Research Article

A Perceptive Approach to Digital Image Watermarking Using a Brightness Model and the Hermite Transform

Boris Escalante-Ramírez ¹ and **S. L. Gomez-Coronel** ²

¹*Facultad de Ingeniería, Departamento de Procesamiento de Señales, Universidad Nacional Autónoma de México, Ciudad de México, Mexico*

²*Departamento de Ingeniería, Instituto Politécnico Nacional, UPIITA, Av. IPN No. 2580., Col. La Laguna Ticomán, 07340 Ciudad de México, Mexico*

Correspondence should be addressed to Boris Escalante-Ramírez; boris@unam.mx

Received 20 September 2017; Revised 10 January 2018; Accepted 28 January 2018; Published 29 April 2018

Academic Editor: Khaled Loukhaoukha

Copyright © 2018 Boris Escalante-Ramírez and S. L. Gomez-Coronel. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This work presents a watermarking technique in digital images using a brightness model and the Hermite Transform (HT). The HT is an image representation model that incorporates important properties of the Human Vision System (HVS), such as the analysis of local orientation, and the model of Gaussian derivatives of early vision. The proposed watermarking scheme is based on a perceptive model that takes advantage of the masking characteristics of the HVS, thus allowing the generation of a watermark that cannot be detected by a human observer. The mask is constructed using a brightness model that exploits the limited sensibility of the human visual system for noise detection in areas of high or low brightness. Experimental results show the imperceptibility of the watermark and the fact that the proposed algorithm is robust to most common processing attacks. For the case of geometric distortions, an image normalization stage is carried out prior to the watermarking.

1. Introduction

The copyright protection of digital contents has become a great problem due to the increase of piracy, cloning of digital documents, and the espionage into the different mass media. The rapid growth of Internet, digital multimedia, and communication systems has increased exponentially the distribution of digital information (voice, data, images, and video), making evident the growing need to protect digital content. A solution for the protection of copyright and intellectual property is watermarking. For digital images, the process consists basically of embedding information into the code related to the author or copyright holder. The quality of an image watermarking technique is measured in terms of robustness, legibility, imperceptibility, and ambiguity [1, 2]. However, it is difficult to find a technique that embraces all of them, since robustness implies introducing stronger image distortions that compromise the watermark imperceptibility. For over a decade, different algorithms seeking to meet the above requirements have been proposed. Some of the

proposed techniques are based on image transformations, using alternative representation models. They attempt to add the mark in the image transform domain [3]. Among these cases, we find the Discrete Fourier Transform (DFT), the Discrete Wavelet Transform (DWT), the Discrete Cosine Transform (DCT), the Contourlet transform, and other techniques. [4–12]. Nevertheless, their use does not guarantee by itself a robust watermarking technique. Some models use the HVS characteristics in order to obtain good results of imperceptibility and robustness, taking advantage of the sensitivity of frequency, luminance, and masking contrast. The watermark that exploits the perceptual information is named a perceptual watermark [9, 10, 13–15]. Transform domain techniques that use perceptual masks based on HVS properties have proved to be more robust since they resist geometric and filter attacks. For example, the algorithm described in [10] uses the DWT and a mask to determine the coefficients of detail where the watermark will be inserted. For the generation of the mask, the texture content and luminance in all frequency bands of the image are considered

according to certain rules based on the HVS. The model by Barni et al. [10] actually considers the reduced eye sensitivity for detecting noise in the borders and the high and low brightness or luminance, as well as heavily texturized regions of an image. The result is a mask that takes into account the content of all the subbands of that image. Their results show that this technique is robust to common processing operations and has been used as a benchmark by several following investigations. In [16], similar ideas to those of scheme [10] have been used, the difference being that the latter works with the Hermite Transform (HT), and some modifications are set for the calculation of the perceptive mask. It can be regarded as the first work of watermarking using the HT. The present proposal is based on a perceptive approach that includes a brightness model and a normalization scheme of the image. Unlike most watermarking algorithms, which insert the mark considering the edges and homogeneous zones of the image, we use the brightness model to generate a perceptive mark and identify the image regions where the watermark detection becomes a difficult task for the human eye, that is, regions that are more likely to be modified without producing perceptive changes, thus assuring the invisibility of the mark. In order to generate the mask, the following elements are considered: luminance to brightness map, contrast sensitivity, and light adaptation threshold. These elements allow identifying the image structures and locations where additional information can be embedded without being perceived by a human observer. The main idea derives from the model proposed in [10] inspired by adaptive quantization in image compression schemes. In this approach, a perceptive mask is built based on the argument of the reduced visual sensitivity to noise in high resolution bands, in areas with high or low brightness, and in textured areas. Our approach shares the same principle; however we include a more elaborated luminance to brightness model that considers the multichannel mechanism that the human visual system uses to build the psychophysical perception of brightness [17]. In order to find textured areas, we use the high order Hermite Transform coefficients, which are known to represent perceptually relevant image structures; however, the visual perception of texture is also determined by the contrast and luminance; therefore we use the light adaptation threshold and contrast sensitivity measures that account for a better identification of textured areas where information can be perceptively hidden. In order to improve robustness against geometric attacks, we propose the use of image normalization techniques [1, 12, 18] that transform the original image so that the orientation and the scale of the objects are fixed. For this purpose, we employ geometric moments and invariants.

The Hermite Transform (HT) is a mathematical tool that allows the local analysis of the visual information of an image and builds scale space relations among its pixel intensities. Moreover, the HT incorporates the Gaussian derivative model of early vision [19–21] that considers the derivatives of Gaussian functions as suitable models for ganglionic and cortical visual cells. Similar to DWT, the HT also decomposes the image in a number of coefficients, where zero order coefficients represent a Gaussian-weighted image average.

Larger order coefficients contain the image details; hence, the watermark data is to be inserted here. The mathematical theory of the HT is introduced in Section 2. The map of luminance brightness for the generation of the perceptive mask is described in Section 3, whereas, in Section 4, the proposed algorithm is detailed. In Section 5, the procedure for the image normalization based on [18] is presented. The results of experimental evaluations of the algorithm and the comparisons with current techniques appear in Section 6. Finally, the conclusions are detailed in Section 7.

2. The Hermite Transform

The Hermite Transform [19–22] is a special case of polynomial transform, which is a technique of signal decomposition. The original signal $L(x, y)$, where (x, y) are the coordinates of the pixels, can be located by multiplying the window function $V(x - p, y - q)$ by the positions p, q that conform to the sampling lattice S :

$$L(x) = \frac{1}{W(x, y)} \sum_{p, q \in S} L(x, y) V(x - p, y - q). \quad (1)$$

The periodic weighting function is then defined as

$$W(x, y) = \sum_{p, q \in S} V(x - p, y - q). \quad (2)$$

The unique condition that allows the polynomial transform to exist is that the weighting function must be different from zero for all coordinates (x, y) .

The local information within every analysis window will then be expanded in terms of an orthogonal polynomial set. The polynomials $G_{m,n-m}(x, y)$, used to approximate the windowed information, are determined by the analysis window function and satisfy the orthogonal condition

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} V^2(x, y) G_{m,n-m}(x, y) G_{j,i-j}(x, y) dx dy \\ = \delta_{ni} \delta_{mj} \end{aligned} \quad (3)$$

for $n, i = 0, 1, \dots, \infty$; $m = 0, \dots, n$; and $j = 0, \dots, i$.

The polynomial coefficients $L_{m,n-m}(p, q)$ are calculated by convolving the original image $L(x, y)$ with the filter function $D_{m,n-m}(x, y) = G_{m,n-m}(-x, -y)V^2(-x, -y)$ followed by a subsampling in the positions (p, q) of the sampling lattice S :

$$\begin{aligned} L_{m,n-m}(p, q) \\ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L(x, y) D_{m,n-m}(x - p, y - q) dx dy. \end{aligned} \quad (4)$$

The orthogonal polynomials associated with $V^2(x)$ are known as Hermite polynomials:

$$G_{n-m,m}(x, y) = \frac{1}{\sqrt{2^n (n-m)!}} H_{n-m} \left(\frac{x}{\sigma} \right) H_m \left(\frac{y}{\sigma} \right), \quad (5)$$

where $H_n(x)$ denotes the Hermite polynomial of order n .

FIGURE 1: Spatial representation of the Hermite Transform coefficients. The diagonals represent the coefficients of order zero $n = 0$; the coefficients of order one $n = 1$; and those of order two $n = 2$ (figure reproduced from S. Gomez-Coronel et al. (2016) [under the Creative Commons Attribution License/public domain]).

In the case of the Hermite Transform, it is possible to demonstrate that the filter functions $D_{m,n-m}(x, y)$ correspond to Gaussian derivatives of order m in x and $n - m$ in y , in agreement with the Gaussian derivative model of early vision [23]. Moreover, the window function resembles the receptive field profiles of human vision:

$$V(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right). \quad (6)$$

Besides constituting a good model for the overlapped receptive fields found in physiological experiments, the choice of a Gaussian window can be justified because it minimizes the uncertainty principle in the spatial and frequency domains. The recovery process of the original image consists in interpolating the transform coefficients through the proper synthesis filters. This process is known as inverse polynomial transform and is defined by

$$\hat{L}(x, y) = \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{p,q \in S} L_{m,n-m}(p, q) P_{m,n-m}(x - p, y - q). \quad (7)$$

The synthesis filters $P_{m,n-m}(x, y)$ of order m in x , and $n - m$ in y , are defined by

$$P_{m,n-m}(x, y) = \frac{G_{m,n-m}(x, y) V(x, y)}{W(x, y)} \quad (8)$$

for $m = 0, \dots, n$ and $n = 0, \dots, \infty$.

It is important to mention that the HT can generate coefficients with and without subsampling. In practice, HT implementation requires the choice of the size that Gaussian window spreads (σ) and a subsampling factor (if used) that defines the sampling lattice S . The resultant Hermite coefficients are accommodated as a set of $N \times N$ equal-sized subbands, as shown in Figure 1 (figure reproduced from S. Gomez-Coronel et al. (2016) [under the Creative Commons Attribution License/public domain]).

In a discrete implementation, the Gaussian window function may be approximated by the binomial window function:

$$\omega^2(x) = \frac{1}{2^N} C_N^x, \quad x = 0, \dots, N, \quad (9)$$

where N is called the order of the binomial window and represents the function length, and

$$C_N^x = \frac{N!}{(N-x)!} x!, \quad x = 0, \dots, N. \quad (10)$$

In this case, the orthogonal polynomials $G_{m,n-m}(x, y)$ associated with binomial window are known as Krawtchouk's polynomials:

$$K_n[x] = \frac{1}{\sqrt{C_N^n}} \sum_{k=0}^n (-1)^{n-k} C_{N-x}^{n-k} C_x^k. \quad (11)$$

For this discrete case, all previous relations hold, with some interesting modifications. First, the window function support is finite (N); as a consequence, expansion with the Krawtchouk polynomials is also finite, and signal reconstruction from the expansion coefficients is perfect. In practice, a Hermite Transform implementation requires the choice of some parameters, that is, the size of the Gaussian window spread (σ), or alternatively, the order N for binomial windows, and the subsampling factor that defines the sampling lattice S . Resulting Hermite coefficients are arranged as a set of $(N \times N)$ equal-sized subbands: one coarse subband $L_{0,0}$ representing a Gaussian-weighted image average and detail subbands $L_{n,m}$ corresponding to higher-order Hermite coefficients, as shown in Figure 1 (figure reproduced from S. Gomez-Coronel et al. (2016) [under the Creative Commons Attribution License/public domain]).

The coefficients of the first diagonal (order 1) are related to convolution with first-order derivatives of a Gaussian in orthogonal directions. These coefficients are useful in detecting edges in the image. In the same way, the coefficients in the following diagonal (order 2) are related to a convolution with second-order derivatives of a Gaussian and so on.

3. Luminance-Brightness Map

Human vision is sensitive to an enormous range of luminance, but the corresponding range of brightness is much lower. As a consequence, and at first sight, the mapping between luminance and brightness can be modeled by a logarithmic compression. Nevertheless, many brightness-related illusions cannot be explained by such a simple model. Brightness perception is a complex psychophysical phenomenon that is influenced by receptive fields, multiresolution channels, and contrast sensitivity among other HVS properties. One of the more thorough brightness models was introduced by Schouten [17]. Schouten's model assumes that brightness is invariant to light source properties and to observation conditions. For the construction of the luminance-brightness map, Schouten divides the algorithm in three stages:

- (1) Scale representation.
- (2) Assembling the scaling signals.
- (3) Local adjustment of the brightness scale.

3.1. Discrete Algorithm of the Luminance-Brightness Map for Images. As preprocessing, the images to be used must be surrounded by a uniform region with a value of constant luminance L_0 , which will be the average value of the image. To avoid unwished variations, it is necessary to normalize the images so that the pixel intensities are in the interval $[0, 1]$. To achieve the first stage of multistage representation, it is indispensable to carry out a sampling with exponentially increasing distances. That is to say, that the expression

$$A(\vec{x}) = \int_{\sigma^-}^{\sigma^+} h_A(\vec{x}, \sigma) d\sigma \quad (12)$$

is a Riemann's sum of terms $h_A(x, y, \sigma_i)$, which are taken in equidistant positions of the scale parameter s .

Since the variations of deployed luminance only happen in an area limited by a homogeneous region, variations can be captured using a controlled number of scales. In our case, we used 512×512 pixel images, and, based on the results of Section 6, we chose to use nine scales. Index $i \in 1, 2, 3, \dots, 9$ denotes the scale, and the scaled signal with index 1 $h_A^{(1)}(x, y)$ represents the finest scale, whereas index 9 corresponds to the coarsest scale.

The central and peripheral answers $V_c(x, y; s_i)$ and $V_s(x, y; s_i)$ are obtained by convolution between the image and the filters that shape the receptive fields. The assembling map $A(x, y)$, of (13), is calculated by adding, point to point, the scaled signals $h_A^{(i)}$ and an offset term A_G , as expressed in (14):

$$A(x, y) = A_G + \ln 2 \sum_{i=1}^9 h_A^{(i)}(x, y) \quad (13)$$

with

$$A_G \approx \beta (L_0 - \delta) \widetilde{A}_G, \quad (14)$$

where $\beta = 0.1$, $\delta = -5.0$, and $A_G = 1.22$ according to [17].

The minimal and maximal values $\widehat{A}_{\min}(x, y)$ and $\widehat{A}_{\max}(x, y)$ are calculated through (15) and (16), respectively:

$$\widehat{A}_{\min}(x, y) = \min \left[\left[\ln 2 \sum_{j=0}^k h_A^{9-j}(x, y) \right] + A_G \right], \quad (15)$$

$$k = 0, 1, \dots, 8$$

$$\widehat{A}_{\max}(x, y) = \max \left[\left[\ln 2 \sum_{j=0}^k h_A^{9-j}(x, y) \right] + A_G \right], \quad (16)$$

$$k = 0, 1, \dots, 8.$$

Finally, the brightness map is obtained after applying the indentation of the brightness with $\lambda = 2$.

3.2. Perceptive Mask. As it has been mentioned before, we aim at constructing a watermarking algorithm that takes advantage of a perceptive mask; which, according to Legge and Foley [24], alludes to the fact that the presence of a signal can conceal the presence of another. The purpose is

to be able to identify the regions in which the detection of the incruited mark becomes difficult. In order to create the embedded mask, the ideas presented in [25] were taken into account, as well as the luminance-brightness map described earlier. The following steps are taken for this purpose:

- (1) Calculate the Hermite Transform coefficients of order n and subsampling factor T of the original image.
- (2) Obtain a smoothed version of the image and histogram by applying a Gaussian filter.
- (3) Calculate the brightness map B of the original image and scale its values taking the lower and upper levels of the histogram of the smoothed image generated in the previous step. The representation of brightness allows the identification of the regions that are of interest, while enabling the generation of the perceptive mask, since it might be modified without any perceptible change or kept unmodified to prevent the recognition of changes related to the brightness perception of the image objects.
- (4) Calculate the contrast with

$$C = \left[\sum_{i=1}^m \sum_{j=1}^{n-m} C_{i,j}^2 \right]^{1/2}, \quad (17)$$

where $C_{i,j}$ represents the Hermite Transform Cartesian Coefficients.

- (5) Calculate the light adaptation threshold, as indicated in

$$C_{\text{thr}} = k_0 \left(C_{\min} + \left| \frac{B^\alpha - L_{\min}^\alpha}{B^\alpha + L_{\min}^\alpha} \right|^{1/\alpha} \right), \quad (18)$$

where k_0 is a constant and $C_{\min}$ is the minimal contrast where a minimal luminance level $L_{\min}$ is found. According to [25] these locations correspond to the maximum contrast sensitivity. And α is a constant $[0, 1]$.

- (6) The perceptive mask M [24] is generated in agreement to the following expression:

$$M = k_1 \max(C_{\text{thr}}, C^\beta C_{\text{thr}}^{1-\beta}), \quad (19)$$

where k_1 is a constant.

4. Watermarking Scheme with the Hermite Transform

The following steps describe how to insert and detect the watermark in an image L of size $(X \times Y)$.

4.1. Watermark Embedding

- (1) Calculate the Hermite Transform coefficients of the original image L .
- (2) Generate the watermark W (pseudorandom binary sequence values $\{-1, 1\}$).

- (3) In order to embed the watermark in the Hermite coefficients, (20) shall be used:

$$\hat{L}_{n,m}(i, j) = L_{n,m} + \gamma M(i, j) W_{n,m}(i, j), \quad (20)$$

where $L_{n,m}$ represents the original Hermite coefficients, $\hat{L}_{n,m}$ represents the embedded coefficients with the watermark, γ is a strength control parameter, and M is the perceptive mask used to adapt the level of watermark strength and invisibility according to the local characteristics of the image. The subscripts n, m indicate which subband of Hermite coefficients $L_{n,m}$ is being marked with $W_{n,m}$.

- (4) Finally, the inverse Hermite Transform is calculated to obtain the marked image.

4.2. Watermark Detection. For the purpose of watermark detection, we adopt the method of comparing the average value $\bar{R}$ of the correlation R , with a threshold value of T . If $\bar{R} > T$, then the mark is present.

$$\begin{aligned} \bar{R} &= \frac{1}{K \cdot f_H \cdot c_H} \sum_n \sum_m R(n, m) \\ R(n, m) &= \begin{cases} \sum_{i=1}^{f_H} \sum_{j=1}^{c_H} \hat{L}_{n,m}(i, j) W_{n,m}(i, j) & \text{if } \hat{L}_{n,m} \text{ is marked} \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (21)$$

f_H, c_H are coefficients dimensions and K are marked Hermite coefficients.

The threshold value T is determined by (22), in compliance with Neyman-Pearson's criterion:

$$T = 3.97\sqrt{2\sigma^2}, \quad (22)$$

where

$$\sigma^2 = \frac{1}{(K \cdot f_H \cdot c_H)^2} \sum_n \sum_m \sum_{i=1}^{f_H} \sum_{j=1}^{c_H} [\hat{L}_{n,m}(i, j)]^2. \quad (23)$$

5. Image Normalization

A normalization step can assure the integrity of the mark, even if the image is submitted to geometric attacks. If an affine transformation is applied to the image, it results in a geometric distortion (translation, rotation, or scaling). This is called a geometric attack and may seriously impair the watermark detection performance. A solution is the use of central geometric moments, as proposed in different researches [18, 26]. In this work, we suggest achieving robustness to geometric attacks by normalizing the image before embedding the mark [12]. We aim at obtaining the necessary parameters to create a normalized template with intensity values up to 255. This template helps the recovered image (after inverse normalization) to preserve quality in

spite of transformations and interpolations suffered during the geometric attack.

Normalization is achieved according to the following steps:

- (1) Determine the normalization parameters of the original image L .
- (2) Generate an image P of the same size of L , with uniform intensity values equal to 255.
- (3) With the parameters obtained in step (1), obtain the normalized image P_{norma} from the image P .
- (4) Calculate the perceptive mask M , from the original image L , and normalize it using the same parameters obtained in step (1).
- (5) Generate the watermark W , as describe before.
- (6) Generate a set of empty Hermite coefficients Y .
- (7) Insert the mark in the coefficients of Y , according to

$$\begin{aligned} \hat{Y}_{n,m}(i, j) &= Y_{n,m}(i, j) \\ &+ \gamma M_{\text{Norma}}(i, j) W_{n,m}(i, j) P_{\text{Norma}}(i, j), \end{aligned} \quad (24)$$

where $Y_{n,m}$ are original coefficients, γ is the strength control parameter, M_{Norma} is a normalized perceptive mask, P_{Norma} is a normalize template, and $\hat{Y}_{n,m}$ are the modified coefficients.

- (8) Calculate the inverse Hermite Transform of the coefficient set $\hat{Y}$, to obtain the corresponding normalized watermark W_{Norma} .
- (9) Apply the inverse normalization to W_{Norma} so that the watermark $\hat{W}$ has the same dimensions and orientation as those of the original image L .
- (10) Generate the marked image through

$$\hat{L} = L + \hat{W}. \quad (25)$$

6. Test and Experimental Results

6.1. Parameter Settings. It is important to determine the values of the different parameters that constitute the perceptive mask, since significant results can be obtained in invisibility and robustness. Diverse experimental tests (with 200 different watermarks) were executed. After executing the algorithm, it was determined that the best combination of parameters was $\alpha = 0.649$, $\beta = 0.8$, $k_0 = 0.08$, and $k_1 = 0.8$. In order to evaluate the algorithm performance, objective as well as subjective tests were carried out using 14 different images.

6.2. Objective Evaluation. We considered the following metrics:

- (1) PSNR (Peak Signal to Noise Ratio in dB), of an image L_n of size $(X \times Y)$, is given by

PSNR

$$= 10 \log_{10} \left(\frac{255^2}{(1/XY) \sum_{X,Y} (L_1(i, j) - L_2(i, j))^2} \right). \quad (26)$$

TABLE 1: Averages values obtained by embedding 200 different watermarks in different images.

Image	PSNR [dB]	MSSIM
Aerial	40.6612	0.9846
Airplane	40.1903	0.9891
Baboon	36.9731	0.9881
Barbara	38.6131	0.9872
Boat	40.2455	0.9864
Cameraman	41.3585	0.9832
Clock	42.6793	0.9816
Elaine	43.0007	0.9872
Gold hill	41.8065	0.9877
Lake	39.6577	0.9869
Lena	42.3998	0.9883
Peppers	41.9539	0.9876
Stream bridge	38.1478	0.9877
Texture	40.8864	0.9893

(2) MSSIM (Mean Structure Similarity Index) is given by

$$\text{MSSIM}(L, N) = \frac{1}{M} \sum_{j=1}^M \text{SSIM}(l_j, n_j), \quad (27)$$

where L, N are original and distorted images respectively, l_j, n_j are the images contents at the j th local window, and M is the number of local windows of the image. If l, n are two images with nonnegative values, SSIM is given by

$$\text{SSIM}(l, n) = \frac{(2\mu_l\mu_n + C_1)(2\sigma_{ln} + C_2)}{(\mu_l^2 + \mu_n^2 + C_1)(\sigma_l^2 + \sigma_n^2 + C_2)}, \quad (28)$$

where μ_l and μ_n are their respective averages and σ_l , σ_n , and σ_{ln} are the standard deviations and covariance, respectively. C_1, C_2 are constants to avoid instability when the denominator is close to zero [27].

Table 1 shows results of the above metrics for 14 different images averaged over 200 watermarks, using the best combination of parameters.

As shown in Table 1, all 14 images always maintained a PSNR average above 38 dB. The MSSIM was always in the range of the value of 0.98. These values show that the watermarked image maintains objective quality.

6.3. Subjective Assessment. We measured perceptive image quality through subjective tests. However, since there is no standard for watermarking techniques, we used Double Stimulus Impairment Scale (DSIS) protocol, which includes some categories and scale (Table 2). This protocol indicates that both the original image and the marked one must be shown, explicitly, to the observers, so that they can evaluate the latter. Different experiment protocols were carried out, but the DSIS protocol was always considered. Three types of evaluations were carried out.

TABLE 2: DSIS categories and scales.

Categories	Value
Imperceptible	5
Perceptible but not annoying	4
Slightly annoying	3
Annoying	2
Very annoying	1

TABLE 3: MOS averaged for all 14 tested images obtained in experiment 1.

Image	MOS	
	Original image	Watermarked image
Aerial	4.4250	3.7433
Airplane	4.0550	4.0400
Baboon	3.3000	3.9400
Barbara	3.6166	3.6550
Boat	3.9166	3.8000
Cameraman	3.6933	3.2966
Clock	4.2433	3.8150
Elaine	4.4500	3.9333
Gold hill	3.6600	3.4433
Lake	3.9866	3.2766
Lena	4.0733	3.9533
Peppers	4.0466	4.1200
Stream bridge	3.9866	3.9000
Texture	4.0700	3.9166

(a) The first experiment was divided into sessions of 15 minutes with 35 observers. In this case we did not let the observer know which of the two images was watermarked. We watermarked fourteen different images using the proposed algorithm. Both the original and the watermarked images were displayed side by side and the observers were asked to judge their quality. For each image an average score given by observes was calculated, thus obtaining the Mean Opinion Score (MOS) for the image under analysis. Table 3 shows these results.

Table 3 results show the MOS values obtained for images. It becomes clear that, since the observers did not know which one was the marked image, both show similar MOS values: in some cases, the observers considered that the original image had lower quality in comparison with the watermarked one. These results confirm that the technique presented in this work produces watermarked images that are difficult to distinguish from the original image.

(b) In the second experiment, each one of the 14 original (O) images and its corresponding watermarked (W) image were organized in pairs in four groups: O-W, W-O, O-O, and W-W, making a total of 56 pairs of images. Each pair was presented in random order three times to the observers. They were asked to choose the image with the lower quality and to scale it according to the DSIS protocol. Average scores were computed for each image. Table 4 shows the average score obtained per group for five observers.

TABLE 4: MOS averaged over all tested images obtained in experiment 2.

Observer	Group							
	1	2	3	4	1	2	3	4
	O	W	O	O	W	W	W	O
1	3.164	3.171	3.187	3.042	3.506	3.199	3.258	3.222
2	3.860	3.900	4.090	4.210	3.490	3.530	3.590	3.910
3	4.280	4.021	4.240	4.694	4.330	4.571	3.840	4.400
4	4.000	3.633	4.000	3.519	4.000	3.473	4.000	3.579
5	4.783	4.810	4.740	4.667	4.910	5.000	4.818	4.778

TABLE 5: MOS averaged over all tested images obtained in experiment 3.

Observer	Group							
	1	2	3	4	1	2	3	4
	O	W	O	O	W	W	W	O
1	0.5714	0.2619	0.0714	1.1428				
2	0.6190	1.0000	0.6904	1.2857				
3	0.1190	0.0238	0.1428	0.0238				
4	0.3571	0.0238	0.0000	0.2619				

According to the average scores, we can conclude that the observers could not distinguish effectively between the marked image and the original one. Moreover, in all cases the scores given either to the original or the marked images are above the value of 3, which, according to the *DSIS* protocol, indicate a good perceived image quality. For example, observer 5 got averages above 4, indistinctly for original and marked images. Observer 1 granted the lowest scores whose averages oscillated between 3.164 and 3.506.

(c) In a third experiment, with a similar stimulus setting of the previous experiment, the observers were asked to choose the image that presented the best quality; then they were asked to rate the difference of perceived quality between the two images in a numerical scale, 5 being the score for the largest quality difference and 1 being the one for the smallest difference. If both images presented the same perceived quality a 0 score was given. Table 5 shows the MOS values obtained per group for four observers. The results of this test showed once more that observers could not tell the difference between the marked image and the original one.

6.4. Embedding and Detection. Once we have got the parameter values involved in the mask generation, we may proceed to the embedding and detection of the watermark. The original image and the marked one are shown in Figure 2, whereas the mask and the difference between the original image and the marked image can be observed in Figure 3.

Table 6 shows the resulting values of the metrics used to test the insertion efficiency of the algorithm.

Subsequently, the detection process was applied to the marked images. It was then repeated for 200 different marks. In Figure 4 we can observe that only the embedded mark overcame the detection threshold. We only show results for Lena, Barbara, and Baboon images, but similar results were obtained with the other 11 images.

TABLE 6: Results obtained when inserting a watermark to the Lena image.

Image	PSNR [dB]	MSSIM
Lena	42.7869	0.98922
Barbara	38.9573	0.98815
Baboon	37.3475	0.98899

6.5. Robustness Evaluation. It is important to evaluate the performance algorithm in terms of its robustness to different processing operations, that is, processing attacks that are meant to extract and/or distort the embedded mark. In order to evaluate this performance, two types of tests were carried out. First, processing operations were applied to marked images using our perceptive algorithm, and a comparison was made with other works to evaluate its efficiency. Second, geometric attacks were executed on normalized images and a similar evaluation was carried out.

Processing Operations. Gaussian and median filters were applied. In the case of the Gaussian filter, the watermark in Lena image was detected up to a filter size of 5 pixels; whereas, with the median filter, the detection was successful up to a filter size of 4 pixels. Results can be observed in Figure 5. *JPEG* compression was also tested for several quality factors ranging from 0 to 100% with increments of 5. Detection was achieved with quality factors of 5% and above. Likewise, image cropping was tested, replacing eliminated regions with black pixels. The watermark was successfully detected with an image variation of up to 90%. Figure 6 presents these results. Finally, Gaussian and salt and pepper noise were added. Gaussian noise variance varied from 0 to 0.5 in steps of 0.05. The watermark was detected up to a variance of 0.25. Regarding the salt and pepper noise, its density was modified

FIGURE 2: Original image: (a) Lena, (c) Barbara, and (e) Baboon. Watermarked image: (b) Lena, (d) Barbara, and (f) Baboon.

and added to the image from 0 to 1 in steps of 0.1. The mark was detected up to a 0.3 noise density. Results are shown in Figure 7.

For the case of Barbara and Baboon images, robustness results were also satisfactory. In both images detection under Gaussian filtering, *JPEG* compression, and Gaussian noise

attacks was successful. With median filter and image cropping attacks, the results were as good as with the Lena image. Finally, with salt and pepper noise the watermark was detected in Barbara image up to a 0.5 noise density and with Baboon image up to 0.6 noise density. Figures 8–13 show the results. These evaluations serve as reference to compare the

FIGURE 3: Original Mask ($\times 8$): (a) Lena, (c) Barbara, and (e) Baboon, and difference between the original image and the marked image ($\times 16$): (b) Lena, (d) Barbara, and (f) Baboon.

proposed scheme with other techniques that use the masking concept. In order to have a baseline between them, an adjustment was made in the embedded mark strength, thus achieving the same value of *MSSIM* among the algorithms.

Geometric Attacks. The geometric attacks that were tested are

- (1) Scaling (factor of 20% to 200%).
- (2) Rotation (counterclockwise rotation, from 0° to 180°).

FIGURE 4: Values of threshold and correlation calculated for the marked images: (a) Lena, (b) Barbara, and (c) Baboon.

- (3) Deformation (shearing in horizontal and vertical orientations).

Figures 14–19 show the obtained results using Lena, Barbara, and Baboon images.

It is clear that, after rotation and shearing attacks in Lena image, the detection of the mark is achieved satisfactorily in all the cases. In the case of scaling, the detection is achieved in 13 out of 16 cases. In the case of the Barbara image, the best detection was obtained with shearing in horizontal orientation. Finally, when we used Baboon image the detection of the mark was achieved in almost all cases after scaling and shearing attacks in horizontal direction. After vertical

shearing, the detection of the mark is successfully achieved only in the first six shearing factors.

6.6. Comparison with Other Watermark Algorithms. Comparing algorithms is necessary to identify common elements; on the contrary it is not a valid comparison. There are many different watermarking algorithms, and differences are about their application, their robustness, type of watermarking, and so on. To validate the algorithm proposed in this paper we consider different recent and old watermarking algorithms.

First of all, we make a comparison with similar algorithms that use perceptive mask. For example, Barni et al. [10] proposed a wavelet transform-based, watermarking method

FIGURE 5: (a) Correlation values and threshold obtained after filtering the watermarked Lena image with a Gaussian filter. (b) Correlation values and threshold obtained after filtering the watermarked Lena image with a median filter.

FIGURE 6: (a) Correlation values and threshold obtained after compressing the watermarked Lena image using JPEG compression. (b) Correlation values and threshold obtained after cropping the watermarked Lena image.

that reports the *PSNR* and *MSSIM* values showed in Table 7, whereas Figure 20 shows the mask and the difference between the original and the watermarked image with this method.

Comparing our results of Table 6 (image Lena) with Barni et al.'s results of Table 7 [10], we note that both methods

show the same value of *MSSIM*, while the *PSNR* value results are better with method [10]. Nevertheless, our method shows better performance regarding processing attacks. For example, the algorithm proposed in [10] only detects the watermark with a size 7 median filter, while our method

FIGURE 7: (a) Correlation values and threshold obtained after adding white Gaussian noise to the watermarked Lena image. (b) Correlation values and threshold obtained after adding salt and pepper to the watermarked Lena image.

FIGURE 8: (a) Correlation values and threshold obtained after filtering the watermarked Barbara image with a Gaussian filter. (b) Correlation values and threshold obtained after filtering the watermarked Barbara image with a median filter.

TABLE 7: Results obtained by embedding a watermark to the Lena image using the scheme proposed in [10].

PSNR [dB]	MSSIM
45.4017	0.98927

TABLE 8: Results obtained by embedding a watermark to the Lena image using the scheme proposed in [16].

PSNR [dB]	MSSIM
41.8394	0.98954

performs well with a size 4 median filter. In the case of Gaussian filter attacks, method in [10] can only recover the mark with filter sizes 1 and 2, while our method recovers the mark with filter sizes up to 4.

The second technique that was compared is described in [16]. Table 8 shows the results obtained with this algorithm; Figure 21 shows the mask and the difference between the original and the marked image.

FIGURE 9: (a) Correlation values and threshold obtained after compressing the watermarked Barbara image using JPEG compression. (b) Correlation values and threshold obtained after cropping the watermarked Barbara image.

FIGURE 10: (a) Correlation values and threshold obtained after adding white Gaussian noise to the watermarked Barbara image. (b) Correlation values and threshold obtained after adding salt and pepper to the watermarked Barbara image.

FIGURE 11: (a) Correlation values and threshold obtained after filtering the watermarked Baboon image with a Gaussian filter. (b) Correlation values and threshold obtained after filtering the watermarked Baboon image with a median filter.

FIGURE 12: (a) Correlation values and threshold obtained after compressing the watermarked Baboon image using JPEG compression. (b) Correlation values and threshold obtained after cropping the watermarked Baboon image.

FIGURE 13: (a) Correlation values and threshold obtained after adding white Gaussian noise to the watermarked Baboon image. (b) Correlation values and threshold obtained after adding salt and pepper to the watermarked Baboon image.

FIGURE 14: (a) Correlation values and threshold obtained after scaling the watermarked Lena image. (b) Correlation values and threshold obtained after rotating the watermarked Lena image.

FIGURE 15: (a) Correlation values and threshold obtained after applying the operation shearing in horizontal direction to the watermarked Lena image; (b) correlation values and threshold obtained after applying the operation shearing in vertical direction to the watermarked Lena image.

FIGURE 16: (a) Correlation values and threshold obtained after scaling the watermarked Barbara image. (b) Correlation values and threshold obtained after rotating the watermarked Barbara image.

FIGURE 17: (a) Correlation values and threshold obtained after applying the operation shearing in horizontal direction to the watermarked Barbara image; (b) correlation values and threshold obtained after applying the operation shearing in vertical direction to the watermarked Barbara image.

FIGURE 18: (a) Correlation values and threshold obtained after scaling the watermarked Baboon image. (b) Correlation values and threshold obtained after rotating the watermarked Baboon image.

FIGURE 19: (a) Correlation values and threshold obtained after applying the operation shearing in horizontal direction to the watermarked Baboon image; (b) correlation values and threshold obtained after applying the operation shearing in vertical direction to the watermarked Baboon image.

FIGURE 20: (a) Mask used in method [10] (×4). (b) Difference between the original and the watermarked image (×16).

FIGURE 21: (a) Mask ($\times 8$) (method [16]). (b) Difference between the original image and the marked image ($\times 16$).

It is clear that the *PSNR* value of [16] is lower than the one obtained through the technique of [10] (Table 7) and the one suggested here (Table 6). Nevertheless, the robustness to processing attacks of [16] is very similar to the one obtained with the algorithm described in this work.

Finally, the use of Contourlet transform was proposed in [15], using some ideas of Barni algorithm. The watermarked image has a *PSNR* = 36.65 dB, which is lower than in our method. Moreover, the author of [15] reports that his method could detect the watermark when attacked by different types of processing, such as median filtered (3×3 and 5×5); filtered Gaussian (3×3 and 5×5); *JPEG* compression with quality factors 75, 50, and 5; Gaussian; and salt and pepper noise, as well as cropping. It is clear that our method outperforms these results.

On the other hand, there are a lot of different watermark algorithms that use other techniques. For example, in [28] a watermark algorithm using Contourlet transform and DCT coefficients was presented. Also they propose using a novel edge detection to detect important edges (Canny), and to maintain the stronger of edges they apply Otsu thresholding. With this, they obtained a perceptive mask, although they highlight the edges. To probe their results they use 5 different typical images, including *Barbara* and *Baboon*. As watermark they use 128 bits of message. In *Barbara* image they report a *PSNR* = 42.89 dB, *SSIM* = 0.9971 and *Baboon* image *PSNR* = 40.06 dB, *SSIM* = 0.9960. To measure the robustness they only use normalized correlation (NC) between the original and the extracted watermark, and they use different attacks: cropping, Gaussian filter, and *JPEG*. Comparing their results with ours, in case of *JPEG compression* our method is better because detection was achieved with all quality factors (Figures 9 and 12) and in technique [28] they report *BER* = 5.47 using *Barbara* image and *Baboon* image using *JPEG* attack 10%. This result shows that the extracted watermark is not the same as the original. With other attacks, results are similar. As can be observed in Figures 8(a), 9(b), 11(a), and 12(b), in both images detection under Gaussian filter attack and cropping attacks was successful. While in technique [28]

they report *NC* = 0.9696 average (watermark images were cropped from 5 to 20%) using *Barbara* image. In case of *Baboon* image they report *NC* = 0.9876 average.

The case of medical application of watermark algorithm is difficult to compare because they use other images and watermarks. In [29] a medical image watermark technique using wavelet transform and applying spread-spectrum concept was proposed. They use a binary image as watermark (50×9). For the extraction of the watermark, statistical profile of DWT coefficients of watermarked image is determined and the obtained probability distribution function (pdf) is utilized for designing the watermark detection procedure. They report as maximum *PSNR* 43.998 dB and minimum *PSNR* value is 33.198 dB. Maximum and minimum correlation (NC) values obtained are 0.9508–0.6871, respectively. As we can observe, these *PSNR* values demonstrate that the watermark image could present variations in reference to original; that is, the original image is clearly modified to human eye. Correlation values indicated that the watermark extracted is different than the original and this is not good for this kind of applications, because a different watermark does not allow identifying the owner. Robustness of the algorithm was probed using different attacks: *JPEG compression*, Gaussian noise, and salt and pepper noise. Results are as follows: *NC* = 0.4066 for *JPEG* (QF = 10), *NC* = 1 for Gaussian attack (mean = 0, var = -0.01), and *NC* = 0.9260 for Gaussian attack (mean = 0, var = -0.05). If we compare these results with ours, it is clear that the technique proposed in this paper is better (Sections 6.4 and 6.5), even when the original images and watermarks are not the same.

7. Conclusions

In this article we present a perceptive approach to digital image watermarking using a brightness model, image normalization, and the Hermite Transform. We use a perceptive model that takes advantage of the masking characteristics of the Human Vision System to build a brightness map. As a result, we show that the presented technique allows the

insertion of a watermark that cannot be easily perceived by a human observer. Previously reported watermark techniques found in the literature, such as [10], use a mask that indicates the locations of the image where additional information can be inserted without producing too much perceptive distortion. However, the human perception elements used to build this map are very limited. Our proposal includes the use of a brightness map and allows working with a perceptive scale instead of luminance values. Furthermore, we also include the contrast sensitivity function, which allows estimating perceived contrasts rather than luminance changes. The brightness map and perceived contrasts allow building a more sophisticated perceptive mask that proves to be a more reliable guide to finding locations of the image where a watermark can be inserted with more confidence. The advantage of the use of the Hermite Transform is that it includes basis functions that shape the profiles of the receptive fields in the Human Vision System, which means that the transform coefficients contain image primitives that represent perceptive relevant structures such as edges, lines, and contours. Objective metrics, namely, *PSNR* and *MSSIM*, show that marked images present very little distortion and three different subjective assessments confirm that marked images are perceptually undistinguishable from the originals. It was verified that the technique is robust to different processing attacks, outperforming several methods reported in the literature such as [10, 15, 16]. Our results with geometric attacks also show that the watermark resisted most of them, the worst case being Lena image under a scaling attack, with only 13 out of 16 satisfactory detection instances. Comparison with other methods demonstrated that using Hermite Transform is a good way to improve watermark algorithms, as we can review in [12, 30].

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors acknowledge Universidad Nacional Autónoma de México (Grant PAPIIT IN116917) and Instituto Politécnico Nacional (IPN).

References

- [1] M. C. Hernández, M. N. Miyakate, and H. P. Meana, *Técnica Robusta De Marca De Agua Basada En Normalización De Imágenes*, vol. 52, Facultad de Ingeniería Antioquia University, 2010.
- [2] J. J. K. O. Ruanaidh, W. J. Dowling, and F. M. Boland, "Watermarking digital images for copyright protection," *IEEE Proceedings Vision, Image and Signal Processing*, vol. 143, pp. 250–256, 1996.
- [3] M. Boreiry and M.-R. Keyvanpour, "Classification of watermarking methods based on watermarking approaches," in *Proceedings of the 7th Conference on Artificial Intelligence and Robotics, IRANOPEN 2017*, pp. 73–76, Iran, 2017.
- [4] M. Urvoy, D. Goudia, and F. Atrusseau, "Perceptual DFT watermarking with improved detection and robustness to geometrical distortions," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 7, pp. 1108–1119, 2014.
- [5] A. K. Singh, M. Dave, and A. Mohan, "Wavelet Based Image Watermarking: Futuristic Concepts in Information Security," *Proceedings of the National Academy of Sciences India Section A - Physical Sciences*, vol. 84, no. 3, pp. 345–359, 2014, <https://doi.org/10.1007/s40010-014-0140-x>.
- [6] N. Baaziz, "Adaptive watermarking schemes based on a redundant contourlet transform," in *Proceedings of the IEEE International Conference on Image Processing 2005, ICIP 2005*, pp. 221–224, Italy, September 2005.
- [7] T.-T. Bai, Z. Liu, and P. Lu, "Digital watermarking scheme in contourlet domain based on qr code," *Journal of Optoelectronics*, vol. 21, no. 4, pp. 76–79, 2014.
- [8] H. Zhou, C. Qi, and X. Gao, "Low luminance smooth blocks based watermarking scheme in DCT domain," in *Proceedings of the 2006 International Conference on Communications, Circuits and Systems, ICCAS*, pp. 19–23, China, June 2006.
- [9] R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, "Perceptual watermarks for digital images and video," *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1108–1126, 1999.
- [10] M. Barni, F. Bartolini, and A. Piva, "Improved wavelet-based watermarking through pixel-wise masking," *IEEE Transactions on Image Processing*, vol. 10, no. 5, pp. 783–791, 2001.
- [11] M. Mundher, D. Muhamad, A. Rehman, T. Saba, and F. Kausar, "Digital watermarking for images security using discrete slantlet transform," *Applied Mathematics & Information Sciences*, vol. 8, no. 6, pp. 2823–2830, 2014.
- [12] S. L. G. Coronel, B. E. Ramírez, and M. A. A. Mosqueda, "Robust watermark technique using masking and Hermite transform," *SpringerPlus*, vol. 5, no. 1, article no. 1830, 2016.
- [13] P.-B. Nguyen, A. Beghdadi, and M. Luong, "Perceptual watermarking using pyramidal JND maps," in *Proceedings of the 10th IEEE International Symposium on Multimedia, ISM 2008*, pp. 418–423, Berkeley, Calif, USA, December 2008.
- [14] P. B. Nguyen, A. Beghdadi, and M. Luong, "Perceptual watermarking using a new Just-Noticeable-Difference model," *Signal Processing: Image Communication*, vol. 28, no. 10, pp. 1506–1525, 2013, <http://dx.doi.org/10.1016/j.image.2013.09.011>.
- [15] H. Song, S. Yu, X. Yang, L. Song, and C. Wang, "Contourlet-based image adaptive watermarking," *Signal Processing: Image Communication*, vol. 23, no. 3, pp. 162–178, 2008, <http://dx.doi.org/10.1016/j.image.2008.01.005>.
- [16] N. Baaziz, B. Escalante-Ramirez, and O. Romero-Hernández, "Image watermarking in the Hermite transform domain with resistance to geometric distortions," in *Proceedings of the Optical and Digital Image Processing*, Strasbourg, France, April 2008.
- [17] G. G. Schouten, *Luminance-Brightness Mapping: The Missing Decades*, Technische Universiteit Eindhoven, 1993, <https://doi.org/10.6100/ir394966>.
- [18] P. Dong, J. G. Brankov, N. P. Galatsanos, Y. Yang, and F. Divoine, "Digital watermarking robust to geometric distortions," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2140–2150, 2005.
- [19] J.-B. Martens, "The Hermite Transform—Theory," *IEEE Transactions on Signal Processing*, vol. 38, no. 9, pp. 1595–1606, 1990.
- [20] J.-B. Martens, "The Hermite Transform—Applications," *IEEE Transactions on Signal Processing*, vol. 38, no. 9, pp. 1607–1618, 1990.
- [21] J. L. Silván-Cárdenas and B. Escalante-Ramírez, "The multiscale hermite transform for local orientation analysis," *IEEE Transactions on Image Processing*, vol. 15, no. 5, pp. 1236–1253, 2006.

- [22] B. Escalante-Ramírez and J. L. Silván-Cárdenas, "Advanced modeling of visual information processing: A multi-resolution directional-oriented image transform based on Gaussian derivatives," *Signal Processing: Image Communication*, vol. 20, no. 9-10, pp. 801–812, 2005, <http://dx.doi.org/10.1016/j.image.2005.05.009>.
- [23] R. A. Young, R. M. Lesperance, and W. W. Meyer, "The Gaussian derivative model for spatial-temporal vision: I. Cortical model," *Spatial vision*, vol. 14, no. 3-4, pp. 261–319, 2001.
- [24] G. E. Legge and J. M. Foley, "Contrast masking in human vision," *Journal of the Optical Society of America*, vol. 70, no. 12, pp. 1458–1471, 1980.
- [25] B. Escalante-Ramírez, P. López-Quiroz, and J. L. Silván-Cárdenas, "SAR-image classification with a directional-oriented discrete hermite transform and markov random fields," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, vol. 6, pp. 3423–3425, July 2003.
- [26] J. Nah and J. Kim, "Digital watermarking robust to geometric distortions," *Communications in Computer and Information Science*, vol. 342, pp. 55–62, 2012.
- [27] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [28] H. R. Fazlali, S. Samavi, N. Karimi, and S. Shirani, "Adaptive blind image watermarking using edge pixel concentration," *Multimedia Tools and Applications*, vol. 76, no. 2, pp. 3105–3120, 2017, <https://doi.org/10.1007/s11042-015-3200-6>.
- [29] D. S. Chauhan, A. K. Singh, A. Adarsh, B. Kumar, and J. P. Saini, "Combining Mexican hat wavelet and spread spectrum for adaptive watermarking and its statistical detection using medical images," *Multimedia Tools and Applications*, pp. 1–15, 2017, <https://doi.org/10.1007/s11042-017-5348-8>.
- [30] S. L. G. Coronel, E. M. Albor, B. E. Ramirez, and J. Brieva, "Watermarked cardiac ct image segmentation using deformable models and the hermite transform," in *Proceedings of the 10th International Symposium on Medical Information Processing and Analysis*, vol. 9287, pp. 9287–9281, 2015, <http://dx.doi.org/10.1117/12.doi:10.1117/12>.