VNU Journal of Science: Comp. Science & Com. Eng., Vol. 32, No. 2 (2016) 49-62

A Watermark Algorithm Against De-Synchronization Attacks☆

Luong Viet Nguyen*, Trinh Nhat Tien, Ho Van Canh

VNU University of Engineering and Technology, Hanoi, Vietnam

Abstract

In this paper, a robust method to the ability of the watermark to resist against attacks is proposed for hiding information into images. The proposed method is blind because the original image is not required at the decoder to recover the embedded data. The robustness of the watermarking scheme is inspired by using a PJND (Pyramidal Just Noticeable Difference) model and the message is inserted into these DoG (Difference of Gaussians) [1, 2]. Our proposal takes into account three main characteristics of Human Visual System, namely: contrast sensitivity, luminance adaptation and contrast marking. Therefore, it not only provides an invisible and robust watermarking but also optimizes watermarking capacity. The performance of the proposed technique is evaluated by a series of experiments with different input images. In terms of transparency, besides using the subjective experiments, eight objective metrics are calculated in comparison with other methods such as PSNR, MSSIM, SVDm, etc. Our approach always presents the outperform values. In terms of robustness, many kinds of attacks from global transformation (rotation, scaling, etc) to local transformation (stirmark, checkmark benchmarks, de-synchronization attacks) are implemented. Many image processing tools are applied to simulate the attacks such as Print-Screen, Using Photo editing software, Camcorder, Print-Scan, etc. The experimental results show an outstanding robustness in resisting these attacks.

Received 04 December 2015, revised 09 January 2016, accepted 14 January 2016

Keywords: Digital Watermarking, Print-Scan process, DoG, De-synchronization attacks, Camcorder.

1. Introduction

Along with the rapid development of the media in communication, it is important and necessary to protect the ownership information of digital images because illegal copying of digital multimedia has become much easier. Recently, many watermarking schemes regarding copyright issue have been proposed for digital media but few methods have been proposed for un-digital content such as the print and scan attack is a challenging one because it not only alters the pixel values but also changes the positions of original pixels. Most of the watermarking systems use a secret key in the embedding phase to encode the watermark. In the detection phase, the same key is required to decode the embedded watermark. The watermarked content is then transmitted via a distribution channel. In transmission process, it may suffer some intentional as well as unintentional manipulations (called attacks) that try to remove or invalid the watermark. There are two major categories of attacks:

• **Unintentional Attacks:** This type of attack consists of all processes that do not initially aim at removing or suppressing watermark. They involve some deteriorations due to compression (Jpeg, Mpeg, etc.), filtering, A/D conversion, changing of coding format or resolution, etc. that a watermarked content may encounter through the transmission process.

☆ This work is dedicated to the 20th Anniversary of the IT Faculty of VNU-UET

* Corresponding author. E-mail.: nguyenlv@vnu.edu.vn

• **Malicious Attacks:** These attacks aim at making the watermark useless, for example, camcorder copy, print-scan, de-synchronization attacks or collusion attacks. However, this group is challenging because they not only alter the pixel intensities but also change the pixel location. In contrast to removal attacks, de-synchronization attacks do not actually remove the embedded watermark itself, but tend to make loss the synchronization between the embedder/detector (i.e alter the location of the watermark in the content). The watermark still exists, but it is undetectable by the detector.

Obviously, in order to preserve robustness of watermark, it is necessary to increase the watermark volume but it accidentally reduces the transparency. This raises a novel problem which is how to tradeoff between robustness and imperceptibility to obtain the best watermark. Many recent physiologic researches show that perceptual factors in HVS (Human Visual System) could be a potential solution to this problem. HVS modeling has become an important issue in image and multimedia processing such as image compression, quality assessment as well as watermarking. Many perceptual watermarking schemes have been proposed [3-6].

In this work, the proposed method deploys a Pyramidal Just Noticeable Difference (PJND) model in [1-2] fixed parameters are tuned as in visual experiments by [2] to attain fidelity. The embedding scheme is similar to the one proposed by [7] but here, the embedding is based on a pyramidal JND (Just-Noticeable-Difference) model in which its strength is controlled by a threshold and the DoG (Difference of Gaussians) representation [1, 2]. Experimental results have proved the performance of our approach in terms of transparency and robustness against severe attacks from Stirmark and Checkmark benchmarks as well as Photo editing software manipulations, especially to de-synchronization, print-scan and camcorder attacks.

For embedding process, we used a template based method where watermark is transformed in a template pattern and the corresponding transform coefficients are used as input message with synchronization and error correction. In detection scheme, the message is extracted from input image based on autocorrelation function after filtering, masking and adaptive line searching with Hough transform. A Hamming coding may be used to ensure that the message can be decoded correctly.

Our paper is structured as follows. Section 2 introduces the related works to summarise the existing methods. Our proposal will be presented in section 3 and 4. In section 3, a detailed embedding scheme is presented and the corresponding extraction scheme is described in section 4. In section 5, the performance of the proposal is evaluated and discussed based on a series of experiments. Finally, the paper ends with a conclusion and states some potential directions for future work.

2. Related Work

De-synchronization attacks are considered as one of the most serious threats for any watermarking system [8]. Therefore, many countermeasures have been introduced in the literature to cope with this type of attack [9] [10], but a perfect robustness to de-synchronization attacks has not been thoroughly obtained and still remains an outstanding area of watermarking research. De-synchronization attacks do not try to eliminate the watermark but aim to make the watermark undetectable although it still remains in the content. In general, we can loosely classify de-synchronization attacks into two categories as below though there is no clear distinction between them:

Global geometric distortion is a parametric transformation which is applied on the whole image. All the pixels are affected in the same manner. An example of typical global geometric transformations which includes affine transforms [11, 12, 13] (including rotation, scaling, translation (RST)) and projective transforms is given in Figure 1.

Such attacks are quite simple to apply but really present a challenge for current watermarking techniques. Indeed, there is no perfect solution for this problem, the robustness to global affine transformations is more or less

handled by using some approaches such as template-based re-synchronization [14], self-synchronizing watermarks [15] and embedding in invariant domains [16, 17].

Figure 1. Example of global geometric distortion. Left: An Affine transform, Right: A Projective transform

Local geometric distortion involves a set of different geometric transforms (with different parameters) applied to different portions of the image so that pixels are warped in different ways. This kind of attack mainly includes random displacement (also known as random jitter attack, introduced in Unzign benchmark), Random Bending Attack (RBA), also called Stirmark attack [18] (see Figure 2) and the two recently reported Desynchronization Attacks proposed by [8]. In the case of random attack, it is almost impossible to estimate the transformation parameters.

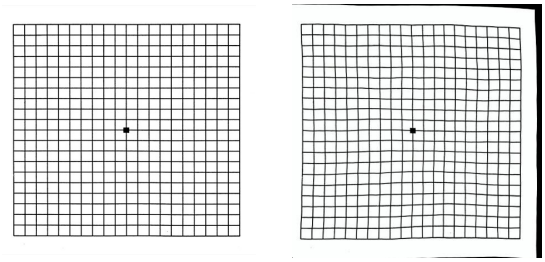


Figure 2. Example of local geometric distortion. Left: Original Image, Right: Local Random Bending Transform from Stirmark

Since the parameters needed to describe the local geometrical transformation are normally much more than those needed for global geometrical transformation, resynchronization from local geometrical distortion is much more difficult than from the global one. In the case that the attack is random, it is almost impossible to estimate the transformation parameters. Furthermore, local geometric attacks are "dangerous" in the sense that they destroy the watermark synchronization without creating significant visual distortion because the Human Visual System (HVS) is less sensitive to slightly local modifications. Hence, resistance to local random alterations like RBAs still remains as an open problem for most of watermarking schemes due to the high complexity of the attack parameter space. Geometrical attacks [11, 12] cause synchronization errors in watermark detection/extraction [19, 20]. Recently, several better approaches dealing with this type of attack have proposed resynchronization using additional template.

Authors embedded [21] an additional template together with the watermark in the DFT domain. This template contains no information about the embedded message but could be later used to recover the transformation undergone by the image. During the detection phase, the template is detected first using inverse transformation before extracting the watermark. However, one major drawback of this approach is that templates can easily be detected and erased by searching local peaks in the transform domain. Furthermore, template-based approach seems to be robust only to some global geometrical transforms such as RST rather than to local geometric distortions. It was also discussed in [22] that local geometrical transforms such as RBAs not only increase the search space and computation significantly for Exhaustive Search Detector, but also raise a serious problem for the template-based watermarking algorithm.

Another alternative approach in [11, 12] [18] inserted a periodic matrix brand in the DWT domain. The estimation of the geometric attacks is evaluated based on the brand autocorrelation function to obtain autocorrelation peaks. If the image is attacked by geometric operations, this plane will undergo the same attack. Indeed, the correlation or cross correlation of a signal by itself detects repeated patterns in a signal as a periodic signal is disturbed by a lot of noise. So, thanks to the periodicity of the brand, the brand's autocorrelation function locates periodic peaks. The mark detector then estimates the

geometrical transformation performed on the image with reference to the plane of the peaks of the extracted mark. The initial state of the image can also be reconstructed.

Among other various attacks the print and scan attack is challenging. The printed images are first converted to digital format (scanning) and the information is extracted from the detected copies; watermark image needs to convert analog data into a digital format in database applications. Both the printing and capturing processes cause various attacks including A/D or D/A conversion, compression, quantification, dithering, filtering, blurring, sampling, noise adding, contrast enhancement, etc. Some of geometrical attacks and distortions caused by devices become challenging problems in watermark extracting. In [7], a template based watermarking embedded multi-bit messages into image spatial blocks using periodic patterns, while one block embeds synchronism information. The embedding scheme is similar to the one proposed by [7] but this is done as follows:

○ In contrast to the above methods, we embed the watermark and JND mask performed into different scales of the DoG scale space, hence reducing the complexity of the method. This model takes into account three main characteristics of the Human Visual System (HVS), namely: contrast sensitivity, luminance adaptation and contrast masking. In [7] only with luminance does not consider neither contrast sensitivity, contrast masking and informed coding or the color channel properties. Furthermore, to ensure transparency, the embedding strength is determined using the pyramidal JND model proposed in [1] [2] and adapted here for the scale space transform.

○ Both embedding and extraction are adaptive, with no need to change parameters settings for different images. It is shown that background variation, or a change of printer/print material (two printers, two materials) has no significant effect on the performance of the method.

○ Completing the features balance between the three categories for a watermarking system namely transparency, robustness and capacity.

○ The message is protected with extended Hamming (64, 57) error correction coding that is capable of correcting three bits.

3. Our Watermark Embedding Scheme

In our proposal, the watermark is embedded into an original image based on a pyramid JND model in order to improve the robustness in countering some attacks on images as well as preserving perceptibility. For the sake of simplicity, we consider only the additive embedding scheme. The detailed diagram of the watermarking method is shows in Figure 3; the order of operations involved is depicted in the following section.

3.1. Computing the Pyramidal JND map

The JND model takes into account only the luminance channel. Hence, in order to apply for color images, the image is first transformed into YCbCr color space, and then only the Y component is watermarked. The JND model in our model is similar to one in [1] [2]. The input image is first decomposed into DoG scale images; the Gaussian image is replaced by the original image and the JND map is then computed for each DoG level. In each level of the scale space, a JND map is computed by incorporating the most relevant HVS's properties such as contrast sensitivity function (CSF), luminance adaptation and contrast masking. Fixed parameters are tuned as in visual experiments by [2] to attain fidelity.

The image and the JND image are then divided into blocks, and several bits are embedded in each block. The following section explains the message encoding and embedding.

3.2. Preparing the message

3.2.1 Hamming encoding message

The watermark is read and processed block by block, and the watermark capacity depends on the number of bits embedded in each block. In our experiments, the image was divided into sixteen blocks and four bits in each block are used for watermarking. Thus, the watermark capacity is of 16x4=64 bits. In this application, we used a (64, 57) extended Hamming code. This is an extension of the original (63, 57) Hamming code by adding an additional redundant bit. A Hamming code [23] is used to

add redundancy to the bits so that the errors can be detected or corrected to a certain extent. Hamming code is a linear block code. The main advantage of linear block codes is the simplicity in implementation and low computational complexity. A linear block code is usually composed of two parts. The first part contains the information bits, the original bits to be transmitted. The second part contains the parity checking bits, which are obtained by summing over a subset of the information bits. A linear block code with length n and k information bits is denoted as a (n; k) code.

The embedded message is protected with an extended Hamming (64,57) which is constructed by a parity bit at the end of the codeword to get even parity error correction coding that is capable of correcting one bit or detecting three erroneous bits. Each Hamming coded sequence of message is transformed into rotation angles by assigning each sequence a value between 0 to 180 degrees. The value is chosen by quantizing the rotation angles and assigning each of the values a number of bits, as illustrated in Figure 4. The quantization angle is determined by equation (1):

$$\alpha = \frac{180}{2^m} \quad (1)$$

where m is the number of bits embedded in each of the blocks.

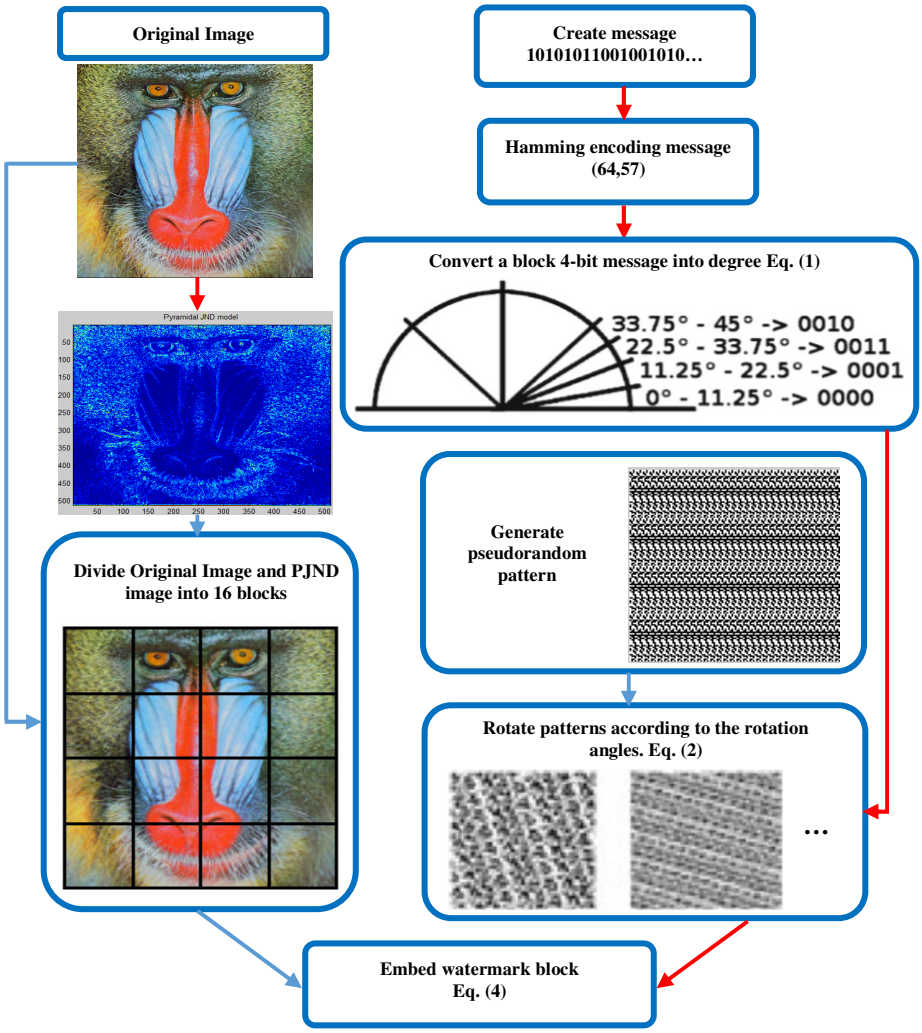


Figure 3. Illustration of watermark embedding process.

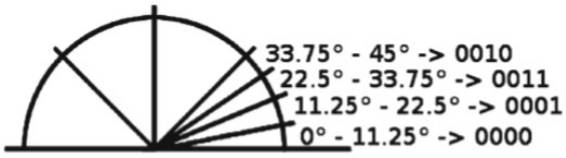


Figure 4. Creating encoding table.

3.2.2. Generating pseudorandom template pattern

The patterns are formed for each block. Each pattern is formed by first repeating a small rectangular pseudorandom sequence until the sequence covers an area the size of the block. A bipolar random sequence $W \in \{-1,1\}$ with mean zero and variance one is generated size 64x64.

Note: in order to get good results, the size of W and the size of one block original image (in this original photo 1/16) should be close together, the closer together the better. The reason is interpolation image will not affect many temples pattern.

When the direction of the template pattern is detected and the direction is erroneously interpreted to the adjacent quantization step, the Hamming coding ensures that the message can be decoded correctly because only one bit changes between adjacent quantization steps.

Each pattern is then rotated according to the message and cut to the size of the block. The process does not affect periodicity. The pattern is embedded in the image block.

3.2.3. Rotating template pattern

It is called the three shear rotation method. The heart of this method is the expansion of the single 2D rotation matrix into three matrices [24]:

$$W^o = \begin{bmatrix} x^* \\ y^* \end{bmatrix} \quad (2)$$

$$= \begin{bmatrix} 1 & -\tan(\theta/2) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \sin\theta & 1 \end{bmatrix} \begin{bmatrix} 1 & -\tan(\theta/2) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Here, bilinear interpolation is used to resize image. Parameter θ defines the rotation angle as in equation (1) of the Hamming coded bit sequence through the results in Figure 4.

- There are some very interesting properties of these three matrices:

- Three matrices are all shear matrices.

- The first and the last matrices are the same.

- The determinant of each matrix is 1.0 (each stage is conformal and keeps the area the same).

As the shear happens in just one plane at a time, and each stage is conformal in area, no aliasing gaps appear in any stage.

3.3. Embedding rule

We evaluate the smoothness of the image area in each block instead of 16×16 sub blocks [7] using average gradient magnitude on an image sharpened with an un-sharp mask. We use linear relationship between β and the average gradient magnitude to place more watermark strength on textured blocks according to equation:

$$\beta = \frac{M_Y}{M_{JND} * M_W} \quad (3)$$

In the equation (3) M_Y, M_{JND} and M_W, respectively stand in average gradient magnitude on original image, Compute Pyramidal JND and interpolate watermark blocks. The watermark is directly embedded in the Pyramidal JND, in which its strength is controlled by using the JND threshold in [2]. By this way, salient regions tend to mask non-salient regions. JND threshold is hence modulated by two masking mechanisms: the contrast masking and the "saliency masking". Recent JND models [7] do not take into account this phenomenon and therefore do not completely exploit HVS limitation.

The embedding of the message in the host image is realized in spatial domain utilizing the equation:

$$I_i(x,y) = Y_i(x,y) + JND_i(x,y).*(\beta * W_i^{\theta}(x,y)) \quad (4)$$

where I_i is the i^{th} watermarked block of the image, Y_i is the original image, JND_i is Multi-scale JND Map and the W_i^{θ} represents the coded watermark information in the form of directed template pattern.

4. Watermark Detection Process

Print and capture involve several distortions, because of the user interaction, the devices and air interface, which are taken into account in designing watermark extraction algorithm. Our detection scheme is shows in Figure 5.

First, the captured image is downsampled by using bilinear interpolation. It was necessary to compromise between the high processing time and the amount of data processed. After capturing a picture with a scan, camcorder the captured image is divided into blocks. The existence of a watermark is processed. The message is read by processing the blocks.

Figure 5. The Detection Scheme.

The un-sharp filter is a simple sharpening operator which derives its name from the fact that it enhances edges (and other high frequency components in an image) via a procedure which subtracts an un-sharp, or smoothed, version of an image from the original image. The un-sharp filtering technique is commonly used in the photographic and printing industries for crispening edges [25]. A signal proportional to the un-sharp or low pass filtered version of the original noisy image is subtracted from the image so that the resulting image is a crisp high-contrast image [26].

For each block, a $\tilde{W}_i(x,y)$ un-sharp filter estimation of the template watermark structure is calculated:

$$\tilde{W}_i(x,y) = Y_i^*(x,y) - Y_{i\ smooth}^*(x,y) \quad (5)$$

where $Y_i^*(x,y)$ is the i^{th} watermarked block and $Y_{i\ smooth}^*(x,y)$ is a low pass filtered version of represents the adaptive Wiener filtering $Y_i^*(x,y)$. Bilinear interpolation is used to resize watermark $\tilde{W}_i(x,y)$ of template pattern.

The autocorrelation function (ACF) is calculated of the Sharpening estimate, and this operation doubles the size of the processing block. Autocorrelation function utilized in order to reveal the periodicity in the extracted watermark estimate can be calculated as:

$$R_{\tilde{W}_i\tilde{W}_i}(a,b) = \int (\tilde{W}_i(x,y)\tilde{W}_i(x+a,y+b))dx \quad (6)$$

The autocorrelation is scaled to the range of [0,1]

$$R^*_{\tilde{W}_i\tilde{W}_i}(a,b) = \frac{\left|R^*_{\tilde{W}_i\tilde{W}_i}(a,b)\right|}{\max(\max(R^*_{\tilde{W}_i\tilde{W}_i}(a,b))} \quad (7)$$

The enhanced autocorrelation peaks are then thresholded, and a binary grid is formed with equation:

$$G^*(a,b) = \begin{cases} 1, & when \quad M(a,b)xR^*_{\tilde{W}_i\tilde{W}_i}(a,b) \geq T \\ 0, & when \quad M(a,b)xR^*_{\tilde{W}_i\tilde{W}_i}(a,b) < T \end{cases} \quad (8)$$

where *M(a, b)* is a circular masking operation. The central area of the autocorrelation image contains noise, which causes errors in line detection. Therefore, the center of the grid is masked out.

Finding edges in intensity image: edge takes an intensity or a binary image $G^*(a,b)$

as its input, and returns a binary image BW of the same size as $G^*(a,b)$, with 1's where the function finds edges in $G^*(a,b)$ and 0's elsewhere. The Sobel method finds edges using the Sobel approximation to the derivative. It returns edges at those points where the gradient of $G^*(a,b)$ is maximum.

The peaks are aligned according to the direction of the pseudorandom sequence pattern and this alignment is detected with Hough transform and line detection. These detected lines then give the angle of the pattern and thus the message.

Lines are searched from the binary grid using Hough Transform. The dominating direction is found by evaluating the number of peaks allocated to the same bin in the Hough transform matrix. Due to the properties of the Hough transform, the possible false peaks in the autocorrelation function have little or no effect. However, it is important to locate as many of the correct peaks as possible for reliable determination of the angle. These peaks are then presented to Hough transform as input to find the dominating direction formed by these peaks and thus giving an angle. Obtained angles are decoded to message nibble using the encoding equation (1).

The order of operations, to extract the message from a captured image, is presented in Figure 5. The message is extracted by analyzing the autocorrelation peaks of the embedded template pattern. Hough transform is used to detect the angle of lines made by these autocorrelation peaks. It helps in correctly identifying the aligned peaks as the small errors in the detection, due to the misleading peaks that appear due to thresholding, are minimized because of the robust properties of Hough transform at the watermark detection side.

The process is repeated for each block and the angle information is quantized. The same quantization step size and encoding table as during embedding is used. Each quantized angle value represents a coded bit sequence which is interpreted using a coding table and decoding. Finally, Hamming (64, 57) error decoding is used to decode the message.

5. Experimental results and discussion

To validate the performance of our method in terms of robustness and imperceptibility, some experiments are carried out on 512x512 scale images. We test the robustness of watermark with some common attacks namely: JPEG compression, Gaussian noise, cropping, and low-pass filtering.

The JND model takes into account only the pixel luminance. Hence, in order to apply for color images, the image is first transformed into YCbCr color space and then, only the Y component is watermarked. Further experiments are also carried out on a variety of natural images to validate the performance of our method in terms of robustness and imperceptibility. Due to the limited space and in order to facilitate the comparison, we only report the results for a set of 10 images, each of which contained 57 bits and error coding.

5.1. Transparency Evaluation

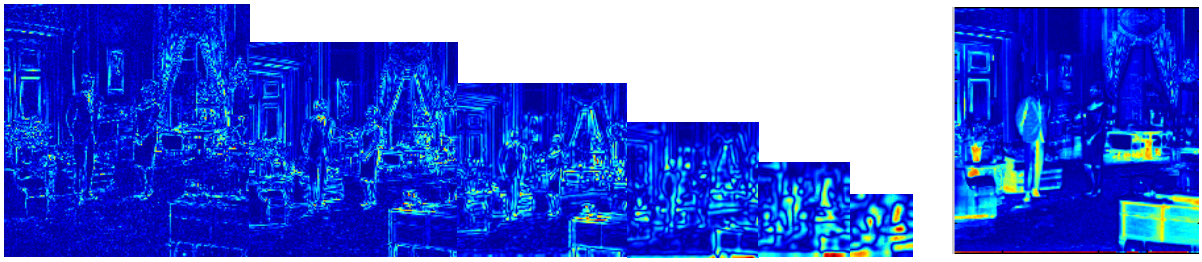
The results in Figure 6 show that the proposed algorithm provides a good imperceptibility the results in Table 2, at the same perceptual quality, the better the model [7] by subjective test. The proposed JND estimator has been compared with JND models [7] in Figure 7.

Although subjective assessment approach is the appropriate and accurate solution to watermark transparency evaluation, it is usually inconvenient, expensive and time-consuming, and not always easy to use. These drawbacks have led to the use of objective assessment as an alternative method. The goal of objective quality evaluation is to assess the quality of image/video by means of an automatic tool (objective metric) without performing any subjective test. In this paper, we have investigated the performance of objective assessment is also done HVS inspired quality metrics, the SSIM (Structural Similarity Index Measure) [27], the so-called Watson metric [28] which can measures the Total Perceptual Error [29] (TPE) between the original and the watermarked image and a wavelet-based metric (PSNR_wav1 and PSNR_wav2) [30], Peak

Signal-to-Noise Ratio (PSNR) comparision for Images Gray , weighted peak signal-to-noise ratio (wPSNR)[28][29] using weighted signal-to-noise ratio (wSNR)[31], Singular value decomposition (SVD)[32]. These metrics have been designed to general image quality assessment, it is therefore necessary to study their performances to the specific purpose of watermark transparency assessment. As mentioned in the first section, there is no objective metric specifically designed for watermarking purpose. It would be then very useful if we can determine, amongst the existing metrics, the one which is the most appropriate for watermark transparency assessment. The results in Table 2 show that the algorithm provides an excellent imperceptibility. However, the objective measures do not fully correlate with the subjective evaluation. It is mainly due to the variation of visual image content. Hence, a specific quality metric for watermarking is still missing.

(a) Original image (b)

Figure 6. Original image middle and Watermarked image: Proposal left (a), Model [7] right (b).



Proposal Model [7]

Figure 7. JND maps of different models.

Experimental results are reported in Table 2. It can be seen that the proposed model yields slightly lower as PSNR, PSNR_wav1, PSNR_wav2, SVD, wPSNR, wSNR scores than [7] for most of test sequences. However, we can see that MSSIM scores of the proposed model are higher than [7], which means that the proposed JND model not only conceals more distortions but also achieves a better quality. To demonstrate the invisibility of the watermark and the advantage of our method with method [7], we use the Structural Similarity Index Measure (SSIM) metric, [27] proposed a multi-scale version of SSIM (MSSIM) where the images are low-passed filtered and down sampled by a factor of two and the contrast and the structure are computed for each sub-sampled level for evaluating the quality of watermarked image of different 512x512 of 10 images. It works under the assumption that human visual perception is highly adapted for extracting structural information from a scene. The SSIM index is based on a combination of luminance, contrast and sensitivity of the

watermarked image with the original. The comparisons are performed on local windows; the overall image quality is averaged on these local windows. SSIM has become a very well-known metric for perceptual image quality assessment and has been extended in various directions. Our results are reported in the following figures and in Table 2. As shows in Figure 6, the watermarked image and the original are visually undistinguishable.

As for the watermark invisibility, quality results of watermarked image of different methods for different images are reported in Table 2. We can observe that the quality of the watermarked images of our method is equivalent or even better than that of other method [7].

5.2. Evaluation of robustness

The robustness of our algorithm is tested via a wide range of attacks including "signal processing" and de-synchronization types (DA). Some of them are very severe attacks like Print-Scan, Print Screen, Camcorder attack, Using Photo editing software and new de-synchronization attacks developed in [8] values are shown in Table 3. To facilitate the task, there are various tools that can test and evaluate watermarking algorithms systematically. Among them, the following two tools are most known to Stirmark [33] and Checkmark [34] benchmarks. However, when regarding the results in [7], we can see that the proposed method has a nearly equivalent robustness for geometric attacks values shown in Table 1. Furthermore, it resists other specific attacks (Camcorder, Print Screen, Using Photo editing software, new DAs…) that the method [7] cannot.

New DAs: these de-synchronization attacks are an extension of classical geometric attacks proposed by [8]. They are proved to be more powerful and less intrusive than the Stirmark attack. We tested three types of these with default parameters as in [8]: the LPCD (Local Permutation with Cancellation and Duplication, C-LPCD (Constraint LPCD), MF (Markov Random field). Watermark detection results are shown in Table 3.

Table 1. Robustness Evaluation (Stirmark [33] and Checkmark [34] benchmarks attack)

Attack Type	**Explicit scheme**	**Method [7]**
Random Cropping	1%	0.8%
Jpeg compression	QF=3%	QF=9%
Jpeg 2000 compression	0.08 bpp	0.1 bpp
Gaussian Noise	σ= 64%	σ= 67%
Wiener filtering	Ok	Ok
Median filtering	5x5	3x3
Sharpening	Ok	Failed
Blurring	Ok	Failed
Bit plan reduction	Ok	Failed
Histogram Equalization	Ok	Ok
Rescale (45%)	Ok	Ok
Affine Transform	Ok	Ok

Print-scan attack: this attack consists of printing image on a classical laser printer: HP LaserJet 4250 PCL6, EPSON Stylus. Scanning: Epson Perfection 4490. The image is printed in color, grayscale level on an A4 paper at 300 dpi resolution (tests were done on image printed on a white paper) and scanned, witch is shows in Figure 8. Watermark detection results are shown in Table 3.

Camcorder attack: we get the picture of image on the computer screen with the Nikon D90 Digital SLR Camera with 18-105mm VR Lens Kit (12.3MP) 3inch LCD. The watermarked test images were captured 10 times with each of the camera setting and each image contained 57 bits and error coding when images were captured by tilting the camera randomly is shows in Figure 9 and values are shown in Table 3.

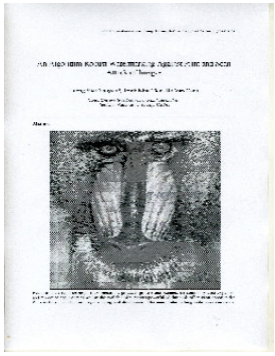
Using Photo editing software: Do you still use Microsoft Paint, or some other under-powered paint packages that allow you to rotate an image by an arbitrary angle (Figure 9).

Table 2. Imperceptibility Evaluation

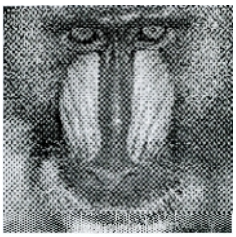
Objective	Method	Image										AVG
		Baboon	Barbara	Boat	Car	Clown	Fruit	Isabe	Lena	Peppers	Plane	
PSNR	Keskinarkaus	25,79	28,41	33,01	32,80	33,66	36,68	36,18	35,45	35,51	35,82	**33,33**
	Proposed	26,10	27,88	32,46	31,98	32,98	34,84	35,25	34,47	34,35	33,61	**32,39**
PSNR wav1	Keskinarkaus	9,36	11,64	15,96	16,93	17,13	19,25	16,78	17,72	19,16	19,65	**16,36**
	Proposed	9,31	10,48	14,66	15,35	15,54	17,02	15,15	15,92	16,97	16,69	**14,71**
PSNR wav2	Keskinarkaus	10,08	12,43	16,33	17,69	18,13	19,68	18,56	19,11	19,99	20,03	**17,20**
	Proposed	9,88	10,87	14,87	15,95	16,45	17,53	16,90	17,14	17,75	17,08	**15,44**
SVDm	Keskinarkaus	37,82	34,82	16,47	17,35	12,51	9,74	9,28	10,70	8,07	11,10	**16,79**
	Proposed	33,36	29,64	13,94	15,20	11,18	9,65	8,27	9,88	7,87	10,82	**14,98**
TPE	Keskinarkaus	0,19	0,10	0,08	0,08	0,08	0,07	0,07	0,07	0,07	0,06	**0,09**
	Proposed	0,19	0,13	0,08	0,09	0,10	0,08	0,07	0,07	0,08	0,07	**0,10**
mssim	Keskinarkaus	0,82	0,88	0,93	0,94	0,93	0,97	0,94	0,94	0,91	0,96	**0,92**
	Proposed	0,85	0,89	0,94	0,94	0,94	0,97	0,95	0,94	0,93	0,96	**0,93**
wPSNR	Keskinarkaus	35,56	37,28	39,25	38,34	38,62	40,46	39,98	39,76	38,95	40,93	**38,91**
	Proposed	35,99	36,93	38,97	37,80	37,90	39,60	39,24	39,31	38,80	39,46	**38,40**
wsnr	Keskinarkaus	33,18	34,85	39,28	38,09	34,50	37,17	40,58	38,69	39,77	41,39	**37,75**
	Proposed	32,16	32,08	37,92	36,41	33,71	35,39	38,82	37,05	37,17	38,08	**35,88**

Table 3. Robustness Evaluation. For some type of attacks, the results showed: X/Y (bit error/bit encoded message) the parameters demonstrate the break-down limit of the method (the strongest attack to which the watermark still survives)

Attack	Method	Baboon	Fruit	Isabe	Lena	Peppers
Camcorder attack	Ours	Ok	Ok	2/64	1/64	3/64
	Keskinarkaus	-	-	-	-	-
Print scan Attack	Ours	Ok	3/64	Ok	2/64	Ok
	Keskinarkaus	Ok	2/48	3/48	Ok	Ok
Photo editing software	Ours	Ok	Ok	2/64	3/64	Ok
	Keskinarkaus	-	-	-	-	-
Print screen Attack	Ours	Ok	Ok	Ok	Ok	Ok
	Keskinarkaus	-	-	-	-	-
DA New	Ours	Ok	Ok	Ok	Ok	Ok
	Keskinarkaus	-	-	-	-	-

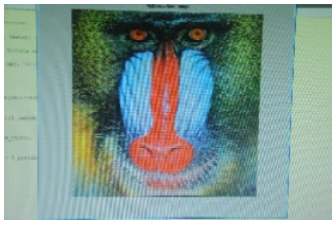


Watermark Image Insert Document and Print



Scanner at 300dpi resolution

Figure 8. The watermark image is printed and scanned.

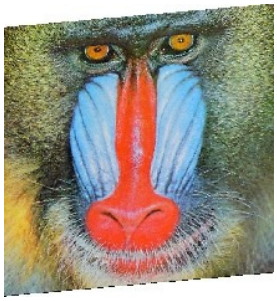


Watermark Image screen with the Nikon

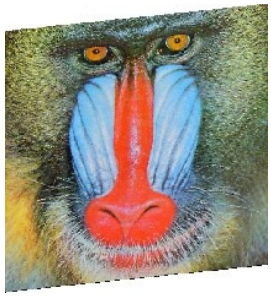
Crop image and Detection

Figure 9. The watermark screen with the Nikon.

We user vertical shear and skew the image a negative number of degrees (1-6 degrees) in the vertical plane which is shows in Figure 10 and values are shown in Table 3.



Vertical shear 5 degrees



Vertical shear 6 degrees

Figure 10. The watermark attack Using Photo editing software.

Print screen Keyboard: When you press it, an image of your screen is copied to the Clipboard. This is called a screen capture or screen shot. You will then need to further edit using some image editing programs values shown in Table 3.

Table 3 shows the average robustness tested for five images (Baboon, Fruit, Isabe, Lena and Peppers). These values denote the breakdown limit of the tested methods, i.e. the strongest level of attacks to which the watermark still survives. Table 1 shows that the watermark survives many severe attacks in both schemes but there are no significant differences in robustness between these two schemes (except for Jpeg compression). Furthermore, robustness against some attacks "like Jpeg" (Jpeg2000) is even slightly improved.

Watermark detection results are shown in Table 1 and Table 3; our method outperformed the method [7] for most attacks. Furthermore, the message protected with Hamming (64, 57) error correction coding that is capable of correcting three bits ensures that the message can be decoded correctly. Especially, in contrast to [7], it survives many severe attacks such as "camcorder", "print-scan" and Stirmark, Checkmark and new DA. However, our method as well as the method [7] are not very robust to "signal processing" attacks such as noise, jpeg compression, etc. Throughout these results, it is clear that using perceptual models helps improve not only transparency but also robustness of a watermarking system. The explicit scheme, once again provides the best robustness amongst the compared methods. The detector outputs for some severe attacks are also displayed in Figure 8, 9 and 11.

6. Conclusion

In this paper, we have presented a novel content based image watermarking operating in the DoG scale space with enhancing robustness against de-synchronization attacks. Such watermarking methods present additional advantages over the published watermarking schemes in terms of detection and recovery from geometric attacks, and with better security characteristics. The experimental results show that the proposed method has a good performance in terms of robustness and imperceptibility. In the future, this method digital watermarking will be extende to used on mobile phones.

References

[1] Luong Viet Nguyen, Trinh Nhat Tien, Ho Van Canh. "Pyramidal JND Model for Grayscale Image and its Application to Watermarking". Proceedings of IEEE International Conference on Computer Science and Automation Engineering Vol.01[C], 2013.

[2] Luong Viet Nguyen, Trinh Nhat Tien, and Ho Van Canh. "A watermarking method robust for copyright protection of images against Print-scan." Information Science and Control Engineering (ICISCE), IET International Conference on. IET, 2012.

[3] Ensaf Hussein, Mohamed A. Belal, "Digital Watermarking Techniques, Applications and Attacks Applied to Digital Media: A Survey," IJERT, ISSN: 2278-0118, Vol. 1 Issue 7, 2012.

[4] S. Voloshynovskiy, A. Herrigel, N. Baumgartner, and T. Pun, "A stochastic approach to content adaptive digital image watermarking," in Proc. of the 3rd International Workshop on Information Hiding, 1999.

[5] Mundher, Myasar, et al. "Digital watermarking for images security using discrete slantlet transform." Applied Mathematics and Information Sciences 8.6: 2823-2830, 2014.

[6] Kekre, H. B., Tanuja Sarode, and Shachi Natu. "Performance of watermarking system using wavelet column transform under various attacks." International Journal of Computer Science and Information Security 12.2 (2014).

[7] A. Keskinarkaus, A. Pramila, T. Seppänen, "Image watermarking with a directed periodic pattern to embed multibit messages resilient to print scan and compound attacks." the Journal of Systems and Software v. 83 (2010) 1715.

[8] Barni, Mauro, Angela D'Angelo, and Neri Merhav. "Expanding the class of watermark de-synchronization attacks." Proceedings of the 9th workshop on Multimedia & security, ACM, 2007.

[9] Wang, Xiang-Yang, and Chang-Ying Cui. "A novel image watermarking scheme against desynchronization attacks by SVR revision." Journal of Visual Communication and Image Representation 19.5(2008) 334.

[10] Schmitz, Roland, et al. "Towards Robust Invariant Commutative Watermarking-Encryption Based on Image Histograms." International Journal of Multimedia Data Engineering and Management 5.4 (2014) 36.

[11] P. Dong, J. G. Brankov, N. P. Galatsanos, Y. Yang, and F. Davoine, "Digital watermarking robust to geometric distortions," IEEE Trans. on Image Processing, vol. 14 (2003) 2140.

[12] M. Alghoniemy and A. H. Tewfik, "Geometric invariance in image watermarking," IEEE Trans. on Image Processing, vol. 13, no. 2 (2004) 145.

[13] Zhao, Yao, RongRong Ni, and ZhenFeng Zhu. "RST transforms resistant image watermarking based on centroid and sector-shaped partition." Science China Information Sciences 55.3, 650-662, 2012.

[14] Wang, Caiyin, and Chao Li. "A Steganography Approach for Printed Image Based on Image Complexity and Template Matching." Open Automation and Control Systems Journal 6 (2014) 84.

[15] Schlauweg, Mathias, et al. "Self-synchronizing robust Texel watermarking in Gaussian scale-space." Proceedings of the 10th ACM workshop on Multimedia and security. ACM, 2008.

[16] Joseph J. K., O' Ruanaidh, J. J. K. et Pun, T. "Rotation, Scale and Translation Invariant Digital Image Watermarking", Image Processing. International Conference, Issue, 26-29 (1997) 536.

[17] Wang, Xiang-Yang, et al. "A new robust digital watermarking based on exponent moments invariants in nonsubsampled contourlet transform domain." Computers & Electrical Engineering 40.3 (2014) 942.

[18] M. Mitrea, F. Preteux, M. Petrescu, and A. Vlad, "The Stirmark watermarking attack in the dwt domain," in Proceedings of the 12th IEEE International Conference on Systems, Signals and Image Processing (IWSSIP'05), Halkida, Greece, pp. 5–9, 2005.

[19] Wang, Xiang-yang, et al. "Robust image watermarking approach using polar harmonic transforms based geometric correction." Neurocomputing, 2015.

[20] N. Merhav, "An information-theoretic view of watermark embedding-detection and geometric attacks," in Proceedings of WaCha'05, Barcelona, Spain, 2005.

[21] D. Zheng, J. Zhao, and A. Saddik, "Rst invariant digital imagewatermarking based on log-polar mapping and phase correlation," IEEE Trans. Circ. Syst. Video Tech., vol. 13 (2003) 753.

[22] S. Pereira and T. Pun, "Robust template matching for affine resistant image watermarks," IEEE Transaction on Image Processing, vol. 9, no. 6, June (2000) 1123.

[23] W. Peterson., "Error-correcting codes," 2nd ed., Cambridge: The MIT Press, 1980. 560p.

[24] Toffoli, Tommaso, and Jason Quick. "Three-dimensional rotations by three shears." Graphical Models and Image Processing 59.2, 89-95, 1997.

[25] Jain A., "Fundamentals of Digital Image Processing," Prentice-Hall, France, 1989.

[26] Amirgholipour S. and Naghsh-Nilchi A., "A New Robust Digital Image Watermarking Technique Based on Joint DWT-DCT Transformation," in Proceedings of the 3rd International Conference on Convergence and Hybrid Information Technology, Busan, vol. 2 (2008) 539.

[27] Wang, Z., Bovik, A. C., Sheikh H. R., Simoncelli, P. E.: "Image quality assessment: from error visibility to structural similarity." IEEE Transactions on Image Processing, 13(4), pp. 600-612, 2004.

[28] Pereira, S., Voloshynovskiy, S., Madueño, M., Marchand-Maillet, S., Pun, T.: "Second generation benchmarking and application oriented evaluation." In Procs of 3rd Int. Workshop on Information Hiding, pp. 219-239, Pittsburgh, PA, USA, 2001.

[29] Pereira, S.: "Robust digital image watermarking", PhD thesis, Genève, Swiss, 2000.

[30] Beghdadi, A. and Pesquet-Popescu, B.: "A New Image Distortion Measure Based on Wavelet Decomposition." In Proc. of 7th IEEE ISSPA, vol. 2, pp. 485-488, Paris, France 2003.

[31] Niranjan, D. V., Thomas, D. K., Wilson, S. G., Brian, L. E., Bovik, A. C.: "Image quality assessment based on a degradation model", In IEEE Trans Image Processing, 9(4) (2000) 636.

[32] Shnayderman, R., Gusev, E., Eskicioglu, A. M.: "An SVD-based gray-scale image quality measure for local and global assessment." In IEEE Transactions on Image Processing, 15(2) (2006) 422.

[33] http://www.petitcolas.net/fabien/watermarking/stirmark/

[34] http://watermarking.unige.ch/Checkmark/index.html