Indian Journal of Science and Technology, *Vol 10(7), DOI: 10.17485/ijst/2017/v10i7/111267, February 2017*
*ISSN (Print) : 0974-6846*
*ISSN (Online) : 0974-5645*

# False Watermark Extraction and Rewatermarking Issues with Image Watermarking Techniques

**Neha Singh*, Sandeep Joshi and Shilpi Birla**

Manipal University, Jaipur - 303007, Rajasthan, India; nneha.singh01@gmail.com, sandeep.joshi@jaipur.manipal.edu, shilpi.birla@jaipur.manipal.edu

## Abstract

With the ease of availability and accessibility of the internet and image processing software, it has become difficult to distinguish between the original and manipulated images. To restore this lost trust in digital images, digital image watermarking is widely used. Some owner's information is embedded within the digital image which is not only required to be robust against the various image processing operations but is required to be extracted when required and verify for the ownership of the information. One of the major issues addressed for image watermarking techniques is false watermark extraction. But overwriting the already watermarked images, that is, rewatermarking needs to be addressed. This paper analyzes three existing solutions to tackle false watermark extraction and rewatermarking.

**Keywords:** Ambiguity in Ownership, Attacks, Digital Image Watermarking, False Watermark Extraction, Rewatermarking

## 1. Introduction

Digital image watermarking embeds some information in the form of some identifying number, string, copyright message or even a gray-scale or colored image into a digital image. The primary purpose of hiding the information is to establish ownership of the image. The image which carries the hidden information, generally termed as the watermark, is known as cover. The embedded watermark may be visible or invisible, but is expected to be retrieved when required to establish the authorship of the cover. If the retrieved watermark is inconsistent with the originally embedded watermark it confirms that the watermarked image has been tampered. Sometimes, watermarking techniques may uncover tampering locations. The major components of an image watermarking system with their respective inputs and outputs are shown in Figure 1. All the inputs shown for each block in the system are not always required. Table 1 summarizes the purpose(s) of each component.

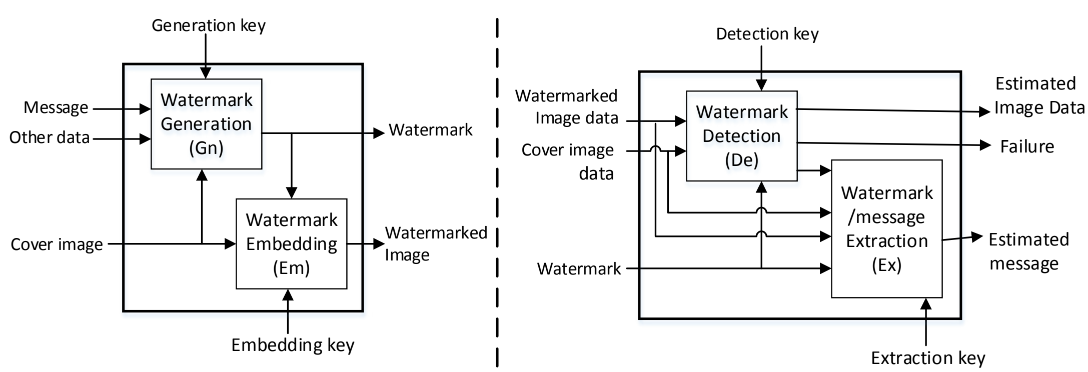



**Figure 1.** General image watermarking system.

*Author for correspondence

**Table 1.** Digital watermarking functions

| Components | Purpose |
|---|---|
| Gn( ) | This component is used to generate watermark. However, if a logo is chosen as watermark, this component is not used in the system. Generation of cover dependent watermark increases security and robustness of the watermarking system. |
| Em( ) | This function identifies the location and procedure for embedding the watermark within the cover to make sure that the embedded information remains unaltered due to any processing on the watermarked image. |
| De( ) | This component is on the retrieval side and is targeted to detect if the image under test carries some hidden information, thereby producing output as true or false. It is useful in steganography to identify the existence of hidden information. |
| Ex( ) | This component extracts the watermark from the watermarked image and may recover the original cover. |

It is necessary for the invisible watermarking technique to ensure that embedding of watermark remains imperceptible and robust against any attack as described by[1].

## 2. Ambiguity Issues

The image watermarking domain has flourished well and many image transforms have been explored individually as well as in combinations. Some commonly used image transforms are Fourier transform, Singular Value Decomposition (SVD), Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) etc. Use of multiple transforms combines the strength of individual for better robustness and imperceptibility as in[2]. Many optimization techniques are also employed with transforms for optimizing location for embedding and weight of watermark as in the work[3,4,6,11]. The choice of watermark for image watermarking techniques has also seen variations from being a binary logo[4], to a text[5], to a randomly generated image, to a grayscale[7] or colored image to cover dependent watermarks.

SVD has been widely used for watermarking because singular values have high stability and a small variation in them does not affect the image quality. This makes the SVD based technique robust against many general image processing operations like image scaling, resizing, rotation, translation, compression and histogram manipulation as shown by [2–9].

Easy availability and usability of software means for image manipulations has resulted in increased variety of possible attacks. The paper[12] presents an overview of general attacks on watermarked images which should be countered while assessing robustness of image watermarking techniques. Many non-blind watermarking techniques which do not use the original cover or watermark during extraction may fail to establish authorship of the content owner thereby defeating the entire purpose of watermarking. Table 2 shows allied possibilities of uncertainty as ambiguity attack possibilities[13] where C is the original cover and WM is the original watermark while C' represents a fake cover and WM' represents fake watermark claimed by attacker. Also, Authentic owner uses function *Em(C, WM)* to generate original watermarked image C_WM and extraction function *Ex(C_WM, WM)*

**Table 2.** Ambiguity attack possibilities

| S.No | Ambiguity | Description | Ambiguity |
|---|---|---|---|
| 1 | False watermark retrieval | Attacker extracts *WM'* using *Ex(C_WM,WM')* claiming WM' to be the originally embedded watermark. | Since both watermarks are successfully extracted from the same image C_WM, attacker is able to establish false claim over the original cover C. |
| 2 | Rewatermarking | Attacker rewatermarks the originally watermarked image by using *Em(C_WM, WM')* to claim *C_WM* to be the original cover | Attacker claims that C_WM was the original cover and not C and the owner claims vice-versa. |
| 3 | False Cover Generation | Attacker reconstructs the originally watermarked image fake cover and watermark with function *Em(C',WM')* | Whether C or C' is the original cover for the watermarked image C_WM. |
| 4 | Unauthorized Watermark Extraction | Attacker extracts original watermark *WM* using *Ex(C_WM)* and embeds it into fake cover as *Em(C', WM)* | Attacker uses the copied watermark as proof to establish that the owner has stolen his data C'. |

to retrieve original watermark WM. Some malicious attacker processes the original watermarked image to produce C_WM'.

The probability of retrieving false watermark is increased if the original watermark or the original cover is used for extraction and if watermark is embedded with small weight or strength factor. Larger strength factor for watermark during embedding results in observable changes and smaller strength factor, on the other hand, results in reduced robustness against attacks. If the attacker uses his fake watermark with larger weight during rewatermarking, he might be able to suppress the original watermark embedded with smaller weight. So, the owner may fail to prove his claim over the cover. So, it is required to use sufficiently large strength factor to ensure that the original watermark is not vanquished by rewatermarking.

## 3. Experimental Results

An attempt to overcome the false watermark detection problem of SVD based technique presented in[3], has been made in[6] by embedding (i) only principal components rather than all singular values of the watermark, and (ii) the complete watermark instead of singular values. This paper extends the scope of analysis for the solutions presented in[6] by testing them against re-watermarking attack too. The method employed by[3] for embedding and extraction of watermark is referred to as scheme-1 in this paper, and the outline is as shown in Figure 2.

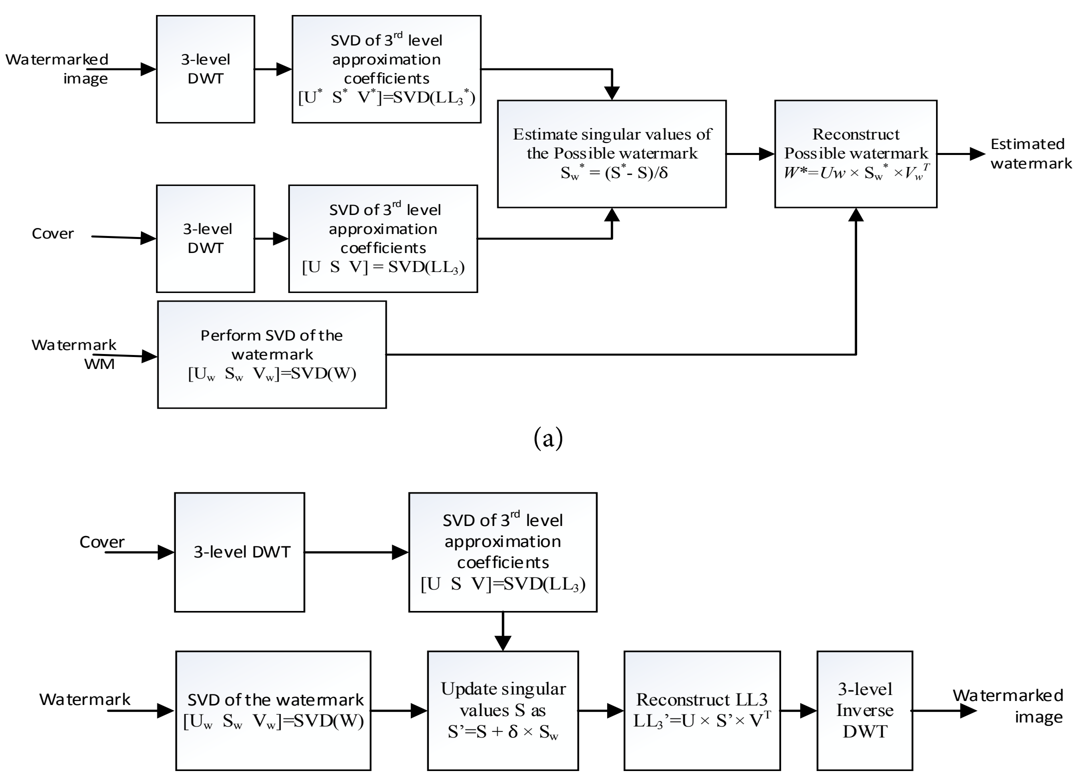
A solution to false watermark extraction problem is proposed in[6], where principal components of the watermark are embedded instead of all the SVs. This scheme is referred to as Scheme-2 in the paper. These principal components were obtained as[6]

$$PCw = Uw \times Sw \quad (1)$$

and embedded into the singular values, S of the cover. Principal components have been used with spread spectrum concept in[7] to tackle false watermark detection issue with SVD based watermarking. The extraction procedure for Scheme-2 is outlined in Figure 3 where principal components PCw* of the approximate watermark are reconstructed.

Another approach to avoid false positive detection, as suggested by[6] is to modify the singular values of the 3rd level approximation coefficients with complete watermark as given by the equation

$$S' = S + \delta \times WM \quad (3)$$




(b)

**Figure 2.** Scheme-1 (a) embedding (b) extraction.

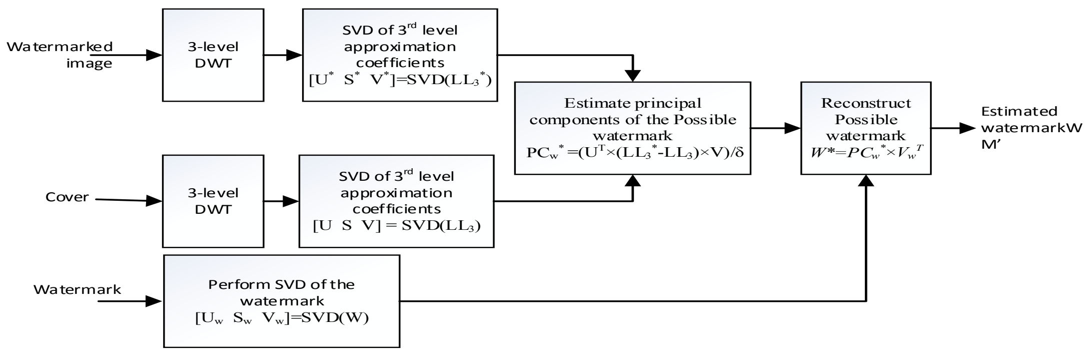



**Figure 3.** Extraction procedure for Scheme-2.

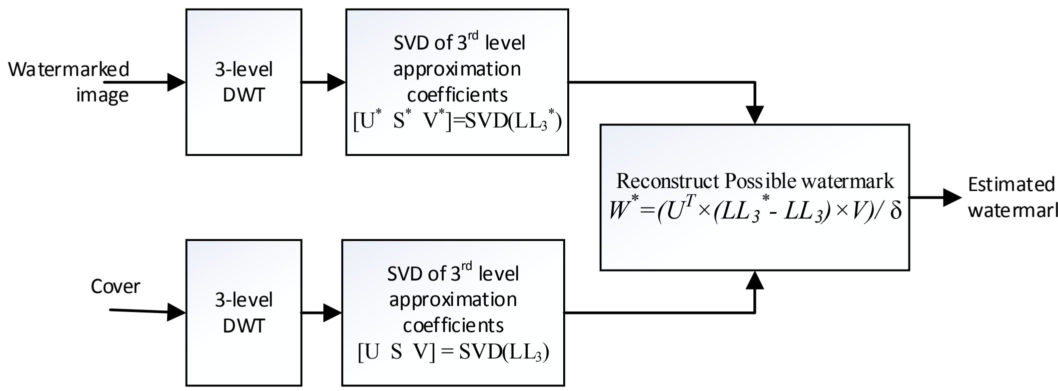



**Figure 4.** Extraction for Scheme-3.

This is referred to as Scheme-3 in this paper. The advantage here is that the need of original watermark during extraction is totally eliminated, thereby reducing the false positive detection possibility. The scheme is referred to as Scheme-3. The extraction procedure for this scheme is shown in Figure 4.

During the experiment, the value of weight or strength factor $\delta$ is chosen so that no perceptible change is observed in the image and its Peak Signal to Noise Ratio (PSNR) is at least 50 dB. The value is kept equal to 0.02 for Scheme-1 and 5 for Scheme-2 and Scheme-3 for a PSNR of around 54 dB. The three schemes are tested against false watermark extraction and rewatermarking attacks. The cover image, original binary watermark and claimed binary watermark used by[6] and re-used in the experiment are shown in Figure 5.

To test these three techniques against false watermark extraction, original watermark is embedded but claimed watermark is used during extraction. Figure 6 shows the retrieved watermarks for each technique against false watermark extraction test.

To test these techniques for rewatermarking attack, the claimed watermark as shown in Figure 5(c) is embedded into the cover image watermarked with the owner's watermark as shown in Figure 5(b). The extraction process for Scheme-1 and Scheme-2 use rewatermarked image, originally watermarked image and the claimed watermark as the inputs. However, Scheme-3 is a blind technique and thus, does not require watermark during extraction.


(a) (b) (c)

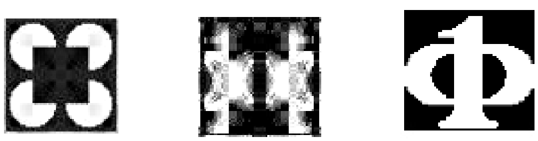
**Figure 5.** (a) Cover image (b) original watermark (c) claimed watermark.

Table 3 shows the result of tests against false watermark extraction and rewatermarking for Scheme-1, Scheme-2 and Scheme-3. The ability of these techniques to resist false watermark extraction being extracted is

**Table 3.** Result for ambiguity tests

| Scheme | False watermark extraction test | Rewatermarking test | |
|---|---|---|---|
| | Correlation with attacker's watermark | Correlation with originalwatermark | |
| | | when watermarked image is used as cover during extraction | when original cover is used during extraction |
| Scheme-1 | 0.9945 | 0.9935 | 0.9558 |
| Scheme-2 | 0.2570 | 0.1520 | 0.7813 |
| Scheme-3 | -0.0456 | -0.0456 | 0.6112 |

measured in terms of correlation of the claimed watermark with the extracted watermark. However, success against rewatermarking attack, of any technique lies in its ability to extract the owner's watermark from the re-watermarked image. Extraction process is done to recover owner's watermark, thus, correlation of the extracted watermark with the owner's watermark is obtained as shown in Table 3.



(a) (b) (c)

**Figure 6.** Retrieved watermarks against false positive detection test for (a) Scheme-1 (b) Scheme-2 (c) Scheme-3.

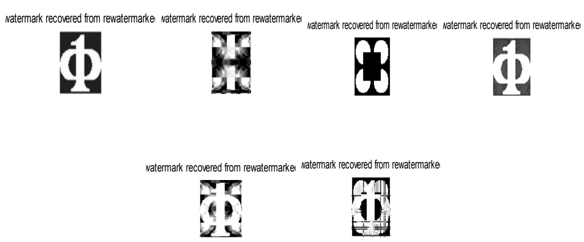


**Figure 7.** Retrieved watermarks against rewatermarking test (i) when watermarked image is used as cover during extraction by (a) Scheme-1 (b) Scheme-2 (c) Scheme-3 ;(ii) when original cover is used during extraction by (d) Scheme-1 (b) Scheme-2 (c) Scheme-3.

## 4. Discussions and Conclusions

Small values of correlation coefficients between the retrieved and embedded watermark for test against false positive detection, in Table 3 show that Scheme-2 and Scheme-3 proposed by Ali and Ahn, provide an improvement over Scheme-1. Embedding the complete watermark clearly worked best due to non-requirement of watermark during extraction. However, the same scheme fails for re-watermarking test. The smallest value of correlation is obtained for both cases- when original cover or watermarked image is used during extraction. Scheme-1 though failed for false positive detection test but displayed highest robustness against rewatermarking attack amongst the three schemes under test. Embedding principal components instead of singular values of the watermark proves to be successful to fail the false watermark detection test as well as rewatermarking attack to some extent. Thus, the Ali & Ahn schemes have scope to be improved for better robustness against ambiguity attacks.

## 5. References


1. H. Nyeem, W. Boles and C. Boyd, EURASIP J. Advances in Signal Processing, (2014) Crossref
2. S. Bekkouch and K.M. Faraoun, J. Inf. Process Syst. 11, 3 (2015)
3. A. Mishra, C. Agarwal, A. Sharma and P. Bedi, Expert Systems with Applications 41, 17 (2015)
4. H. Tsai, Y. Jhuanga and Y. Lai, Applied Soft Computing 12, (2012)
5. A.A. Mohammad, A. Alhaj and S. Shaltaf, Signal Processing 88, (2008)
6. M. Ali and C.W. Ahn, Expert Systems with Applications 42, (2015)
7. J.M. Guo and H. Prasetyo, J. Vis. Commun. Image R. 25, (2014)
8. H. Shi, F. Lv and Y. Cao, J. Software 9,7 (2014)
9. H. Huang, D. Chen, C. Lin, S. Chen and W. Hsu, EURASIP J. Image and Video Processing, (2015) Crossref, Crossref, Crossref, Crossref, Crossref, Crossref PMid:26259244
10. M. Ali, C.W. Ahn, M. Pant and P. Siarry, Discrete Dynamics in Nature and Society (2016) 5.2. Conference Proceedings

11. H. Dong, M. He and M. Qiu. Optimized Gray-scale Image Watermarking Algorithm Based on DWT-DCT-SVD and Chaotic Firefly Algorithm. Proceedings of the 7th International IEEE Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, (2015) September 17-19; Xi'An, China Crossref
12. M. Tanha, S. Torshizi, S. Dawood, M.T. Abdullah and F. Hashim. Overview of Attacks against Digital Watermarking and their Respective Countermeasures. Proceedings of International IEEE Conference on Cyber Warfare and Digital Forensic, (2012) June 26-28; Kuala Lampur, Malaysia Crossref
13. N. Singh, S. Joshi. Ambiguity Attacks on SVD Based Watermarking Technique. Proceedings of International Conference on Smart Trends for Information Technology and Computer Communications. SmartCom 2016. Communications in Computer and Information Science, vol 628. Springer, Singapore (2016) August 6-7; Jaipur, India. Crossref