Single-Channel Color Image Encryption Using the Reality-Preserving Fractional Discrete Cosine Transform in YCbCr Space

Jianhua Wu, Fangfang Guo
Nanchang University/Department of Electronic Information Engineering, Nanchang 330031, China
Email: jhwu@ncu.edu.cn (JH Wu), hellosugger@qq.com (FF Guo)

Nanrun Zhou*
Nanchang University/Department of Electronic Information Engineering, Nanchang 330031, China
Beijing University of Posts and Telecommunications/Information Security Center, Beijing 100876, China
Email: nrzhou@ncu.edu.cn

Abstract—A novel single-channel color image encryption algorithm is proposed, which utilizes the reality-preserving fractional discrete cosine transform in YCbCr space. The color image to be encrypted is decomposed into Y, Cb, and Cr components, which are then separately transformed by Discrete Cosine Transform (DCT). The resulting three spectra sequences, obtained by zig-zag scanning the spectra matrices, are truncated and the lower frequency coefficients of the three components are scrambled up into a single matrix of the same size with the original color image. Then the obtained single matrix is encrypted by the fractional discrete cosine transform, which is a kind of encryption with secrecy of pixel value and pixel position simultaneously. The encrypted image is convenient for display, transmission and storage, thanks to the reality-preserving property of the fractional discrete cosine transform. Additionally, the proposed algorithm enlarges the key space by employing the generating sequence as an extra key in addition to the fractional orders. Simulation results and security analysis demonstrate the proposed algorithm is feasible, effective and secure. The robustness to noise attack is also guaranteed to some extent.

Index Terms—single-channel, color image encryption, reality-preserving fractional discrete cosine transform, generating sequence, spectrum truncation

I. INTRODUCTION

The reason of the use of color in image processing not only is that color is a powerful descriptor to provide beauty in vision, but also is that humans can discern thousands of color shades and intensities compared with about only two dozen shades of gray. Thus, color images contain more information than gray images do and are widely used in real life. Color image encryption has become a major task for information security since the issues about illegal data access on Internet are becoming more and more serious.

Past two decades we have witnessed the appearance of various encryption methods for gray images. Amongst, the most famous and widely used one is the double random phase encoding (DRPE) given by Refregier and Javidi[1], which applies two random phase masks arranged separately in the input and the Fourier planes to encrypt the image into a stationary white noise. The two random phase masks are uniformly distributed in the interval $[0, 2\pi]$ and the second one is taken as the main cipher key. Except the Fourier domain, other different domains such as fractional Fourier transform (FrFT) domain[2-5], Fresnel transform domain [6,7], Hartley transform domain[8,9] and Gyrator transform domain[10-12] are explored for more new encryption methods. However, the decrypted images resulting from optical cryptosystems would lose their color information, which makes these encryption algorithms inappropriate to encrypt color images. In response to this demand, many image encryption schemes especially for color images have been designed, where three components of color image are encrypted using the traditional gray image encryption methods separately[13]. In that case, it renders the cryptosystems sophisticated, since three channels must be involved. To address this problem, various single-channel color image encryption techniques have been put forward successively[14-17]. Zhou et al.[14] proposed a single-channel color image encryption algorithm based on chaotic scrambling and the FrFT in HSI space, the output of the encryption system is not a color image but a gray and a phase matrix. Wu et al.[16] made full use of the complex number mode to realize a single-channel color image encryption in fractional Fourier domain.

Although the above discussed encryption methods belong to single-channel, the encrypted images are complex-valued possessing amplitude information as well as phase information, which makes them inconvenient to...
display, transmit and store. In this paper, a novel single-channel color image algorithm based on the fractional discrete cosine transform (FrDCT)\cite{18}, which inherits the reality of the discrete cosine transform (DCT) matrix, is proposed. The original color image is converted into the YCbCr space, where the Y component denotes the brightness, and the Cb and the Cr components respectively denote the color differences of red and blue \cite{16}. Since human eyes are more attuned to brightness and less to color differences, hence the YCbCr color model allows more attention to be paid to the Y component, and less to the others. It is well known that the discrete cosine transform (DCT)\cite{9} has the property of energy concentration, namely, the energy of an image after DCT concentrates towards the top left corner — the low frequency, which human visual is more sensitive to. With the help of spectrum truncation, the low frequency spectra truncated from the corresponding cosine spectra in accordance with the ratio of 2:1:1 are scrambled up into one single matrix and sequentially encrypted by the FrDCT, which has a character of altering the pixel value and the pixel position simultaneously. The resulting cipher-text is a real gray-scale image, which is convenient for display, transmission and storage, and has camouflage property to some extent. Furthermore, generating sequence (GS), which results from the multiplicity of DCT matrices’ roots, is introduced as an extra cipher key. Spatiotemporal chaotic map\cite{20} is utilized to generate the random GS. Thus the high sensitiveness to initial values and system parameters inherent in any chaotic system provides high security naturally. Since the fractional orders are not so sensitive compared with the chaotic maps, they can be abandoned or be used merely as auxiliary keys. Simulation results and security analysis verify the effectiveness and feasibility of the algorithm. Robustness to noise attack is also validated.

The rest of this paper is organized as follows. Section II describes the theoretical background about the reality-preserving fractional discrete cosine transform. Section III gives the details of the proposed algorithm including the color model, spectrum truncation and the spatiotemporal chaotic map. The procedures of the proposed algorithm are also described in Section III. Simulations and discussions are given in Section IV. Finally, conclusion is drawn in final section followed.

II. THEORETICAL BACKGROUND

The fractional discrete cosine transform (FrDCT) is a generalization of the DCT. In current literatures, even though several versions of fractional cosine transform have been derived, the FrDCT\cite{18} different from those defined in [21,22] possesses the mathematical properties of reality in addition to linearity, unitarily and additivity. And the reality is of importance for image encryption, which ensures the outputs are real for real inputs.

The FrDCT is derived based on the eigen-decomposition and eigenvalue substitution of the DCT-II kernel denoted as:

$$C = \left[ \frac{1}{\sqrt{N}} e^{-j\pi \frac{(2n+1)k}{4N}} \right]$$

where \(n, k = 0, 1, \ldots, N-1\) and \(\epsilon_0 = 1, \epsilon_k = \sqrt{2}\) for nonzero \(k\).

The eigen-decomposition of an \(N \times N\) DCT-II matrix \(C\) is:

$$C = U \Lambda U^* = \sum \limits_n U_n e^{j\mu_n}$$

where \(U\) is a unitary matrix, composed of columns (eigenvectors) \(U_n\), \(U_n^* U_n = \delta_{mn}\), and \(\Lambda\) is the diagonal matrix with diagonal entries, i.e. eigenvalues \(\lambda_n\):

$$\lambda_n = e^{j\mu_n} \quad \text{with} \quad 0 < \varphi_n < \pi.$$  \(\mu_n\) represents the \(n\)th power \(\lambda_n = e^{j\mu_n}\), with its \(n\)th powers \(\lambda_n^a\), i.e., the matrix \(\Lambda\) by its \(n\)th power \(\Lambda^a\):

$$C^a = U \Lambda^a U^*$$

The matrix \(C^a\), given by (3) can be rewritten in an alternative form according to the eigenstructure of \(C\):

$$C^a = 2R \left( \sum \limits_n U_n \lambda_n^a \right) + V_1 \left( 1^a \right) + V_{-1} \left( -1^a \right)$$

where \(K = (N - \mu_1 - \mu_2)/2\), \(\mu_1\) and \(\mu_2\) represents the multiplicities of the eigenvalues 1 and -1, respectively. \(V_1\) collects the \(\mu_i\) matrices \(U_n\) corresponding to the eigenvalue 1 and similarly for \(V_{-1}\).

For \(N = 4N_o\) and real \(\alpha\), the FrDCT matrix \(C^a\) becomes a real-valued matrix because of the absence of the eigenvalues \(\pm 1\) and can be written as:

$$C^a = 2R \left( \sum \limits_n U_n \lambda_n^a \right) = 2R \left( \sum \limits_n e^{j\lambda_n^a n q_n} \right)$$

$$= \sum \limits_{n=1}^{N/2} \left( A_n \cos \omega_n + B_n \sin \omega_n \right) \quad \text{for} \quad n = 1, 2, \ldots, N/2$$

$$\omega_n = \varphi_n + 2\pi q_n \quad \text{where} \quad 0 < \varphi_n < \pi$$

where \(U_n = u_n u_n^*\), \(A_n = 2R \left( U_n \right)\), \(B_n = -2\Im \left( U_n \right)\) and \(q = \left( q_1, q_2, \ldots, q_{N/2} \right)\), introduced due to the multiplicity of the \(n\)th power of \(\lambda_n\) and called as generating sequence (GS) of the FrDCT, is an arbitrary sequence of integers depending on the nature of the fraction and has strong effect on the results of the FrDCT since different \(q\) leads to different \(C^a\), so by taking the GS \(q\) as secret key can provide a huge key space. Readers can refer to the [18] for more information about \(q\). The
expansion of the FrDCT for a two-dimensional signal is straightforward and simple through two FrDCTs successively by rows and by columns.

III. DESCRIPTION OF THE METHOD

A. Color Image Model

In essence, a color model is a specification of a coordinate system and a subspace where each color is represented by a single point. There are numerous color models in use today due to the fact that color science is a broad field that encompasses many areas of application. The most widely used is the RGB model, which is based on a Cartesian coordinate system. Images represented in the RGB color model consist of three component images, one for each primary color. Since the color image is not simply synthesized by three primary-color images to one image in practice, RGB model cannot be adapted well to the color understood by human [13]. In addition, the R, G and B components are equivalent and have strong correlation, which makes the changes in one component will affect the others. While the YCbCr model, where Y represents the brightness component, Cb and Cr components respectively represent the color difference of red and blue, has the ability to vary each component independently without affecting the others. Furthermore, human vision is more sensitive to Y component than to the other two. Thus, the key Y component can be encrypted with high-strength encryption methods, and the Cb, Cr components are subsidiary encryption components. The two models can be transformed with each other and the mathematical transform formulations are given by the following equations [16].

\[
\begin{bmatrix}
Y \\
Cb \\
Cr
\end{bmatrix} =
\begin{bmatrix}
0.299 & 0.587 & 0.144 \\
-0.16875 & -0.33126 & 0.5 \\
0.5 & -0.41869 & -0.08131
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

(7)

\[
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 1.402 \\
1 & -0.34413 & -0.71414 \\
1 & 1.772 & 0
\end{bmatrix}
\begin{bmatrix}
Y \\
Cb \\
Cr
\end{bmatrix}
\]

(8)

B. Spectrum Truncation

Since human vision is more sensitive to the low frequency, which contains most of the image information, the low spectrum is usually used to reconstruct the original image, thus to realize the image compression. The idea will be introduced into the proposed algorithm in the preprocessing stage. Because of the reality and energy concentration, the discrete cosine transform (DCT) is a more reasonable choice for the spectrum truncation. And since the energy of an image concentrates towards the left top corner in the DCT domain, the zigzag scanning is utilized prior to spectrum truncation to extract the low frequency of the 2-D image spectrum. The lower indexed coefficients of a 1-D array obtained from the 2-D DCT coefficient matrix through the zigzag scanning mean the lower frequencies. Then the 1-D array can be truncated at an appropriate position to realize the low frequency extraction. As described before, the Y component contains more image information, thus, the three corresponding spectra, of Y, Cb, and Cr, are truncated by a ratio of 1/2, 1/4, and 1/4, respectively, and scrambled up into a single combined spectrum for encryption. Thus a single combined spectrum for encryption can be realized.

C. Spatiotemporal Chaotic System

Spatiotemporal chaotic system maintains more complex behavior and more abundant characteristics, which makes it excellent candidate for encryption compared with the low-dimensional chaotic systems. A coupled map lattice with time delays (DCML) [20] consisting of the logistic map is adopted to construct the spatiotemporal chaotic system. Its mathematical representation is expressed as:

\[
x_{i,k+1} = \text{mod}((1-\varepsilon)g(x_i^k)+\varepsilon g(x_i^{k-1})+(1-\gamma)g(x_{i-1}^k)+\gamma g(x_{i-1}^{k-1}))
\]

(9)

where \( x_i^k \) represents the state variable for the \( i \)th site at time \( k \), \( i=1,2,\ldots,L \) ( \( L \) is the length of the DCML) is the lattice site index, \( \varepsilon \) and \( \gamma \) are the coupling coefficients ranging in \([0,1]\). The periodic boundary condition \( x_i^0 = x_i^L \) is assumed for any valid \( k \). \( \tau \) is the time delay and \( \tau = 5 \) in this paper. Given an initial sequences of length \( L \), a spatiotemporal chaotic matrix can be generated by (9). Logistic map is used as the nonlinear map \( g(x) \) to generate the initial sequences and given by:

\[
x_{k+1} = \mu \cdot x_k \cdot (1-x_k)
\]

(10)

with the system parameter \( \mu \in [3.5699456,4] \) and the initial value \( x_0 \in [0,1] \), the system exhibits chaotic state.

D. Encryption Process

The flowchart of the proposed color image encryption algorithm is illustrated in Fig. 1; the whole encryption procedure includes two stages: preprocessing and encryption. The specified encryption procedures for the original color image of size \( M \times N \times 3 \) are explained hereafter:

1) Preprocessing stage. The original color image \( I \) based on RGB space is firstly converted into YCbCr space, then a spectrum matrix \( I' \) of size \( M \times N \) is constituted from three low frequency parts, truncated by a ratio of 1/2, 1/4 and 1/4, respectively, from the Y, Cb and Cr components gotten by the discrete cosine transform.

2) Encryption stage. The interim result \( I' \) is transformed by the reality-preserving FrDCT, where two generating sequences are needed for the row and the column transform respectively. The procedures are described as follows:
Step 1: Iterate (9)-(10) to obtain two random sequences whose lengths respectively are $K + M/2$ and $K + N/2$, using the two initial values $x_0^1, x_0^2$ and the coupling coefficients $\varepsilon, \gamma$. The system parameter $\mu$ of the Logistic map is set to be 3.9999. Then discarding the previous $K$ values to avoid the harmful effect, we obtain two sequences $X_1 = \{x_i(n) | n = 1, 2, \ldots, M/2\}$, $X_2 = \{x_i(n) | n = 1, 2, \ldots, N/2\}$, $x_0^1, x_0^2$, the random coupling coefficients $\varepsilon, \gamma$ and the arbitrary integer $K$ are used as cipher keys.

Step 2: Generate the random GS $q_1$ for the rows with integers limited to 0 and 1 for the sake of brevity by defining a threshold function:

$$q_i(n) = \begin{cases} 0, & 0 < x_i(n) \leq 0.5 \\ 1, & 0.5 < x_i(n) < 1 \end{cases} \quad n = 1, 2, \ldots, M/2$$

Step 3: Generate the random GS $q_2$ for the columns in a similar way.

Step 4: Perform two 1-D FrDCTs with the fractional order $\alpha$ and GS $q_1$, the fractional order $\beta$ and GS $q_2$ for each row and each column of $I'$, respectively. $\alpha, \beta$ are the given fractional orders for the rows and the columns respectively. The obtained result is denoted as $C$.

The decryption process is similar to that of the encryption process but in the reversed order, and the fractional orders need to be modified as $\alpha' = -\alpha, \beta' = -\beta$.

IV. DIGITAL SIMULATION AND DISCUSSION

The typical color image Lena of size $512 \times 512$ and 3 8-bit R, G and B components is chosen as the plain-text.

A. Encryption and Decryption Simulations

Two fractional orders of the rows and the columns are fixed as 0.7689 and 0.4578, respectively. Two generating sequences are generated under the coefficients $\varepsilon=0.2, \gamma=0.92$; and $x_0^1=0.345678921, x_0^2=0.456789321$. The constant $K$ is set to be 1000. Fig. 2(b) shows the encrypted output, which is completely rough-and-tumble and does not reveal any information about the original image. Besides, the encrypted image is a single component rather than three components, which can bewilder others in a sense. The decrypted image with all correct keys is shown in Fig. 2(c).

For a color image, the peak signal-to-noise ratio (PSNR) is calculated as:

$$\text{PSNR}=10\log_{10}\left(\frac{255\times255\times3}{\text{MSE}_R+\text{MSE}_G+\text{MSE}_B}\right) \quad (12)$$

where the MSE represents the mean square error between the decrypted component and the corresponding original component. The PSNR is 37.0238 dB in this simulation, which declares a sufficiently good visual quality of the decrypted image and the feasibility of the proposed algorithm.

B. Key Sensitivity

The high sensitive to initial conditions and system parameters is inherent to any chaotic system. Fig. 3 shows the decrypted images using incorrect keys. As illustrated in Figs. 3(a)-(d), one cannot recognize the content of the decrypted image visually, even when the deviation of the initial values is so small as $10^{-16}$, which is mainly due to the high sensitivity to initial values and system parameters of chaotic maps. Moreover, Figs. 3(g) and (h) show that even if the parameter $K$ is 1 less or more than the correct value, the decrypted images still are completely incomprehensible and do not leak any information about the original image. Thus the cipher keys are highly sensitive in the proposed algorithm.
C. Key Space Analysis

A good cryptosystem should provide large key space to make any brute-force attack ineffective. The fractional orders are not used as cipher keys because of its low sensitivity and small key space. Thus our cryptosystem actually has the following secret keys: (1) $x_1^0$, $x_2^0$; (2) $\varepsilon$, $\gamma$; and (3) $K$. From Fig. 3, it is easy to know that these parameters maintain 13 and 16 digits after decimal point respectively. Thus, the key space comes to be about $10^{38}$. It is worth to mention that $K$ also has an effect on the quality of decrypted image. Therefore, a sufficiently large key space is ensured in the proposed algorithm.

D. Statistical Analysis

Two aspects are tested on the histograms and correlations of the original image and its cipher-text, demonstrating its superior confusion and diffusion properties which strongly resist statistical attacks. Figs. 4(a) and (b) show the histograms of the cipher-texts of the color image “Lena” and “Peppers” respectively under the same conditions. They are quite similar and the cipher-texts of different plain images have such similar histograms that illegal attackers cannot obtain any useful information with this statistical analysis.

The correlation coefficients of horizontally, vertically and diagonally adjacent pixels are calculated respectively. The formula is expressed as:

$$C_{x,y} = \frac{\sum_{i=1}^{l} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{l} (x_i - \bar{x})^2 (y_i - \bar{y})^2}}$$ (13)

where $x$, $y$ are the intensity levels of two adjacent pixels, and

$$\bar{x} = \frac{\sum_{i=1}^{l} x_i}{l}, \quad \bar{y} = \frac{\sum_{i=1}^{l} y_i}{l}$$

with $l$ the number of samples obtained from the image.

Table 1 indicates clearly that the correlations of adjacent pixels in the original and the encrypted images, which indicates that the correlation coefficient of the original image is significant while that of the encrypted image is very small. So the proposed algorithm reduces the correlation of the adjacent pixels in the original image.

<table>
<thead>
<tr>
<th>Correlation coefficients</th>
<th>Original image</th>
<th>Encrypted image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal $R$</td>
<td>0.9747</td>
<td>0.9689</td>
</tr>
<tr>
<td>Vertical $G$</td>
<td>0.9864</td>
<td>0.9656</td>
</tr>
<tr>
<td>Diagonal $B$</td>
<td>0.9698</td>
<td>0.9267</td>
</tr>
</tbody>
</table>

E. Robustness to Noise

To show the influence of interference in transmission, the robustness of the proposed method against noise is considered. The model of noise attack is defined as follows:

$$C' = C(1 + kG)$$ (14)

where $C$ and $C'$ represent the original color image component before and after adding noise, respectively. $k$ is the coefficient of the noise intensity of the Gaussian random noise $G$ whose mean value and standard deviation equal 0 and 1, respectively. Decrypted images with noise intensities 0.1, 0.5 and 1 are illustrated in Fig. 5. As observed, the decrypted images can still be visible despite of some noise interference, and even when the noise intensity increases to 1. Consequently, the proposed encryption algorithm can resist the noise attack to some extent.
We have represented a single-channel color image encryption by use of the reality-preserving fractional discrete cosine transform in YCbCr space, which is a kind of encryption with secrecy of pixel value and pixel position simultaneously. Unlike the RGB model, the YCbCr color model allows more attention to be paid on the Y component and less to the Cb and Cr components. Thus with spectrum truncation, a single combined cosine discrete cosine transform in YCbCr space, which is a encryption by use of the reality-preserving fractional encryption is robust to noise attack to some extent. Besides, the generating sequences determined by spatiotemporal chaotic map are introduced to enlarge the key space. The simulation results indicate that the proposed encryption algorithm is feasible and effective. Performance in noisy channel demonstrates the encryption is robust to noise attack to some extent.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (Grant Nos. 61262084 and 61141007), the Foundation for Young Scientists of Jiangxi Province (Grant No. 20122BAB201029) and the Natural Science Foundation of Jiangxi Province (Grant No. 20122BAB201029).

REFERENCE


Jianhua Wu, born in Jinxian County of Jiangxi Province, China, on September 9, 1956, was graduated from Harbin Institute of Technology in 1982 and got a Bachelor’s degree in information engineering. In 1985, he was graduated from South China University of Technology, Guangzhou, China and got a Master’s degree of science majored in communication and electronic systems. In 2005, he got the Ph.D. from the University of Poitiers, Poitiers, France majored in image and signal processing.

He is currently a professor with Department of Electronic Information Engineering, Nanchang University, China. He has published more than twenty papers in journals such as Optics Communications, Optics and Laser Technology, etc. His research interests include image and signal processing, image encryption, pattern recognition, etc.

Dr. Wu is a member of the IEEE.

Fangfang Guo was born in Suzhou, Anhui province, China on 13rd Oct. 1988. She received a Bachelor’s degree in communication engineering from Nanchang University in 2011. She is now in pursuit for the Master’s degree in communication and information systems in Nanchang University. Her research interests include image processing, image encryption and information security.