Send Orders for Reprints to reprints@benthamscience.ae

The Open Automation and Control Systems Journal, 2015, *7*, 2215-2222

A Contourlet-Based Image Watermarking Resisting to JPEG Compression

Junpeng Zhang*, Huaijing Qu and Hongjuan Yang

Shandong Jianzhu University, School of Information & Electrical Engineering, Fengming Road, Lingang Development Zone, Jinan, 250101, P.R. China

Abstract: A novel image authentication algorithm resisting to JPEG compression is proposed. The authentication watermark is generated based on the stability of relationship of adjacent Contourlet transform coefficients before and after JPEG. According to the random block mapping, the watermark is embedded into the discrete cosine transform (DCT) coefficients of other image blocks. The tampering identification matrix is determined by the difference between extracted watermark and reconstructed watermark, and adjacent region comparison is adopted to determine the authenticity of image blocks. Experimental results show that the algorithm can not only achieve blind detection with better resistance to JPEG compression, but also is able to position malicious tampering more accurately.

Keywords: Contourlet transform, Image authentication, JPEG compression, Tampering detection.

1. INTRODUCTION

As an open system for the general public, Internet makes multimedia information encounter a variety of tampering attacks, which can destroy the integrity of information. Image authentication based on digital watermarking technology has become a hot research field [1].

There are two types of image authentication techniques: integrity authentication and robust authentication [2, 3]. Integrity authentication is mainly based on cryptography or fragile watermarks, as long as any slight changes happen, the image will be incomplete and unreliable. On the other hand, the robust authentication is usually based on a semi-fragile watermark which is able to distinguish malicious tampering and accidental operation, where information will be incomplete or unreliable only when image content changes. Most of modifications are not allowed in some important legal evidence images, medical images, *etc*., therefore integrity authentication is required. However, the major part of images will always face some accidental operations such as compression, slight noise, filtering, and image enhancement and so on, which will not affect these image applications in human vision. If authentication cannot pass indiscriminately, the information will be retransmitted, so that the transmission efficiency of digital images will be reduced, therefore robust authentication is required.

During transmission process, images will inevitably be compressed. JPEG compression is widely used and then image authentication resisting to JPEG compression has been the key difficulty in research. Self-embedding watermarking is an important method which requires the authentication watermark being robust to accidental operation but fragile to malicious tampering. The authentication watermark [4] is based on the invariance of the relationships between discrete cosine transform (DCT) coefficients at the same position in separate blocks of an image and the schemes can prevent malicious manipulations but allow JPEG lossy compression. According to semi-fragile characteristics of Zernike moments magnitudes of the low frequency sub-band in discrete wavelet transform (DWT), the algorithm [5] can distinguish malicious tampering from accidental modification, however, has low perceptual invisibility. A semi-fragile watermarking algorithm [6] is proposed which is based on the fact that most of the wavelet coefficients of high frequency have the same relative energy relations before and after JPEG compression, while the scheme cannot achieve blind detection. An effective watermarking system [7] based on a quantization index modulation algorithm is proposed and the approach can survive JPEG compression with good invisibility and robustness. Partial energy relations [8] between groups of 8×8 DCT blocks are used to generate the feature which has the semi-fragile property so that the scheme has good robustness to JPEG compression. An effective self-embedding watermarking scheme [9, 10] is used only for integrity authentication but fragile to JPEG compression.

2. CONTOURLET TRANSFORM

2.1. Contourlet Transform of Image

Contourlet transform [9] is a multi-resolution method with many directions, which can effectively represent an important and complex geometry structure in visual information. Contourlet transform is different from wavelet transform, of which high-frequency information in each layer can be decomposed into three directions, namely LH, HL, and HH in image decomposition. However, with Contourlet

*Address correspondence to this author at the Shandong Jianzhu University, School of Information & Electrical Engineering, Fengming Road, Lingang Development Zone, Jinan, P.R. China, 250101; Tel: +86-13864097756; Fax: +86-531-86367066; E-mail: zjpeng1234@163.com

1874-4443/15

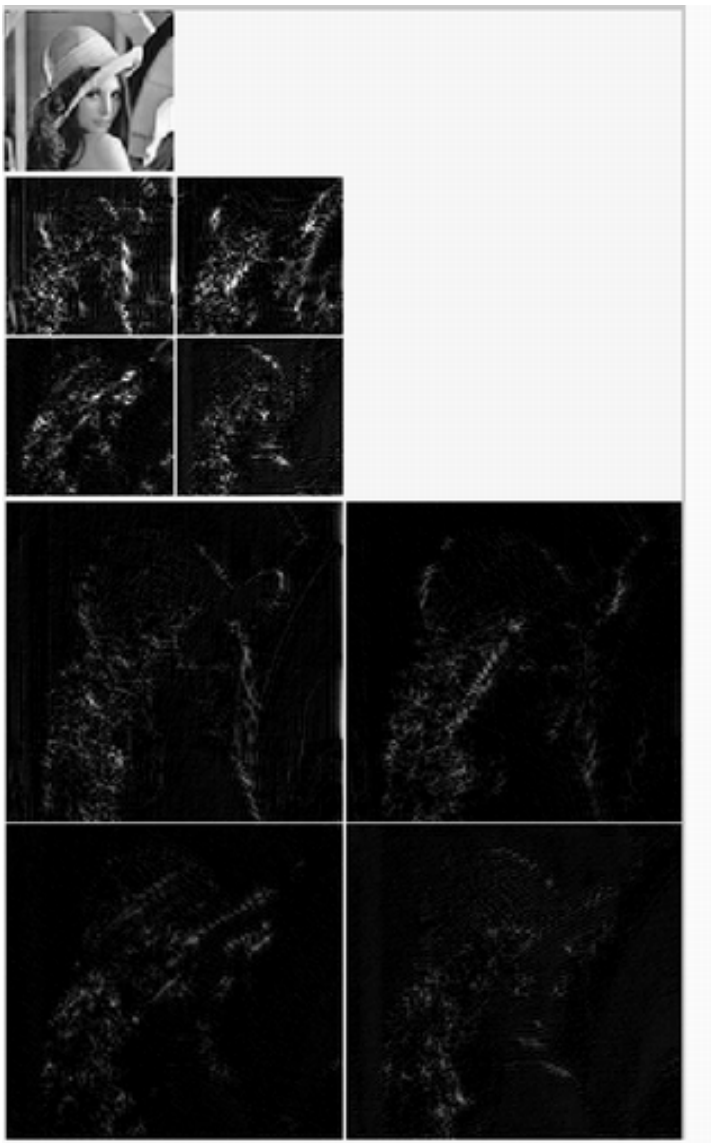


Fig. (1). Contourlet representation of Lena image.

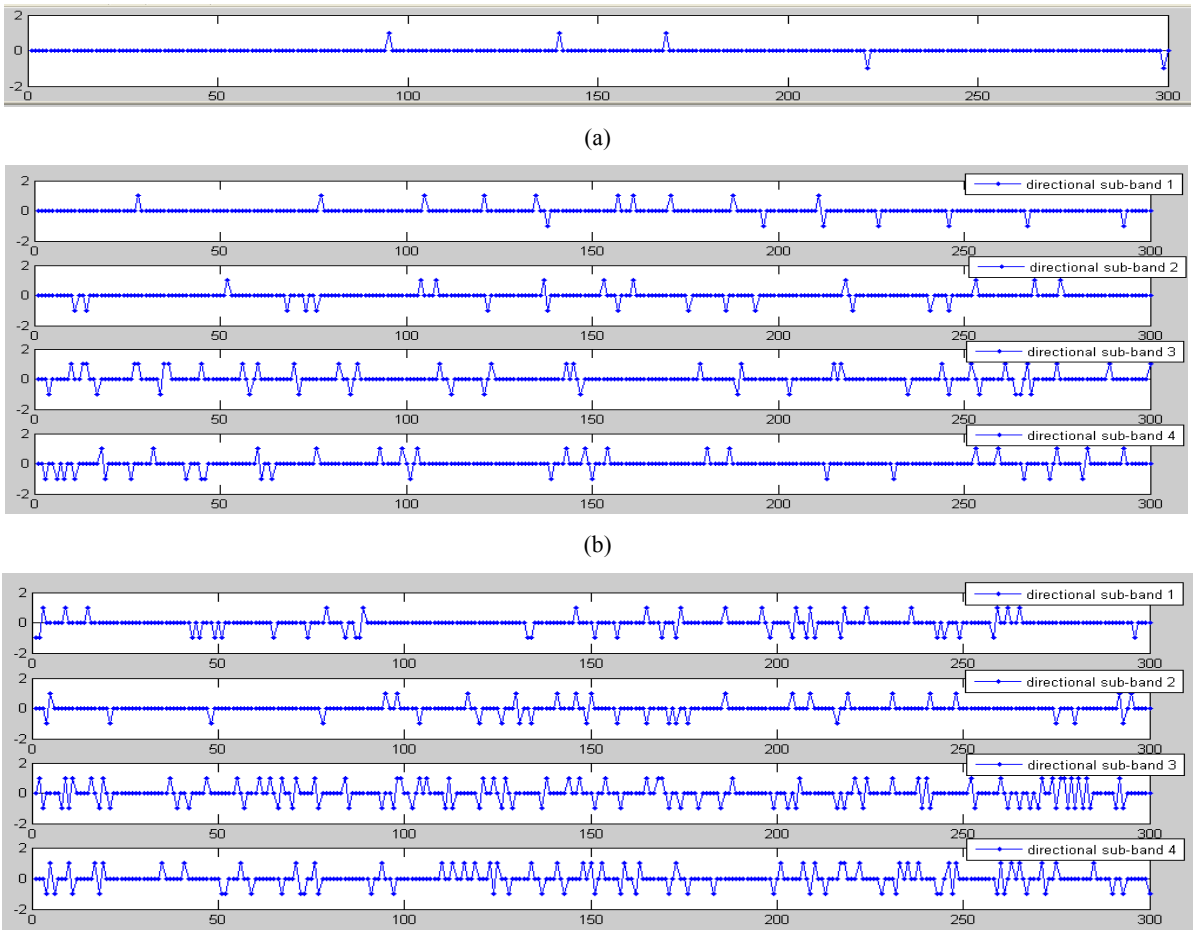
transform, high-frequency information of each layer can be divided into 2^n (n is non-zero positive integer) directions, which is the multi-directional character of Contourlet transform. As a result, with Contourlet transform, we can conduct more detailed image analysis and processing, so as to achieve better processing results.

Contourlet representation of Lena image is given in Fig. (**1**). Two layers of LP (Laplacian pyramid) decomposition is conducted on 512×512 Lena image, where each high-frequency image is divided into 4 directions of sub-bands by DFB (directional filter bank). The Contourlet coefficients can be expressed as $C_{s,d}(i,j)$, where s and d respectively refer to the scale and direction of Contourlet sub-band, $s \in \{0,1,2\cdots S\}$, S is the minimum resolution order, $d \in \{1,2,3,4\}$. $C_{s,d}(i,j)$ is Contourlet coefficient of original image in s-level resolution and d direction sub-band.

2.2. Impact of JPEG Compression on the Relationship Between Contourlet Coefficients of Image

The relationship of Contourlet coefficients can be used to generate the authentication watermark. To prove the feasibility of this approach, characteristics of the image Contourlet coefficients before and after JPEG will be explored as follows.

Take Lena image as an example to conduct Contourlet transform, where "9-7" biorthogonal filter is applied for LP decomposition, because the linear phase and approximation orthogonal characteristics are more suitable for image processing; "pkva" filter is applied in directional filter. Two layers are done in LP decomposition and each high frequency sub-band is divided into 4 directional sub-bands by DFB. Meanwhile, the same Contourlet transform is conducted for JPEG compressed image (QF=90), and respectively observe the changes of coefficients in low-frequency, sub-high



(a)

(b)

(c)

Fig. (2). The change in relationship of the adjacent Contourlet coefficients under JPEG: (**a**) Low Frequency Sub-band; (**b**) 4 Directional Sub-bands of the Sub-high Resolution; (**c**) 4 Directional Sub-bands of the Highest Resolution.

Table 1. The change in relationship of the adjacent Contourlet coefficients of Lena under JPEG (QF=90).

low frequency	sub-high resolution				the highest resolution			
	$(C_{2,1})$	$(C_{2,2})$	$(C_{2,3})$	$(C_{2,4})$	$(C_{1,1})$	$(C_{1,2})$	$(C_{1,3})$	$(C_{1,4})$
0.9812	0.9452	0.9267	0.9049	0.9102	0.7922	0.7666	0.7353	0.7547

resolution and the highest resolution sub-bands after compression. For each sub-band before and after JPEG compression, 1×2 non-overlapping window is used to slide and intercept coefficients, so as to calculate the difference of two coefficients, denoted as d. If the difference $d>0$ (or $d<0$) at the same position before and after compression, then the original is considered to have not changed after JPEG compression. If $d>0$ (or $d<0$) of the original image while $d<0$ ($d>0$) of the compressed image, then the relationship is considered as have changed.

Fig. (**2**) shows the relationship changes of Contourlet coefficients in each sub-band before and after JPEG compression. For ease of observation, the figure shows only the changes in former 300 coefficients in each sub-band. Value "0" means that there are no changes in the relationships of coefficient pairs after compression, while "1" means there are changes after compression.

Table **1** gives the changes in various sub-band coefficients of Lena image after JPEG compression (compression quality factor QF = 90). The data in Table **1** is the ratio between the unchanged coefficient pairs without changing and the total coefficient pairs. The larger ratio means that the impact of JPEG compression on the coefficient of sub-band is smaller. As can be seen from Fig. (**2**) and Table **1**, the majority relationships of adjacent coefficients are unchanged before and after JPEG compression of low frequency and sub-high resolution sub-bands, while the relationships are seriously affected by JPEG compression in the highest resolution sub-bands.

Table 2. The change in relationship of the adjacent Contourlet coefficients in the sub-high resolution and low frequency sub-bands under JPEG compression (QF=90).

Q / Image	Q=90			Q=70			Q=50			Q=30		
	C_0	$C_{1,1}$	$C_{1,2}$	C_0	$C_{1,1}$	$C_{1,2}$	C_0	$C_{1,1}$	$C_{1,2}$	C_0	$C_{1,1}$	$C_{1,2}$
Pepper	0.9812	0.9452	0.9267	0.9609	0.9245	0.8994	0.9471	0.8911	0.8575	0.9177	0.8568	0.8389
Cammerman	0.9776	0.9402	0.9304	0.9543	0.9233	0.9001	0.946	0.8992	0.8564	0.9012	0.8499	0.8465
Girl	0.9855	0.9545	0.9373	0.9572	0.9267	0.9014	0.9488	0.9003	0.8614	0.9205	0.8734	0.8452
Couple	0.9731	0.9679	0.9456	0.9426	0.9318	0.8996	0.9313	0.9013	0.8551	0.8962	0.8727	0.8226
Boat	0.9666	0.9372	0.9269	0.9410	0.9289	0.8954	0.9214	0.8978	0.842	0.8852	0.8468	0.8379

To further demonstrate the characteristics, tests are conducted for five standard images: Peppers, Girl, Cameraman, Couple and Boat. The relationships of low frequency and 1, 2 direction sub-bands of sub-high resolution are observed and the data is shown in Table **2**.

As can be seen from Table **2**, most of the relationships of adjacent Contourlet coefficients in low-frequency and 1, 2 direction sub-bands of sub-high resolution have no change after compression, while the proportion of unchanged coefficients decreases with the increase of JPEG compression strength. The algorithm in this paper will use this feature to generate authentication information.

3. DESCRIPTION OF SEMI-FRAGILE WATERMARKING ALGORITHM

3.1. Watermark Embedding

Taking into account the visual characteristics of watermarked image and demand for tampering positioning, the low frequency and 1, 2 direction sub-bands of sub-high resolution are used to generate the watermark. Authentication watermark is embedded into the direct current component of image DCT coefficients. Suppose the size of original image X is m×n, and assume that m and n are both the integral multiple of 4.

(1) Generate block mappings. Pseudo-random sequence $S=\{s_i \mid i=1,2,\mathrm{L}\ N\}$ is generated with the secret key, where N=m×n/(4×4), is the number of blocks. An ordered index sequence $(t_1,t_2,\mathrm{L}\ ,t_N)$ is obtained, satisfying $s_{t_1} \le s_{t_2} \le \mathrm{L} \le s_{t_N}$, by sorting the random sequence S. For each block X_i, let the index of its mapping block be $f(i)=t_i$.

(2) Generate authentication information. "9-7" biorthogonal filter and "pkva" directional filter are applied to conduct two-layer decomposition for the original image, and each layer component is decomposed by 4 directional sub-bands. Using low frequency coefficient and sub-high resolution sub-bands $C_{2,1}$ and $C_{2,2}$, the authentication information is generated with the same method that in section 2.2 is applied to generate, and the former N bits are taken as the watermark $W=\{w_i \mid i=1,2\mathrm{L}\ N\}$ to be embedded.

(3) Select position for embedding. Divide the original image X into non-overlapping 4×4 image blocks and number blocks from top to bottom and from left to right: $X=\{X_i \mid i=1,2\cdots N\}$, conduct DCT on image block X_i, $\boldsymbol{D}_i = \mathrm{DCT2}(\boldsymbol{X}_i)$, and number the direct current (DC) coefficients corresponding to the blocks and generate a DC coefficient vector $Q=\{q_i \mid i=1,2\cdots N\}$.

(4) Watermark embedding. Each watermark w_i is embedded into the direct current coefficient $q_{f(i)}$ of image block $X_{f(i)}$ according to the block mapping generated in step(1). The watermark embedding processing is described by the equation as:

$$\beta_i = \mathrm{mod}(\lfloor q_{f(i)} \rfloor, 2^4) - \mathrm{mod}(\lfloor q_{f(i)} \rfloor, 2^3) \quad (1)$$

$$q_{f(i)} = q_{f(i)} - \beta_i + 2^3 \times w_i \quad (2)$$

3.2. Watermark Extraction

Watermark extraction is an inverse process of embedding. Assume that image Y^* is the image to be measured, and divide Y^* into non-overlapping 4×4 image block $Y^*=\{Y_i^* \mid i=1,2\cdots N\}$. Conduct DCT on image block Y_i^*, $D^*_i = \mathrm{DCT2}(Y_i^*)$, number the direct current coefficients according to the corresponding blocks and generate a direct current coefficient vector $Q_i^*=\{q_i^* \mid i=1,2\cdots N\}$. Watermark $W^*=\{w_i^* \mid i=1,2\mathrm{L}\ N\}$ is extracted by:

$$w_i^* = (\mathrm{mod}(\lfloor q_i \rfloor, 2^4) - \mathrm{mod}(\lfloor q_i \rfloor, 2^3))/2^3 \quad (3)$$

3.3. Tampering Detection

Tamper detection is to determine the authenticity of image, give instructions about whether the image has been tampered and can localize the tampered area. The processing of tampering detection is described as:

(1) Generate the block mapping. According to method in section 3.1, take the same key to generate the block mapping.

(2) Reconstruct the authentication watermark. Conduct the same Contourlet transform on image Y^* and reconstruct authentication information according to step (2) in section 3.1, then take the former N bits as the reconstructed authentication watermark $W' = \{w_i' | i = 1,2\text{L } N\}$.

(3) Extract watermark w_i^* from each 4×4 image block Y_i^*.

(4) Tampering judgment. Compare each reconstructed authentication watermark w_i' with the extracted watermark $w_{f(i)}^*$ from image block $Y_{f(i)}^*$, and define the tampering identification matrix $D = \{d_i | i = 1,2\text{L } N\}$,

$$d_{f(i)} = \begin{cases} 0 & w_{f(i)}^* = w_i' \\ 1 & \text{otherwise} \end{cases} \quad (4)$$

$d_{f(i)} = 1$ means that the watermark extracted from image block $Y_{f(i)}^*$ is inconsistent with the reconstructed watermark, then the image block $Y_{f(i)}^*$ is called invalid block, and $d_{f(i)} = 0$ means that $Y_{f(i)}^*$ is valid block.

There will still be a few non-zero points in the tampering identification matrix D, even the watermarked image is not tampered because the watermark embedding will affect the relationship among high-frequency sub-bands coefficients. The watermark is embedded into other blocks with a certain random relation; therefore, the non-zero points in tampering identification matrix are bound to be in random distribution. When the image is subject to JPEG compression or slight noise and other occasional attacks, the image tampering is randomly distributed, so the non-zero points in tampering identification matrix will present random distribution as well. When regional tampering occurs in the image, some non-zero points will be centered in the tampered region, while the randomly distributed non-zero points will exist as well. Definitions of sparse points and dense points in Ref. 6 are taken for statistics of the total number S_{sparse} of sparse points and total number S_{dense} in tampering identification matrix D, define $\delta = S_{sparse} / S_{dense}$, and select the threshold T, if $\delta \geq T$, then it is considered accidental attack. On the contrary, it is considered to be malicious tampering.

(5) Tampering localization. For every image block Y_i^* which is marked valid or invalid, the3×3 block-neighborhood of the central block Y_i^* can be denoted by the compass directions [10]:

NW N NE

W Y_i^* E

SW S SE

If the image block Y_i^* is invalid and all blocks of its 3 × 3 neighborhood are marked as valid, then mark the block Y_i^* valid.

If the image block Y_i^* is invalid, check the following four triples of neighboring block situation: (N, NE, E), (E, SE, S), (S, SW, W), (W, NW, N). If all blocks in any of the four triples are marked invalid, mark the block Y_i^* invalid.

For every block Y_i^* mark it valid if there are three or less invalid blocks in its 3 × 3 neighborhood, otherwise, mark the block invalid.

4. PERFORMANCE ANALYSIS

4.1. Performance of Invisibility

Invisibility is one of the important measures for watermarking performance. The Peak Signal to Noise Ratio (PSNR) is defined to measure the differences of watermark image and original image.

$$PSNR = 10\log 10\left[\frac{255*255}{\frac{1}{m*n}\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j)-I^w(i,j)]^2}\right] \quad (5)$$

Where I is the original image, I^w is the watermarked image of which size is m×n. The worst case after embedding is that every second lowest bit in watermarked images is different from that in original image, namely $\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j)-I^w(i,j)]^2 = 4\times m\times n$. Then calculate PSNR =42.11 dB, which means the PSNR of watermarked image and original image must be greater than 42.11 dB. Because every bit is independent, and the probabilities of $\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j)-I^w(i,j)]^2=0$ and $\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j)-I^w(i,j)]^2=2$ are both 0.5, namely $E([I(i,j)-I^w(i,j)]^2)=2$, and then the mathematical expectation of PSNR is E(PSNR)=45.12 dB. The original image Lena and watermarked image are given in Fig. (**3**), and the PSNR is 45.18 dB. It is basically consistent with the theoretical value, and higher than the PSNR 43.42 dB in Ref.6, indicating that the algorithm can better satisfy the invisible requirements of watermarking.

4.2 Performance of Resisting JPEG Compression

Semi-fragile watermarking algorithm should be able to ensure the precise positioning of malicious tampering, while the occasional attacks like JPEG compression and so on have certain robustness. To test the ability of resisting JPEG compression, Lena image is taken as an example to conduct JPEG compression on watermarked images respectively with compression quality factor (QF) 90, 80, 70, 60, 50, 40 and 30. Calculate the tampering identification matrix D and collect statistics for the number of dense points and sparse points, calculate value δ, and compare with the threshold T (T is 0.5 in this paper) so as to determine the type of tampering and authentication results.

(**a**) (**b**)

Fig. (3). Original image and watermarked image: (**a**) original image; (**b**) watermarked image.

Data in Table **3** refers to δ data under different compression quality factors. In the Table **3**, Y indicates that the authentication result is in accidental tampering which can pass authentication, while N means the authentication result is malicious attack. Compared with data in Ref. 6, the value of δ in this paper is lower, which will not be able to pass authentication when QF = 60. The main reason is that tampering identification matrix D in Ref.6 is obtained from the difference between extracted watermark and original watermark, while a self-embedding watermarking algorithm is applied in this paper in which tampering identification matrix D is obtained from extracted watermark and reconstructed watermark, so the embedding processing will affect the quality of reconstructed watermark in a certain degree, thus resulting in less precise matrix D. However, the biggest advantage of self-embedding watermarking algorithm is that original image and original watermark are not required in authentication.

Table 3. The values of δ under different compression quality factors.

Quality factor （QF）	90	80	70	60	50
δ (This paper)	1.8990 Y	0.9452 Y	0.7324 Y	0.4201 N	0.2053 N
δ (Ref.6)	2.0652 Y	1.6330 Y	1.0041 Y	0.5656 Y	0.4295 N

Fig. (**4**) shows the results of tampering positioning under different compression quality factors. The black region indicates that the image has not been tampered, while the tampered regions are marked white. As a result, when QF = 90, there are no misjudgment which means that the affection of JEPG compression can be negligible; when QF is 80 and 70, there are several misjudged blocks in the positioning, but the impact of compression result is not significant; when QF is 60, there are more misjudged blocks existed in the positioning images which are not able to pass the tamper detection. But we know, at a lower compression quality factor, the image has lost more content information, and greater changes have occurred in visual quality of image. At this time, the JPEG compression attack test has lost its meaning. Experimental results show that the algorithm has certain robustness for JPEG compression.

4.3. Recognition for Partial Malicious Tampering

When malicious tampering occurs in part of the image, non-zero points concentrated in the tampered region and will be randomly distributed in intact region. Tampering positioning result of image under malicious attack is given in Fig. (**5**), where (a) is the tampering result after cutting the face region of Fig. (**3b**), (**c**) is the tampering results of collage attacks. (**b**) and (**d**) respectively are the tampering positioning results of (**a**) and (**c**). As can be seen, in the tampering results, apart from some misjudged blocks existing on the edge of tampered parts, the tamper positioning is more accurate.

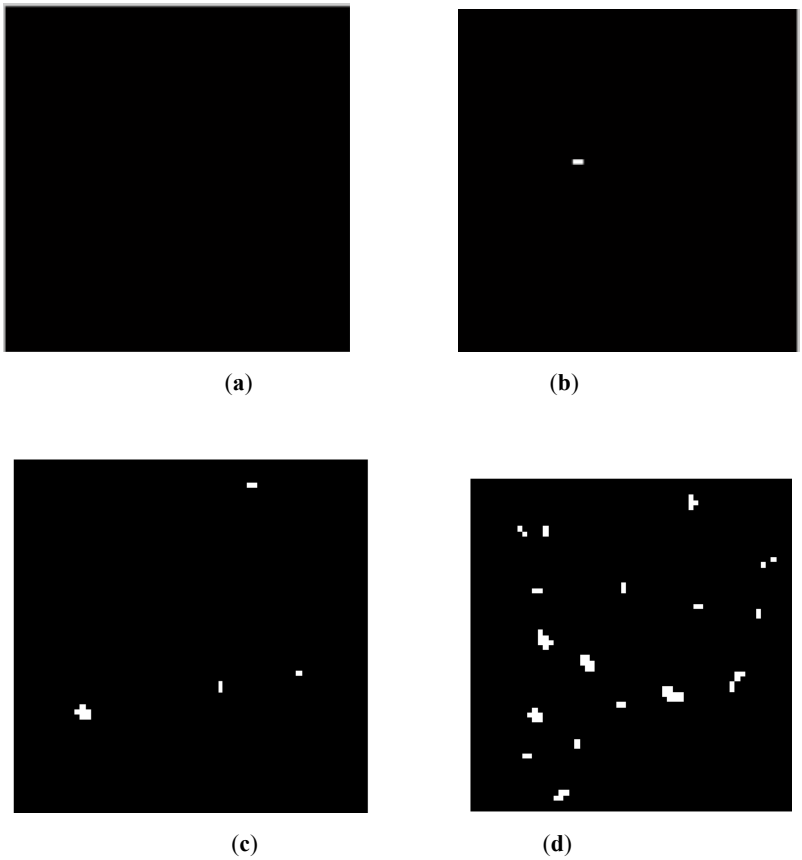
4.4. Comparison with Other Algorithms

In order to verify the performance of the image authentication algorithm, the comparison of experiment results with [6, 8] is shown in Table **4**, where results are all obtained when 512×512 Lena is the carrier image.

Table 4. Performance comparison with other algorithms.

	Size of original image	Precision of localization	PSNR （dB）	QF of resisting to JPEG	Blind authentication
This paper	512*512	exactly	45.18	70	Y
Ref.6	512*512	exactly	43.42	60	N
Ref.8	512*512	exactly	36.70	70	Y

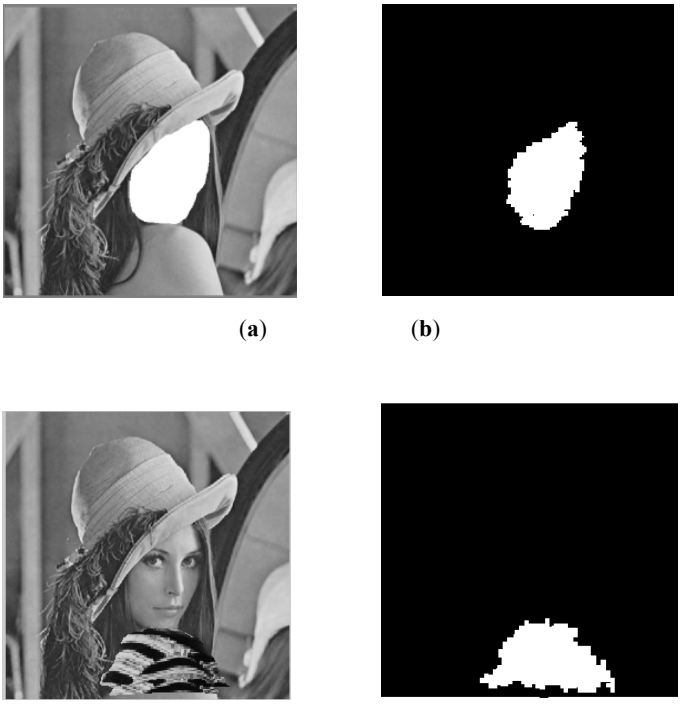
Authentication watermark is generated according to the relationship of adjacent Contourlet transform coefficients before and after JPEG compression. The PSNR of watermarked image is higher than [6, 8]. Although the performance of anti-JPEG compression is slightly higher than that in this paper, the shortcoming in [6] is that authentication needs original watermark and cannot achieve blind



(**a**) (**b**)

(**c**) (**d**)

Fig. (4). Tampering localization under different compression quality factors: (**a**) QF=90; (**b**) QF=80; (**c**) QF=70; (**d**) QF=60



(**a**) (**b**)

(**c**) (**d**)

Fig. (5). The tampering localization results for malicious local modification: (**a**) Cropping Attack; (**b**) Tampering Localization for Cropping Attack; (**c**) Collage Attack; (**d**) Tampering Localization for Collage Attack

detection because watermark is irrelevant with the image content. The algorithm in Ref. 8 can detect blindly which has the same performance of resisting JPEG compression as our algorithm. However, the impact of embedding on image visual is great in [8]. On the whole, the algorithm in this paper has obvious advantages.

CONCLUSION

The robustness of adjacent Contourlet transform coefficients on the JPEG compression is analyzed in this paper, where authentication watermark is generated on this basis. According to pseudo-random sequence, watermark is embedded into DCT coefficients of image blocks. The algorithm can better ensure the watermark invisibility and distinguish accidental tampering from malicious attacks with good anti-JPEG compression performance. Also, it can conduct more accurate positioning on malicious tampering.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

This work is supported by the National Natural Science Foundation of China (No. 61303087) and the Shandong Provincial Natural Science Foundation of China (No. ZR2014FM016).

REFERENCES

[1] D. Kundur, and D. Hatzinakos, "Digital watermarking for telltale tamper proofing and authentication," *Proceeding of the IEEE*., vol. 87, no. 7, pp. 1167-1180, 1999.

[2] Y. Hao, L. Chuang, and Q. Feng, "A survey of digital watermarking," *Journal of Computer Research and Development*, vol. 42, no. 7, pp. 1093-1099, 2005.

[3] M. Yeung, and F. Mintzer, "Invisible watermarking for image verification," *Journal of Electronic Imaging*, vol. 7, no. 3, pp. 578-591, 1998.

[4] L. Ching-Yung, "A robust image authentication method distinguishing JPEG compression from Malicious Manipulation," *IEEE Tran on circuits and system of video technology*, vol. 11, no. 2, pp. 153-168, 2001.

[5] X. Wu, X. Liang, H. Liu, J. Huang, and G. Qiu, "Reversible semi-fragile image authentication using zemike moments and integer wavelet transform," *Digital Rights Management: Technologies, Issues, Challenges and Systems*, vol. 3919, pp. 135-145, 2006.

[6] L. Chun, and Huang Jiwu, "A semi-fragile image watermarking resisting to JPEG," *Journal of Software,* vol. 17, no. 2, pp. 315-324, 2006.

[7] C. Fan, H. Huang, and W. Hsu, "A robust watermarking technique resistant JPEG compression," *Journal of Information Science and Engineering*, vol. 27, pp. 163-180, 2011.

[8] W. Jinshen, D. Yuewei, and W. Zhiquan, "JPEG image authentication based on content feature watermarking," *Journal of Image and Graphics*, vol. 13, no. 7, pp. 1265-1271, 2008.

[9] L. Yingjiang, S. Jianhong, and L. Chunyan, "Robust image watermarking algorithm based on contourlet transform and SVD decomposition," *Computer Science & Application,* vol. 2, no. 5, pp. 262-269, 2012.

[10] J. Zhang, Q. Zhang and H. Lv, " A novel image tamper localization and recovery algorithm based on watermarking technology," *International Journal for Light and Electron Optics*, vol. 124, no. 23, pp. 6367-6371, 2013.

Received: June 16, 2015 Revised: August 10, 2015 Accepted: September 19, 2015