Multimedia Tools and Applications (2023) 82:22275–22295
https://doi.org/10.1007/s11042-023-15248-6

1187: RECENT ADVANCES IN MULTIMEDIA INFORMATION SECURITY: CRYPTOGRAPHY AND STEGANOGRAPHY

A comprehensive analysis of the false-positive problem in SVD-based image watermarking

Narges Safizadeh[1] · **Sayyed Hamid Reza Ahmadi**[1]

Received: 11 October 2020 / Revised: 29 January 2023 / Accepted: 6 April 2023 /
Published online: 18 April 2023

Abstract
In recent years, two similar formulations have been proposed for semi-blind robust SVD-based watermarking. Several research articles reported some faults in those formulations, which are known collectively as the False-Positive Problem. Other researchers investigated different aspects of this problem from different viewpoints, and proposed solutions to overcome the faults. This paper, first categorizes those researchers' viewpoints into three groups, and shows that while focusing on certain aspects of the two algorithms, they did not succeed in finding the real cause of the False-Positive Problem. Secondly, this paper locates the actual cause of the faults of the two algorithms by directly examining the formulations. Our experiments show that the algorithms fail to correctly embed a distinguishable part of the watermark into the cover image. Further analysis shows that the algorithms embed some arbitrary data in the cover image, and send most of the watermark data as the side information in a semi-blind watermarking setup. Finally, this paper reviews the existing solutions to the False-Positive Problem, and shows that the correct solution to this problem is embedding the principal component of the watermark into the cover image.

Keywords Watermarking · SVD · False-Positive Problem · Ownership · Copyright

1 Introduction

The rate of production of multimedia assets such as images and videos has been growing exponentially. Multimedia plays a critical role in today's advertising and attention economies. The ease of copying, using, sharing, and modifying images, makes reliable copyright protection more important than ever before. Watermarking is one of the common techniques for the copyright protection of digital images. Watermarking a digital image

✉ Narges Safizadeh
n.safizadeh@ut.ac.ir

Sayyed Hamid Reza Ahmadi
hrahmadi@ut.ac.ir

[1] Faculty of New Sciences and Technologies, University of Tehran, Tehran, Iran

consists of embedding additional data in the image that can later be matched to a copyright signature. Additional data embedded in the image is used to determine the original owner of the image. Counterfeit copies of the image could be made by modifying the image in such a way that the original signature cannot be extracted. To achieve this goal, signal processing algorithms are used to make small modifications to the watermarked image such that the embedded watermark is removed or heavily modified. These modifications, which are considered "watermarking attacks", may include contrast adjustment, gamma correction, filtering, histogram equalization, introducing noise, etc. For a watermarking algorithm to be used in copyright protection, it must be designed in a way that attacks do not succeed in removing or modifying the watermark. Such algorithms are called "robust" against the attacks. Several categories of robust watermarking algorithms exist, and a large number of research articles have been published on this subject [1, 3, 4, 8, 9, 15–17, 30]. There are also other articles that have analysed and applied robust watermarking algorithms in specific application fields such as medical images [7, 13, 14, 19, 33, 35].

One of the categories of robust watermarking techniques is based on the mathematical properties of singular value decomposition (SVD), which was proposed in 2002 [20]. Compared to other categories, the main advantage and *"raison d'être"* of SVD-based watermarking is that it provides robustness against geometric attacks (i.e., cropping, rotation, resizing, and flipping) [1]. Like other types of watermarking techniques, the SVD-based algorithms may be categorized as blind, semi-blind, or non-blind. This paper focuses on semi-blind SVD-based watermarking algorithms.

For the purpose of watermarking, a 2-D image I is treated as a 2-D matrix, which is first decomposed by SVD into the product of three separate matrices, as:

$$I = USV^* \tag{1}$$

where S is the singular values matrix of I, and U and V are the principal components of I due to the mathematical relationship between SVD and principal component analysis [20]. After the decomposition in (1), the watermark data is embedded into the S component of the image. Because of the properties of principal components, any geometric attack applied to the image I, will only affect the principal components (i.e., U and V); hence, leaving the S component and the embedded watermark unchanged and robust against the applied attacks [20]. Based on the robustness properties mentioned above, several embedding techniques in the literature have been proposed [1, 3, 4, 8, 9, 20, 21, 24, 27, 31, 32], which provide a watermark robust against the geometric attacks. All of these proposed techniques use semi-blind watermarking, in which, some side-information generated during embedding must be present to make the watermark extraction possible.

In 2005, Zhang and Li [37] revealed a structural problem in the formulae proposed in [20] and named it "the false-positive problem". They showed that the watermark extraction process in [20], gives an attacker the opportunity to extract any desired image from the watermarked image, therefore making the copyright protection ineffective. Despite the results published in [37], still many other papers published in the following years that used the same flawed formulae and suffered from the same problem [1, 3, 8, 10, 21, 24, 27, 31]. Other research papers presented several different analyses of the problem and tried to solve it based on their given analysis using different solutions [2, 12, 22, 25, 28].

The contributions of this paper are:

- Focusing on the embedding and extraction formulae proposed for SVD-based watermarking to find the actual cause of the false-positive problem,

- Examining the structure of the singular values matrix (S),
- Reporting a certain class of simple images with a special type of S matrix, and
- Discussing the correctness of the previously presented analyses and solutions to the false-positive problem.

To make the above contributions, we studied the mentioned problems thoroughly and performed different experiments. To help the reader better understand these experiments and the structure of the paper, a flowchart of our study is shown in Fig. 1, which shows the steps of our study and refers each part to the corresponding section of the paper.

The rest of this paper is organized as follows. Section 2 reviews primary approaches used in SVD-based image watermarking. In Sect. 3, we use an experiment-based approach to investigate the root cause of the problems with the SVD-based algorithm. In Sect. 4, the

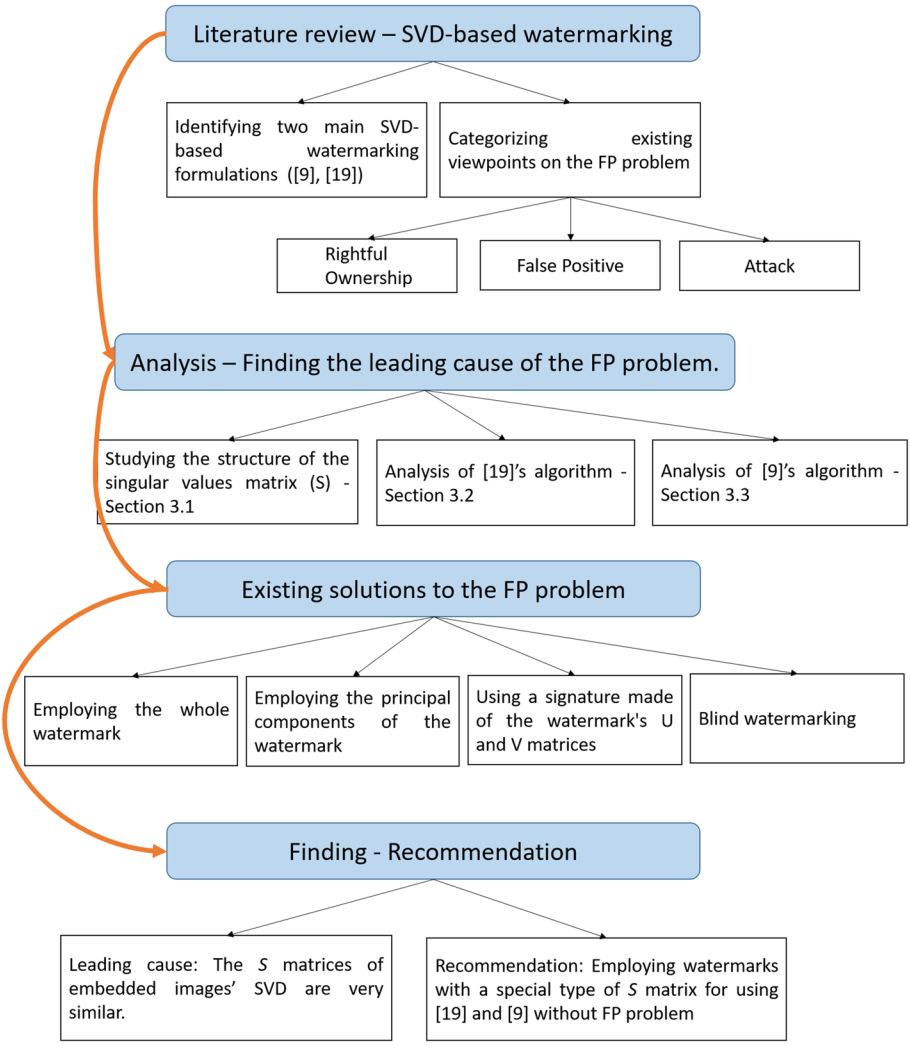


Fig. 1 Flowchart of the present study

existing solutions for solving the false-positive problem are described. In Sect. 5, we show how using principle components is the only reliable solution to the false-positive problem in semi-blind SVD-based watermarking.

2 Related Works

The SVD and some of its applications were first introduced in 1965 by Golub and Kahan (Golub & Kahan, 1965). In 2002, SVD based watermarking was first used to achieve robustness against geometric attacks [20]. Shortly after, other researchers also proposed similar techniques [9, 32]. The false-positive (FP) problem was identified as an underlying issue of the algorithm proposed in [9] in the work of Zhang and Li [37]. The same problem was also reported by Rykaczewski [29] in 2007, albeit from a different viewpoint. In later years, many papers have been published on the SVD-based watermarking, which may be divided into two groups depending on their approach towards the FP problem. The first group of papers, surprisingly, have neglected the FP problem without any effort to find a solution for it [1, 3, 8, 10, 21, 24, 27, 31]. The second group of pairs cited here as [2, 12, 22, 25, 28] provided various solutions to solve the FP problem. Another point which is worth mentioning here is that some other papers only reported, repeatedly, the FP problem of the papers in the first group, e.g., [27, 28]. Surprisingly, some of these repeated reports are only 1-page or 2-page papers [29, 31].

2.1 The Original SVD-based Watermarking Formulations

Two similar formulations were proposed in [20] and [9] using SVD-based watermarking for semi-blind copyright protection. Table 1 shows the embedding and extraction phases of the algorithms in [20] and [9]. Almost all other papers published on semi-blind SVD-based watermarking use one of these two formulations. Note that in this formulation the right arrows show SVD decompositions.

In [20], the embedding of watermark is done by first calculating the SVD of the cover image A. Then, the watermark image is added to the singular values matrix of A (i.e., S) using a scaling factor α, and the SVD of the resulting matrix is calculated. The singular values obtained from the second SVD is used along with the U and V matrices of the

Table 1 SVD-based watermarking proposed in [20] and [9] (Right-arrow shows SVD decomposition)

Reference Article	Embedding phase	Extraction phase
[20]	$A \rightarrow USV^H$	$A^*_W \rightarrow U^* S^*_W V^{*H}$
	$S + \alpha W \rightarrow U_W S_W V^H_W$	$D^* = U_W S^*_W V^H_W$
	$A_W = US_W V^H$	$W^* = \frac{1}{\alpha}(D^* - S)$
[9]	$A \rightarrow USV^T$	$A^*_W \rightarrow U^* \Sigma^* V^{*\mathrm{T}}$
	$W \rightarrow U_W S_W V^T_W$	$S^*_W = (\Sigma^* - S)/\alpha$
	$\Sigma = S + \alpha S_W$	$W^* = U_W S^*_W V^T_W$
	$A_W = U\Sigma V^T$	

cover image, to compose the watermarked image. The 4-tuple (U_W, V_W, S, α) must be provided as side-information to extract the watermark in the extraction phase. The U_W and V_W matrices are the principal components obtained from the second SVD. The α scaling factor determines the strength of the embedded watermark and affects its robustness and usually has a small value (e.g., $\alpha = 0.2$ in [20]). The extraction phase is exactly the reverse of the embedding phase.

The formulae proposed in [9] are very similar to [20]. The difference is that in [9], the watermark image is first decomposed by SVD, and then the obtained singular values are added to the singular values of the cover image after scaling by α.

In addition to the difference mentioned above, [20] and [9] also differ in their watermarking domain. In [20], the watermark is embedded directly in the pixel values of the cover image, which is called spatial-domain watermarking. However, in [9], frequency-domain watermarking is used. That is, they first transformed the cover image into the frequency domain using the DWT and then used the DWT coefficients to embed the watermark. Hence, in the formulae of [9] shown in the last row of Table 1, A represents the DWT coefficients of the cover image, and A_W is the same coefficients after embedding of the watermark.

The benefit of using the SVD in the frequency domain is that the resulting watermark will be robust against frequency attacks in addition to the geometric attacks. For this same reason, combining the SVD with some type of frequency-domain transform has been also reported in many other publications. Even there are many papers that combined more than one transform with the SVD. Table 2 shows some other articles using multiple transforms besides the SVD.

As can be seen in Table 2, all of these watermarking techniques use either the formulae of [20] or the formulae of [9], and therefore suffer from the same false-positive problem.

2.2 Categorizing Existing Viewpoints on the False-Positive Problem

There are several works in the literature that reported, analysed, or tried to solve the problem of the algorithms proposed in [20] and [9]. While these papers collectively refer to the problem as the False-Positive problem, they looked at this problem from different points of view:

Table 2 SVD-based watermarking combined with multiple frequency-domain transforms

Ref. #	Transform combination	SVD algorithm used
[3]	FHT-DWT-SVD	[9]
[8]	RT-DWT-SVD	[9]
[10]	RDWT-SVD	[9]
[21]	DWT-SVD	[9]
[24]	DCT-DWT-SVD	[20]
[27]	RIDWT-SVD	[20]
[31]	DCT-DWT-SVD	[20]

1. "Rightful Ownership" point of view: Focusing only on the application of copyright protection and checking whether the algorithms fulfil their intended goal of protecting image ownership.
2. "False Positive" point of view: Focusing on the performance measure of false positive error and checking whether the algorithms fulfil this performance goal. Every watermarking algorithm should have zero or small false positive error.
3. "Attack" point of view: Focusing on the robustness of the algorithms against specific types of attacks and testing the algorithms' success in the presence of those attacks.

In the following, we categorize and describe the findings of these papers according to their specific viewpoints:

"Rightful Ownership" viewpoint: The first paper to address the faults of these algorithms, i.e. [37], tested the extraction phase of the algorithm of [20] and found that this algorithm cannot be used for copyright protection. The reason was that, because of the inherit design of the extraction formulae, any other person could extract their desired watermark from the cover image in the extraction phase, regardless of the actual embedded watermark image. The same fault was reported by [29] for the watermarking algorithm proposed by [9].

Both [37] and [29] only examined the applicability of the *extraction* algorithms for ownership protection, and reported the fault, without trying to analyze the *embedding* formulae or trying to propose a way to solve this problem. In this paper we name this the "rightful ownerships" viewpoint, since the only observation was that the extraction algorithms were not suitable for copyright protection and ownership management issues.

"False Positive" viewpoint: Some other researches like [5] and [34], focused on a security aspect of watermarking algorithms, namely the false positive measure. Ideally a watermarking algorithm should be free of false positive, i.e. the extraction process never validates a fake watermark. In case of semi-blind algorithms like the ones proposed by [20] and [9], the only way to extract the embedded watermark is to provide the side-info generated during the embedding phase, and any other side-info should not be able to extract a valid watermark. In [5] and [34], the researchers showed that the algorithm presented by [20] will always generate a valid watermark in the extraction phase when provided with suitable side-info, regardless of the actual embedded watermark. The algorithm showed a 100 percent false positive probability, hence it was useless. In this paper we name this the "false positive" viewpoint, since the focus was on this aspect of the fault and the fact that a 100 percent probability renders the algorithms completely useless.

"Attack" viewpoint: The algorithms of [20] and [9] were designed for semi-blind watermarking. This means that in a correct situation, the side-info generated during embedding phase must be available to obtain the watermark in the extraction phase. But as explained above (and reported by [5] and [34]), the extraction phase of these algorithms generate valid watermarks for any suitable side-info other than that generated in the embedding phase.

Researchers in [36] and then also in [22, 23, 26], interpreted the problem of these algorithms as being susceptible to attacks. That is, they "supposed" that the algorithms are working normally, but lack robustness against certain attacks. In this case the harmful attack is when an attacker changes the side-info and inputs his/her desired side-info (U_W and V_w matrices) in the extraction phase, and causes the algorithm to produce his/her desired watermark.

In this paper we name this the "attack" viewpoint, since the researchers described the problem based on attacks. These researchers also proposed counter-attack solutions to the problem of algorithms in [20] and [9], which will be analysed with more detail in Sect. 4. They tried to make a reliable connection between the embedding and extraction phases by adding authentication to the side-information, hence preventing image manipulation with the wrong data in the watermark extraction.

2.2.1 Concluding Remarks

Table 3 summarizes the different points of view used by researchers in their analyses of [20] and [9]. A common aspect of these analyses, is that the researchers only focused on the extraction phase of [20] and [9] and checked whether specific criteria are met by these algorithms. Although the researchers have found that these algorithms do not meet some necessary criteria, they have not identified the actual cause of the problem. In the next section, the actual cause of the problems will be found by examining the embedding formulae of [20] and [9] in detail.

3 Analysis of the False-Positive Problem

SVD-based watermarking algorithms proposed in [20] and [9] are designed to be semi-blind. That is, certain information generated in the embedding process is needed to extract the watermark during the extraction phase, as depicted in Fig. 2. For the algorithms in [20] and [9], this side-information is the 4-tuple (U_W, V_W, S, α) that was explained in Sect. 2–1. Careful analysis of the side-information and how it is generated is the key to understand the actual cause of the False-Positive problem.

3.1 The role of the S Matrix in SVD-based watermarking

Refer to the approaches expressed in Table 1, the watermark image in [20]'s algorithm is multiplied by alpha coefficient then added to the S matrix of the cover image. Then, SVD decomposition applied to the result. On the other hand, in [9]'s algorithm, SVD is applied to the watermark image. In both algorithms, after applying SVD decomposition, the diagonal S matrix is the part that embeds in the cover image in the embedding process. In both algorithms, the S matrix is the key part of the embedding process, while other parts are

Table 3 Viewpoints of different articles about the SVD based watermarking faults

Viewpoint	Ref	SVD algorithm discussed in Ref
Rightful ownership	[29]	[9]
	[37]	[20]
False-positive	[5]	[20]
	[34]	[20]
Attack	[22]	[20]
	[23]	[9] & [20]
	[26]	[20]
	[36]	[20]

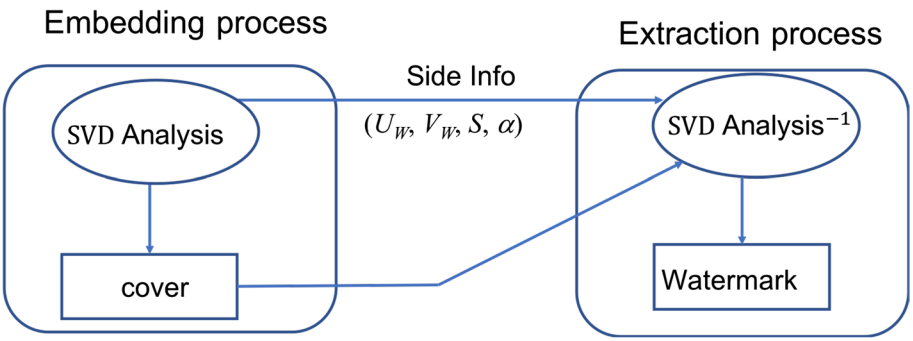


Fig. 2 Semi-blind watermarking technique in [9]; the 4-tuple (U_W, V_W, S, α) is sent independently as side-information

sent as side information. That is why careful analysis of the S matrix, the representor of the watermark in the embedding process, is necessary.

In an SVD procedure, U and V matrices contain more information from an image. They contain the structure and detail of the image [37]. Nevertheless, the S matrix is not an appropriate representor. No matter what the value of the S matrix is, the extracted watermark depends on side information which contains the U and V matrices. Although this point is correctly mentioned in [37], only the rightful ownership problem is investigated, and the main problem of singular value decomposition watermarking is disregarded in this article. This section demonstrates the role of the S matrix in SVD decomposition through a few experiments.

3.1.1 Experiment 1: Pattern of matrix S in the SVD

In experiment 1, we extracted matrix S from the SVD of 50 different 512*512 natural images. The results are plotted in Fig. 3. The 512 diagonal values of 50 images make vertical values in Fig. 3, while numbers 1–512 makes horizontal values. As can be seen, all 50 images have the same S matrix pattern with approximately similar values. The difference

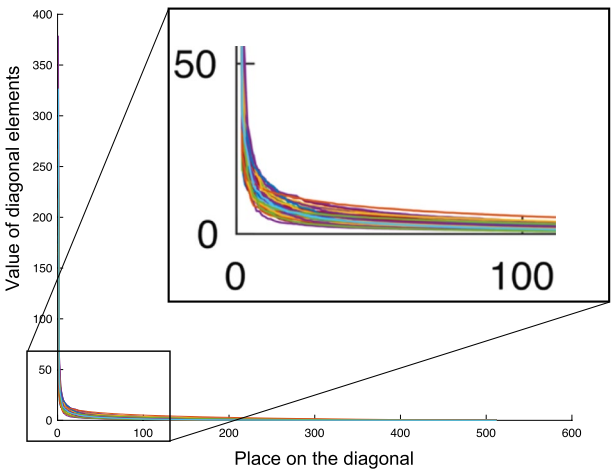


Fig. 3 Diagonal values of matrix S of 50 different 512*512 natural images' SVD

between the maximum and minimum values of the S matrix for 50 images in four places is shown in Fig. 3. For example, in the area of 50, this difference is 6.02. Considering that the maximum value for the brightest picture is 362 and for the darkest picture is 142, the differences are insignificant to be a watermark. This experiment demonstrates that the pattern of the S matrix does not vary significantly in different images. As a result, the S matrix in the SVD cannot provide distinguishable information from an image.

3.1.2 Experiment 2: Exchanging the S matrix of two different images

This experiment is to examine the importance of the S matrix as a unique representor of an image. In this experiment, the SVD of two different images A_1 and A_2 are computed and the S matrices are extracted (Eqs. 2 and 3, below). Then, the matrix S_1 is used to reverse the SVD procedure in image 2 and vice versa (Eqs. 4 and 5, below). The procedure is also shown in Fig. 4.

$$A_1 \rightarrow U_1 S_1 V_1 \tag{2}$$

$$A_2 \rightarrow U_2 S_2 V_2 \tag{3}$$

$$U_1 S_2 V_1 \rightarrow A_1^{'} \tag{4}$$

$$U_2 S_1 V_2 \rightarrow A_2^{'} \tag{5}$$

We exchanged the S matrices on a collection of test images [6]. Figure 5 shows some examples from our standard image dataset. In the part (a) of Fig. 5, for example, the S matrix of the cameraman is used for SVD-reverse of the pirate and vice versa. It can be seen that the new pirate after SVD-reverse is very similar to the original one, with only a little change in its luminance. This experiment demonstrates that low-amplitude variations in the S matrix (which is used in [9] and [20]) will have little impact on the watermarking process.

To further examine this matter, images with the most significant difference in their S matrix's value were selected from experiment 1, to test an extreme situation. Among them, the S matrices' biggest-first-element is 362, which belongs to an airplane image that is also the brightest image in the dataset. Nevertheless, among them, the S matrices'

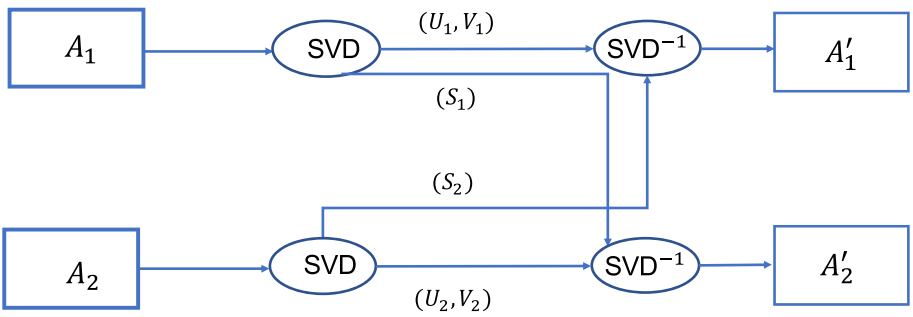


Fig. 4 Replacing S matrices of two images

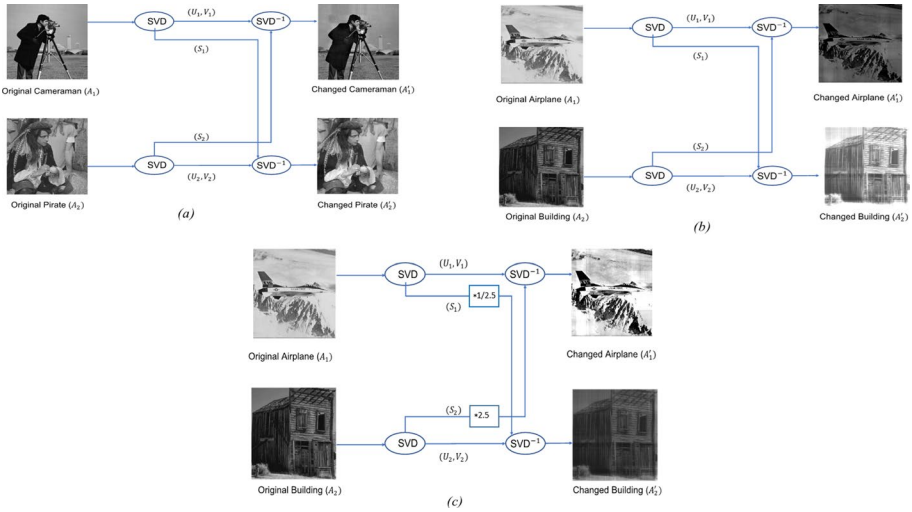


Fig. 5 **a** Exchanging the S matrix of the cameraman and the pirate. **b** Exchanging the S matrix of the airplane and the building. **c** Correcting the coefficient of S matrices of airplane and building

smallest-first-element is 142 and belongs to a building image. Although the S matrices of these two images have the same pattern, the first-elements are much different. Completing our test, we replaced the S matrix of airplane's SVD with the S matrix of building's SVD and vice versa. The result is shown in Fig. 5b. It can be seen that the main structure of the images stayed fixed, and only the luminance of the images was affected by changing their S matrices.

In the next step, we show that the S matrix approximately plays the role of a coefficient. In this step, we normalize the matrices using a coefficient, 362 is 2.5*(142). We multiply the S matrix of airplane by 1/2.5 and then combine it with U and V matrices of the building image. We do this for S matrix of building's SVD but multiply it by 2.5 as well. The result in Fig. 5c shows that by correcting the S matrix by a fixed coefficient, even the luminance of the images are corrected. This experiment proves that the pattern of the S matrix is the same between different images and by finding the proper coefficient, the luminance of the image could also be corrected.

Our experiments show that the S matrix not only is not a distinguishable part of an image's SVD but also its impact on image watermarking is trivial. The identifiable parts of the image are the principal component values (U and V matrices). In this section, we demonstrated that the lack of information in the S matrix is the underlying issue of the SVD-based watermarking proposed in [20] and [9].

3.2 Analysis of the watermarking algorithm proposed in [20]

In this section, we analyze the watermarking algorithm proposed in [20] based on the findings in the previous section. First, we focus on embedding phase. The S matrix of cover image's SVD is added to the watermark image scaled by a small value α, which results in a non-diagonal matrix ($S + \alpha W$). The non-diagonal elements have small values, since α is always a small value. Also, most of the diagonal elements have values close to zero. SVD is applied to the resulting matrix and S_W obtained. S_W does not have a noticeable difference

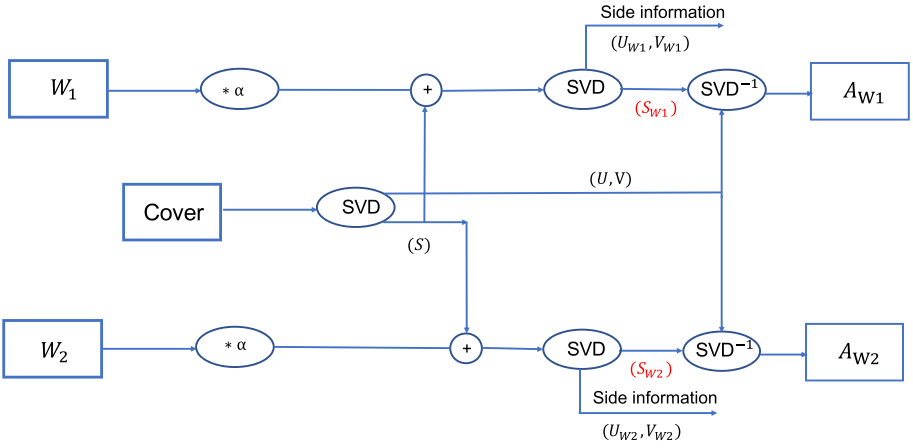


Fig. 6 The experiment of embedding two watermarks in a unique cover image

compared to the S matrix of the cover image's SVD. Figure 6 shows the experiment of embedding two different watermarks in one cover image. The Cameraman is used as the cover image while Lena and Mandrill used as watermark images. Figure 7 shows the different parts of the watermark after multiplying with alpha coefficient 0.05 and adding the S matrix of the cover image. Because the alpha coefficient is small, we multiply every element of Fig. 7a by 20, and show them again in part (b) in order to magnify the small values and emphasize the differences. While the watermark images are entirely different from each other, the S_{W1} and S_{W2} matrices which are embedded in the cover image, are very similar even after magnification. In contrast, U_{W1} and U_{W2} and V_{W1} and V_{W2} which are sent as side information are different.

While S_{W1} and S_{W2} matrices stayed the same, the difference between U_{W1} and U_{W2} and V_{W1} and V_{W2} became significant. These experiments demonstrate that the embedded components of the cover image are independent of the watermark.

The extraction of the watermark in the algorithm of [20], requires five elements. The 4-tuple (U_W, V_W, S, α) is sent as side information, and only S_W is needed from the watermarked cover image to complete the extraction phase. As shown in Sect. 3.1, the S matrix of an image' SVD is not noticeable and could be replaced with another matrix without destructive damage.

To experiment how important is the side information of the algorithm of [20], we want to reconstruct the watermark with only side information and S matrix of a desired image' SVD. In this experiment cameraman is the cover image, and Lena is the watermark. We

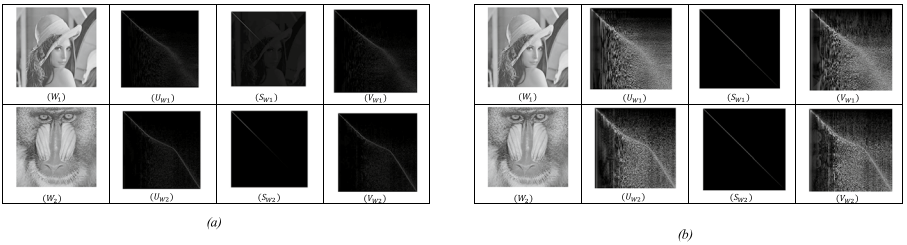


Fig. 7 **a** The different elements of the SVDs of S + αW_1 and +αW_2. **b** Same data of part a, scaled by 20 to emphasize the differences

also use the House image as an arbitrary image to get the S matrix pattern. Note that the result is the same with any arbitrary image. The steps of reconstructing the watermark are shown in Fig. 8 and also listed below.

1- An arbitrary image multiplied to alpha and added to S matrix of the cover image from the side information named X.

$$\mathrm{X} = (\mathrm{ArbitraryImage} * \alpha) + \mathrm{S}$$

2- The SVD of X is computed.

$$\mathrm{X} \rightarrow U_A S_A V_A^*$$

3- U_W and V_W from the side information multiplied with S_A.

$$\mathrm{ArbitraryCoverWatermark} \leftarrow U_W S_A V_W^H$$

4- S matrix of cover image subtracted from ArbitraryCoverWatermark divided to alpha. The Reconstructed watermark obtained.

$$\mathrm{ReconstructedWatermark} \leftarrow \frac{1}{\alpha}(\mathrm{ArbitraryCoverWatermark} - S)$$

The reconstructed watermark is shown in Fig. 9. As can be seen, the part of the watermark which is embedded in the cover image is not an appropriate representor of the watermark, and most of the watermark information is transmitted as side information. In other words, the extraction phase is independent of the cover image and embedded watermark and is only dependent on the side information. There is one point that should be mentioned about the result shown in Fig. 9. In Fig. 9, a diagonal line can be seen overlaid on the extracted watermark. This diagonal line is caused by the mismatch between the S matrix of the original watermark (i.e. the original Lena image) and the S matrix of the arbitrary image. As shown in Fig. 3, the S matrices of different images are very similar with only

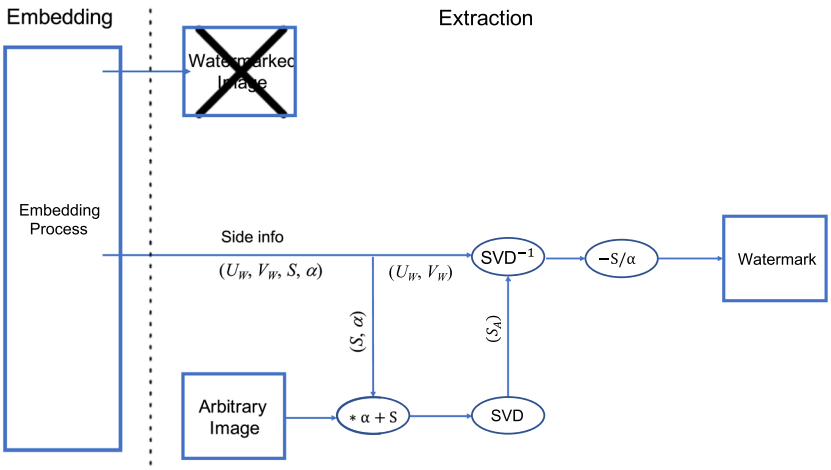


Fig. 8 Reconstruct the watermark without the watermarked image

Fig. 9 Lena watermark recreated by only side information of method [20]

small differences in their values. These small differences, cause the diagonal pixels to vary slightly from their original values.

Through this section, we mentioned the main issues of [20]'s approach to semi-blind watermarking. In conclusion, there is a significant difference between [20]'s algorithm and other semi-blind watermarking algorithms. In such algorithms, the whole watermark, or a considerable part of it, is embedded in the cover image. At the same time, side information is generated and sent to combine with the embedded part in the extraction process. Nevertheless, in [20], a small part of the watermark is used for embedding phase, while, the major part is sent as side information. Moreover, in [20] independent of the cover image and embedding phase, the side information is enough for the extraction process.

3.3 Analysis of The Algorithm of [9]

In this section, the algorithm of [9] is analyzed. As shown in Table 1, in the embedding phase of [9], the watermark image is first decomposed into a singular values matrix S_W and two principal component matrices U_W and V_W. The two principal component matrices, are not included in the embedding and are sent as side-information for the extraction process. The only part of the watermark which is embedded in the cover image, is the S_W matrix, after multiplying by the small scaling factor. Based on the point mentioned in 3.1, S_W matrix has the same pattern with other images. Scaling down by multiplying with an alpha factor makes the values very small. After adding these small values to the S matrix of cover image' SVD, the resulting diagonal matrix has a small difference with the initial matrix. Like the last section, we embed two watermarks in the cover image using algorithm [9] and compare Σ_{W1} and Σ_{W2}. The detail of the algorithm visualized in Fig. 10. We use cameraman as the cover image, Lena as W_1 and Mandrill image as W_2. The difference between Σ_{W1} and Σ_{W2} is 0.01 percentage. Figure 11 shows

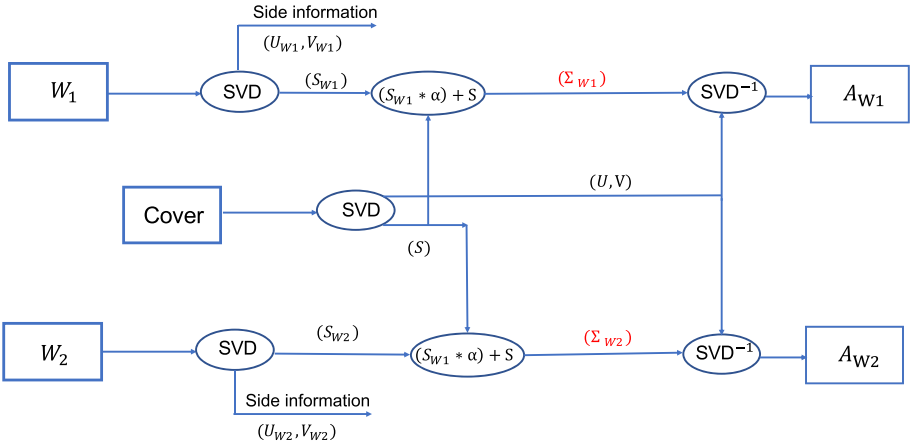
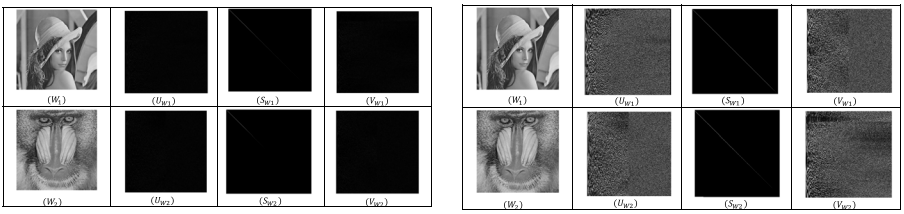


Fig. 10 Embedding two watermarks with the algorithm of [9]



(a) *(b)*

Fig. 11 **a** Side information and embedded parts of two watermarks with the algorithm of [9]. **b** Same data of part a, Scaled by 20 to compensate for the effect of alpha

Σ_{W1} and Σ_{W2} matrices versus U_{W1}, U_{W2} and V_{W1}, V_{W2}. Part b shows the differences better. As can be seen in part b, the side information of two watermarks is different from each other, while the embedded part is similar. This experiment is repeated for different covers and watermarks, and the result is the same. Considering the result of 3.1 experiences, the S matrix of different image' SVD could replace with each other without significant change. Because the difference is insignificant, we could say that the embedded part is independent of the watermark. This is the problem of embedding process of the algorithm [9].

In the extraction phase, there is also the same problem of the algorithm of [20], which was mentioned in 3.2. Based on the formula of extraction of the algorithm [9] in Table 1, U_W and V_W are sent as side information of the watermark. Without concerning the two first line of extraction, instead of recovering the S matrix from the embedded image, we use an arbitrary image S matrix. The watermark extracts simply by using the following steps.

1. An arbitrary image' SVD computed.

$$\text{ArbitraryImage} \rightarrow U_A S_A V_A^*$$

2. Multiplying U_W and V_W from side information with S_A. Reconstructed watermark obtained.

$$\text{Reconstructedwatermark} \leftarrow U_W S_A V_W^H$$

Figure 12 is reconstructed watermark using U_W, V_W and S matrix gained from SVD decomposition of an arbitrary image. Note that for reconstructing the watermark in this algorithm, even S matrix and alpha coefficient from the side information is not necessary. It can be seen that in the algorithm of [9], most of the information contained in the watermark image is not embedded in the cover image. The extraction phase is independent of the cover image and embedding phase, and the required information comes from side information. So, the analysis comes for algorithm [20] is also true for this algorithm.

3.4 The Actual Cause of the FP problem

Section 2.2 went through a review on researchers' viewpoints on [20] and [9] problems. The first and third viewpoints are rightful ownership and attack respectively. Based on first, the result of the extraction phase, independent of the embedding phase, is the image that we used its principal components as side information. The third viewpoint also considers the attack on the side information to change the result in the extraction phase. The reason for both issues is that the significant part of the watermark is sent as side information instead of being embedded in the cover image. Moreover, the second viewpoint, which refers to the FP problem, says although we use a noisy image instead of the watermarked image in the extraction phase, having proper side information helps the exact watermark extraction.

Fig. 12 Lena watermark recreated by only side information in method [9]

In Sects. 3.1, 3.2 and 3.3, we focused on the properties of the S matrix in SVD decompositions and also examined the *embedding* phase of the algorithms proposed by [20] and [9]. The actual reason of the problems of [20] and [9], is that the embedded data (which is the S matrix) has the same pattern in natural images and is nearly identical for all images. This means that in these algorithms, the output of the *embedding* phase does not carry any meaningful data about the watermark.

3.5 Simple Watermark Images with a Special Type of *S* Matrix

All of the articles mentioned in Sect. 2, use natural images as watermark. In Sect. 3, it was shown that the S matrix of these natural images share the same pattern and are nearly identical. In this section, we want to report an observation of some images which have special SVD properties that could be used in watermarking based on [20] and [9] formulae. Since these images have simple structures, in this section they are referred to as "Logo". In the rest of this section, it will be shown how different is the S matrix of these logos from the S matrix of other natural watermark images; and how this difference helps in using the primary formulae of SVD based watermarking without problem. Figure 13 shows a sample of a logo.

Figure 14 compares the diagonal values of the S matrix of the logo (shown in Fig. 13) to the corresponding values of the Cameraman image. It can be seen that this simple logo has only a few non-zero values on the diagonal of its S matrix. Compared to what was shown in Fig. 3, it is evident that the simple structure of this logo translates to this special form of the S matrix.

In Fig. 15 the result of using the simple logo as the watermark is shown. The S matrix of the logo is different from the corresponding matrix of the Panda image. Therefore, when the logo is used for embedding and the extraction is performed with the side information of the Panda image, the extracted watermark will be corrupted.

Fig. 13 Sample logo

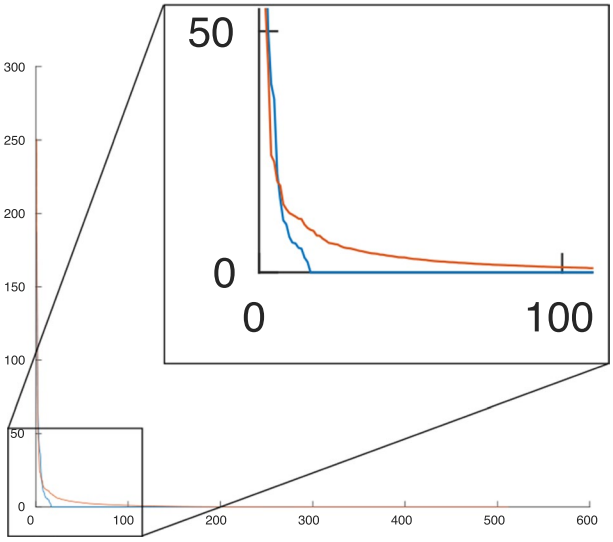


Fig. 14 Diagonal values of matrix S of logo image compare to Cameraman

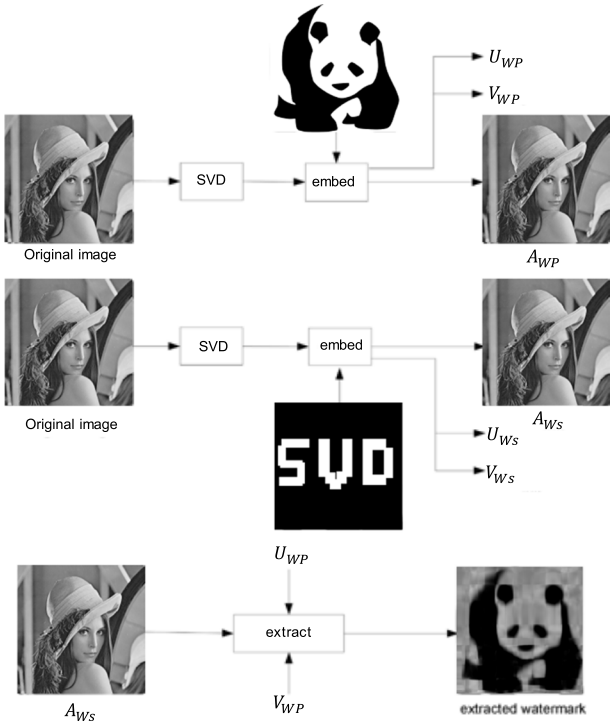


Fig. 15 Exchanging the S matrix of the SVD logo and the panda

The effect of using this kind of simple logos in SVD-based watermarking of [20] and [9], further reinforces our reported finding that the actual cause of the so-called False-Positive problem is the bad choice of embedded data.

4 Existing solutions to the false-positive problem

In this section, we briefly review the literature regarding the different solutions to the so-called false-positive problem.

In 2008, Mohammad, Alhaj and Shaltaf [25] presented the first solution to the problem of the algorithm [20]. They embedded the whole watermark in the cover image and omitted the second SVD decomposition (see Table 1). It means that, instead of embedding the S_W matrix in line 2 of the embedding phase in Table 1, they embedded $S+\alpha W$ into the cover image.

In 2010, Ling, Phan and Heng [18] casted doubt on the solution of [25] and showed its failure under a type of attacks on the watermarked image. The attack occurs when after the embedding phase, one embeds a second watermark in the cover image. Thus, the original watermark cannot be detected in the extraction phase. The problem comes from embedding the whole watermark in the cover image.

The other solution for [20] is using the principal components of the watermark for embedding in the cover image. The first article addressed this was Jain, Arora and Panigrahi [12]. They manipulated U_W and S_W of the watermark by a coefficient, alpha, and added that to the S matrix of the cover image's SVD. While this solution is the most popular, there are four approaches for choosing the embedding part of the watermark. Guo and Prasetyo [11] embed the watermark's principal components ($U_w S_w$, $S_w V_W^T$) and its eigenvector (U,V^T) to the cover image and compare their robustness. Their approach solves the false-positive problem, which mentioned in Sects. 3.2 and 3.3. It can be stated that what is embedded in [20] and [9]'s algorithms is not a good representative of the watermark and cause the false-positive problem, which can be solved by using the principal component in the embedding phase.

Mishra, Agarwal, Sharma and Bedi [24] and also Fazli, Moeini [8] w.r.t. [20] and [9] claim that using a signature made of the watermark's U and V matrices is a novel solution for the false-positive problem. They add the signature to the watermarked image in the embedding phase. Then this signature is used to prove the owner's identification before the extraction phase. Based on our categorization in 2.2, their solution comes from an attack viewpoint. They state that using signature in the authentication process is the key to solving the FP problem. However, their solution is just a well-known cryptography-based solution. In other words, their algorithm hides the issue of the FP problem instead of solving it. Liu and Tan [20] try to address watermarking as a solution for keeping rightful ownership, while the [24] and [8]'s algorithm checks the ownership by authentication, not the watermark itself. To better show our point regarding the solutions of [24] and [8], Fig. 16 gives

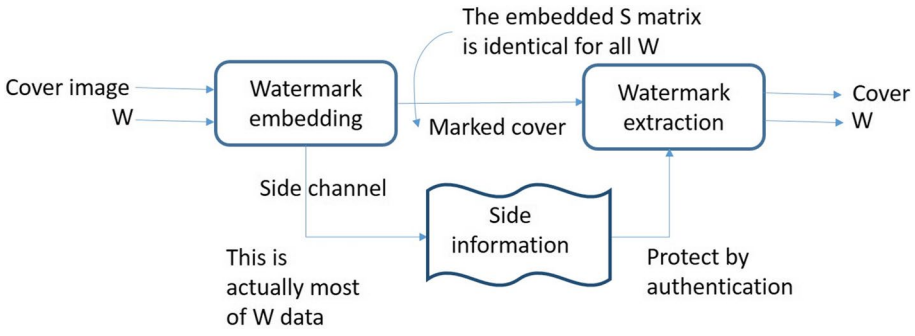


Fig. 16 Solutions presented in [24] and [8]

a sketch of these solutions. It can be seen that the main information of the watermark (i.e., U and V matrices) are sent via the side channel and protected by authentication, while the embedding and extraction phases are not changed and still suffer from the false-positive problem.

Some of the researchers recommend a blind watermarking algorithm based on SVD as the solution for the false-positive problem like [25]. Such blind algorithms have different formulation and properties, e.g. the capacity of blind algorithms is in order of bits. In contrast, the watermark size in Liu and Tan [20] may be equal to the cover image. Golub and Kahan [10] published the blind approach even before 2002, which is the publishing time of [20]. As a result, the blind formulation may not be considered as a solution to the false-positive problem.

In this section, we did a brief review of different approaches to solve the false-positive problem. Since no sign of the watermark remains to authenticate it in the extraction phase, embedding the whole watermark in the cover image is not an acceptable solution. Moreover, while side information authentication brings security, it does not solve the FP problem of [20]. The only suitable approach is storing the watermark's principal component in the cover image. There is also a need for several analyses to check whether this approach meets the primary goals of the SVD based watermarking, which is not the subject of this paper and will be presented separately in the future.

5 Conclusions

In this article, we focused on the false-positive problem in semi-blind SVD-based watermarking which was reported by other researchers. The first contribution of this paper was analyzing the papers that reported this problem (Sect. 2.2), based on their respective points of view and also analyzing their solutions for this problem (Sect. 4).

The second and main contribution of this paper was a detailed analysis of the algorithms proposed in [20] and [9] (Sect. 3). In this analysis we showed that the actual cause of the problems in [20] and [9], was the bad choice of the data which is embedded in the cover image (i.e., the S matrix). We showed that the S matrix of an image's SVD is very similar for almost all of the images, and different images cannot be distinguished based on the S matrix. This means that for any given cover image, the output of the embedding phase in [20] and [9] is nearly the same for all watermarks, or in other words nothing meaningful is embedded at all. We also showed that, although the embedded cover image does not carry a meaningful part of the watermark, the extraction phase is still able to find the watermark. This was shown to be the result of the fact that the algorithms of [20] and [9] are semi-blind and the extraction phase uses the side information to build the watermark. We also showed that the extraction phase is independent of the cover image and the watermark image. Also in Sect. 3.5, we observed that for a certain class of simple watermark images, the same algorithms will work correctly due to the special form of the S matrix of these simple watermarks. This observation is also a further proof that the actual cause of the so-called False-Positive problem is the similarity of the S matrix for different watermark images.

In Sect. 4, we reviewed the existing solutions proposed for the false-positive problem. It seems that the only practical solution which really changes the formulae of [20] and [9] to eliminate the actual cause of the false-positive problem, is the use of the principal components of the watermark in the embedding phase (proposed e.g. by [12] and [11]).

While this solution succeeds in eliminating the false-positive problem, the robustness of this solution against different attacks was not analyzed in the literature. Our future work will address the robustness of this solution against watermarking attacks, especially the geometric attacks which were supposed to be the competitive advantage of the SVD-based robust watermarking. Another direction for future work, will be a study of many papers which combined several transforms with SVD-based watermarking, which we briefly mentioned in Table 2 and in Sect. 2.1.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Abdallah E, Hamza A, Bhattacharya P (2007) Improved image watermarking scheme using fast Hadamard and discrete wavelet transforms. J Electron Imaging 16(3):033020
2. Ali M, Ahn C (2014) An optimized watermarking technique based on self adaptive DE in DWT–SVD transform domain. Signal Process 94:545–556
3. Ali M, Ahn C, Pant M, Siarry P (2015) An image watermarking scheme in wavelet domain with optimized compensation of singular value decomposition via artificial bee colony. Inf Sci 301:44–60
4. Ayesha S, Manikandan V, Masilamani V (2015) A combined SVD-DWT watermarking scheme with multi-level compression using sampling and quantization on DCT followed by PCA. In: Indian Institute of Information Technology Design and Manufacturing. Springer International Publishing, pp 141–149. https://doi.org/10.1007/978-3-319-11933-5_16
5. Changzhen X, Fenhong G, Zhengxi L (2009) Weakness analysis of singular value based watermarking. In: International Conference on Mechatronics and Automation. IEEE, pp 2596–2601. https://doi.org/10.1109/ICMA.2009.5246754
6. Dataset of standard 512x512 grayscale test images (n.d.) Retrieved from https://decsai.ugr.es/cvg/CG/base.htm. Accessed 20 Sep 2019
7. Fang Y, Liu J, Li J, Yi D, Cui W, Xiao X, Han B, Bhatti UA (2021) A novel robust watermarking algorithm for encrypted medical image based on Bandelet-DCT. In: Innovation in medicine and healthcare.. Springer, Singapore, pp 61–73. https://doi.org/10.1007/978-981-16-3013-2_6
8. Fazli S, Moeini M (2016) A robust image watermarking method based on DWT, DCT, and SVD using a new technique for correction of main geometric attacksA robust image watermarking method based on DWT, DCT, and SVD using a new technique for correction of main geometric attacks. Optik-Int J Light Electron Opt 127(2):964–972
9. Ganic E, Eskicioglu A (2004) Robust DWT-SVD domain image watermarking: embedding data in all frequencies. In: Proceedings of the 2004 Workshop on Multimedia and Security. https://doi.org/10.1145/1022431.1022461
10. Golub G, Kahan W (1965) Calculating the Singular Values and Pseudo Inverse of a Matrix. J Soc Ind Appl Math 2(2):205–224
11. Guo J, Prasetyo H (2014) False-positive-free SVD-based image watermarking. J Vis Commun Image Represent 25(5):1149–1163
12. Jain C, Arora S, Panigrahi P (2008) A reliable SVD based watermarking scheme. arXiv preprint arXiv. https://arxiv.org/abs/0808.0309. Accessed 26 Jun 2019
13. Li T, Li J, Liu J, Huang M, Chen Y-W, Bhatti U (2022) Robust watermarking algorithm for medical images based on log-polar transform. EURASIP J Wire Commun Network 24(2022):1–11. https://doi.org/10.1186/s13638-022-02106-6
14. Li Y, Li J, Shao C, Bhatti U, Ma J (2022) Robust Multi-watermarking Algorithm for Medical Images Using Patchwork-DCT. International Conference on Artificial Intelligence and Security. Springer, Cham, pp 386–399
15. Liao X, Shu C (2015) Reversible data hiding in encrypted images based on absolute mean difference of multiple neighboring pixels. J Vis Commun Image Represent 28:21–27

16. Liao X, Wen Q, Song T, Zhang J (2013) Quantum steganography with high efficiency with noisy depolarizing channels. IEICE Trans Fundam Electron Commun Comput Sci 96(10):2039–2044
17. Liao X, Wen Q, Zhang J (2013) Improving the adaptive steganographic methods based on modulus function. IEICE Trans Fundam Electron Commun Comput Sci 96(12):2731–2734
18. Ling H, Phan R, Heng S (2010) Analysis on the improved SVD-based watermarking scheme. In: Advances in Computer Science and Information Technology. Springer Berlin Heidelberg, pp 143–149. https://doi.org/10.1007/978-3-642-13577-4_12
19. Liu W, Li J, Shao C, Ma J, Huang M, Bhatti UA (2022) Robust Zero Watermarking Algorithm for Medical Images Using Local Binary Pattern and Discrete Cosine Transform In International Conference on Artificial Intelligence and Security (pp. 350-362). Springer, Cham
20. Liu R, Tan T (2002) An SVD-based watermarking scheme for protecting rightful ownership. IEEE Trans Multimed 1(4):121–128
21. Makbol N, Khoo B (2013) Robust blind image watermarking scheme based on Redundant Discrete Wavelet Transform and Singular Value Decomposition. Int J Electron Commun 67(2):102–112
22. Makbol N, Khoo B (2014) A new robust and secure digital image watermarking scheme based on the integer wavelet transform and singular value decomposition. Digital Signal Processing 33:134–147
23. Makbol N, Khoo B (2018) Security analyses of false positive problem for the SVD-based hybrid digital image watermarking techniques in the wavelet transform domain. Multimed Tools Appl 77(20):26845–26879
24. Mishra A, Agarwal C, Sharma A, Bedi P (2014) Optimized gray-scale image watermarking using DWT–SVD and firefly algorithm. Expert Syst Appl 41(17):7858–7867
25. Mohammad A, Alhaj A, Shaltaf S (2008) An improved SVD-based watermarking scheme for protecting rightful ownership. Signal Process 8(9):2158–2180
26. Naderahmadian Y, Beheshti S (2015) A realistic attack on SVD based watermarking scheme. In: Canadian Conference on Electrical and Computer Engineering. IEEE, pp 1238–1242. https://doi.org/10.1109/CCECE.2015.7129455
27. Rastegar S, Namazi F, Yaghmaie K, Aliabadian A (2011) Hybrid watermarking algorithm based on singular value decomposition and radon transform. AEU-Int J Electron Commun 65(7):658–663
28. Run R, Horng S, Lai J, Kao T, Chen R (2012) An improved SVD based watermarking technique for copyright protection. Expert Syst Appl 39(1):673–689
29. Rykaczewski R (2007) Comments on 'an SVD-based watermarking scheme for protecting rightful ownership". IEEE Trans Multimed 9(2):421–423
30. Salim M, Abboud A, Yildirim R (2022) A Visual Cryptography-Based Watermarking Approach for the Detection and Localization of Image Forgery. Electron 11(1):136
31. Singh A, Dave M, Mohan A (2014) Hybrid technique for robust and imperceptible image watermarking in DWT–DCT–SVD domain. National Acad Sci Lett 37(4):351–358
32. Sverdlov A, Dexter S, Eskicioglu A (2005) Robust DCT-SVD domain image watermarking for copyright protection: embedding data in all frequencies. In: 2005 13th European Signal Processing Conference. IEEE, pp 1–4
33. Xiao X, Li J, Yi D, Fang Y, Cui W, Bhatti UA, Han B (2021) Robust Zero Watermarking Algorithm for Encrypted Medical Images Based on DWT-Gabor In Innovation in Medicine and Healthcare (pp. 75-86). Springer, Singapore
34. Xiong C, Ward R, Xu J (2008) On the security of singular value based watermarking. In: 2008 15th IEEE International Conference on Image Processing. IEEE, pp 437–440. https://doi.org/10.1109/ICIP.2008.4711785
35. Yi D, Li J, Fang Y, Cui W, Xiao X, Bhatti UA, Han B (2021) A Robust Zero-Watermarkinging Algorithm Based on PHTs-DCT for Medical Images in the Encrypted Domain In Innovation in Medicine and Healthcare (pp. 101-113). Springer, Singapore
36. Yongdong W (2005) On the security of an SVD-based ownership watermarking. IEEE Trans Multimed 7(4):624–627
37. Zhang X, Li K (2005) Comments on "An SVD-based watermarking scheme for protecting rightful ownership". IEEE Trans Multimed 7(2):593–594