Electronic Research Archive

ERA, 31(7): 4102−4118.
DOI: 10.3934/era.2023209
Received: 22 March 2023
Revised: 16 April 2023
Accepted: 08 May 2023
Published: 26 May 2023

http://www.aimspress.com/journal/ERA

Research article

Color image steganalysis based on quaternion discrete cosine transform

Meng Xu[1], Xiangyang Luo[1,*], Jinwei Wang[2] and Hao Wang[3]

[1] State Key Laboratory of Mathematical Engineering and Advanced Computing, Henan 450001, China
[2] Nanjing University of Information Science and Technology, Department of Computer and Software, Nanjing 210044, China
[3] Nanjing University of Science and Technology, School of automation, Nanjing 210014, China

* **Correspondence:** Email: luoxy_ieu@sina.com.

Abstract: With the rapid development and application of Internet technology in recent years, the issue of information security has received more and more attention. Digital steganography is used as a means of secure communication to hide information by modifying the carrier. However, steganography can also be used for illegal acts, so it is of great significance to study steganalysis techniques. The steganalysis technology can be used to solve the illegal steganography problem of computer vision and engineering applications technology. Most of the images in the Internet are color images, and steganalysis for color images is a very critical problem in the field of steganalysis at this stage. Currently proposed algorithms for steganalysis of color images mainly rely on the manual design of steganographic features, and the steganographic features do not fully consider the internal connection between the three channels of color images. In recent years, advanced steganography techniques for color images have been proposed, which brings more serious challenges to color image steganalysis. Quaternions are a good tool to represent color images, and the transformation of quaternions can fully exploit the correlation among color image channels. In this paper, we propose a color image steganalysis algorithm based on quaternion discrete cosine transform, firstly, the image is represented by quaternion, then the quaternion discrete cosine transform is applied to it, and the coefficients obtained from the transformation are extracted to design features of the coeval matrix. The experimental results show that the proposed algorithm works better than the typical color image steganalysis algorithm.

Keywords: steganalysis; steganography; quaternion

1. Introduction

Digital media has now become an important carrier for the military, business and other fields as well as for individuals to obtain information and transmit it with the development of information technology. Since, at the same time, because digital communication in the Internet is vulnerable to malicious threats such as eavesdropping and malicious interference, people are more concerned than ever about security issues such as privacy protection and data integrity in the process of information transmission. The traditional solution is to use encryption technology to convert the information to be transmitted into cipher text for transmission. However, the drawback is that the ciphertext after encryption is usually disorganized and disorderly, which can easily make attackers notice the existence of secret communication and lead to the failure of information transmission. In other words, communication security not only means that the content of the transmitted message is secure, but also requires that the act of transmitting the secret message itself is also agnostic. Therefore, steganography, which is characterized by camouflage in the act of transmitting information, is receiving more and more attention. However, steganography is also a veritable double-edged sword that, while providing a reliable and secure means of Internet communication, may also facilitate organizations and individuals with malicious intent or improper purposes. The illegal or malicious use of digital steganography has already brought serious harm to national information security, business and personal privacy and property security. Under such circumstances, it has become an urgent need for the military and security departments of various countries to effectively monitor the use of steganography in real life, and to prevent or intercept the malicious or illegal use of steganography immediately. Therefore, steganalysis technology, which is a countermeasure to steganography, has emerged and received great attention from governments and research institutions. The research of steganalysis is of great importance for preventing the leakage of confidential information, combating terrorism and criminal activities, and maintaining Internet security.

Digital steganography embed additional secret information into digital media to achieve covert communication without unduly affecting the multimedia signal. Steganographic covert communication in steganography can be established by using text, audio, images, or video as a cover medium [1]. Image steganography techniques can be divided into two main categories: spatial domain steganography and transform (frequency) domain steganography. The first type hides secret information directly in the carrier image pixels, while the second type hides in the transform coefficients of the carrier image. Spatial domain steganography are widely used in steganography software, and many steganography tools use the least significant bit to hide secret messages, which is favored by researchers due to the high capacity and easy implementation of algorithms based on the lowest bit plane. The principle of spatial domain replacement technique is that any digital media information, when scanned and sampled, generates physical random noise to which the human sensory system is insensitive. Most spatial domain steganography algorithms are based on LSB (Least Significant Bit) [2−6]. The spatial domain algorithms have their own shortcomings in terms of security, because the secret message is mostly embedded in the insignificant part of the carrier.

The lossy compression of images, signal filtering processing, carrier storage format changes, etc., may remove the insignificant information of carriers, which affects the integrity and accuracy of ciphertext extraction. Therefore, scholars propose to embed information in the transform domain, and the idea of transform domain steganography is to hide information in some prominent areas of the carrier to make it imperceptible to human senses while resisting attacks. JPEG images are the most

widely used and popular images, so JPEG images become the most commonly used images in steganography algorithms. Common JPEG image based steganography JSteg [7], JPHide & Seek [8], Out Guess [9] and model-based steganography for the MB (Model-based) system [10]. In the last few years, steganography has improved in complexity and performance, especially adaptive steganography based on image content. In the literature [11], the authors proposed the first adaptive steganography algorithm HUGO (Highly Undetectable stego) in the architecture of an embedding distortion function and STC (Syndrome Trellis Codes) coding based on the third-order symbiotic matrix features that portray the correlation of neighboring pixels in the spatial domain image, which improves the security of the steganography algorithm. The literature [12] and [13] construct the embedding distortion function based on the quantization rounding error of DCT coefficients in the JPEG compression process, and achieve the embedding of secret information by interfering with the rounding direction of DCT coefficients, and propose PQ (Perturbed Quantization) and MMe (Modified Matrix Encoding) steganography algorithms based on wet paper coding and matrix coding, respectively. The PQ (Perturbed Quantization) and MMe (Modified Matrix Encoding) steganography algorithms are the typical representatives of the early design of JPEG image adaptive steganography algorithms based on the embedding distortion function and steganography encoding architecture. In contrast, steganalysis, which determines whether a carrier contains steganographic information based on its visual and statistical properties, has also developed rapidly. Steganalysts currently focus on steganalysis of steganographic algorithms with grayscale images as carriers [14]. Steganalysis techniques can be classified according to the scope of adaptation into those that are specific to a particular steganographic tool or a class of embedding techniques and general methods that do not target specific embedding methods. Dedicated steganalysis algorithms exploit the security vulnerabilities present in a specific steganography algorithm for targeted detection. Representative methods include cardinality attack methods [15], RS methods [16], SPA methods [17], and WS methods [18], which are designed to address the security vulnerability of the histogram statistical asymmetry phenomenon caused by LSB steganography to carrier images. Dedicated steganalysis methods generally have higher detection rates, but are not very practical because new steganographic algorithms are emerging in practical applications, and not all steganographic algorithms can be exhausted. In this case, the importance of general-purpose steganalysis is becoming more and more prominent and has become the focus of research within the field of steganalysis in recent years. In recent years, some steganalysis techniques based on deep learning have also been proposed [19,20].

However, color images are more widely used in our daily life and work. Some researchers have also proposed some steganalysis algorithms for color image steganography. These algorithms include steganalysis algorithm based on the number of pixel color changes [21,22], steganalysis algorithm based on inter channel texture consistency [23], steganalysis algorithm based on inter channel co-occurrence matrix [24−26], steganalysis algorithm based on inter channel prediction error [27−29], and steganalysis algorithm based on different channel feature combinations [30]. Compared with simply applying the gray image steganography detection algorithm to the three channel then fuse the results, the above algorithm improves the accuracy of color image steganography detection. Because the pixels in different color channels have strong correlation and interaction, the current steganalysis algorithms for color images do not fully develop this correlation. Extracting features from the differences of different channels can further improve the accuracy of steganalysis of color images. The color image steganalysis algorithm based on deep learning is difficult to fully develop the correlation between the three channels of color images, and is easily deceived by adversarial examples. The

manually designed color image steganalysis feature preserves the strong correlation of three channels, and slight modifications to the pixels in each layer of channels can affect the entire feature. This means that the manually designed color image steganalysis algorithm can amplify the steganalysis traces and improve the steganalysis efficiency. Combining hand-crafted features with deep learning features can enhance the robustness of steganalysis [31]. Some algorithms also try to use deep learning to improve the performance of steganalysis [32,33].

In order to solve the problem that the correlation between the channels of color image is not fully utilized, a color image steganalysis algorithm based on quaternion is proposed. Firstly, the image is represented by quaternion, and then it is transformed by Quaternion Discrete Cosine Transform. With the idea of spatial rich model and JPEG rich model, the co-occurrence matrix is extracted from the transformed coefficients, and the 11,255 dimensional feature of color image steganalysis is designed. The main contributions of this paper are as follows:

1) The color image is represented by quaternion, and the Quaternion Discrete Cosine Transform is performed in the transform domain, therefore the three-dimensional data of the color image is processed as a whole, combines quaternions with color image steganalysis for the first time.

2) Processes the three channels of a color image as a whole, the internal correlation of color image is more fully developed, and an effective color image steganalysis algorithm is designed.

The rest of this article is organized as follows. The second section briefly introduces the classical spatial rich model, JRM algorithm, quaternion representation of color image, and the shortcomings of the rich model. In the third section, a color image steganalysis algorithm based on quaternion is proposed. The fourth section gives the experimental results. Finally, this paper summarizes in the fifth section.

2. Related works

2.1. SRM

SRM (Spatial Rich Model) feature is a feature based on multiple residuals. The SRM is based on the SPAM (Subtractive Pixel Adjacency Matrix) feature, that is, the residual image generated by the high pass filter is characterized by SPAM. The process of image processing by high pass filter can be regarded as the process of obtaining the difference between the pixel and its predicted value, that is, obtaining the residual value. The formula for obtaining the residuals is as follows:

$$\mathrm{R_{i,j}} = \widehat{\mathrm{X}}_{\mathrm{i,j}}(\mathrm{N_{i,j}}) - \mathrm{cX_{i,j}} \tag{1}$$

Where $N_{i,j}$ is a number of domain pixels of $X_{i,j}$, $\hat{X}_{i,j}(N_{i,j})$ predicts $cX_{i,j}$ through this field. Since the number of pixels in $N_{i,j}$ is equal to c, c is called the residual order. The residuals $R_{i,j}$ are then quantized and truncated as follows.

$$\mathrm{R_{i,j}} \leftarrow \mathrm{trunc_T}(\mathrm{round}(\frac{\mathrm{R_{i,j}}}{\mathrm{q}})) \tag{2}$$

Where q is the quantization step and the truncation function is defined as

$$\mathrm{trunc}_{\mathrm{T}}(\mathrm{x}) = \begin{cases} \mathrm{x}, \mathrm{x} \in [-\mathrm{T}, \mathrm{T}] \\ \mathrm{Tsign}(\mathrm{x}), \mathrm{x} \notin [-\mathrm{T}, \mathrm{T}] \end{cases} \tag{3}$$

The significance of truncation operation is that, on the one hand, steganalysis is more interested in regions with strong correlation and small residuals; on the other hand, truncating the residuals is helpful to reduce the dimension of the last extracted feature. The partial residual pattern of the rich model is shown in the following figure.

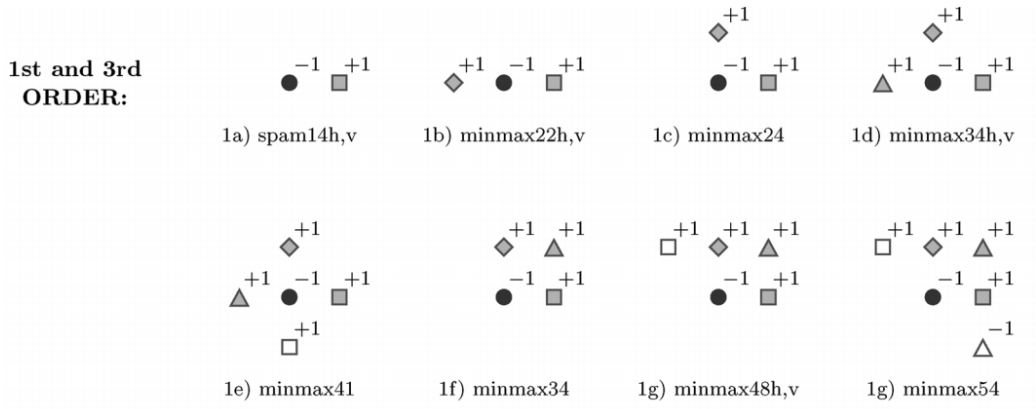


Figure 1. Partial residual pattern of rich model.

In the SRM feature, the co-occurrence matrix is mainly used to characterize the residual image, and the co-occurrence matrix is used as the feature of steganalysis. The key of the algorithm is how to generate features. According to the steganography characteristics of Hugo algorithm, secret information is embedded into the texture and edge region of the image. Therefore, in order to reflect the steganography changes well, the algorithm uses 106 sub models to generate residuals, so as to increase the diversity of features and improve the accuracy of feature detection. In the process of generating residuals, the algorithm also uses symmetry (Symbolic Symmetry and Spatial Symmetry) to reduce the feature dimension and increase the robustness of the feature. Due to the large number of sub models and the large feature dimension of a single model, the feature dimensions of the final two versions of SRM are 12,753 and 34,671 based on different quantization steps.

2.2. JRM

JPEG steganography is mainly performed on JPEG coefficients. Using the idea of rich model for reference, Kodovsky proposed the JRM (JPEG Rich Model) steganalysis feature [34]. The JRM feature also contains several sub models, and each sub model is constructed according to the statistical correlation of a class of block JPEG coefficients. Therefore, the JRM feature considers the intra block and inter block correlations of various JPEG coefficients, and detects the disturbance of steganography on the steganography coefficients from as many aspects as possible, thus improving the detection accuracy of JPEG image steganography. JRM features are mainly composed of block JPEG coefficients and the second-order co-occurrence matrix of their differences. Define the following matrix:

$$A_{i,j}^{x} = \left|D_{i,j}\right|, i = 1, \ldots, M, j = 1, \ldots, N$$

$$A_{i,j}^{\rightarrow} = \left|D_{i,j}\right| - \left|D_{i,j+1}\right|, i = 1, \ldots, M, j = 1, \ldots, N-1$$

$$A_{i,j}^{\downarrow} = \left|D_{i,j}\right| - \left|D_{i+1,j}\right|, i = 1, \ldots, M-1, j = 1, \ldots, N$$

$$A_{i,j}^{\searrow} = \left|D_{i,j}\right| - \left|D_{i+1,j+1}\right|, i = 1, \ldots, M-1, j = 1, \ldots, N-1$$

$$A_{i,j}^{\rightrightarrows} = \left|D_{i,j}\right| - \left|D_{i,j+8}\right|, i = 1, \ldots, M, j = 1, \ldots, N-8$$

$$A_{i,j}^{\downdownarrows} = \left|D_{i,j}\right| - \left|D_{i+8,j}\right|, i = 1, \ldots, M-8, j = 1, \ldots, N$$

In JPEG coefficient matrix A^{*}, $*\in \{x, \rightarrow, \downarrow, \searrow, \rightrightarrows, \downdownarrows\}$, the 2nd order coeval matrix $C_T^{\star}(x, y, \triangle x, \triangle y)$ is constructed from the elements on the (x,y) and $(x + \triangle x, y + \triangle y)$ modes, form a single sub-model in JRM, (x,y) are also the coordinates within the block, Δx and Δy are a fixed positional offset, whose offset range can reach other blocks. The elements in $C_T^{\star}$ are calculated by the following equation.

$$C_{kl}^{\star}(x, y, \triangle x, \triangle y) = \frac{1}{Z}\sum_{i,j} \left|\left\{T_{xy}^{(i,j)} \middle| T = trunc_T(A^{\star}); T_{xy}^{(i,j)} = k, T_{x+\triangle x, y+\triangle y}^{(i,j)} = l\right\}\right| \quad (4)$$

Where Z is the normalization constant used to ensure that $\sum_{k,l} c_{kl}^{*} = 1$; $trunc_T(\cdot)$ denotes the truncation operation.

$$trunc_T(x) = \begin{cases} T \cdot sign(x), & |x| > T \\ x, & other \end{cases} \quad (5)$$

The total dimension of JRM features is 8635 + 2620 = 11,255, and the analysis effect of JRM features can be further improved through modifications such as Descartes calibration.

2.3. Quaternion

In order to use the color information of a color image and the correlation information between color channels, Quaternion is used to represent a color image. Quaternion q is composed of one real part and three imaginary parts, and its form is as follows:

$$q = a + bi + cj + dk\{a, b, c, d \in R\} \quad (6)$$

where i, j, k need to satisfy

$$\begin{cases} i^2 = j^2 = k^2 = -1 \\ ij = -ji = k \\ jk = -kj = i \\ ki = -ik = j \end{cases} \tag{7}$$

Color image is a special quaternion. The specific representation is that the three RGB channels correspond to the three imaginary coefficients of the quaternion, and the real coefficient is 0, that is, each pixel of the color image is a pure quaternion.

QDCT (Quaternion Discrete Cosine Transform) has the characteristics similar to the DCT transform that transforms the energy dispersion in the spatial domain to the energy concentration in the frequency domain, and the coefficients obtained from the color image after the frequency domain transform have the color information and the inter-channel information of the color image. After the color image is transformed by QDCT, the original image can be restored by inverse transform of IQDCT. Take the airplane image as an example. After QDCT transformation, the images of the three channels restored by inverse transformation are as follows:

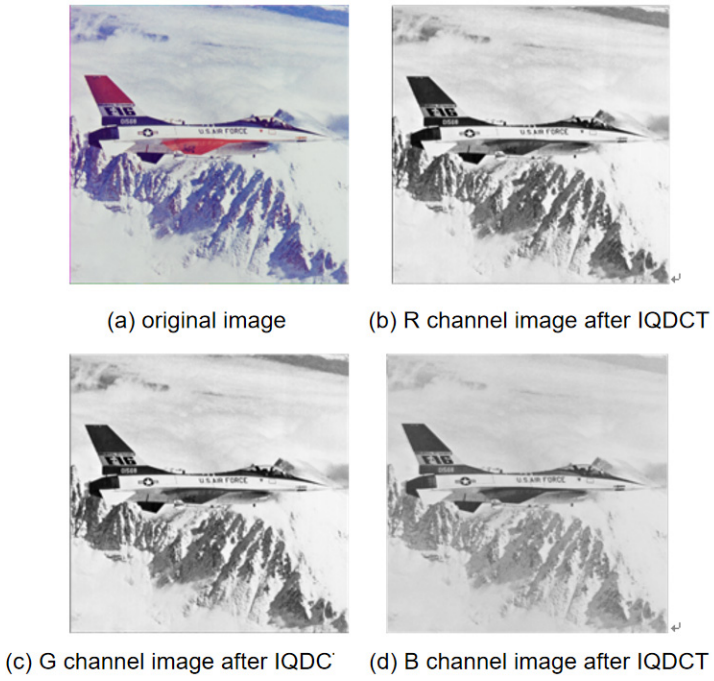


Figure 2. Three channels restored by inverse transformation.

3. The proposed method

This paper designs a steganalysis algorithm for color images. First, the color images are represented by quaternions, and then Quaternion Discrete Cosine Transform is performed [35]. With the help of the idea of spatial rich model and JPEG rich model, the co-occurrence matrix is calculated for the coefficients after Quaternion Discrete Cosine Transform, and the steganalysis characteristics of color images are finally obtained according to the co-occurrence matrix. The algorithm framework of this paper is as follows:

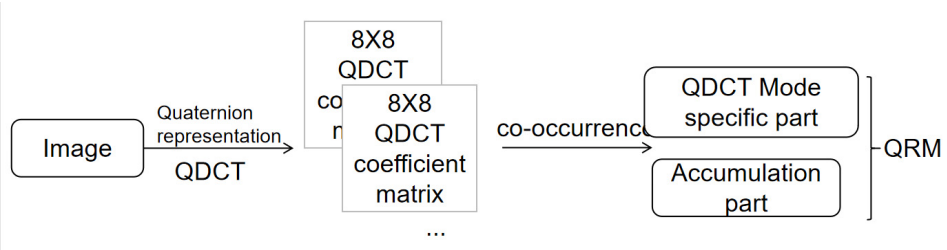


Figure 3. Framework of the proposed algorithm.

The specific steps of the algorithm are as follows.

1) The image is represented by quaternions.

2) Divide the image into blocks and perform QDCT transformation on each block.

Color image Quaternion Discrete Cosine Transform has two forms, Quaternion Discrete Cosine Left Transform (L-QDCT) and Quaternion Discrete Cosine Right Transform (R-QDCT). The specific transformations are:

$$\mathrm{QDCT_q^L(p,s)} = \alpha(\mathrm{p})\alpha(\mathrm{s})\sum_{\mathrm{m=0}}^{\mathrm{M-1}}\sum_{\mathrm{n=0}}^{\mathrm{N-1}} \mathrm{u_q}\cdot \mathrm{f_q(m,n)}\cdot \mathrm{N(p,s,m,n)} \tag{8}$$

$$\mathrm{QDCT_q^R(p,s)} = \alpha(\mathrm{p})\alpha(\mathrm{s})\sum_{\mathrm{m=0}}^{\mathrm{M-1}}\sum_{\mathrm{n=0}}^{\mathrm{N-1}} \mathrm{f_q(m,n)}\cdot \mathrm{N(p,s,m,n)}\cdot \mathrm{u_q} \tag{9}$$

$f_q(m,n)$ presents an $M\times N$ quaternion matrix, u_q presents pure unit quaternion ($|u_q|=1$ and $s(u_q)=0$, in this article for $\frac{\sqrt{3}}{3}i+\frac{\sqrt{3}}{3}j+\frac{\sqrt{3}}{3}k$). $QDCT_q^L(p,s)$ and $QDCT_q^R(p,s)$ present transformed matrix, $\alpha(p)$, $\alpha(s)$ and $T(p,s,m,n)$ are defined as follows.

$$\alpha(\mathrm{p}) = \begin{cases}\sqrt{1/\mathrm{M}}, \mathrm{p}=0\\ \sqrt{2/\mathrm{M}}, \mathrm{p}\neq 0\end{cases} \tag{10}$$

$$\alpha(\mathrm{s}) = \begin{cases}\sqrt{1/\mathrm{N}}, \mathrm{s}=0\\ \sqrt{2/\mathrm{N}}, \mathrm{s}\neq 0\end{cases} \tag{11}$$

$$\mathrm{T(p,s,m,n)} = \cos\left[\frac{\pi(2\mathrm{m}+1)\mathrm{p}}{2\mathrm{M}}\right]\cos\left[\frac{\pi(2\mathrm{n}+1)\mathrm{s}}{2\mathrm{N}}\right] \tag{12}$$

Corresponding to the two positive transformations of QDCT, the two inverse transformations of QDCT (IQDCT) are as follows.

$$\mathrm{IQDCT_q^L(m,n)} = \sum_{\mathrm{p=0}}^{\mathrm{M-1}}\sum_{\mathrm{s=0}}^{\mathrm{N-1}} \alpha(\mathrm{p})\alpha(\mathrm{s})\cdot \mathrm{u_q}\cdot \mathrm{C_q(p,s)}\cdot \mathrm{N(p,s,m,n)} \tag{13}$$

$$\mathrm{IQDCT_q^R(m,n)} = \sum_{\mathrm{p=0}}^{\mathrm{M-1}}\sum_{\mathrm{s=0}}^{\mathrm{N-1}} \alpha(\mathrm{p})\alpha(\mathrm{s})\cdot \mathrm{C_q(p,s)}\cdot \mathrm{N(p,s,m,n)}\cdot \mathrm{u_q} \tag{14}$$

$C(p,s)$ represrnts coefficient matrix after Discrete Dosine Transformation of Quaternions. The positive and negative conversion relations are as follows.

$$\mathrm{f_q(m,n) = IQDCT_q^L[QDCT_q^L(f_q(m,n))] = IQDCT_q^R[QDCT_q^R(f_q(m,n))]} \tag{15}$$

The image can be restored by IQDCT after QDCT transformation. Taking the airport.jpg image as an example, the 8 × 8 block coefficients of imaginary component i after QDCT transformation are as follows.

Table 1. 8 × 8 block coefficients of imaginary component i after QDCT.

131	92.16	85.16	78.26	65.75	56.63	36.05	15.98
-45.88	-15.25	-9.8	-9.16	-10.89	-8.08	-7.27	-4.66
-7.2	7.48	4.5	5.79	4.28	2.65	1.89	2.49
-7.59	2.92	-0.03	1.31	-0.05	-0.97	0.90	-1.8
-7.87	0.21	1.17	-0.88	-0.51	-0.67	0.37	1.43
-9.40	-1.33	0.61	3.22	-0.62	-0.44	-0.19	-2.51
-6.63	2.47	1.89	0.91	0.34	0.96	1.13	-0.28
-5.31	-1.91	0.24	0.81	0.25	-0.25	0.80	-0.62

3) QRM (Quaternion Rich Model) mainly consists of a 2nd order co-occurrence matrix of the block JPEG coefficients of each RGB channel and the difference between them. $D_R \in Z_R^{M\times N}$ represents the R-channel data of a $M\times N$ color JPEG image, $D_{R,xy}^{(i,j)}$ represents JPEG coefficients on the (x,y) pattern in the (i,j) block, $(x,y)\in\{0,...,7\}^2, i=1,...,\lceil M/8\rceil$, and $j=1,...,\lceil N/8\rceil$. $D_{R,xy}$ represents the (x,y) coefficient in the large matrix consisting of the whole R-channel blocks, $i=1,...,M, j=1,...,N$.

$$\mathrm{A_{R,i,j}^{x} = |D_{R,i,j}|, i = 1,\ldots,M, j = 1,\ldots,N}$$

$$\mathrm{A_{R,i,j}^{\rightarrow} = |D_{R,i,j}| - |D_{R,i,j+1}|, i = 1,\ldots,M, j = 1,\ldots,N-1}$$

$$\mathrm{A_{R,i,j}^{\downarrow} = |D_{R,i,j}| - |D_{R,i+1,j}|, i = 1,\ldots,M-1, j = 1,\ldots,N}$$

$$\mathrm{A_{R,i,j}^{\searrow} = |D_{R,i,j}| - |D_{R,i+1,j+1}|, i = 1,\ldots,M-1, j = 1,\ldots,N-1}$$

$$A^{\rightrightarrows}_{R,i,j} = \left|D_{R,i,j}\right| - \left|D_{R,i,j+8}\right|, i = 1,\ldots,M, j = 1,\ldots,N-8$$

$$A^{\downdownarrows}_{R,i,j} = \left|D_{R,i,j}\right| - \left|D_{R,i+8,j}\right|, i = 1,\ldots,M-8, j = 1,\ldots,N$$

In the R-channel JPEG coefficient matrix A^{*}_{R}, $*\in \{x, \rightarrow, \downarrow, \searrow, \rightrightarrows, \downdownarrows\}$, the 2nd order coeval matrix $C^{\star}_{R,T}(x, y, \triangle x, \triangle y)$ is constructed from the elements on the (x, y) and (x+Δx,y+Δy) modes, form a single sub-model in QRM, (x, y) are also the coordinates within the block, Δx and Δy are a fixed positional offset, whose offset range can reach other blocks. The JPEG coefficient co-occurrence matrices of the three RGB channels are combined to obtain the 3D QRM features: $(C^{\star}_{R,T}(x, y, \triangle x, \triangle y), (C^{\star}_{G,T}(x, y, \triangle x, \triangle y), (C^{\star}_{B,T}(x, y, \triangle x, \triangle y), 0)$. Then find the co-occurrence matrix for each coefficient block, and the specific method is beyond the design method of the co-occurrence matrix of JRM. The features of the three channels are accumulated, and the final QRM features are combined by the DCT mode specific part and the accumulation part. Consistent with the JRM feature, the total feature of QRM is 8635 + 2620 = 11,255 dimensions, and the relevant sub models and their feature dimensions are shown in the table.

The calculation method of the specific co-occurrence matrix is consistent with the JRM algorithm introduced previously, and will not be repeated here. It should be noted that the coefficients obtained by QDCT are four-dimensional, while JRM is one-dimensional. The processing method in this paper is to accumulate the three-dimensional features of discarding the real part, so as to reduce the computational complexity.

Most of the proposed color image steganalysis algorithms simply split RGB three channels, and then process each channel separately to obtain steganalysis features, ignoring the correlation between the three channels of color image. The algorithm proposed in this paper uses quaternion to operate the information of all three color channels. The Quaternion Discrete Cosine Transform has characteristics similar to the DCT transform that transforms the energy dispersion in the spatial domain to the energy concentration in the frequency domain, and the coefficients obtained from the color image after the frequency domain transform have the color information and make use of the inter-channel information of the color image. The steganalysis features designed on this basis have more diversity, and develop the internal correlation of color images, which improves the performance of steganalysis of color images.

Table 2. Relevant sub models and feature dimensions.

Accumulator sub model					Left splice	Right splice
180,I^{x}					180,I^{x}	
224,$I^{\rightarrow}_{f}$	224,$I^{\downarrow}_{f}$	224,$I^{\searrow}_{f}$	224,$I^{\rightrightarrows}_{f}$	224,$I^{\downdownarrows}_{f}$	1220,I_{f}	2620,I
224$I^{\rightarrow}_{s}$	224,$I^{\downarrow}_{s}$	224,$I^{\searrow}_{s}$	224,$I^{\rightrightarrows}_{s}$	224,$I^{\downdownarrows}_{sf}$	1220,I_{x}	

Most of the existing color image steganalysis features deal with the three channels of the color image separately, thus ignoring the correlation between the three channels. In order to solve the above problems, we innovatively apply quaternions to color image steganalysis. Quaternion treats the image as a whole, and Quaternion Discrete Cosine Transform can effectively exploit the correlation between the three channels of color images, which can effectively improve steganography detection efficiency. Therefore, we use QDCT to design a new algorithm for steganalysis of color images. Technically, we perform QDCT on color images, process the obtained four dimensional QDCT coefficients, and obtain color image steganalysis features. Our designed feature preserves the strong correlation of three channels, and slight modifications to the pixels in each layer of channels can affect the entire feature. This means that the manually designed color image steganalysis algorithm can amplify the steganalysis traces and improve the steganalysis efficiency.

In addition, high dimensional steganalysis features increase computational complexity, while excessive features can affect steganalysis detection results. Therefore, we have designed a feature optimization selection strategy. The four-dimensional coefficients obtained from the QDCT of an image consist of one real part c_r and three imaginary parts c_i, c_j, c_k. The three imaginary coefficient parts store the main information of the color image, while the real coefficient part can be discarded. Excessive redundant features will reduce steganography efficiency. We compared the experimental results of features that include real coefficients $F_{r,i,j,k}$ with features that discard real coefficients $F_{i,j,k}$, 10,000 color images with size of 512 × 512 from BossBase image library are selected, the steganalysis algorithms are LSBM, and the embedding rates are 0.05, 0.1, 0.2, 0.3, 0.4, 0.5 bpc (bits per channel) respectively. The experimental results are shown in the table 3 below, results show that the detection error of features discarding the real part is lower than those containing the real part. The experiment verified that the real part of the coefficients reduce the efficiency of steganalysis, so the real part of the coefficients was discarded in the designed features.

Table 3. Detection errors between features $F_{r,i,j,k}$ and features $F_{i,j,k}$.

Payload (bpc)	0.05	0.1	0.2	0.3	0.4
$F_{r,i,j,k}$	0.339	0.252	0.160	0.101	1.069
$F_{i,j,k}$	0.331	0.244	0.152	0.093	0.061

At the same time, we also compared the experimental results of concatenating three imaginary features $F_{i,j,k}$ and combining three imaginary features F_{i+j+k}.

$$F_{i,j,k} = F_i, F_j, F_k \tag{16}$$

$$F_{i+j+k} = F_i + F_j + F_k \tag{17}$$

The same configuration as the above experiment was used, and the experimental results are shown in the Table 4 below, results show that the detection error of features combining three imaginary features is lower than those concatenating three imaginary features. The experiment showed that concatenating three imaginary features not only increased algorithm complexity, but also reduced detection efficiency. Therefore, we adopted F_{i+j+k} to obtain a feature of 11,255 dimensions. The experimental results demonstrate the superiority of our designed feature technology. The algorithm we proposed is not only a simple combination of quaternions and steganalysis, but also

involves the selection of steganalysis features.

Table 4. Detection errors between feature $F_{i,j,k}$ and feature F_{i+j+k}.

Payload (bpc)	0.05	0.1	0.2	0.3	0.4
$F_{i,j,k}$	0.331	0.244	0.152	0.093	0.061
F_{i+j+k}	0.325	0.233	0.140	0.086	0.054

4. Experiment

For the experiment, 10,000 color images with size of 512 × 512 from BossBase image library are selected, to compare the proposed algorithm with SCRM, the state-of-the-art rich models for color image steganalysist. The steganalysis algorithms are LSBM and WOW [36], and the embedding rates are 0.05, 0.1, 0.2, 0.3, 0.4, 0.5 bpc (bits per channel) respectively. The PE (Prediction Errors) of different algorithms are compared. The experimental results of the proposed algorithm and the comparison algorithm are as follows.

The embedding algorithm is LSBM:

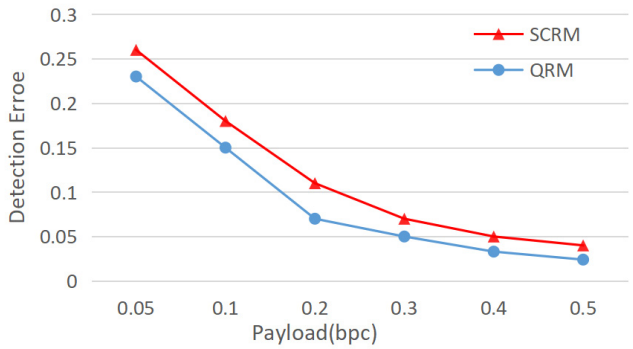


Figure 4. Detection errors with the embedding algorithm is LSBM.

The embedding algorithm is WOW:

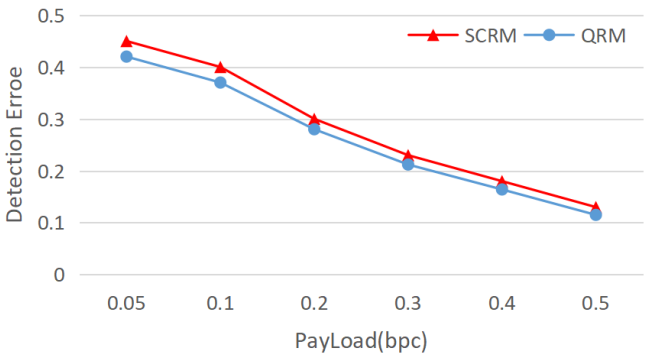
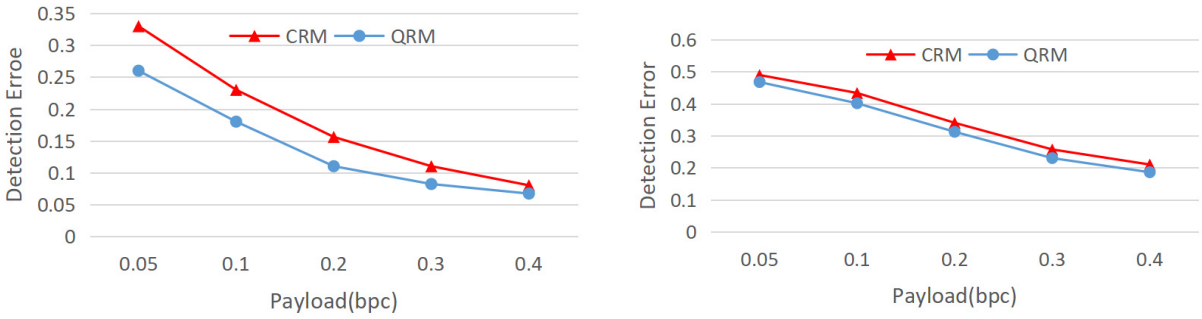


Figure 5. Detection errors with the embedding algorithm is WOW.

The detection errors of proposed QRM algorithm and classic SCRM color image steganalysis algorithm for LSBM with the payloads are 0.05, 0.1, 0.2, 0.3, 0.4, 0.5 bpc respectively are showed in Figure 4. Figure 5 shows the detection results for the WOW. It can be found that because QRM algorithm is improved on the basis of the classical SCRM algorithm, the correlation between various channels of color image is considered by using quaternion impulse division, and the experimental effect is better than that of SCRM algorithm.

The SRM gray steganalysis algorithm is used for the three channels respectively. The settings are consistent with the above experimental settings. Under different payloads, the experimental comparison is as follows. Figure 6 shows the detection errors of our proposed algorithm and gray image steganalysis algorithm for LSBM and WOW embedding algorithm with different payloads.



(a) (b)

Figure 6. Experimental comparison with gray steganalysis algorithm, (a): embedding algorithm is LSBM, (b): embedding algorithm is WOW.

The experiments further show that the gray image steganalysis algorithm is used directly in three channels, ignoring the correlation among the three channels of the color image. The steganalysis algorithm proposed in this paper is better than this algorithm. Some content-adaptive color image steganography techniques select and modify different pixels of different channels, each channel has different characteristics, thus analyzing each channel separately will reduce the accuracy of the steganalysis algorithm. The three channels of a color image have a strong correlation, which can be used to amplify tampering traces and thus effectively improve the detection rate. If the pixels of a single channel are modified, this modification is amplified to all three channels by the mapping of the QDCT, which improves the sensitivity of detection features to tampering.

Through the above experimental settings, the performance of the SCRM algorithm [24], the SCRM+SGF algorithm proposed by Abdulrahman et al. [22], the SCRM+SGF+DSRMQ1+DSGF algorithm proposed by Feng and Hu [35], and the steganalysis algorithm proposed in this paper are tested. The average test error of three steganalysis algorithms of WOW and S-UNIWARD under different loads is shown in the Tables 5 and 6.

Table 5. Average test error of three steganalysis algorithms of WOW.

Algorithm	Payload (bpc)	0.05	0.1	0.2	0.3	0.4
WOW	SCRM	0.374	0.269	0.156	0.101	0.069
	SCRM+SGF	0.387	0.276	0.161	0.104	0.070
	SCRM+SGF+DSRMQ1+DSGF	0.334	0.243	0.144	0.087	0.060
	QRM	0.325	0.233	0.140	0.086	0.054

Table 6. Average test error of three steganalysis algorithms of S-UNIWARD.

Algorithm	Payload(bpc)	0.05	0.1	0.2	0.3	0.4
S-UNIWARD	SCRM	0.370	0.269	0.156	0.096	0.063
	SCRM+SGF	0.379	0.267	0.156	0.097	0.064
	SCRM+SGF+DSRMQ1+DSGF	0.333	0.237	0.134	0.084	0.055
	QRM	0.325	0.225	0.127	0.072	0.050

It can be seen that the proposed steganalysis algorithm is better than the first two WOW steganalysis algorithms under different payloads. Specifically, when the payload is small, the steganalysis algorithm proposed in this paper has more significant advantages. Compared with the comparison algorithm, the average test error of WOW and S-UNIWARD can be reduced by 4% to 5%. In addition, with the increase of payload, the success rate of the three algorithms is improved, which leads to the advantage of the algorithm proposed in this paper is not as significant as that of low payload.

5. Conclusions

Most of the currently proposed color image steganalysis algorithms simply split RGB three channels, and then process each channel separately to obtain steganalysis features, ignoring the correlation between the three channels of color image. The algorithm proposed in this paper uses quaternion to operate on the information of all three color channels. Quaternion Discrete Cosine Transform has characteristics similar to the DCT transform that transforms the energy dispersion in the spatial domain to the energy concentration in the frequency domain, and the coefficients obtained from the color image after the frequency domain transform have the color information and the inter-channel information of the color image. The steganalysis features designed on this basis have more diversity, and make use of the internal correlation of color images, which improves the performance of steganalysis of color images.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. U1804263, No. U1736214, No. 62172435).

Conflict of interest

All authors declare that there are no competing interests.

References

1. I. Cox, M. Miller, J. Bloom, J. Fridrich, T. Kalker, *Digital Watermarking and Steganography*, Morgan Kaufmann, 2007. https://doi.org/10.1016/B978-0-12-372585-1.X5001-3
2. W. Luo, F. Huang, J. Huang, Edge adaptive image steganography based on LSB matching revisited, *IEEE Trans. Inf. Forensics Secur.*, **5** (2010), 201−214. https://doi.org/10.1109/TIFS.2010.2041812
3. E. Kawaguchi, R. O. Eason, Principles and applications of BPCS steganography, in *Proceedings of the Photonics East (ISAM, VVDC, IEMB)*, International Society for Optics and Photonics, **3528** (1999), 464−473. https://doi.org/10.1117/12.337436
4. V. K. Sharma, V. Shrivastava, A steganography algorithm for hiding image in image by improved LSB substitution by minimize detection, *J. Theor. Appl. Inf. Technol.*, **36** (2012), 1−8. Available from: http://www.jatit.org/volumes/Vol36No1/1Vol36No1.pdf.
5. B. C. Nguyen, S. M. Yoon, H. K. Lee, Multi bit plane image steganography, in *Digital Watermarking*, Springer, (2006), 61−70. https://doi.org/10.1007/11922841_6
6. M. Omoomi, S. Samavi, S. Dumitrescu, An efficient high payload ±1 data embedding scheme, *Multimedia Tools Appl.*, **54** (2011), 201−218. https://doi.org/10.1007/s11042-010-0517-z
7. D. Upham, Jpeg-jsteg-v4, 2008. Available from: http://www.funet.fi/pub/crypt/steganography/jpeg-jsteg-v4.diff.gz.
8. F. A. Petitcolas, R. J. Anderson, M. G. Kuhn, Information hiding-a survey, *Proc. IEEE*, **87** (1999), 1062−1078. https://doi.org/10.1109/5.771065
9. N. Provos, Defending against statistical steganalysis, in *Proceedings of the Usenix Security Symposium*, **10** (2001), 323−336. https://dl.acm.org/doi/10.5555/1251327.1251351
10. P. Sallee, Model-based steganography, in *Digital Watermarking*, Springer, (2004), 154−167. https://doi.org/10.1007/978-3-540-24624-4_12
11. T. Pevný, T. Filler, P. Bas, Using high-dimensional image models to perform highly undetectable steganography, in *International Workshop on Information Hiding*, **6387** (2010), 161−177. https://doi.org/10.1007/978-3-642-16435-4_13
12. J. Fridrich, M. Goljan, D. Soukal, Perturbed quantization steganography, *Multimedia Syst.*, **11** (2005), 98−107. https://doi.org/10.1007/s00530-005-0194-3
13. X. Song, F. Liu, C. Yang, X. Luo, Y. Zhang, Steganalysis of adaptive JPEG steganography using 2D gabor filters, in *Proceedings of the 3rd ACM Workshop on Information Hiding and Multimedia Security*, (2015), 15−23. https://doi.org/10.1145/2756601.2756608
14. A. Westfeld, A. Pfitzmann, Attacks on steganographic systems, in *International Workshop on Information Hiding*, Springer, Berlin, Heidelberg, (2000), 61−76. https://doi.org/10.1007/10719724_5
15. J. Fridrich, M. Goljan, D. Rui, Reliable detection of LSB steganography in color and grayscale images, in *Proceedings of the 2001 Workshop on Multimedia and Security: New Challenges*, 2002. https://doi.org/10.1145/1232454.1232466
16. S. Dumitrescu, X. Wu, W. Zhe, Detection of LSB steganography via sample pair analysis, *IEEE Trans. Signal Process.*, **51** (2002), 1995−2007. https://doi.org/10.1109/TSP.2003.812753
17. J. Fridrich, M. Goljan, On estimation of secret message length in LSB steganography in spatial domain, in *Security, Steganography, and Watermarking of Multimedia Contents VI*, San Jose, California, USA, 2004. https://doi.org/10.1117/12.521350

18. J. Fridrich, M. Long, Steganalysis of LSB encoding in color images, in *Proceedings of IEEE International Conference on Multimedia and Expo*, (2000), 1279−1282. https://doi.org/10.1109/ICME.2000.871000
19. W. You, H. Zhang, X. Zhao, A siamese CNN for image steganalysis, *IEEE Trans. Inf. Forensics Secur.*, **16** (2020), 291−306. https://doi.org/10.1109/TIFS.2020.3013204
20. K. Wei, W. Luo, S. Tan, J. Huang, Universal deep network for steganalysis of color image based on channel representation, *IEEE Trans. Inf. Forensics Secur.*, **17** (2022), 3022−3036. https://doi.org/10.1109/TIFS.2022.3196265
21. G. Su, H. Lu, H. Fang, B. Yang, A steganalysis algorithm based on statistic characteristics of the color images, in *Proceedings of International Conference on Computer Science and Network Technology*, (2011), 2294−2297. https://doi.org/10.1109/ICCSNT.2011.6182432
22. H. Abdulrahman, M. Chaumont, P. Montesinos, B. Magnier, Color images steganalysis using RGB channel geometric transformation measures, *Secur. Commun. Netw.*, **9** (2016), 2945−2956. https://doi.org/10.1002/sec.1427
23. M. Goljan, J. Fridrich, R. Cogranne, Rich model for steganalysis of color images, in *Proceedings of IEEE International Workshop on Information Forensics and Security*, (2014), 185−190. https://doi.org/10.1109/WIFS.2014.7084325
24. M. Goljan, J. Fridrich, CFA-aware features for steganalysis of color images, in *Proceedings of SPIE 9409, Media Watermarking, Security, and Forensics*, (2015), 1−13. https://doi.org/10.1117/12.2078399
25. X. Liao, G. Chen, J. Yin, Content-adaptive steganalysis for color images, *Secur. Commun. Netw.*, **9** (2016), 5756−5763. https://doi.org/10.1002/sec.1734
26. S. Lyu, H. Farid, Steganalysis using color wavelet statistics and one-class support vector machines, in *Proceedings of SPIE 5306, Security, Steganography, and Watermarking of Multimedia Contents VI*, (2004), 35−45. https://doi.org/10.1117/12.526012
27. Q. Liu, A. H. Sung, J. Xu, B. M. Ribeiro, Image complexity and feature extraction for steganalysis of LSB matching steganography, in *Proceedings of the 18th International Conference on Pattern Recognition*, (2006), 267−270. https://doi.org/10.1109/ICPR.2006.684
28. F. Y. Li, X. P. Zhang, J. Yu, Steganalysis for color JPEG images based on ensemble proportion training, *J. Electron. Inf. Technol.*, **36** (2014), 114−120. https://doi.org/10.3724/SP.J.1146.2013.00443
29. H. Abdulrahman, M. Chaumont, P. Montesinos, B. Magnier, Color image steganalysis based on steerable gaussian filters bank, in *Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security*, (2016), 109−114. https://doi.org/10.1145/2909827.2930799
30. Y. Kang, F. Liu, C. Yang, L. Xiang, X. Luo, P. Wang, Color image steganalysis based on channel gradient correlation, *Int. J. Distrib. Sens. Netw.*, **15** (2019). https://doi.org/10.1177/1550147719852031
31. C. Qin, Y. Chen, K. Chen, X. Dong, W. Zhang, Y. He, et al., Feature fusion based adversarial example detection against second-round adversarial attacks, *IEEE Trans. Artif. Intell.*, **2022** (2022), 1−12. https://doi.org/10.1109/TAI.2022.3190816
32. J. Zeng, S. Tan, G. Liu, B. Li, J. Huang, WISERNet: wider separate-then-reunion network for steganalysis of color images, *IEEE Trans. Inf. Forensics Secur.*, **14** (2019), 2735−2748. https://doi.org/10.1109/TIFS.2019.2904413

33. T. Wu, W. Ren, D. Li, L. Wang, J. Jia, JPEG steganalysis based on denoising network and attention module, *Int. J. Intell. Syst.*, **37** (2021), 5011−5030. https://doi.org/10.1002/int.22749
34. J. Kodovsky, J. Fridrich, Steganalysis of JPEG images using rich models, in *Proceedings of SPIE - Media Watermarking, Security, and Forensics*, 2012. https://doi.org/10.1117/12.907495
35. W. Feng, B. Hu, Quaternion discrete cosine transform and its application in color template matching, in *Congress on Image and Signal Processing*, Sanya, (2008), 252−256. https://doi.org/10.1109/CISP.2008.61
36. V. Holub, J. Fridrich, Designing steganographic distortion using directional filters, in *2012 IEEE International Workshop on Information Forensics and Security (WIFS)*, (2012), 234−239. https://doi.org/10.1109/WIFS.2012.6412655

AIMS Press