Comments on “A Semi-blind Digital Watermarking Scheme Based on Singular Value Decomposition”

一种基于单值分解的半盲数字水印算法评价

1 Introduction
In recent years, many image watermarking schemes based on SVD have been proposed. In [1], Chung et al. proposed an image hiding scheme based on SVD and VQ (Vector Quantization) which leads to good compression ratio and satisfactory image quality by means of plugging the VQ technique into the SVD-based compression method. In [2], Chang et al. proposed a SVD-based image watermarking scheme using both of the U and S matrices to embed the watermark. In [3], a hybrid non-blind watermarking scheme combining SVD and DWT (Discrete Wavelet Transform) was proposed by Ganic et al. The host image is first decomposed into four bands using DWT, and then the SVD operation is performed on each band. The same process is performed on the watermark too. Finally, the singular values of the host image are modified with those of the watermark.

In [4], the authors proposed a robust semi-blind digital watermarking scheme for hiding grayscale watermarks into digital images. This method embeds the singular values (SVs) of the original image into the watermark by plugging the codebook concept into the SVD. We argue that this watermarking method is non-effective and has a very high probability of false positive detection. Finally, we draw a conclusion in Section 4 summarizing the weaknesses of this category of watermarking schemes.

2 An Overview of the Method
In the approach proposed in [4], the original image and the watermark are both grayscale and of the same size N×N.

The SVD of an N×N image matrix A is computed by:

\[ A = U S V^T \]  \hspace{1cm} (1)

where U and V are N×N orthogonal matrices, and S is an N×N diagonal matrix built with the singular values of matrix A. A brief overview of the algorithm is as follows.

2.1 Embedding Procedure
Step 1. Chaotic mixing of the images
For an image, the chaotic transformation is calculated as follows:

\[ \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ k & k+1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \pmod{N} \]  \hspace{1cm} (2)

where \( k \in \{1, N\} \in \mathbb{Z} \), \((x, y)\),and \((x', y')\) are the positions of the pixels before and after the transformation. Then Eq. (2) is used to obtain the scrambled versions of AC and WC of the original image A and the watermark W respectively.
where $key_A$ and $key_W$ are the encryption keys of the chaotic mixing process for $A$ and $W$ respectively.

**Step 2. Block-based SVD of the images**

\[
AC_i = \text{Scramble}(A, key_A) \quad (3)
\]
\[
WC_i = \text{Scramble}(W, key_W) \quad (4)
\]

where $AC_i$ and $WC_i$ are the $i$th and $j$th blocks of $AC$ and $WC$ respectively and $1 \leq i, j \leq (N/M) \times (N/M)$.

**Step 3. Embedding of the chaos-permuted watermark**

The smallest squared Euclidean distance between $AC_i$ and $WC_j$ is calculated as:

\[
d(AC_i, WC_j) = \min \sum_{r=1}^{M} (AC_{i_r} - WC_{j_r})^2 \quad (7)
\]

The first matched vector $\hat{WC}_j$ is replaced with the vector $AC_i$ and the replaced-vector $\hat{WC}_j$ will not be modified in the remaining search. Then the one-to-one mapping relationship of $AC_i \leftrightarrow WC_j$ is recorded as a file called BM for extraction.

**Step 4. Obtaining the secret image**

\[
WC_j^p = U_{WC_j} S_{AC_i} V_{WC_j}^T \quad (8)
\]

where $S_{AC_i}$ is the modified $S_{WC_j}$ obtained by Eq. (7).

## 2.2 Extracting Procedure

First, perform Chaotic mixing and Block-based SVD on the suspected image $A'$ and use the mapping file BM to attain $\hat{S}_{AC_i}$, then the possible distorted watermark $\hat{W}$ is extracted as follows:

\[
\hat{WC}_j^p = U_{WC_j} S_{AC_i} V_{WC_j}^T \quad (9)
\]
\[
\hat{W} = \text{Unscramble}(\hat{WC}_j^p, \hat{key}_W) \quad (10)
\]

where $1 \leq i, j \leq (N/M) \times (N/M)$ and $\hat{key}_W$ is the related decryption key in Eq.(4). Note that the matrices $U_{WC_j}$, $V_{WC_j}$ and BM are required in watermark extraction.

## 2.3 Argument

In the embedding procedure, the SVs of the watermark image are replaced with those of the original one according to Eq. (7). It is obvious that the modified $S_{WC_j}$ is only slightly different from before. As discussed in [5], singular values do not contain necessary information for face recognition, instead, all the useful information is contained in the orthogonal matrices of the SVD-decomposition of the image. We can say that the matrices of $U_{WC_j}$ and $V_{WC_j}$ represent the SVD subspace of the WC and can preserve the most useful information of the image as used in Eqs. (8) and (9). Therefore, in the watermark extraction, no matter what value the diagonal matrix $S_{AC_i}$ takes, the extracted watermark is in the same SVD subspace defined by the scrambled original watermark WC. In other words, Eq. (9) effectively preserves the major information of watermark $W$ on $\hat{WC}_j^p$ no matter what the suspected image $A'$ is. Hence, the extracted watermark $\hat{W}$ is determined by the reference watermark $W$ whatever $A'$ is. In fact, the great robustness of the proposed approach in [4] is a result of false algorithm design.

Analogical problem can be found in paper [6]. The authors propose a novel watermarking algorithm based on SVD by embedding the watermark matrix $W$ into the matrix $S$ of the original image by means of SVD decomposition. The authors also argue that it performs well in both security and robustness. But it has been proven to be inaccurate. The experiments in [7] and [8] show that the extracted watermark is not the embedded one but determined by the reference watermark because of false conception to insert watermark information into image singular values.

## 3 Experiments and Discussion

In this section, three groups of experiments are carried out to demonstrate the improper design of the proposed approach. To be consistent with [4], all the images used here are grayscale and of the same size 256×256, which are then divided into blocks of size 32×32. The image “Baboon” is used as the original image $A$ and “Lake” is used as the watermark $W$, as shown in Fig. 1(a) and Fig. 1(b) respectively.

![Fig.1. (a) Original image “Baboon”; (b) Watermark image “Lake”](image)

### 3.1 Experiments According to the Proposed Method by Shich et al.

In this experiment, we follow the steps described in Section 2. The results are shown in Fig. 2. The suspected images are listed in the left column, with the PSNRs over...
the original image “Baboon” given below each image. The corresponding extracted watermarks together with their PSNRs over the original watermark are shown in the right column.

Ideally, the watermark “Lake” should not be obtained from Fig. 2(e) and Fig. 2(g). However, as shown in Fig. 2(f) and Fig. 2(h), the watermarks “Lake” have been extracted all the same. In fact, Fig. 2 shows that highly recognizable watermarks can be extracted from whatever suspected images. Surprisingly, the watermark extracted from the image “Lena” has an even higher PSNR value than the watermark extracted from the original image “Baboon”.

Note that the experiments reported in [4] only use the “Baboon” image and its attacked version as the input of the detector, while more experiments on various images are carried out in this paper. Therefore the experiments employed in [4] are obviously flawed.

### 3.2. The Influences of the Chaotic Mixing on Extraction

To find out how the chaotic mixing step influences the extraction process, we intentionally skip the chaotic mixing process of both the original image and the suspected image. Then we carry out experiments similar to those of Section 3.1. The extracted watermarks together with their PSNRs are shown in Fig. 3.

As can be seen from Fig. 3, the fidelity of the retrieved watermarks has all been degraded. However, the watermarks are still recognizable although the watermark
extracted from the cropped image “Baboon” has suffered the most serious degradation.

Note that the experiments in [4] have performed the chaotic transformation on the original image and then the block-based SVD. The scrambled image distributes nearly randomly. Therefore the values of $S_{dc}$ and $S_{wc}$ have similar distribution and $S_{wc}$ changes slightly after performing Eq. (7). The properties of SVD of a digital image indicate that each singular value (SV) specifies the luminance of the SVD image layer, whereas the respective pair of singular vectors specifies image geometry. Since we skip over the chaotic transformation process in our experiment, the $S_{wc}$ is much different from before by performing Eq. (7). However, since the most useful information is preserved in the orthogonal matrices, the watermarks can still be extracted.

3.3 The Influences of the key $A$ on Extraction

To validate the effect of the key used in the method proposed in [4], we intentionally use different keys in the embedding and extraction procedures respectively. The extracted watermarks together with their PSNRs are shown in Fig. 4. Surprisingly, the quality of the extraction is almost the same as that observed in Section 3.1. It is obvious that the $key_A$ used for encryption does not work well. The keys can not be used to protect the copyright of the objective.

4 Conclusion

In this paper, we have made a thorough analysis of the watermarking scheme proposed in [4]. Extensive experiments are implemented to support our arguments. According to the experiments described in the previous section, we come to the following conclusions:

1) Based on the scheme proposed in [4], the extraction of watermark is in fact independent on the input suspected images. We can extract a recognizable watermark from whatever suspected images.

2) The scheme of [4] proposes to perform a chaotic transformation processing on the original image and the suspected image. Our experimental results show that such processing has very limited influence on the extraction of watermark.

3) The scheme of [4] proposes to use an encryption key to protect the copyright of the image. Unfortunately, we have found in our experiments that the extraction results are always desirable despite the use of a wrong key, which means that the key doesn’t work in the scheme.

Evidently, the alleged great robustness of the scheme of [4] against a variety of attacks does not exist. The scheme actually relies on the information that is dependent on the reference watermark during the extraction process. The scheme could result in a very high false positive detection rate and thus can not be used to protect the copyright of the objective.

REFERENCES


Fig.4. (a) watermark obtained from Fig. 2(a); (b) watermark obtained from Fig. 2(c); (c) watermark obtained from Fig. 2(e); (d) watermark obtained from Fig. 2(g).

作者简介

郑为民 哈尔滨工业大学博士，副研究员，硕士生导师，科研处副处长。2001年在哈尔滨工业大学自动化测试与控制系获测试计量技术及仪器专业工学博士学位。2001年7月至2003年7月在中国科学院计算技术研究所系统结构研究所龙芯处理器研制组做博士后研究。2004年由于在龙芯CPU研制中的贡献获得中国科学院杰出成就奖。2007至2008年在香港科技大学电子与计算机工程系做访问学者。已有10项发明专利被受理，在国内外期刊及会议上发表论文28篇。