Multiple and color images compression-encryption schemes with balanced qualities based on the multiple-order discrete fractional cosine transform

Qingmin Zhao (赵庆敏), Xianzhe Luo (罗賢哲), Nanrun Zhou (周南润), and Jianhua Wu (吴建华)*

Department of Electronic Information Engineering, Nanchang University, Nanchang 330031, China

**Corresponding author: jhwu@ncu.edu.cn*

Received October 12, 2011; accepted January 18, 2012; posted online May 9, 2012

By making use of the discrete fractional cosine transform, spectrum cutting and combining, rate-distortion control and color space transform, a joint compression and encryption scheme for multiple images and color images is proposed based on the multiple-order discrete fractional cosine transform (MODFrCT). The spectra coefficients of images gotten by the discrete cosine transform are scanned in a way of zigzag, cut at an appropriate position, and combined into a single spectrum image sequentially encrypted by the MODFrCT. Rate-distortion control is utilized during the spectrum cutting to balance the qualities of the multiple reconstructed images. A color image can be decomposed into Y, Cb, and Cr components prior to the encryption, and these three components are then encrypted in the same way as that for multiple images. The numerical simulations demonstrate the validity and efficiency of these schemes, and the robustness of the schemes against occlusion attack is examined.

OCIS codes: 100.0100, 350.1260.

doi: 10.3788/COL201210.S11006.

Image encryption is a way to transform the original images into the encrypted images and the unauthorized users cannot get the information of the original images from the encrypted images without the cipher keys. Over near two decades, many image encryption technologies based on the discrete fractional Fourier transform (DFrFT) have been presented[1−5], which regards the fractional order of the DFrFT as the cipher key.

To achieve greater efficiency of image encryption, the technology of multiple images encryption (MIE) becomes a focus[6−13]. Situ *et al.* firstly proposed an algorithm of MIE by using wavelength multiplexing[6], where the original images are directly added into one image. The quality of the decrypted image in this scheme is not so good due to the cross-talk effect. Liu *et al.* presented a double image encryption scheme based on the phase retrieval algorithm and fractional Fourier transform[7]. In Ref. [7], two images can be simultaneously encrypted into a single one as the amplitudes of fractional Fourier transform with different orders and this method can be extended to multiple-image encryption, but how to select the proper phase is a time-consuming work. Ran Tao *et al.* proposed an algorithm which encodes two original images into the amplitude and phase and combines them into one image. However, this method has the drawbacks in optical implementations since the decrypted phase image cannot be easily manipulated[8]. A new kind of double image encryption by use of cutting spectrum in the FrFT domain is used by Liu *et al.*[11], in which the part spectra of two images are scrabbled up a new spectrum that is encrypted with double random phase encoding technique. This method can encrypt multiple images simultaneously by cutting central spectrum and combining them, but the quality of decrypted images will be lowered, for the higher frequency coefficients are lost in the process of spectrum cutting. In addition, other researchers present some multiple-image encryption methods based on spectrum truncating and combing[12,13]. In these multi-image encryption methods, different reconstructed images will have different qualities, for the multiple images have different image details. Color images are widely used in modern society and several color image encryption methods are also proposed recently[14−19]. Joshi *et al.* proposed a technique for four color images encryption, where the four input RGB images are firstly converted into their indexed image formats and subsequently multiplexed into a single image through elementary mathematical steps prior to the encryption[14]. The information of original images can be recovered only under the help of color map, and it increases the transmission load. The encryption method of color image based on RGB mode is proposed by Ge *et al.*[19], and color image hiding and encryption with wavelength multiplexing is proposed by embedding and encryption in R, G, and B three channels. This scheme is not convenient in the practical application, for three channels must be used in the transmission process.

In order to meet the real-time requirement of image data transmission, we propose a general mode of multi-image encryption based on the multiple-order discrete fractional cosine transform (MODFrCT) in this letter, which is a reality-preserving transform and still retains many characteristics of DFrFT. The technology of image compression is introduced into this encryption mode, for single encrypted image must contain the important information of all the original images. The energy of an image concentrates in the low frequency part of the discrete cosine transform (DCT) domain[20] and the zigzag scanning that is used in JPEG is introduced to facilitate the spectrum cutting of the multi-image encryption. Quality balance is used for multiple images during the spectrum cutting and combining operations. The number of the fractional orders in the MODFrCT of both encryption and decryption equals the sum of image columns

and rows. Based on the theory that the Cb and the Cr components can be down-sampled without visual quality loss, we propose a color image encryption scheme in the YCbCr space. The color image to be encrypted in RGB format is firstly transformed into the YCbCr space. Both the Cb and the Cr components are decimated in column (DIC), and these two images after decimation are scrabbled up a combined image Cbr, and the Y and Cbr components are then encrypted in the same way as that for multiple images when the number of images is 2.

The discrete fractional cosine transform (DFrCT) is a generalization of the discrete cosine transform (DCT). The DFrCT uses the eigen decomposition of the DCT-I kernel that is written as[21]

$$\sqrt{\frac{2}{N-1}}K_m K_n \cos\left(\frac{mn\pi}{N-1}\right), \tag{1}$$

for $m, n = 0, 1, \cdots, N-1$, where K_m is defined as

$$K_m = \begin{cases} 1/\sqrt{2}, & m = 0, N-1 \\ 1, & \text{otherwise.} \end{cases}$$

The eigen decomposition of an $N \times N$DCT-I transform matrix $\mathbf{M}$ can be expressed by

$$\mathbf{M} = \mathbf{V}\mathbf{\Lambda}\mathbf{V} \tag{2}$$

where $\mathbf{V}_N = [\mathbf{v}_1|\mathbf{v}_2|\cdots|\mathbf{v}_N]$, $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_N)$.

With these eigen values and eigen vectors, the fractional cosine transform matrix $\mathbf{M}_a$ is constructed by

$$\mathbf{M}_a = \mathbf{V}\mathbf{\Lambda}_a\mathbf{V}. \tag{3}$$

The block-diagonal matrix $\mathbf{\Lambda}_a$ is computed by

$$\mathbf{\Lambda}_a = \begin{bmatrix} \mathbf{G}1_{2/N}[\theta(a)] & 0 \\ 0 & \mathbf{G}2_{2/N}[\eta(a)] \end{bmatrix}, \tag{4}$$

where $\mathbf{G}1_{2/N}$ and $\mathbf{G}2_{2/N}$ are block-diagonal Givens matrices whose blocks are

$$\mathbf{G}1_2[\theta(a)[= \begin{bmatrix} \cos[\theta(a)] & \sin[\theta(a)] \\ -\sin[\theta(a)] & \cos[\theta(a)] \end{bmatrix}.$$

$$\mathbf{G}2_2[\eta(a)] = \begin{bmatrix} \cos[\eta(a)] & \sin[\eta(a)] \\ -\sin[\eta(a)] & \cos[\eta(a)] \end{bmatrix}.$$

It can also be verified that the DFrCT has the mathematical properties of linearity, unitarily, additivity of rotations, periodicity, zero rotation and reality[21]. The property of reality is of much importance for image encryption. If $x(n)$ is real, its DFrCT is also real. The decryption process does not increase the storage or transmission load as in the case of non reality preserving transform.

For a given image $\mathbf{F}$ of size $M \times N$, the encryption and the decryption steps of the multiple-order DFrCT (MODFrCT)[22] are as follows:

1) Generate an order vector $\boldsymbol{p} = [p_0, p_1, p_2, \cdots, p_{M-1}]$ and an order vector $\boldsymbol{q} = [q_0, q_1, q_2, \cdots, q_{N-1}]$, whose members are uniformly distributed in the range of (0, 2).

2) Each row of image is transformed with one-dimensional (1D) DFrCT, with fractional order π for the ith row, the resulting image is $\mathbf{F}_1$; then each column of $\mathbf{F}_1$ is transformed with 1-dimensional DFrCT, with fractional orderq_j for the jth column, the resulting image is $\mathbf{F}_2$. $\mathbf{F}_2$ is regarded as the encrypted image, and $\boldsymbol{p}$, $\boldsymbol{q}$ are regarded as the row and the column cipher key vectors, respectively.

3) Each column of $\mathbf{F}_2$ is transformed by 1D DFrCT with the fractional order vector $\boldsymbol{q}' = 2 - \boldsymbol{q}$, and the resulting image is $\mathbf{F'}_1$. Then each row of $\mathbf{F'}_1$ is transformed with 1D DFrCT with the fractional order $\boldsymbol{p}' = 2 - \boldsymbol{p}$, and the resulting image $\mathbf{F}'$ is taken as the decrypted image.

To encrypt multiple images, image spectrum cutting operation is introduced into the proposed algorithm. The low frequency part of spectrum includes main information of an image, which is considered for describing and recording the image approximately[11]. The DCT followed by the zigzag scanning is used prior to the spectrum cutting operation to extract the low frequency part of the two-dimensional image spectrum, which is widely used for image compression, as in the well-known international standard JPEG. It is known that the DCT has a characteristic of energy concentration towards the left-up part in the DCT domain. In a DCT transformed block, the first coefficient is called the direct current (DC) coefficient while the rest coefficients are called the alternating current (AC) ones[20]. DCT concentrates most of the perceptually important information into DC and a few (low frequency) AC coefficients. In other words, the higher frequency coefficients contain relatively less visual information. After the zigzag scanning, the two-dimensional (2D) DCT coefficient matrix becomes a one-dimensional (1D) array in which the lower indexed coefficients mean the lower frequency. This facilitates the frequency extraction by cutting the 1D array at an appropriate position, as shown in Fig. 1.

Generally, the multiple images have different image details. If an equal space is assigned to each image, different reconstructed images will have different qualities, among which some are good enough and some are not. In order to balance the qualities of the recovered images as much as possible, the rate-distortion control is performed as follows:

1) perform the discrete cosine transform on m images and yield m 2D transform coefficient arrays.

2) Scan the m 2D transform coefficient arrays in the way of zigzag to form m 1D arrays.

DC coefficient AC coefficients

1	2	6	7	15	16	28	29
3	5	8	14	17	27	30	43
4	9	13	18	26	31	42	44
10	12	19	25	32	41	45	54
11	20	24	33	40	46	53	55
21	23	34	39	47	52	56	61
22	35	38	48	51	57	60	62
36	37	49	50	58	59	63	64

Fig. 1. Zigzag scanning order of a DCT coefficient matrix of 8×8 block.

3) Initially assign a size of $(M \times N)/m$ to each 1D array. Temporarily, the former $(M \times N)/m$ coefficients of each 1D array are retained, the others are cut away.
4) Compute the mean square errors (MSEs) in the DCT domain for these m images, which are equal to those computed in the spatial domain because the discrete cosine transform is the norm preserving transform.
5) Find the maximum MSE and the minimum MSE among m images; denote respectively their corresponding image numbers as $m_{\max}$ and $m_{\min}$.
6) Simultaneously increase the size for the 1D array of coefficients of image $m_{\max}$ by 1 and decrease the size for the 1D array of coefficients of image $m_{\min}$ by 1, this will decrease the MSE of image $m_{\max}$ and increase the MSE of image $m_{\min}$. Repeat until these two MSEs are equal or their relative values are just reversed.
7) Go to step (v) to find the new maximum and minimum MSEs. If the difference between the two new MSEs is less than a predefined value, the process of rate-distortion control stops.

For the encryption of m images, the images are firstly transformed with DCT and the resulting coefficients are processed with spectrum cutting and combining operations described previously, in order to generate one coefficient matrix. The scrambling operation as suggested in Ref. [23] is applied in order for the coefficients to be more disordered. Then this resulting matrix is encrypted with the MODFrCT and the resulting matrix is regarded as the encrypted image. The decryption process is a reverse version of the encryption process and these two processes are shown in Fig. 2.

We use two, three, four, five, six, seven, and eight images to show the performance of the multi-image encryption method. The following eight images Lena, Cameraman, Plane, Lake, Peppers, Baboon, Milkdrop, and House are chosen as test images shown in Figs. 3(a)–(h), respectively, to demonstrate the performance of this algorithm. Figure 3(i) shows the combined spectrum which consists of the eight images' spectra. Figure 3(j) shows the encrypted image of (i) with the MODFrCT, in which the row cipher key vector $\boldsymbol{p} = [0.4979, 0.9516, 0.7982, \cdots, 1.2896]$ and the column cipher key vector $\boldsymbol{q} = [1.7928, 0.9645, 0.0282, \cdots, 0.6162]$, whose elements have all been randomly chosen in the range of (0, 2). Figures 3(k)–(r) are the decrypted images from (j), using the correct decryption parameters $\boldsymbol{p}' = 2 - \boldsymbol{p}$ and $\boldsymbol{q}' = 2 - \boldsymbol{q}$. The technology of rate-distortion control is used in this simulation to balance the quality of decrypted images.

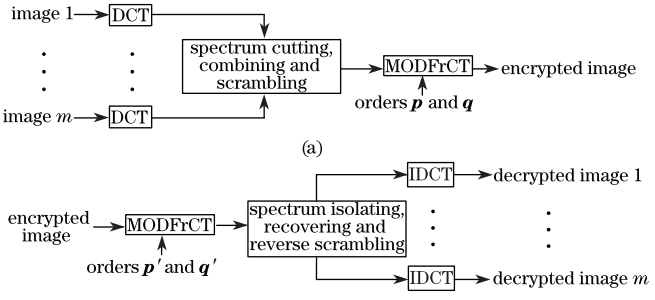

Fig. 2. Process of multi-image encryption: (a) encryption process; (b) decryption process.

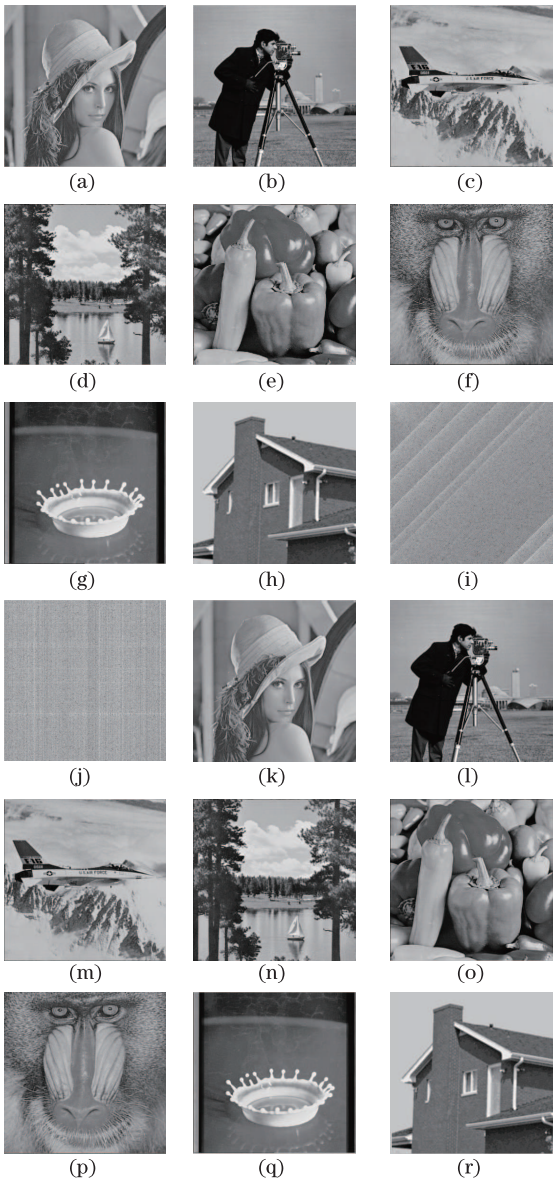

Fig. 3. Eight-image encryption and decryption results: (a)–(h) original images of 512×512 pixels; (i) the combined spectrum of eight original images; (j) encrypted image; (k)–(r) decrypted images.

To measure the performance of the encryption method, we calculate the peak signal-to-noise ratio (PSNR) of the decrypted images, mathematically, as

$$\mathrm{PSNR} = 10 \lg \frac{255 \times 255}{\mathrm{MSE}}, \tag{5}$$

where MSE is the mean square error between the decrypted image and the original one.

The PSNRs of the decrypted images in the case of multiple image encryption are shown in Table 1.

From Table 1, one can see that the qualities of the decrypted images are gradually degraded with the increase of number of the images encrypted. In fact, the more images are simultaneously encrypted, the more coefficients are cut out in order to piece up a single coefficient matrix with the remaining coefficients, as shown in Fig. 3. Consequently, the reconstructed image quality becomes worse. It is worthwhile to note that when six images are

Table 1. PSNR Values/Compression Rate (CR) of the Multiple Image Encryption System

Number of images	2	3	4	5	6	7	8
Lena	44.82	39.77	35.35	33.84	30.40	30.19	30.14
Cameraman	45.02	39.57	35.28	33.84	30.36	30.21	30.14
Plane		39.60	35.29	33.94	30.38	30.21	30.13
Lake			35.14	33.83	30.36	30.20	30.08
Peppers				33.83	30.37	30.22	30.08
Baboon					30.40	30.19	30.08
Milkdrop						30.23	30.11
House							30.08
CR	2:1	3:1	4:1	5:1	6:1	7:1	8:1

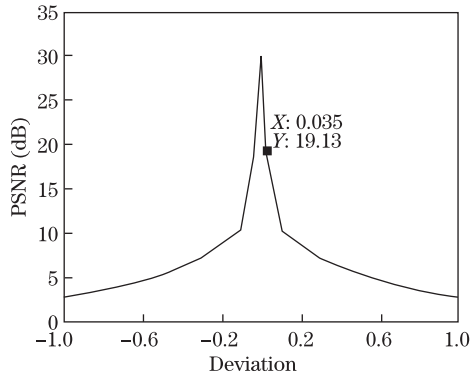

Fig. 4. PSNR versus the deviation of order parameters.

encrypted, the PSNRs decrease fast. This is mainly because the image Baboon has much more details than the others.

Take the eight-image encryption as an example, the average values of PSNRs for different fractional order parameters are graphically shown in Fig. 4. It can be seen from Fig. 4 that the average PSNR is up to about 30 dB when the fractional orders approach to correct key values and the images' quality is visually acceptable. However, when there is a deviation from the correct fractional order parameter, the images' quality decreases rapidly: down to 19.13 dB with a relative deviation of 0.035/2. This implies that the proposed algorithm possesses a high sensitivity versus the changes of decryption parameters.

A color image is usually stored in memory as a raster map, a 2D array of small integer triplets. For visually acceptable results, it is necessary (and almost sufficient) to provide three samples (color channels) for each pixel, which are interpreted as coordinates in some color space. Typical color mode is based on RGB space. A color image based on RGB space is the composition of R, G, and B components and each component can be seen as a gray image. The RGB space is very common in the field of color image processing, however other spaces such as YCbCr, HSI, etc., are often used in some contexts.

The RGB color space exits some deficiencies: the RGB space is a color display space, which is not suitable for human visual features[23]; there is a strong correlation among R, G, and B components. In YCbCr color mode, the Y component denotes the intensity, and the Cr and Cb components respectively denote the color differences of red and blue. YCbCr color mode is used to separate out an intensity Y that can be expressed with higher resolution and two chroma components Cb and Cr that can be bandwidth-reduced and hence subsampled. In the decryption, the full sized Cb and Cr components can be restored by interpolation, say, bilinear interpolation (BI).

The conversion from RGB to YCbCr color space is formulated as follows[24]:

$$\begin{bmatrix} \mathrm{Y} \\ \mathrm{Cb} \\ \mathrm{Cr} \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ -0.16875 & -0.33126 & 0.5 \\ 0.5 & -0.41869 & -0.08131 \end{bmatrix} \begin{bmatrix} \mathrm{R} \\ \mathrm{G} \\ \mathrm{B} \end{bmatrix}. \tag{6}$$

The inverse transformation is simply expressed by[25]

$$\begin{bmatrix} \mathrm{R} \\ \mathrm{G} \\ \mathrm{B} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.402 \\ 1 & -0.34413 & -0.71414 \\ 1 & 1.772 & 0 \end{bmatrix} \begin{bmatrix} \mathrm{Y} \\ \mathrm{Cb} \\ \mathrm{Cr} \end{bmatrix}. \tag{7}$$

The color image based on RGB space is firstly converted into YCbCr space. The Cb and the Cr components are then decimated by half in column respectively and the proportion among the Y, Cb, and Cr components is 4:2:2. Then these two components are scrabbled up a single component. Next, the Y component and the combined component are encrypted with the mentioned multiple image encryption mode when the number of images is 2. The process of encryption and decryption of color images are shown in Fig. 5.

The typical color image Lena of 512×512 is chosen to demonstrate the performance of this scheme. The encryption results are shown in Fig. 6. Here, the cipher vectors $\boldsymbol{p} = [0.3612, 0.0901, 1.4463, \cdots, 0.8612]$ and $\boldsymbol{q} = [1.9231, 1.5248, 0.0147, \cdots, 1.3517]$, whose elements were all chosen from realizations of a random variable uniformly distributed in (0, 2).

Any color image can be decomposed into three components and each component can be seen as a gray image. For a color image, the PSNR is calculated as

$$\mathrm{PSNR} = 10 \lg \frac{255 \times 255 \times 3}{\mathrm{MSE(R)} + \mathrm{MSE(G)} + \mathrm{MSE(B)}}, \tag{8}$$

where MSE is the mean square error between the decrypted component and the original one.

For the test image, after decryption, the PSNR for Fig. 13(f) is 35.30 dB; such a high value normally means that the visual quality is good enough and our color encryption scheme is feasible and effective.

In conclusion, we have proposed a method for encryption and decryption of multiple images and color image based on MODFrCT combined with image compression. By means of the DCT and zigzag scanning, one can easily truncate the coefficient array in DCT domain and realize image compression. The truncated coefficients in DCT domain are scrabbled up before the application of MODFrCT for encryption. There is not cross-talk interference among multiple images reconstructed and the quality of decrypted images is balanced as much as possible. The color image is processed based on the YCbCr color space, which is used in many image compression algorithms such as the JPEG 2000. The reconstructed image quality is strongly dependent on the number of images encrypted. Further work is suggested to optimize the image compre-

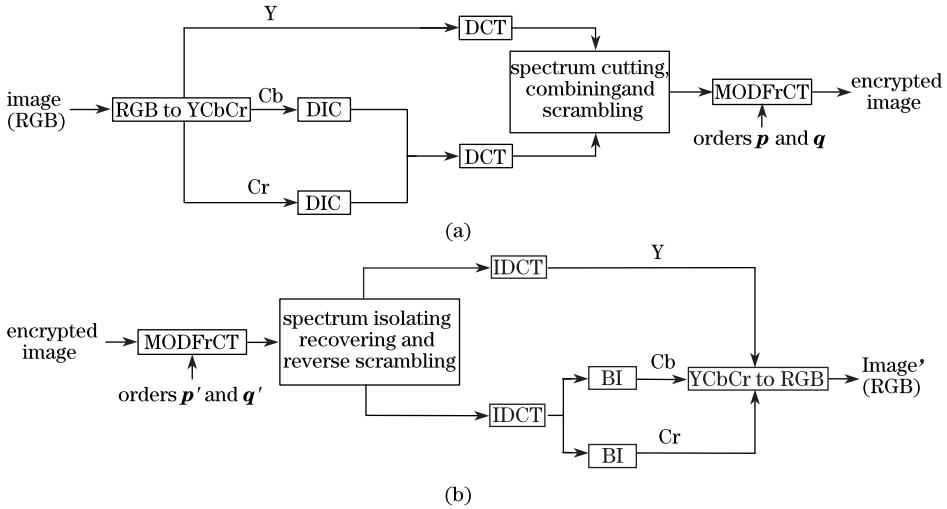

Fig. 5. Processes of encryption and decryption of color images: (a) encryption; (b) decryption.

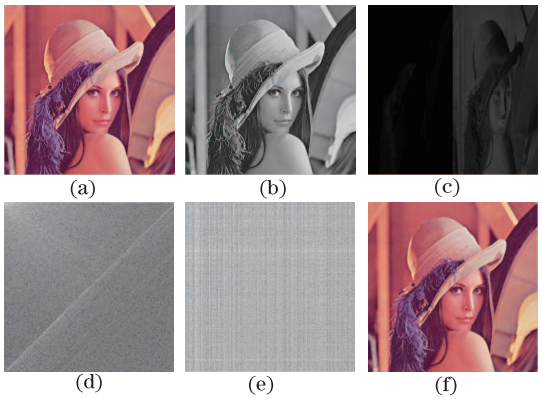

Fig. 6. Results of color image encryption: (a) original color image Lena; (b) the Y component; (c) the combined image contained the Cb and the Cr components after decimation; (d) the combined spectrum contained the low spectrum parts of (b) and (c); (e) encrypted image; (f) decrypted image.

ssion, for example, to apply entropy coding for the truncated coefficients.

This work was supported by the National Natural Science Foundation of China (No. 61141007) and the Natural Science Foundation of Jiangxi Province (Nos. 2009GQS0080 and 2010GZS0160). We would also thank the reviewers and the journal editors for their helpful comments.

References

1. P. Refregier and B. Javidi, Opt. Lett. **20,** 767 (1995).
2. B. M. Hennelly and J. T. Sheridan, Optik. **114,** 251 (2003).
3. L. Chen and D. Zhao, Opt. Commun. **282,** 3433 (2009).
4. M. Joshi, C. Shakher, and K. Singh, Opt. Eng. **46,** 522 (2008).
5. R. Tao, J. Lang, and Y. Wang, Opt. Lett. **33,** 581 (2008).
6. G. Situ and J. Zhang, Opt. Lett. **30,** 1306 (2005).
7. Z. Liu and S. Liu, Opt. Commun. **275,** 324 (2007).
8. R. Tao, Y. Xin, and Y. Wang, Opt. Express. **15,** 16067 (2007).
9. H. Li and Y.Wang, Opt. Commun. **281,** 5745 (2008).
10. X. F. Meng, L. Z. Cai, M. Z. He, G. Y. Dong, and X. X. Shen, Opt. Commun. **269,** 47 (2007).
11. Z. Liu, Q. Li, J. Dai, X. Sun, S. Liu, and M. A. Ahmad, Opt. Commun. **282,** 1536 (2009).
12. Z. Liu, Y. Zhang, H. Zhao, M. A. Ahmad, and S. Liu, Optik **122,** 1010 (2011).
13. A. Alfalou, A. Loussert, A. Alkholidi, and R. E Sawda, Future Generation Communication and Networking **2,** 590 (2007).
14. M. Joshi and K. Singh, Opt. Commun. **283,** 2496 (2010).
15. W. Wang, F. Liu, X. Ge, and Y. You, in *Proceedings of the 2nd IEEE International Conference on Information Management and Engineering* **15,** 271 (2010).
16. N. O. Abokhdair, A. B. A. Manaf, and M. Zamani, in *Proceedings of the International Conference on Digital Content, Multimedia Technology and its Application* **20,** 23 (2010).
17. H. Liu and X. Wang, Comput. Math. Appl. **59,** 3320 (2010).
18. P. Yin and L. Min, in *Proceedings of the International Conference on Intelligent Control and Information Processing* **433,** 447 (2010).
19. F. Ge, L. Chen, and D. Zhao, Opt. Commun. **281,** 4254 (2008).
20. S. Panchanathan and A. Jain, Comput. Commun. **19,** 1001 (1996).
21. I. Venturini and P. Duhamel, in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing* **5,** 205 (2004).
22. J. Wu, L. Zhang, and N. Zhou, Opt. Commun. **283,** 1720 (2010).
23. H. Li and Y. Wang, Opt. Laser Eng. **49,** 753 (2011).
24. M. Xie, J. Wu, L. Zhang, and C. Li, in *Proceedings of 2009 IEEE International Conference on Information and Automation* **138,** 143 (2009).
25. F. Douak, R. Benzid, and N. Benoudjit, AEU-International Journal of Electronics and Communications **65,** 16 (2011).