Wavelet-based watermarking algorithms: theory, applications and critical aspects

S. Agreste* and L. Puccio

Department of Mathematics, Messina University, Messina, Italy

(Received 8 September 2009; revised version received 17 March 2010; second revision received 11 June 2010; accepted 12 June 2010)

The copyright protection of digital images is a critical element for multimedia applications on Web, electronic libraries and virtual image galleries. Since the middle of the nineties, digital watermarking techniques have been proposed as a streamlined solution to solve the copyright problem. In this article we present the difficulties occurred during the construction and the experimentation of our wavelet-based watermarking algorithm. In fact we faced the classical watermarking problems such as computational cost, imperceptibility and robustness. In addition to these critical aspects we had to resolve problems concerning the application of watermarking techniques on a real database Web gallery.

Keywords: copyright protection; digital colour image processing; wavelets; watermarking algorithm; adaptive technique

2010 AMS Subject Classifications: 65T60; 94A08

1. Introduction

The popularization and representation of multimedia document systems have known a rapid development due to the increase of technology. This is the cause of some problems of copyright protection. The copyright protection and the safeguard of the objects subordinated to such right are guaranteed by international laws. The World Intellectual Property Organization (WIPO) is an international organization dedicated to promoting the use and protection of works of the intellectual property. This organization counts 183 nations as member states. It manages several multilateral agreements about the legal and administrative acts concerned with intellectual property [17]. The most important agreement has been stipulated in the Diplomatic Conference on Certain Copyright and Neighboring Rights Questions in Geneva in 1996. Subsequently, the Congress of USA issued the Digital Millennium Copyright Act (DMCA) and the European Parliament drew up the directive on the harmonisation of certain aspects of copyright and related rights in the information society. Copyright laws have the purpose to protect the authors from the not-authorized use of their intellectual products. Despite several laws, the digital piracy phenomenon is constantly growing. Therefore it becomes strategic and very much important to be...
able to create and develop methods and numerical algorithms to resolve copyright problems. Such methods must be stable and with low computational cost. Since the middle of the nineties digital watermarking techniques have been proposed as a streamlined solution. A watermark, or an electronic mark, is a code or an identified logo. It can hold some information about the authors, the owner, the authorized distributor and the user of the multimedia document. The watermarking techniques do not prevent the copy of the digital object. Their use discourages a user from illegally redistributing copies of the objects because the watermark identifies the owner in an unambiguous way.

Most of the watermarking algorithms are based on a dual process: coding and detection. These two phases are temporarily independent, but without the embedding phase the detection phase would not have reason to be. In the coding process the watermark is embedded in a digital image to generate the watermarked image. Let \( I \) be a digital image and let \( W = \{\omega_1, \omega_2, \ldots, \omega_n\} \) be the watermarking signal; then the watermarked image \( \tilde{I} \) is obtained by a coding function \( C: C(I, W) = \tilde{I} \). In the detection process, a decoding function \( D \) extracts the watermark \( W' \) from the watermarked image \( \tilde{I}' = D(I, \tilde{I}') = W' \). Then the extracted watermark \( W' \) is compared with \( W \) by means of a similarity function \( F_\delta \): \( F_\delta(W, W') = c \). The Boolean function \( F_\delta \) indicates if the two watermark signals correspond (1) or do not correspond (0) by means a suitable threshold \( \delta \). The processes are shown in Figure 1. In order to conjugate the defense of the intellectual property and the quality of the image, this mark must be robust, unperceivable, unambiguous proof of ownership. Robust means resistant to digital attacks that can modify, delete, or substitute the mark, either intentionally or not. Unperceivable means that the watermark must not damage the visual quality of the image. Unambiguous proof of ownership means that the watermark detection must be identified by the owner of the image in an unambiguous way: the detection of another watermark (false-positive) and the not-detection of the embedding watermark (false-negative) must be impossible. Unambiguous proof of ownership is critical for copyright protection.

In this article we summarize and emphasize the difficulties occurred during the wide experimentation of our wavelet-based watermarking algorithm described in [2,3]. We faced the classification watermarking problems in: function selection in frequency domain algorithms, embedding of the watermarking signal, alignment between original and watermarked image for watermark detection and detection of the watermarking signal. Some of these critical aspects are related to the application of watermarking techniques on a real database Web gallery. In fact we have increased our sample set. The differences of historical-artistic images, considered during the wide experimental tests, have given a diversified and complex survey. Therefore the algorithm has been tested on full and heterogeneous set of images. This testing has permitted us to improve our algorithm, increasing the robustness of the watermark.

The paper is organized as follows. We start by presenting the classification of the watermarking algorithms and the different kind of attacks. In Section 3 we describe some critical aspects concerning the application of watermarking techniques on a real database Web gallery. Moreover we expose our solutions to these problems.
2. Watermarking algorithm classifications and attacks

There are different ways to classify the watermark algorithms. The first distinction that one needs to do in the study of watermarking for digital images is the notion of visible watermarks versus invisible ones. The visible watermark is used to mark, obviously in a clearly detectable way, a digital image in order to give a general idea of what it looks like while preventing any commercial use of that particular image [9]. It aims to discourage any unauthorized use of an image by adding an obvious identification element, which removes the image commercial value; while the purpose of the invisible watermark is to indicate whether a specified image has been used without the owner’s formal consent or if the image has been altered in any way.

A classification criterion distinguishes watermark schemes into spatial domain and frequency domain techniques in relation to the approach followed to process the original image in the embedding phase. Spatial domain techniques are less complex because they work directly in the spatial domain and no transform is involved [12,13]. In frequency domain techniques the watermark is encoded by altering some frequency bins obtained by transforming the image by means of discrete cosine transform (DCT [11]) or discrete wavelet transform (DWT [4,7]) and other techniques [8].

Moreover, it is possible to classify watermark algorithms according to the detection process into public and private. In the first case the detection function \( D \) identifies the watermark \( W \) from the watermarked images \( \tilde{I} \) (or \( \tilde{I}' \)), without the necessity to have the original image \( D(\tilde{I}) = W' \). The survey happens by computing the correlation between the elements involved in the embedding phase. In the second case the original image is essential in the detection phase. In fact the decoding function extracts the watermark \( W' \) from the watermarked images \( \tilde{I} \) compared to the original image \( I \). The survey happens by computing the correlation between the elements involved in the embedding phase. This method is more robust but it needs the original image.

We define attacks the alterations brought to watermarked images. The attacks can modify, delete, or substitute the embedded watermark. The main attacks can be assembled into four classes [18] (Figure 2): removal attacks remove the watermark from the watermarked image by means of typical image processing algorithms (i.e. denoising, lossy compression, quantization); geometrical attacks do not remove the embedded watermark but distort it through spatial or geometrical alterations of the watermarked image, so that the detection process is not aligned with the watermark signal embedded (i.e. StirMark [14,15]); cryptographic attacks are brute force attacks whose aim is to find the secret information used to generate the watermark signal (i.e. the brute-force search for the embedded secret information); protocol attacks aim at attacking the concept of the watermarking application [6].

3. Critical aspects and our solution

Many watermarking algorithms are proposed in the literature. The development of new watermarking algorithms arises from the need to create a robust one. Generally the watermark is simply
said to be ‘robust against common signal processing algorithms and geometric distortions when used on some standard images’ [10]. Our aim is to build robust watermarking algorithms for an applicative use in a real database whose images are different in colours, textures, patterns, shapes, process of acquisition and lightning (see Figure 3). In this article we present some critical aspects of the watermarking algorithm implementation for several image sets with different features. In particular we explain the following critical points:

- function selection in frequency domain algorithms;
- embedding of the watermarking signal;
- alignment between the original and watermarked images for watermark detection;
- detection of the watermarking signal.

### 3.1 Function selection in frequency domain algorithms

The use of frequency-based transforms allows the direct understanding of the content of the image. Frequency domain techniques perform a special transformation to the original image, embedding the watermark into part of the transformed coefficients. Embedding the watermark into the transform domain generally helps to increase imperceptibility, security, and robustness. Therefore, this approach is very common. DCT, DWT are the two main transform methods used. Recently the wavelet transform has often been used because it allows one to analyse different frequencies of the images [5]. The human visual system (HVS) study shows that the eye is less sensitive to noise in high frequencies; in particular, it is least sensitive at oblique angle orientation. Furthermore the peculiarity of the DWT is to obtain information on the spatial domain and on the frequency domain at the same time. It is possible to obtain information about the features of the image, as edges, texture, etc., to adapt the watermark according to the HVS study. Low frequency components have larger perceptual capacity compared to high frequency components because they have large magnitudes and can be used to embed stronger watermarks without introducing any visible artifacts. Moreover, embedding a watermark in the higher level sub-bands increases the robustness of the watermark; however the image visual fidelity, which can be measured by the peak signal to noise ratio (PSNR), may be lost. We can apply DWT to square matrices whose number of columns and rows are divisible by \(2^l\), where \(l\) is the decomposition level. In a real Web gallery, images have generally different sizes and can be square or rectangular with high and low resolution (Figure 3). It is necessary to transform the matrix associated to the image into a new matrix whose dimension is coherent with the mathematical condition of the DWT. In embedded and detection processes, we have introduced an additional, very important step:
the pre-processing step. If the images have much difference in length and width, for a major distribution of the watermark, we divide the image into two blocks and apply the algorithm to each of them. This phase has been described in [1,3]. For the sake of simplicity we describe it by an example.

Let $V$ be the matrix representing the value plane of input image of dimension $1714 \times 824$ (Figure 4(a)). From $V$, two blocks $V_1$ and $V_2$ of dimension $857 \times 824$ (Figure 4(b)) are obtained. Figure 4(c) represents the operations to have the matrices $C_1$ and $C_2$ of order 512. $C_1$ and $C_2$ are obtained by cutting the first 345 rows in $V_1$ and the last 312 rows in $V_2$.

### 3.2 How to embed the watermarking signal

There are different ways to insert the watermarking signal in the wavelet coefficients. We describe one of our solutions. Let $C$ be a matrix of dimension $n = 2^k$, coherent with the DWT mathematical conditions. The DWT decomposition is applied to $C$ to obtain the four sub-matrices $C_{LL}^k, C_{HL}^k, C_{LH}^k$ and $C_{HH}^k$ of order $n_l = 2^{k-l}$, where $l$ is a suitable decomposition level chosen according to the order $n$ of matrix $C$ (Figure 5). Watermark embedding is calculated with the following casting formulas:

$$\tilde{C}_\theta^i(i, j) = C_\theta^i(i, j) + \omega \ast \alpha^\theta(i, j) \quad i, j = 1, \ldots, n_l, \quad \theta = HL, LH, HH,$$

where $\omega$ is a global parameter related to the watermark strength:

$$\omega = \text{std} \left( \text{mean} \left( \text{mean} (C_\theta^i(i, j)) \right) \right);$$

$\alpha^\theta$, where $\theta = HL, LH, HH$, are the following weight matrices:

$$\alpha^\theta(i, j) = \begin{cases} 0 & \text{if } C_\theta^i(i, j) \in [-\omega, \omega], \\ 1 & \text{if } C_\theta^i(i, j) \geq T_\alpha \quad i, j = 1, 2, \ldots, n_l, \\ -1 & \text{if } C_\theta^i(i, j) < T_\alpha, \end{cases}$$

and $T_\alpha$ is a suitable, statistic and experimental threshold.
Figure 5. Sub-matrices $C^\theta_k$, $\theta = LL, LH, HL, HH$, $k = 1, \ldots, l$, obtained applying $l$ levels of DWT.

![Figure 5](image.png)

Figure 6. Example of different choices of $\alpha^\theta$-entries: (a) $\alpha^\theta(i, j) \in \{1\}$; (b) $\alpha^\theta(i, j) \in \{1, -1\}$; (c) $\alpha^\theta(i, j) \in \{1, 0, -1\}$.

![Figure 6](image.png)

Since $\omega$ represents the robustness property, while $\alpha$ guarantees the invisibility property, it is important to make an optimal selection of these elements.

In Figure 6 an example of different choices of the weight matrix entries is shown. From Figures 6(a) and (b) it is possible to observe that a bad choice of $\alpha^\theta(i, j)$, with $i, j = 1, 2, \ldots, n_l$, causes a visible watermark that damages the image quality, while Figure 6(c) represents the results obtained using our $\alpha^\theta$ solution.

One delicate point in the embedding phase is the storage of the watermarked image in JPEG format. In fact, this operation produces an alteration of pixels that can destroy the embedding watermark. In the first row of Figure 7 the difference between the original and watermarked images before the storage of the watermarked image is shown. It is interesting to observe the difference (second row) after having compressed the original image with the same factor (85) of the watermarked one: the error introduced by JPEG storage is reduced and the alteration of the watermarked process appears on all planes.

3.3 **Alignment between the original and watermarked image**

In the detection process, it is necessary to apply a check procedure verifying if the size of watermarked image has been modified by attacks as resize or cut [2]. To align the original and watermarked images we consider the following steps:

1. take out the central block $\tilde{B}$ of the watermarked image;
2. compare $\tilde{B}$ with some blocks $B_{ij}$ of the original image by means of the mean squared error (MSE) function;
3. resize the original image in order to make both $\tilde{B}$ and $B_{ij}$ congruent.

The alignment procedure is schematized in Figure 8. In Figure 9 an example of the alignment procedure is shown. Figure 9(a) represents the watermarked image modified by cut attack. In this figure the block in the red line is the sub-matrices $\tilde{B}$ of order $p$. Usually the value of $p \in$
Figure 7. Difference between original and watermarked images before and after storage in JPEG format.

Figure 8. Scheme of the alignment step.

\{4, 8, 12, 16\} represents the accuracy of the method. In our experimental test \( p = 8 \), there is a good agreement between the computational cost and the experimental results. We compare \( \tilde{B} \) with some blocks \( B_{ij} \) of the original image. \( B_{ij} \) are sub-matrices of order \( p \) of the block in the red line in Figure 9(b).

### 3.4 How to detect the watermarking signal

The detection method that we use is similar to the method proposed in [4]. Watermark detection is accomplished without referring to the original image. We have detected the watermark by computing the correlation \( \rho \) between the watermarked/original coefficients and the watermark,
in comparison to the threshold $T_\rho$, where

$$\rho = \frac{1}{M} \sum_\theta \left[ \sum_{i=1}^{n_k} \sum_{j=1}^{n_k} \alpha^\theta(i, j)(\tilde{C}^\theta_k(i, j) - C^\theta_k(i, j)) \right].$$

$k$ is the level of wavelet decomposition, $n_l$ is the order of $C^\theta_k$, $\theta = \text{HL, LH, HH}$, and $M$ is the number of non-zero entries of $\alpha^\theta$. The method used is based on the Neyman–Pearson statistic criterion [16]. It determines a detection threshold minimizing the probability of missing detection with respect to the given probability of false alarm.

To compute $T_\rho$, we have supposed that the probability of false positive detection $P_\ell$ is fixed to $10^{-8}$:

$$P_\ell \leq \frac{1}{2} \text{erfc} \left( \frac{T_\rho}{\sqrt{2\sigma^2}} \right) \quad P_\ell \approx 10^{-8},$$

where $\sigma^2$ represents the variance computed by

$$\sigma^2 = \frac{1}{M} \sum_\theta \left[ \sum_{i=1}^{n_l} \sum_{j=1}^{n_l} (\tilde{C}^\theta_l(i, j) - C^\theta_l(i, j))^2 \right].$$

Then if $\rho > T_\rho$, the watermarking signal is detected, otherwise not.
4. Conclusion

The copyright protection of digital images is a critical element for multimedia applications on Web, electronic libraries and virtual image galleries. Its aim is to answer the copyright problem of colour digital images by means of developing appropriate watermarking algorithms. In particular, we are interested in the analysis, development, implementation and application of watermarking algorithms to digital colour images from a viewpoint of an effective usability and availability in

Figure 10. Example of original and watermarked images attacked by add noise and rotation with cut and zigzag deformation: (a) original image; (b) watermarked image attacked by add noise; (c) original image; (d) watermarked image attacked by rotation and cut; (e) original image; (f) watermarked image attacked by zigzag deformation.
a real applicative background. In this article we have presented some critical aspects of wavelet-based algorithms. Moreover we have exposed our solutions to these problems. To analyze the robustness of the algorithms we have applied some attacks on a set of colour digital images belonging to the database BEWEB (http://www.chiesacattolica.it/beweb/). They are obtained by different acquisition processes (photo CD, digital camera, scanning from negative film, scanning of photographs). All the images are in JPEG format. As far as the attack simulations executed to verify the level of robustness of our algorithm we have used the *StirMark* tool with default parameters and the commercial software Adobe Photoshop: in particular we have executed using Adobe Photoshop the following geometric attacks: better blurring with range = 24.2 and threshold = 3.0; distortion: deform with parameter = 1 and zigzag with factor = −1 and relief = 1 (concentric); rotation degree=1; adding noise with parameter = 6; cutting of image part; resize of −1%; resize of +1%. In Figure 10 we present some examples of original images and the watermarked images attacked by add noise, rotation with cut and zigzag deformation, respectively.

Note that the proposed attacks introduce a practically unnoticeable quality loss in the image, but their application de-aligns the original and watermarked images. In Table 1 we present the results obtained on a set of 500 images in high resolution of BEWEB database with dimension greater than 300 Kb. Furthermore it is very important to note that the high percentage of success with respect to StirMark attack is 91% compared to the results relative to digital watermarking commercial software in [15].

By means of the proposed solutions we have developed a wavelet-based watermarking algorithm that is robust according to a new definition: ‘*robust* against common signal processing algorithms and geometric distortions when used on *real* images’.

### Acknowledgements

This work was supported by PRA Projects of University of Messina and GNCS of INdAM. We thank the IDS-Unitelm company of Messina and the Italian Catholic Bishop Conference for giving us the opportunity to use some images from BEWEB database in the experimentation phase.

### References


