Cryptanalysis of an Image Scrambling Scheme without Bandwidth Expansion

Shujun Li, Chengqing Li, Kowk-Tung Lo, *Member, IEEE* and Guanrong Chen, *Fellow, IEEE*

Abstract

Recently, a novel image scrambling (i.e., encryption) scheme without bandwidth expansion was proposed based on two-dimensional (2-D) discrete prolate spheroidal sequences (DPSS). This paper gives a comprehensive cryptanalysis of the image scrambling scheme, and draw a conclusion that it is not sufficiently secure against various cryptographical attacks, including ciphertext-only attack, known/chosen-plaintext attack and chosen-ciphertext attack. The cryptanalytic results suggest that the image scrambling scheme can only be used to realize perceptual encryption, instead of provide content protection for digital images.

Index Terms

discrete prolate spheroidal sequences (DPSS), image scrambling, encryption, cryptanalysis, ciphertext-only attack, known-plaintext attack, chosen-plaintext attack, chosen-ciphertext attack, Hadamard matrix.

I. Introduction

The content protection of multimedia data (especially digital images and videos) through encryption has attracted more and more attention as the rapid development of multimedia and network technologies in past decades. During last decades, various image/video encryption (or scrambling[1]) schemes have been proposed, but some of which have been successfully cryptanalyzed. To offer a reasonable background knowledge on the content of this paper, in the following a very brief introduction to some existing image/video encryption schemes will be given. For a more comprehensive survey of the state-of-the-art of this topic, readers are suggested to refer to [1]–[4].

The most straight-forward idea of image/video encryption is to consider the entire 2-D multimedia data as a 1-D textual bit-stream and then apply any conventional cipher that has been established in modern cryptography, such as DES, IDEA, AES, etc. [5], [6]. This solution is called naive encryption in some literature [7]. The major problem of naive encryption lies in the following two aspects: 1) the encryption speed may be too slow; 2) it does not consider the information redundancies existing in uncompressed images/videos and the syntax structures of compressed ones. A possible way to overcome the above problem is to encrypt part of the given plain-image/video, which is called selective (or partial) encryption. For example, for MPEG videos, only sign bits of the DCT coefficients and the motion vectors can be selected for encryption. Even though partial encryption may not provide a high level of security, it is still useful to realize perceptual encryption [2], [8] and format-compliant encryption [9], two interesting security requirements of some multimedia applications.

As one of frequently-used approaches to encrypt images and videos, some schemes were designed by secretly permuting pixels in the plain-image or each frame of the plain-video [10]–[12]. This idea can also be generalized to frequency domain, while in this case the encryption is achieved by permuting transform coefficients (and/or nodes for some transforms with a tree-like structure)[2] [15]–[18]. However, a large number of cryptanalysis work has shown that these permutation-based image/video encryption schemes are not sufficiently secure from a cryptographical point

The corresponding author is Shujun Li. Contact him by accessing his person web site: http://www.hooklee.com.

Shujun Li and Kowk-Tung Lo are with the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR, P. R. China.

Chengqing Li and Guanrong Chen are with the Department of Electronic Engineering, City University of Hong Kong, Kowloon Toon, Hong Kong SAR, P. R. China.

[1]The term "scrambling" was used instead of "encryption" by some designers of multimedia encryption schemes, especially by those early designers who intended to encrypt analogue signals by "scrambling" them in some way. From a cryptographical point of view, we just consider "scrambling" as a synonym of "encryption".

[2]Though speech encryption is not the focus of this paper, it deserve mentioning that many speech scrambling schemes were also developed based on this idea working in frequency domain [13], [14].

of view [7], [19]–[25]. The main security problems include: 1) the plain-image/video may be partially recovered due to the large information redundancies existing in natural images/videos; 2) secret permutations are always insecure against known/chosen-plaintext attack. As a general result of existing cryptanalysis, secret permutations must be combined with other techniques to design a secure image/video encryption scheme.

Another idea of designing encryption schemes is to scramble all the pixels and/or transform coefficients with one or more multiplicative or additive matrices. This idea can be considered as a generalization of the permutation-only encryption, since secret permutations can be formulated with a permutation matrix as shown in the theoretical models of some permutation-based speech scrambling schemes [13], [14]. Many optical image encryption have been developed in this way, by introducing double random phase matrices (keys) [4], [26], [27], which are used to scramble the plain-image in spatial and frequent (Fourier transform) domains, respectively. However, some recent cryptanalysis work [28]–[30] show that optical image encryption schemes of this kind is not sufficiently secure against known-plaintext attack and chosen-ciphertext attack.

Additionally, a large number of image encryption schemes were designed by combining different encryption techniques. For example, some image encryption schemes are based on the multi-round combination of secret permutations and pixel-value substitutions [31]–[33]. There are also some attempts of using chaos to design image/video encryption schemes [2], while some chaos-based image encryption schemes have been successfully cryptanalyzed [34]–[40].

This paper focuses on a new image scrambling scheme proposed in [41], which is a 2-D extension of a speech scrambling scheme proposed by Wyner in [42]. Compared with other existing image scrambling methods, this scheme does not introduce much high-frequency components into the cipher-spectrum, but results in a negligible expansion of bandwidth. This feature is useful in some real applications, as the cipher-image can be transmitted with a band-limited channel that carries the plain-image. The encryption process is mainly achieved by scrambling low-frequency components of 2-D DPSS (discrete prolate spheroidal sequences) transform with a multiplicative matrix, which serves as the secret key of the scheme. To further enhance the security, random swapping of some high-frequency components and multiple secret matrices were also suggested. When multiple secret matrices are used, each of which corresponds to the encryption of a single block of the plain-image.

In this paper, we make a complete investigation on the security of the image scrambling scheme in [41], and point out that it is not sufficiently secure against various cryptographical attacks, including ciphertext-only attack, known/chosen-plaintext attack and chosen-ciphertext attack. Some other security defects have also been found when a fixed secret matrix is used to encrypt all blocks of the plain-image. Based on the cryptanalytic results, we conclude that the image scrambling scheme should only be used to realize perceptual encryption, i.e., to degrade the visual quality of plain-images.

The rest of this paper is organized as follows. We first give a description of the image scrambling scheme in next section. Then, Section III focuses on the cryptanalytic findings, with both theoretical and experimental results given. In Sec. IV we discuss how to use the image scrambling scheme in practice. Finally the last section concludes this paper.

II. Image Scrambling Scheme without Bandwidth Expansion

The image scrambling scheme proposed in [41] is a 2-D extension of Wyner's signal scrambling scheme [42] based on discrete prolate spheroidal sequences (DPSS), which are defined as the normalized eigenvectors of the following real and symmetric matrix

$$\mathbf{V} = \left[\frac{\sin(2\pi W(m-n))}{\pi(m-n)}\right]_{0\le m,n\le N-1}. \tag{1}$$

Assume that the DPSS, i.e., the N eigenvectors of $\mathbf{V}$, are $\{\phi_j\}_{j=0}^{N-1}$, where $\phi_j = \begin{bmatrix}\phi_j(0) & \cdots & \phi_j(N-1)\end{bmatrix}^T$, and that the corresponding eigenvalues are $\{\lambda_j\}_{j=0}^{N-1}$. Slepian [43] showed that $\{\phi_j(n)\}_{j=0}^{N-1}$ form an orthonormal basis that span the subspace of sequences with an energy concentration in a certain band $[-W, W]$. Thus, for any sequence $\boldsymbol{a} = \begin{bmatrix}a(0) & \cdots & a(N-1)\end{bmatrix}^T$, one can use the DPSS to get another sequence $\boldsymbol{\alpha} = \begin{bmatrix}\alpha_1 & \cdots & \alpha_{N-1}\end{bmatrix}^T$ such that $\boldsymbol{\alpha} = \mathbf{S}\boldsymbol{a}$ (or $\boldsymbol{a} = \mathbf{S}^T\boldsymbol{\alpha}$), where $\mathbf{S} = \begin{bmatrix}\phi_0 & \cdots & \phi_{N-1}\end{bmatrix}^T$. Based on such a 2-D DPSS transformation, one can scramble $\boldsymbol{\alpha}$ and then perform an inverse transform as an alternative way of encrypting the original sequence $\boldsymbol{a}$. The encryption process can be described as $\boldsymbol{a}' = \mathbf{S}^T\mathbf{M}\boldsymbol{\alpha} = \mathbf{S}^T\mathbf{M}\mathbf{S}\boldsymbol{a}$, where $\mathbf{M}$ is the secret matrix that

scrambles $\boldsymbol{\alpha}$. For such a scrambling scheme, Wyner [42] showed that the bandwidth expansion will be negligible if the lowest eigenvalue corresponding to scrambled coefficients in $\boldsymbol{\alpha}$ is sufficiently high. Concretely, assuming that all coefficients in $\boldsymbol{\alpha}$ are ranked by the eigenvalues and only the v lowest coefficients $\{\alpha_j\}_{j=0}^{v-1}$ are scrambled, Wyner deduced that the energy concentration of the scrambled sequence $\boldsymbol{a}'$ differs at most $1-\lambda_{v-1}$ compared to the concentration of the original sequence $\boldsymbol{a}$. In this case, the secret matrix $\mathbf{M}$ is in the following form:

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_v & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N-v} \end{bmatrix}, \tag{2}$$

where $\mathbf{M}_v$ is the sub-matrix scrambling the v lowest coefficients and $\mathbf{I}_{N-v}$ is the $(N-v)\times(N-v)$ identity matrix. One can see that this scrambling scheme is a selective encryption algorithm, as some coefficients in $\boldsymbol{\alpha}$ are left unchanged.

In [41], Van De Ville et al. extended 1-D DPSS to the case of 2-D square passband region as follows:

$$\phi_{j_1,j_2}^{(2D)}(n_1,n_2) = \phi_{j_1}^{(1D)}(n_1)\phi_{j_2}^{(1D)}(n_2), \tag{3}$$

where $0 \le n_1, j_1 \le N_1 - 1$ and $0 \le n_2, j_2 \le N_2 - 1$. Accordingly, the eigenvalues corresponding to $\phi_{j_1,j_2}^{(2D)}$ are $\lambda_{j_1,j_2}^{(2D)} = \lambda_{j_1}^{(1D)}\lambda_{j_2}^{(1D)}$. Then, scanning all elements in each $\phi_{j_1,j_2}^{(2D)}$ to form a $N_1N_2 \times 1$ vector $\phi_j^{(2D)}$ ($0 \le j \le N_1N_2 - 1$) and sorting these eigenvectors such that $\lambda_0^{(2D)} \ge \cdots \ge \lambda_{N_1N_2-1}^{(2D)}$, one gets an $N_1N_2 \times N_1N_2$ matrix $\mathbf{S} = \begin{bmatrix} \phi_0^{(2D)} & \cdots & \phi_{N_1N_2-1}^{(2D)} \end{bmatrix}^T$. Next, given an $N_1N_2 \times N_1N_2$ secret matrix $\mathbf{M}$ and a $N_1N_2 \times 1$ vector $\boldsymbol{a}$, the 2-D scrambling scheme has the same encryption formula as the 1-D one:

$$\boldsymbol{a}' = \mathbf{S}^T\mathbf{M}\mathbf{S}\boldsymbol{a} = \mathcal{M}\boldsymbol{a}. \tag{4}$$

Each $N_1 \times N_2$ block of a digital image is scanned row by row (or column by column) to form a $N_1N_2 \times 1$ vector and then is encrypted following the above equation. After the encryption is done, all elements in the $N_1N_2 \times 1$ vector $\boldsymbol{a}'$ are placed back into the $N_1 \times N_2$ image block in the same scanning order.

Besides scrambling v lowest coefficients in $\boldsymbol{\alpha}$, the authors of [41] also suggested another encryption operation: swapping coefficients that correspond to the same eigenvalues. This swapping operation is dependent on the fact that $\lambda_{j_1,j_2}^{(2D)} = \lambda_{j_2,j_1}^{(2D)}$ and is unavailable in 1-D case. In [41], it was not explicitly how to perform the swapping operation on valid coefficients. However, if all pairs of coefficients with the same eigenvalue are swapped, then an attacker can swap all of them to totally cancel this encryption operation. So, in this paper, we assume that a secret pseudo-random keystream is used to randomly select some pairs of coefficients for swapping.

When the scrambling scheme is exerted on digital images with L gray scales, the input and output have to be calibrated to make the scrambling more efficient. Assuming that one plain-block in the plain-image is I and the corresponding cipher-block is I' (both are $N_1N_2 \times 1$-vectors), the encryption process becomes

$$I' = \mathrm{round}\left(\frac{\mathcal{M}\left(I - 2/L\right)}{\gamma} + 2/L\right), \tag{5}$$

where $\mathrm{round}(\cdot)$ converts the input real number into the nearest integer in $\{0, \cdots, L-1\}$ and

$$\gamma = \max_n \left(\sum_{j=0}^{N_1N_2-1} |\mathcal{M}_{n,j}|\right). \tag{6}$$

Accordingly, assuming the recovered image is $\hat{I}$, the decryption process is as follows

$$\hat{I} = \mathrm{round}\left(\mathcal{M}^T\left(\gamma\left(I' - 2/L\right)\right) + 2/L\right). \tag{7}$$

Due to the rescaling effect embedded in the encryption/decryption processes, it is obvious that the original plain-image cannot be exactly recovered in most cases. Another problem is that the use of γ may enlarge the noise added to I'. In [41], experimental results were given to show that γ may not be determined by Eq. (6), and an "optimal" value was found for a set of test images: $\gamma = 3$, in the sense of MAE (mean absolute error) and MSE (mean squared error).

On the choice of the secret sub-matrix $\mathbf{M}_v$ that scrambles the v lowest coefficients in $\boldsymbol{\alpha}$, Van De Ville et al. suggested deriving it from a Hadamard matrix $\mathbf{H}$, which is a $v \times v$ $(-1, 1)$-matrix[3] whose rows and columns are orthogonal [44]. Its inverse matrix is $\mathbf{H}^{-1} = \frac{1}{v}\mathbf{H}^T$. By permuting the rows/columns of $\mathbf{H}$, and/or multiplying some rows/columns by -1, one can get an H-equivalent matrix. In this way one can get $(v!2^v)^2$ H-equivalent matrices, where some of them are identical. Each H-equivalent matrix $\mathbf{H}^*$ can be scaled to get a secret sub-matrix $\mathbf{M}_v = \frac{1}{\sqrt{v}}\mathbf{H}^*$.

Generally the plain-image is much larger than $N_1 \times N_2$, so there are many $N_1 \times N_2$ blocks for encryption. To further enhance the security of the image scrambling scheme, in [41] it was also suggested that one change the secret matrix for each block, under the control of a cryptographical pseudo-random number generator (PRNG). In this case, the key of the scrambling scheme becomes the seed of the PRNG. To facilitate the following cryptanalysis, we use change_key=1 to denote this encryption configuration and change_key=0 to denote the basic configuration with a fixed secret matrix.

III. Cryptanalysis

Cryptanalysis is a major part of modern cryptology and focuses on the security analysis of different kinds of cryptographical algorithms [5]. Generally, the following four types of attacks should be considered when evaluating the security of a cryptosystem:

- *ciphertext-only attack*, in which an attacker can only observe a number of ciphertexts;
- *known-plaintext attack*, in which an attacker can observe a number of plaintexts and the corresponding ciphertexts;
- *chosen-plaintext attack*, in which an attacker can deliberately choose a number of plaintexts and observe the corresponding ciphertexts;
- *chosen-ciphertext attack*, in which an attacker can deliberately choose a number of ciphertexts and observe the corresponding plaintexts.

Among the four attacks, ciphertext-only attack is the simplest one and every cryptosystem should resist this attack. The other three attacks are much more advanced, but become more and more popular in today's digital and networked world. Known-plaintext attack is very common for modern ciphers, since most binary files and data packets transmitted over network have some fixed segments, such as the leading headers and frequently-used syntax elements. Chosen-plaintext and chosen-ciphertext attacks are possible when the attacker gets a temporary access to the encryption/decryption machine, or he/she can seduce the target user to store some chosen files or transmit some chosen data. An imaginary scenario of chosen-plaintext attack is as follows: Eve sends an interesting photo to Alice, who encrypts it with her secret key and then forwards it to Bob for sharing, which makes it possible for Eve to mount a chosen-plaintext attack after he observes the encrypted photo on the public transmission channel.

If a cryptosystem can resist only ciphertext-only attack, it has to be used very carefully to avoid any possibility of the other three attacks. In this section, we will given a comprehensive investigation on the security of the image scrambing scheme in [41] against all the four attacks.

Throughout this section, without loss of generality, we employ the scrambling parameters used in Sec. V of [41] for demonstrating the experimental results: $N_1 = N_2 = N = 8$, $W = 0.25$, $v = 8$, $\gamma = 3$. The secret matrix and the random swapping operations are both controlled by the `rand()` function with a random seed. All the experiments were made with Mathwork's Matlab 6.5, based on a series of programs derived from the reference codes that Dr. Van De Ville (the first the author of [41]) sent to us.

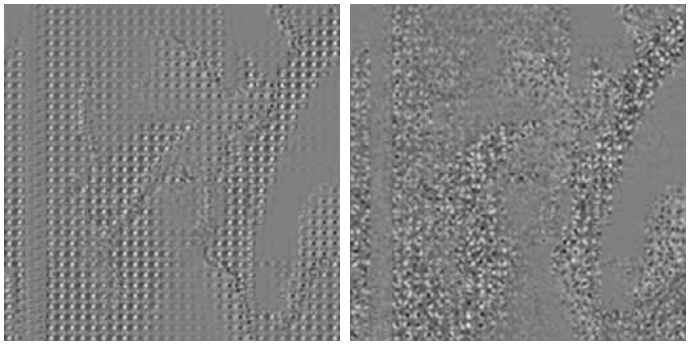
A. Ciphertext-Only Attack

At first, let us see the encryption performance of the image scrambling scheme. For a 256×256 plain-image with 256 gray scales (i.e., $L = 256$) shown in Fig. 1, the encryption results when change_key=0 and 1 are given in Figs. 2a and 2b, respectively. It can be seen that some smooth areas in the plain-image is still recognizable after encryption. This problem was also noticed by the authors of [41] and considered as a minor security problem that can be further remedied with some other techniques.

In the following let us investigate how to get more visual information from the cipher-images than that leaking in Fig. 2. There are several different ways to do this task.

[3]The order of a Hadamard matrix (i.e., the value of v) cannot be an arbitrary value, but be 1, 2, or $4n$, where $n \in \mathbb{Z}$. Here, without loss of generality, we assume that v satisfies this requirement.

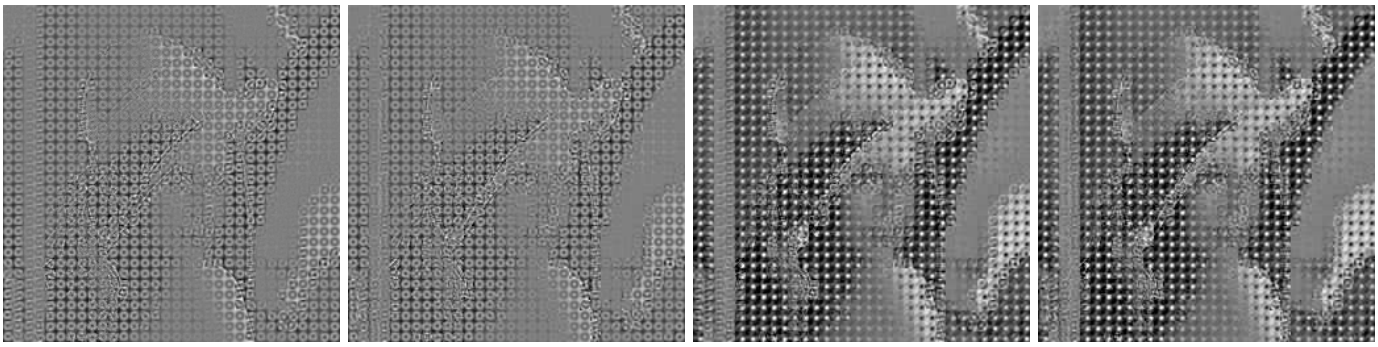
Fig. 1. The plain-image "Lenna".



a) b)

Fig. 2. The encryption results of "Lenna" when change_key=0 (a) and change_key=1 (b).

1) Error-Concealment Based Attack [9]: As the image scrambling scheme is a selective encryption algorithm, we can try to recover the plain-image from these unencrypted coefficients. This error-concealment based attack (ECA in short) is a common attack for all selective encryption methods. As pointed out in [2], [45], for selective encryption based on any orthogonal transform, there is always some visual information leaking from the unencrypted transform coefficients. Though the corresponding energy of these unencrypted coefficients may be rather small, some important visual information may still be distinguished by human eyes. It is true that 2-D DPSS also form an orthogonal transform, so an attacker can try to carry out an ECA on the image scrambling scheme by setting the v scrambled low coefficients to be some fixed values. For the cipher-images shown in Fig. 2, the breaking results are shown in Fig. 3 when the fixed values are 0 and α_v (the lowest unencrypted coefficient). Comparing Figs. 2 and 3, one can see that a rough view of the original plain-image has emerged.



a) $\mathrm{MAE} = 44.0384$ b) $\mathrm{MAE} = 43.5369$ c) $\mathrm{MAE} = 33.6299$ d) $\mathrm{MAE} = 33.0244$

Fig. 3. The breaking performance of ECA on the cipher-images shown in Fig. 2 (measured by MAE – mean absolute error), by setting the $v = 8$ scrambled low coefficients of each block as follows: a-b) $\alpha_0 = \cdots = \alpha_7 = 0$; c-d) $\alpha_0 = \cdots = \alpha_7 = \alpha_8$. The left column (a,c) corresponds to the case of change_key=0, and the right column (b,d) to change_key=1.

The breaking results shown in Fig. 3 can be further enhanced by investigating the statistical relationship between

the average values of all the coefficients in α. For the plain-image shown in Fig. 1, we calculated the histograms of all the 2-D DPSS coefficients and the lowest 10 densities are given in Fig. 4. Among the 10 lowest 2-D DPSS coefficients, one can see that the mean values of α_1, α_2, α_3, α_6, α_7 and α_9 are all close to 0, while those of α_0, α_4, α_5 and α_8 are not (see also Fig. 5 for a plot of all the $N^2 = 64$ mean values). Dividing all the mean values by that of α_8, we can get 64 ratios, $\{r_i = E(\alpha_i)/E(\alpha_8)\}_{i=0}^{63}$, as shown in Fig. 6. If these ratios keep stable for most natural images, one can use them to statistically optimize the breaking performance of ECA. For 1,200 natural images falling into four different categories, "people", "wild animals", "textures" and "city life and China", we calculated the mean values and covariance of $\{r_i\}_{i=0}^{63}$ as shown in Fig. 7. One can see that the mean value of these ratios are really stable for the 1,200 test images (though the variances is not very small for some ones): $r_0 \approx 2.8$, $r_4 \approx r_5 \approx 1.68$, $r_1 \approx r_2 \approx r_3 \approx r_6 \approx r_7 \approx 0$.

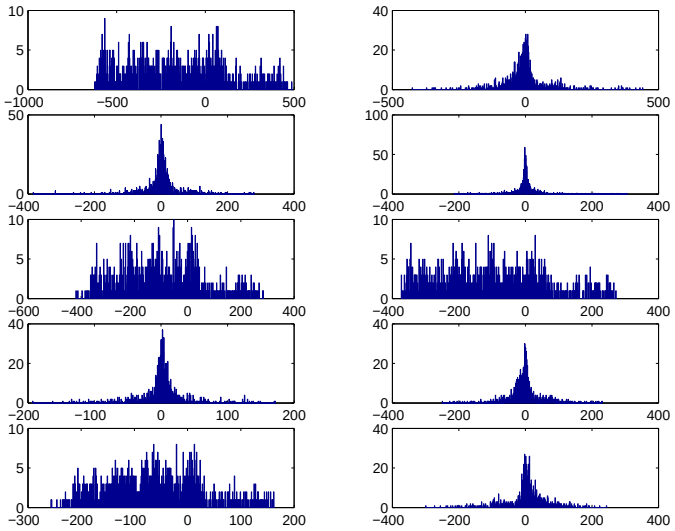


Fig. 4. The histograms of $\alpha_0 \sim \alpha_9$ estimated from all blocks in the plain-image "Lenna" (order: from left to right, from top to bottom).

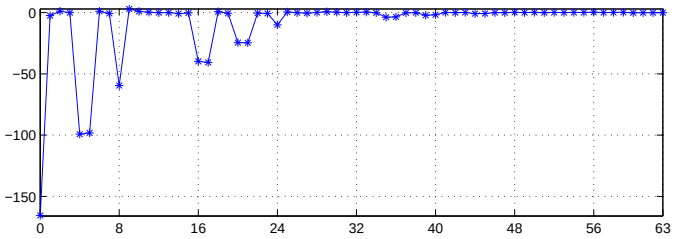


Fig. 5. The mean values of the 64 2-D DPSS coefficients of all blocks in the plain-image "Lenna".

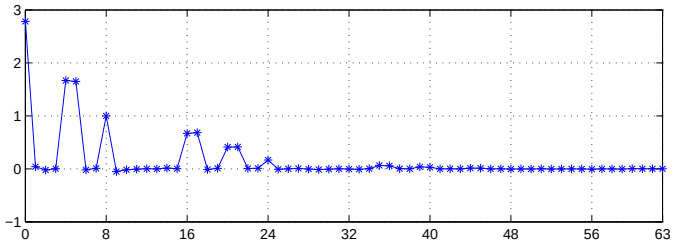
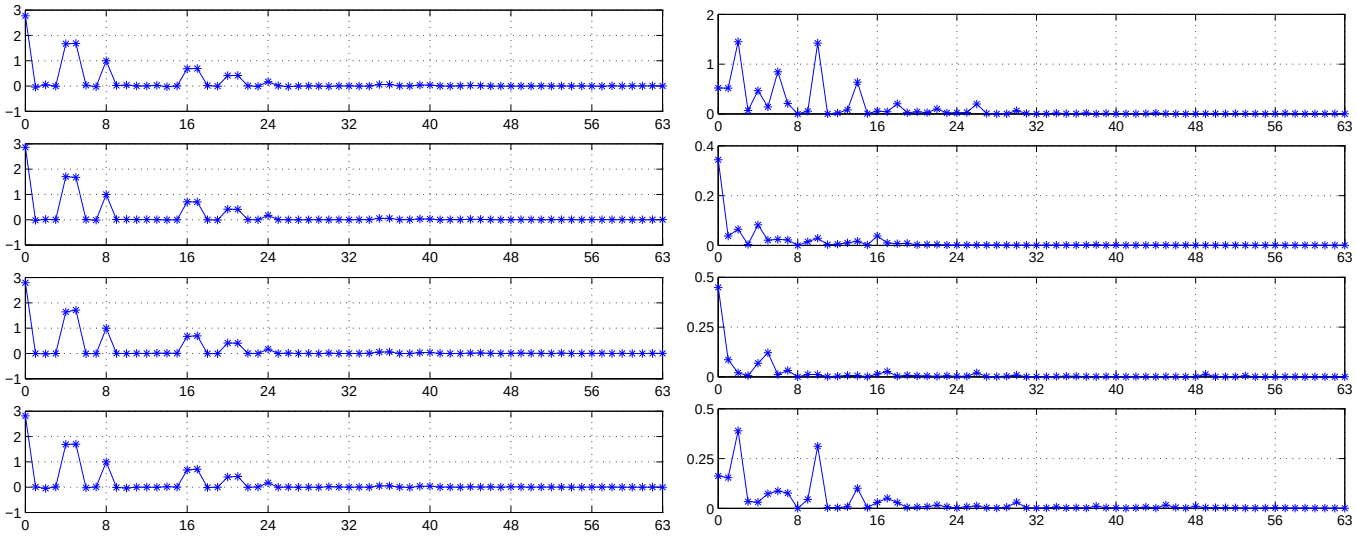


Fig. 6. The ratios of the mean values of the 64 2-D DPSS coefficients to the mean value of α_8 of all blocks in the plain-image "Lenna".

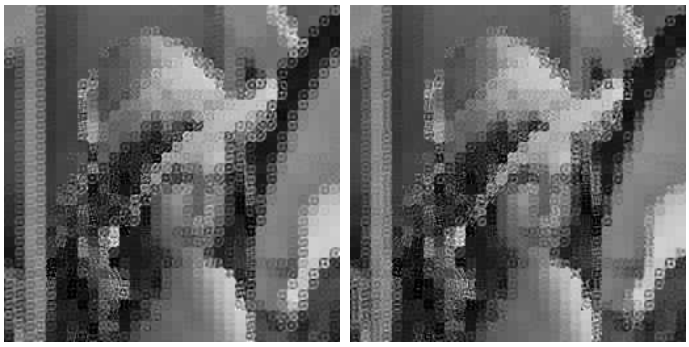
The above fact implies that the following setting of the 8 lowest coefficients is optimal to achieve the best breaking performance of ECA in a statistic sense: $\alpha_0 = 2.8\alpha_8$, $\alpha_4 = \alpha_5 = 1.68\alpha_8$, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_6 = \alpha_7 = 0$.



a) Mean values b) Variances

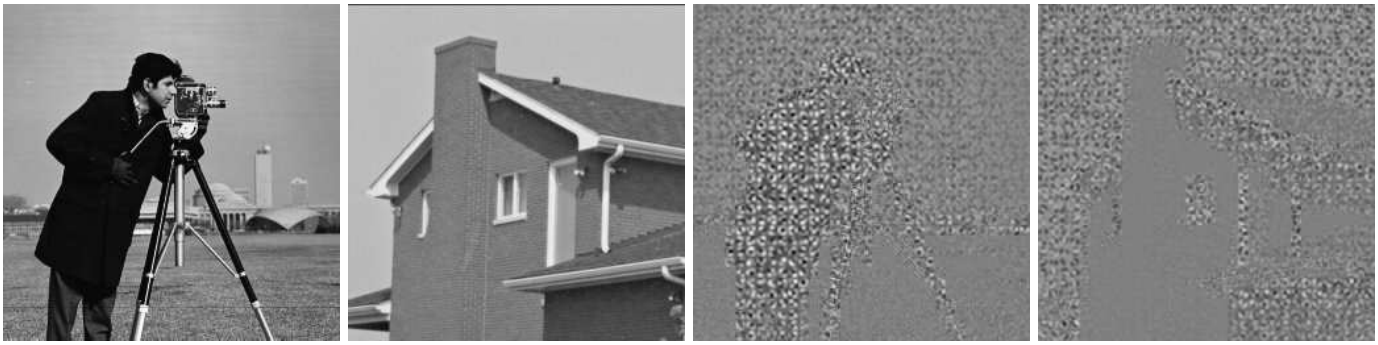
Fig. 7. The mean values and covariances of $\{r_i = E(\alpha_i)/E(\alpha_8)\}_{i=0}^{63}$ of 1,200 test images in four different categories (from top to bottom: "people", "wild animals", "textures", "city life and China").

For the two cipher-images in Fig. 2, the performance of such an optimized ECA is shown in Fig. 8. For another two plain-images and their cipher-images (see Fig. 9), the results of the optimized ECA are given in Fig. 10. Considering the variances of r_i shown in Fig. 7b, in a real attack one can further adjust the values of r_1, r_4 and r_5 to get a better breaking result.



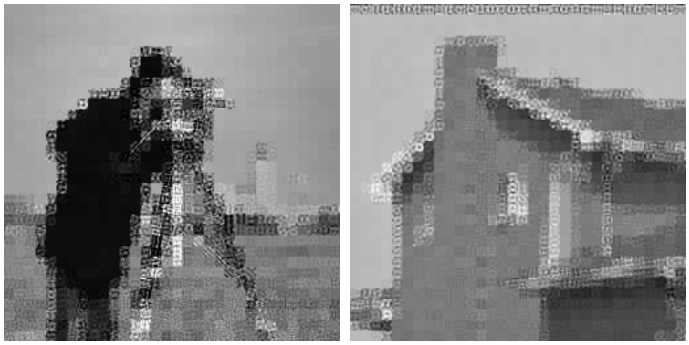
a) MAE $= 18.2867$ b) MAE $= 17.4056$

Fig. 8. The optimized ECA of the cipher-images shown in Fig. 2.



a) "cameraman" b) "house" c) scrambled "cameraman" d) scrambled "house"

Fig. 9. Two plain-images, "cameraman" and "house", and the corresponding cipher-images when change_key=1.



a) MAE $= 15.9883$ b) MAE $= 12.3437$

Fig. 10. The optimized ECA for the cipher-images corresponding to the two plain-images shown in Fig. 9.

2) Breaking Random Swapping: When change_key=0, the random swapping coefficients may be exhaustively guessed and then verified by observing the breaking performance of the optimized ECA. When $v = 8$, there are $8(8-1)/2-6 = 22$ pairs of coefficients with equal eigenvalues in $\{\alpha_i\}_{i=8}^{63}$, so the complexity of guessing all random coefficients operations is not greater than $O(2^{22})$. Considering the 2-D DPSS coefficients $\{\alpha_i\}_{i=32}^{63}$ play a minor role in representing the visual information of an image, one can only guess the random swapping coefficients in $\{\alpha_i\}_{i=8}^{31}$, in which only $14-6 = 8$ valid pairs of coefficients for possible swapping. In this case, the guessing complexity is reduced to be $O(2^8)$ and becomes feasible for an attacker to carry out in practice. Figure 11 shows the breaking result when all random swapping operations of $\alpha_8 \sim \alpha_{31}$ are removed. Comparing Fig. 11 with Fig. 8a, one can see that the former contains more recognizable details along edges.

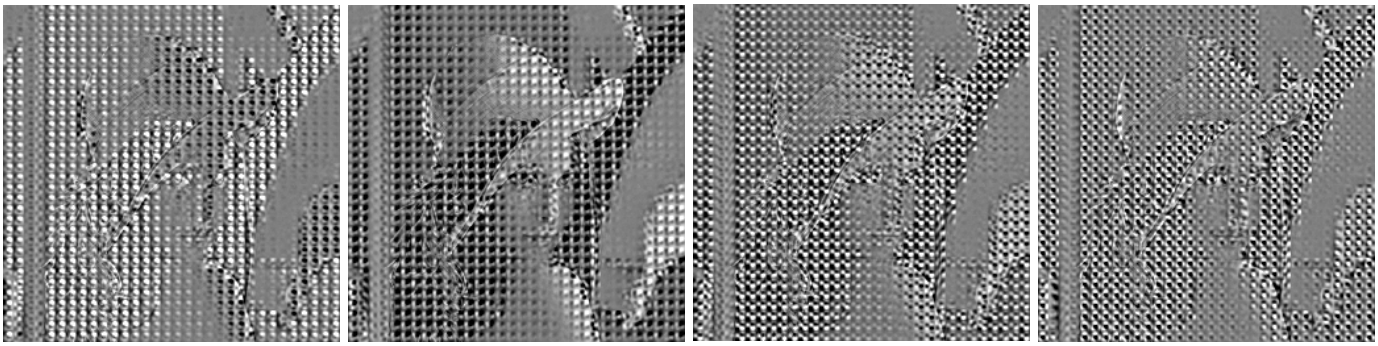
Fig. 11. The breaking performance of the optimized ECA when random swapping operations of $\alpha_8 \sim \alpha_{31}$ are removed: MAE=14.2199.

To overcome this security defect, one can either enlarge the block size or always set change_key=1. The latter remedy is better since it works for any block size.

3) Insecurity of Hadamard-Based Matrices: When the secret sub-matrix $\mathbf{M}_v$ is generated from a $v \times v$ Hadamard matrix $\mathbf{H}$ as suggested in [41], our experiments showed that the decryption is not sufficiently sensitive to the key mismatch, which is an undesirable property for a good cryptosystem and generally leads to a dramatic reduction of the key space [5]. In our experiments, we exerted some fundamental matrix transformations on the original sub-matrix $\mathbf{M}_v$ to get some mismatched matrices, which are then used as a replacement of $\mathbf{M}$ to decrypt the cipher-image Fig. 2a. Some selected results are given in Fig. 12, from which one can see many severely mismatched keys can recover the plain-image with an acceptable quality.

The low sensitivity of decryption to key mismatch means that a randomly-generated key may be capable to roughly recover the plain-image. Figure 13 gives the best recovery result of one experiment, in which 100 randomly-generated keys were used to decrypt the cipher-image Fig. 2a. By testing 100,000 random keys, an estimated PDF (probability density function) of MAE was obtained as shown in Fig. 14. From this empirical PDF, we can calculate that the probability of MAE$\leq$35 is about 0.03. Thus, it is a high-probability event to get a similar result like Fig. 13 by guessing 100 random keys, making the random-guess attack feasible in practice. In some sense we can say that the size of the Hadamard-based key space is dramatically reduced to be smaller than 100.

a) MAE=34.8139 b) MAE=12.2794 c) MAE=37.3327 d) MAE=30.8321



e) MAE=58.8623 f) MAE=38.3956 g) MAE=49.2140 h) MAE=52.1581

Fig. 12. The decryption results (when change_key=0) corresponding to the plain-image "Lenna" with some mismatched keys by processing the sub-matrix $\mathbf{M}_v$ as follows: a) reversing the signs of all elements; b) swapping Rows 1, 8; c) swapping Columns 1, 8; d) swapping Rows 1, 8 and Columns 1, 8; e) reversing the order of all rows; f) reversing the order of all rows and the signs of all elements; g) reversing the order of all columns; h) reversing the order of all rows and that of all columns.

Fig. 13. The best recovery result of the plain-image "Lenna" decrypted with 100 randomly-generated keys: MAE=27.9677.

Note that this security flaw is not so severe when change_key=1. In this case, each block is encrypted by different secret matrices. If the number of blocks in a plain-image is sufficiently large, it will be impossible to randomly guess all the secret matrices to reveal the whole image. Of course, it remains possible for an attacker to guess several selected blocks and roughly recover a small windows of the plain-image.

B. Known-Plaintext Attack

In known-plaintext attack, one can get a number of plain-images and the corresponding cipher-images. According to Eq. (5), the encryption matrix $\mathcal{M}$ can be derived as follows when N_1N_2 known plain-blocks form an invertible $N_1N_2 \times N_1N_2$ matrix $(\boldsymbol{I} - 2/L)$:

$$\mathcal{M} = \gamma(\boldsymbol{I}' + \Delta_{\boldsymbol{I}'} - L/2)(\boldsymbol{I} - 2/L)^{-1}, \tag{8}$$

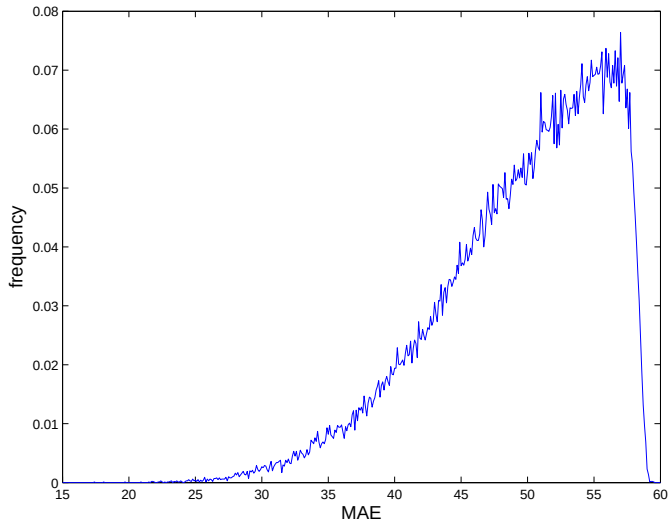


Fig. 14. The empirical PDF of MAE of decrypting the cipher-image Fig. 2a, estimated from 100,000 random keys.

where $\boldsymbol{I}'$ is the $N_1N_2 \times N_1N_2$ cipher-matrix corresponding to $\boldsymbol{I}$, and $\Delta_{\boldsymbol{I}'}$ denotes the error matrix induced by the $\mathrm{rand}(\cdot)$ function. Apparently, ignoring the error matrix $\Delta_{\boldsymbol{I}'}$, $\mathcal{M}$ can be estimated by the following equation:

$$\widetilde{\mathcal{M}} = \gamma(\boldsymbol{I}' - L/2)(\boldsymbol{I} - 2/L)^{-1}, \tag{9}$$

and the estimation error is

$$\Delta_{\mathcal{M}} = \widetilde{\mathcal{M}} - \mathcal{M} = \gamma\Delta_{\boldsymbol{I}'}(\boldsymbol{I} - 2/L)^{-1}. \tag{10}$$

Then, $\widetilde{\mathcal{M}}^T$ can be used as a replacement of $\mathcal{M}^T$ for decryption.

It is not easy to theoretically analyses the relationship between $\boldsymbol{I}$ and $\Delta_{\mathcal{M}}$ (i.e., the relationship between $\boldsymbol{I}$ and the decryption performance of $\widetilde{\mathcal{M}}$), so we made a large number of experiments to investigate the real decryption performance of $\widetilde{\mathcal{M}}^T$ by choosing some sets of N_1N_2 plain-blocks to construct $\boldsymbol{I}$. In the following, we report our experimental results for two different cases according to the value of change_key.

1) When change_key=0: In this case, all blocks of a plain-image are encrypted with the same matrix $\mathcal{M}$, so generally one known plain-image is enough for an attacker to choose many sets of N_1N_2 plain-blocks, some of which may correspond to a good estimation of $\mathcal{M}$ (i.e., to an acceptable recovery performance of any given plain-image).

When the known plain-image is "Lenna" (Fig. 1), the best results of decrypting the cipher-image of "Lenna" (Fig. 2a) in two separate attacks are given in Fig. 15, with 1,000 and 10,000 set of N_1N_2 blocks[4], respectively. One can see that the decryption performance is good enough to reveal almost all visual information in the plain-image.

Further experiments showed that some known plain-images can even get a much better performance than "Lenna". When the known plain-image "Lenna" is replaced by another two images shown in Fig. 16, respectively, the decryption results of the cipher-image Fig. 2a are given in Fig. 17. It can be seen that the decryption performance is nearly perfect.

2) When change_key=1: In this case, each block of a plain-image corresponds to a distinctive encryption matrix $\mathcal{M}_{i,j}$, so one plain-image is not capable of supporting the known-plaintext attack. Instead, $m > N_1N_2$ plain-images should be known such that for each block N_1N_2 plain-blocks (one in each plain-image) can be chosen to estimate each $\mathcal{M}_{i,j}$. When $m = 200$, for the four different categories of natural images used in last subsection, "wild animals", "people", "textures", and "city life and China", we tested the decryption performances of the known-plaintext attack with 500 valid sets of N_1N_2 plain-blocks for each $\mathcal{M}_{i,j}$. The decryption results are shown in Fig. 18, ranked by MAE. By choosing more valid sets of N_1N_2 plain-blocks for each encryption matrix $\mathcal{M}_{i,j}$, the performance can be further improved.

[4]To ensure the invertibility of the formed matrix $\boldsymbol{I} - L/2$ and increase the attacking efficiency, in our experiments we first ranked all valid blocks by their variances and then randomly chose the N_1N_2 blocks from the 100 ones with larger variances for attacking. A similar but slightly different measure was also used for the experiments given in next sub-subsection when change_key=1.

a) MAE=19.7005 b) MAE=16.5549

Fig. 15. The best decryption results in two experiments of known-plain attack with the known plain-image "Lenna" (when change_key=0): a) 1,000 sets of N_1N_2 plain-blocks; b) 10,000 sets of N_1N_2 plain-blocks.

a) b)

Fig. 16. Another two images for testing the performance of known-plaintext attack when change_key=0.

a) MAE=7.4056 b) MAE=5.6989

Fig. 17. The decryption results of known-plain attack when when Figs. 16a and 16b serve as the known plain-image, respectively, where change_key=0 and 1,000 sets of N_1N_2 plain-blocks are processed.

C. Chosen-Plaintext Attack

Compared with known-plaintext attack, in chosen-plaintext attack one can freely choose some plain-blocks to optimize the breaking performance. Now let us choose $\boldsymbol{I} - L/2 = s\mathbf{I}$, where $\mathbf{I}$ denotes the $N_1N_2 \times N_1N_2$ identity matrix. Then, Eq. (10) can be simplified as follows:

$$\Delta_{\mathcal{M}} = \gamma \Delta_{\boldsymbol{I}'}(s\mathbf{I})^{-1} = \frac{\gamma \Delta_{\boldsymbol{I}'}}{s}. \tag{11}$$

On the range of each element in $\Delta_{\boldsymbol{I}'}$, we have the following proposition.

Proposition 1: $\forall(i,j)$, $|\Delta_{\boldsymbol{I}'}(i,j)| \leq 1/2$.

Proof: Observing Eq. (11), one can deduce that

$$\boldsymbol{I}' + \Delta_{\boldsymbol{I}'} - L/2 = \frac{\mathcal{M}(\boldsymbol{I} - L/2)}{\gamma} = \frac{s\mathcal{M}}{\gamma}.$$

a) MAE=32.9514 b) MAE=17.4211 c) MAE=7.2946 d) MAE=6.3547

Fig. 18. The decryption results of known-plain attack when change_key=1, with $m = 200$ known plain-images lying in four different categories of natural images: a) "people", b) "wild animals", c) "city life and China", d) "textures". For each value of (i, j), 500 valid sets of plain-blocks are chosen to estimate each $\mathcal{M}_{i,j}$.

Since $\mathcal{M}$ is an orthogonal matrix, so $-1 \leq \mathcal{M}(i, j) \leq 1$. Thus,

$$\frac{-|s|}{\gamma} + L/2 \leq \boldsymbol{I}'(i, j) + \Delta_{\boldsymbol{I}'}(i, j) \leq \frac{|s|}{\gamma} + L/2.$$

Considering $|s| \leq L/2$ and $\gamma \geq 1$, one has

$$\boldsymbol{I}'(i, j) + \Delta_{\boldsymbol{I}'}(i, j) \in [0, L - 1],$$

which immediately leads to the fact that $|\Delta_{\boldsymbol{I}'}(i, j)| \leq 1/2$ and proves this proposition. ■

Then, from Proposition 1, one can get

$$|\Delta_{\mathcal{M}}(i, j)| = \left| \frac{\gamma \Delta_{\boldsymbol{I}'}(i, j)}{s} \right| \leq \frac{\gamma}{2|s|}, \tag{12}$$

which means that the best breaking performance is reached when $|s|$ is maximized, i.e., $s = -L/2$ when $\boldsymbol{I} = (1 - \mathbf{I})L/2$. With this chosen value of s, some experiments have been made to confirm this theoretical result as shown in Fig. 19. Note that this attack needs only $N_1 N_2$ plain-blocks when change_key=0 and $N_1 N_2$ plain-images when change_key=1.

a) MAE=2.3627 b) MAE=2.4110

Fig. 19. The decryption results of chosen-plain attack when $s = -L/2 = -128$ under the following two cases: a) change_key=0; b) change_key=1.

D. Chosen-Ciphertext Attack

In this attack, one can choose cipher-images instead of plain-images, so the target for reconstruction changes from $\mathcal{M}$ to its transpose matrix $\mathcal{M}^T$. When $N_1 N_2$ cipher-blocks form an invertible $N_1 N_2 \times N_1 N_2$ matrix $\boldsymbol{I}' - L/2$, one can get the following equation from Eq. (7):

$$\mathcal{M}^T = \frac{(\hat{\boldsymbol{I}} - 2/L + \Delta_{\hat{\boldsymbol{I}}})(\boldsymbol{I}' - 2/L)^{-1}}{\gamma}. \tag{13}$$

Removing the quantization error $\Delta_{\hat{\boldsymbol{I}}}$, one has

$$\widetilde{\mathcal{M}^T} = \frac{(\hat{\boldsymbol{I}} - 2/L)(\boldsymbol{I}' - 2/L)^{-1}}{\gamma}, \tag{14}$$

and

$$\Delta_{\mathcal{M}^T} = \frac{\Delta_{\hat{\boldsymbol{I}}}(\boldsymbol{I}' - 2/L)^{-1}}{\gamma}. \tag{15}$$

Similarly, choosing $\boldsymbol{I}' - 2/L = s\mathbf{I}$, one further gets

$$\Delta_{\mathcal{M}^T} = \frac{\Delta_{\hat{\boldsymbol{I}}}(s\mathbf{I})^{-1}}{\gamma} = \frac{\Delta_{\hat{\boldsymbol{I}}}}{s\gamma}. \tag{16}$$

On the range of each element in $\Delta_{\hat{\boldsymbol{I}}}$, we have the following proposition.

Proposition 2: $\forall(i,j)$, $|\Delta_{\hat{\boldsymbol{I}}}(i,j)| \leq 1/2$ if and only if

$$-\frac{L+1}{2\gamma\mathcal{M}^T(i,j)} \leq s \leq \frac{L-1}{2\gamma\mathcal{M}^T(i,j)}. \tag{17}$$

Proof: From Eq. (7), one has

$$\hat{\boldsymbol{I}} + \Delta_{\hat{\boldsymbol{I}}} = \mathcal{M}^T(\gamma(\boldsymbol{I}' - L/2)) + L/2 = s\gamma\mathcal{M}^T + L/2.$$

Note that $|\Delta_{\hat{\boldsymbol{I}}}(i,j)| \leq 1/2$ if and only if $-1/2 \leq \hat{\boldsymbol{I}}(i,j) + \Delta_{\hat{\boldsymbol{I}}}(i,j) \leq (L-1) + 1/2$, which is equivalent to $-1/2 \leq s\gamma\mathcal{M}^T(i,j) + L/2 \leq L - 1/2$. Solving the two inequalities, this proposition is proved immediately. ■

From the above proposition, when $-\frac{L+1}{2\gamma\max(\mathcal{M}^T)} \leq s \leq \frac{L-1}{2\gamma\max(\mathcal{M}^T)}$,

$$|\Delta_{\mathcal{M}^T}| = \left|\frac{\Delta_{\hat{\boldsymbol{I}}}}{s\gamma}\right| \leq \frac{1}{2|s|\gamma}. \tag{18}$$

When $s > \frac{L-1}{2\gamma\max(\mathcal{M}^T)}$ or $s < -\frac{L+1}{2\gamma\max(\mathcal{M}^T)}$, it is not easy to directly estimate the range of $|\Delta_{\mathcal{M}^T}|$, but it is expected to be greater than $\frac{1}{2|s|\gamma}$. For a randomly generated key, Figure 20 gives the experimental relationship between the value of s and the breaking performance of chosen-ciphertext attack, from which one can see that the best decryption performance is achieved when $|s| \approx 47$ (see Fig. 21 for the decryption results when $s = -47$). Due to the symmetry of curve shown in Fig. 20, in a real attack one can try all positive (or negative) values of s to determine an optimal value as the outcome of the cryptanalysis. This means that this attack needs no more than $N_1N_2L/2$ plain-blocks when change_key=0 and $N_1N_2L/2$ plain-images when change_key=1. Note that in most cases it is actually sufficient to achieve a nearly optimal result by fixing the value of $|s|$ around 50. In this case, only N_1N_2 plain-blocks/images are enough to support this attack.

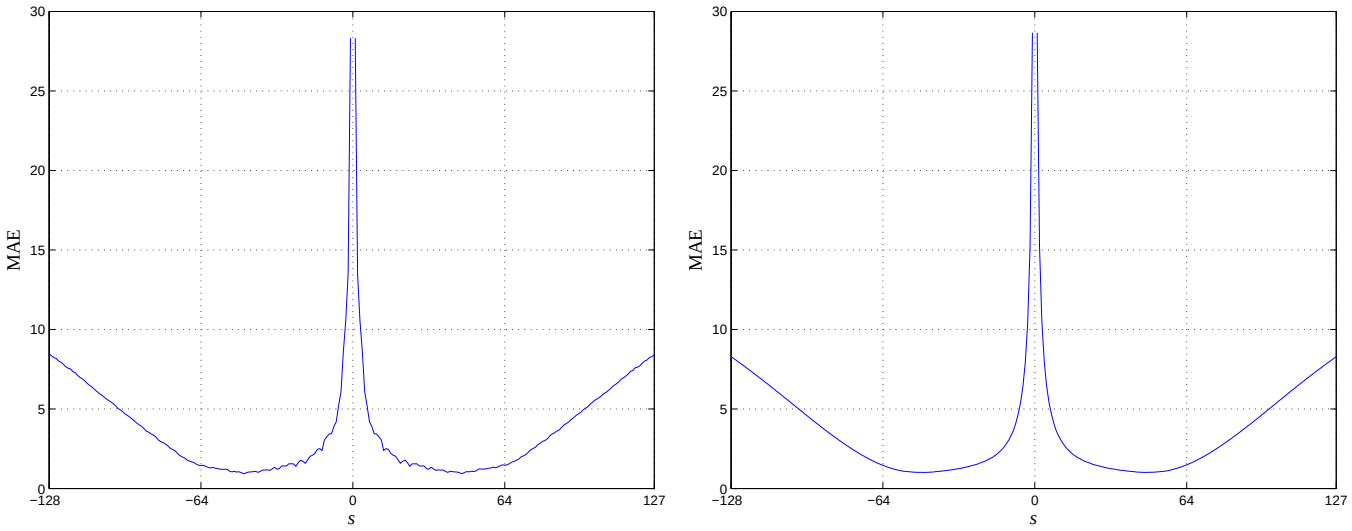
IV. Discussion

Recalling the cryptanalysis given in last section, one can see that one of essential reasons of all the attacks is due to the low sensitivity of the decryption to key mismatch. Actually, this feature is not specific for the original 2-D DPSS basis set shown in Fig. 1 of [41]. We also tested some other basis sets and similar results have been reached (but with some differences on the details, such as the histograms of $\alpha_0 \sim \alpha_9$ shown in Fig. 4). This implies that the low sensitivity to key mismatch is a common feature of all orthogonal transforms[5]. It can be explained by the low sensitivity of matrix computation to small quantization errors and the marvelous capability of human eyes to resist noises in natural images.

In Tables I and II, we give a summary of all the cryptanalytic results obtained through last section. It is clear that the image scrambling scheme proposed in [41] is not secure against all the four types of attacks, regardless it does not suffer from two security flaws when change_key=1.

From the experimental results given in last section, the breaking performance of the four attacks can be ranked as follows (from the best to the worst): chosen-ciphertext attack > chosen-plaintext attack > known-plaintext attack > ciphertext-only attack (error-concealment based attack). For the worst attack – error-concealment based attack,

[5]It is an analogue of the fact that selective encryption working with any orthogonal transform cannot conceal all visual information [2], [45].



a) change_key=0 b) change_key=1

Fig. 20. The experimental relationship between the value of s and the breaking performance of chosen-ciphertext attack (measured by MAE), when the test plain-image is "Lenna".

a) MAE=0.9937 b) MAE=1.0094

Fig. 21. The decryption results of chosen-ciphertext attack (measured by MAE) when $s = -47$ and the test plain-image is "Lenna": a) change_key=0; b) change_key=1.

TABLE I

A BRIEF SUMMARY OF CRYPTANALYTIC RESULTS GIVEN IN SECTION III.

	change_key=0	change_key=1
error-concealment based attack (ciphertext-only attack)	insecure	insecure
random swapping breaking	insecure	secure
insecurity of Hadamard-based key	insecure	secure
known-plaintext attack	insecure	insecure
chosen-plaintext attack	insecure	insecure
chosen-ciphertext attack	insecure	insecure

TABLE II

THE NUMBERS OF PLAINTEXTS/CIPHERTEXTS NEEDED IN SOME ATTACKS.

	change_key=0 (blocks)	change_key=1 (images)
known-plaintext attack	$O(N_1N_2)$	
chosen-plaintext attack	N_1N_2	
chosen-ciphertext attack	$\leq N_1N_2L/2$	

only a low-resolution view of the plain-image can be successfully recovered, and most high-resolution details are missing (see Fig. 8). As a result, we have the following recommendations on how to apply the image scrambling scheme in real applications:

- Use it ONLY for the purpose of perceptual encryption.
- NEVER use the same key to encrypt more than one plain-image[6]. Or, NEVER repeatedly use the same key for more than one plain-images if known/chosen-plaintext or chosen-ciphertext attack is available.
- ALWAYS set change_key=1 if the secret matrix is generated from a Hadamard matrix.

Perceptual encryption is a technique of multimedia encryption that is used to degrade the perceptible quality of multimedia data, under the control of a secret key K and a quality-degradation factor q [2], [8]. Here, the secret key is used to avoid any illegal attempt of reconstructing the multimedia data at a higher quality, and the quality-degradation factor determines the degradation degree induced by the perceptual encryption. Apparently, for the image scrambling scheme under study, the degradation on the visual quality of the plain-image should not be measured by the cipher-image, but by the recovered plain-image via the optimized ECA (error-concealment based attack) discussed in Sec. III-A.1.

Despite of the above security problems and limitations, this image scrambling scheme has some advantages on realizing a **lossy** perceptual encryption scheme, i.e., an encryption scheme that works well with any lossy compression algorithm. This is mainly because that this scrambling scheme does not incur significant bandwidth expansion, which is generally not true for many other image encryption schemes. Our experiments showed that the encryption has only a negligible influence on the compression efficiency of a standard JPEG algorithm, as expected from the bandwidth preservation feature. Another important factor is that the decryption is not very sensitive to errors in cipher-images, due to the same reason that it is not very sensitive to key mismatch as discussed in last section. Figure 22 gives the decryption results when the cipher-image Fig. 2b is compressed by the standard JPEG algorithm with the parameter "Quality" equal to $40 \sim 90$, respectively. One can see that the lossy compression really leads to a lossy decryption result, but the recovery performance remains acceptable as long as the compression ratio is not very high.

a) MAE=11.6895 b) MAE=10.1958 c) MAE=8.9733

d) MAE=7.6228 e) MAE=6.1377 f) MAE=4.2761

Fig. 22. The lossy decryption results when the cipher-image Fig. 2b is compressed by the standard JPEG algorithm: a) Quality=40; b) Quality=50; c) Quality=60; d) Quality=70; e) Quality=80; f) Quality=90.

[6]To do so, a key-management system is generally needed to generate the secret key for each plain-image [5].

V. Conclusion

This paper presents a comprehensive investigation on the security of an image scrambling scheme recently proposed in [41]. As a result, it has been found that the image scrambling scheme is not sufficiently secure against various types of attacks: ciphertext-only attack, known-plaintext attack, chosen-plaintext attack, chosen-ciphertext attack. We also pinpointed another two major security flaws when a fixed secret matrix is used to encrypt the whole image. Based on the cryptanalytic results, we conclude that the image scrambling scheme can only be used to realize (lossless or lossy) perceptual encryption, instead of providing a full protection on all visual information in the plain-image.

Acknowledgement

This research was partially supported by The Hong Kong Polytechnic University's Postdoctoral Fellowships Program under grant no. G-YX63.

References

[1] B. Furht, D. Socek, and A. M. Eskicioglu, "Fundamentals of multimedia encryption techniques," in *Multimedia Security Handbook*, B. Furht and D. Kirovski, Eds. CRC Press, 2004, ch. 3, pp. 93–132.

[2] S. Li, G. Chen, and X. Zheng, "Chaos-based encryption for digital images and videos," in *Multimedia Security Handbook*, B. Furht and D. Kirovski, Eds. CRC Press, 2004, ch. 4, pp. 133–167, preprint is available at http://www.hooklee.com/pub.html.

[3] A. Uhl and A. Pommer, *Image and Video Encryption: From Digital Rights Management to Secured Personal Communication*. Boston: Springer Science + Business Media Inc., 2005.

[4] B. Javidi, *Optical and Digital Techniques for Information Security*. New York: Springer Science + Business Media Inc., 2005.

[5] B. Schneier, *Applied Cryptography – Protocols, Algorithms, and Souce Code in C*, 2nd ed. New York: John Wiley & Sons, Inc., 1996.

[6] National Institute of Standards and Technology (US), "Specification for the advanced encryption standard (AES)," Federal Information Processing Standards Publication 197 (FIPS PUB 197), November 2001.

[7] L. Qiao, "Multimedia security and copyright protection," Ph.D. dissertation, Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA, 1998.

[8] S. Li, G. Chen, A. Cheung, and B. Bhargava, "On the design of perceptual mpeg-video encryption algorithms," arXiv e-print, cs.MM/0501014, available at http://arxiv.org/abs/cs.MM/0501014, 2005.

[9] J. Wen, M. Severa, W. Zeng, M. H. Luttrell, and W. Jin, "A format-compliant configurable encryption framework for access control of video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 545–557, 2002.

[10] Y. Matias and A. Shamir, "A video scrambing technique based on space filling curve (extended abstract)," in *Advances in Cryptology – Crypto'87*, ser. Lecture Notes in Computer Science, vol. 293, 1987, pp. 398–417.

[11] N. Bourbakis and C. Alexopoulos, "Picture data encryption using scan patterns," *Pattern Recognition*, vol. 25, no. 6, pp. 567–581, 1992.

[12] A. Kudelski, "Method for scrambling and unscrambling a video signal," U.S. Patent 5375168, 1994.

[13] F. F. transform based speech encryption system, "Design and cryptanalysis of transform-based analog speech scramblers," *IEE Proc. I – Comm., Speech & Vision*, vol. 138, no. 3, pp. 215–223, 1991.

[14] B. Goldburg, S. Sridharan, and E. Dawson, "Design and cryptanalysis of transform-based analog speech scramblers," *IEEE J. Select. Areas Commun.*, vol. 11, no. 5, pp. 735–744, 1993.

[15] L. Tang, "Methods for encrypting and decrypting MPEG video data efficiently," in *Proc. 4th ACM Int. Conference on Multimedia*, 1996, pp. 219–229.

[16] K.-L. Chung and L.-C. Chang, "Large encrypting binary images with higher security," *Pattern Recognition Letters*, vol. 19, no. 5-6, pp. 461–468, 1998.

[17] S. U. Shin, K. S. Sim, and K. H. Rhee, "A secrecy scheme for MPEG video data using the joint of compression and encryption," in *Information Security: Second Int. Workshop (ISW'99) Proc.*, ser. Lecture Notes in Computer Science, vol. 1729, 1999, pp. 191–201.

[18] W. Zeng and S. Lei, "Efficient frequency domain selective scrambling of digital video," *IEEE Trans. Multimedia*, vol. 5, no. 1, pp. 118–129, 2003.

[19] M. Bertilsson, E. F. Brickell, and I. Ingemarson, "Cryptanalysis of video encryption based on space-filling curves," in *Advances in Cryptology – EuroCrypt'88*, ser. Lecture Notes in Computer Science, vol. 434, 1989, pp. 403–411.

[20] J.-K. Jan and Y.-M. Tseng, "On the security of image encryption method," *Information Processing Letters*, vol. 60, no. 5, pp. 261–265, 1996.

[21] M. G. Kuhn, "Analysis for the nagravision video scrambling method," Online document, available at http://www.cl.cam.ac.uk/~mgk25, 1998.

[22] H. C. H. Cheng, "Partial encryption for image and video communication." Master's thesis, Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada, Fall 1998.

[23] T. Uehara and R. Safavi-Naini, "Chosen DCT coefficients attack on MPEG encryption schemes," in *Proc. IEEE Pacific-Rim Conference on Multimedia (IEEE-PCM'2000)*, 2000, pp. 316–319.

[24] C.-C. Chang and T.-X. Yu, "Cryptanalysis of an encryption scheme for binary images," *Pattern Recognition Letters*, vol. 23, no. 14, pp. 1847–1852, 2002.

[25] S. Li, C. Li, G. Chen, D. Zhang, and N. G. Bourbakis, "A general cryptanalysis of permutation-only multimedia encryption algorithms," IACR's Cryptology ePrint Archive: Report 2004/374, available online at http://eprint.iacr.org/2004/374, 2004.

[26] P. Refregier and B. Javidi, "Optical image encryption based on input plane and Fourier plane random encoding," *Optics Letters*, vol. 20, no. 7, pp. 767–769, 1995.

[27] X. Tan, O. Matoba, T. Shimura, K. Kuroda, and B. Javidi, "Secure optical storage that uses fully phase encryption," *Applied Optics*, vol. 39, no. 35, pp. 6689–6694, 2000.

[28] A. Carnicer, M. Montes-Usategui, S. Arcos, and I. Juvells, "Vulnerability to chosen-cyphertext attacks of optical encryption schemes based on double random phase keys," *Optics Letters*, vol. 30, no. 13, pp. 1644–1646, 2005.

[29] X. Peng, P. Zhang, H. Wei, and B. Yu, "Known-plaintext attack on optical encryption based on double random phase keys," *Optics Letters*, vol. 31, no. 8, pp. 1044–1046, 2006.

[30] ——, "Known-plaintext attack on double random encoding encryption technique," *Acta Physica Sinica*, vol. 55, no. 3, pp. 1130–1136, 2006.

[31] J. Scharinger, "Fast encryption of image data using chaotic Kolmogorov flows," *J. Electronic Imaging*, vol. 7, no. 2, pp. 318–325, 1998.

[32] J. Fridrich, "Symmetric ciphers based on two-dimensional chaotic maps," *Int. J. Bifurcation and Chaos*, vol. 8, no. 6, pp. 1259–1284, 1998.

[33] Y. Mao, G. Chen, and C. K. Chui, "A symmetric image encryption scheme based on 3D chaotic cat maps," *Chaos, Solitons & Fractals*, vol. 21, no. 3, pp. 749–761, 2004.

[34] S. Li and X. Zheng, "Cryptanalysis of a chaotic image encryption method," in *Proc. IEEE Int. Symposium on Circuits and Systems*, vol. II, 2002, pp. 708–711.

[35] ——, "On the security of an image encryption method," in *Proc. IEEE Int. Conference on Image Processing*, vol. 2, 2002, pp. 925–928.

[36] C. Li, S. Li, D. Zhang, and G. Chen, "Cryptanalysis of a chaotic neural network based multimedia encryption scheme," in *Advances in Multimedia Information Processing - PCM 2004: 5th Pacific Rim Conference on Multimedia, Tokyo, Japan, November 30 - December 3, 2004. Proceedings, Part III*, ser. Lecture Notes in Computer Science, vol. 3333. Springer-Verlag, 2004, pp. 418–425.

[37] S. Li, C. Li, G. Chen, and X. Mou, "Cryptanalysis of the RCES/RSES image encryption scheme," IACR's Cryptology ePrint Archive: Report 2004/376, available online at http://eprint.iacr.org/2004/376, 2004.

[38] C. Li, X. Li, S. Li, and G. Chen, "Cryptanalysis of a multistage encryption system," in *Proc. IEEE Int. Symposium on Circuits and Systems*, 2005, pp. 880–883.

[39] C. Li, S. Li, G. Chen, G. Chen, and L. Hu, "Cryptanalysis of a new signal security system for multimedia data transmission," *EURASIP J. Applied Signal Processing*, vol. 2005, no. 8, pp. 1277–1288, 2005.

[40] C. Li, S. Li, D.-C. Lou, and D. Zhang, "On the security of the Yen-Guo's domino signal encryption algorithm (DSEA)," *J. Systems and Software*, vol. 79, no. 2, pp. 253–258, 2006.

[41] D. V. D. Ville, W. Philips, R. V. de Walle, and I. Lemanhieu, "Image scrambling without bandwidth expansion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 6, pp. 892–897, 2004.

[42] A. D. Wyner, "An analog scrambling scheme which does not expand bandwidth, Part I: Discrete time," *IEEE Trans. Inform. Theory*, vol. 25, no. 3, pp. 261–274, 1979.

[43] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty–V: The discrete case," *Bell Syst. Tech. J.*, vol. 57, no. 5, pp. 1371–1430, 1978.

[44] E. W. Weisstein, "Hadamard matrix," From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/HadamardMatrix.html, 2005.

[45] C.-P. Wu and C.-C. J. Kuo, "Design of integrated multimedia compression and encryption systems," *IEEE Trans. Multimedia*, vol. 7, no. 5, pp. 828–839, 2005.