J. Vis. Commun. Image R. 25 (2014) 805–813

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

A content security protection scheme in JPEG compressed domain

Yanyan Xu *, Lizhi Xiong, Zhengquan Xu, Shaoming Pan

State Key Lab of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan, Hubei 430079, China

ARTICLE INFO

Article history:
Received 22 August 2013
Accepted 3 January 2014
Available online 16 January 2014

Keywords:
Content security
Compressed domain
Encryption
Fingerprint
Format compliance
Variable length coding
Space mapping
Compressed data stream

ABSTRACT

The access and distribution convenience of public networks opens a considerable content security threat when sending, receiving, and using multimedia information. In this paper, a content security protection scheme that integrates encryption and digital fingerprinting is proposed to provide comprehensive security protection for multimedia information during its transmission and usage. In contrast to other schemes, this method is implemented in the JPEG compressed domain with no transcoding or decompression, therefore, this scheme is highly efficient and suitable for multimedia information, which is seldom available in an uncompressed form. In addition, a variable modular encryption method is proposed to solve the invalid variable length coding (VLC) problem when a compressed data stream is encrypted directly. Experimental results demonstrate improved security and the efficiency provided by the proposed scheme. The experiments also demonstrate imperceptibility and collusion resistance of fingerprints.

1. Introduction

With the rapid development of information and communication technologies, it has become commonplace to distribute multimedia information over the internet. However, this data access and distribution convenience increases security risks when sending and receiving sensitive information and images. Eavesdropping, unauthorized duplication, and release are a growing threat which must be effectively contained. In some instances, exposure and leakage of this information will cause damage to personal privacy or even national security. Therefore, effective content security measures must be adopted to guarantee the safety of sensitive or proprietary multimedia information.

Two requirements must be satisfied to guarantee the security of multimedia information, confidentiality and proper usage [1,2,26]. Encryption is a common method for ensuring data confidentiality, but data security after decryption cannot be ensured because the data can be duplicated and distributed improperly by legal users. Fingerprinting is an emerging technology that imperceptibly embeds a unique user-dependent identity number into media content. If users distribute data improperly, the hidden fingerprints can be extracted from the copied media and used to trace unauthorized users. But, when fingerprinting digital data, the problem of collusion attacks must be considered [3]. Colluders compare their fingerprinted copies of data, modify differences and generate a new copy to avoid discovery. However, digital fingerprinting is a passive form of security and works only after the content is received and has been made available to the user [4]. Therefore, only a combination of encryption and fingerprinting can provide comprehensive content security protection for multimedia information, because both confidentiality and proper usage vulnerabilities are addressed.

Several intermediate goals must be achieved to meet the security requirements for multimedia information content protection and are listed as follows:

1. Encryption security. Different from text/binary encryption, multimedia encryption requires both cryptographic security and perceptual security [5]. The former refers to security against cryptographic attacks, and the latter means that the encrypted multimedia content is unintelligible to human perception.
2. Format compliance. Format information is generated after encoding multimedia data, such as file headers and synchronization information. This information will be used by the decoder to successfully recover data and to keep multimedia communication synchronized [5]; therefore, format information must not be effected by encryption. The ciphertext is considered format-compliant if the encrypted data stream can be decoded by a standard decoder.
3. Imperceptibility. The embedded fingerprint information must be invisible and have little perceptual impact on the image quality.
4. Robustness against collusion attacks. The embedded fingerprint code must be robust against collusion attacks.

* Corresponding author.
E-mail address: xuyy@whu.edu.cn (Y. Xu).

http://dx.doi.org/10.1016/j.jvcir.2014.01.005
1047-3203/

5. Efficiency. Encryption and fingerprinting operations should be highly efficient because multimedia information is huge and the number of users is extremely large. It is widely believed that video information must be distributed efficiently because of its data rate and large size. However, some image information is also massive [15]. For example, a typical hyperspectral remote sensing image covering a small region of a few kilometers contains millions of pixels, and each pixel is represented by several bands [27]. Thus, the data volume can be several GB or even several hundred GB [28].
6. Compression ratio. In all cases, multimedia encryption algorithms should not change compression ratio or should at least keep the changes in a small range [5].
7. Compressed domain implementation. Because most multimedia signals are available in a compressed form, it is desirable to encrypt and embed fingerprints directly into the bitstream of compressed media with no transcoding, decompression or even partial versions of such computations [6–8,15].

The existing research on content security protection for multimedia information can be classified into three types, but the methods discussed have significant shortcomings. The first type [3,24] embeds each user's fingerprint into the plaintext and then encrypts it separately on the sender side, leading to low efficiency and poor scalability. The second type encrypts data on the sender side and embeds fingerprints on the receiver side with tamper-proof hardware [9] or trusting network nodes [10,23]. These solutions can save a lot of computation time and bandwidth usage, but they often prove to be insecure and inflexible in application. Another type of solution integrates decryption and fingerprinting on the receiver side. Anderson proposed a Chameleon scheme [11] that encrypts uncompressed plaintext audio data at the source. Different users decrypt the same ciphertext with slightly different keys and obtain slightly different least significant bits (LSB) of the plaintext audio data. This scheme though, is not very efficient and the fingerprint in LSB is not robust against common signal processing operations. Adelsbach et al. proposed a modified Chameleon scheme in order to embed spread spectrum watermarks [1]. But, this approach still only considers uncompressed baseband signals. Celik et al. also improved the Chameleon scheme by using algebraic operations during encryption/decryption and then embedding robust spread spectrum watermarks [25]. This scheme can be modified to handle joint decryption and watermarking on vector quantized images [26]. Kundur et al. proposed a joint fingerprinting and decryption (JFD) scheme [4] that encrypts perceptually relevant components by scrambling on the sender side, and receivers partially decrypt media content to obtain fingerprinted copies. This scheme is highly efficient, but it also has disadvantages. The encrypted media content is not secure from perception and the robustness against collusion attacks cannot be confirmed. Lemma et al. presented a scheme based on additive encryption [12]; its perceptual security is better than [4] but it is not secured against cryptographic attacks. Furthermore, its robustness against collusion attacks was not a focus of this research. Lian et al. proposed an improved scheme in [13,14], where media content is encrypted by additive modulation. A cipher-video was decrypted by controllable demodulation controlled by fingerprint codes. Although these methods can meet most requirements for multimedia information content protection, they are not implemented in a strictly "compressed domain" environment. Most of these methods can be partially decompressed to gain access to transformational coefficients so they are not strictly compressed domain methods [15]. Only one real compressed domain method has been proposed [16]; however, this method also has some problems, since its encryption security has not been proven, its ciphertext cannot be kept format compliant, and its robustness against collusion attacks has not been tested.

JPEG is a widely used multimedia compression standard. A novel content security protection scheme for the JPEG compressed domain is proposed in this paper. A variable-modular encryption method based on space mapping is used to encrypt a compressed data stream directly, so as to obtain a format compliant ciphertext. A unique fingerprinted image is generated naturally for each user by decrypting the encrypted data stream with different decryption keys. In contrast to other schemes, this scheme is implemented in the JPEG compressed domain with no transcoding or decompression; therefore it is highly efficient and suitable for multimedia information seldom available in an uncompressed form. Experimental results illustrate the security benefits of the proposed scheme, the imperceptibility of fingerprint embedding, and its robustness against collusion attacks.

The organization of this paper is as follows: Section 2 discusses the related research, and Section 3 proposes our scheme. Section 4 provides experimental results and a performance analysis, and Section 5 presents conclusions.

2. Background

The general architecture for JPEG compression is shown in Fig. 1. An original image is first transformed and quantized. The resulting quantized coefficients are further entropy coded to form a compressed stream. According to this process, the potential encryption locations are listed as follows: (A) raw data encryption; (B) transformed coefficient encryption before or after quantization; (C) encryption by entropy encoding; and (D) encryption of the compressed data stream. These encryption locations are shown in Fig. 1.

In raw data encryption, the media data are encrypted before compression. Because the encryption operation changes the adjacent relations of the image pixels, the compression ratio can be decreased greatly and thus format compliance cannot be maintained [17]. The second and third encryption types implement an encryption operation during compression; these techniques are codec dependent. Because the adjacency relations of the transformation coefficients are changed by encryption, the compression ratio is also decreased [18–19]. The fourth type encrypts the compressed data stream directly with some significant advantages. It provides better security because the compressed data has almost no redundancy. The fourth type is also more efficient because the length of the plaintext is shorter than for any of the other types. Its compression ratio and the format compliance are easily kept; and easier to integrate with different application systems because it is codec independent [20]. This approach is the mainstream research direction in the field of visual media encryption.

Digital fingerprinting is a special form of digital watermarking and can be embedded either in a spatial or transform domain. Embedding information in the spatial domain has the problem of insufficient robustness against common operations such as slight noise and compression; therefore, embedding fingerprints in selected coefficients in the transform domain is a more widely used method. However, this operation is not considered to be strictly a "compressed domain" method and will lead to low efficiency. It is therefore highly desirable to develop watermarking algorithms that work entirely in the compressed domain. Until now, few methods have been proposed to embed information directly in the compressed data stream [6–8,15].

According to these reasons, we can draw the conclusion that encryption or fingerprinting of the compressed data stream directly is more suitable because most multimedia signals are transmitted or saved in a compressed form. The efficiency is high because the time-

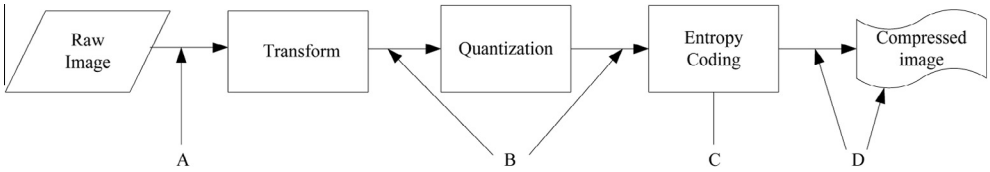


Fig. 1. Encryption locations in the process of compression encoding.

intensive decompression or transcoding processes are also avoided. However, one problem caused by operating on a compressed data stream directly is the difficulty of keeping it format-compliant. In the image and video compression standards, transformation and quantization are followed by a variable length coding (VLC) process to obtain a higher compression ratio. A characteristic feature of a VLC codeword is that its length is variable and the whole codeword space is not occupied; in other words, the aggregate of an n-bit valid VLC codeword is not equal to V_n, and V_n is the aggregate of all of the possible n-bit binary combinations. Taking a 3-bit codeword for example, there are 8 different binary codewords in total; however, only 2 VLC codewords are valid, specifically 100 and 101, as shown in the dark gray region in Fig. 2. The encryption operation for a codeword is actually a mapping from the plaintext space p to ciphertext space $c = E(p)$. Because of the randomness in the encryption operation $E()$, a valid plaintext codeword will most likely map to a random position in ciphertext space, as shown by the dark gray region and translucent gray region in Fig. 2. If a valid plaintext codeword is mapped to the translucent gray region, then the ciphertext will not conform to the syntax of the compression standard and cannot be kept format-compliant.

3. The proposed scheme

This section details the proposed content security protection method that solves the limitations of existing approaches. The general architecture of the proposed scheme is shown in Fig. 3. On the far left, raw image X is compressed and the variable modular encryption is implemented in a compressed data stream on the sender side to obtain a format compliant ciphertext. On the lower right side of the diagram, fingerprinted copies are generated by decrypting the ciphertext with different keys on the client's side. The extraction and detection of fingerprint are analyzed on the lower left side.

3.1. A variable modular encryption method based on space mapping

In JPEG encoding process, after quantization, a DC coefficient and 63 AC coefficients are prepared for entropy encoding. For the DC coefficient, the previously quantized DC coefficient is used to predict the currently quantized DC coefficient. The size of the differences is encoded as a VLC codeword of CODELENGTH bits, followed by binary code of SIZE bits for the amplitude. For the AC coefficients, a composite value is used to describe the run length of the zero coefficients (Run) and the amplitude of the next non-zero coefficient (Level); it is Huffman coded as a CODELENGTH bits VLC codeword followed by the binary code of SIZE bits for the amplitude, which specifies the amplitude of the coefficient. Therefore, regardless of the values for the DC and AC coefficients, the total length of the compressed data stream is TOTALLENGTH = CODELENGTH + SIZE.

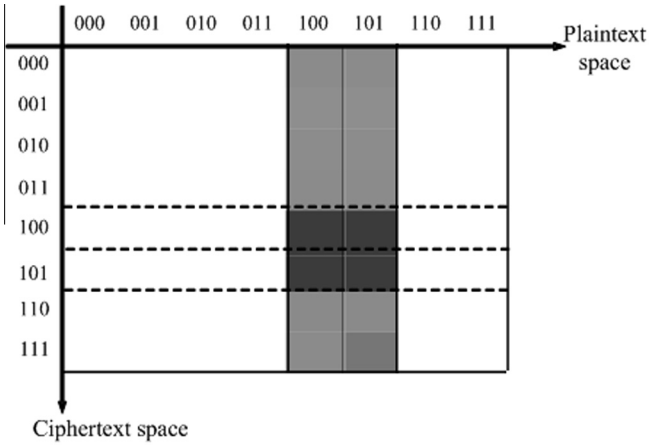


Fig. 2. The mapping from plaintext space to ciphertext space.

According to this entropy encoding process, a variable modular encryption based on space mapping is proposed as follows:

First, different alphabetic tables are constructed. We put all the VLC codewords of the same length into one alphabetic table, and we define the number of VLC codewords as its module M. Table 1 includes several alphabetic tables that are constructed for luminance AC coefficients according to Table K.5 in [21]; because there is only one VLC codeword for total lengths 3, 4, and 5, we put these three codewords into the same alphabetic table.

We assume that the ith element of the alphabetic table corresponds to an independent space $V(i)$ that is composed of n points, i.e., $V(i) = \{P(i+1)\cdots, P(i+n), \cdots\}$ for the alphabetic table having the same TOTALLENGTH, $n \in [1, 2^{SIZE}]$; otherwise $n \in [0, 2^{\text{MAXTOTALLENGTH} - \text{TOTALLENGTH} + \text{SIZE}}]$, where MAXTOTALLENGTH is the element in the alphabetic table that has the maximum TOTALLENGTH. For example, for alphabetic table V with the same TOTALLENGTH, the size of the 1st element is 5, $n = 2^5 = 32$, $V(1) = \{1101000000, 1101000001, \cdots, 1101011111\}$; for alphabetic table I with a another TOTAL LENGTH, the MAX TOTAL LENGTH is 5, and the TOTAL LENGTH of the 1st element is 3, $n = 2^{5-3+1} = 8$; then, 8 points are included in $V(1)$, $V(1) = \{00000, 00001, \cdots, 00111\}$.

For the space $V(i)$, assuming that the original codeword is X_i, the codeword corresponds to a point $P(i+r)$ in $V(i)$ according to the mapping rule $S()$, as shown in Eq. (1):

$$P(i+r) = S(X_i) \tag{1}$$

Then, the point $P(i+r)$ is encrypted with a module add operation by the random integer K_i:

$$P(j+s) = E(P(i+r), K_i) = (P(i+r) + K_i) \, Mod \, M \tag{2}$$

where K_i is a random integer that is distributed uniformly in $[0, M-1]$. The module M of each alphabetic table is different; therefore, the proposed method is called the variable modular encryption method.

After the encryption operation, point $P(i+r)$ in $V(i)$ is mapped into a random point $P(j+s)$ in $V(j)$. $P(j+s)$ corresponds to ciphertext C_j, which is also the jth valid VLC codeword in the same alphabetic table:

$$C_j = S^{-1}(P(j+s)) \tag{3}$$

The encryption process is shown in Fig. 4.

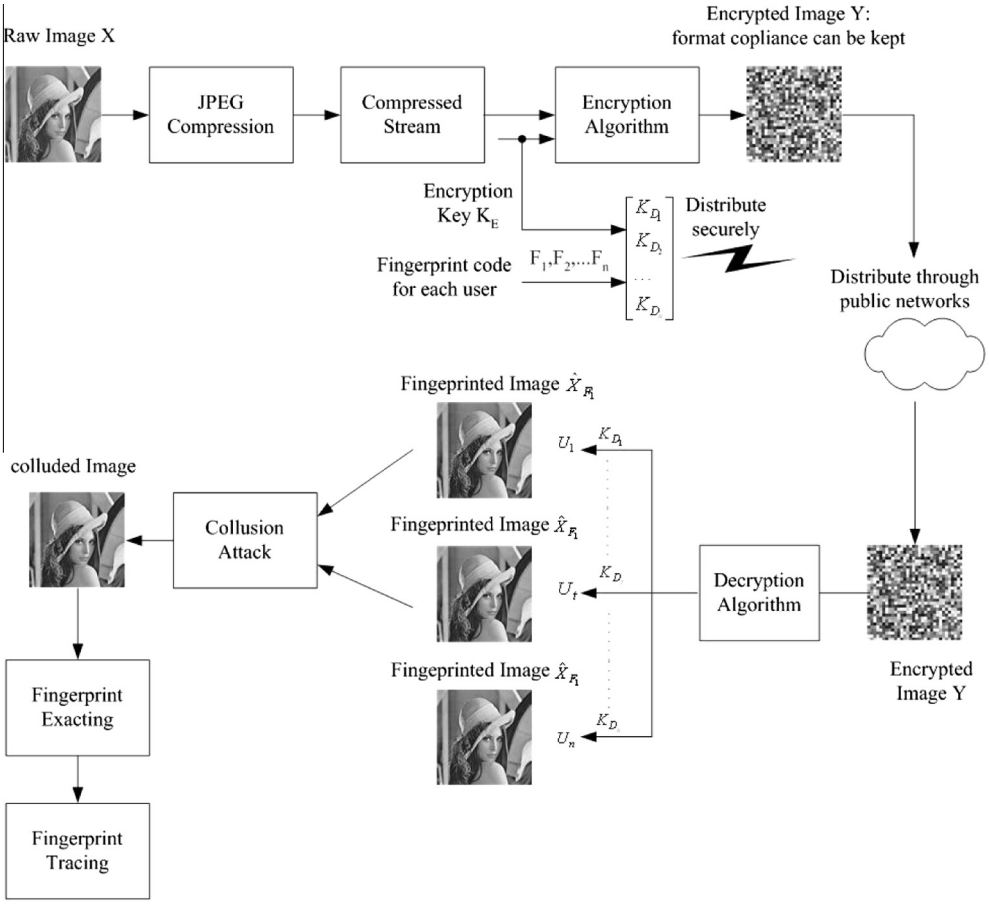


Fig. 3. The general architecture of the proposed scheme.

Table 1
Alphabetic tables for VLC codeword.

Alphabetic table	Run/size	Code length	Code word	Total length	Module
I	0/1	2	00	3	3
	0/2	2	01	4	
	1/1	4	1100	5	
II	0/3	3	100	6	2
	2/1	5	11100	6	
III	1/2	5	11011	7	3
	3/1	6	111010	7	
	4/1	6	111011	7	
IV	0/4	4	1011	8	4
	5/1	7	1111010	8	
	6/1	7	1111011	8	
	7/1	8	11111010	9	
V	0/5	5	11010	10	6
	1/3	7	1111001	10	
	2/2	8	11111001	10	
	8/1	9	111111000	10	
	9/1	9	111111001	10	
	A/1	9	111111010	10	
...	.....	...	...		...

3.2. Joint decryption and fingerprinting

Users decrypt ciphertext C_i with the decryption key K'_i, where $K'_i = K_i - W_i$, and where W_i is the fingerprint sequence. K'_i is generated and transmitted securely by the sender.

After receiving the ciphertext C_j, the user maps C_j to a unique point $P(j+s)$ in the space $V(j)$ according to the same rules as on the sender side, as shown in Eq. (4):

$$P(j+s) = S(C_j) \tag{4}$$

The decryption and fingerprint embedding process is:

$$D(P(j+s), K'_i) == (P(i+r) + K_i - K_i + W_i) \, Mod \, M = P(i+t) \tag{5}$$

$P(i+t)$ is a point in $V(i)$; after re-mapping the compressed stream with the embedded fingerprint, X'_i is generated, as follows:

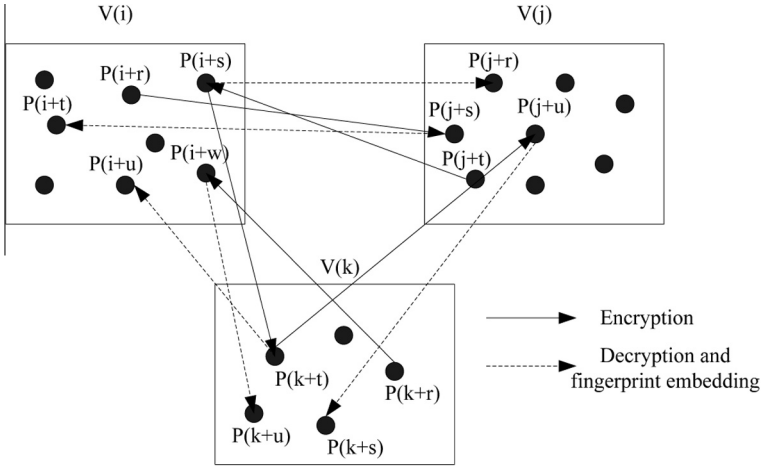


Fig. 4. Variable modular encryption based on space mapping.

$$X_i' = S^{-1}(P(i+t)) \tag{6}$$

X_i' is similar to X_i and is only slightly different from X_i in its amplitude value.

3.3. Fingerprint detection and traitor tracing

Suppose that a colluded stream is z'. We can compare z' with the original stream x, extract the fingerprint sequence, and obtain a colluded fingerprint codeword w'. We correlate w' with each user's fingerprint sequence w_i using Eq. (7):

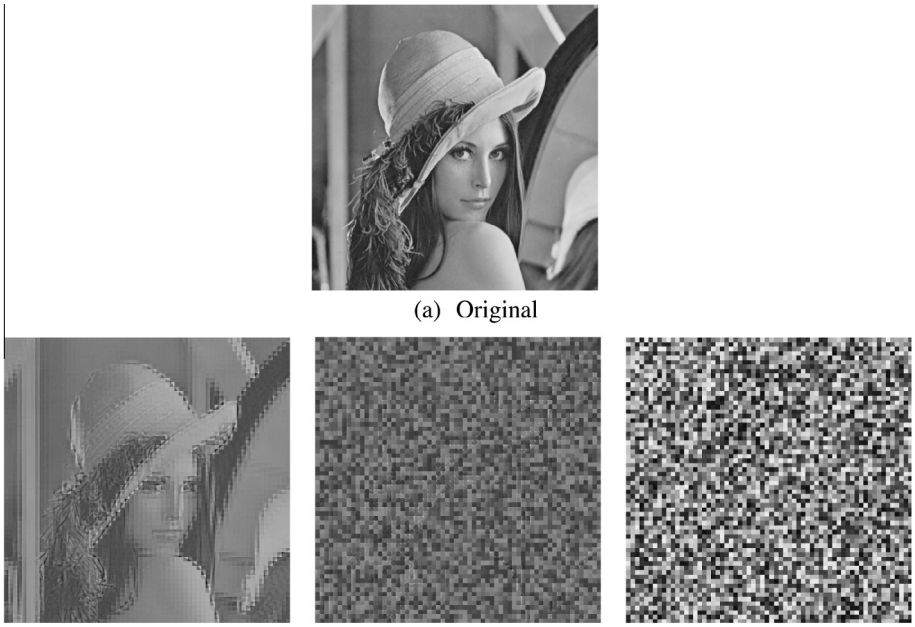
$$T_N(\mathrm{i}) = \frac{w' w_i}{\sqrt{\|w_i\|^2}}, \quad i = 1, 2, \ldots, N \tag{7}$$

The user whose fingerprint has the highest correlation value $T_N(i)$ is identified as the colluder. By averaging collusion, $T_N(i)$ follows an N_u -dimensional Gaussian distribution:

$$\begin{aligned} T &= [T_N(1), \ldots, T_N(N_u)]^T \sim N\left([m_1, m_2]^T, \sigma_d^2\right) \\ m_1 &= \|s\|\left(\frac{1}{c} + \left(1 - \frac{1}{c}\right)\rho\right), \quad m_2 = \|s\|\rho \end{aligned} \tag{8}$$

where m_1 is the mean vector for colluders, m_2 is the mean vector for innocent users, and ρ is the average correlation between two different fingerprints. According to [22], for an L-tuple q-ary Reed-Solomon code with dimension t, $\rho = \frac{t-1}{L}$.

We define the maximum detection statistic for colluders and innocent users as T_1 and T_2, respectively:



(a) Original

(b) Kundur et al's scheme (PSNR = 21.5489 dB)

(c) Lian et al's scheme (PSNR = 12.8220 dB)

(d) proposed scheme (PSNR = 10.1123 dB)

Fig. 5. Encrypted Image Quality.

$$T_1 = \max_{j \in S_c} T_N(j), \quad T_2 = \max_{j \notin S_c} T_N(j) \tag{9}$$

where S_c is the aggregate of colluders. The probability of catching one colluder using the maximum detector can be expressed as:

$$P_d = P_r(T_1 > T_2) \approx \int_{-\infty}^{+\infty} P_r(T_1 > t) f_{T_2}(t)dt = \int_{-\infty}^{+\infty} \left(\int_{x}^{\infty} f_{T_1}(z)dz \right) f_{T_2}(t)dt$$
$$f_{T_i}(x) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x-m_i)^2}{2\sigma_i^2}} \tag{10}$$

where $P_r()$ is the distribution function of T_1, and $f_{T_1}()$ and $f_{T_2}()$ is the probability density function of T_1 and T_2, respectively.

4. The experimental results and analysis

The experiment included three parts: a performance analysis of the encryption, imperceptibility of the fingerprints, and robustness of the fingerprints.

4.1. Performance analysis of the encryption

The performance of the proposed encryption method was analyzed from the standpoint of confidentiality, compression efficiency, and format compliance.

(1) Perceptual security

The proposed scheme, Kundur et al's scheme in [4] and Lian et al's scheme in [14] are used to encrypt the 512×512 Lena image, respectively. Perceptual security was evaluated by the PSNR value. In general, the lower the PSNR value is, the lower the intelligibility of the encrypted information and the higher the perceptual security. The encryption results are shown in Fig. 5.

From this figure, we can see that the proposed scheme has a higher perceptual security than the other two schemes. This is because Kundur et al's scheme encrypts only the sign of the DCT coefficients; therefore, perceptual security is poor. Lian et al's scheme encrypts the DC coefficients and the sign of the AC coefficients; thus, it obtains a higher perceptual security than Kundur et al's scheme. Our proposed scheme encrypts VLC codewords that correspond to all of the DC and AC coefficients, thus it has the highest perceptual security of the three schemes.

Fig. 6 shows quality of recovery images from encrypted images with a typical perceptual attack Error Concealment Attack (ECA) [29]. As the results shown, compared with other schemes, the proposed scheme is more secure and robust again perceptual attack.

Table 2 shows the comparative results of encryption with different size images in the USC-SIPI image database. The results show that Lian et al's scheme has better security than Kundur et al's scheme, while our proposed scheme obtains better perceptual security than the other two schemes.

(2) Cryptographic security

In the proposed encryption method, each VLC codeword is replaced by another valid VLC codeword with the same length under the control of a random sequence K_i which determines the randomness of the ciphertext. If truly random sequences are used in K_i, then the encryption of the VLC codeword is equal to a one-time pad cipher, and is robust against ciphertext-only, known-plaintext, and chosen-plaintext attacks.

The VLC codeword is related not only to the previous codeword but also to the next codeword; thus, knowing only parts of the information could be of little use for unauthorized viewing of the image content. Thus, attackers must analyze all of the data that affects the previous codeword and the next codeword, which will make an exhaustive attack more difficult. Taking the 512∗512 Lena image as an example, the number of encrypted VLC codewords is 262144 and the total length of encrypted VLC codewords is 20581 bytes; if an attacker wants to recover a cipher-image to an intelligible image, the attacker will have to exhaust at least 50% of the key space; then, the number of bits to be exhausted is $20581 \times 8 \times 50\% = 82324$ bits, the number of calculations required

(a) Kundur et al's scheme (PSNR = 23.8807dB) (b) Lian et al's scheme (PSNR = 18.1297dB) (c) Proposed scheme (PSNR = 14.2903dB)

Fig. 6. Quality of recovery image from encrypted image with ECA.

Table 2
Comparison of the Encrypted Image Quality.

File description	Size	Encrypted Image Quality (PSNR)			Recovered Image Quality (PSNR)		
		Lian's	Kundur	proposed	Lian's	Kundur	proposed
Airplane	256 × 256	8.2947	10.0474	8.0232	12.2143	15.6766	9.3566
Aerial	256 × 256	10.9257	12.2578	10.1060	15.8671	17.3904	13.9134
Clock	256 × 256	8.1208	12.8811	7.6901	10.9337	17.3904	9.7407
Chemical plant	256 × 256	12.6905	16.2263	9.9436	15.7578	19.3218	13.3204
Couple	512 × 512	14.8102	16.2980	11.5198	15.0135	17.1532	14.6816
Aerial	512 × 512	9.9296	15.5154	9.3301	12.3162	17.2841	10.7584
Tank	512 × 512	13.2237	15.9683	12.7572	17.0231	19.9842	14.8188
Man	1024 × 1024	12.0986	14.1028	8.9977	15.9341	18.4382	12.1644
Airport	1024 × 1024	14.2722	17.2226	9.6540	16.4676	21.5606	13.4611

Table 3
Compression ratio of encrypted image.

Method	Kundur's	Lian's	Proposed
Changed compression ratio	Unchanged	33%	Unchanged

is $2^{82324} \approx 9.85*10^{24781}$, so the cost spent in exhausting the calculation would be much larger than the value of the image itself. If the attacker wants to recover a clearer image or an encrypted image at a larger size, then the time spent in exhausting the calculations will grow exponentially.

(3) Format compliance of the ciphertext

To achieve format compliance, the original VLC codeword is replaced by another valid VLC codeword with the same length according to the constructed alphabetic table. Even if the decoder does not know the decryption key, it still can find the ending of the codeword and maintain synchronization, and therefore, format-compliance of the cipher-text is maintained. For the few codewords that have a different TOTALLENGTH in the same alphabetic table, for example, the codewords in Alphabetic table I, the extra bits can fill in the end of the shorter codewords, and the TOTALLENGTH of the table can be kept the same.

(4) Compression efficiency

We use the changed compression ratio (CCR) [5] to evaluate compression efficiency. Table 3 shows the comparative results of different methods. Kundur's method did not change the compression ratio because it only flipped the sign of AC coefficients. In the proposed method, the original codeword is encrypted by the replacement of another valid codeword with the same length; therefore, the compression ratio will not be changed. However, In Lian's method, the adjacency relations of the transformation coefficients are changed by encryption and the compression ratio is changed greatly.

4.2. Imperceptibility of fingerprint embedding

Fig. 7 shows a comparison of a fingerprinted image generated by different methods when the length of fingerprint sequence is 1778. As seen in the figure, our scheme obtains better image quality after fingerprint embedding.

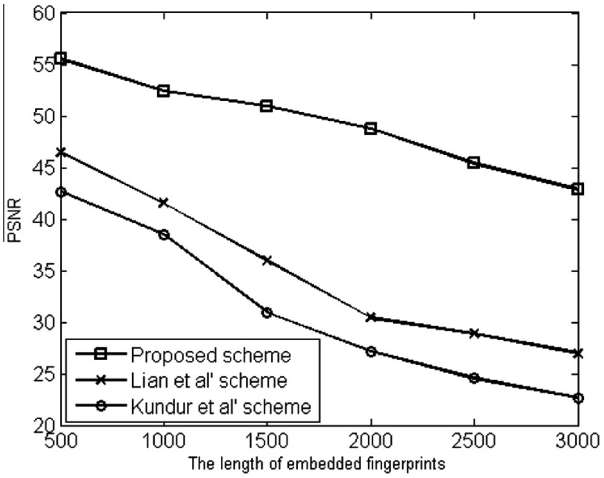


Fig. 8. Relation between imperceptibility and the length of embedded fingerprint sequences.

In our scheme, the fingerprint is embedded in the AC value, which has a smaller effect on the image quality compared with the other two schemes; therefore, the fingerprint is embedded imperceptibly. This finding arises because in Kundur's scheme, the fingerprint is generated by partially decrypting the sign bit of the AC coefficients, which affects the image quality in an obvious manner; in Lian et al's scheme, the fingerprint is embedded in the DC coefficient by decrypting under the control of the key and fingerprint, and the change in the DC coefficient also degrades the image quality in an evident fashion; and the degradation of Kundur's scheme is stronger than Lian et al's scheme.

Fig. 8 compares PSNR value of the fingerprinted image when the length of fingerprint sequences increases continuously. As the figure shows, the fingerprinted image generated by our method has better image quality than other two methods. Thus using our method can obtain larger embedding space and achieve better robustness against collusion attacks.

Fig. 7. Imperceptibility of fingerprinting image.

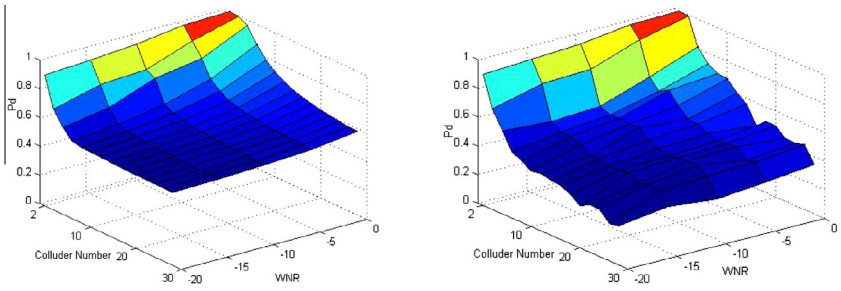


Fig. 9. (a) Analytical approximation of the fingerprint under an average attack (b) Simulation results of the fingerprint under an average attack.

4.3. Collusion resistance of fingerprints

To test the collusion resistance of fingerprints, we chose $(14, 2)_{16}$ Reed-Solomon code as collusion-resisting code, where $L = 14$, $t = 2$, $q = 16$, and the number of users is $N_u = 256$. The gold sequence with the length of $l = 127$ was used to perform a spread-spectrum operation; the length of the constructed fingerprint code was 1778, and the total number of users was 256. Supposing that the watermark-to-noise-ratio (WNR) ranged from 0 to -20 dB, including scenarios ranging from severe distortion to mild distortion; then, according to Eq. (9), the theoretical results of colluder identification performance under typical attack conditions is shown in Fig. 9(a). As shown in this figure, the fingerprint can resist at least a dozen colluders under a high WNR and a half dozen colluders under a low WNR. We estimated the probability of catching one colluder (P_d) for different colluder numbers c. The results of 100 iterations are shown in Fig. 9(b). It can be seen that the simulation results verify the analytical approximation.

5. Conclusions

A content security protection scheme that integrates encryption and digital fingerprints for the JPEG compressed domain is proposed in this paper. On the sender side, VLC codewords in the compressed domain are encrypted directly and data confidentiality can be guaranteed; on the receiver side, collusion-resistant fingerprint codes are embedded in the image naturally after decryption by the user, which can be used to trace colluders who illegally distribute copies of the media thus ensuring proper usage of the media. The proposed scheme is highly efficient in two aspects: First, only a single encrypted copy of media content is distributed, users who decrypt the data with different decryption keys obtain different fingerprinted copies. Second, the encryption and fingerprinting operation are implemented in a compressed domain, which avoids time-consuming decompression, encryption, or fingerprint embedding and recompression processes. The proposed variable modular encryption method solves the problem of having an invalid VLC codeword stemming from compressed data stream encryption, while format compliance of the ciphertext can be kept. Experimental results and analyses show the effectiveness and robustness of our proposed scheme. The proposed scheme is suitable for content security protection for multimedia information.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 41101416), the National Basic Research Program of China (No. 2011CB302204), and the Open Research Fund of The Academy of Satellite Application (No. 20121689).

References

[1] A. Adelsbach, U. Huber, A. Sadeghi, Fingercasting-joint fingerprinting and decryption of broadcast messages, in: ACISP, LNCS, 4058, Springer-Verlag, Berlin Heidelberg, 2006, pp. 136–147.
[2] A. Sadeghi, The marriage of cryptography and watermarking-beneficial and challenging for secure watermarking and detection, LNCS, 5041, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 2–18.
[3] D. Boneh, J. Shaw, Collusion-secure fingerprinting for digital data, IEEE Trans. Info. Theory 44 (5) (1998) 1897–1905, http://dx.doi.org/10.1109/18.705568.
[4] D. Kundur, K. Karthik, Video fingerprinting and encryption principles for digital rights management, Proc. IEEE 92 (6) (2004) 918–932.
[5] S. Lian, Multimedia Content Encryption: Techniques and Applications, Auerbach Publications, New York, 2009.
[6] B. Mobasseri, R. Berger. Watermarking in the JPEG bitstream, Proc. SPIE 5681, Security, Steganography, and Watermarking of Multimedia Contents VII.
[7] F. Hartung, B. Girod, Digital watermarking of MPEG-2 coded video in the bitstream domain, in: ICASSP, Proc. IEEE, 4, IEEE Computer Society, Washington, DC, USA, 1997, pp. 2621–2624.
[8] A.V. Subramanyam, Sabu Emmanuel, Mohan S. Kankanhalli, Robust watermarking of compressed and encrypted JPEG2000 images, IEEE Trans. Multimedia 14 (3) (2012) 703–716.
[9] J. Bloom, "Security and rights management in digital cinema", Proc. ICME 2003 1, 621–624, IEEE, Piscataway, NJ (2003).
[10] J. Pegueroles et al., A practical solution for distribution rights protection in multicast environments computational science and its applications, in: ICCSA, LNCS, 3982, Springer-Verlag, Berlin Heidelberg, 2006, pp. 527–536.
[11] J. Anderson, C. Manifavas, Chameleon-a new kind of stream cipher, Proc. FSE, 1267, Springer-Verlag, 1997, pp. 107–113.
[12] A.N. Lemma et al., Secure watermark embedding through partial encryption, in: Proc. IWDW, LNCS, 4283, Springer-Verlag, Berlin Heidelberg, 2006, pp. 433–445.
[13] S. Lian, Z. Wang, Collusion-traceable secure multimedia distribution based on controllable modulation, IEEE Trans. Circuits Syst. Video Technol. 18 (10) (2008) 1462–1467.
[14] S. Lian, X. Chen, Secure and traceable multimedia distribution for convergent mobile TV services, Comput. Commun. 33 (2010) 1664–1673, http://dx.doi.org/10.1016/j.comcom.2010.03.015.
[15] B. Mobasseri, R. Berger, M. Marcinak, Y. NaikRaikar, Data embedding in JPEG bitstream by code mapping, IEEE Trans. Image Process. 19 (4) (2010) 958–966.
[16] S. Lian, Z. Liu, Z. Ren, H. Wang, Secure distribution scheme for compressed data streams, Proc. IEEE ICIP 1953–1956 (2006) 2006.
[17] Tosun A S, Feng W C. On error preserving encryption algorithms for wireless video transmission. Proceedings of the ACM International Multimedia Conference and Exhibition. Ottawa, Ont, 2001, 302–308.
[18] Wu C P, Kuo C C J. Fast encryption methods for audiovisual data confidentiality. Proc. of SPIE International Symposia on Information Technologies 2000. Boston, USA, 2000. 284–295.
[19] Wu C P, Kuo C CJ. Efficient multimedia encryption via entropy codec design. SPIE International Symposium on Electronic Imaging 2001. San Jose, USA, 2001, 128–138.
[20] A. Boho, G. Wallendael, A. Dooms, J. Cock, G. Braeckman, P. Schelkens, B. Preneel, R. Walle, End-To-End security for video distribution, the combination of encryption, watermarking, and video adaptation, IEEE Signal Process. Mag. 30 (2) (2013) 97–107.
[21] Int. Telecommunication Union, CCITT Recommendation T.81, Information Technology-Digital Compression and Coding of Continuous tone Still Images-Requirements and Guidelines 1992.
[22] Shan. He, Wu. Min, Joint coding and embedding techniques for multimedia fingerprinting, IEEE Trans. Inf. Secur. 1 (2) (2006) 231–247.
[23] I. Brown, C. Perkins, J. Crowcroft, Watercasting: Distributed watermarking of multicast media, in: In Proc. NGC, LNCS, 1736, Springer, Heidelberg, 1999, pp. 286–300.
[24] H.V. Zhao, K.J. Liu, Fingerprint multicast in secure video streaming, IEEE Trans Image Process. 15 (1) (2006) 12–29.

[25] M.U. Celik, A.N. Lemma, S. Katzenbeisser, M.V.D. Veen, Lookup-table-based secure client-side embedding for spread-spectrum watermarks, IEEE Trans. Inf. Forensics Secur. 3 (3) (2008) 475–487.
[26] C.-Y. Lin, P. Prangjarote, L.-W. Kang, W.-L. Huang, T.-H. Chen, Joint fingerprinting and decryption with noise-resistant for vector quantization images, Signal Process. 92 (9) (2012) 2159–2171.
[27] Jordi. Serra-Ruiz, David. Megias, A novel semi-fragile forensic watermarking scheme for remote sensing images, Int. J. Remote Sens. 32 (19) (2011) 5583–5606.
[28] Joan. Serra-Sagristà, Francesc. Aulí-Llinàs, Remote sensing data compression, Comput. Intell. Remote Sens., SCI 133 (2008) 27–61.
[29] Jiangtao Wen, M. Severa, Wenjun Zheng, M.H. Luttrell, Wenyin Jin, A format-compliant configurable encryption framework for access control of video, IEEE Trans. Circuits Syst. Video Technol. 12 (6) (2002) 545–557.