An SVD-Based Watermarking Scheme for Protecting Rightful Ownership

Ruizhen Liu and Tieniu Tan, *Senior Member, IEEE*

***Abstract*—Digital watermarking has been proposed as a solution to the problem of copyright protection of multimedia documents in networked environments. There are two important issues that watermarking algorithms need to address. First, watermarking schemes are required to provide trustworthy evidence for protecting rightful ownership. Second, good watermarking schemes should satisfy the requirement of robustness and resist distortions due to common image manipulations (such as filtering, compression, etc.). In this paper, we propose a novel watermarking algorithm based on singular value decomposition (SVD). Analysis and experimental results show that the new watermarking method performs well in both security and robustness.**

I. Introduction

THE advent of the Internet and the wide availability of computers, scanners, and printers make digital data acquisition, exchange, and transmission simple tasks. However, making digital data accessible to others through networks also creates opportunities for malicious parties to make salable copies of copyrighted content without permission of the content owner. Digital watermarking techniques have been proposed in recent years as methods to protect the copyright of multimedia data [1]–[4].

In general, an effective watermarking scheme should satisfy the following basic requirements.

1) *Imperceptibility*: the perceptual difference between the watermarked and the original documents should be unnoticeable to the human observer, namely, watermarks should not interfere with the media being protected.
2) *Trustworthiness* [5]–[8]: a satisfactory watermarking scheme should also guarantee that it is impossible to generate counterfeit watermarks and should provide trustworthy evidence to protect the rightful ownership.
3) *Robustness* [9]–[12]: given a watermarked document, an unauthorized party should not be able to destroy the watermark without also making the document useless, that is, watermarks should be robust to common signal processing and intentional attacks. In particular, they should still be detectable or extractable even after common signal processing operations have been applied to the watermarked image (such as digital-to-analog, analog-to-digital conversions, resampling, filtering, compression, geometric transformation, cropping, scaling, rotation, etc.).

Most existing watermarking schemes focus on robust means to make the watermark imperceptible rather than on addressing the important issue of how to resolve the rightful ownership of an image embedded with multiple signatures (or watermarks, labels, etc.) [5], [8], [13]. Craver *et al.* [13] were among the first to note that resolving the rightful ownership of watermarked images is a very important issue. They simulated the cases of attacking existing watermarking techniques by providing counterfeit watermarking schemes that can be performed on a watermarked image to allow multiple claims of ownership. Their proposed attack is sometimes known as the IBM ambiguity attack (or protocol level attack). An example is the inversion attack by Craver *et al.* that attempts to discredit the authority of the watermark by embedding one or several additional watermarks, so that it is unclear which was the first authoritative watermark embedded by the IPR owner.

Currently only a few methods have been reported that try to solve the ownership problem. The first such method is probably the one presented by Craver *et al.*. They design a noninvertible scheme which is a modified version of the watermarking method proposed by Cox *et al.* [9]. However, the noninvertibility of their scheme is based on an invalid assumption [8].

The second method is based on time-stamping, as proposed by Wolfgang and Delp [14]. The owner with the earliest timestamp is the true owner of the watermarked documents. However, time-stamping has the disadvantages of requiring ongoing involvement of a third party and being unsuitable for time-sensitive applications. In addition, timestamps can be manipulated by anyone else besides the true owners. This scheme can thus be easily defeated.

The third method is proposed by Zeng *et al.* [15]. Because their watermarking scheme detects the embedded watermark without using the original image, it actually cannot resolve rightful ownership [8]. In their algorithm, the watermarking scheme protects the embedded watermark rather than the ownership of digital images. The key issue is that, because watermark detection in their algorithm does not need original images, an attacker can always create counterfeit original images and claim ownership (see [8] for further discussions).

Qiao *et al.* [8] proposed the fourth scheme to try to solve the ownership problem. They combine their scheme with cryptography and use a standard encryption algorithm (i.e., DES) to generate the randomized watermarks from the original image or video chip. The embedded watermark is a function of the encryption key and the original image. An obvious drawback of their scheme is that the algorithm cannot insert semantically

Manuscript received December 22, 1999; revised May 22, 2001. This work was supported by the Chinese NSF (Grant 59825105) and the Chinese Academy of Sciences. The associate editor coordinating the review of this paper and approving it for publication was Dr. Chung-Sheng Li.

The authors are with the National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, 100080, China (e-mail: liurz@nlpr.ia.ac.cn; tnt@nlpr.ia.ac.cn).

Publisher Item Identifier S 1520-9210(02)01402-5.

1520–9210/02$17.00 © 2002 IEEE

meaningful watermarks and the capacity of watermarks is seriously restricted. In addition, it is a nonlinear watermarking model, so it is difficult to estimate the inserted watermark's energy and capacity and control the visual quality of the watermarked images.

In this paper, we present a new digital image watermarking method based on singular value decomposition (SVD). We will show later that, because SVD is in fact a one-way decomposition algorithm and is an optimal matrix decomposition in a least-square sense, the new method performs well both in resolving rightful ownership and in resisting common attacks.

The rest of this paper is organized as follows. Section II outlines the problem of rightful ownership. Section III describes the principles of SVD and the proposed method. Section IV presents further analysis and discussion on the method. Section V discusses experimental results. The paper is concluded in Section VI.

II. Problem of Rightful Ownership

One of the main purposes of a watermark is to protect the owner's copyright. However, for many existing watermarking schemes, an attacker can easily confuse one by manipulating the watermarked image (or video, audio) and claim that he or she is the legitimate owner [13]. Some watermarking schemes require the original image (or video chip) to perform watermark verification. However even with the presence of the original image, the rightful ownership problem still exists. The class of watermarking schemes that can be attacked by creating a "counterfeit original" is called invertible.

Craver *et al.* [13] define the concept of noninvertibility to address the issue of rightful ownership. Informally, noninvertibility means that it is computationally unfeasible for an attacker to find a faked image and a watermark, such that the pair can result in the same watermarked image created by the real owner. We will review the definition of the class of invertible watermarking schemes in the following.

Denote the original image by A. The owner of A uses the watermark (denoted as W) to make a watermarked image A_W. That is

$$A \oplus W \Rightarrow A_W \tag{1}$$

or, in function representation

$$A_W = E(A, W) \tag{2}$$

where the function $E(\cdot)$ denotes the watermark-embedding algorithm.

An attacker who obtains the published A_W without knowing A, can create a counterfeit watermark W_F and a counterfeit image A_F that satisfy the following equation:

$$A_W = E(A_F, W_F). \tag{3}$$

Then the attacker can use A_F as his "original" to claim the ownership of A_W.

If (3) holds, the watermarking scheme is called invertible. Otherwise, it is called noninvertible [7], [8]. We regard (2) and (3) as the basic equations to define noninvertible watermarking schemes (though there may be some alternative definitions). Detailed descriptions can be found in [8] and [13].

III. SVD-Based Watermarking

SVD is a numerical technique used to diagonalize matrices in numerical analysis. It is an algorithm developed for a variety of applications.

The main properties of SVD from the viewpoint of image processing applications are: 1) the singular values (SVs) of an image have very good stability, that is, when a small perturbation is added to an image, its SVs do not change significantly; and 2) SVs represent intrinsic algebraic image properties.

In this section, we describe a watermark casting and detection scheme based on the SVD.

A. SVD

From the viewpoint of linear algebra, we can observe that a discrete image is an array of nonnegative scalar entries, which may be regarded as a matrix. Let such an image be denoted by A. Without loss of generality, we assume in the subsequent discussions that A is a square image, denoted by $A \in \mathbf{F}^{N\times N}$, where $\mathbf{F}$ represents either the real number domain $\mathbf{R}$ or the complex number domain $\mathbf{C}$. The SVD of A is defined as

$$A = USV^H \tag{4}$$

where $U \in \mathbf{F}^{N\times N}$ and $V \in \mathbf{F}^{N\times N}$ are unitary matrices and $S \in \mathbf{F}^{N\times N}$ is a diagonal matrix.

We notice that the unique property of the SVD transform is that the potential N^2 degrees of freedom (DOF) or samples in the original image now get mapped into

$$\begin{aligned} S &\Rightarrow N \quad \text{DOF} \\ U &\Rightarrow \frac{N(N-1)}{2} \quad \text{DOF} \\ V &\Rightarrow \frac{N(N-1)}{2} \quad \text{DOF}. \end{aligned}$$

totaling N^2 DOF.

SVD has many good mathematical characteristics. For the sake of space, we will not discuss them in this paper. Further details on SVD may be found in [16] and [17].

B. Watermark Casting and Detection

In watermark casting, the SVD of an $N \times N$ image $A \in \mathbf{F}^{N\times N}$ is computed to obtain two orthogonal matrices $U \in \mathbf{F}^{N\times N}$ and $V \in \mathbf{F}^{N\times N}$ and one diagonal matrix $S \in \mathbf{F}^{N\times N}$. Although we assume A to be a square matrix (image) for convenience, other nonsquare images can be processed in exactly the same way. We maintain that this is one of the advantages of the SVD method over some other popular watermarking schemes which cannot directly handle nonsquare matrices [9], [18], [19].

We add a watermark W (also represented as a matrix) into the matrix S and perform SVD on the new matrix $S + aW$ to get U_W, S_W, and V_W ($S + aW = U_W S_W V_W^H$), where the positive constant a is the scale factor which controls the strength of the watermark to be inserted. We then obtain the watermarked image A_W by multiplying the matrices U, S_W, and V^T. That is, if the matrices A and W represent the original image and the watermark, respectively, we obtain the watermarked image A_W through the following three steps:

$$\begin{aligned} A &\Rightarrow USV^H \\ S + aW &\Rightarrow U_W S_W V_W^H \\ A_W &\Leftarrow US_W V^H. \end{aligned} \tag{5}$$

In watermark detection, given matrices U_W, S, V_W, and the possibly distorted image A_W^*, one can extract a possibly corrupted watermark W^* by essentially reversing the above steps. That is

$$\begin{aligned} A_W^* &\Rightarrow U^* S_W^* V^{*H} \\ D^* &\Leftarrow U_W S_W^* V_W^H \\ W^* &\Leftarrow \frac{1}{a}(D^* - S). \end{aligned} \tag{6}$$

Note that the total number of DOF of matrices U_W, S, and V_W is N^2, which is the same as that of a $N \times N$ matrix. While many existing watermarking algorithms require the original image (whose DOF are also N^2) to extract the watermark, our method uses three matrices to extract the watermark and needs no additional information.

The similarity between W (the original watermark) and W^* (the extracted watermark) is measured by means of correlation. For convenience we regard W and W^* as one-dimensional (1-D) vectors and compute their correlation coefficient $c(W, W^*)$ in the standard manner. For a two-dimensional (2-D) watermark (such as the image of a company logo), we can simply map the watermark into a 1-D vector or compute the 2-D correlation coefficient directly.

IV. Analysis and Discussion

A. Noninvertible Watermarking Scheme

If an attacker attempts to create a "counterfeit original" image A_F, (2) and (3) must be satisfied simultaneously by the watermarked image A_W and the created counterfeit watermark W_F. In the following, we will analyze the method presented in the preceding section in order to show that it is a one-way and noninvertible watermarking algorithm.

Lemma 1: Suppose the SVD of image $A \in \mathbf{R}^{N\times N}$ is performed as in (4), then the mapping of matrix A to S is many-to-one and nonlinear.

Proof: Equation (4) yields

$$s_i(A) = \left[\lambda_i\left(A^H A\right)\right]^{1/2} \tag{7}$$

where $\lambda_i(\cdot)$ are the eigenvalues of matrix $A^H A$, $i = 1, 2, \ldots, N$. It is easy to prove that $S = \{s_1, s_2, \ldots, s_N\}$ is unique for a given A. But the inverse is untrue. In fact, there may exist many matrices whose SV matrices equal S. For a given S, we can illustrate this by constructing another matrix $\tilde{A} \in \mathbf{R}^{N\times N}$ as follows:

$$\tilde{A} \Leftarrow \tilde{U} S \tilde{V}^H. \tag{8}$$

If unitary matrices $\tilde{U} \neq U$ and/or $\tilde{V} \neq V$, then $\tilde{A}$ is different from A.

Equation (7) also shows that $s_i(\cdot)$ can be completely defined with matrix eigenvalues $\lambda_i(\cdot)$. Because the computation of $\lambda_i(\cdot)$ in domain $\mathbf{R}$ is not in closed-form and the eigenvalues cannot be obtained in a finite number of steps, we can conclude that the mapping of matrix A to S is nonlinear. □

We are now in a position to deduce an important result.

Theorem 1: If watermarking is carried out as in (5), then we have the following two equivalences:

$$\begin{aligned} A_W = E(A, W) &\Leftrightarrow S + aW = U_W S_W V_W^H \\ A_W = E(A_F, W_F) &\Leftrightarrow S_F + aW_F = U_{F,W} S_W V_{F,W}^H \end{aligned} \tag{9}$$

where $S_F = diag(s_{F,1}, \ldots, s_{F,N})$ and $S_W = diag(s_{W,1}, \ldots, s_{W,N})$ are the SV matrices of A_F and A_W, respectively; W_F is the created counterfeit watermark; a is the scale factor as described in (5); and $U_{F,W} \in \mathbf{F}^{N\times N}$ and $V_{F,W} \in \mathbf{F}^{N\times N}$ are unitary matrices.

Proof: The proof is simple and the above two equivalences are similar. Therefore, we only need to prove one of them. From the definition of SVD and the proposed watermarking algorithm, we have

$$\begin{aligned} A_W &= E(A_F, W_F) \\ \Leftrightarrow A_W &= U S_W V^H \\ \Leftrightarrow U S_W V^H &= U U_{F,W}^H (S_F + aW_F) V_{F,W} V^H \\ \Leftrightarrow S_W &= U_{F,W}^H (S_F + aW_F) V_{F,W} \\ \Leftrightarrow U_{F,W} S_W V_{F,W}^H &= S_F + aW_F \\ \Leftrightarrow S_F + aW_F &= U_{F,W} S_W V_{F,W}^H. \end{aligned}$$

Similarly, it is obvious that the following equivalence holds:

$$A_W = E(A, W) \Leftrightarrow S + aW = U_W S_W V_W^H.$$

□

Based on the definition of noninvertibility and on the conclusion obtained above, we can deduce that if the new method is noninvertible, it should not satisfy the following two conditions simultaneously:

$$S + aW = U_W S_W V_W^H \tag{10}$$

$$S_F + aW_F = U_{F,W} S_W V_{F,W}^T \tag{11}$$

where matrices $S \in \mathbf{F}^{N\times N}$, $S_F \in \mathbf{F}^{N\times N}$, and $S_W \in \mathbf{F}^{N\times N}$ are all diagonal and represent the original image A, the "counterfeit original" image A_F, and the watermarked image A_W, respectively. $W = \{w_{ij}\} \in \mathbf{F}^{N\times N}$ and $W_F = \{w_{F,ij}\} \in \mathbf{F}^{N\times N}$ are watermarks. Here, for simplicity and without loss of generality, we neglect the scale factor a. In the following, we will show that for our proposed method, a suitable "counterfeit original" image A_F and watermark W_F cannot be created if certain constraints on the watermarks are imposed.

Proposition 1: a) Given S_W, (10) is an ill-posed problem, that is, it is impossible to find the original S and W. b) Given S_W and S, it is impossible to obtain W. c) Given S_W and W, it is computationally unfeasible to obtain S.

Proof:

a) It holds obviously because there is no unique solution for S and W if only S_W is given.

b) Equation (10) can be expanded into a set of N^2 equations. If we compute $W = \{w_{ij}\}$ with N^2 unknowns, under the $2N$ constraints provided by S_W and S, there is an insufficient number of constraints. Therefore the solution for the original watermark W cannot be found.

c) Given S_W and W, we have enough constraints and the solution of S in (10) is unique. Because $U = U(W, S)$ and $V = V(W, S)$ are all functions of S, (10) is a high-dimensional and highly nonlinear equation. So it is computationally almost unfeasible to obtain the solution. □

Proposition 2: For (11), we have similar observations. a) Given S_W and W_F, it is computationally difficult to create S_F that satisfies (11).b) Given S_W, S_F and by imposing some constraints on W_F (but without knowing W_F), it is difficult to find the solution of W_F.

Proof:

a) From the conclusion of Proposition 1(c), given S_W and W_F, there is a unique solution for S_F, but it is difficult to find it.

b) For given S_W and S_F, if there is no constraint on the watermark W_F, then (11) can be satisfied easily, as done in the following:

$$\begin{aligned} W_F =& U_{F,W} S_W V_{F,W}^H - S_F \\ \Leftrightarrow S_F + W_F =& U_{F,W} S_W V_{F,W}^H. \end{aligned} \tag{12}$$

This makes it possible for an attacker to create counterfeit watermarks. However, if we impose some constraints on W_F, the real W_F cannot be obtained. In general, it is reasonable to require the watermark to satisfy some conditions instead of being randomly selected. For example, we can require that the watermarks be semantically meaningful. For enough constraints on the watermarks, there is a unique solution for W_F, but finding the solution is computationally almost impossible. □

Proposition 1 shows that (10) is an ill-posed problem, which means that a unique solution does not exist. From the above discussion, we can see that the new method is a safe watermarking scheme, especially when compared with many existing watermarking algorithms. In these watermarking schemes, given any two of A, A_W, and W, the third can easily be obtained, whereas this is not true for our method, as Proposition 1 shows. Furthermore, our method involves no cryptography (though cryptography can be incorporated to further enhance safety).

Proposition 2 shows that an attacker cannot create the counterfeit image A_F and watermark W_F while satisfying (11) under certain given constraints on the watermark. For example, we can require that the watermark take values form $\{1, 0\}$ or be semantically meaningful.

B. Error Estimation

When we add a watermark into an image, two questions should be asked: "What is the difference between the original image and the watermarked image?" and "How much energy or watermark information can be inserted?" These questions are not isolated and all in fact related to error estimation, a subject that has mostly been overlooked in the existing literature.

Definition 1: We define the spectral norm of a metrix $A = \{a_y\} \in \mathbf{F}^{M\times N}$ as follows:

$$\|A\|_2 = \sqrt{\lambda_{\max}} = s_{\max} \tag{13}$$

where $\lambda_{\max}$ and $s_{\max}$ denote the maximum eigenvalue of $A^T A$ and the maximum SV of A, respectively.

Lemma 2: If $U \in \mathbf{F}^{M\times M}$ and $V \in \mathbf{F}^{N\times N}$ are orthogonal matrices and $A \in \mathbf{F}^{M\times N}$, then

$$\|UAV\|_2 = \|A\|_2 . \tag{14}$$

Lemma 3: Let $A \in \mathbf{F}^{N\times N}$, let δA be is the disturbance on A, we denote $A_W = A + \delta A$. The ith SV in descending order of A and A_W are $s_i(A)$ and $s_i(A_W)$. Then

$$|s_i(A) - s_i(A_W)| \le \|\delta A\|_2 \tag{15}$$

where $i = 1, 2, \ldots, N$.

The above conclusions can be easily found in many textbooks on matrix theory. Lemma 3 is also called the singular value disturbance theorem. Then, it is easy for us to prove the following result.

Theorem 2: If I, A_W, W, and $s_i(\cdot)$ are defined as above, then

$$|s_i(A_W) - s_i(A)| \le a \|W\|_2 , \; i = 1, \ldots, n. \tag{16}$$

Proof: From (5), (6), (14), and (15), we get

$$\begin{aligned} |s_i(A_W) - & s_i(A)| \\ &= |s_i(S_W) - s_i(S)| \\ &= |s_i(S + aW) - s_i(S)| \\ &\le a \|W\|_2 . \end{aligned}$$

□

From (16) we see that $\|W\|_2$ can be used as a measure to determine the error between I and $\hat{A}$. Therefore we can adjust the watermark's spectral norm to an acceptable level to tradeoff between robustness and perceptibility. The simplest way is to adjust the value of the scale factor a. Theorem 2 provides theoretical guidance for us to select watermarks, control watermark location, and determine the watermarking energy inserted. Such information is typically unavailable in existing watermarking algorithms, but is of great importance in practical applications.

C. Watermark Selection

From (16), we can see that, given scale factor a, the error between the original image and the watermarked image is controlled by the spectral norm of the watermark $\|W\|_2$. From the viewpoint of imperceptibility of the watermark, we want $\|W\|_2$ to be as small as possible. Smaller value of $\|W\|_2$ suggests better similarity between the original image and the watermarked image. Therefore, watermark selection is an important issue to be considered. Many watermarking algorithms select pseudo-Gaussian random sequences as watermarks and use them to prove the existence of watermarks by means of correlation detection or by exploiting the statistical characteristics [9]. However, they hardly consider semantically meaningful visual watermarks. In fact, such watermarks are very common in many practical applications.

Consider two watermark matrices $A = \{a_{ij}\} \in \mathbf{F}^{N\times N}$ and $B = \{b_{ij}\} \in \mathbf{F}^{N\times N}$, where $i, j = 1, \ldots, N$. A is a matrix whose elements are real random numbers, while B is a grayscale image. If a_{ij} is Gaussian, for example, $a_{ij} \sim N(0, 1)$, and $b_y \subset [0, 1]$ is normalized to the range of [0,1], then in general we have

$$\|B\|_2 \gg \|A\|_2 . \tag{17}$$

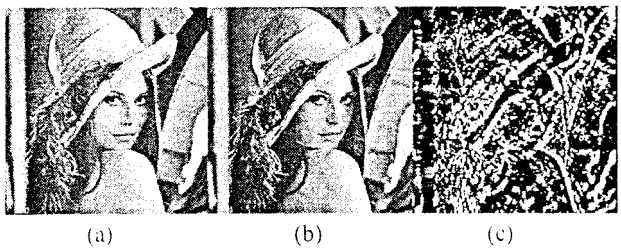

Fig. 1. Digital watermarking for image Lena by the SVD method. (a) Original image. (b)Watermarked image. (c) Absolute error image.

For most images, the energy is mainly concentrated in a small number of large SVs. From (16), it is obvious that if we choose random matrix as the watermark, better results can be obtained. For nonrandom meaningful watermarks, we can preprocess the watermark to reduce the value of the watermark's 2-norm. One effective approach is to randomize the watermark.

V. Experimental Results

In this section, we demonstrate the robustness of our watermarking method. The resistance of the proposed watermarking algorithm to various distortions was studied in a series of experiments on grayscale images.

We compare our method with the Spread Spectrum Communication method proposed by Cox [9] in order to put the performance investigation of our algorithm in proper context. The results show that our method is much more robust. The algorithm is tested on a variety of images, but for the sake of space, here we only give the results obtained using the 200×200 grayscale image Lena and test robustness under six practical conditions: adding noise, low-pass filtering, JPEG compression, scaling, image cropping, and rotation.

Similar to the Cox method, the watermark used is a 2500×1 vector consisting of pseudo-Gaussian random numbers. In watermark casting with the SVD method, we represent the watermark vector as a 50×50 matrix, while in the Cox method, the watermark is directly added to the first 2500 highest magnitude DCT coefficients of the image. The factor a of the Cox method, which controls the watermark energy to be inserted, is set to 0.1 (a typical value used by Cox [9]). We use a set of 50 2500×1 random vectors as watermarks for testing, of which only the tenth is the correct one. The similarity of the original image between the watermarked one is evaluated by their 2-D correlation coefficient e_c. The value of e_c^{cox} produced by the Cox method is 0.9957.

Fig. 1 shows the result of digital watermarking on Lena by the SVD method. The original image Lena is in Fig. 1(a). The watermarked image is shown in Fig. 1(b) and (c) is the absolute error image scaled up by a factor of 64. The value of a is set to 0.2 to ensure that the two images watermarked by the SVD method and the Cox method have comparable visual appearance. The correlation coefficient value e_c^{SVD} is 0.9966. Note that the absolute error image shows the texture characteristics of the original image.

Fig. 2 shows the result of adding Gaussian noise. We first obtain the watermarked image according to (5) and then add Gaussian noise to it. The mean of the additive Gaussian white noise is zero and its variance is 0.05. By performing the watermark detection, we obtain the corrupted watermark W^*. Then the correlation coefficient $c(W, W^*)$ between W (original watermark) and W^* is computed. The watermarked image produced by the SVD method after adding noise is shown in Fig. 2(a). The response of watermark correlation detection is shown in Fig. 2(b). The ordinate axis represents the value of the correlation coefficients and the abscissa axis represents a set of 50 2500×1 watermarks. It is obvious that only the tenth watermark (the correct one) achieves a meaningful correlation value (about 0.7).

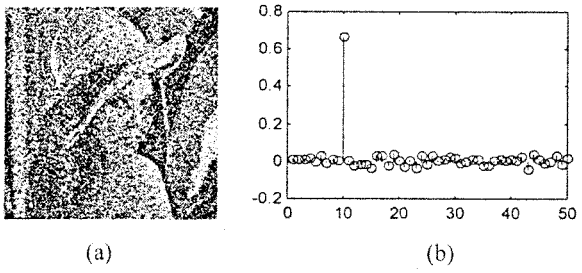

Fig. 2. Noise robustness of the SVD method. (a) Noisy image. (b) Watermark correlation coefficient.

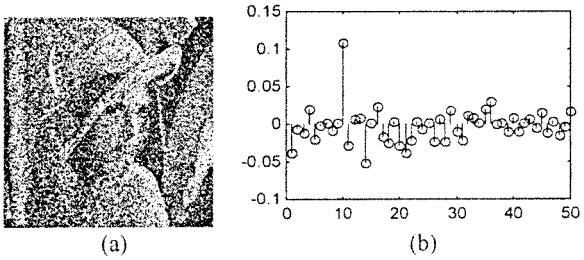

Fig. 3. Noise robustness of the Cox method. (a) The noisy image. (b) Watermark correlation coefficient.

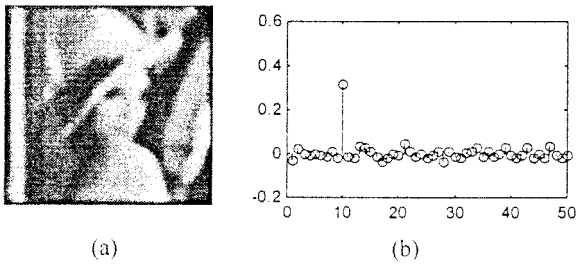

Fig. 4. Robustness test against low-pass filtering for the SVD method. (a) The blurred image. (b) Watermark correlation coefficient.

Fig. 3 shows similar results of the Cox method for comparison. Again, we first add watermark to the original image, then corrupt it by adding the same noise described above. The correlation coefficients $c(W, W^*)$ are also computed. The corrupted image produced by the Cox method is shown in Fig. 3(a) and the response of watermark detection is shown in Fig. 3(b). We note that, although the correct watermark (the tenth) shows the highest response among all watermarks, the highest value in this case is only a mere 0.11.

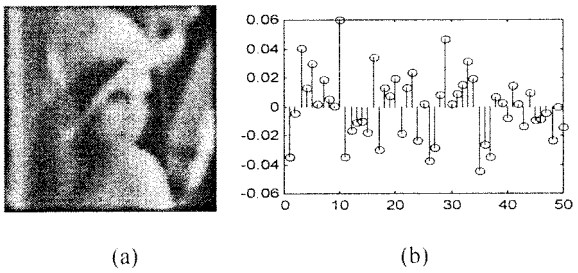

Fig. 5. Robustness test against low-pass filtering for the Cox method. (a) The blurred image. (b) Watermark correlation coefficient.

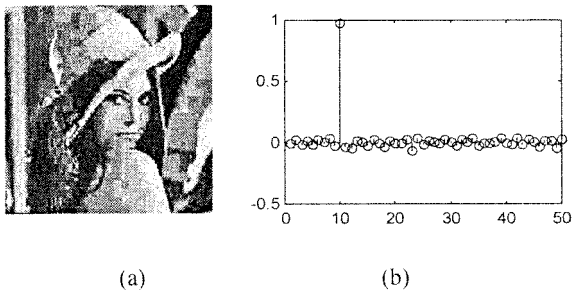

Fig. 6. Robustness test against JPEG compression using the SVD method. (a) The compressed-decompressed image. (b) Watermark correlation coefficient.

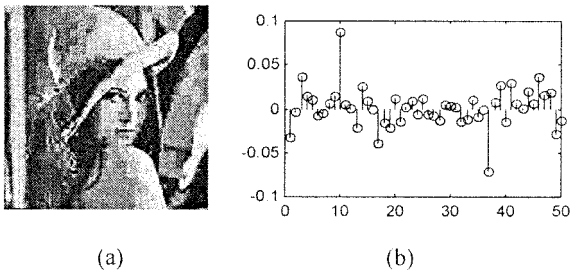

Fig. 7. Robustness test against JPEG compression using the Cox method. (a) The compressed-decompressed image. (b) Watermark correlation coefficient.

Fig. 4 shows the result of detecting the watermark with the SVD method on low-pass filtered images. The filter is a Gaussian low-pass filter. Its size is 16×16 and the variance σ is 4. We use the filter to perform 2-D FIR filtering on the watermarked image. After filtering, the image is heavily smoothed. The smoothed image is shown in Fig. 4(a). The responses of watermark detection from the blurred image are shown in Fig. 4(b). We can see that after heavy smoothing, the SVD method can still reliably detect the correct watermark. The value of the correlation coefficient of the correct watermark is about 0.3, which is much higher than that of the Cox method, where the result is almost meaningless (see Fig. 5).

We also perform lossy compression for the watermarked images. Fig. 6 shows the robustness test against JPEG compression for the SVD method. We apply heavy compression to the watermarked image. The quality of JPEG compression is five with a compression ratio of 18. Fig. 6(a) shows lossy version of the image. Fig. 6(b) displays the responses of correlation detection. We can see that, ever after heavy compression, the watermark is almost unchanged. The value of $c(W, W^*)$ is 0.9812, which suggests that our method is extremely robust against image compression.

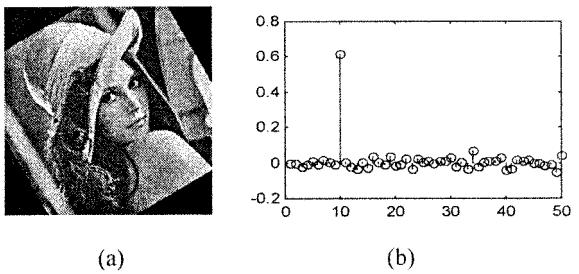

Fig. 8. Robustness test against image rotation for the SVD method. (a) The rotated image. (b) Watermark correlation coefficient.

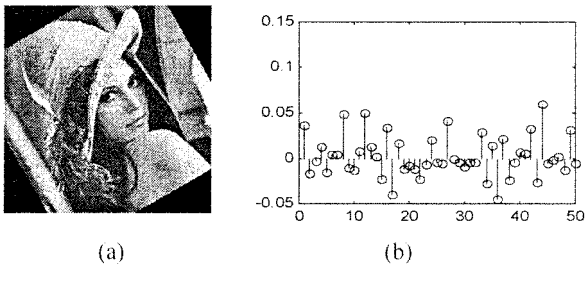

Fig. 9. Robustness test against image rotation for the Cox method. (a) The rotated imagee. (b) Watermark correlation coefficient.

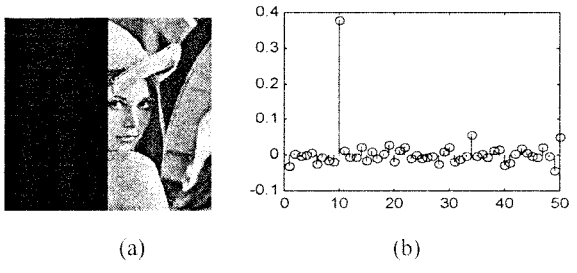

Fig. 10. Robustness test against image cropping for the SVD method. (a) The clipped image. (b) Watermark correlation coefficient.

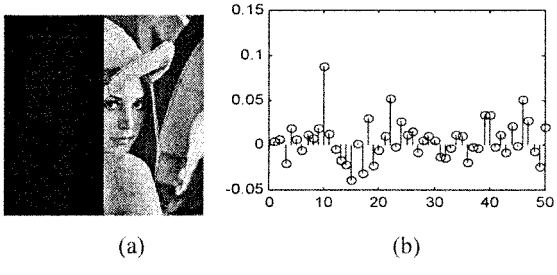

Fig. 11. Robustness test against image cropping for the Cox method. (a) The clipped image. (b) Watermark correlation coefficient.

Fig. 7 shows the Cox method against JPEG under the same compression conditions. Fig. 7(a) shows the resulting image after JPEG compression. Fig. 7(b) displays the responses of correlation detection. Again, the result is almost of no use.

Fig. 8 shows the test for robustness to image rotation for the SVD method. We use bilinear interpolation to perform the rotation of the watermarked image. The rotation angle is 30°. After rotation, we crop the four corners of the rotated image in order to keep the same size as the original image. Fig. 8(a) shows the rotated image. Fig. 8(b) displays the correlation coefficients of watermark detection. Note that rotation by angles of multiples of 90° and image transpose have no effect on the method.

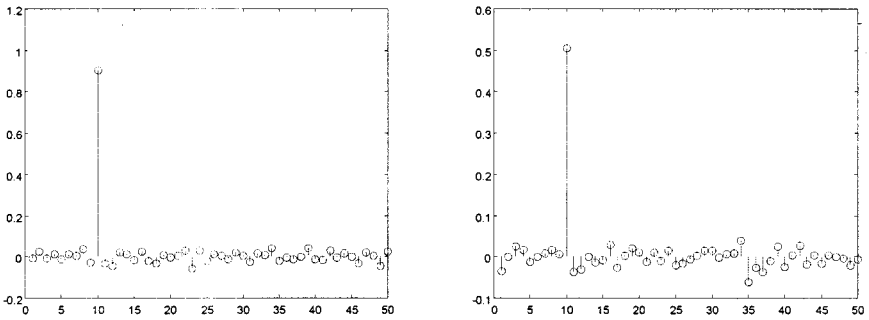

Fig. 12. Robustness test against rescaling and resampling of the Cox method and the SVD method; rescaling and resampling factors are 2. (a) Test result of the SVD method. (b) Test result of the Cox method.

Fig. 9 shows the Cox method against image rotation. The watermarked image is also rotated by 30°. Fig. 9(a) is the rotated image. Fig. 9(b) displays the responses of correlation detection. We can see that the watermark is destroyed completely after image rotation. In fact, the Cox method has little resistance to image rotation.

Fig. 10 shows the test for robustness to image cropping. We remove the left half of the watermarked image produced by the SVD method and then detect existence of the watermark from the remaining data. Fig. 10(a) is the clipped image. Fig. 10(b) displays detection results. We can see that the correct watermark can still be reliably detected. The correlation coefficient value is 0.3786. One interesting observation is that, when we also perform the test of removing the image's right half, the lower case correlation coefficient value becomes 0.1548, which means that the image's right half contains more information than the left half as far as image watermarking is concerned.

Fig. 11 is the same robustness test for the Cox method. Fig. 11(a) is the clipped image. Fig. 11(b) shows the responses of correlation detection. The maximum correlation coefficient value is 0.0877. Note that the Cox method requires the original image to extract the watermark. Therefore, if we replace the watermarked image's left part with the corresponding part of the original image, a much better result can be obtained. The correlation coefficient value becomes 0.7178, while if the same processing is applied to our method, the correlation coefficient value becomes 0.9912.

Fig. 12 shows the results of the test for robustness against rescaling and resampling. By comparing the correlation coefficient values of the SVD method (0.90) and the Cox method (0.50), we can see that our method is more robust to rescaling and resampling attacks.

Finally, we demonstrate the results of using visual watermarks. In our method, the added watermark is represented by a matrix, so it is convenient for us to embed a visual watermark directly into the image without additive processing. Here, we add a 50×50 grayscale image into the Lena image. The watermark is shown in Fig. 14(a). The value of scaling factor a is also 0.2 and the error value e_e^{SVD} is 0.9989. Fig. 13 shows the experimental results.

Robustness tests have also been done for the watermarked image with a visual watermark. Fig. 14 shows the test results. (a) is the original watermark. (b) is the extracted watermark after the watermarked image is contaminated by white noise. The noise's mean is 0 and its variance is 0.05. The error value

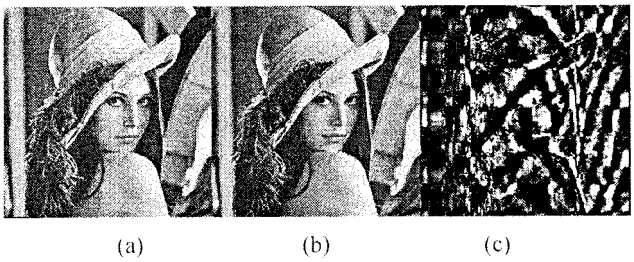

Fig. 13. Embedding visual watermarks by the SVD method. (a) Original image. (b) Watermarked image. (c) The absolute error of the two images scaled up by 64.

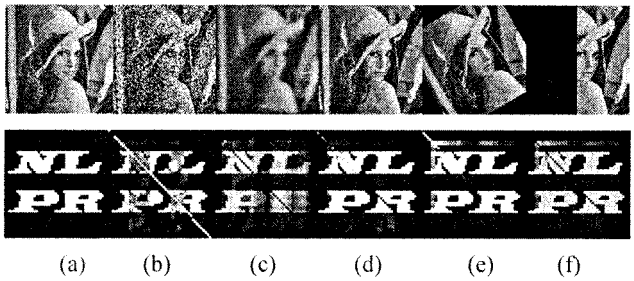

Fig. 14. Robustness test of visual watermark. The corresponding Lena images of (b) $\sim$ (f) are shown on the top row.

(i.e., the value of 2-D correlation coefficient of the original watermark and the corrupted watermark) is 0.2725. (c) shows the result of low-pass filtering. The filter is 16×16 Gaussian filter, with its variance being 4. The error value is 0.1503. (d) is the JPEG compression result. The JPEG compression quality is 5. The correlation coefficient is 0.9425. We can see that heavy compression has nearly no effect on our method. (e) is the corrupted watermark extracted after we rotate the watermarked image for 30°, with bilinear interpolation and cropping. The correlation coefficient is 0.925. (f) shows the cropping robustness test. The image's left half part is removed, but the logo can still be identified. The correlation coefficient is 0.3300.

These experimental results show that, even if the watermarked image has undergone severe physical distortions, the SVD method can still detect the correct watermark and determine the existence of the watermark. The results also clearly demonstrate that the novel method is considerably more robust than the popular Cox method.

VI. Conclusions

In this paper, a new watermarking method for digital images has been presented. The watermark is added to the SVD domain of the original image. The mathematical background of this method is very clear and the error between the original image and the watermarked image can be estimated. As a result, important questions such as how to determine the location of the watermark and how much energy to be inserted can be answered easily. Unlike some other unitary transformations, which adopt fixed orthogonal bases (such as discrete Fourier transform (DFT), discrete cosine transform (DCT), etc.), SVD uses nonfixed orthogonal bases. It is a one-way, nonsymmetrical decomposition. These properties lead to the good performance of the novel algorithm in both security and robustness. Furthermore, the algorithm does not require encryption to resolve rightful ownership and can provide more powerful security for rightful ownership if combined with encryption. Extensive experiments and comparisons with the Cox method have been made. Results show that the new method is very robust against image distortion and is considerably more robust than the Cox method.

Acknowledgment

Some parts of the work discussed in this paper have been filed for patent (Patent no. 99 107 964.7). The authors would like to thank the anonymous referees for their thorough review of the paper and many constructive comments.

References

[1] B. R. Macq and I. Pitas, "Special issue on water making," *Signal Process.*, vol. 66, no. 3, pp. 281–282, 1998.
[2] M. D. Swanson, M. Kobayashi, and A. H. Tewfik, "Multimedia data-embedding and watermarking technologies," *Proc. IEEE*, vol. 86, pp. 1064–1087, June 1998.
[3] S. H. Low, N. F. Maxemchuk, and A. M. Lapone, "Document identification for copyright protection using centroid detection," *IEEE Trans. Commun.*, vol. 46, pp. 372–383, Mar. 1998.
[4] J. M. Acken, "How watermarking adds value to digital content," *Commun. ACM*, vol. 41, no. 7, pp. 74–77, 1998.
[5] J. Zhao, E. Koch, and C. Luo, "Digital watermarking in business today and tomorrow," *Commun. ACM*, vol. 41, no. 7, pp. 67–72, 1998.
[6] W. Zeng, "Digital watermarking and data hiding: Technologies and applications," in *Proc. ICISAS*, vol. 3, 1998, pp. 223–229.
[7] S. Craver, N. Memon, B. L. Yeo, and M. M. Yeung, "Resolving rightful ownerships with invisible watermarking techniques: Limitations, attacks and implications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 4, pp. 573–586, 1998.
[8] L. T. Qiao and K. Nahrstedt, "Watermarking schemes and protocols for protecting rightful ownership and customer's rights," *J. Vis. Commun. Image Represent.*, vol. 9, no. 3, pp. 194–210, 1998.
[9] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Processing*, vol. 6, pp. 1673–1687, Dec. 1997.
[10] N. Nikolaidis and I. Pitas, "Copyright protection of images using robust digital signatures," in *Proc. ICASSP*, vol. 4, May 1996, pp. 2168–2171.
[11] ——, "Robust image watermarking in the spatial domain," *Signal Process.*, vol. 66, no. 3, pp. 385–403, 1998.
[12] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain system for robust image watermarking," *Signal Process.*, vol. 66, no. 3, pp. 357–372, 1998.
[13] S. Craver, N. Memon, B. Yeo, and M. Yeung, "Can invisible watermarks resolve rightful ownership?," IBM Research Division, Tech. Rep. RC 20509, 1996.
[14] R. B. Wolfgang and E. J. Delp, "A watermark technique for digital imagery: Further studies," in *Proc. ICISST*, Las Vegas, NV, 1997.
[15] W. Zeng and B. Liu, "A statistical watermark detection technique without using original images for resolving rightful ownerships of digital images," *IEEE Trans. Image Processing*, vol. 8, pp. 1534–1548, Nov. 1999.
[16] H. C. Andrews and C. L. Patterson, "Singular value decomposition (SVD) image coding," *IEEE Trans. Commun.*, vol. COM-24, pp. 425–432, Apr. 1976.
[17] G. H. Golub and C. Reinsch, "Singular value decomposition and least squares solutions," *Numer. Math.*, vol. 14, pp. 403–420, 1970.
[18] C. T. Hsu and J. L. Wu, "Multiresolution watermarking for digital images," *IEEE Trans. Circuits and Systems II*, vol. 45, no. 8, pp. 1097–1101, 1998.
[19] J. J. K. O'Ruanaidh and T. Pun, "Rotation, scale, and translation invariant spread spectrum digital image watermarking," *Signal Process.*, vol. 66, no. 3, pp. 303–317, 1998.

Ruizhen Liu received the B.Sc. degree in aerocraft design from Beijing University of Aeronautics and Astronautics, Beijing, China, in 1990, the M.Sc. degree from Fuzhou University, Fuzhou, China, in 1998, and the Ph.D. degree from National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2001.

He is currently an R&D Engineer in MediaSec Technologies, GmbH, Germany.

Tieniu Tan (M'92–SM'97) received the B.Sc. degree in electronic engineering from Xi'an Jiaotong University, China, in 1984, and the M.Sc. and Ph.D. degrees in electronic engineering from Imperial College of Science, Technology and Medicine, London, U.K., in 1986 and 1989, respectively.

In October 1989, he joined the Computational Vision Group at the Department of Computer Science, University of Reading, Reading, U.K., where he worked as a Research Fellow, Senior Research Fellow, and Lecturer. In January 1998, he returned to China to join the National Laboratory of Pattern Recognition, Institute of Automation of the Chinese Academy of Sciences, Beijing, China, where he is currently Professor and Director of the National Laboratory of Pattern Recognition, as well as the President of the Institute of Automation. He has published widely on image processing, computer vision, and pattern recognition. His current research interests include speech and image processing, machine and computer vision, pattern recognition, multimedia, and robotics.

Dr. Tan was an elected member of the Executive Committee of the British Machine Vision Association and Society for Pattern Recognition (1996–1997). He serves as referee for many major national and international journals and conferences. He is an Associate Editor of the *International Journal of Pattern Recognition*, the Asia Editor of the *International Journal of Image and Vision-Computing*, a founding co-chair of the IEEE International Workshop on Visual Surveillance, and the Program Chair for the Third International Conference on Multimodal Interfaces (ICMI'2000).