Int. J. Electron. Commun. (AEÜ) 68 (2014) 816–834

Contents lists available at ScienceDirect

International Journal of Electronics and Communications (AEÜ)

journal homepage: www.elsevier.com/locate/aeue

Security analyses of the watermarking scheme based on redundant discrete wavelet transform and singular value decomposition

Jing-Ming Guo [1], Heri Prasetyo *

Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan

ARTICLE INFO

Article history:
Received 30 May 2013
Accepted 13 March 2014

Keywords:
False positive problem
Image watermarking
Redundant discrete wavelet transform (RDWT)
Singular value decomposition (SVD)
Vulnerable attack

ABSTRACT

This study analyzes the recent image watermarking schemes based on redundant discrete wavelet transform (RDWT) and singular value decomposition (SVD), and shows that in fact they are insecure and cannot be used for protecting the rightful ownership. The RDWT-SVD watermarking directly embeds a grayscale watermark image of the same size with the host image into the singular value matrix of the RDWT-transformed host image, then produces the left and right orthogonal matrices as side information which is later used in the watermark extraction stage. The RDWT-SVD approach enjoys the advantage of the RDWT redundancy to achieve a high embedding capacity, and preserves the watermark imperceptibility by exploiting the SVD stability properties. It is claimed that RDWT-SVD watermarking is robust against several common image processing and geometrical attacks, yet a fundamental flaw in the RDWT-SVD scheme is found, which leads to severe the false positive issue. Three vulnerable attacks should be considered in the RDWT-SVD scheme: (1) An attacker can easily claim the owner watermarked image; (2) the owner has the ambiguity because of the wrong side information usage, and (3) the owner can extract the correct watermark from arbitrary image. Thus, it is important to highlight these attacks when implementing the RDWT-SVD watermarking scheme.

1. Introduction

The main goal of image watermarking is to hide specific information into host image [1] such that the presence of the watermark cannot be realized by human visual. The good watermarking algorithm design should meet three criteria, i.e. the rightful ownership protection, robustness against the image manipulations, and watermark imperceptibility. In the ideal situation, only the corresponding/real owner can extract watermark correctly from a watermarked image. The watermarking proposed in [1] is pioneer in the SVD-based image watermarking domain. Numerous researches have devoted to improve the SVD watermarking performance [7–22]. In the SVD watermarking, the watermark information is embedded into the singular value matrix of the host image. In [7–15,17,19], the host image is transformed first (most of them using the wavelet transform and its family) before performing the SVD. Conversely, the methods in [1,16,18,20–22] directly compute the SVD of the host image.

In SVD watermarking techniques [7–15,17,19], the visual watermark is embedded into the singular value matrix of the host image in the transformed domain by selecting the suitable scaling factor. The discrete wavelet transform and its family are chosen in [7–15,17,19] to achieve the imperceptibility aspect in the watermarked image. Two common approaches are used to embed the watermark information into a host image: (1) directly insert a watermark image into host image singular value matrix [1,15–22], and (2) only inject the watermark singular value into the singular value matrix [7–14] of host image. Both types of these methods yield a good quality watermarked image in terms of the PSNR, and robust against various attacks in terms of the correlation coefficient as documented in [1,7–22].

In the SVD watermarking approaches, the scaling factor plays an important role to control the robustness and imperceptibility of the watermark. By setting a higher value of scaling factor, the watermarked image is more robust against attacks. Yet, the image quality is dramatically degraded. In contrast, the transparency of the watermark is improved by setting a lower value of scaling factor, with the trade off that the watermarked image is less robust against various geometric distortions and image processing attacks. Some efforts have been addressed to find the suitable scaling factor by regarding the scaling factor determination as an optimization problem [16,18,20]. The metaheuristic algorithm, namely

* Corresponding author. Tel.: +886 2 27303241; fax: +886 2 27376699.
E-mail addresses: jmguo@seed.net.tw (J.-M. Guo), heri_inf_its_02@yahoo.co.id (H. Prasetyo).
[1] Senior member, IEEE.

http://dx.doi.org/10.1016/j.aeue.2014.03.008
1434-8411/

genetic algorithm, is used to find the optimum scaling factor [16]. The tiny genetic algorithm [18] and differential evolution algorithm [20] iteratively search and update the suited scaling factor to achieve the robustness aspect, and to improve the image quality of watermarked image. However, a major flaw is found in the SVD watermarking as reported in [2–6].

The RDWT-SVD watermarking [15] utilizes the RDWT transformation in the watermark embedding stage because of its spatio-frequency localization property. The RDWT overcomes the discrete wavelet transform (DWT) problem in the shift invariant, and avoids the downsampling process of each level in the DWT filtering. The RDWT sub-band maintains the same size as the original image, and keeps the important texture of an original image at the same spatial location in each sub-band. Thus, the RDWT can hide information with invisibility constraint. Compared with the DWT transformation, the RDWT also removes the upsampling and downsampling for its coefficients. At each level, RDWT produces output coefficients which have the same size as its input. It is the main reason that the RDWT can achieve a higher watermark embedding capacity compared to the DWT.

As demonstrated in [15], the RDWT-SVD watermarking is not only robust against several image processing attacks and geometric distortions, but also it yields high PSNR for the watermarked image. It means that the RDWT-SVD can successfully render the visual watermark image into the host image and achieve satisfied watermark robustness aspect against malicious attacks. However, the RDWT-SVD is not vigorous against three vulnerable attacks presented in this paper. An attacker can easily destroy the watermark information from the real owner watermarked image. Some ambiguity may easily occur in the real owner or counterfeit attacker part because of wrong usage of the side information in the watermark extraction stage. Additional serious problem also occurs as the real owner can extract correct watermark from an arbitrary image. These vulnerable attacks should be considered when implementing the RDWT-SVD watermarking for the ownership rightful applications in the future.

The rest of this paper is organized as follows. The RDWT-SVD image watermarking is briefly reviewed in Section 2. Section 3 presents the three vulnerable attacks for RDWT-SVD scheme. The analysis of the RDWT-SVD watermarking is given in Section 4. Experimental results for RDWT-SVD attacks are reported in Section 6. Finally, the conclusions are drawn at the end of this paper.

2. RDWT-SVD image watermarking

This section gives a review of the singular value decomposition (SVD) and brief introduction for the RDWT-SVD image watermarking. The watermark embedding and extraction strategy from [15] is reviewed, and subsequently show the robustness and imperceptibility result from the RDWT-SVD scheme.

2.1. SVD overview

The SVD is a numerical tool to diagonalize and decompose a matrix into its eigenvectors and eigenvalues [1]. The SVD of an image $A \in \mathbb{R}^{N\times M}$ is defined as:

$$A \Rightarrow U\Sigma V^T, \tag{1}$$

where $U \in \mathbb{R}^{N\times r}$ and $V \in \mathbb{R}^{M\times r}$ are left and right singular vectors, and $\Sigma \in \mathbb{R}^{r\times r}$ is a singular value matrix. The U and V are unitary matrices which satisfy $UU^T = U^TU = I_N$ and $VV^T = V^TV = I_M$. The Σ is a diagonal matrix which has nonnegative entries in its diagonal ($\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$). When r denotes a rank of matrix A or the number of non-zero singular value, the singular value matrix is

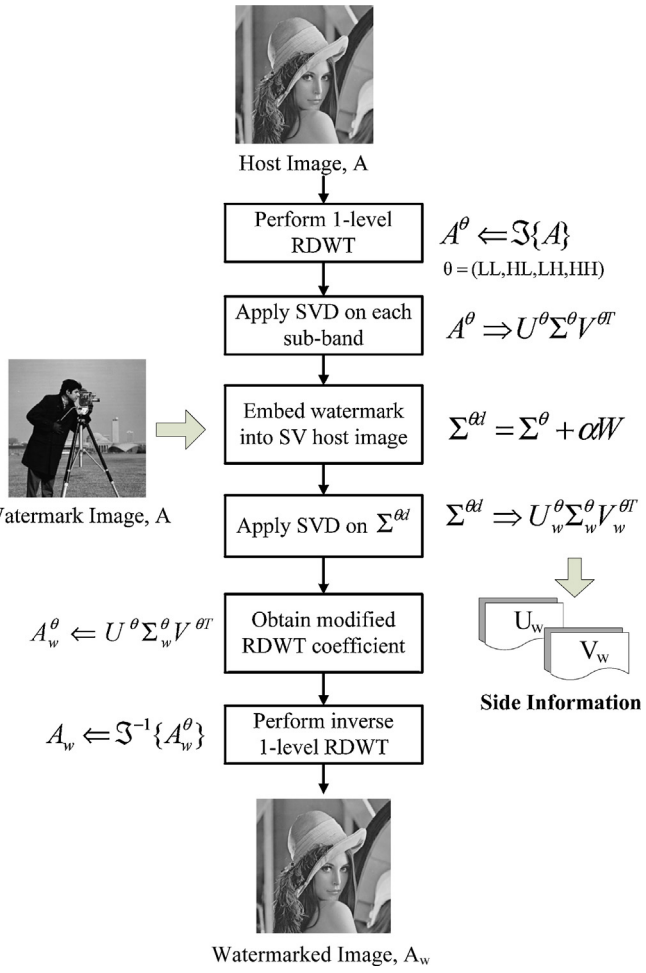

Fig. 1. Flowchart of watermark embedding in RDWT-SVD scheme [15].

$\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$ and $\Sigma_r = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$. The matrix A can be rewritten in the compact form as:

$$A = \sum_{i=1}^{r} u_i \sigma_i v_i^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T. \tag{2}$$

The multiplication of AA^T and A^TA yield the following result:

$$AA^T = U\Sigma V^T [U\Sigma V^T]^T = U\Sigma V^T V\Sigma^T U^T = U\Sigma\Sigma^T U^T = U\Sigma^2 U^T, \tag{3}$$

$$A^TA = [U\Sigma V^T]^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V\Sigma^T\Sigma V^T = V\Sigma^2 V^T. \tag{4}$$

The multiplication results of AA^T and A^TA are symmetric matrix, then the problem in (3) and (4) can be regarded and solved by performing the eigenvalue-decomposition of AA^T and A^TA. The matrices U and Σ^2 are eigenvector and eigenvalue of AA^T, respectively. The matrix V is the eigenvector from A^TA. It is shown in [1] that the singular value matrix of an image has good stability when a small perturbation is added into an image, which is the major reason that most of SVD-based image watermarking attempts to embed watermark into the singular value matrix of a host image.

2.2. Watermark embedding

In this subsection, the watermark embedding strategy as proposed in [15] is reviewed. Let A be a grayscale host image of size $M \times N$, and W is the visual grayscale watermark of the same size with the host image. Fig. 1 illustrates the watermark

embedding process in the RDWT-SVD scheme. First, the host image is decomposed using the RDWT into four sub-bands, and then the SVD is applied for each sub-band. The visual watermark is directly embedded into the singular value matrix of each sub-band on the RDWT transformed domain by employing a suited scaling factor. The RDWT-SVD scheme takes the main advantage of RDWT, in which the watermark can be of the same size as the host image. Formally, the watermark embedding with the RDWT-SVD scheme is as below:

- E1. Perform the 1-level RDWT on the host image, and obtain four sub-bands:

$$\mathrm{A}^{\theta} \Leftarrow \Im\{\mathrm{A}\}, \tag{5}$$

where $\Im\{\cdot\}$ denotes the RDWT operation, and θ = (LL, HL, LH, HH). We can substitute the $\Im\{\cdot\}$ with the other transformation such as DWT, Fourier, DCT, Hadamard transform, etc.
- E2. Apply the SVD to each sub-band:

$$\mathrm{A}^{\theta} \Rightarrow \mathrm{U}^{\theta}\Sigma^{\theta}\mathrm{V}^{\theta T}, \tag{6}$$

- E3. Embed the watermark W directly into the singular value matrix of the host image, and then perform SVD on the modified singular value matrix of the host image:

$$\Sigma^{\theta} + \alpha\mathrm{W} \Rightarrow \mathrm{U}_{\mathrm{w}}^{\theta}\Sigma_{\mathrm{w}}^{\theta}\mathrm{V}_{\mathrm{w}}^{\theta\mathrm{T}}, \tag{7}$$

where α denotes the scaling factor which controls the strength of the watermark. In [15], the scaling factor for LL sub-band and the other sub-band are set at 0.05 and 0.005, respectively.
- E4. Obtain the modified RDWT coefficient for each sub-band $\mathrm{A}_{\mathrm{w}}^{\theta}$, by replacing the singular value matrix from A^{θ} with singular value matrix $\Sigma_{\mathrm{w}}^{\theta}$:

$$\mathrm{A}_{\mathrm{w}}^{\theta} \Leftarrow \mathrm{U}^{\theta}\Sigma_{\mathrm{w}}^{\theta}\mathrm{V}^{\theta\mathrm{T}}, \tag{8}$$

- E5. Perform the inverse RDWT for all modified coefficient sub-bands to obtain the watermarked image A_{w}:

$$\mathrm{A}_{\mathrm{w}} \Leftarrow \Im^{-1}\{\mathrm{A}_{\mathrm{w}}^{\theta} | \theta = (\mathrm{LL}, \mathrm{HL}, \mathrm{LH}, \mathrm{HH})\}, \tag{9}$$

The left and right singular vectors ($\mathrm{U}_{\mathrm{w}}^{\theta}$ and $\mathrm{V}_{\mathrm{w}}^{\theta}$) obtained from Step. E3 are considered as side information, and singular value matrix Σ^{θ} from Step. E2 for all sub-bands, which will later be used in the watermark extraction stage.

2.3. Watermark extraction

In the watermark extraction, the RDWT-SVD scheme extracts the watermark W^{*} from the possibly corrupted watermarked image $\mathrm{A}_{\mathrm{w}}^{*}$. Some common image processing attacks and geometric distortions may degrade the image quality. Fig. 2 depicts the watermark extraction flowchart of the RDWT-SVD image watermarking scheme. By employing the $\mathrm{U}_{\mathrm{w}}^{\theta}$, $\mathrm{V}_{\mathrm{w}}^{\theta}$, and Σ^{θ} matrices obtained from the embedding step as side information, the watermark extraction step can be formally defined as follows:

- X1. Perform 1-level RDWT on the possibly corrupted watermarked image $\mathrm{A}_{\mathrm{w}}^{*}$ to obtain its four sub-bands:

$$\mathrm{A}_{\mathrm{w}}^{\theta *} \Leftarrow \Im\{\mathrm{A}_{\mathrm{w}}^{*}\}, \tag{10}$$

where θ = (LL, HL, LH, HH). The transformation $\Im\{\cdot\}$ in this step is identical to that used in the watermarking embedding step.
- X2. Apply the SVD for each sub-band $\mathrm{A}_{\mathrm{w}}^{\theta *}$:

$$\mathrm{A}_{\mathrm{w}}^{\theta *} \Rightarrow \mathrm{U}^{\theta *}\Sigma^{\theta *}\mathrm{V}^{\theta *\mathrm{T}}, \tag{11}$$

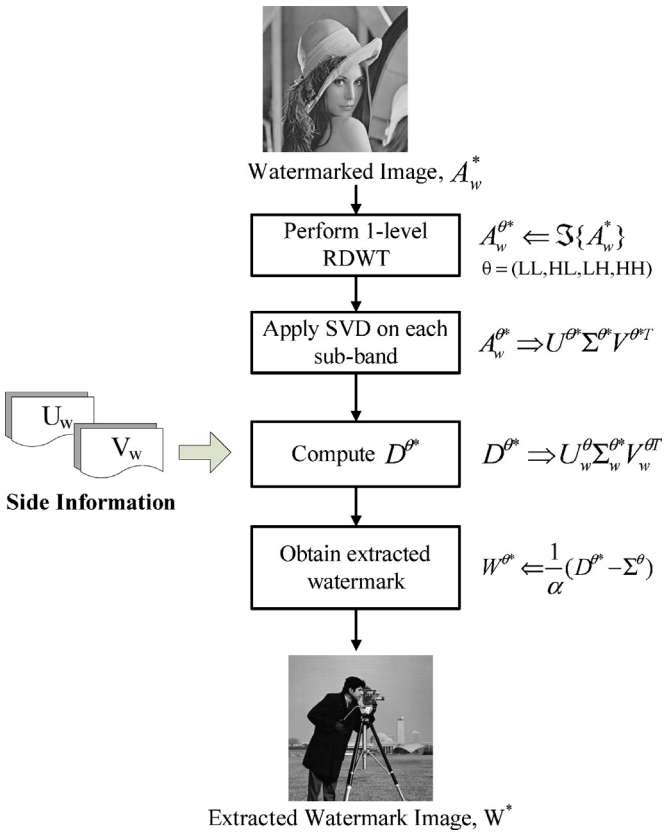

Fig. 2. Flowchart of watermark extraction in RDWT-SVD scheme [15].

- X3. Compute the matrix $\mathrm{D}^{\theta *}$ by multiplying $\Sigma^{\theta *}$ from Step X2 with the left and right singular vectors $\mathrm{U}_{\mathrm{w}}^{\theta}$ and $\mathrm{V}_{\mathrm{w}}^{\theta}$ from watermark embedding step as:

$$\mathrm{D}^{\theta *} \Leftarrow \mathrm{U}_{\mathrm{w}}^{\theta}\Sigma^{\theta *}\mathrm{V}_{\mathrm{w}}^{\theta\mathrm{T}}, \tag{12}$$

- X4. Obtain the extracted watermark by subtracting $\mathrm{D}^{\theta *}$ with Σ^{θ} from the watermark embedding step:

$$\mathrm{W}^{\theta *} \Leftarrow \frac{1}{\alpha}(\mathrm{D}^{\theta *} - \Sigma^{\theta}). \tag{13}$$

2.4. Imperceptibility and robustness of the RDWT-SVD image watermarking

The RDWT-SVD watermarking scheme has been evaluated under various experiments to demonstrate the robustness and imperceptibility [15]. The RDWT-SVD robustness is measured using the normalized cross correlation (NCC) as the similarity between the original watermark image W and the extracted watermark W^{*}:

$$\rho(\mathrm{W}, \mathrm{W}^{*}) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(w(i,j) - \mu_{\mathrm{w}})(w^{*}(i,j) - \mu_{\mathrm{w}}^{*})}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}(w(i,j) - \mu_{\mathrm{w}})^2}\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}(w^{*}(i,j) - \mu_{\mathrm{w}}^{*})^2}}, \tag{14}$$

where μ_{w} and μ_{w}^{*} indicate the mean value of the original and extracted watermarks, respectively. The symbol $w(i,j)$ and $w^{*}(i,j)$ denote the original and extracted watermark at pixel position (i,j), respectively. The higher value of NCC indicates that the watermark is more robust against various attacks. It means that the image quality of extracted watermark is still acceptable for human vision when some image manipulations and geometric distortions are injected into the host image.

The other criteria to judge the effectiveness of the RDWT-SVD watermarking scheme is in the imperceptibility aspect. The visual quality of host image should not be degraded too much after the

Fig. 3. (a) Lena; (b) Baboon; (c) Peppers; and (d) Cameraman.

watermark insertion. The PSNR can be adopted to objectively measure the similarity between the original and watermarked images. The PSNR is formally defined as

$$\mathrm{PSNR} = 10\log_{10}\frac{255^2}{\frac{1}{MN}\sum_{i=1}^{M}\sum_{i=1}^{N}[A(i,j)-A_w(i,j)]^2}. \quad (15)$$

where $A(i, j)$ and $A_w(i, j)$ represent the original image and watermarked image at pixel position (i, j), respectively.

Fig. 3 shows the grayscale images, Lena, Baboon, Peppers, and Cameraman, of size 512×512 used in [15] to examine the RDWT-SVD scheme performance. Fig. 4 shows the Lena, Baboon, and Peppers watermarked image with its corresponding PSNR when the cameraman image is turned as the watermark image. The RDWT-SVD scheme yields very high PSNR for all watermarked image and performs well to hide the visual watermark image into host image. Fig. 5 shows the watermarked image and its corresponding PSNR after various attacks. Fig. 6 presents the extracted watermark and its corresponding NCC value from the corrupted watermarked image. From this figure, we can see that the RDWT-SVD scheme is robust against various malicious attacks as indicated with high NCC values.

3. Attacks on RDWT-SVD image watermarking

In this section, three vulnerable attacks are evaluated for the RDWT-SVD watermarking. The RDWT-SVD scheme seems robust against various image processing attacks and geometric distortions; yet, a fundamental flaw is found in the RDWT-SVD approach. This flaw induces the false-positive issue, and thus causes the ambiguity situation. The vulnerable attacks are related to rightful ownership protection and can be defined as follows.

3.1. Attack type I

The image owner tries to embed a watermark image to protect the image ownership before distributing into public digital media. Fig. 7 shows the schematic diagram for attack type I. This attack is regarded as the ambiguity on the owner side. Suppose that the owner has two different watermark images, namely W^1 and W^2, which are then embedded into a host image A. The owner conducts the watermark embedding and extraction twice.

First, the watermark W^1 is inserted into the host image A yielding the watermarked image A_w^1 and the side information (U_w^1 and V_w^1). At the second embedding process, the owner embeds the watermark W^2 into the host image A resulting in the watermarked image A_w^2 and the side information (U_w^2 and V_w^2) as key in the watermark extraction stage.

At the watermark extraction stage, the false positive problem occurs when the owner extracts watermark information from the watermarked image A_w^1 using incorrect side information (U_w^2 and V_w^2). The owner obtains the possibly corrupted watermark image W^{2*} which is visually similar to W^2. However, the watermarked image A_w^1 contains the watermark information W^1.

The same problem happens when the owner tries to extract the watermarked image A_w^2 using the wrong side information, i.e. U_w^1 and V_w^1. As a result, the owner can receive the possibly corrupted watermark image W^{1*} which is apparently identical to the watermark image W^1. As we know that the watermarked image A_w^2 conceives the watermark information W^2. Apparently, these ambiguity situations in the RWDT-SVD watermarking cannot be accepted for real applications which need the ownership validity.

3.2. Attack type II

Two parties are involved in this type of attack, i.e. real owner and attacker. Fig. 8 illustrates the schematic diagram for attack type II. The real owner embeds the watermark image W into the host image A yielding the watermarked image A_w and the side

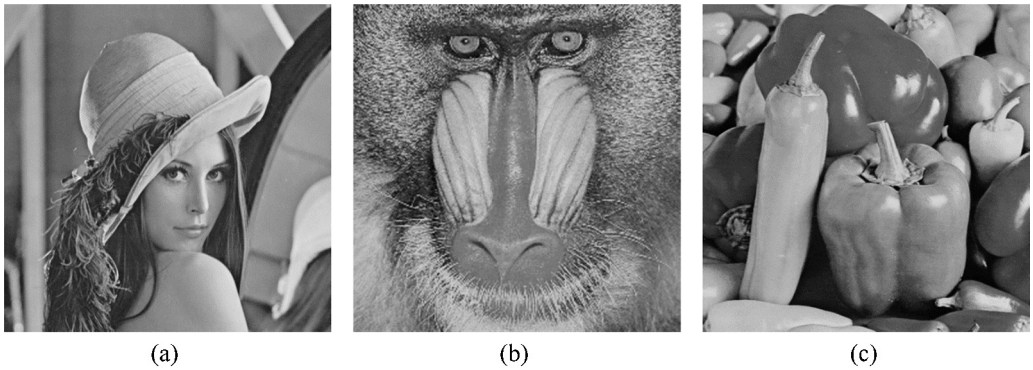

Fig. 4. Watermarked images. (a) Lena (PSNR 53.9851 dB); (b) Baboon (PSNR 54.7021 dB); and (c) Peppers (PSNR 54.0673 dB).

Fig. 5. Watermarked images under: (a) speckle noise (0.4) (PSNR 10.69 dB); (b) Gaussian noise (0, 0.005) (PSNR 23.03 dB); (c) pepper and salt noise (0.001) (PSNR 25.45 dB); (d) median filter (3×3) (PSNR 35.50 dB); (e) JPEG compression ($Q = 40$) (PSNR 35.08 dB); (f) histogram equalization (PSNR 19.13 dB); (g) rotation (angle 50°) (PSNR 10.35 dB); (h) shift 50% (PSNR 10.95 dB); and (i) cut attack (PSNR 12.97 dB).

information (U_w and V_w). Then, the real owner publishes the watermarked image A_w, and keeps the side information image as a tool to proof the image ownership. An attacker can easily obtain and modify the watermarked image A_w from publicly available digital media. In this attack, an attacker tries to embed the counterfeit watermark image W_f into the owner watermarked image A_w by the following procedure:

Watermark embedding

- E1-II. Perform 1-level RDWT on the owner watermarked image A_w as similar to Step E1:

$$A_f^{\theta} \Leftarrow \Im\{A_w\} \quad \text{where} \quad \theta = (\text{LL, HL, LH, HH}), \tag{16}$$

- E2-II. Perform SVD on each sub-band as similar to Step E2:

$$A_f^{\theta} \Rightarrow U_f^{\theta} \Sigma_f^{\theta} V_f^{\theta T}, \tag{17}$$

- E3-II. Insert the counterfeit watermark image W_f and then apply SVD as similar to Step E3:

$$\Sigma_f^{\theta} + \alpha W_f \Rightarrow U_{wf}^{\theta} \Sigma_{wf}^{\theta} V_{wf}^{\theta T}, \tag{18}$$

- E4-II. Obtain the modified RDWT coefficient as similar to Step E4:

$$A_{wf}^{\theta} \Leftarrow U_f^{\theta} \Sigma_{wf}^{\theta} V_f^{\theta T}, \tag{19}$$

- E5-II. Obtain the attacker watermarked image as similar to Step E5:

$$A_{wf} \Leftarrow \Im^{-1}\{A_{wf}^{\theta} | \theta = (\text{LL, HL, LH, HH})\}. \tag{20}$$

At the end of watermark embedding stage, an attacker obtains the watermarked image A_{wf}, and keeps the counterfeit side information (U_{wf} and V_{wf}). Using this side information, an attacker attempts to extract the counterfeit watermark W_f from the real owner watermarked image A_w with the following watermark extraction procedure:

Watermark extraction

- X1-II. Perform 1-level RDWT on the real owner watermarked image A_w as similar to Step X1:

$$A_w^{\theta *} \Leftarrow \Im\{A_w\} \quad \text{where} \quad \theta = (\text{LL, HL, LH, HH}), \tag{21}$$

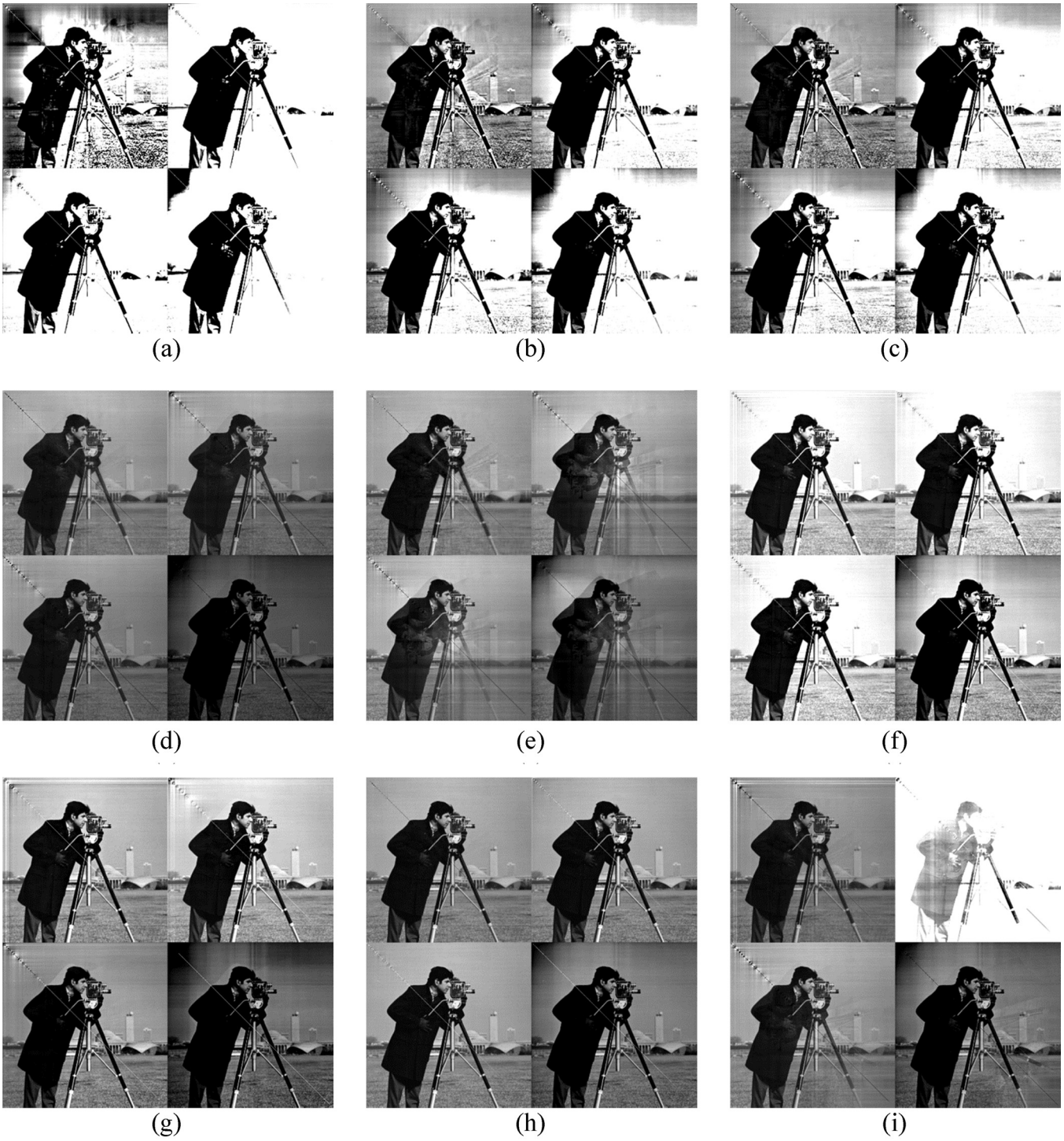

Fig. 6. Extracted watermark after: (a) speckle noise (0.4) (NCC = 0.70, 0.89, 0.90, 0.81); (b) Gaussian noise (0, 0.005) (NCC = 0.92, 0.86, 0.82, 0.83); (c) pepper and salt noise (0.001) (NCC = 0.95, 0.87, 0.82, 0.84); (d) median filter (3×3) (NCC = 0.98, 0.96, 0.97, 0.89); (e) JPEG compression ($Q = 40$) (NCC = 0.99, 0.94, 0.93, 0.86); (f) histogram equalization (NCC = 0.99, 0.99, 0.98, 0.91); (g) rotation (angle 50°) (NCC = 0.98, 0.99, 0.97, 0.77); (h) shift 50% (NCC = 1, 0.99, 0.99, 0.92); and (i) cut attack (NCC = 0.98, 0.77, 0.97, 0.90). The NCC is computed for each LL, LH, HL, and HH sub-band.

- X2-II. Perform SVD operation as similar to Step X2:

$$A_w^{\theta*} \Rightarrow U^{\theta*}\Sigma^{\theta*}V^{\theta*T}, \tag{22}$$

- X3-II. Obtain the matrix $D_{wf}^{\theta*}$ using the attacker side information (U_{wf} and V_{wf}) as similar to Step X3:

$$D_{wf}^{\theta*} \Leftarrow U_{wf}^{\theta}\Sigma^{\theta*}V_{wf}^{\theta T}, \tag{23}$$

- X4-II. Obtain the extracted counterfeit watermark image as similar to Step X4:

$$w_f^{\theta*} \Leftarrow \frac{1}{\alpha}(D_{wf}^{\theta*} - \Sigma_f^{\theta}). \tag{24}$$

Using this attack, an attacker can easily claim and prove the real owner watermarked image. An attacker can successfully extract the correct counterfeit watermark image with a high NCC value. Both

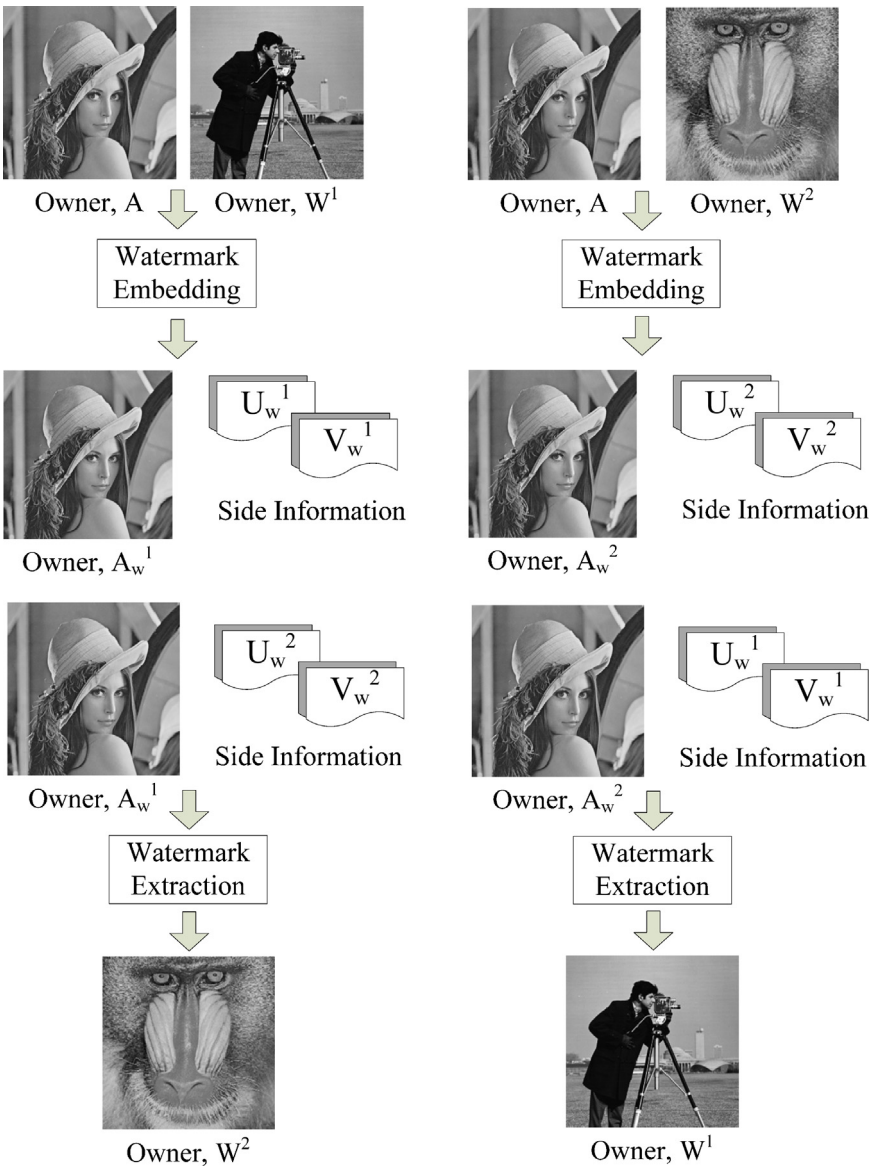

Fig. 7. Schematic diagram for attack type I.

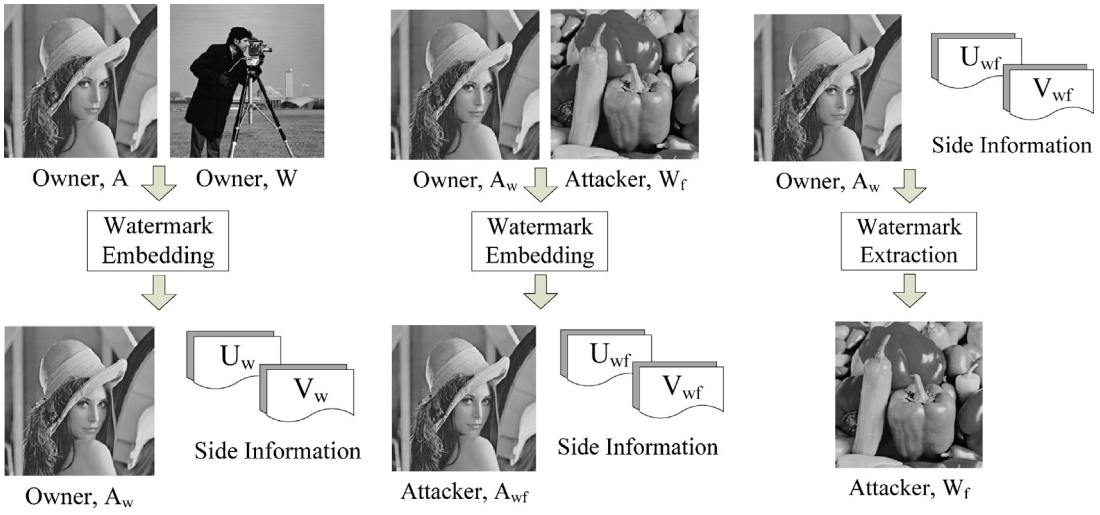

Fig. 8. Schematic diagram for attack type II.

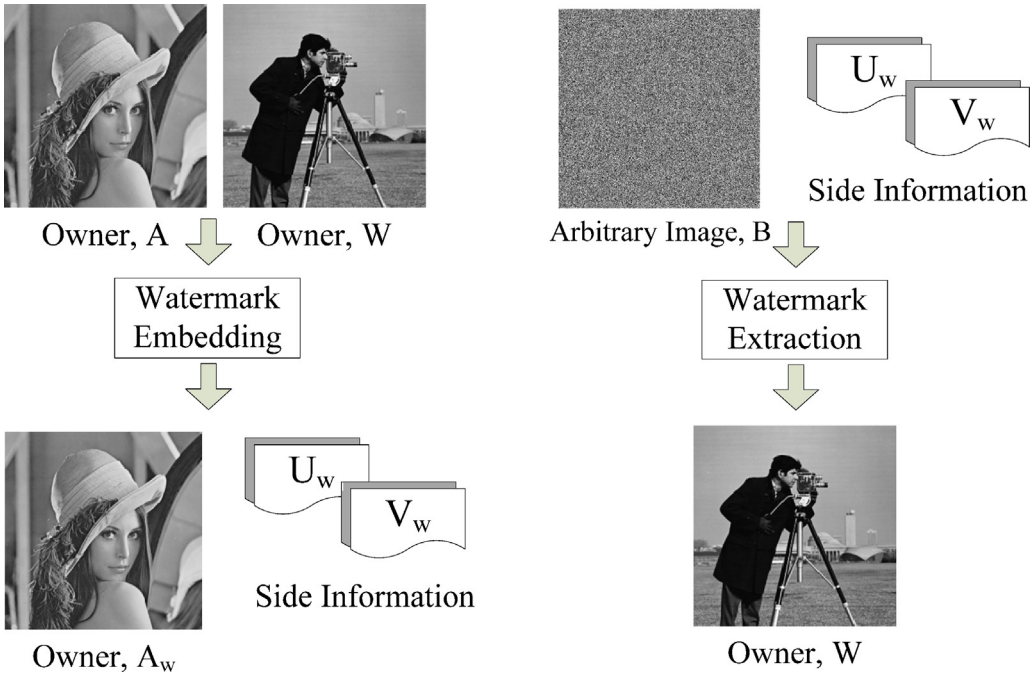

Fig. 9. Schematic diagram for attack type III.

of the real owner and attacker can claim the same image A as their image, and declare as the owner. This type of attack proves that the RDWT-SVD based image watermarking cannot solve the copyright protection problem.

3.3. Attack type III

In this type of attack, an image owner embeds the watermark image W into the host image A to obtain the watermarked image A_w. The side information (U_w and V_w) is also produced at the end of the watermark embedding process. Fig. 9 shows the schematic diagram for attack type III. The owner executes the following procedure to insert the watermark image:

Watermark embedding

- E1-III. Apply 1-level RDWT to the host image A as similar to Step E1:

$$A^{\theta} \Leftarrow \Im\{A\} \quad \text{where} \quad \theta = (LL, HL, LH, HH), \tag{25}$$

- E2-III. Apply SVD for all sub-bands as similar to Step E2:

$$A^{\theta} \Rightarrow U^{\theta}\Sigma^{\theta}V^{\theta T}, \tag{26}$$

- E3-III. Embed the watermark image W and then perform SVD as similar to Step E3:

$$\Sigma^{\theta} + \alpha W \Rightarrow U_w^{\theta}\Sigma_w^{\theta}V_w^{\theta T}, \tag{27}$$

- E4-III. Compute the modified RDWT coefficient as similar to Step E4:

$$A_w^{\theta} \Leftarrow U^{\theta}\Sigma_w^{\theta}V^{\theta T}, \tag{28}$$

- E5-III. Obtain the watermarked image as similar to Step E5:

$$A_w \Leftarrow \Im^{-1}\{A_w^{\theta} | \theta = (LL, HL, LH, HH)\}. \tag{29}$$

From the watermark embedding process, the owner obtains the side information (U_w and V_w) which can be used to extract the watermark information. In this attack, the owner makes effort to extract the watermark image using the side information (U_w and V_w) from an arbitrary image B using the following procedure.

Watermark extraction

- X1-III. Apply 1-level RDWT on arbitrary image B as similar to Step X1:

$$B^{\theta} \Leftarrow \Im\{B\} \quad \text{where} \quad \theta = (LL, HL, LH, HH), \tag{30}$$

- X2-III. Apply SVD operation as similar to Step X2:

$$B^{\theta} \Rightarrow U_B^{\theta}\Sigma_B^{\theta}V_B^{\theta T}, \tag{31}$$

- X3-III. Calculate the matrix $D_{wf}^{\theta *}$ using owner side information (U_w and V_w) as similar to Step X3:

$$D^{\theta *} \Leftarrow U_w^{\theta}\Sigma_B^{\theta}V_w^{\theta T}, \tag{32}$$

- X4-III. Generate the extracted watermark as similar to Step X4:

$$W^{\theta *} \Leftarrow \frac{1}{\alpha}(D^{\theta *} - \Sigma^{\theta}). \tag{33}$$

The extracted watermark W^{θ^*} has high correlation and visual similarity with the watermark W. However, the watermark W^{θ^*} is extracted from an arbitrary image which not actually contains the watermark information W. Using this attack, the owner can claim and extract the correct watermark from arbitrary image with high NCC value. Again, the RDWT-SVD watermarking fails to protect the copyright and ownership.

4. Analysis of RDWT-SVD image watermarking

The theoretical analysis of the RDWT-SVD image watermarking is presented in this section. The RDWT-SVD scheme is an improved version of the SVD-based watermarking proposed in [1]. Both methods embed a watermark directly into the singular value matrix of a host image. The SVD-based watermarking [1] performs SVD computation directly on a host image, whereas the RDWT-SVD scheme [15] applies SVD operation on the RDWT transform domain. Even though SVD-based watermarking in [1] seems robust against

attacks, a major flaw in the algorithm design was reported in [2,3]. The false positive problem and three vulnerable attacks presented in this paper also highly occur in [1].

The RDWT-SVD scheme [15] exhibits good performance against various attacks, yet it cannot withstand the three vulnerable attacks presented in this paper. In the watermark embedding strategy at Steps E3, E3-II, and E3-III, the results from $\Sigma + \alpha W$ is dominated by the geometrical feature of the watermark information W. Then, the side information (U_w and V_w) yielded from Eqs. (7), (18) and (27) obviously contains the eigenimages of the watermark information. In the normal situation, only the real owner can have and record the side information U_w and V_w. However, an attacker can easily generate the counterfeit side information (U_{wf} and V_{wf}) using his own watermark W_f as suggested in the attack type II.

The most important steps in the RDWT-SVD scheme are at Steps X3, X3-II, and X3-III in the watermark extraction stage. It is clear that the estimation of matrix D^* in Eqs. (12), (23) and (32) is geometrically determined by the side information matrices, U_w and V_w. Using an arbitrary singular value matrix Σ^*, the matrix $D^* \Leftarrow U_w \Sigma^* V_w^T$ will highly correlate with the watermark W. The extracted watermark W^* is simply computed by subtracting the matrix D^* with Σ. The matrix D^* is full matrix while Σ is diagonal matrix. Subtracting Σ from D^* only introduces little distortion on the diagonal value of the matrix D^*, meaning that a good extracted watermark estimation is still obtained using the singular value matrix from any arbitrary image.

Let A and B be a full matrix of size $N \times N$. The SVD operation decompose matrix A to be $A \Rightarrow U\Sigma_A V^T$. The matrix B is composed by changing the singular value matrix of A with any arbitrary singular value matrix, yielding $B \Leftarrow U\Sigma_B V^T$. The squared error sum between matrix A and B can be trivially derived as:

$$\begin{aligned}\left\|A, B\right\|_2^2 &= \left\|A(i,j) - B(i,j)\right\|_2^2,\\ &= \sum_{i=1}^{N}\sum_{j=1}^{N}(U\Sigma_A V^T - U\Sigma_B V^T)^2,\\ &= \sum_{i=1}^{N}\sum_{j=1}^{N}[\Sigma_A^2(i,j) + \Sigma_B^2(i,j) - 2\Sigma_A(i,j)\Sigma_B(i,j)],\\ &= \sum_{i=1}^{N}\sum_{j=1}^{N}[\Sigma_A(i,j) - \Sigma_B(i,j)]^2,\\ &= \left\|\Sigma_A, \Sigma_B\right\|_2^2.\end{aligned} \quad (34)$$

From Eq. (34), the squared error sum for the two matrices A and B is equal to the squared error sum between the two singular value matrices Σ_A and Σ_B. When Σ_A and Σ_B simply have a little difference, the value $\left\|A, B\right\|_2^2$ is close to zero, meaning that the two matrices A and B are nearly identical. In the special case, when A^T is the transposed version of matrix A, then the squared error sum between A and A^T is zero. Since the matrix A and its transpose have the same non-zero singular values, i.e. $A^T = [U\Sigma_A V^T]^T = V\Sigma_A U^T$. In the RDWT-SVD watermarking, the squared error sum between matrix $\Sigma + \alpha W$ and D^* is close to zero when the diagonal matrix is slightly different by employing the same side information (U_w and V_w) for both matrices.

The multiplication of A and B using element-wise operation in the SVD form is given as

$$\begin{aligned}\sum_{i=1}^{N}\sum_{j=1}^{N}A_{ij}B_{ij} &= \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{r}\lambda_k^A u_{ik} v_{jk}^T \sum_{k=1}^{r}\lambda_k^B u_{ik} v_{jk}^T,\\ &= \sum_{k=1}^{r}\lambda_k^A\lambda_k^B \sum_{i=1}^{N}u_{ik}u_{ik}\sum_{j=1}^{N}v_{jk}v_{jk}.\end{aligned}$$

The norm of A and B in the SVD domain is expressed as

$$\begin{aligned}\left\|A\right\| &= \sum_{i=1}^{N}\sum_{j=1}^{N}A_{ij}A_{ij},\\ &= \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{r}\lambda_k^A u_{ik} v_{jk}^T \sum_{k=1}^{r}\lambda_k^A u_{ik} v_{jk}^T,\\ &= \sum_{k=1}^{r}\lambda_k^A\lambda_k^A \sum_{i=1}^{N}u_{ik}u_{ik}\sum_{j=1}^{N}v_{jk}v_{jk}.\end{aligned}$$

$$\begin{aligned}\left\|B\right\| &= \sum_{i=1}^{N}\sum_{j=1}^{N}B_{ij}B_{ij},\\ &= \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{r}\lambda_k^B u_{ik} v_{jk}^T \sum_{k=1}^{r}\lambda_k^B u_{ik} v_{jk}^T,\\ &= \sum_{k=1}^{r}\lambda_k^B\lambda_k^B \sum_{i=1}^{N}u_{ik}u_{ik}\sum_{j=1}^{N}v_{jk}v_{jk}.\end{aligned}$$

As a result, the correlation coefficient between A and B can be computed as [4,5]:

$$\begin{aligned}\rho(A, B) &= \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}a_{i,j}\cdot b_{i,j}}{\left\|A\right\|\cdot\left\|B\right\|},\\ &= \frac{\sum_{r_1=1}^{r}\sigma_{r_1}^A\sigma_{r_1}^B\sum_{i=1}^{N}u_{i,r_1}^2\sum_{j=1}^{N}v_{j,r_1}^2}{\sqrt{\sum_{r_1=1}^{r}\sigma_{r_1}^A}\sqrt{\sum_{r_1=1}^{r}\sigma_{r_1}^B}\sum_{i=1}^{N}u_{i,r_1}^2\sum_{j=1}^{N}v_{j,r_1}^2},\\ &= \frac{\sum_{r_1=1}^{r}\sigma_{r_1}^A\sigma_{r_1}^B}{\sqrt{\sum_{r_1=1}^{r}\sigma_{r_1}^A}\sqrt{\sum_{r_1=1}^{r}\sigma_{r_1}^B}},\\ &= \rho(\Sigma_A, \Sigma_B).\end{aligned} \quad (35)$$

The correlation coefficient between A and B is equal to the correlation coefficient between Σ_A and Σ_B. When the diagonal matrix of Σ_A and Σ_B is slightly different, the $\rho(\Sigma_A, \Sigma_B) \approx 1$. The condition $A \approx B$ reveals that the two matrices are highly correlated. Thus, the matrix $\Sigma + \alpha W$ and D^* are slightly different when the side information (U_w and V_w) is identical and diagonal matrix is slightly different.

5. Secure RDWT-SVD image watermarking scheme

A new method, namely Secure RDWT-SVD (S-RDWT-SVD) image watermarking, is presented for avoiding three vulnerable attacks occurring in the RDWT-SVD image watermarking [15]. In this method, a visual watermark image is directly embedded into the principal component of RDWT-transformed host image. The host image is firstly transformed using RDWT. Afterward, the SVD is applied on the RDWT-transformed sub-band of host image to obtain the left and right singular matrix as well as the singular value matrix. The principal component matrix of host image can be easily computed by performing multiplication between left singular matrix with the singular value matrix. In this scenario, the right singular matrix obtained from watermark embedding is stored as side information which will later be used in the watermark extraction process to decode the watermark information. Let A and W be host image and visual watermark image of the same size $M \times N$, respectively. As similar with RDWT-SVD scheme [15], our proposed method employs a scaling factor in the watermark embedding. Formally, our watermark embedding and extraction strategy can be illustrated as follows:

5.1. Watermark embedding

In our proposed S-RDWT-SVD scheme, we avoid the SVD computation on Step E3 of the RDWT-SVD watermark embedding [15]. In fact, this step is the major flaw causing the weakness of RDWT-SVD watermarking scheme against three malicious attacks presented in Section 3. Fig. 13 illustrates the flowchart of the S-RDWT-SVD watermark embedding scheme. The watermark embedding of S-RDWT-SVD is as follows:

- E1-S: Firstly, we perform 1-level RDWT transformation on host image as:

$$A^\theta \Leftarrow \Im\{A\}, \quad (36)$$

where $\Im\{\cdot\}$ denotes the RDWT operation, and $\theta = (LL, HL, LH, HH)$. We may also substitute the RDWT transformation with the other orthogonal transformation such as DCT, DFT, etc.

- E2-S: We decompose A^{θ} using SVD operation to yield the following result:

$$A^{\theta} \Rightarrow U^{\theta}\Sigma^{\theta}V^{\theta T}, \tag{37}$$

- E3-S: Obtain the principal component of host image as:

$$A^{\theta}_{U\Sigma} \Leftarrow U^{\theta}\Sigma^{\theta}, \tag{38}$$

- E4-S: Embed the watermark image W into the principal component of host image by selecting a suitable scaling factor α as:

$$A^{d\theta}_{U\Sigma} \Leftarrow A^{\theta}_{U\Sigma} + \alpha W, \tag{39}$$

where $A^{d\theta}_{U\Sigma}$ denotes the distorted principal component of host image.

- E5-S: Obtain the modified RDWT coefficient for each sub-band as:

$$A^{\theta}_{w} \Leftarrow A^{d\theta}_{U\Sigma}V^{\theta T}, \tag{40}$$

- E6-S: Perform 1-level inverse RDWT for all sub-bands to obtain watermarked image as:

$$A_w \Leftarrow \Im^{-1}\{A^{\theta}_{w}|\theta = (LL, LH, HL, HH)\}, \tag{41}$$

where $\Im^{-1}\{\cdot\}$ denotes the inverse transformation process.

In proposed S-RDWT-SVD scheme, we only keep the right singular matrix (V^{θ}) and matrix $A^{\theta}_{U\Sigma}$ as side information for watermark extraction purpose. Using this strategy, the watermarked image can resist against three vulnerable attacks presented in Section 3.

5.2. Watermark extraction

As opposite to the embedding process, we perform watermark extraction from possibly corrupted watermarked image A^{*}_{w} using right singular matrix V^{θ} and $A^{\theta}_{U\Sigma}$ matrix as side information. Fig. 14 shows the watermark extraction process of the proposed S-RDWT-SVD scheme. Using this strategy, the S-RDWT-SVD watermark extraction scenario can be formally defined as:

- X1-S: Decompose a watermarked image A_w using 1-level RDWT transformation as follows:

$$A^{\theta *}_{w} \Leftarrow \Im\{A^{*}_{w}\}, \tag{42}$$

where $\Im\{\cdot\}$ denotes the RDWT operation, and $\theta = (LL, HL, LH, HH)$.

- X2-S: Using the right singular matrix V^{θ} as side information, we may obtain the distorted principal component as:

$$A^{d\theta *}_{U\Sigma} \Leftarrow A^{\theta *}_{w}V^{\theta}, \tag{43}$$

In this stage, we do not need to perform the transpose matrix or inverse matrix calculation for V^{θ} since of unitary (orthogonal) property in the left and right singular matrix.

- X3-S: Obtain the extracted watermark from distorted principal component as:

$$W^{\theta *} \Leftarrow \frac{1}{\alpha}\{A^{d\theta *}_{U\Sigma} - A^{\theta}_{U\Sigma}\}, \tag{44}$$

The matrix $A^{\theta}_{U\Sigma}$ is obtained from watermark embedding Step E3-S. The proposed S-RDWT-SVD scheme needs less step in the watermark embedding and extraction process compared with the former RDWT-SVD scheme [15]. In addition, the S-RDWT-SVD method is more secure than RDWT-SVD scheme against three vulnerable attacks presented in this paper.

6. Experimental results

In this section, several experiments are conducted to examine the robustness of the RDWT-SVD watermarking [15] against the three vulnerable attacks. The experiment is carried out using four images, Lena, Baboon, Peppers, and Cameraman, as shown in Fig. 3. The host and watermark images are the visual grayscale image of size 512×512. The quality of the watermarked image and extracted watermark are objectively measured in terms of the PSNR and NCC, respectively.

6.1. Attack type I

An experiment is conducted to investigate the robustness of the RDWT-SVD watermarking against attack type I. The Lena image is chosen as the host image. The cameraman and baboon are selected for watermark images. Two different experiments are conducted in the watermark embedding: (1) Embedding the cameraman watermark into Lena host image, and (2) inserting Baboon watermark into Lena host image. The watermarked Lena image and the side information are produced for each experiment. The side information is used to extract watermark from each watermarked Lena image.

Fig. 10(a) shows the watermarked image (PSNR = 53.77) when the Baboon is embedded into Lena host image. Fig. 10(c) shows the watermarked image (PSNR = 53.99) when Lena and cameraman are set as host and watermark image, respectively. The RDWT-SVD scheme performs well to render the watermark information into a host image as indicated with high PSNR value. Fig. 10(b) shows the extracted watermark from watermarked Lena image (a) using the side information from figure (c). The extracted watermark for each sub-band has very high NCC value. Yet, it is very important to highlight that Fig. 10(a) contains Baboon in the watermarked image, while the Cameraman is obtained in extraction stage, inducing ambiguity on the owner side when the wrong side information is used to extract the watermark.

A similar case also occurs when we extract the watermarked Lena image as shown in Fig. 10(c) using the side information from Fig. 10(a). The extracted watermark for each sub-band also yields high NCC value. The same ambiguity also appears as Fig. 10(c) containing Cameraman information, while the owner obtaining the Baboon watermark image. Thus, the RDWT-SVD scheme should not be used to protect the rightful ownership because of the ambiguity on the owner side.

6.2. Attack type II

In this experiment, the attack type II is applied to the RDWT-SVD image watermarking. Two parties are involved in this experiment, i.e. the real image owner and the counterfeit attacker. The real owner embeds the Cameraman into Lena host image for ownership protection applications. At the other side, the attacker only holds the Peppers as the counterfeit watermark. The attacker intends to claim the real owner's Lena watermarked image.

Fig. 11(a) and (b) shows the owner's and attacker's watermarks, respectively. Fig. 11(c) shows the watermarked Lena image (PSNR = 53.99) when the owner embeds Cameraman into Lena host image. An attacker can easily obtain and modify the Lena watermarked image which is already publicly published by the owner. An attacker tries to embed the counterfeit watermark (Fig. 11(b)) into the owner's watermarked Lena image (Fig. 11(c)). Fig. 11(d) shows the attacker's watermarked Lena image with PSNR = 49.43. An attacker keeps the side information produced from watermark embedding process to extract the watermark. Fig. 11(e) shows the extracted watermark when the attacker extracts the watermark from the attacker's watermarked Lena image (Fig. 11(d))

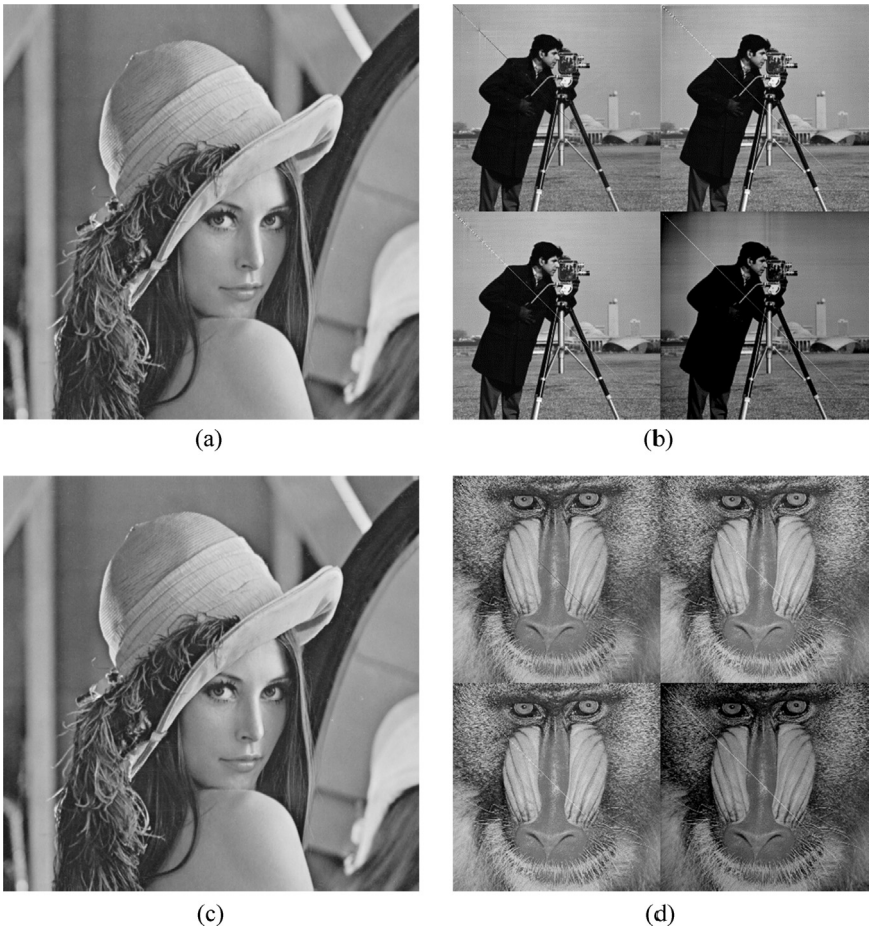

Fig. 10. Results of attack type I. (a) Watermarked image (host image: Lena, watermark: baboon); (b) extracted watermark using side information from (c) (NCC = 1, 0.99, 1, 0.93); (c) watermarked image (host image: Lena, watermark: cameraman); (d) extracted watermark using side information from (a) (NCC = 0.99, 0.99, 0.99, 0.95).

Fig. 11. Results of attack type II. (a) Owner's watermark image; (b) attacker's watermark image; (c) owner's watermarked image (PSNR 53.99 dB); (d) attacker's watermarked image (PSNR 49.43 dB); (e) extracted watermark from attacker's watermarked image (NCC = 0.99, 1, 1, 0.96); (f) extracted watermark from owner's watermarked image (NCC = 1, 0.99, 0.99, 0.97). The NCC is computed for each LL, LH, HL, and HH sub-band.

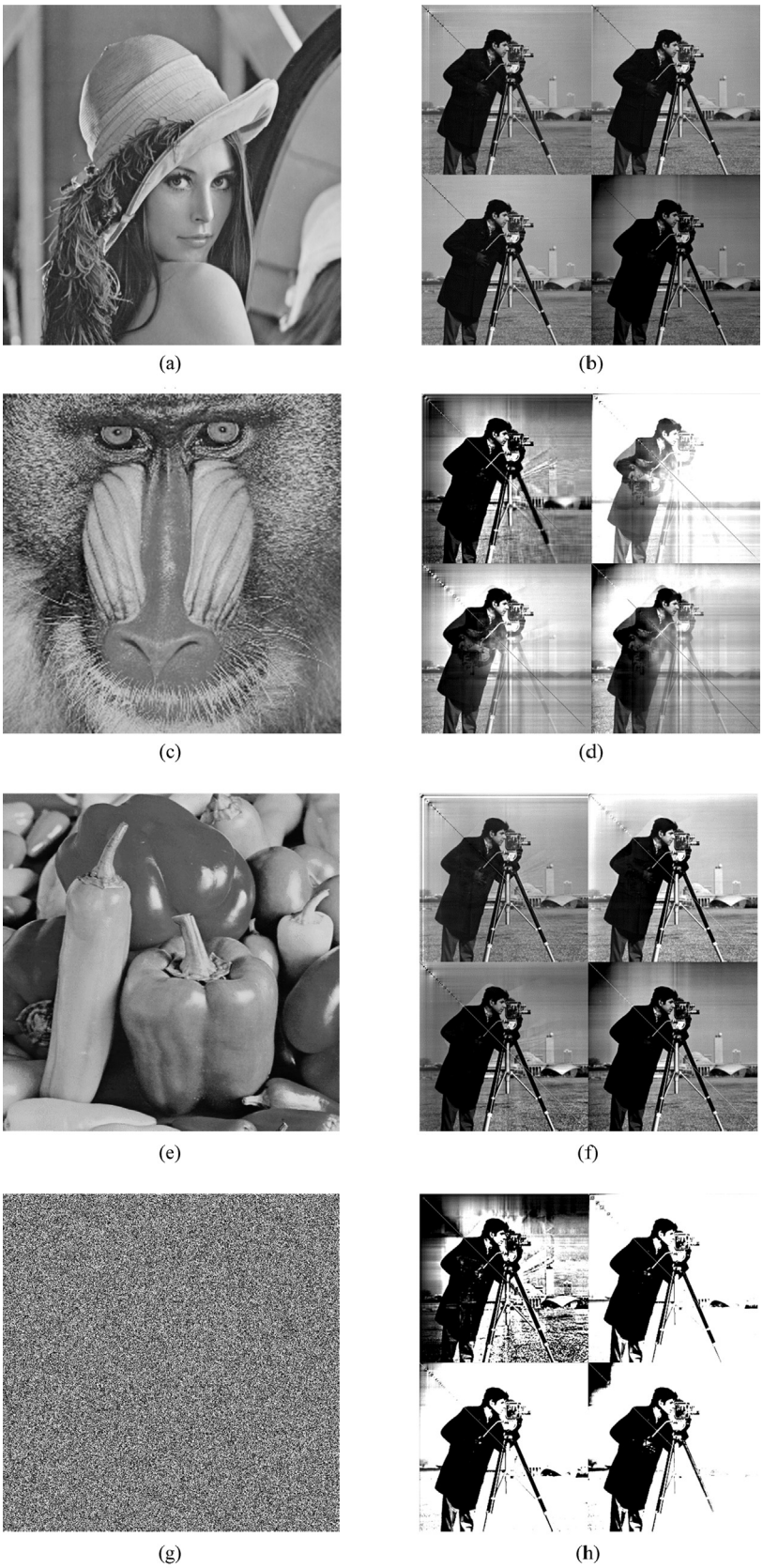

Fig. 12. Results of attack type III. (a) Unwatermarked Lena image; (b) extracted watermark from image (a) (NCC = 0.99, 1, 1, 0.90); (c) unwatermarked baboon image; (d) extracted watermark from image (c) (NCC = 0.81, 0.80, 0.85, 0.82); (e) unwatermarked peppers image; (f) extracted watermark from image (e) (NCC = 0.96, 0.98, 0.93, 0.85); (g) random image; (h) extracted watermark from image (g) (NCC = 0.59, 0.89, 0.90, 0.82).

using the attacker's side information. This serious problem occurs when the attacker extracts the owner's watermarked Lena image using attacker's side information. An attacker receives the extracted watermark, i.e. Peppers image, as shown in Fig. 11(f). Attacker can easily claim the owner watermarked image when the attacker has the counterfeit side information. The counterfeit side information can be generated from counterfeit watermark. From this experiment, it is clear that the RDWT-SVD scheme fails to protect the rightful ownership.

6.3. Attack type III

In this experiment, another drawback of the RDWT-SVD watermarking is demonstrated. The Lena and Cameraman are chosen as the host and watermark image, respectively. The RDWT-SVD scheme embeds the Cameraman into Lena host image to yield watermarked Lena image and the corresponding side information. Subsequently, we try to extract the watermark from an arbitrary image to judge the robustness and correctness of the RDWT-SVD watermarking. The effectiveness of the RDWT-SVD is determined whether it is able or not to extract the watermark from an arbitrary image.

Fig. 12(b) shows the extracted watermark from the unwatermarked Lena image (Fig. 12(a)) using the side information obtained from watermark embedding stage. Fig. 12(d) and (f) shows the extracted watermarks from unwatermarked Baboon image (Fig. 12(c)) and unwatermarked Peppers image (Fig. 12(e)), respectively. From Fig. 12(d) and (f), one can still obtain the correct extracted watermarks (Cameraman) with high NCC values. Fig. 12(h) shows the extracted watermark from the uniformly random-generated image as shown in Fig. 12(g). From Fig. 12(h), it still can be visually recognized that the extracted watermark is the Cameraman. Thus, we conclude that the RDWT-SVD watermarking has major fundamental flaw causing the false positive problem, as the owner or attacker can easily claim and extract the correct watermark from an arbitrary image. Consequently, the RDWT-SVD image watermarking scheme cannot solve the rightful ownership problem.

6.4. Performance test of the proposed S-RDWT-SVD watermarking scheme

We conducted the proposed S-RDWT-SVD image watermarking using Lena and cameraman images as host and watermark image, respectively. The size of the host and watermark image are identical, i.e. 512×512. For simplicity and without losing generality, we only perform watermark embedding in the LL sub-band of the RDWT transformed host image. However, the similar strategy can be easily extended into the other RDWT-transformed sub-bands, i.e. LH, HL, and HH. Figs. 15 and 16 show the S-RDWT-SVD watermarking result when the watermark image is inserted into the principal component of the RDWT-transformed host image by selecting scaling factor $\alpha = 0.1$ and $\alpha = 0.5$, respectively.

As shown in Figs. 15 and 16, the watermarked image quality is less acceptable for human vision even though it achieves good PSNR value, i.e. 35.0322. The presence of the watermark image can be easily perceived. The extracted watermark only produces a low NCC value (about 0.5) as shown in Figs. 15 and 16. From this experiment, the new strategy cannot meet the imperceptibility and robustness requirement for a good image watermarking scheme.

In the SVD-based watermarking scheme, the scaling factor plays an important role to control robustness as well as imperceptibility of the watermark image. We also conduct an additional experiment to further investigate the effect of scaling factor in the watermarked result (in terms of PSNR value) and the extracted watermark (in terms of NCC value) of the S-RDWT-SVD

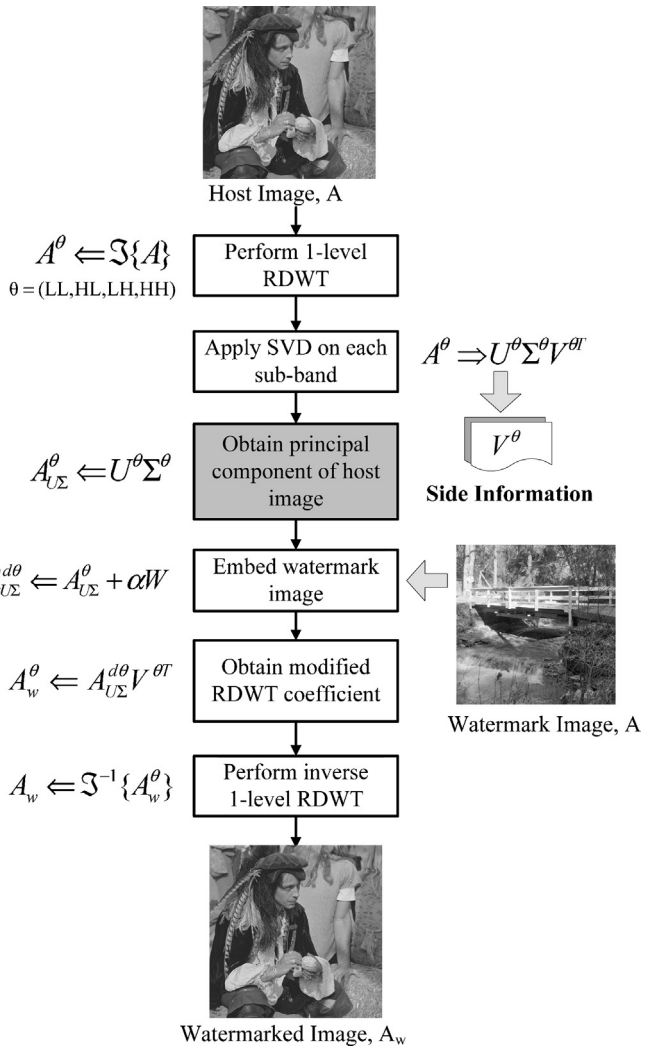

Fig. 13. Flowchart of the proposed S-RDWT-SVD watermark embedding scheme.

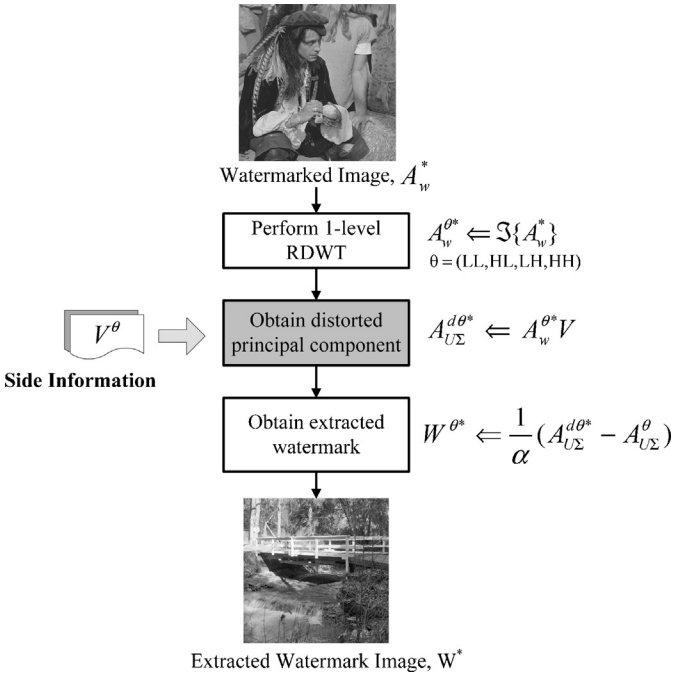

Fig. 14. Flowchart of the proposed S-RDWT-SVD watermark extraction scheme.

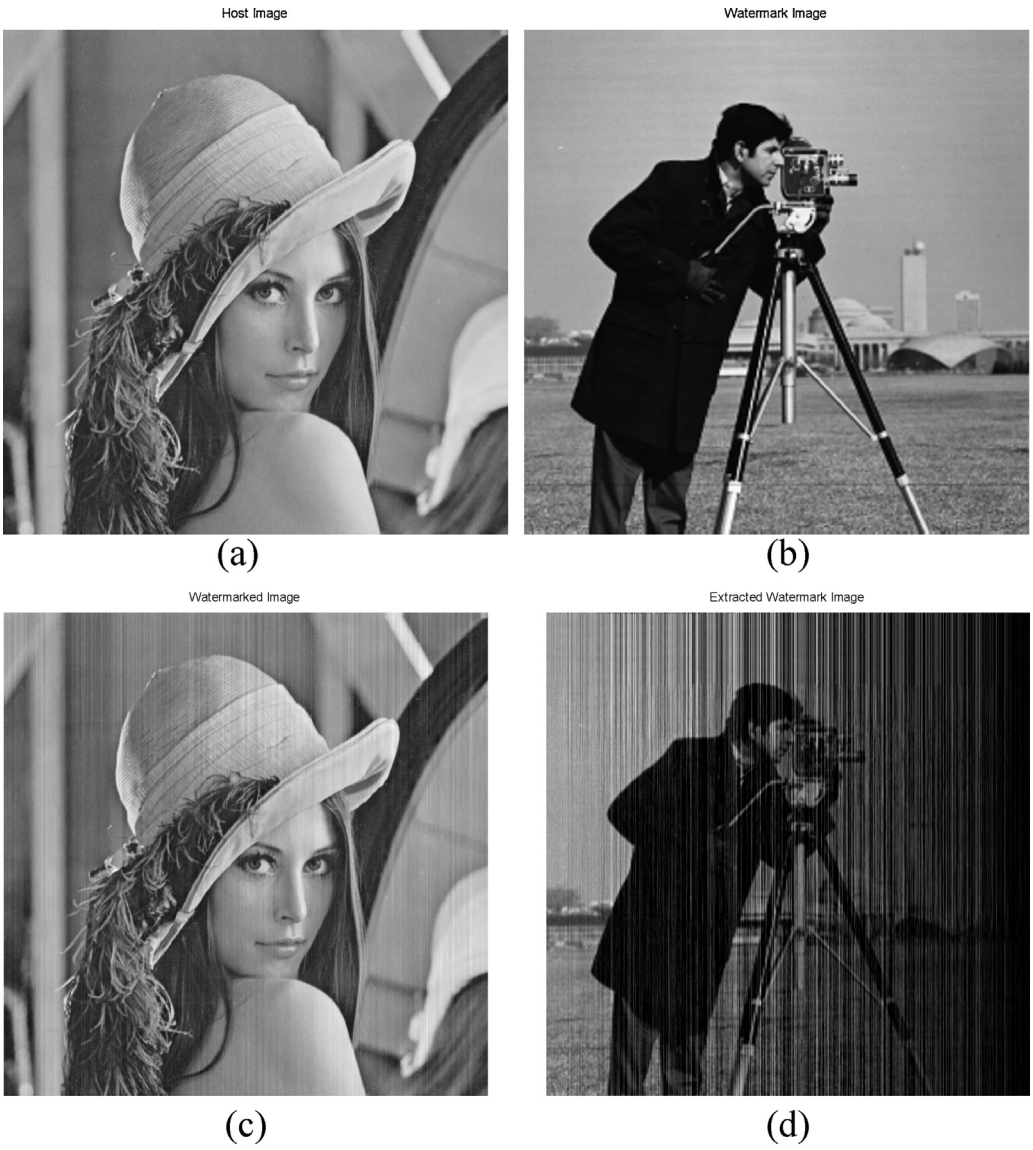

Fig. 15. Watermarking result using $\alpha = 0.1$: (a) original image, (b) watermark image, (c) watermarked image (PSNR = 35.0322), and (d) extracted watermark (NCC = 0.5067).

watermarking scheme. Fig. 17 shows the PSNR value of the watermarked image and NCC value of extracted watermark under different scaling factor settings, i.e. $-5 \leq \alpha \leq 5$. The PSNR value of watermarked image has acceptable PSNR when the scaling factor is in interval $0 < \alpha \leq 0.5$ or $-0.5 \leq \alpha < 0$. However, the NCC value is below or around 0.5 for different scaling factor parameter causing the extracted watermark image is less acceptable for human vision as shown in Fig. 17(b). Based on this experiment, it is very hard to find a suitable scaling factor to meet the imperceptibility and robustness aspects.

Subsequently, we justify the robustness of the new S-RDWT-SVD image watermarking scheme under three malicious attacks. Fig. 18 presents the result of attack type I. Fig. 18(a) and (c) shows the watermarked image when Baboon and cameraman, respectively, are embedded into Lena host image. Fig. 18(b) gives the extracted watermark (i.e. Baboon image) when Fig. 18(a) is extracted using side information from Fig. 18(c) (Lena + cameraman). In fact, the watermarked image in Fig. 18(a) contains the Baboon image information. Herein, we can see that the extracted watermark is Baboon image implying that we can solve the false positive problem in Attack I. The same conclusion can also be drawn for Fig. 18(d) which depicts the extracted watermark when the watermark image is extracted using side information from Fig. 18(a) (Lena + Baboon). We obtain the cameraman watermark image from Fig. 18(c). From this experiment, we conclude that the new S-RDWT-SVD watermarking scheme is robust against Attack I.

We also conduct an additional experiment to investigate the robustness of the new S-RDWT-SVD image watermarking scheme under Attack II. The experiment setting is set as similar to Section 6, part B. Fig. 19(a)–(c) shows the owner's watermark image, attacker's watermark image, and owner's watermarked image (host image: Lena, watermark image: cameraman), respectively. In this experiment, an attacker embeds the Peppers watermark image (as shown in Fig. 19(b)) into the owner's watermarked image (which is already available for public usage) resulting an attacker's watermarked image as given in Fig. 19(d). Fig. 19(e) shows an extracted watermark when an attacker extracts the watermark image from attacker's watermarked image. We can see that an attacker can obtain the correct watermark image (i.e. Peppers image). In this case, an attacker does not violate Attack II since the watermark image in Fig. 19(d) contains Peppers watermark information. However, an attacker gets nothing when he tries to extract the watermark from owner's watermarked image (Fig. 19(c)) as

Fig. 16. Watermarking result using $\alpha = 0.5$: (a) original image, (b) watermark image, (c) watermarked image (PSNR = 21.1126), and (d) extracted watermark (NCC = 0.5063).

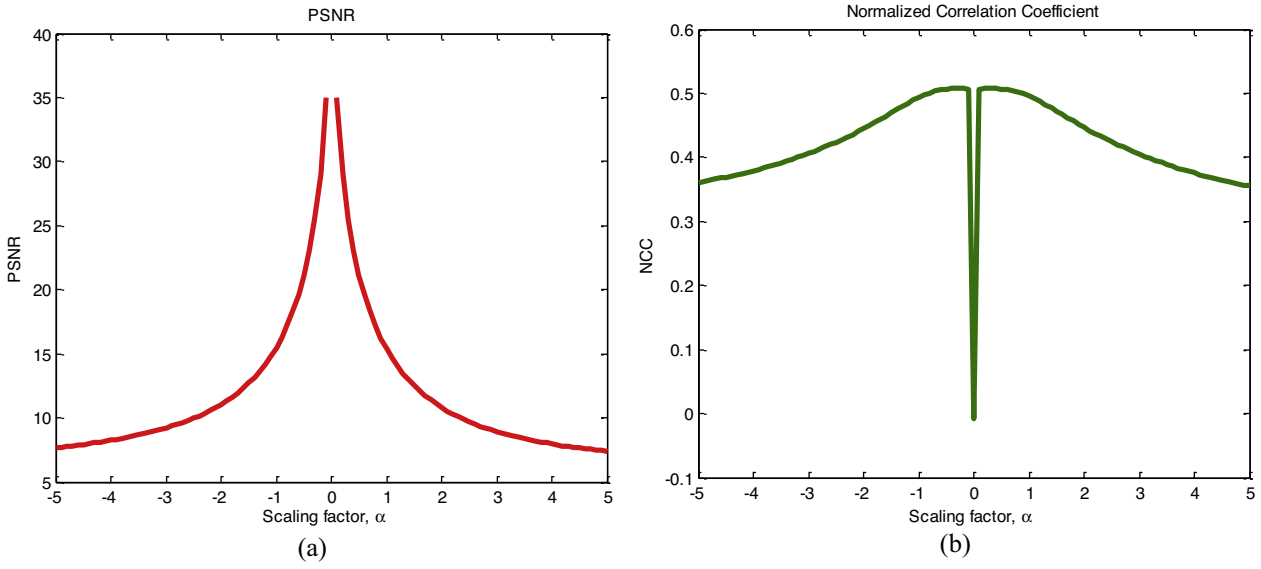

Fig. 17. Watermarking result: (a) PSNR and (b) NCC under different scaling factor.

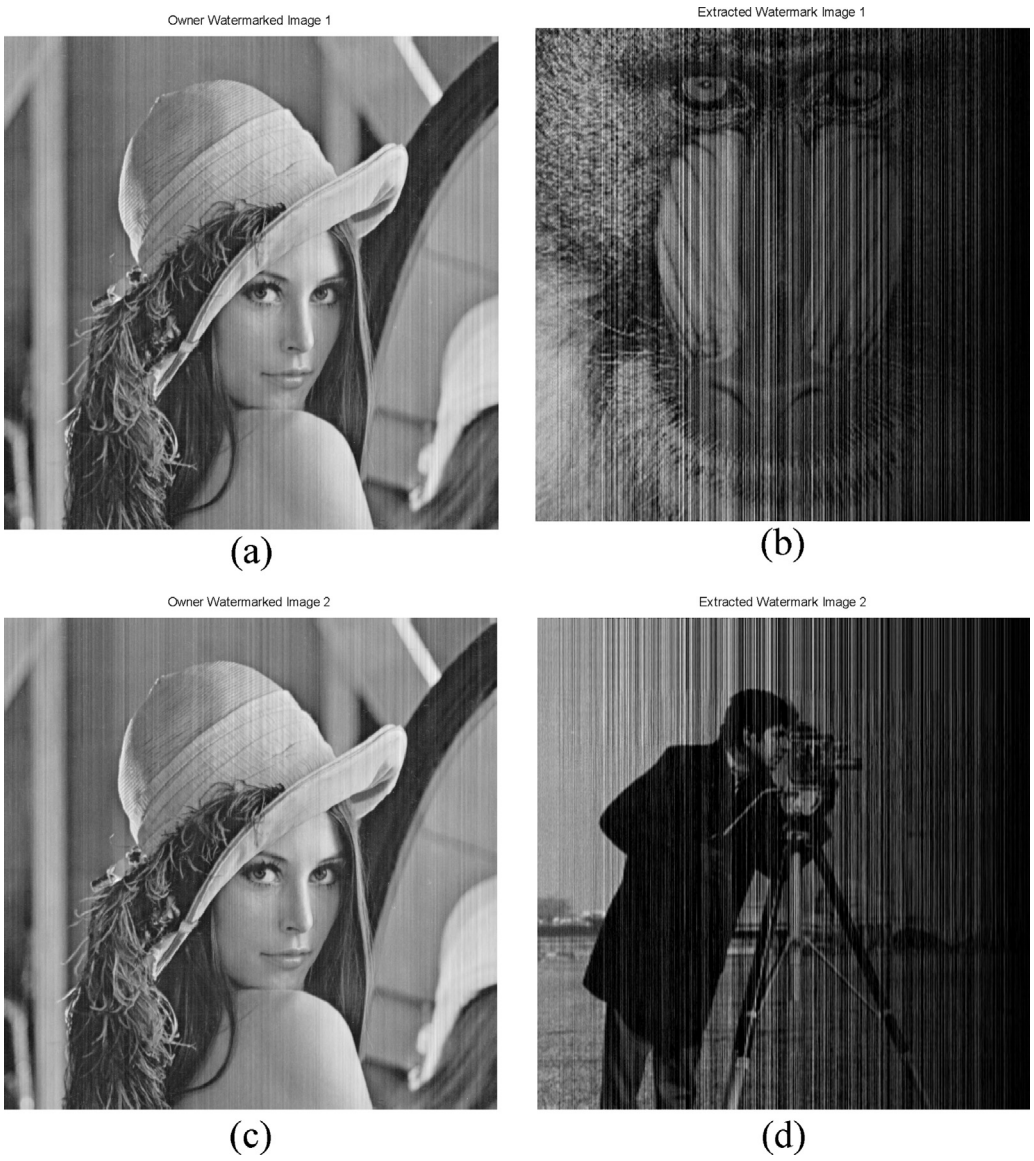

Fig. 18. Results of attack type I. (a) Watermarked image (host image: Lena, watermark: baboon); (b) extracted watermark using side information from (c); (c) watermarked image (host image: Lena, watermark: cameraman); (d) extracted watermark using side information from (a).

illustrated in Fig. 19(f). Based on this experiment, we can conclude that the new S-RDWT-SVD watermarking scheme is robust against Attack II.

An additional experiment was carried out to further investigate the robustness of the new S-RDWT-SVD image watermarking strategy against Attack III. In this experiment, an owner embeds the owner watermark image (i.e. cameraman image) into the Lena host image resulting in the side information (Lena + cameraman). An owner tries to extract the watermark information using this side information (Lena + cameraman) from Lena unwatermarked image, Baboon unwatermarked image, Peppers unwatermarked image, and random image as illustrated in Fig. 20(a), (c), (e) and (g). An owner gets nothing from this watermark extraction as shown in Fig. 20(b), (d), (f) and (h). From this experiment, we can avoid an Attack III using the new S-RDWT-SVD image watermarking strategy.

6.5. Comparison with other methods

Some literature reviews are conducted to show that the three vulnerable attacks presented in this paper also occur in the recently published SVD-based watermarking schemes [7–22]. The SVD-based image watermarking embeds watermark information into the singular value matrix of a host image. The host image is first transformed into wavelet transformed domain [7–15,17,19] before performing the SVD, or directly applies SVD to the host image [16,18,20–22]. Two possible ways are used to render the watermark image: (1) Only embedding the watermark singular value [7–14], or (2) hiding the watermark into the singular value matrix of a host image [15–22]. The suited scaling factor is chosen to meet the robustness and imperceptibility aspect in the SVD-based watermarking.

The SVD image watermarking is not only robust against several image manipulations and geometric distortions, but also it is able to hide the watermark information with high payload as reported in [7–22]. Table 1 shows comparisons on various SVD-based image watermarking schemes. The RDWT-SVD based watermarking has the highest watermark embedding capacity [9,14,15] compared with the other schemes by exploiting the redundancy in RDWT. The RDWT-SVD scheme has good performance which inserts the visual grayscale watermark image into all subbands of the RDWT-transformed host image.

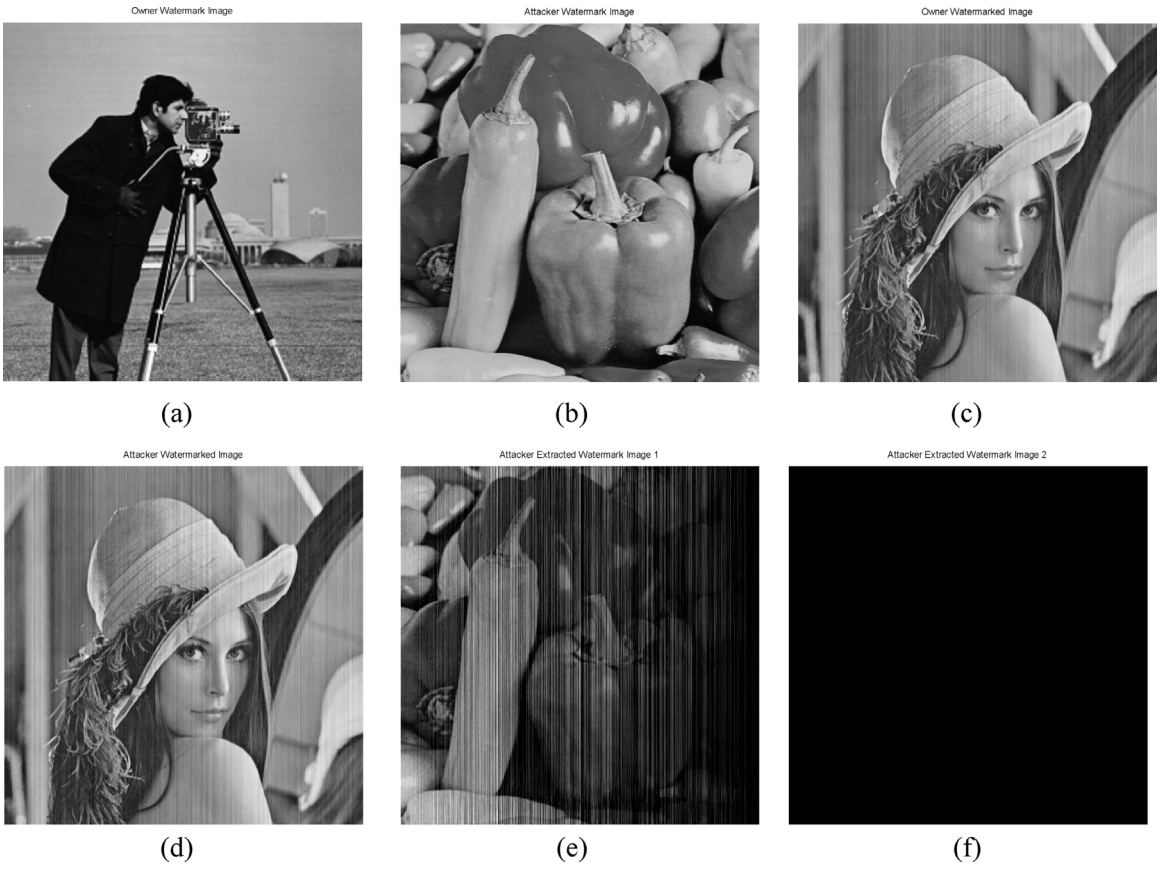

Fig. 19. Results of attack type II. (a) Owner's watermark image; (b) attacker's watermark image; (c) owner's watermarked image; (d) attacker's watermarked image; (e) extracted watermark from attacker's watermarked image; (f) extracted watermark from owner's watermarked image.

Table 1
Comparisons on various SVD image watermarking schemes.

Method	Type of transform	Embedding sub-bands	Perform SVD on watermark	Size of host image	Size of watermark	Type of watermark	Watermark false positive problem	Attack type I	Attack type II	Attack type III
[7]	Finite Radon Transform + DWT + SVD	All, LH and HL	Yes	257×257	33×33	Binary	Yes	No	No	No
[8]	DWT + SVD	All	Yes	512×512	256×256	Grayscale	Yes	No	No	No
[9]	Chaotic map + Hessenberg decomposition + RDWT + SVD	LL	Yes	256×256	256×256, 512×512	Binary + Grayscale	Yes	No	No	No
[10]	Fractional Dual Tree Complex Wavelet Transform + SVD	All	Yes	256×256	64×64	Grayscale	Yes	No	No	No
[11]	DWT + SVD	LH	Yes	1024×1024	256×256	Grayscale	Yes	No	No	No
[12]	Fractional Wavelet Packet Transform + SVD	All	Yes	512×512	128×128	Grayscale	Yes	No	No	No
[13]	Redundant Fractional Wavelet Transform + SVD	LL	Yes	256×256	256×256	Binary + Grayscale	Yes	No	No	No
[14]	RDWT + SVD	All	Yes	512×512	512×512	Grayscale	Yes	No	No	No
[15]	RDWT + SVD	All	No	512×512	512×512	Grayscale	No	Yes	Yes	Yes
[16]	SVD	No	No	256×256	32×32	Grayscale	No	Yes	Yes	Yes
[17]	DWT + SVD	HH, HL, LH	No	352×288 (video)	198×54, 108×33, 66×18, 144×42, 84×24, 54×12	Binary	No	Yes	Yes	Yes
[18]	SVD	No	No	256×256	64×64	Grayscale	No	Yes	Yes	Yes
[19]	DWT + SVD	LH, HL	No	256×256	128×128	Grayscale	No	Yes	Yes	Yes
[20]	SVD	No	No	256×256	32×32	Grayscale	No	Yes	Yes	Yes
[21]	SVD	No	No	–	–	–	No	Yes	Yes	Yes
[22]	SVD	No	No	– (audio)	256×256	Grayscale	No	Yes	Yes	Yes
Proposed	RDWT-SVD	LL	No	512×512	512×512	Grayscale	No	No	No	No

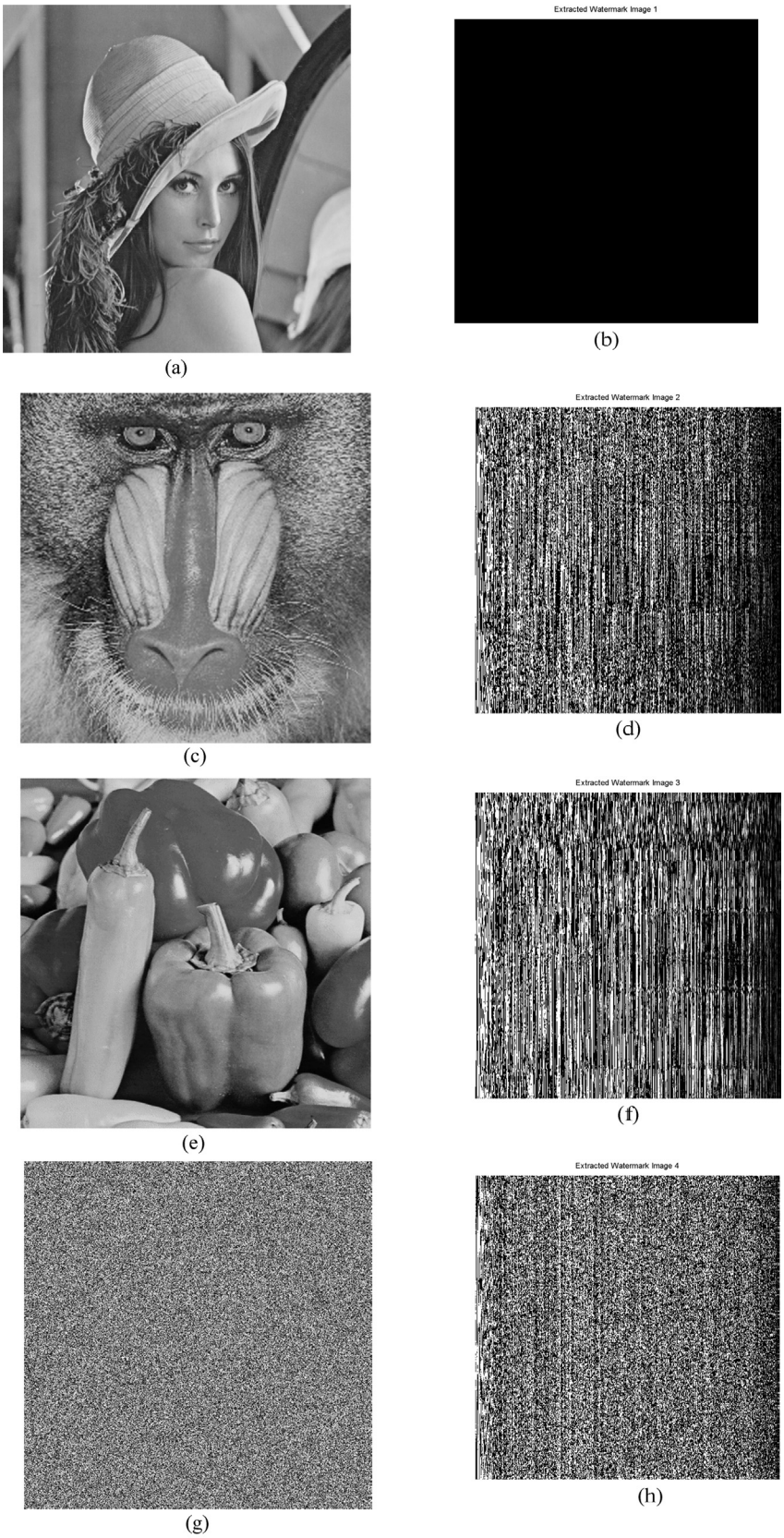

Fig. 20. Results of attack type III. (a) Unwatermarked Lena image; (b) extracted watermark from image (a); (c) unwatermarked baboon image; (d) extracted watermark from image (c); (e) unwatermarked peppers image; (f) extracted watermark from image (e); (g) random image; (h) extracted watermark from image (g).

However, the SVD-based image watermarking fails to meet the requirement for proof of ownership and copyright protection. The false positive problem occurs in the SVD-based image watermarking. When the watermark singular value is inserted into the singular value matrix of the host image [7–14], the false positive problem and the vulnerable attacks presented in [6] occur with high probability. Yet, the three vulnerable attacks presented in this paper were not happened in [7–14]. Consequently, the SVD-based image watermarking in [7–14] cannot provide the trusty guidance for ownership and copyright protection. The new S-RDWT-SVD based image watermarking can resist against three vulnerable attacks presented in this paper, but selecting the proper scaling factor is still the main problem in the S-RDWT-SVD scheme to satisfy the imperceptibility and the watermark robustness aspect.

If the watermark is directly embedded into the singular value matrix of a host image [15–22], the three vulnerable attacks presented in this paper will definitely occur. As confirmed in Table 1, the false positive problem in the watermark extraction [6] were not happened in [15–22]. Thus, the SVD-based image watermarking proposed in [15–22] is not suited for real applications which require the copyright and ownership protection. In addition, the robustness of the SVD-based image watermarking does not exist because of incorrect design in the watermark embedding and extraction strategy.

7. Conclusion

Three vulnerable attacks have been presented to show the weakness of the redundant discrete wavelet transform and singular value decomposition image watermarking scheme. The RDWT-SVD scheme can embed the watermark information with a high capacity and little degradation on the image quality. It meets the imperceptibility and robustness aspects for good watermarking design requirements. As reported in former studies, the RDWT-SVD watermarking is robust against various common image manipulations (e.g. filtering, compression, noise additive, etc.) and geometric distortions (e.g. rotation, shifting, cutting, etc.). Yet, the RDWT-SVD scheme cannot resist against the three vulnerable attacks presented in this paper. As documented in Section 6, the severe false positive issue occurs in the RDWT-SVD watermarking extraction. Thus, the RDWT-SVD scheme in fact cannot provide trustworthy evidence in rightful ownership protection. In addition, the robustness advantage of the RDWT-SVD scheme in fact is a result of improper algorithm design.

References

[1] Liu RZ, Tan. TN. An SVD-based watermarking scheme for protecting rightful ownership. IEEE Trans Multimedia 2002;4(1):121–8.
[2] Zhang XP, Li. K. Comments on an SVD-based watermarking scheme for protecting rightful ownership. IEEE Trans Multimedia 2005;7(3):593–4.
[3] Rykaczewski R. Comments on SVD-based watermarking scheme for protecting rightful ownership. IEEE Trans Multimedia 2007;9(2):421–3.
[4] Changzhen X, Fenhong G, Zhengxi L. Weakness analysis of singular value based watermarking. In: Int. Conf. Mechantronics and Automation. 2009. p. 2596–601.
[5] Changzhen X, Ward RK, Xu J. On the security of singular value based watermarking. In: IEEE Int. Conf. Image Proc. 2008. p. 437–40.
[6] Sadek RA. Blind synthesis attack on SVD based watermarking techniques. In: Int. Conf. Computational Intelligence for Modeling Control and Automation. 2008. p. 140–5.
[7] Rastegar S, Namazi F, Yaghmaie K, Aliabadian A. Hybrid watermarking based on singular value decomposition and Radon transform. AEU-Int J Electron Commun 2011;65(7):658–63.
[8] Ganic E, Eskicioglu AM. Robust embedding of visual watermarks using DWT-SVD. J Electron Imaging 2005;14(4):043004.
[9] Bhatnagar G. A new facet in robust digital watermarking framework. AEU-Int J Electron Commun 2012;66(4):275–85.
[10] Bhatnagar G, Wu QMJ. Biometrics inspired watermarking based on a fractional dual tree complex wavelet transform. Future Gener Comput Syst 2013;29(1):182–95.
[11] Song C, Sudirman S, Merabti M. A robust region-adaptive dual image watermarking technique. J Vis Commun Image Represent 2012;23(3):549–68.
[12] Bhatnagar G, Wu QMJ, Raman B. A new robust adjustable logo watermarking scheme. Computer Security 2012;31(1):40–58.
[13] Bhatnagar G, Wu QMJ, Raman B. A new logo watermarking based on redundant fractional wavelet transform. Math Comput Model 2012, http://dx.doi.org/10.1016/j.mcm.2012.06.002.
[14] Lagzian S, Soryani M, Fathy M. Robust watermarking scheme based in RDWT-SVD: embedding data in all subbands. In: Int. Symp. Artificial Intelligence and Signal Processing. 2011. p. 48–52.
[15] Makbol NM, Khoo. BE. Robust blind image watermarking scheme based on redundant discrete wavelet transform and singular value decomposition. AEU-Int J Electron Commun 2013;67(2):102–12.
[16] Aslantas V. A singular-value decomposition-based image watermarking using genetic algorithm. AEU-Int J Electron Commun 2008;62(5):386–94.
[17] Faragallah OS. Efficient video watermarking based on singular value decomposition in the discrete wavelet transform domain. AEU-Int J Electron Commun 2013;67(3):189–96.
[18] Lai. CC. A digital watermarking scheme based on singular value decomposition and tiny genetic algorithm. Digit Signal Process 2011;21(4):522–7.
[19] Lai CC, Tsai. CC. Digital image watermarking using discrete wavelet transform and singular value decomposition. IEEE Trans Instrum Meas 2010;59(11):3060–3.
[20] Aslantas V. An optimal robust digital image watermarking based on SVD using differential evolution algorithm. Opt Commun 2009;282(5):769–77.
[21] Dogan S, Tuncer T, Avci E, Gulten A. A robust color image watermarking with singular value decomposition method. Adv Eng Softw 2011;42(6):336–46.
[22] Al-Nuaimy W, El-Bendary MAM, Shafik A, Shawki F, Abou-El-azm AE, El-Fishawy NA, et al. An SVD audio watermarking approach using chaotic encrypted images. Digit Signal Process 2011;21(6):764–79.