Abstract-- Two vulnerable attacks on the wavelet transform (WT) and Singular Value Decomposition (SVD) based image watermarking scheme are presented in this paper. The WT-SVD-based watermarking is robust against various common image manipulations and geometrical attacks; however, it cannot resist against two security attacks, i.e. an attacker attack which successfully claims the real owner's watermarked image, and owner attack which correctly extracts watermark from any arbitrary image. As proved in this study, the SVD watermarking scheme cannot provide trustworthy evidence in rightful ownership protection. In addition, the robustness of the SVD watermarking scheme in is a result of improper algorithm design.

Index Terms: false positive problem, image watermarking, Wavelet Transform, Singular Value Decomposition (SVD), vulnerable attack.

I. INTRODUCTION

The good watermarking algorithm design should meet three criteria, i.e. the rightful ownership protection, robustness to image manipulations, and watermark imperceptibility. Numerous researches have been elaborated in order to improve the SVD watermarking performances [1-8]. The watermark information is embedded into the singular value matrix of the host image. The host image is firstly transformed using wavelet transform or its family [1, 3, 5] before performing the SVD operation. The discrete wavelet transform and its family are chosen to achieve the imperceptibility aspect in the watermarked image. The scaling factor plays an important role to control the robustness and imperceptibility of watermark.

The Redundant Discrete Wavelet Transform (RDWT)-SVD watermarking scheme [1] directly embeds a grayscale watermark image of the same size with 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 will later be used in the watermark extraction stage. The RDWT-SVD approach takes the advantage of the RDWT redundancy to achieve high embedding capacity and preserves the watermark imperceptibility by exploiting the SVD properties. As reported in [1], the RDWT-SVD watermarking scheme is not only robust against several image processing attacks and geometric distortions, but also at the same time it yields the high PSNR value for the watermarked image. It means that the RDWT-SVD can successfully render the watermark image into the host image and achieve the robustness aspect. However, the RDWT-SVD is not vigorous against two vulnerable attacks presented in this paper.

II. ATTACKS ON WAVELET TRANSFORM-SVD IMAGE WATERMARKING

Two vulnerable attacks for the RDWT-SVD watermarking scheme is delivered in this section. The similar attacks can be extended into another wavelet transform. The vulnerable attacks are related to rightful ownership protection and can be defined as follow:

A. Attack I

There are two parties involved in this attack, i.e. real owner and attacker. The real owner embeds the watermark image \( W \) into the host image \( A \) yielding the watermarked image \( A_w \) and side information \( (U_w, V_w) \). Then, the real owner publishes the watermarked image \( A_w \) and keeps the side information to extract the watermark image as a tool to prove 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 \( W_t \) into the owner watermarked image \( A_w \). At the end of watermark embedding stage, an attacker obtains the watermarked image \( A_{w_f} \) and keeps the counterfeit side information \( (U_{w_f}, V_{w_f}) \). Using this side information, an attacker attempts to extract the counterfeit watermark \( W_t \) from the real owner watermarked image \( A_w \). Using this attack, an attacker can easily claim and prove the real owner watermarked image.

B. Attack II

In this attack, an image owner embeds the watermark image \( W \) into the host image \( A \) to obtain the watermarked image \( A_w \). The side information \( (U_w, V_w) \) is also produced at the end of watermark embedding process. The owner makes effort to extract the watermark image using the side information \( (U_w, V_w) \) from any arbitrary image \( B \).

The extracted watermark image \( W^{0*} \) has high correlation and visual similarity with the watermark image \( W \). However, the watermark \( W^{0*} \) 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.

III. EXPERIMENTAL RESULTS

In this section, several experiments were conducted to examine the robustness of the RDWT-SVD watermarking scheme [1] against two vulnerable attacks. The host and watermark image are the visual grayscale image of size \( 512 \times 512 \).
A. Attack I

Figure 1 (c) shows the Lena watermarked image (PSNR = 53.99) when the owner embeds cameraman image (figure 1 (a)) into Lena host image. Figure 1 (d) shows the attacker’s Lena watermarked image with PSNR=49.43, when the counterfeit watermark (figure 1 (b)) is embedded into owner watermarked image (figure 1 (c)).

Figure 1 (e) presents the extracted watermark when attacker extracts the watermark from attacker’s Lena watermarked image (figure 1(d)) using the attacker’s side information. Figure 1 (f) gives extracted watermark from owner’s Lena watermarked image (figure 1 (c)). The RDWT-SVD scheme is unsuccessful to protect the rightful ownership.

B. Attack II

The RDWT-SVD scheme embeds the cameraman image into Lena host image to obtain Lena watermarked image and side information. Figure 2 (d-f) shows the extracted watermark from the baboon, peppers, uniform randomly image (figure 2 (a-c)) using cameraman side information. We can visually recognize the extracted watermark as cameraman image (figure 2 (d-f)). The RDWT-SVD watermarking has major fundamental flaw causing the false positive problem.

C. Comparison with the other methods

We conduct some literature review to show that two vulnerable attacks presented in this paper also occur in the recently published SVD-based image watermarking scheme [1-8]. Table I shows comparison on several SVD-based image watermarking scheme. The RDWT-SVD based watermarking has the highest watermark embedding capacity compared with the other scheme. However, the two vulnerable attacks will definitely occur in [1-8]. Thus, the SVD-based image watermarking fails to satisfy the requirement as a tool of ownership and copyright protection.

CONCLUSIONS

Two vulnerable attacks have been presented in this study for wavelet transform and singular value decomposition watermarking scheme. The former RDWT-SVD scheme can embed the watermark information with high capacity while little degradation on the watermarked image quality. It meets the imperceptibility and robustness aspects, and thus meets the good watermarking design requirements. However, as proved in this study, in fact there is a fundamental flaw in the RDWT-SVD scheme which leads to the serious false positive problem.

REFERENCES


