Robust Digital Image Watermarking based on complex wavelet transform

TERZIJA NATAŠA, GEISSELHARDT WALTER
Institute of Information Technology
University Duisburg-Essen
Bismarckstr. 81, 47057 Duisburg
GERMANY

Abstract: - In this paper a new robust digital image watermarking method is presented. It is based on the Complex Wavelet Transform (CWT) and combines the image feature extraction. The goal is to resist both geometric distortion and signal processing attacks. The proposed method is performed in the spatial domain. By this, the original image is decomposed into four level of CWT decomposition. Every level of CWT decomposition is further separately reconstructed using the Inverse Complex Wavelet Transform (ICWT), in order to form its spatial representation (the *channel*),. For the embedding purpose the channel of the fourth level of decomposition is selected. With the Harris feature detector the salient points of the fourth channel are detected and around the salient points a pseudorandom sequence used as the watermark is repeatedly embedded. Algorithm robustness is tested on the following attacks: JPEG, JPEG2000 compressions, geometrical attacks (rotation, scaling, cropping) and different filtering attacks carried out by using the benchmark software Checkmark.

Key-Words: - Watermarking, Complex wavelet transform, feature point detectors, robust algorithms, attacks.

1 Introduction

Digital watermarking is the process that embeds data called a watermark into a multimedia object such that watermark can be detected or extracted later to make an assertion about the object [1]-[4]. The up to date known watermarking applications considered in the open literature are as follows: *Copyright Protection* - for the protection of intellectual property; *Fingerprinting* - to trace the source of illegal copies; *Copy protection* - the information stored in watermark can directly control digital recording devices for copy protection purposes; *Broadcast monitoring* - by embedding watermark in commercial advertisements, an automated monitoring system can verify whether the advertisements are broadcasted as contracted; *Data authentication* - fragile watermarks can be used to check the authenticity of data; *Indexing* - indexing of video mail; indexing of movies and news items, where markers and comments can be inserted that can be used by search engines; *Medical safety* - embedding the date and the patient's name in medical images could be useful safety measure, *Data Hiding* - watermark techniques can be used for the transmission of secret messages. The ideal properties of a digital watermark include features like imperceptibility and robustness. A watermarked image should retain as closely as possible the quality of the original image and at the same time the watermark should be robust to various types of image processing techniques or attacks applied to remove the watermark. Watermarking algorithms can be performed either in spatial or in transform domain. The spatial-domain techniques modify directly the intensities or color values of some selected pixels. The transform-domain techniques modify the values of some transformed coefficients. That means that the transform domain coefficients must be modified to embed the watermark and finally the inverse transform should be applied to obtain the marked image. The following transformations commonly used for watermarking purposes are respectively: Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT) and Discrete Wavelet Transform, Complex Wavelet Transform (CWT), or Fourier-Mellin Transform (FMT). In this paper, a new transformation domain watermarking algorithm will be presented. It uses the Dual-tree Complex Wavelet Transform (DT-CWT). The embedding technique is performed in the spatial channel number four, obtained in the same way as in [5]. The channel four it obtained by applying the inverse CWT transform only on the coefficients of the fourth level of CWT decomposition. By this, the coefficients of all other levels of decomposition, including the approximation level are set to zero. In that way, the channel four presents the spatial representation of the CWT coefficients from the fourth level of the CWT decomposition. It is used as an embedding domain because in [5] was shown that the watermark embedded in spatial channel four was more robust on image processing operations than watermark embedded in other spatial channels. In many watermarking techniques the watermark pattern is embedded relative to image size. When the geometrical attack is occurred, the image size is changed and the watermark detector/decoder is not more able to detect/extract the watermark without using the synchronization technique. In our approach the watermark pattern is embedded relative to salient points of the channel four and in the watermark detection procedure the watermark pattern can be easily detected only by finding the salient points in the channel four of the tested image. By this, the salient points are computed using the Harris corner detector [6]. In order to be robust on cropping attacks the same watermark is repeatedly embedded in every non-overlapping channel block located around the salient point of the fourth channel.

This paper is organized as follows. In Section 2 the most important properties of the DT-CWT transform for the watermarking techniques are briefly described. In Section 3, the watermarking embedding technique is introduced. This scheme is further tested in Section 4, and the paper conclusion is given in Section 5.

2 DT-CWT Properties

In this Section the most important properties of the DT-CWT transform for the watermarking application are briefly described. The existing watermarking approaches which are based on the CWT are described, as well.

The DT-CWT is one of the transformations which belongs to the group of shift-invariant transformations. By this, a transformation is shift-invariant if it produces subbands such that the total energy of the coefficients in any subband is unaffected by translations applied to the original image. Complex wavelets have not been used widely in image processing due to the difficulty in designing complex filters which satisfy the perfect reconstruction (PR) property. Kingsbury [7] proposed a dual tree implementation of the CWT (DT-CWT) which is approximately shift-invariant Two fully decimated trees are produced by downsampling effect by taking first the even and then the odd samples after the first level filters. To get the uniform intervals between the two trees' samples, the subsequent filters in one tree must have delays that are half a sample different. Two wavelet trees have the real filters which operate in parallel giving the real and imaginary parts of a complex filter. Using the two wavelet trees for 1-D signals instead of one tree like in the standard DWT, the redundancy of 2:1 is introduced. DT-CWT offers both magnitude and phase information. To compute the 2-D CWT of images, these two trees are applied to the rows and than columns of the image, as in the conventional DWT. This operation results in 6 subbands per resolution instead of 3 as in the DWT. The subbands are oriented with the orientation of: ±15, ±45 and ±75 degree. The magnitude of the DT-CWT is approximately shift-invariant. The main features of the DT-CWT are:

- *Approximate shift-invariance*
- *Good directional selectivity* in 2-D
- *Perfect reconstruction* using short linear-phase filters
- *Limited redundancy*: 2:1 in 1-D and 4:1 in 2-D
- *Low computation* comparing to other shift-invariant transformations.

The DTCWT is inherently sensitive to rotation, but the sum of the energies of coefficients across all 6 directional subbands is reasonably invariant to rotation. In [8] Hill extracts the rotation-invariant texture feature using the DT-CWT transform. His approach is based on the Fourier analysis of the 6 CWT oriented subbands energies. The rotation-invariant features are extracted from the energies of the DT-CWT shift-invariant oriented subbands.

Typically a standard transform watermarking technique is based on the addition of a pseudorandom sequence to the host image coefficients in one of the transformation domains like DCT, DFT, DWT, etc. A watermarked image is obtained by taking the inverse transform. Applying the same transformation on the watermarked image and extracting the watermark sequence from the transformation coefficients, it is expected that the extracted watermark is the same as the embedded watermark. Using the CWT transformation this is not a case. Due to the redundancy of 4:1 some parts of the watermark sequence will be lost. Such technique is not suitable for the watermarking in the CWT domain. There are several existing watermarking approaches which are based on the CWT transformation. In [9], the authors based their approach on the addition of the CWT coefficient of the pseudorandom image to the CWT coefficients of the original image. The watermark embedding is performed in every level of decomposition separately and the CWT coefficients of the pseudorandom image are weighted with different visual masks. The standard correlation based watermark detector has been applied; the robustness of the presented method on the watermarking attacks was not investigated. In [10] Loo performs embedding in the spatial domain. A random image of ±1 of the same size as the host image is generated and the CWT coefficients of both images are computed. Independently for each subband, the scaling factors (the *visual mask*) are computed from the host image's CWT coefficients. The random image coefficients are modulated by the payload and further scaled with the computed visual mask. After that, the modulated coefficients are further inverse transformed to form a watermark. Finally, in order to obtain the marked image, the watermark is added to the host image in the spatial domain. The robustness of the method is tested on compression attacks (JPEG, JPEG2000), AWGN (Additive White Gaussian Noise), mean and median filtering, denoising and remodulation attack [11]. For the yes/no watermarking detection principle (where *yes* denotes that the watermark is detected in the tested image), the algorithm showed very good performance for all tested attacks except in a case of remodulation attack. In [10] Loo based his approach on the Chen quantisation based watermarking algorithm [12]. Here the *spread transform*, which is a combination of spread spectrum technique and quantisation based watermarking, is used. The algorithm has the following steps: The host image is divided into blocks and the payload is divided into equal-sized portions, with each portion being embedded in a separate block. A pseudorandom vector of ±1 (the same size as one block) is generated for each bit of the payload to be embedded. The forward CWT of this vector is computed, and scaled according to local image activity. The scaled coefficients of this vector are frequency partitioned into 3 vectors, with each one reconstructed back into the spatial domain using only one level of CWT coefficients. From this instant the watermarking process takes place in the spatial domain, the CWT domain is only used for adapting the random vector to the image. The method is tested on AGWN attack and compression (JPEG and JPEG2000) attack.

In the previous work of authors [5], the following approach is proposed: the original image is decomposed into four levels of decomposition. In order to form its spatial representation (the *channel*) every level of CWT decomposition is separately reconstructed back into the spatial domain by using ICWT. The bipolar watermark sequence is embedded in the embedding channels 3 and 4, reconstructed from the level three and four of the CWT coefficients, respectively. The largest and the smallest values of the embedding channels are modified according to the watermark sequence with the same strength factor. By adding the watermarked channels to the approximation and channels reconstructed from levels 1 and 2 of the CWT coefficients, the watermarked image is obtained. The method is robust on non-geometrical distortion like JPEG, JPEG2000 compressions, filtering (gaussian, median, sharpening) attacks, as well as on dithering and thresholding. This technique showed that the watermark robustness was greater in the embedding channel four than in the other embedding channels. This result was expected due to the fact that the embedding channel four presents the spatial representation of the CWT coefficients which are close to the lower frequency part of the image and they are small affected by the image transformations. In the case of geometrical attacks the watermark detector was not able to detect the watermark because the watermark pattern was embedded relative to image size.

3. Embedding Method

In the following paragraphs the embedding method will be presented.

3.1 Embedding procedure

To improve the robustness on geometrical attacks redundancy is added into the embedding procedure. A pseudorandom sequence is used as the watermark. The same watermark is embedded into the different image blocks using an additive spread spectrum technique. Watermark embedding procedure can be described using the following steps.

1. The original image $\mathbf{I}(x,y), x=1,..,N_1, y=1,..,N_2$ (where N_1 and N_2 are image dimensions) is firstly decomposed into the four-level CWT. Every level of CWT decomposition is separately ICWT reconstructed to form the spatial representation (a channel). The channels are denoted as $\mathbf{X}_1(x,y)$, $\mathbf{X}_2(\mathrm{x,y})$, $\mathbf{X}_3(x,y)$, $\mathbf{X}_4(x,y)$, $\mathbf{X}_\mathrm{a}(x,y)$, or shortly as $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$, $\mathbf{X}_4$ for detail channels and $\mathbf{X}_\mathrm{a}$ for approximation channel.
2. For the embedding purpose select the channel $\mathbf{X}_4$. Apply the Harris corner detector [6] to the channel $\mathbf{X}_4$ and find the set V of the salient points of that channel.
3. Generate a pseudorandom sequence with length of l bits. The pseudorandom sequence will be arranged in the form of 2-D matrix $\mathbf{W}$, used as the watermark. The watermark dimension will be $M_1 \mathrm{x} M_2$ and $l < M_1 M_2$.
4. Divide the channel $\mathbf{X}_4$ into the $M_1 \mathrm{x} M_2$ blocks in such a way that the first point in every block will be one of the salient point obtained in the step 2.
5. Embed the watermark into the block according to the following formula:
$$\mathbf{X}_{4\mathrm{w}}(u+i,v+j) = \mathbf{X}_4(u+i,v+j) + \alpha\mathbf{W}(i+1,j+1)\,, \qquad (1)$$
where $i=0,\ldots,M_1-1, j=0,\ldots,M_2-1$. Here $\mathbf{X}_4(u,v)\in V$ presents the starting salient point in a block. α is the strength parameter which controls the level of the watermark $\mathbf{W}$.
6. Adding the watermarked channel $\mathbf{X}_{4\mathrm{w}}$ to the other channels $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$ and $\mathbf{X}_\mathrm{a}$ the watermark image $\mathbf{I}_\mathrm{w}$ is obtained.

Due to noise added to the watermarked image by attacks or transmission over the communication channel, the watermarked image can be destroyed. In the extraction procedure such an image will be called the received image, $\mathbf{I}_\mathrm{r}$, instead the watermark image, $\mathbf{I}_\mathrm{w}$.

3.2 Detection procedure

Detection procedure consists of two stages. If the watermarked image is altered with an affine geometrical distortion, firstly the parameters of affine transformation will be computed using the original and received image. For that purpose the scale invariant feature detector [13], [14] will be used and the pairs of the affine-invariant points on original and received image will be detected. Scale Invariant Feature Transform (SIFT) is an approach developed by Lowe [14] for detecting and extracting the local feature descriptors that are invariant to image rotation and scaling, and partially invariant to change in illumination, image noise and small changes in viewpoint. Performing the SIFT on original and received image the local descriptors of all affine invariant points are computed. Computing the correlation coefficient between the local descriptors of both original and received image and comparing it with the threshold, the pairs of affine invariant points are detected. Using the positions of affine invariant feature pairs, the parameters of the affine transformations can be computed. After that, the image is inverted and the watermark detection procedure is performed. In the case of cropping attacks or non-geometrical attacks, the watermark detection procedure can be directly implemented to the received image.

The first stage in the watermark detection procedure is to select the channel block which possibly contains the watermark. After that, the watermark detector takes the selected channel block, possibly altered by the attack, and the watermark. It measures the correlation coefficient between the watermark and the channel block and compares it with the threshold. If the value of the correlation coefficient is greater than the threshold, the watermark is detected. The correlation coefficient is used because it does not depend on the magnitude of the input arguments. On the other hand, as an output it delivers a single bit indicating whether an image contains the watermark, or not. The threshold depends on the probability of false positive. A false positive or false detection occurs when the detector incorrectly concludes that an unwatermarked image contains a given watermark.

Detection procedure is performed by the following steps:

1. Apply the four-level CWT transform on the received image $\mathbf{I}_{\mathrm{r}}$. To get the spatial channels $\mathbf{X}_{\mathrm{r}1}$, $\mathbf{X}_{\mathrm{r}2}$, $\mathbf{X}_{\mathrm{r}3}$, $\mathbf{X}_{\mathrm{r}4}$, reconstruct each level of decomposition separately with ICWT.
2. Apply the Harris corner detector to the channel $\mathbf{X}_{\mathrm{r}4}$ and find the set V_r of the salient points of that channel.
3. For every salient point $\mathbf{X}_{r4}(u,v) \in V_r$ from the set V_r, starting from that point extract the $M_1 \times M_2$ block $\mathbf{W}_e$ from the $\mathbf{X}_{\mathrm{r}4}$ and compare it with original watermark $\mathbf{W}$. Apply the correlation:

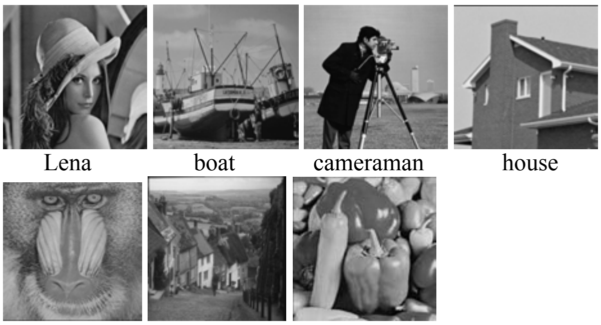
$$corr(\mathbf{W}, \mathbf{W}_{\mathrm{e}}) \geq \eta(P_{fp}) \qquad (2)$$

where $\eta(P_{fp})$ represents a threshold depending on the false-positive probability. The probability of false positive will be obtained by running the detector on the unwatermarked image $\mathbf{I}(x,y)$. The highest correlation value between the unwatermarked image block and the watermark, computed for all image blocks, will be used as a detection threshold.

$$P_{fp} = P\{T_{\max} > T\} \qquad (3)$$

4 Algorithm Testing

For the testing purposes of the watermarking method presented in this paper a few standard gray-scale images (256x256) are used: *Lena, boat, cameraman, house, baboon, goldhill, peppers* (see Fig. 1).



Lena boat cameraman house

baboon goldhill peppers

Fig. 1. Standard watermarked test images

In every tested image the pseudorandom sequence of 50 bits is used to form a watermark. The watermark is embedded into the channel $\mathbf{X}_4$ (see equation (1), Section 3). The selected block sizes were 20×20. The strength parameter α is selected for every image differently in such a way that the computed Peak Signal to Noise Ratio (PSNR) is greater than 45 dB. This parameter presents the tradeoff between the invisibility constant and robustness and it is adaptively selected. For the greater value of parameter α the watermark energy can be moved to the other CWT subbands, so the watermark cannot be detected from the channel $\mathbf{X}_{\mathrm{r}4}$. This feature will be used here for the watermark detection and the watermark detection procedure will be also performed for the channels $\mathbf{X}_{\mathrm{r}1}$, $\mathbf{X}_{\mathrm{r}2}$ and $\mathbf{X}_{\mathrm{r}3}$. Detection threshold is empirically determined by performing the detection procedure on unwatermarked image and computed the correlation coefficient between the unwatermarked channel block and the watermark sequence. Maximal correlation value computed for all image blocks was used as the detection threshold. For every channel $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$ and $\mathbf{X}_4$ the different thresholds are obtained. The watermark will be successfully detected in one channel if the value of the correlation coefficient between the tested image block from the received image and original watermark is greater then the detection threshold for that channel.

In order to test the robustness of the presented algorithm various Checkmark attacks are performed using the Checkmark software [11], [15]. Among them are:

- compression attacks: JPEG and JPEG2000 compression attacks with various quality factors;
- filtering attacks: gaus1 (Gaussian filtering with a 3×3 window size), gaus2 (Gaussian filtering with a 5×5 window size), medfilt1 (median filtering with window size 2×2), medfilt2 (median filtering with window

size 3×3), trimean1 (trimmed mean filtering with a 7×7 window), trimean2 (trimmed mean filtering with a 11×11 window), sharpening (a 3×3 high pass filter is applied to the image to accentuate edges), wiener1 (wiener filtering with window size 3×3), wiener2 (wiener filtering with window size 5×5);

- geometrical attacks: combination of rotation, scaling and cropping attack is performed and cropping attack with the different percentage of cropped image size.

Detection results after different attacks are given in Table 1. "*One*" in the table means that the watermark is detected whereas "*zero*" means that it is not detected. It can be observed from the Table 1 that for the compression attacks with lower bitrates for JPEG2000 or quality factors for JPEG the method gives very good results. The method is effective for JPEG2000 bitrates till 0.4 for most of the images. Till the quality factors of 30 for JPEG compression the tested method showed good results. It can be shown from the table that the algorithm has a good performance for all types of filtering attacks except for trimmedmean attack.

In Table 1 the cropping attack is also analyzed. The percentage figures denote the size of image after cropping. Only few images with less then 65% of image size are robust on the cropping attack. It strongly depends on the image content and the distribution of the salient points.

Three geometrical attacks are here performed. Geom1-3 from the table denote the attacks which consist of rotation, followed by the scaling and cropping. Geom1 presents rotation of 20 degrees, scaling 1.2 of the image size and cropping, geom2: rotation of 45 degrees, scaling 1.3, cropping and geom3: rotation of 8 degrees, scaling 2.2 and cropping. These attacks are presented on the Figures 2-4 for Lena image. In order to compute the parameters of affine transformation the SIFT procedure is performed. The local descriptors of the affine invariant regions of the original (or watermarked) image and geometrically distorted image are computed. Using the simple correlation (with correlation threshold 0.9) of the descriptors of two images, the matched pairs of affine-invariant image points are found. When the matched pairs are found the parameters of the affine transformation are computed and the image is inverted. The detection procedure is than performed and the watermark was successfully detected (see Table 1).

In Fig. 2-4 the watermarked Lena image and the three geometrically distorted images with geom1-3 attack are presented with the affine-invariant matched points. From the Fig. 2-4 it can be seen that for different distortion the different invariant-image pairs are found.

5 Conclusion

In this paper a new digital image watermarking method developed in the CWT domain is presented and tested. In the testing procedure seven test images are processed. In every image the watermark sequence was embedded. The computed PSNR value was greater than 45 dB. The robustness of methods against various attacks is investigated, as well. Algorithm was robust on all tested filtering attacks: gaussian, median, wiener, sharpening, trimmed mean filtering. For the compression attack with lower bitrates for JPEG2000 or quality factors for JPEG the method gives very good results, except for the baboon image. JPEG2000 compression was more effective in destroying the watermark then JPEG compression. Robustness on the geometrical attacks was improved using the affine-invariant feature detector. The watermark was successfully detected after geometrical attacks which are combination of scaling, rotation and cropping attacks. Robustness on the cropping attack is investigated. Only few images can resist the cropping attack with less then 65% of the image size. It strongly depends on the image content and on distribution of the selected image blocks. All aforementioned results suggest that this technique might be a promising watermarking technique.

References:

[1] S. Katzenbeisser, *Information Hiding* (Boston, London: Artech House, 2000).

[2] G. Voyatzis, I. Pitas, Protecting digital image copyrights: A framework, *IEEE Comput. Graph. Applicat.*, vol. 19, pp. 18-23, Jan. 1999.

[3] *Signal Processing*, vol. 66, no. 3, 1998. Special issue on watermarking.

[4] *Proc. IEEE*, vol. 87, July 1999. Special issue on identification and protection of multimedia information.

[5] N. Terzija, W. Geisselhardt, Digital Image Watermarking Using Complex Wavelet Transform, ACM Multimedia and Security Workshop 2004, Magdeburg, Germany, September 2004.

[6] C. Harris, M. Stephen, A combined corner and edge detector, in *4th Alvey Vision Conf.*, pages 147-151, 1988.

[7] N G Kingsbury, Image Processing with Complex Wavelets, Phil. Trans. Royal Society London A, 357:2543-2560, September 1999.

[8] P. R. Hill, D. R. Bull, C. N. Canagarajah, Rotationally invariant texture features using the dual-tree complex wavelet transform, *Proceedings of Intern. Conf. of Image Processing,* September 2000.

[9] P. Loo, N.G. Kingsbury, Digital watermarking with complex wavelets, Proc IEE Colloquium on Secure Images and Image Authentication, IEE, London, 10 April, 2000.

[10] P Loo, N G Kingsbury, Digital Watermarking using Complex Wavelets, Proc. IEEE Conf. on Image

Processing, Vancouver, September 11-13, 2000, paper 3608.

[11] S. Voloshynovskiy, S. Pereira, V. Iquise and T. Pun, Attack Modelling: Towards a Second Generation Watermarking Benchmark, *Signal Processing*, Special Issue: Information Theoretic Issues in Digital Watermarking, May 2001.

[12] B. Chen, G. Wornell, Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding, *IEEE Trans. on Information Theory,* 47(4): 1423-1443, May 2001.

[13] K. Mikolajczyk, C. Schmid, Scale and Affine invariant interest point detectors, In *International Journal of Computer Vision* 1(60): 63-86, 2004.

[14] D. Lowe, Distinctive image features from scale invariant keypoints. In *International Journal of Computer Vision* 2(60): 91-110, 2004.

[15]http://watermarking.unige.ch/Checkmark

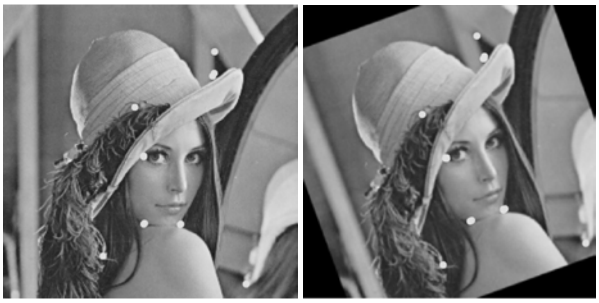


Fig. 2. Left: watermarked Lena image. Right: Watermarked Lena image after geom1 attack. The matched pairs of affine-invariant feature points are presented on the both images.

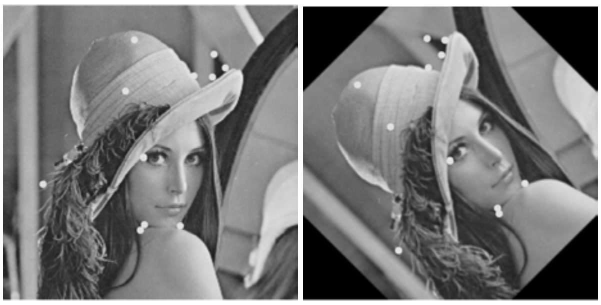


Fig. 3. Left: watermarked Lena image. Right: Watermarked Lena image after geom2 attack. The matched pairs of affine-invariant feature points are presented on the both images.

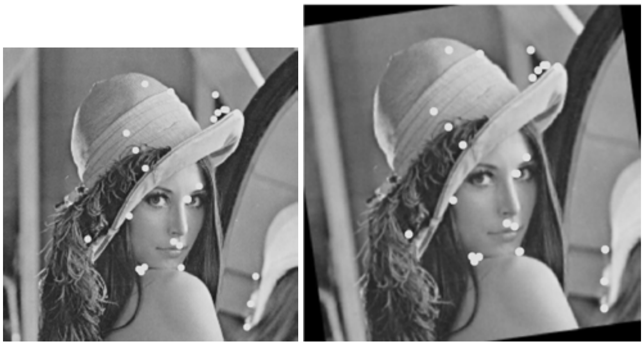


Fig. 4. Left: watermarked Lena image. Right: Watermarked Lena image after geom3 attack. The matched pairs of affine-invariant feature points are presented on the both images.

Table 1. Detection results after compression attacks

Compression attacks	lena	boat	cameraman	house	baboon	goldhill	peppers
JPEG 40	1	1	1	1	1	1	1
JPEG 30	1	1	0	1	1	1	1
JPEG 20	1	1	0	1	0	1	0
JPEG2000 0.4	1	1	1	1	0	1	1
JPEG2000 0.3	0	0	0	0	0	0	0
gaussian1	1	1	1	1	1	1	1
gaussian2	1	1	1	1	1	1	1
median1	1	1	1	1	0	1	1
median2	1	1	1	1	1	1	1
sharpening	1	1	1	1	1	1	1
wiener1	1	1	1	1	1	1	1
wiener2	1	1	1	1	1	1	1
trimmean1	1	1	1	1	1	1	1
trimmean2	0	1	0	1	1	0	1
cropp>65%	1	1	1	1	1	1	1
cropp 40%	1	0	0	0	0	0	1
geom1	1	1	1	1	1	1	1
geom2	1	1	1	1	1	1	1
geom3	1	1	1	1	1	1	1