ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Review on Image Watermarking Attacks.

S Priya*, P Swaminathan , and B Santhi.

School of Computing, SASTRA University, Thanjavur, Tamil Nadu India.

ABSTRACT

With the rapid development of information technology, multimedia data are transmitted through digital media. Always the transmission media is not secure. Hence there is a need to protect the information during transmission. To transmit the image information in a secured way, digital image watermarking is one of the best data hiding technique. In image watermarking technique there is a chance to destroy the information intentionally or uinintentionally by the attacker. This will affect the robustness of the entire image watermarking system. In this paper recent image watermarking works are reviewed for different attacks based on image quality, size, reterived data, etc . In this paper, the different category of image watermarking attacks are analyzed based on attacker knowledge. This paper categorizes the image watermarking attacks into blind and non-blind image watermarking attacks. Then image watermarking techniques are analysed for various attacks using different quality measures. Also, this paper specifies the four different benchmark tools such as Stirmark, Optimark, Checkmark and Certimark to evaluate the image watermarking system. The image watermarking attacks are measured with the help of various performance measures.

Keywords: Image watermarking, Watrmarking attacks, Performace measures, Stirmark, CheckMark, Optimark, Certimark.

Corresponding author

ISSN: 0975-8585

INTRODUCTION

Image watermarking is a technique to protect the images in a secured way. The original image is considered as a cover image [1] and the hidden data or authenticated data is considered as watermark. In image watermarking, at the source side, the watermark information is embedded within a cover image to form a watermarked image and at the destination side, watermark information is extracted and the original image is reconstructed. It is mainly used for copyrights, authentication, broadcast monitoring, etc.

The basic mechanism of image watermarking process [2] are watermark generation, embedding, detection or extraction. In watermark generation, suitable watermark is generated based on application requirements. Then the generated watermark is embedded within a cover image using an embedding algorithm without affecting the image perceptual quality. This embedding process takes watermark and cover image as input and watermarked image as the output. Watermark generation and embedding processes are done on the sender side. Watermark detection is done at the receiver side. In this process, the embedded watermark is extracted using an extraction algorithm and the original image is reconstructed.

Image watermarking is classified as different types like spatial domain and transform domain watermarking based on domain, Visible and invisible watermarking based on watermark visibility [2] [27], robust and fragile watermarking is based on security as shown in Figure 1, and application based image watermarking.

One of the main requirements of an image watermarking technique is robustness[1-4]. The watermarked image is transmitted through multimedia channel, which leads to data loss by the attack. The attack is a process which impairs the watermarked image by different attack function to give the wrong result in the extraction side.The watermark attack is mainly classified into intentional and unintentional attack.

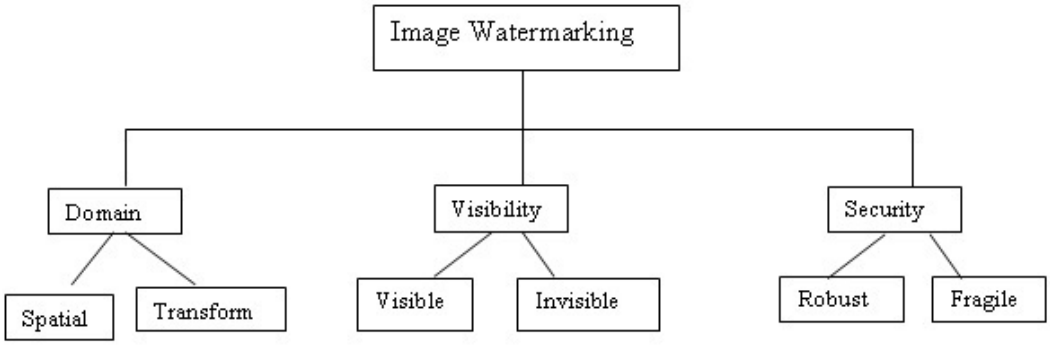

Figure 1. Image watermarking types

In [3] watermarking attacks are classified into four different categories like removal attack, geometric attack, cryptography attack and protocol attack. Removal attack is used to remove the watermark information from the transmitted watermarked image and to affect the image quality. This type of attack will not find out the watermarking technique. Geometric attack is defined as a set of operations performed on the watermarked image. In this paper, later we discuss about this type of attack. Cryptography attack aims to break the watermark system in order to relax the security level by removing or modifying the watermark. Protocol attack is another type of watermarking attack. It tries to affect the entire watermarking system without destroying the watermark.

Passive attack and active attck are two types [4] of image watermarking attack . The passive attack does not modify or destroy the watermark, but tries to detect the secret (watermark) information and watermarking algorithm. This type of attack is not easily identified. Active attacks do the modification in the watermarking system. This type of attack is easily identified.

In this paper, watermarking attack is mainly classified into two categories, based on the information available to the attacker, such as Blind and Non-Blind attack as shown in Figure-2.

ISSN: 0975-8585

WATERAMRKING ATTACK CLASSIFICATION

BLIND ATTACK

The attackers modify or destroy the watermark system without knowing any information about the watermarking process and additional information (key, cover image, etc.). This will affect the security of the watermarking system by changing or unchanging the image quality. Blind attack is again classified into Common signal processing attack, Pixel reallocation attack and Synchronization attack.

Signal Processing attack:

This type of attack occurs unintentionally when the image is transmitted through non secured channel and when signal processing techniques are used to remove the additional noises.
The common signal processing attacks involve:

(i) Noisy attack
(ii) Filtering attack
(iii) Compression attack

Noisy attack:

The watermarked image is transmitted to the receiver from sender through communication channels. Some channel noises (Gaussian, salt-pepper, etc.,) are added to the original watermarked image. It will affect the original and the watermark image after extraction.

Filtering attack:

At the receiver side, in order to remove the noises in the watermarked image and to enhance the image quality, different types of filters are used like Gaussian, median filter, etc [5]. Along with noises, these filters also remove some of the watermarked image values, which are belonging to the filter frequency ranges.

Compression Attack:

The watermarked image is transmitted to the receiver from the sender, but the file size is very large. In order to reduce the size of the watermarked image during transmission, various lossy compression techniques are used like JPEG, JPEG2000. It will automatically affect the watermarked image. Decompression technique is used to reconstruct the watermarked image from the compressed image. The decompression technique does not always give the exact match of the original image and some data are lost.

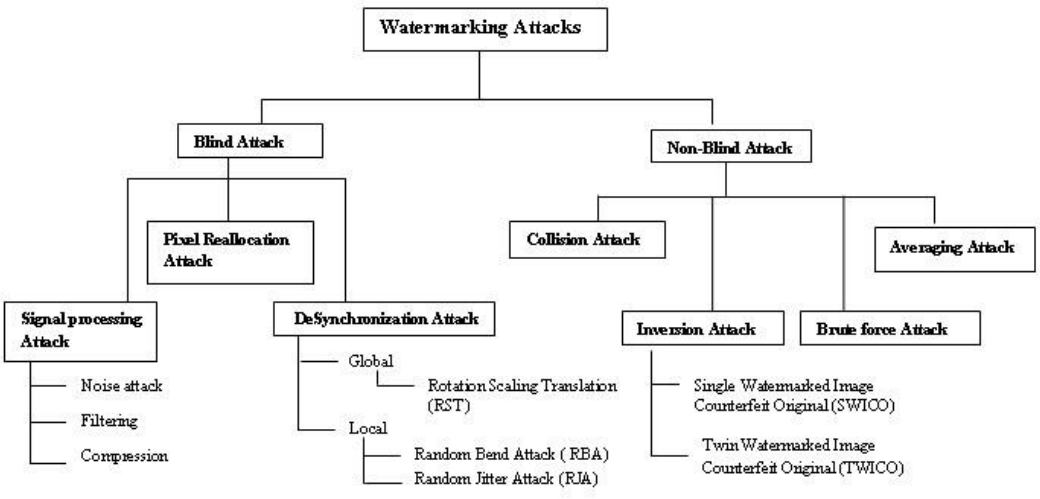

Figure 2. Watermarking Attacks

Desynchronization Attack

Desynchronization attack [7-10] or geometric attack is a special type of attack which does not remove the watermark. But it modifies the watermarked image in such a way that it introduces the synchronization error between original and extracted image. So the detector is not able to find the exact watermark. Geometric attack is again classified into two types such as global and local transformation.

Global transformation:

Global transformation or global distortion is a set of operation that is performed over an entire image. This affects all the image pixel values of an image uniformly. Rotation, Scaling and Translation (RST) performs global geometric distortion on the watermarked image. RST affects the entire watermarked image. At the extraction side, it is not easy to extract the original watermark.

Local Transformation:

In this type of attack, the attack function affects only a particular part or subset of the watermarked image and not an entire image. The following two attacks are coming under local geometric distortion attack.

(a) Random Bend Attack (RBA):

RBA [6] affects the image by doing a small bend or displacement over the small local region of the image or the image pixel grid. The size of regions is not constant because the attack distortion frequency varies randomly .

(b) Random jitter attack (RJA)

RJA [6,11]is defined as distortion or displacement of the pixel location of the image sampling grid followed by the interpolation. In some of the image pixels, random amount of jitter effect is added . Under this attack, rows and columns of an image are randomly removed.

Pixel Reallocation attack (PRA):

Pixel Reallocation attack PRA [12] is a simple attack which aims to modify the watermarked image and will not remove the watermark. The attacker selects the neighbor pixel of the current pixel randomly. Then, the threshold value is compared with their absolute difference value. If the difference value is lesser than the threshold value, then the current pixel is replaced with that neighbor pixel .

NON-BLIND ATTACK

In this type of attack, the attacker has the knowledge about any one of the following information such as watermarking technique, watermark, original image, key, etc. Non blind attack is again classified into collision attack, statistical averaging attack, inversion attack and brute force attack.

Collision attack:

Collision attack [13] is aimed to remove the watermark. Attackers already know some of the watermark templates. Using these templates, the attacker removes the watermark from the watermarked image and inserts their own information within a watermarked image. Then, attacker transmits modified watermarked image to the destination and claims to be the owner .

Statistical averaging attack:

Statistical averaging attack [14] is similar to collision attack. The attacker has a number of different watermarked images with same watermark information. But the attacker does not know the watermark. Now the attacker easily identifies the watermark by averaging these entire watermarked images. Similarly, the original image is identified by using a number of different watermarks.

ISSN: 0975-8585

Inversion Attack:

This attack comes under protocol attack. The main goal is to claim the ownership of the document. The attacker creates fake watermarked images or original image in order to get confusion about the owner of the images by embedding his own watermark within the original watermarked image. Using the counterfeit watermarking system, this attack creates an invertible watermarking system to allow multiple claims[15-17]. Two types of inversion attacks are present, they are single watermarked image counterfeit original and twin watermarked image counterfeit original.

Single Watermarked Image Counterfeit Original (SWICO)

In SWICO[14,15] attack, the attacker modifies the ownership of the original image (I) by embedding the fake watermark (w ') within an original watermarked image (I_w) to generate a fake watermarked image ($I_w^{'}$).

Twin Watermarked Image Counterfeit Original (TWICO)

In TWICO [14,15] attack, Using the quasi invertible technique, the attacker modifies the ownership of the original image (I) by embedding the fake watermark (w') within a counterfeit watermarked image ($I_w^{'}$) to generate a fake watermarked image ($I_w^{''}$) .

Brute force attack:

This type of attack also comes under protocol attack and it is not the watermark removal attack. The main aim of this attack is identifying the watermark information. The attacker first knows the watermark detection algorithm. Now the attacker tries all possible keys to find out the watermark without modifying or destroying the watermarked image.

BENCHMARKS

Image watermarking technique is evaluated based on different watermarking requirements such as robustness, security, availability, imperceptibility, etc. using various image watermarking benchmark systems [18,19] . The major four different benchmarks are Stirmark, Optimark, Checkmark and Certimark.

Stirmark:

This is the first benchmark developed by Petitcolas in 1996 in order to provide an automatic evaluation of watermarking system. It uses windows operating system. This benchmark divides the various attacks into 9 different categories. It mainly concentrates on geometric attack and checks whether a message is decoded or not under each attack.

Optimark:

It is developed by Solachidis et al and uses windows operating system. They use Graphical User Interface (GUI). It evaluates image watermarking system with its statistical characteristics. It mainly checks the detection or decoding performance on the receiver side. The statistical dependency of the keys and watermarking message is also tested. This includes geometry, filtering and compression attacks for evaluation.

Checkmark:

This benchmark evaluates watermarking system with broad range image watermarking attack. It is related to certimark. It uses UNIX and windows operating system to evaluate the image watermarking technique. It uses wavelet transform to test removal and denoising attacks which are not included in Stirmark.

Certimark

Certimark (Certification for watermarking) benchmark is used to design, develop and publish a benchmark suite for image watermarking. It labels the image watermarking system with international certificate. It evaluates the watermarking system and certified based on different parameters and attacks.

IMAGE WATERMARKING METHODS AGAINST ATTACKS

Today different image watermarking techniques are developed to protect the transmitted image from different image watermarking attacks. Mostly transform domain is used in watermarking techniques than spatial domain. Table-1 lists few image watermarking techniques resistant to different attacks. In section 2, different watermarking attacks are studied. In order to evaluate the robustness of the each image watermarking technique, various image watermarking attacks are applied to the watermarked image. At the receiver side, from the attacked watermarked image, whether or not the original watermark information is extracted will be verified. The original reconstructed image quality is also verified. In Table-1, signal processing, filtering, compression and local geometric attacks are discussed for each method. The robustness of the each watermarking method are evaluated using different performance measures like Peak Signal to Noise ratio (PSNR), Detection Rate (DR), Bit Error rate (BER), Normalized Correlation (NC), Structural Similarity Index Measure (SSIM).

Table-1: Robust image watermarking techniques

Methods	Attacks	Measures value
Harris–Laplace detector, local characteristic region, discrete Fourier transform [12]	Median filter (3X3)	Detection Rate = 3/6
	Gaussian noise	Detection Rate =2/6
	JPEG compression	Detection Rate = 4/6
	Scaling (1.4)	Detection Rate =1/6
	Rotation (30)	Detection Rate =2/6
	Translation (0.6)	Detection Rate =1/6
	Remove 8 rows & 16 columns	Detection Rate =5/6
	Local Random Bending	Detection Rate =3/6
DWT, GMM [20]	Salt & pepper Noise	BER=8.31
	JPEG	BER=0 to 0.07 (Quality factor >50%)
	Median filter (5X5)	BER=7.89
	Gaussian filter (5X5)	BER=0.55
	Scaling	BER=0.20
Image Texture feature and DWT [21]	Rotation	SSIM=0.33
	JPEG	PSNR= 51.65
Wavelet Transform, Significant Amplitude Difference (SAD),Dither modulation [23]	Median Filtering	NC=0.93
	Gaussian Filtering	NC=0.99
	Cropping (25%)	NC=0.88
	Resize (0.5)	NC=0.97
	Rotation (-0.25) (0.25)	NC=0.47, 0.42
	JPEG	NC=1
	JPEG 2000	NC=1
Slantlet transform (SLT) , Mean value difference [24]	Histogram equalization	BER=0.0068
	Median filtering	BER=0.0371
	Sharpening	BER=0
	JPEG	Φ=0.8
	JPEG 2000	Φ=0.9
	Additive Gaussian Noise (AGN)	Φ=0.9
DWT, Gradient vectors, Scrambling, Difference angle	Gaussian Filtering	BER=0
	Cropping (20%)	BER=0.2 %

quantization index modulation(DAQM) [22]	Scaling (0.75)	BER = 0%
	JPEG Quantization Noise (20)	BER=1.42 %
Local polar harmonic transform, Speeded up robust feature, affine invariant local feature [25]	Median Filter	Detection rate = 8/11
	Gaussian Filter	Detection rate = 8/11
	JPEG compression	Detection rate = 10/11
	Scaling 150%	Detection rate = 7/11
	Rotation 30°	Detection rate = 7/11
	Removed 5 rows and 17 columns	Detection rate = 6/11
Fractional wavelet packet, non linear chaotic map, Singular value Decomposition [26]	Gaussian Noise	Watermark extracted 100%
	Salt and Pepper	Watermark extracted 100%
	JPEG compression	Up to CR 100:1
	Row–column deletion	Up to 20-R & 20-C
	Contrast adjustment	Up to 80%
Contourlet Transform, Arnold transform, Singular Value decomposition, Triangular Number Generation [17]	Inversion Attack	Robust

CONCLUSION

In this paper different image watermarking technique attacks are studied. It is classified into blind and non blind image watermarking attacks based on the knowledge of the attacker. Blind watermarking attacks mainly occur unintentionally because of preprocessing, transformation and compression process. Also, it modifies the watermarked image, but does not remove the watermark. But Non-blind attacks are performed intentionally with the existing knowledge about the image watermarking system to remove or modify the watermarking system. In this paper, to evaluate the image watermarking system with various attacks, four benchmark tools are also studied. The strength of the image watermarking algorithm is analyzed with few measures and in geometric attack, only particular cases are discussed. So in future, a generalized algorithm is needed to meet out these existing limitations.

REFERENCES

[1] PotdarV M, SongH, Chang E. A survey of digital image watermarking techniques.2005. 3rd International conference on Industrial Informatics(INDIN' 05); 2005, pp.709-16.

[2] Priya S,Santhi B,Swaminathan P.Image watermarking techniques - a review. Research journal of applied sciences, engineering, technology.2012,4(14), pp.2251-254.

[3] Voloshynovskiy S,Pereira S,Pun T,Eggers JJ,Su, JK. Attacks on digital watermarks: classification, estimation based attacks and benchmarks.IEEE Communications Magazine. 2001,39(8),pp.118-26.

[4] Nyeem H,Boles W,Boyd C. Digital image watermarking: its formal model, fundamental properties and possible attacks. EURASIP Journal on Advances in Signal Processing. 2014,135(August), pp.1-22.

[5] Chung Y D,Kim C H. Robust image watermarking against filtering attacks. SCIE annual conference;Vol. 3;2003,pp. 4-6.

[6] Licks V,Jordan R. Geometric Attacks on Image Watermarking Systems. IEEE multimedia. 2005,12(3),pp.68-78.

[7] Zheng D, Liu Y, Zhao J, Saddik A E.A survey of RST invariant image watermarking an Algorithm. ACM Computing. Surveys(CSER),2007,39(2),pp.1–91.

[8] Tang C W,Hang H M. A feature-based robust digital image watermarkingscheme. IEEE Transactions on Signal Processing,2003,51(4),pp.950–59.

[9] Dong P, BrankovJ G,Galatsanos N P,Yang Y, Davoine F. Digital watermarking robust to geometric distortion. IEEE transcation on image processing. 2005,14(12),pp.2140-150.

[10] Guo J, Zheng P, Huang J. Secure watermarking scheme against watermark attacks in the encrypted domain. Journal of Visual Communication and Image Represantation.2015,30(c),pp.125–35.

[11] Licks K, Ourique F,Jordan R,Pirez-Gonzalez P. The Effect Of The Random Jitter Attack On The Bit Error Rate Performance Of Spatial Domain Image Watermarking.2003.International Conference on Image Processing,2003,vol. 2,ICIP'03,pp. 455-58.

[12] Tsang K F, Au O C, A review on attacks, problems and weakness of digital watermarking and the pixel reallocation attack. Proc. SPIE—Security and Watermarking of Multimedia Contents, vol.4314,2001,pp.385–93.
[13] Mitekin V A, Timbay E I.A new watermarking sequence generation algorithm for collision-free digital Watermarking. 2012 8th International conference on Intelligent Information Hiding and Multimedia Signal Processing,Piraeus,2012,pp.256-60.
[14] Freire L P,Comesana P,Ramon J,Pastoriza T. Watermarking security-A survey.Transactions on Data Hiding and Multimedia Security I, Springer-Verlag: Heidekberg, 2006,pp.46-72.
[15] Craver S,Memon N,Yeo B L,Yeung M. Resolving rightful ownership with invisible watermarking techniques : limitations, attacks and implementations. IEEE journal of Selected Areas in communications,1998,pp.573-86.
[16] Velumani R. Inversion attack.Encyclopedia of cryptography and security. 2011,pp.160-80.
[17] Velllaisamy S,Ramesh V. Inversion attack resilient zero-watermarking scheme for medical image authentication. IEEE journal on Image processing,2014,8(12),pp.718-27.
[18] Macq B,Dittmann J,Delp E J.Benchmarking of image watermarking algorithms for digital rights management. Proceedings of the IEEE, 2004,92(6),pp.971–84.
[19] Solachidis D V,Tefas A,Nikolaidis N,Tsekeridou S,Nikolaidis A,Pitas I. A benchmarking protocol for watermarking methods. International Conference onn Image Processing,2001,pp.1-4.
[20] M Amirmazlaghani , M Rezghi, H Amindavar.A novel robust scaling image watermarking scheme based on Gaussian Mixture Model. Expert Systems with Applications.2015:42:1960–1971.
[21] Nambakhsh M S,Ahmadian A, Zaidi H. A contextual based double watermarking of PET images by patient ID and ECG signal. computer methods and programs in biomedicine. 2011,104(3),pp418–25.
[22] Cai N,Zhu N,Weng S, Ling B W-K. Difference angle quantization index modulation scheme for image watermarking. Signal Processing: Image Communication. 2015,34(May),,pp.52–60.
[23] Li C,Zhang Z,Wang Y,Ma B,Huang D. Dither modulation of significant amplitude difference for wavelet based robust watermarking. Neurocomputing, 2015,166 (October),pp. 404-15.
[24] Thabit R,EeKhoo B. A new robust lossless data hiding scheme and its application to color medical images. Digital Signal Processing.2015,38,(March),pp. 77–94.
[25] Wang X Y,Liu Y N, Li S, Yang H Y,NiuP P, Zhang Y. A new robust digital watermarking using local polar harmonic transform. Computers and Electrical Engineering.2015,pp. 1-16.Article in press.
[26] Bhatnagar G,WuQM J,Atrey P K. Robust logo watermarking using biometrics inspired key generation. Expert Systems with Applications. 2014,41(10), 4563–578.
[27] Thanki R M,Borisagar K R. Experimental study of sparse watermarking techniques for Multibiometric system. Indian journal of Science and technology,2015,8(1),pp. 42-48