Pakistan Journal of Information and Technology, 2 (3): 213-230, 2003
ISSN 1682-6027

A Robust Wavelet-Based Watermarking Scheme Using Quantization and Human Visual System Model

Tzung-Her Chen, Gwoboa Horng and Sang-Hao Wang
Institute of Computer Science, National Chung-Hsing University
250 Kuo-Kuang Road, Taichung 40227, Taiwan R.O.C.

Abstract: In this paper, a highly robust watermarking scheme based on Discrete Wavelet Transform (DWT) is proposed. It has the following features. First, unlike conventional proposed watermarking techniques, a watermark is embedded into the lowest frequency components. Second, the Human Visual System (HVS) model is adopted to improve the transparency of the watermarked image. Several experiments are conducted to demonstrate the robustness property of this watermarking scheme. Furthermore, it is worthwhile to mention that the proposed scheme is resistant to the counterfeiting attack. In addition, comparison of the proposed scheme with other state-of-the-art watermarking schemes based on DWT shows that our scheme is more effective.

Key words: Digital watermarking, copyright protection, discrete wavelet transform, human visual system, quantization, digital signature, digital timestamp

Introduction

Internet is very popular in recent years because its convenience in conducting daily works and business affairs. Unfortunately, new threats to information security are growing at the same time, for example, illegal distribution, malicious manipulation and ownership authentication. While some of the threats can be prevented by cryptosystems, the copyright protection of digital documents, e.g., audio, video and images, can not. This is because the cryptographic tools can only protect the digital documents during transmission. Once the encrypted messages are decrypted, the decrypted documents, which is identical to the origin ones, may be illegal duplicated or redistributed. Hence, we can't protect the copyright of digital documents any more. Moreover, there is no distortion allowed in an encrypted message. This requirement is, however, not always necessary for a digital documents. A manipulated document with little distortion is still acceptable, since human visual systems are not sensitive enough to detect small levels of distortion.

Digital watermarking

In order to overcome the copyright-protection issue, digital watermarking techniques have received considerable attentions and have been under development for several years. A watermark can be a logo or a binary sequence which can indicate a legal owner uniquely.

We embed the watermark into the digital documents to be protected and, subsequently, use extraction or detection algorithms to show the ownership of the digital documents.

Watermarking schemes can be robust or fragile. A robust watermarking scheme is designed to resist to malicious or intentional distortions, such as general image processing and geometric distortions; while a fragile watermarking scheme is designed for the purpose of authentication and verification. We can also classify watermarking schemes according to operation domain: the spatial domain and frequency domain. The simplest watermarking technique embeds a watermark directly into the spatial domain by modifying the least-significant bit plane of the original image. By modifying the Least Significant Bit (LSB), the human eyes are not sensitive enough to identify the difference because the change in intensity of certain pixels is slight. One advantage of such watermarking techniques is the low computational complexity. However, such techniques are weak against image processing operations and geometric distortions.

The watermarking scheme based on the frequency domains can be further classified into the Discrete Cosine Transform (DCT) (Cox ***et al.***, 1997; Hsu and Wu, 1999), discrete wavelet transform (DWT) (Hsu and Wu, 1998; Tsai ***et al.***, 2000; Barni ***et al.***, 2001) and Discrete Fourier Transform (DFT). There are many DCT-based schemes proposed. But more and more researches focus on the DWT approaches because DWT is used in the upcoming JPEG2000 standard. The watermarking scheme proposed in this paper is also based on the DWT approachs.

An effective watermarking scheme should process the following essential properties:

Robustness

The watermarks embedded into images must be hard to destroy under the common image processing operations such as image compression, blurring, noising, sharpening, etc. That is, a watermarking scheme must be robust against those attacks under the assumption that the distorted images are still acceptable.

Transparency

In order to preserve commerce value, the quality of watermarked images should be perceptibly identical to the original ones in the human eyes.

Unambiguity

The watermark extracted or detected from the watermarked image must identify the ownership unambiguously. This implies that the error of the extracted watermark must be as low as possible. For example, (Craver ***et al.***, 1997) argued that a pirate can confuse the ownership by simply embedding an illegal watermark to the watermarked image. This is the problem of multiple claims of ownership (also called the counterfeiting attack).

Security

According to Kerckhoff's principle, the security of a cryptosystem must not depend on keeping the cryptographic algorithm secret (Schneier, 1996). Instead, security depends only upon keeping the corresponding key secret. Similarly, the watermarking algorithm must be public while the embedded watermark is irremovable without knowing the secret key.

Blindness

For more effective applications, a good watermarking scheme should not need the original image in order to detect or extract the embedded watermark in extracting algorithm. The main advantage is that the owner doesn't require extra space to store/maintain more and more original images. This is an essential property for the watermarking techniques to be practical.

Today, very few of the proposed watermarking schemes meet all these requirements, especially against counterfeiting attacks.

We note that even a scheme satisfying all the essential properties mentioned above is still not enough. Craver claimed that robustness is necessary but not sufficient to guarantee security (Craver ***et al.***, 1998). Moreover, Katzenbeisser claimed that watermarking alone is not sufficient to resolve rightful ownership of digital data; a protocol relying on the existing public-key infrastructure which is also used for digital signatures is necessary (Katzenbeisser, 2001). In watermarking schemes, attacks are considered successful if any attacker can argue against the existence of the watermarks in the watermarking extraction stage. Since the processor of an image has the complete control of watermark embedding and extracting, the counterfeiting attack, first demonstrated by (Craver ***et al.***, 1997), is possible. There are two methods to thwart this attack: non-invertibility and digital timestamp. The former (such as Craver, 1997; Qian and Nahrstedt, 1998) tries to propose non-invertibility watermarking schemes. Unfortunately, the user loses the freedom to choose the watermark logo. The latter (such as Voyatzis and Pitas, 1999; Chang, ***et al.***, 2002) introduces a trusted third party to issue the digital timestamp indicating the data and time of generating the watermarked image. Since the public key infrastructure is well-defined, the timestamp technology is ready to solve the counterfeiting attack.

Related work

One of the most cited watermarking methods is proposed in (Cox ***et al.***, 1997). Cox adopted the watermark $W = w_1, w_2, \ldots, w_n$, where w_i is selected according to $N(0,1)$, where $N(\mu,\sigma^2)$ denotes a normal distribution with mean μ and variance σ^2. According to their algorithm, the watermark sequence w_i is placed on the n highest magnitude DCT coefficients:

$$I_i' = I_i + \alpha I_i m_i = I_i (1+ \alpha m_i),\ I = 1,2,\ldots,n$$

where I_i, I_i' denote frequency components of the original and watermarked image, respectively and α is a scalar factor, for example, the authors choose α to be 0.1 in their experiments. Watermark detection is done by evaluating the similarity between the extracted watermark W^* and the original watermark W:

$$\mathrm{sim}(W, W^*) = \frac{W.W^*}{\sqrt{W^*.W^*}}$$

Since the watermark is embedded in the highest magnitude DCT coefficients which correspond to the most perceptually significant regions of the original image, it is robust against various common image processing and geometric distortion. However, the word of "perceptually significant" is somewhat subjective.

In (Hsu and Wu, 1998; Hsu and Wu, 1999), discrete cosine/wavelet transform approaches were proposed to embed a binary watermark, a visually recognizable pattern, by modifying the middle-frequency coefficients. These schemes are resistant to common image processing; but not to geometric distortions. The main common drawback in (Cox ***et al.***, 1997; Hsu and Wu, 1998; Hsu and Wu, 1999) is that they need the original images to detect/extract the watermarks.

In (Tsai ***et al.***, 2000), the authors proposed a blind watermarking scheme based on DWT. A visually recognizable watermark is embedded into the frequency components besides the lowest subband. Note that the selected coefficients to be embedded are quantized by a constant factor. The scheme does not require the original image to extract the watermark. However, only JPEG compression is conducted as the experiment for evaluating robustness. Moreover, a constant quantization factor is not suitable for the full image. For example, textured areas allow larger quantization factor since they are tolerant to large modification. On the contrary, for smooth areas, only smaller factors are allowable.

There are more and more watermarking schemes adopting the HVS model. A blind watermarking scheme based on DWT and the HVS model is proposed in (Barni ***et al.***, 2001). It embeds the watermark, consisting a binary pseudorandom sequence, into the three largest detail subbands (the highest frequency components). To avoid affecting robustness, high level of watermark strength is necessary. The watermark strength is modulated according to the local characteristics based on the HVS model. This scheme is resistant to JPEG compression, DWT-based compression, cropping and morphing in their experiments. It is worth noticing that an attacker can easily remove the watermark by discarding the highest subbands in DWT domain without seriously distorting the watermarked image.

Overview of our approach

Without loss of generality, we consider only the robust watermarking schemes for digital images although they are suitable for other document formats, such as audio and video. In fact, almost all frequency-domain watermarking schemes try to insert the watermark into the middle-range frequency components. This is because that the low frequency components are sensitive to the human visual system. However, the lower frequency components can sustain under various attacks, which is a good property for a watermarking scheme to be robust.

In contrast to wavelet-based watermarking schemes, which embed the watermarks into the middle-range frequency components, our proposed scheme embeds the watermark into the lowest frequency components and adopt the HVS model to reduce the incurred distortion. Hence, it has both of the advantages of robustness and transparency.

Although cryptographic techniques are not suitable for data allowing distortion, they are the best tools for providing well-defined security services such as secrecy and authentication (Surety web site; VeriSign web site). In this paper, we will introduce cryptographic tools into the watermarking process.

Based on the above ideas, we propose a watermarking scheme that not only meets all of the watermarking properties mentioned above, but also has the following properties:

Belief

To see is to believe. The watermark ought to be treated as a recognizable pattern in certain applications such that a human observer can unambiguously judge the extracted watermark. In fact, people are pleased to use a meaningful and visually recognizable pattern as their copyright logo. It seems to us that an extractable watermarking scheme is better than the detectable one. This is because that the latter cannot completely satisfy all of the people's requirements.

Resistant to counterfeit attacks

A private can confuse the ownership by simply appending a forged watermark into the already-watermarked image. Hence, there are two different watermarks in the watermarked image. It fails while verifying the identical ownership of the watermarked image. Our scheme can resist the counterfeit attack.

DWT-based

There are several reasons to adopt the DWT domain:

- The DWT domain is the kernel technique of JPEG-2000.
- The DWT is highly integrable with JPEG-2000.
- The goal to be robust against JPEG-2000 compression is achieved.
- The DWT based approaches usually produce watermarked images with the best tradeoff between transparency and robustness while the DFT and DCT domain approaches have blocking artifacts (Loo and Kingsbury, 2000).

Organization of the paper

The rest of this paper is organized as follows. In Section 2, we briefly introduce the discrete wavelet transform, the HVS model, the concept of digital signature and timestamp. Section 3 describes the embedding and extracting algorithms. In Section 4, we present the experimental results. Further discussions and conclusions are given in Section 5 and 6, respectively.

Preliminaries

Discrete wavelet transform

The discrete wavelet transformation is a tool that can transfer a 2-D still image from spatial domain to frequency domain (Lewis and Knowles, 1992; Shapiro, 1993). After 1-scale level DWT, an image I can be decomposed into four subbands, as shown in Fig. 1. I_0^0, I_0^1 and I_0^2 are the higher resolution subbands and the less sensitive components.

I_0^3 is the lowest resolution subband, which contains the most of energy in the image. If we apply DWT to the lowest subband I_0^3, then I_0^3 will be further decomposed into another four subbands I_1^0, I_1^1, I_1^2 and I_1^3. Then we can obtain 2-scale transformation. Repeating the same way, we also can obtain N-scale level discrete wavelet transformations for any N>2. Without loss of generality, we assume that the 4-scale level wavelet transformation is performed.

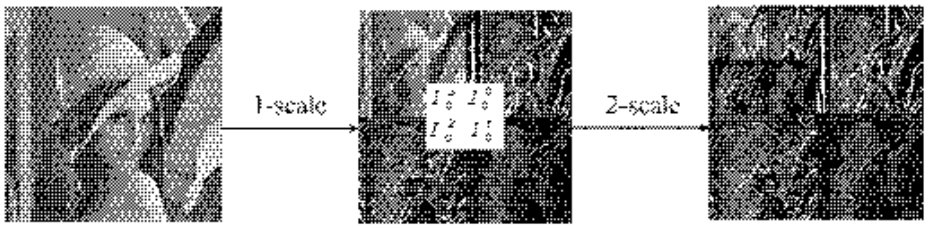


Fig. 1: Discrete Wavelet Transform

HVS model

In order to design a more transparent and robust watermarking scheme, we apply the HVS model to analyze the DWT coefficients. In (Barni *et al*., 2001) and (Lewis and Knowles, 1992), the authors exploited three visual psychophysics functions to analyze the DWT coefficients.

Frequency

The human visual system is less sensitive to distortions in the high frequency components.

Luminance

The human visual system is less sensitive to distortion in the area with higher or lower brightness.

Texture masking and edge proximity

The human visual system is less sensitive to distortion if there is high variation in local coefficients, while the human visual system is more sensitive to distortion near the edges.

Throughout the paper, the following notations are used:

L:l ∈ {0, 1, 2, 3} , DWT resolution level,

θ:θ ∈ {0, 1, 2, 3}, DWT orientation,

· ¨ ¨¨ : Subband at level l, orientation θ and DWT coefficient (l, j),

F(l, θ): Frequency analysis at level l and orientation,

L(l, I, J): Luminance analysis at level l and DWT coefficient (i,j),

T(l, I, j): Texture masking and edge proximity analysis at level l and DWT coefficient(l, j),

$q_1^{\theta}(i,j)$: The quantization factor of DWT coefficients at level l and orientation θ.

Before obtaining the quantization factor, the three kernel items are computed as follows:

$$F(l,\theta)=\begin{Bmatrix}\sqrt{2} & \text{if } \theta=1\\ 1 & \text{otherwise}\end{Bmatrix}\cdot\begin{Bmatrix}1.00 & \text{if } l=0\\ 0.32 & \text{if } l=1\\ 0.16 & \text{if } l=2\\ 0.10, & \text{if } l=3\end{Bmatrix} \quad (1)$$

$$L(l,i,j)=1+\frac{1}{256}l^{3}\left(1+\left[\frac{i}{2^{3-\frac{1}{l}}}\right],1+\left[\frac{j}{2^{3-\frac{1}{l}}}\right]\right) \quad (2)$$

$$T(l,i,j)=\sum_{k=0}^{3-l}\frac{1}{16^k}\sum_{\theta=0}^{2}\sum_{r=0}^{l}\left[I_{k+1}^{\theta}\left(y+\frac{i}{2^k},x+\frac{j}{2^k}\right)\right]^2.\mathrm{var}\left\{I_3^3\left(1+y+\frac{i}{2^{3-l}},\ 1+x\frac{j}{2^{3-l}}\right)\right\}_{x=0,1,y=0,1} \quad (3)$$

where Var(.) is a variance function.
Finally, we compute the quantization factor

$$q_l^0(i,j)=F(l,\theta):L(l,i,j,).T(l,i,j,)^{0.2} \quad (4)$$

$q_l^0(i,j)$ is subsequently used as the quantization step for a low DWT coefficient at location (l, j).

Digital signature and digital time-stamping

Digital time-stamping is a technique used to verify whether digital data was created or signed at a certain time. Because the date and time on a computer is easily alterable, the timestamp technique will be done by a trusted Certification Authority (CA). A digital signature is a piece of information based on both the signed data and the signer's private key (Schneier, 1996). Anyone can verify the signature with the corresponding public key of the signer. The digital signature technique provides the proof of identifying the rightful author. Hence, a combination of the digital timestamp and digital signature provide an effective solution to protect digital documents.

Proposed scheme

The lowest frequency subband contains the most energy in an image. Hence, for more robust, we should embed the watermark in the lowest frequency subband. However the lowest frequency subband is too sensitive to modify coefficients too much. The HVS model is adopted to embed a watermark into the lowest frequency subband without introducing too much distortion. Therefore, our scheme can achieve both robustness and transparency. The embedding and extracting algorithms are illustrated in Fig. 2 and 3, respectively. We describe them in details in the following subsections.

Embedding algorithm

Assume that the original image I is a gray-level image with 8-bit per pixel and the watermark W is a binary image. They are defined as follows:

$$I=\{I_{i,j}|0\le I_{i,j}\le 255,0\le i<I_w 0\le j<I_H\} \quad (5)$$

where I_W and I_H is the width and height of I, respectively.

$$W=\{W_{i,j}|W_{i,j}\in\{0,1\}0\le i<W_w 0\le j<W_H\} \quad (6)$$

where W_w and W_H is the width and height of W, respectively.
Without loss of generality, let I_W, I_H, W_W and be powers of 2.

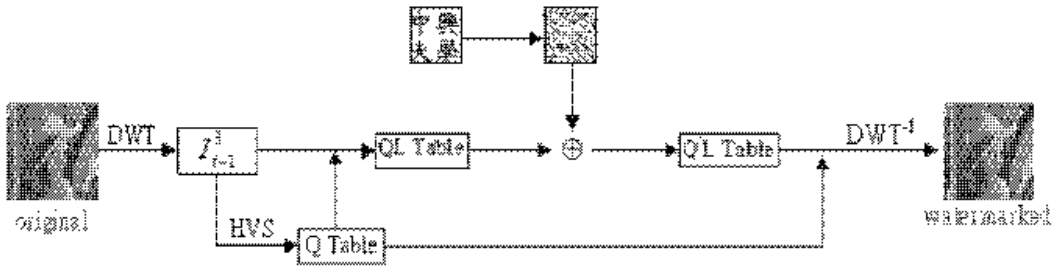


Fig. 2: Flow chart of the embedding algorithm

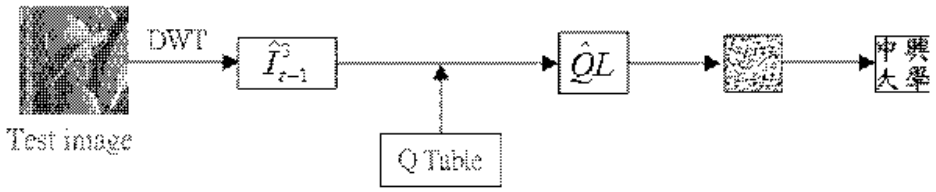


Fig. 3: Flow chart of the extracting algorithm

Step 1: Wavelet transforming original image

The original image I is decomposed by t-scale wavelet transform to obtain the lowest frequency subband. I^3_{t-1} The size of I^3_{t-1} is equal to the one of the watermark W. Here, t is predetermined according to the tradeoff between the transparency and robustness.

Step 2: Perceptual weighing

Apply the functions mentioned in Section 2.2 to analyze the DWT coefficients of I^3_{t-1}. We can obtain a quantization table Q.

$$Q = [Q_{i,j}] Q_{i,j} = round[(\alpha . q^3_{t-1}(i,j)) 0 \le i < W_w, 0 \le j < W_H], \quad (7)$$

where α is a positive constant and pre-determined by the owner. Q is kept as secret information. Note that we can use α to control the tradeoff between robustness and transparency.

Step 3: Quantifying the DWT coefficients

After obtaining the quantization table Q, we can quantify the DWT coefficients by computing

$$QL = [QL_{i,j}] QL_{i,j} = round\left(\frac{I^3_{t-1}(i,j)}{Q_{ij}}\right), 0 \le i < W_w, 0 \le j < W_H \quad (8)$$

Step 4: Embedding watermark

Before embedding, the watermark W should be randomly permutated into W' by a seed s and pseudo-random permutation function RPF (.), where

$$W' = \{W'_{i,j} | W'_{i,j} = RPF(W_{i,j}), 0 \le i, i' < W_w, 0 \le i, j' < W_H\} \quad (9)$$

The seed s is the secret key.
Consequently, we compute

$$Q'L_{i,j} = \begin{cases} QL_{i,j}, & \text{if } (QL_{i,j} \bmod 2) = W_{i,j} \\ QL_{i,j} \pm 1, & \text{otherwise} \end{cases} \quad (10)$$

The operation of "mod 2" turns the odd/even value of the quantified coefficient into that of the watermark bit. In the second case, if $I^3_{t-1}(i,j) \ge QL_{i,j} \cdot QL_{i,j}, Q'L_{i,j}$, is equal to $Ql_{i,j}$ + 1 ; else $Q'L_{i,j}$ is equal to $Q'L_{i,j}$ - 1.

Step 5: Inversing Quantization and DWT to obtain watermarked image

We can inverse the DWT coefficients to generate the watermarked image by inversing the quantization operation as follows:

$$I'^3_{t-1} = \{I'^3_{t-1}(i,j) | I'^3_{t-1}(i,j) = Q'L_{i,j} \times Q_{i,j}, 0 \le i < W_w, 0 \le j < W_H\} \quad (11)$$

Finally, the seed s, the DWT scale parameter t, the scalar factor $\alpha$and the quantization table Q, are kept as the verification keys.

Step 6: Digital signing and time-stamping

The verification keys are signed by the owner with the existing digital signature tools. Let $DS = Sign_{OPK}(t, s, \alpha, Q)$, where $Sign_{OPK}(.)$ denotes a digital signature function using the owner's private key OPK. Subsequently, the owner sends the signature DS to a trusted Certification Authority (CA) in a secure way. The CA digitally time-stamps it by appending the date and time received. And compute $TS = TS_{CAPK}(DS)$, where TS_{CAPK} denotes the time-stamping technology with the private key CAPK. After receiving the TS, the owner stores it together with the verification keys.

Extracting algorithm

In order to verify the copyright of the test image, the verifier checks the validation of the timestamp TS and the signature DS using the CA's public key and the owner's public key, respectively. If it succeeds, the verification holds and the extraction operation goes on. Otherwise the algorithm stops and returns failure. In the extracting algorithm, the scheme does not need the original image. Only parameters t, s and Q need to be used. Let $\hat{I}$ be the test image and $\hat{W}$ be the extracted watermark.

Step 1: Wavelet transforming test image

The test image $\hat{I}$ is also decomposed by t-scale wavelet transformation to obtain the lowest frequency subband $\hat{I}^3_{t-1}$.

Step 2: Quantifying the DWT coefficients

The coefficients in $\hat{I}^3_{t-1}$ is then quantified by computing

$$\hat{Q}L=\{\hat{Q}L_{i,j}|\hat{Q}L_{i,j}=\text{round}(\frac{\hat{I}^3_{t-1}(i,j)}{Q_{i,j}}),0\le i<W_w,0\le j<W_H\} \quad (12)$$

Step 3: Extracting watermark

Next, the temporary binary table P is generated.

$$P=\{P_{i,j}|P_{i,j}=\hat{Q}L_{i,j}\bmod 2,0\le i<W_W,0\le j<W_H\} \quad (13)$$

Finally, the extracted watermark $\hat{W}$ is obtained by inversing the permutation function in Eq. (9) according to the seed s as follows.

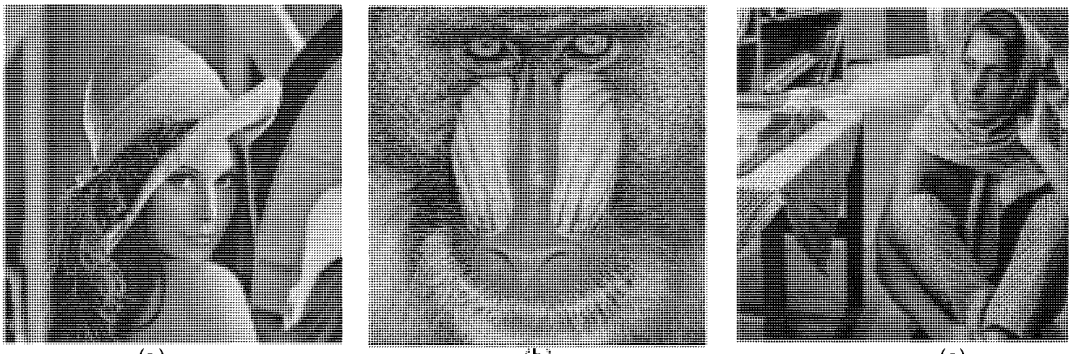
$$\hat{W}=\{\hat{W}_{i,j}|\hat{W}_{i,j}=RPF^{-1}(P_{i',j'})\ 0\le i,i'<W_W,0\le j,j'<W_H\} \quad (14)$$

Results and Discussion

To demonstrate the feasibility of the proposed scheme, we conduct several experiments. Three "classical" images Lena, Baboon and Barbara shown in Fig. 4(a)-(c), are employed as the original images I and Fig. 4(d) "National Chung-Hsing University", represented by Chinese characters is the watermark W. The original images are 256 gray level images with the size of 512 x 512 pixels and the watermark is a 32 x 32 pixels binary image. The original images are 4-scale level wavelet transformed and the size of the lowest frequency subband is 32 x 32 pixels, i.e., t=4. α is set to be 20, 24 and 28 in the following experiments, respectively.

(a) (b) (c) (d)

Fig. 4: (a) Lena, (b) Baboon, (c) Barbara, (d) Watermark



(a) (b) (c)

Fig. 5: The watermarked images under the parameter α:(a) the watermarked Lena (PSNR= 39.5 dB); (b) the watermarked Baboon (PSNR= 39.9 dB); and (c) the watermarked Barbara (PSNR= 39.0 dB)

We evaluate the quality between the attacked image and the original image using their Peak Signal-to-Noise Ratio (PSNR). The PSNR formula is defined as follows:

$$PSNR=10\log_{10}\frac{E_{max}^2\times I_H\times I_W}{\sum[I_{i,j}-I'_{i,j}]^2},$$

where I_H and I_W are the image's height and width, respectively. $I_{i,j}$ is the original value of the coordinate (i,j). $I'_{i,j}$ is the distorted value of the coordinate (i,j). E_{max} is the largest energy of the image pixels (e.g. E_{max} =255 for 8 bits/pixel). Generally speaking, when the PSNR value of a watermarked image is greater than 30 dB, the quality is still acceptable to the human eyes (Chen *et al.*, 1998). The watermarked images and their corresponding PSNR values are shown in Fig. 5, respectively.

The watermark retrieval rate is computed as the ratio of the number of accurate pixels to all of the pixels of the retrieved watermark. That is, the retrieval ratio is defined as follows:

$$R=\frac{\sum_{i=1}^{W_W}\sum_{j=1}^{W_W}\overline{W_{i,j}\oplus\hat{W}_{i,j}}}{W_W\times W_H}$$

where the notations are the same as Eq. (6) and (14).

The experiments are divided into three classes:

Various JPEG compressions

A series of experiments of different JPEG compression quality is conducted to show the robustness of the scheme under JPEG compression. Table 1, 2 and 3 show the PSNR values of the attacked images and the high corresponding ratio values of the extracted watermarks under

different and JPEG quality. Not surprisingly, while the JPEG quality is retained 10%, the watermark-extraction ratio is still 94.0%. At the same time, the PSNR value of the watermarked image is still above 36 dB.

Several attacks including image processing and geometric distortion

- JPEG compression: A JPEG compression version with parameters of 10% quality and 0% smoothing. The retrieved watermarks are recognizable (Table 4).
- Image Blurring: We blurred the images, in which their PSNR values are reduced under 30 dB. However, the retrieved watermarks are still recognizable to the human eyes (Table 4).
- Image Sharpening: We sharpened the images, in which their PSNR values are reduced under 30 dB. The retrieved watermarks are clearly recognizable (Table 4).
- Image Noising: We added Gaussian noise to the images, in which its PSNR value is reduced to about 30 dB. The retrieved watermarks are recognizable (Table 4).
- Image Scaling: We scaled the images from 512 × 512 to 128 × 128 pixels. Then the scaled images are rescaled back to 512 × 512 pixels. The PSNR values in those images are reduced far under 30 dB. It is clear that some details have been lost in the scaling process; however, the retrieved watermarks are still recognizable (Table 4).
- Image Cropping: The cropped version discards the left-top corner of the images. Since the binary watermark is pseudo-randomly permutated, the malicious attacker could not remove the watermark by cropping an area. The reason is that the distorted pixels of the watermark are spread around and thus not concentrated in a specific area. After cropping, the PSNR value declines to 11 dB. Fortunately, we can still extract the watermark (Table 4).

Re-computing the quantization table

If the quantization table is not stored, it can be reconstructed from the test image directly. This saves the storage space at the cost of reducing the watermark-extraction ratio. Table 5 shows the extraction ratio by reconstructing the quantization table using the watermarked image.

To summarize, the experimental results show that the extracted watermarks are recognizable. That is, the extracted ratios are very high, even if the watermarked images are seriously distorted.

The experimental results demonstrate that our watermarking scheme has the following properties:

Discussions and Analysis

If a bigger α is chosen, then the scheme is more robust. However, the distortion increases at the same time. Hence, α can be used to control the tradeoff between robustness and transparency. Although low resolution subband is sensitive to alteration, we adopt the HVS model to improve the quality of the watermarked image. Therefore, the owner can maximize the watermark robustness while satisfying the transparency property. The experimental results show that the proposed scheme is robust under various image processing and geometric distortion attacks, such as JPEG compression, blurring, noising, sharpening, scaling and cropping.

Security

The security of the proposed scheme is based on the verification keys including the DWT scaling parameter t, quantization scalar factor, α the permutation seed and the quantization table. Moreover, the integrity of the verification keys and the timestamp of the creation of the watermark are guaranteed under the digital signature and timestamp mechanisms, respectively. It seems possible that an attacker can construct the quantization table. Since the attacker has no knowledge of the secret information t and, α he can not construct the exact quantization table. Furthermore, there are several different HVS models to select in practice. Even through the attacker can guess a quantization table and try to remove the embedded watermark; he still can not be sure whether the embedded watermark is already removed or not. Another possible attack is to disturb the extraction by modify the odd/even value of the quantified coefficients in the DWT domain according to randomly selecting a half of the coefficients. Note that this disturbance attack will distort seriously the watermarked image if the attacker guesses the larger quantization values; and the extraction is not obviously disturbed if the attacker guesses the smaller quantization values. Based on the two assumptions: watermarking is to deter the behavior of copying and attacking must not loss the commercial value of the image, the proposed scheme is still secure. Because the attacker does not know the secret parameters t, s and α, he hardly distorts the embedded watermark in the watermarked image while preserving the quality of the commercial value.

Table 1: The PSNR values of JPEG-compression-attacked Lena images and the corresponding ratio values (%) of the extracted watermarks

		20		24		28	
Image PSNR		39.5		37.9		36.5	
JPEG Quality	JPEG PSNR	Extract Rate(%)		Extract Rate(%)		Extract Rate(%)	
100	58.5		100		100		100
90	43.1		100		100		100
80	40.7		100		100		100
70	39.4		100		100		100
60	38.3		100		100		100
50	37.6		100		100		100
40	36.7		100		100		100
30	35.7		100		100		100
20	34.1		99.1		99.9		100
10	31.2		85.3		90.2		94.0

Table 2: The PSNR values of JPEG-compression-attacked Baboon images and the corresponding ratio values (%) of the extracted watermarks

		20		24		28	
Image PSNR		39.9		38.4		37.1	
JPEG Quality	JPEG PSNR	Extract Rate(%)		Extract Rate(%)		Extract Rate(%)	
100	58.4		100		100		100
90	37.1		100		100		100
80	32.6		100		100		100
70	30.5		100		100		100
60	29.1		100		100		100
50	28.2		100		100		100
40	27.4		100		100		100
30	26.4		99.9		100		100
20	25.3		98.0		100		100
10	23.4		87.4		95.4		96.7

Table 3: The PSNR values of JPEG-compression-attacked Barbara images and the corresponding ratio values (%) of the extracted watermarks

		20		24		28	
Image PSNR		39.0		37.9		36.2	
JPEG Quality	JPEG PSNR	Extract Rate(%)		Extract Rate(%)		Extract Rate(%)	
100	58.4		100		100		100
90	40.2		100		100		100
80	36.9		100		100		100
70	35.0		100		100		100
60	33.6		100		100		100
50	32.5		100		100		100
40	31.5		100		100		100
30	30.2		99.9		100		100
20	28.3		99.2		100		100
10	25.7		89.3		94.7		96.8

Table 4: Several attacks including image processing and geometric distortion

	JPEG(10%)	Blurring	Sharpening	Noise	Scaling	cropping
Lena	31.2 dB	32.0 dB	28.8 dB	30.2 dB	29.8 dB	11.3 dB
						
	94.0%	99.2%	99.4%	99.3%	99.9%	86.9%
Baboon	23.4 dB	24.2 dB	23.8dB	29.7 dB	21.1 dB	11.1 dB
						
	96.7%	99.7%	99.9%	100%	98.7%	86.8%
Barbara	25.7 dB	26.1 dB	26.0dB	30.4 dB	23.6 dB	11.3 dB
						
	96.8%	99.8%	98.2%	99.8%	99.9%	90.4%

Table 5: With/without storing the quantization table

	No Attack	JPEG(20)	Blurring	Sharpening	Noising	Scaling	Cropping
With Q table	 100%	 99.1%	 99.2%	 99.4%	 99.3%	 99.9%	 86.9%
Without Q table	 97.0%	 95.5%	 96.3%	 95.8%	 96.6%	 96.9%	 78.3%

Table 6: Comparisons of various proposed methods based on DWT

	Hsu, 1998	Tsai, 2000	Barni, 2001	Proposed Method
Blindness	No	Yes	Yes	Yes
Quantization coefficients	No	Constant	Variable	Variable
HVS model	No	No	Yes	Yes
Middle/low	Middle-low	Middle-low	High	Low
Counterfeit attack	Undescribed	Undescribed	Undescribed	Timestamp
Extraction/ Detection	Extraction	Extraction	Detection	Extraction
Visual recognizable pattern	Yes	Yes	No	Yes
Robustness	Compression	Compression	Compression	Compression
	Blurring		Cropping	Blurring
			Morphing	Noising
				Sharpening
				Scaling
				Cropping

Unambiguity

Regarding the experimental results, the retrieval ratios are very high. The watermark is a visually recognizable pattern that is convincingly related to the owner of the image.

Blindness

The watermark extraction phase does not require the original image to retrieve the logo. In practice, this is an essential property of the watermarking scheme. Hence, the owner is not required to spend extra space to store the original image but only the secret information including the corresponding timestamp. In fact, the timestamp may be distributed with the image or stored as the special field in the head of image file format.

Belief

In many proposed watermarking schemes, the watermark is a copyright owner ID or a sequence number selected according to a normal distribution. Subsequently, if the dispute occurs, the detector is used to detect whether the copyright watermark is embedded into the image in the court and then yields a "yes" or "no". If the similarity is greater than a predetermined threshold value, we say that the watermark is indeed embedded into the image and vice versa. However, how to decide a proper threshold value is very difficult. A small threshold will presume the existence of a watermark if not so. Oppositely, a lager threshold will reject the existence of a watermark although it is indeed embedded. The extraction of the embedded visually recognizable watermark convinces us of its existence. This is because people have more belief in judging the retrieved visually recognizable watermark than a detected result.

Counterfeit attacks

An attacker can claim his ownership of the image by forging the verification keys (t', s', α' Q'), Digital Signature (DS') and Time Stamp (TS'). However, the date and time of forged TS' is always later than that of the owner. Hence, our scheme is secure and resistant to copy attacks.

Space requirement

The size of the quantization table is $1 \times 32 \times 32$ Bytes = 1KB, which is only 1/256 of the original image. In contrast to intellectual property loss, the extra cost for the secret key is worthwhile. If the storage is limited, we have the alternative of not storing the quantization table. Table 5 shows the extraction ratio, still very high, by re-reconstructing the quantization table using the watermarked image.

Generalization and comparisons

Gray-level watermark

Both a binary and a gray-level image can be regarded as a watermark in our scheme. A gray-level image, for example a 256-gray image, is decomposed into 8 bitplanes, with each bitplane considered as a binary image. The gray-level watermark is more feasible than the binary one and has wide applications in practice. A gray-level watermark has a greater chance of survival than

a binary one under attacks because the former can always preserve a certain degree of "contextual relationship". The more significant bitplane of a gray-level watermark is embedded into the lower components of the original image. On the contrary, the less significant bitplane is embedded into middle-low components while the least bitplane may be discarded alternatively. Hence, the gray-level watermark has a greater chance of survival.

Quantization

In our scheme, each component in lowest subbands is quantized by different quantization coefficient respectively. Namely, the embedding operation is to modify the component with different watermark strength by weighting parameter. Based on the properties of local characteristics, maximum robustness, invisibility and security are satisfied simultaneously.

Color image

Color images are also suitable if they are transformed into RGB or YUV domains. This is because our scheme is not depended on the specific image format.

Table 6 demonstrates the comparison of our scheme with other state-of-the-art watermarking schemes based on DWT. It shows that our approach is more effective.

For the sake of high robustness, the proposed watermarking scheme embeds the watermark into the low frequency components by modifying them based on the HVS model. The operations of perceptual weighing provide the balance between robustness and transparency. We demonstrate experimentally that this new scheme is robust against different attacks. The extracted watermarks, extracted without involving the original image, are unambiguously recognizable even under low JPEG compression quality, i.e., the high compression ratio. In addition, the experiments also show that our scheme is resistant to the blurring, noising, sharpening, scaling and cropping attacks. It is worthwhile to mention that our scheme is also resistant to the counterfeiting attacks.

References

Barni, M., F. Bartolini and A. Piva, 2001. Improved Wavelet-based Watermarking Through Pixel-Wise Masking. IEEE Transactions on Image Processing, 10: 783-791.

Chang, C.C., K.F. Hwang and M.S. Hwang, 2002. Robust Authentication Scheme for Protecting Copyrights of Images and Graphics. IEE Proc.-Vis. Image Signal Process, 149: 43-50.

Chen, T.S., C.C. Chang and M.S. Hwang, 1998. A Virtual Image Cryptosystem Based upon Vector Quantization. IEEE Transactions on Image Processing, 7: 1485-1488.

Cox, I.J., J. Kilian, J.F.T. Leighton and T. Shamoon, 1997. Secure spread spectrum watermarking for multimedia. IEEE Transactions on Image Processing, 6: 1673-1687.

Craver, S., N. Memon, B.L. Yeo, and M. Yeung, 1997. Can invisible watermarks resolve rightful ownership? Proc. SPIE Storage and Retrieval for Still Image and Video Databases V., pp: 310-321.

Craver, S., B.L. Yeo and M. Yeung, 1998. Technical Trials and Legal Tribulations. COMMUNICATIONS OF THE ACM., 41: 45-54.

Hsu, C.T. and J.L. Wu, 1998. Multiresolution Watermarking for Digital Images. IEEE Transactions on Circuits and System—II: Analog and Digital Signal Processing, 45: 1097-1101.

Hsu, C.T. and J.L. Wu, 1999. Hidden Digital Watermarks in Images. IEEE Transactions on Image Processing, 8: 58-68.

Katzenbeisser, S., 2001. On the Design of Copyright Protection Protocols for Multimedia Distribution Using Symmetric and Public-Key Watermarking, in 12th International Workshop on Database and Expert systems Applications, 5th International Query Processing and Multimedia Issues in Distributed Systems Workshop, IEEE Computer Society Press, pp: 815-819.

Kutter, M., S. Voloshynovskiy and A. Herrigel, 2000. The Watermark Copy Attack. Proceedings of SPIE: Security and Watermarking of Multimedia Contents II, vol. 3971, San Jose, California.

Lewis, A.S. and G. Knowles, 1992. Image Compression Using the 2-D Wavelet Transform, IEEE Transactions on Image Processing, 1: 244-250.

Loo, P. and N.G. Kingsbury, 2000. Digital Watermarking using Complex Wavelets. IEEE Conference on Image Processing, Vancouver, pp: 29-32.

Qian, L. and K. Nahrstedt, 1998. Watermarking Schemes and Protocols for Protecting Rightful Ownership and Customer's Rights, Journal of Visual Communication and Image Representation, 9: 194-210.

Schneier, B., 1996. Applied Cryptography. WILEY, 2nd Ed.

Shapiro, J.M., 1993. Embedded image coding using zerotrees of wavelet coefficients. Signal Processing, IEEE Transactions on Signal Processing, 41: 3445 -3462.

Tsai, M.J., K.Y. Yu and Y.Z. Chen, 2000. Joint Wavelet and Spatial Transformation for Digital Watermarking, IEEE Transactions on Consumer Electronics, 46: 241-245.

VeriSign web site, http://www.verisign.com.

Voyatzis, G. and I. Pitas, 1999. Protecting Digital -image Copyrights: a framework. IEEE Computer Graphics and Applications, 19: 18-24.