further works will be devoted to this specific detection threshold selection problem in the context of JND mask stretched watermarks.

## 4 Results

As the visual model performance has already been proven and assessed elsewhere (33), (6), this section is devoted to the watermark's robustness assessment.

#### 4.1 Preliminary tests

In order to optimize the data detection, several watermark sizes have been tested. The detection reliability is closely related to the watermark length. Too small watermarks usually cause important false detection results, whereas the reliability increases with the watermark size. We tested the watermarking detection reliability with four images, each underwent a typical attack for seven different watermark lengths. Square shaped watermarks of size:  $8 \times 8$ ,  $12 \times 12$ ,  $16 \times 16$ ,  $20 \times 20$ ,  $24 \times 24$ ,  $28 \times 28$  and  $32 \times 32$  were independently embedded into four images' spectrums (horizontal middle frequency range as depicted on figure 3(a)), Stirmark distortions were applied and the detection process was performed. Figure 4 shows the cross-correlation results (Y-axis) according to the watermark size (X-axis). It clearly appears here that for a watermark size above  $16 \times 16$  (256 coefficients), the scheme's robustness sensibly decreases for most attacks. This figure also confirms a possible weakness of the method against strong geometrical distortions (rotation and scaling are considered here). Once the watermark size defined, we tested the watermarking

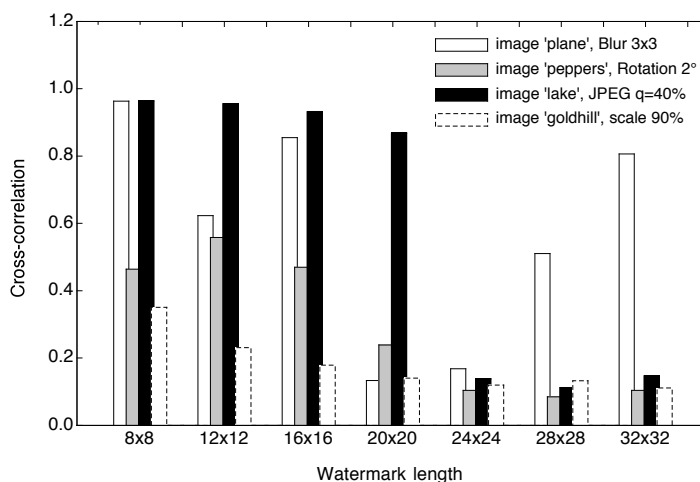


Fig. 4. Watermark length optimization. Images 'plane', 'peppers', 'lake' and 'gold-hill' respectively attacked by a  $3 \times 3$  blurring, a 2 degrees rotation, a q=40% JPEG coding, and a 90% Scaling

scheme's behavior for various frequency range, i.e. the watermark was respectively embedded in the PCD crowns labeled *III*, *IV* and *V* (Fig. 1) and its robustness was tested for each sub-band. Figure 5 shows the obtained normalized cross-correlation coefficient for each of the 89 Stirmark attacks (20), (27), (26). Cross-correlation of all Stirmark output images were sorted in alphabetical order. Associations between the distortion name and the corresponding attack index (X-axis) are given on the top of figure 5. These associations are



more precisely detailed in table 2. The figure 5 clearly shows the best detection

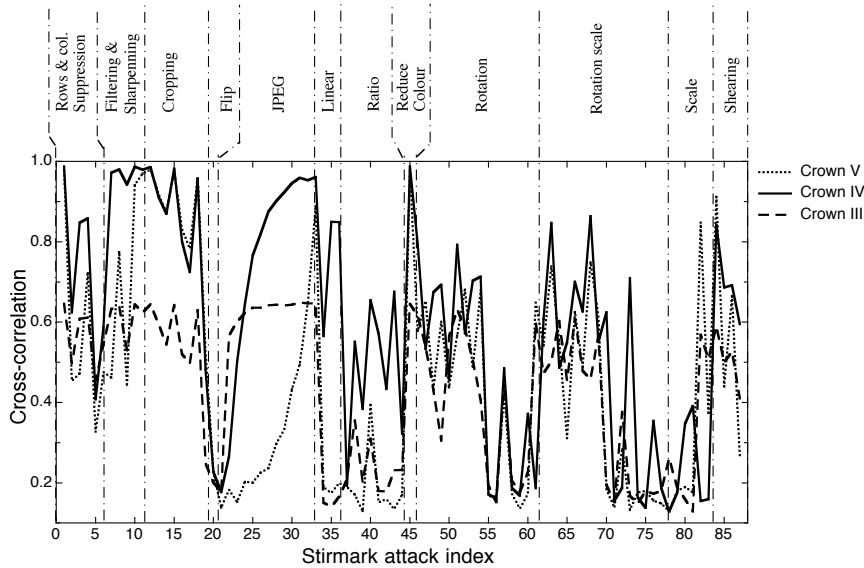


Fig. 5. Cross-correlation coefficient for the *barbara* image against the 89 Stirmark attacks when the III (dashed line), IV (solid line) and V (dotted line) PCD crown are respectively marked.

results for mid-frequency range coefficients (results for crown IV are represented by the solid line). As explained in section 2, the local mean luminance being the same for all sectors of a specific crown (denominator in equation 2), the only difference between the JND masks issued from a single crown will come from the quantization steps (1), which usually present weak variations. Hence, from now on, the chosen embedding crown is the  $(IV, 1)$ . Evidently, one can choose to embed multiple watermarks, either in the same crown or in independent crowns. The authors are presently working on an objective quality assessment technique based on multiple watermarks embedding using the presented embedding framework. Based on this selection of the  $(IV, 1)$  sub-band, the presented perceptual watermarking algorithm have been tested on 7 gray level images (whose names are given in Table 3) against the 89 Stirmark attacks (this makes a total of 623 tested images). Figures 4 and 5 respectively shows the optimum watermark length and the optimum embedding frequency range ( $16 \times 16$  watermarks in crown IV). Before performing detection tests on the Stirmark benchmark, a detection threshold must be defined. We then tested the watermark's robustness against false alarm. The detection process was tested for six hundred different watermarks into an unmarked image, the normalized cross correlation was computed for every trial. This process was repeated for three images. For the image *lena* degraded by a one degree rotation, the 600 cross correlations coefficients are plotted in figure 6. According to this detector response, three empirically chosen detection threshold ( $T=0.3, 0.4$

and 0.5) are tested in the following (see table 3). The choice of  $T=0.4$  seems reasonable and is used in the next figures to assess the detection power of this algorithm. This latter allows a good detection rate while minimizing the false alarm. Once again, it is important to notice here that our main goal was to provide the best detection correlation for a HVS model based watermarking and to resist to geometrical distortions, thus, we did not focus on the optimal detection threshold selection. This particular topic will be studied further in a another work dealing with image watermarking in a quality assessment framework.

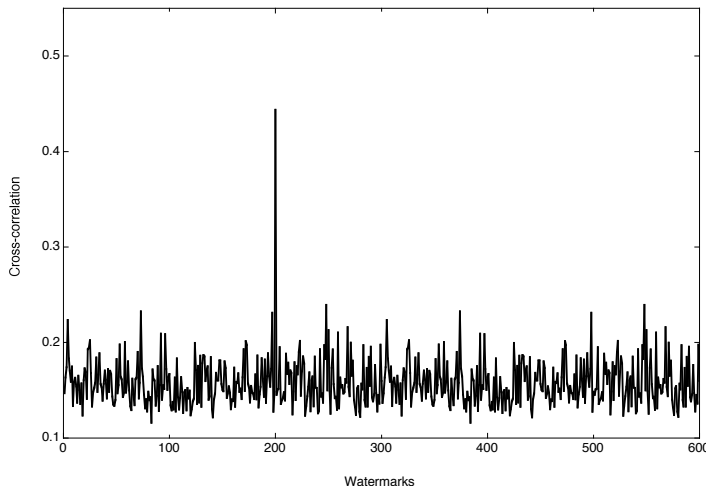
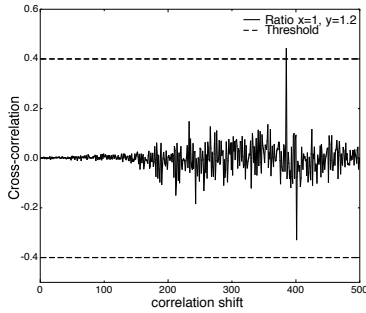


Fig. 6. Detector response. Only one matching watermark was found at position 200 (image "lena", 1 degree rotation).

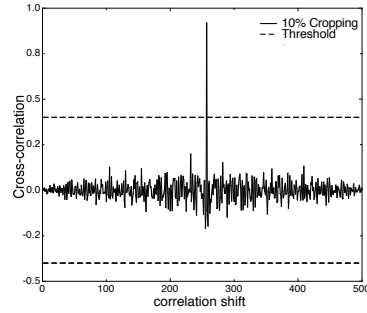
#### 4.2 Detection results

The one-dimensional cross-correlation was computed for seven images, and for all Stirmark attacks. Four cross-correlation plots are given in figure 7. In these figures, the horizontal axis represents the shift parameter ( $d$ ) in equation 7, whereas the vertical axis is the cross-correlation value ( $r(d)$ ). The dashed lines on figure 7 represent a detection threshold empirically set to 0.4. As the cross-correlation function is given on the interval  $[-1, 1]$ , the threshold's negative value is also represented. Figure 7 and 8 respectively show the one-dimensional and two-dimensional cross-correlation graphs when the image underwent a ratio ( $x=1.0, y=1.2$ ) (figures 7(a) and 8(a)), a 10% cropping (figures 7(b) and 8(b)), a 2 degrees rotation (figures 7(c)) and 8(c)), and finally when 5 rows and 1 column have been removed (figures 7(d)) and 8(d)). As theoretically expected from the data retrieval technique, although the detection threshold always allow a successful retrieval (in figure 7), the presented watermark-

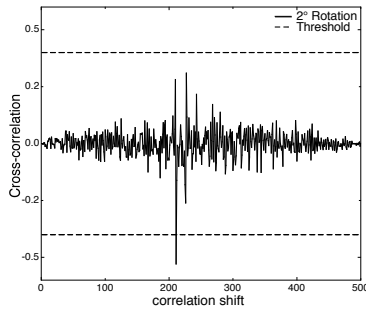
ing scheme presented a weaker robustness against geometrical distortions.



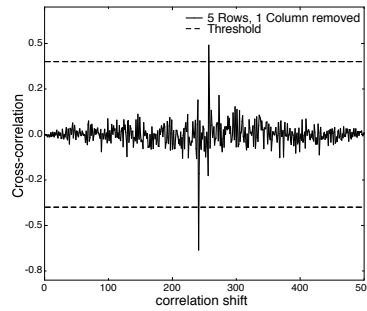
(a) Ratio 1.0, 1.2



(b) 10% Cropping



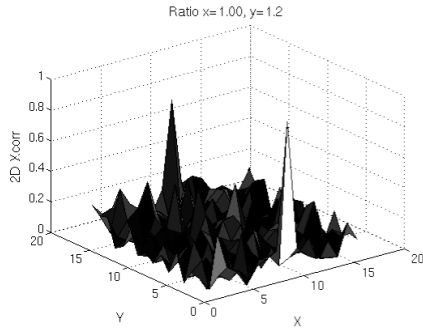
(c) 2 degrees rotation



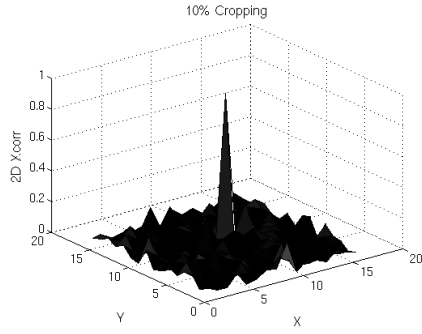
(d) 5 rows, 1 col removed

Fig. 7. 1D cross-correlation results

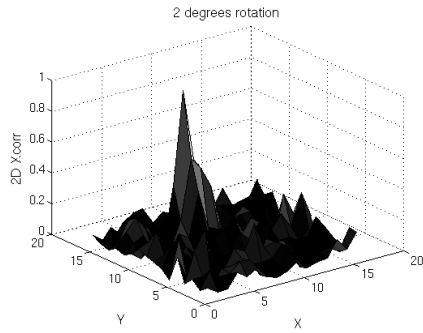
For such distortions, operating a mark de-synchronization, on the 2D correlation plots (figure 8(d)), the correlation peaks clearly moved compared to the watermark carrier (supposedly placed at the center of the map). The cross-correlation results for the whole set of Stirmark attacks are given in figure 9 for the 7 tested images. The readers may refer to figure 5 or table 2 for more details on the attacks index (x axis on figure 9). We can easily observe that with the previously defined detection threshold (set to 0.4, represented by the dashed lines), the watermark is still detected for most of the Stirmark attacks, including several geometrical distortions. As expected from the detection process, the watermark is usually detected as long as the distortion keeps at least a part of the  $16 \times 16$  noise like sequence within the  $16 \times 16$  checked Fourier coefficients. Considering the 623 tested images (7 test images and 89 attacks), the mean detection rate is found to be about 62% with a detection threshold set to  $T=0.5$ , it equals 70% with a threshold set to 0.4. And finally, this detection rate reaches 77% with a threshold set to 0.3 (see



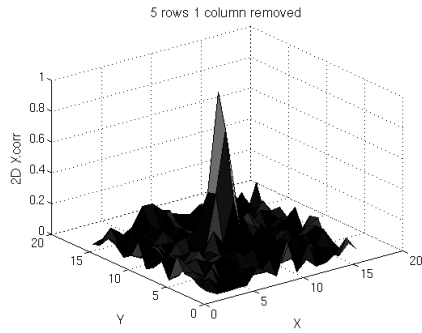
(a) Ratio 1.0, 1.2



(b) 10% Cropping



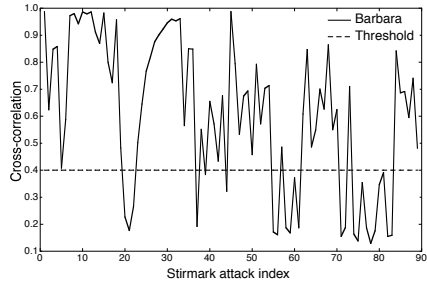
(c) 2 degrees rotation



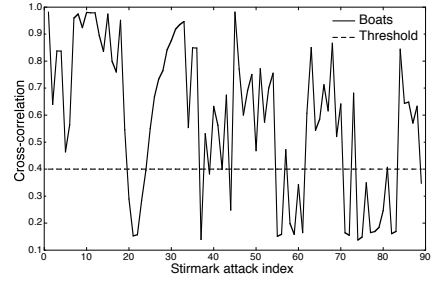
(d) 5 rows, 1 col removed

Fig. 8. 2D cross-correlation results

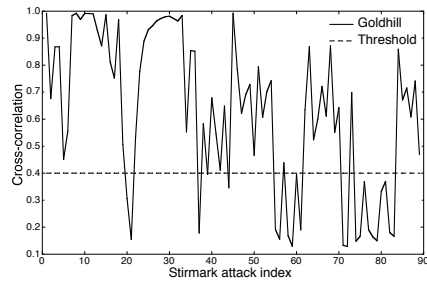
table 3), which is still a fair threshold regarding to the false alarm rate (figure 6). More detailed detection rates are given in Table 3, which shows both the number of detected watermarks for the 89 attacked images (lines 1, 3 and 5), and the corresponding detection rate (lines 2, 4 and 6). The last column gives the total number of detected watermarks along with the appropriate rate for the 623 tested images. During the detection process, attacked images have to be scaled to the original image's resolution, in order to seek into the same frequency range, a zero padding is usually performed for cropped versions. Note that using an interpolation technique may improve the detection results. The storage of the frequency carrier as a percentage of the spectrum size could also be used to perform the detection process, and would in fact avoid the zero padding technique. However, this might not sensibly increase the overall detection rate.



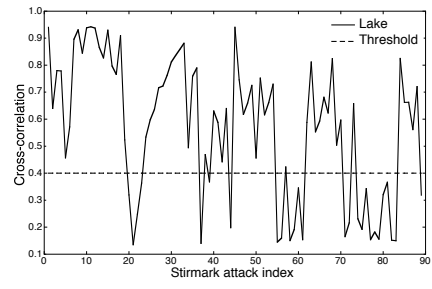
(a) *Barbara* image



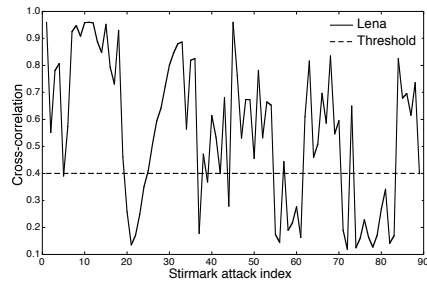
(b) *Boats* image



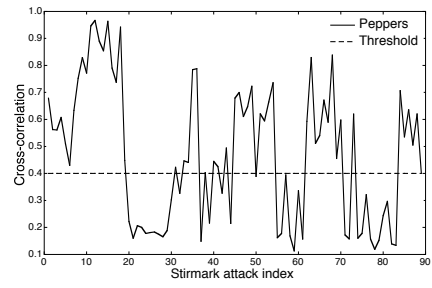
(c) *Goldhill* image



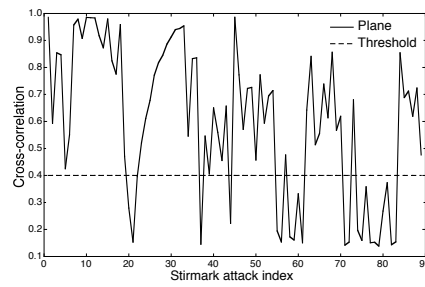
(d) *Lake* image



(e) *Lena* image



(f) *Peppers* image



(g) *Plane* image

Fig. 9. Cross-correlation coefficient for the 89 Stirmark attacks for 7 test images.

1	Marked image	24	JPEG 20	47	rotation -0,50	70	rot scale 1,0
2	17row 5col rem	25	JPEG 25	48	rotation -0,75	71	rot scale 10,0
3	1row 1col rem	26	JPEG 30	49	rotation -1,0	72	rot scale 15,0
4	1row 5col rem	27	JPEG 35	50	rotation -2,0	73	rot scale 2,0
5	5 row 17 col rem	28	JPEG 40	51	rotation 0,25	74	rot scale 30,0
6	5 row 1 col rem	29	JPEG 50	52	rotation 0,5	75	rot scale 45,0
7	4x4 med filter	30	JPEG 60	53	rotation 0,75	76	rot scale 5,0
8	2x2 med filter	31	JPEG 70	54	rotation 1,0	77	rot scale 90,0
9	3x3 med filter	32	JPEG 80	55	rotation 10,0	78	scale 0,50
10	Gauss filt 3x3	33	JPEG 90	56	rotation 15,0	79	scale 0,75
11	Sharp 3x3	34	linear Filt 1	57	rotation 2,0	80	scale 0,9
12	cropping 1	35	linear Filt 2	58	rotation 30,0	81	scale 1,1
13	cropping 10	36	linear Filt 3	59	rotation 45,0	82	scale 1,5
14	cropping 15	37	ratio (0,8 1,0)	60	rotation 5,0	83	scale 2,0
15	cropping 2	38	ratio (0,9 1,0)	61	rotation 90,0	84	shear (0,0 1,0)
16	cropping 20	39	ratio (1,0 0,8)	62	rot scale -0,25	85	shear (0,0 5,0)
17	cropping 25	40	ratio (1,0 0,9)	63	rot scale -0,50	86	shear (1,0 0,0)
18	cropping 5	41	ratio (1,0 1,10)	64	rot scale -0,75	87	shear (1,0 1,0)
19	cropping 50	42	ratio (1,0 1,20)	65	rot scale -1,0	88	shear (5,0 0,0)
20	cropping 75	43	ratio (1,1 1,00)	66	rot scale -2,0	89	shear (5,0 5,0)
21	flip	44	ratio (1,2 1,0)	67	rot scale 0,25		
22	JPEG 10	45	reduce colour	68	rot scale 0,50		
23	JPEG 15	46	rotation -0,25	69	rot scale 0,75		

Table 2

List of Stirmark attacks along with the corresponding index

## 5 Conclusion

This work presents a new watermarking technique for copyright protection using an advanced HVS model. The used perceptual model ensures the best invisibility capacity and robustness trade-off. Contrary to most perceptual JND masks, the visibility thresholds used in this study are computed for complex signals rather than for simple gratings. The HVS model provides a spatially defined Just Noticeable Difference mask, which gives the appropriate

Threshold	image	Barbara	Boats	Goldhill	Lake	Lena	Peppers	Plane	Total
0.5	Nb marks	58	58	61	57	54	42	59	389
0.5	Rate (%)	65	65	69	64	61	47	66	62
0.4	Nb marks	65	62	66	63	60	53	66	435
0.4	Rate (%)	73	70	74	71	67	60	74	70
0.3	Nb marks	71	68	73	71	66	60	70	479
0.3	Rate (%)	80	76	82	80	74	67	79	77

Table 3

Detection percentage against Stirmark attacks

weighting coefficient for a given watermark whose optimal length have been assessed. The weighted watermark is finally embedded into the image, and the detection process takes place in the Fourier domain. The watermark, a square pseudo-random sequence, is embedded on Fourier domain coefficients and its amplitude is stretched to best match the spatially defined perceptual mask. The detection process performs a normalized cross correlation function between the marked (and possibly attacked) part of the Fourier spectrum and the original weighted watermark. This detection process provided very good robustness results for a large set of attacks, including soft geometric distortions. However, for stronger geometrical attacks, the detection process showed weaker results. In fact, the presented detection scheme is able to detect the mark as long as a geometric distortion keeps a small part of the watermark into its original location, i.e. as long as the new watermark location (after being attacked) overlaps its original position. The main drawback such attacks would induce is a mark de-synchronization problem : the detection process only looks for the mark at a selected carrier and may not be able to detect it when this latter moved far away from the carrier. The detector scheme was tested on the 89 distorted versions of 7 tested images, the obtained overall detection rate was about 70% for a detection threshold set to 0.4. Furthermore, the 2D-crosscorrelation detection was proven to detect the degradation type, and future works will be devoted to compensate such distortions before using the detection process. Evidently, for several Stirmark attacks the presented detection process is not able to detect the mark, but usually, these distortions severely degrade the images' quality or semantic content (rotations greater than 10 degrees, important scaling, 10 to 15 percent JPEG quality compression). Such important attacks modify the images so severely that the obtained images don't represent any commercial interest.

## References

- [1] A. Z. Tirkel, G. A. Rankin, R. van Schyndel, W. J. Ho, N. Mee, C. F. Osborne, Electronic Watermark, *Digital Image Computing, Technology and Applications*, Sydney Australia, 1993, pp. 666-672.
- [2] F. Hartung, J. K. Su and B. Girod", Spread Spectrum Watermarking: Malicious Attacks and Counterattacks, in *Security and Watermarking of Multimedia Contents, Proc. SPIE 3657*, pp. 147-158, Jan. 1999.
- [3] M. Kutter, S. Winkler, A vision-based masking model for spread-spectrum image watermarking, *IEEE Transactions on Image Processing*, vol.11, no.1, pp.16-25, January 2002
- [4] Florent Autrusseau, Jeanpierre Guédon, Image Watermarking in the Fourier Domain Using the Mojette Transform. *Santorini Greece 14th IEEE International Conference on Digital Signal Processing (DSP2002)*, vol. II, pp. 725-728, July 2002.
- [5] Florent Autrusseau, Abdelhakim Saadane, and Dominique Barba. Psychovisual selection of auspicious sites for watermarking. *SPIE PICS*, pages 326–329, 2000.
- [6] Florent Autrusseau, Abdelhakim Saadane, and Dominique Barba. A psychovisual approach for watermarking. *SPIE Electronic Imaging*, 4314(41), pages : 378–386, 2001.
- [7] W. Bender, D. Gruhl, N. Morimoto, and A. Lu. Techniques for data hiding. *IBM systems journal*, 35(3/4), pages : 313–336, 1996.
- [8] Gordon W. Braudaway. Protecting publicly-available images with an invisible image watermark. *Proceedings of the International Conference on Image Processing*, 1, pages : 524–527, 1997.
- [9] Patrick Le Callet, Dominique Barba A robust quality metric for color image quality assessment. *Int. Cont. on Image Proc.* 2003, (1) pp. 437-440.
- [10] F.W. Campbell and J.G. Robson. Application of fourier analysis to the visibility of gratings. *Journal of Physiology*, 197, pages : 551–566, 1968.
- [11] S. Comes and B. Macq. Human visual quality criterion. *Proceedings of the SPIE, Visual Communications and Image Processing*, 1360, pages : 2–13, 1990.
- [12] I. J. Cox, J. Kilian, and T. Shamoan. Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing*, 6(12), pages : 1673–1687, 1997.
- [13] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoan. Secure spread spectrum watermarking for images, audio and video. *Proceedings of the International Conference on Image Processing*, pages 243–246, 1996.
- [14] I. J. Cox, M. L. Miller The First 50 Years of Electronic Watermarking. *EURASIP Journal of Applied Signal Processing*, 2, 126-132, 2002. Ingemar J. Cox and Matt L. Miller
- [15] J.F. Delaigle, C. De Vleeschouwer, and B. Macq. Watermarking algorithm based on a human visual model. *Signal Processing*, 66, pages : 319–335, 1998.



- [16] J. M. Foley. Human luminance pattern mechanisms : Masking experiments require a new model. *Journal of the Optical Soc. of America*, 11(6), pages : 1710–1719, 1994.
- [17] J. Fridrich. Applications of data hiding in digital images. In *Tutorial for the ISSPA'99 Conference*, pages 22–25, Brisbane, Australia, August 1999.
- [18] D. J. Heeger. Normalisation of cells responses in cat striates cortex. *Visual Neuroscience*, 9, pages : 181–198, 1992.
- [19] Mohan S. Kankanhalli, Rajmohan, and K. R. Ramakrishnan. Content based watermarking of images. *ACM Multimedia - Electronic Proceedings*, 1998.
- [20] Martin Kutter and Fabien A. P. Petitcolas. A fair benchmark for image watermarking systems. *Electronic Imaging'99, Security ans Watermarking of Multimedia Contents*, 3657, Janunary 1999.
- [21] G. E. Legge and J. M. Foley. Contrast masking in human vision. *Journal of the Optical Soc. of America*, 70, pages : 1458–1471, 1980.
- [22] J. Lubin. The use of psychophysical data and models in the analysis of display system performance. *Digital Images and Human Vision*, pages 162–178, 1993.
- [23] E. Peli. Contrast in complex images. *J.O.S.A.*, 7(10), pages : 2032–2040, 1990.
- [24] A. Bradley, E. Switkes, K. K. De Valois. Contrast dependence and mechanisms of masking interactions among chromatic and luminance gratings *Journal of the optical society of America*, vol. 5 (7), pp. 1149-1159, 1988.
- [25] A. Bradley. A Wavelet Visible Difference Predictor *IEEE Transactions on Image Processing*, vol. 8 (5), pp. 717-730, 1999.
- [26] Fabien A. P. Petitcolas. Watermarking schemes evaluation. *I.E.E.E. Signal Processing*, vol. 17, no. 5, pp. 5864, September 2000
- [27] Fabien A.P. Petitcolas, Ross J. Anderson, and Markus G. Kuhn. Attacks on copyright marking systems. *Second Workshop on information hiding*, 1525, pages : 218–238, 1998.
- [28] Christine I. Podilchuk and Wenjun Zeng. Image-adaptive watermarking using visual models. *IEEE Journal on Selected Areas in Communications*, 16(4), pages : 525–539, May 1998.
- [29] M.B. Sachs, J. Nachmias, and J.G. Robson. Spatial frequency channels in human vision. *Journal of the optical society of America*, 61(9), pages : 1176–1186, 1971.
- [30] D.J. Sakrison. On the role of the observer and a distortion measure in image transmission. *IEEE Trans. on Com.*, 25(11), pages : 1251–1267, 1977.
- [31] R.G. Van Schyndel, A.Z. Tirkel, and C.F. Osborne. A digital watermark. *Proceedings of the International Conference on Image Processing*, 2, pages : 86–90, 1994.
- [32] H. Senane, A. Saadane, Dominique Barba *Image coding in the context of a psychovisual image representation with vector quantization* Int. Conf. on Image Proc. 1995, pp. 97-100, 1995.
- [33] N. Bekkat, A. Saadane, D. Barba: Masking effects in the quality as-

- assessment of coded images, SPIE Human Vision and Electronic Imaging V (2000), Vol. 3959, pp. 211-219.
- [34] P.C. Teo and D. J. Heeger. Perceptual images distortion. *Proc. of SPIE*, 2179, pages : 127–141, 1994.
- [35] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor. Visual thresholds for wavelet quantization error. *SPIE Proceedings, Human Vision and Electronic Imaging*, 1996.
- [36] Andrew B. Watson. The cortex transform : Rapid computation of simulated neural images. *Computer Vision, Graphics, And Image Processing*, 39, pages : 311–327, 1987.
- [37] Andrew B. Watson. Image data compression having minimum perceptual error. *US Patent 54629780*.
- [38] A. B. Watson, J. A. Solomon. Model of visual contrast gain control and pattern masking. *Journal of the Optical Society of America*, vol. 14 (9), pp. 2379-2391, 1997.
- [39] Andrew B. Watson. Efficiency of a model human image code. *J. O. S. A.*, 4(12), pages : 2401–2417, 1987.
- [40] Andrew B. Watson. DCT quantization matrices visually optimized for individual images. *Proc. of the SPIE Conference on Human Vision, Visual Processing and Digital Display IV*, 1913, pages : 202–216, February 1993.
- [41] W. Zeng, S. Daly, S. Lei. An overview of the visual optimization tools in JPEG 2000 *Signal Processing: Image Communication*, 17 (2002) pp. 85-104.
- [42] C. Zetsche and G. Hauske. Multiple channel model for the prediction of subjective image quality. *Proceedings of the SPIE, Human Vision, Visual Processing and Digital Display*, pages 209–216, 1989.
- [43] M. L. Miller, J. A. Bloom, Computing the probability of false watermark detection *Workshop on information hiding, Dresden Germany, Sept. 29 - Oct. 1, 1999*
- [44] JP. Linnartz, T. Kalker, G. Depovere Modelling the false alarm and missed detection rate for electronic watermarks *2nd Workshop on Information Hiding, Portland, OR, 15-17 April, 1998*
- [45] A. Piva, M. Barni, V. Cappellini Threshold selection for correlation-based watermark detection *Proceedings of COST 254 Workshop on Intelligent Communications, L'Aquila, Italy, June 4-6, 1998, pp. 67-72*
- [46] F. Bartolini, M. Barni, V. Cappellini, and A. Piva. Mask building for perceptual hiding frequency embedded watermarks. *Proceedings of the International Conference on image Processing*, 1, pages : 450–454, 1998.
- [47] P.G.J. Barten. Contrast sensitivity of the human eye and its effects on image quality. *SPIE, Bellingham*, 1999.
- [48] S. Daly. The visible difference predictor : An algorithm for the assessment of image fidelity. *Proc. of the SPIE, Human Vision, Visual Processing and Digital Display*, III, pages : 2–15, 1992.