An Improved Joint Fingerprinting and Decryption Content Security Protection Scheme for Visual Media

Yanyan Xu Zhengquan Xu Yuxia Zhang
State Key Lab of Information Engineering in Surveying, Mapping & Remote Sensing, WuHan University, WuHan
E-mail: xuyy@whu.edu.cn, xuzq@whu.edu.cn, zhangyuxia0819@163.com

***Abstract*—Only combining encryption with fingerprinting together can provide comprehensive content security protection for visual media. Joint fingerprinting and decryption (JFD) framework solves encryption and fingerprinting simultaneously and has high efficiency, but several problems still remain to be tackled in JFD, including poor encryption security, severe fingerprinted image distortion, etc. An improved JFD scheme is presented in the paper, where a new encryption strategy based on selective content encryption is put forward to enhance security, and a new method for choosing fingerprint embedding area by structural distortion is proposed to reduce fingerprint's influence to image quality. The experiment results show our scheme's effectiveness.**

I. Introduction

While the rapid development of network technology and information technology contributes to the further application of visual media, and makes it easy to be transmitted, it also causes media data easy to be copied and spread out. Therefore effective methods should be taken to protect the content security of visual media.

Two requirements should be satisfied in data's content security protection, that is, security in the process of data transmission or storage, and security in the process of data usage [1, 2]. Encryption is considered as a regular method to protect data's confidentiality while transmission and storage, but encryption will lose its capability when encrypted data is decrypted. Therefore data's security in the process of usage cannot be ensured because data can be duplicated and distributed arbitrarily after it is decrypted by legal users. Fingerprint is a kind of technology to solve the security problem after data's decryption, this passive form of protection can be achieved by imperceptibly embedding a unique user-dependent identity number into the digital media. Once users distribute data illegally, the hidden fingerprints can be used to trace the illegal users [3]. But it should be noted that digital fingerprinting is a passive form of security and works only after the content has been received and made available to the user [4]. In a word, security protection based on encryption or fingerprint independently is not complete. Only a combination of encryption and fingerprint can provide comprehensive content security protection for visual media.

Encryption and fingerprinting are regarded as two seemingly orthogonal processing fields [4], many problems will arise if the two operations are just superposed simply. If encryption is done first, the value of cipher text will be modified by fingerprint embedding, and the plaintext is possible to be irretrievable because of the diffusion effect. If fingerprinting is followed by encryption, the plaintext must be fingerprinted, encrypted and transmitted respectively for each user because of fingerprints' uniqueness, which will result in low data processing efficiency and high transmission cost, and it is an unpractical way in many cases, especially for visual media data which is often of large volumes. So the combination of encryption and fingerprinting has been the key issue in the field.

At present, research on the integration of two processes can be classified into three types: transmitter-side encryption and fingerprinting, transmitter-side encryption and receiver-side fingerprint embedding, joint fingerprinting and decryption. The first method embeds a unique fingerprint copy for each user, then encrypts and transmits it respectively [3], [5]. This kind of trivial combination of fingerprinting and encryption will lead to low processing efficiency, high transmission bandwidth consuming and poor scalability.

The second type encrypts data at source, decrypts and embeds fingerprints at receiver side by reliable tamper-proof hardware or reliable software engine technology [7, 8]. The problem is its security relying on extra hardware equipments or nodes, and security attacks may be taken place after decryption and before fingerprinting, as a consequence it is not a secure solution.

Another kind of solution is combining decryption with fingerprinting in the receiver-side. Transmitter-side only distributes a single encrypted copy of the media, users decrypt data using different decryption keys and get different fingerprint copies. Anderson proposed a Chameleon scheme, which encrypts plaintext audio data in uncompressed form at source [9], and different users decrypt ciphertext by slightly different keys, then each user can get slightly different least significant bits of the plaintext audio data. Kundur et al. proposed a classic joint fingerprinting and decryption architecture (JFD) [4], the idea is that the perceptually relevant components is encrypted by scrambling at source, and the ciphertext is partially decrypted by receivers such that the un-decrypted parts constitute the fingerprint. In the scheme proposed by Lian [12], data is encrypted by additive modulation, and the cipher-video is decrypted by controllable demodulation under the control of fingerprint codes. A JFD based on vector quantization is proposed by Lin [6], where JFD method is applied to VQ compressed images equipped with the ability to remove noise automatically.

This work is supported by National Basic Research Program of China(No. 2011CB302204) and LIESMARS Special Research Funding.

Joint decryption and fingerprinting, which is characterized by JFD framework can provide comprehensive content security protection, also it has high efficiency both in encryption and fingerprint embedding, and with significant advantages compared to two other types. But in this framework, there exists contradictory relationship between encryption and fingerprinting, and it causes some problems such as lower encryption security, severe distortion in fingerprinted image, etc. These problems reduce its practicability, especially in some fields which have high requirements to precision, e.g., medical and remote sensing image, etc. In this paper, we will present an improved JFD scheme to solve these problems, and enable it suitable for content security protection for visual media.

The organization of the paper is arranged as follows, Section 2 introduces and analyzes a classic JFD scheme, and section 3 proposes our scheme. In Section 4, the experiments and results are analyzed in detail. Finally, some conclusions are drawn and future work is given in Section 5.

II. Overview of Kundur's JFD Scheme

Integrating decryption and fingerprinting in receiver-side is a kind of practical scheme, and the scheme presented by Kundur is a representative one [4]. A detail introduction and analysis of this classic scheme is as below.

The main idea of JFD is , if the multimedia content X is encrypted by K_S , the algorithm of encryption and decryption is denoted as E and D , the encrypted content Y is denoted as $Y = E_{k_s}[X]$, the decryption key set K_i is designed jointly with the K_S , $K_i \subseteq K_S$, $K_i \cong K_S$, users decrypt Y using K_i and get $\hat{X}_{F_i} = D_{k_i}[Y]$, as a result an imperceptible and slightly different fingerprint is embedded in the content after decryption, Fig.1 summarizes the process.

Kurdur's method is applied to encoded image through DCT and quantized. A set of perceptually significant DCT coefficients in the low and midfrequency region are identified and partitioned into n subsets. The members of each subset are all encrypted by sign scrambled. The user is given a unique subset of keys K_i for decrypting only a p/n fraction of the n encrypted subsets, the remaining $k = n - p$ subsets are distinct from each user, and different positions of the undecrypted subsets determine the image copy's uniqueness and constitute the unique fingerprint.

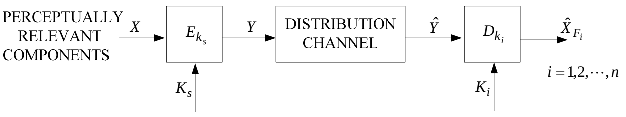
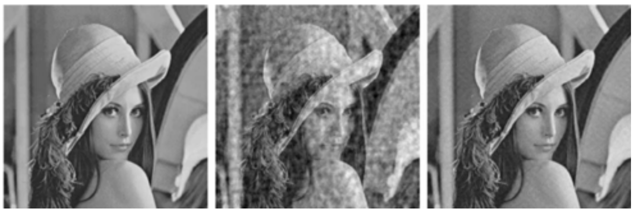


Fig 1. Architecture of JFD

The feature of JFD scheme is, the media is encrypted only once at the source, the same encrypted copy is shared by all users and can be multicast to every user. Each user can get his private fingerprinted copy through decryption naturally. Thus the advantages of JFD include: it resolves the problem of integrating encryption and fingerprinting fundamentally through generating fingerprints by partial decryption, and it is very high efficient because of its "one encryption, different decryption" mechanism which enables the media is encrypted and transmitted only one time instead of encrypted and embedded fingerprint for each user separately.

But three disadvantages also exist in this scheme. The first is the encrypted media content is not secured. Because of partial decryption, in order to decrease its influence to image quality, a lightweight encryption method is used in JFD, and it results in poor encryption effect, and encrypted image is still intelligible, as shown in Fig.2(b). The second, even a lightweight encryption way is used, the quality of fingerprinted image is still not satisfied, as shown in Fig.2(c). This is because JFD executes partial decryption in encrypted media data randomly, which degrades the quality of fingerprinted image, especially when scrambled coefficient subsets k increase or the scrambled coefficients in k increase, the fingerprinted image's quality degrades obviously. The third, the fingerprint's robustness against collusion attack is not considered.

By the detailed analysis of JFD scheme, we can conclude that encryption and fingerprinting are combined effectively in JFD, massive visual media's high requirement to data processing and transmission efficiency also can be satisfied. But the lightweight encryption method makes it vulnerable to attacks in the process of transmission and storage, and the severe fingerprinted image distortion restricts its application in the fields which have high requirement to precision. If these problems can not be solved, JFD scheme's further application will be influenced.



(a) (b) (c)

Fig 2. (a)Original Image (b)Encrypted Image (c)Fingerprinted Image

III. An Improved JFD Scheme

Aiming at JFD's problems, we will present an improved JFD scheme which inherits JFD scheme's high efficiency and overcomes its shortcomings by two ways. That is, a new encryption/decryption strategy based on selective content encryption is proposed to enhance encryption security and solve the contradictory relationship between encryption security and fingerprinted image quality in JFD; a new structural distortion method as the measurement for choosing fingerprint embedding area is proposed to reduce partial decryption's influence on image quality. By these ways, JFD's problems can be solved in order to enable it suits for the content security protection of visual media.

A. *Encryption/Decryption Strategy Based on Selective Content Encryption*

There is a compromise between encryption's security and the fingerprinted image quality in JFD framework. The trouble will be allayed from two aspects. At first, selective content encryption method is used to ensure encryption security, and a partial decryption strategy is put forward to decrease the influence to image quality. Secondly, fingerprint embedding areas are chosen to make an utmost reduction of the partial decryption's influence to image quality.

Considering fingerprinted image quality, DCT coefficients are only sign scrambled during encryption in JFD scheme, which is a lightweight encryption method and clearly has low security. In order to ensure encryption effect, we present a new encryption scheme, which encrypts all perceptually significant DCT coefficients by selective content encryption method. Selective content encryption has become the mainstream way in visual media's encryption, and its security has been proved widely. In order to reduce partial decryption's influence to image quality, we design a strategy where DC coefficients are decrypted entirely, and AC coefficients which affect image quality less are partially decrypted for the purpose of ensuring encryption security as well as fingerprinted image quality.

B. Selection of Fingerprint embedding Area

Because of image's particular structural feature, different regions have different influences to image quality, and it will cause image distortion if encrypted image is decrypted partially without differential. Especially when the parts left encrypted are becoming large, it will result in remarkable increasing of image distortion. Taking account of image pixels' structural change in encrypted image, we propose a new way to choose fingerprint embedding area using visual security assessment based on structural distortion, by which the influence of fingerprint embedding to image quality is greatly decreased.

The visual security assessment to visual media's ciphertext is usually used to evaluate its unintelligible degree, the greater the unintelligible degree, the higher security [13]. According to this feature, a new method for choosing fingerprint embedding area is proposed in our scheme, where sub-image blocks having higher visual confidentiality are decrypted completely, and more intelligible parts in perception are decrypted partially and embedded fingerprints to reduce the image distortion.

The research of visual security can be classified into two types, that is, subjective assessment and objective assessment. The subjective one can be influenced by the measuring environment and subjective sensation, the single use of it is not effective in practical application [13]. A video quality assessment with peak signal noise ratio (PSNR) is considered as an effective objective evaluation of visual security, where PSNR values of cipher-images is used to judge the cipher-images' unintelligible degree [11], [14]. In this method, the differences between image's pixels are only considered and the pixel is regarded as the independent point, while the correlation between adjacent pixels and the structural feature of image are ignored, therefore the PSNR value curve of cipher-images is not consistent with the conclusions of user's subjective judgments in some cases [13].

An objective video quality assessment method based on Structure Similarity (SSIM) is presented by Wang et al. in 2002 [10]. SSIM has a good performance on the evaluation of statistic image quality since image's structure information and its change are well considered in this scheme. Based on the conception of SSIM, Structure Distortion (SD) is introduced and used to choose the fingerprint embedding area in our scheme. The definition of SD is shown as (1):

$$\begin{aligned} SD &= 1 - SSIM \\ &= 1 - \frac{(2\mu_{\mathrm{x}}\mu_{\mathrm{y}} + C_1)(2\sigma_{\mathrm{xy}} + C_2)}{(\mu_{\mathrm{x}}{}^2 + \mu_{\mathrm{y}}{}^2 + C_1)(\sigma_{\mathrm{x}}{}^2 + \sigma_{\mathrm{y}}{}^2 + C_2)} \end{aligned} \quad (1)$$

Where μ_x and μ_y are the mean of image x and image y respectively, σ_x and σ_y are the standard deviation of x and y respectively, and σ_{xy} is the correlation coefficient between x and y. The small constants C_1, C_2 are used to avoid instability of zero denominator, and these coefficients are defined as follows:

$$\mu_x = \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (2)$$

$$\sigma_x = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)^2} \quad (3)$$

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y) \quad (4)$$

In our scheme, the SD value is used to indicate the disordered degree of the structural information in cipher-image, i.e., the unintelligible degree of the cipher-image. The bigger the SD value, the higher the unintelligible degree. Therefore sub-blocks with smaller SD value are chosen as the fingerprint embedding area to form the fingerprints. The detailed process is described as follows:

Assuming the original image is X, X is divided into N sub-image blocks $X_1, X_2, \ldots, X_i, \ldots$, $1 \le i \le N$. The encryption/decryption strategy is applied to each block, then partially decrypted sub-blocks are: $C(X_i) = E_{k_s}[X_i]$. Calculating the SD value of each block and getting $SD[C(X_i)]$. Supposing the threshold value of SD is δ_p, the fingerprint embedding area is H_1 , other area is H_2 , for $\forall X_i$, if $SD[C(X_i)] < \delta_p$, $X_i \in H_1$, and if $SD[C(X_i)] \ge \delta_p$, $X_i \in H_2$.

C. The Construction and Distribution of Decryption Key

In this paper, users get fingerprinted image through partial decryption, thus the decryption key plays an important role in the formation of fingerprint. A key generation and distribution method is presented in this section, where decryption key matrix is constructed according to encryption key, binary random sequence, and the location information of sub-image blocks, then the matrix is distributed securely. The steps are given as follows:

Step1: Partition the image into several sub-image blocks $X_1, X_2, ..., X_M$, where M is the total number of blocks. Generate encryption key of each block randomly, and get a $1\times M$ encryption key matrix $K_E = \left|K_1 K_2 \cdots K_M\right|$;

Step2: Generate binary random sequence w_j^i for each user, $1\le j\le L$, $1\le i\le N$, and L is the length, N is the number of users. Construct a $M\times 1$ binary location matrix I through the location information of sub-image blocks and the value of the random sequence codeword, let $I=\left|I_1 I_2 ... I_j ... I_M\right|^T$, $1\le j\le M$, suppose the fingerprint embedding area is H_1 , other area is H_2 , if I_j lies in H_1 , then it's value is w_j , if I_j lies in H_2 , then it's value is 1, as shown in (5):

$$I_j = \begin{cases} w_j & \text{if } I_j \in H_1 \\ 1 & \text{if } I_j \in H_2 \end{cases} \qquad (5)$$

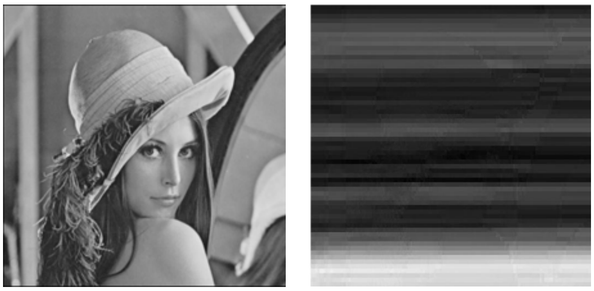
Step3: Calculate decryption key matrix K_{D_i} by $K_{D_i} = K_E \times I$, and use CDMA's(Code Division Multiple Access) theory to distribute it securely through public channel.

IV. Experimental Results and Analysis

Two parts are included in the experiment, the security of encryption scheme is verified in the first part, and the influence of choosing fingerprint embedding area to the image quality is evaluated in the second part.

A. Experiment1: Encryption Security

The 512 × 512 image is divided into 4096 8 × 8 blocks, each block is DCT encoded and quantized, afterwards it is encrypted by the method we described before. Because the security of selective content encryption has been proved widely, so we only give the result of encryption effect, as Fig. 3(b) shows, it can be seen that the encryption effect is good and the encrypted image is unintelligible.

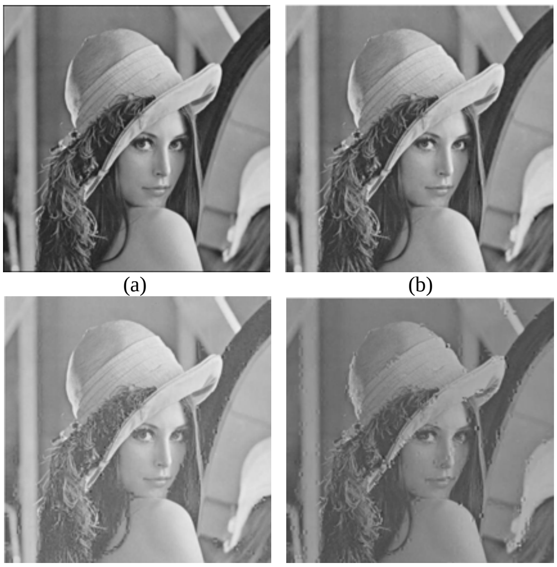


(a) (b)

Fig 3. (a)Original Image (b)Encrypted Image

B. Experiment 2: Choosing Fingerprint Embedding Area

DC coefficients of the sub-blocks in cipher-image are decrypted totally while AC coefficients are kept encrypted; and different methods are used to choose fingerprint embedding area. When 1000 blocks are included in this area, and they are all decrypted partially, the effect of the fingerprinted image is shown in Fig.4, from which we can get the conclusion that using SD value as a measurement obviously has better image quality than choosing fingerprint embedding area randomly or using PSNR value to choose it.



(a) (b)

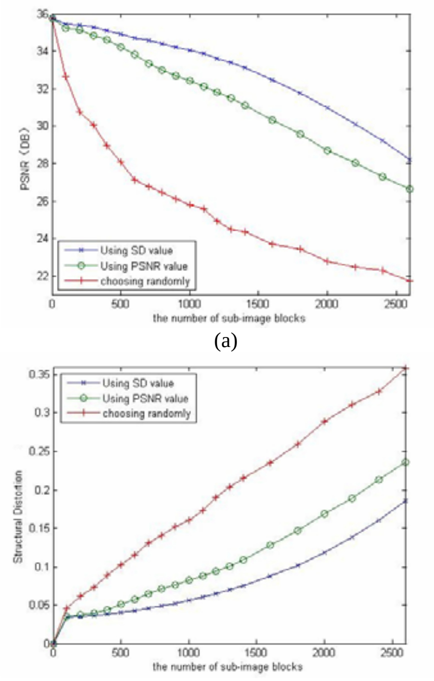
(c) (d)

Fig. 4 Comparison of fingerprinted image quality when the number of blocks in fingerprint embedding area keeps constant: (a)original image (b)using SD to choose(c)using PSNR to choose (d)choosing randomly

Fig. 5 shows with the continuously increasing of sub-blocks in fingerprint embedding area, the comparison of PSNR value and SD value of the fingerprinted image. Fig. 5(a) is the result when PSNR value is used to evaluate the fingerprinted image quality, we can see the image quality is worst when embedding area is chosen in random, it is becoming better when PSNR is used as a measurement, and the image quality is the best when SD value is employed to choose embedding area. Fig. 5(b) shows the result when SD value is used to evaluate the image quality, as can be seen from the figure, by this method the fingerprinted image has the lowest structural distortion.

The experiments we mentioned above are done in one image, we do experiment using 500 different images, get similar results, and have an ideal effects.

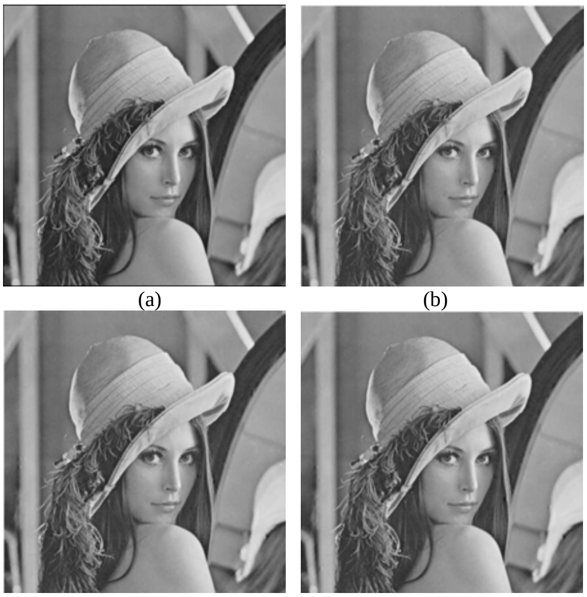
The fingerprinted images are shown in Fig.6, (a) is the original image, (b),(c),(d) are different fingerprinted image copies generated by user50, user100, and user200 respectively, it can be seen that each user's copy only has a slight difference from each other and is similar to original image.



(a)

(b)

Fig 5. Comparison of fingerprinted image quality when sub-blocks in fingerprint embedding area are increasing (a)PSNR value of fingerprinted image (b)SD value of fingerprinted image



(a)

(b)

(c)

(d)

Fig 6. a) Original Image b) Fingerprinted Image of User 50 c) Fingerprinted Image of User 100 d) Fingerprinted Image of User 200

V. CONCLUSION

The biggest advantage of JFD framework is the comprehensive security protection combining encryption with digital fingerprint, and the fingerprint is generated naturally by partial decryption and has a high efficiency. In this paper, aiming at the contradictory relationship between encryption security and fingerprinted image quality, we propose a new encryption/decryption strategy based on selective content encryption to enhance encryption security and solve the contradictory relationship between encryption security and fingerprinted image quality. In order to further reduce partial decryption's influence on image quality, we use the structural distortion as a measurement to choose fingerprint embedding area. The experimental results show that the fingerprinted image still has good quality after high secure selective content encryption and partial decryption. This scheme also preserves enough space for anti-collusion encoding of fingerprint, which is also our future work.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions.

REFERENCES

[1] A. Adelsbach, U. Huber, A. Sadeghi, "Fingercasting-joint fingerprinting and decryption of broadcast messages". In *ACISP 2006. LNCS*, vol.4058, pp136-147

[2] A. Sadeghi. The Marriage of Cryptography and Watermarking — Beneficial and Challenging for Secure Watermarking and Detection, *Lecture Notes in Computer Science*, 2008, vol.5041, pp2-18

[3] D. Boneh and J. Shaw, Collusion-secure fingerprinting for digital data, *IEEE Trans. Inform. Theory*, 1998, vol. 44, pp1897–1905

[4] D. Kundur, K. Karthik, Video Fingerprinting and Encryption Principles for Digital Rights Management, *Proc. IEEE Special Issue on Enabling Security Technologies for Digital Rights Management*, 2004, Vol.92, No.6, pp918-932.

[5] F. Hartung, B. Girod, Digital watermarking of MPEG-2 coded video in the bitstream domain, in *Proc. Int. Conf. Acoustics, Speech and Signal Processing*, 1997, vol.4, pp 2621–2624.

[6] Chih-Yang Lin, Wei-Lun Huang, Tzung-Her Chen. Noise-resistant joint fingerprinting and decryption based on vector quantization. *2010 International Conference on Broadband, Wireless Computing, Communication and Applications*, 2010, pp463-468

[7] J. Bloom, "Security and rights management in digital cinema," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2003, vol.4, pp712–715.

[8] J. Pegueroles, M. Fernández, F. Rico-Novella and M. Soriano, A Practical Solution for Distribution Rights Protection in Multicast Environments, *Computational Science and Its Applications-ICCSA 2006, Lecture Notes in Computer Science*, 2006, Vol. 3982, pp527-536

[9] J. Anderson and C. Manifavas, Chameleon – A new kind of stream cipher, *Proc. 4th Workshop on Fast Software Encryption*, 1997, pp107-113.

[10] Wang Zhou, Lu Liang, Bovik Alan C. Video Quality Assessment using Structural Distortion Measurement. *Proc. 2002 International Conference on Image Processing*, Rochester, NY, 2002, vol.3, pp65-68.

[11] T. Stutz, A. Uhl. On Efficient Transparent JPEG2000 Encryption. *MM&Sec'07*, 2007, Dallas, Texas, USA.

[12] Shiguo Lian, Zhiquan Wang, Collusion-Traceable Secure Multimedia Distribution Based on Controllable Modulation, *IEEE Trans. On Circuits and Systems for Video technology*, 2008, Vol. 18, No.10, pp1462-1467

[13] Jing Sun, Zhengquan Xu, Jin Liu, Ye Yao. An objective visual security assessment for cipher-images based on local entropy, *Multimed Tools Appl*, 2010, vol. 53, No.1, pp75-95

[14] Shiguo Lian, Zhongxuan Liu, Zhen Ren, Haila Wang. Secure Advanced Video Coding Based on Selective Encryption Algorithms. *IEEE Transactions on Consumer Electronics*, 2006, Vol.52, No.2, pp621-629.