Combined Encryption and Watermarking Approaches for Scalable Multimedia Coding

Feng-Cheng Chang, Hsiang-Cheh Huang, and Hsueh-Ming Hang

Department of Electronics Engineering, National Chiao Tung University,
Hsinchu 300, Taiwan, R.O.C.,
hmhang@mail.nctu.edu.tw

Abstract. Intellectual Property (IP) protection is a critical element in a multimedia transmission system. Conventional IP protection schemes can be categorized into two major branches: *encryption* and *watermarking*. In this paper, a structure to perform layered access protection by combining encryption and robust watermarking is proposed and implemented. By taking advantage of the nature of cryptographic schemes and digital watermarking, the copyright of multimedia contents can be well protected. We adopt the scalable transmission method over the broadcasting environment. The embedded watermark can be thus extracted with high confidence. Then, the next-layer secrets can be perfectly decrypted and reconstructed. Finally, the media contents are recovered.

1 Introduction

With the widespread use of multimedia broadcasting, the digital media, including images, audio and video clips, are easily acquired in our daily life. The current network environments make scalable coding of multimedia a necessary requirement when multiple users try to access the same information through different communication links [1,2]. Scalability means that a multimedia data bitstream is partitioned into layers in such a way that the base layer is independently decodable into a content with a reduced quality. The reduction may be in spatial resolution, temporal resolution, or signal-to-noise ratio (SNR). To reproduce the original content, enhancement layers provide additional data to restore the original quality from the base layer. Enhancement layers represent the scalability of the content coding, namely, spatial, temporal, or SNR scalability. Therefore, scalable coding of multimedia is suitable for delivering digital contents to different users and devices with various capabilities [3].

In many cases, it requires to deliver multimedia content securely. However, the channel for multimedia broadcasting is an open environment; thus, if the user data and information are not protected, it might be illegally used and altered by hackers To protect privacy and intellectual property (IP) rights, people often use cryptographic techniques to encrypt data, and the contents protected by encryption are expected to be securely transmitted over the Internet [4,5].

In cryptography, the contents to be encrypted are called *plaintext*, and the encrypted contents are called *ciphertext*. Although cryptographic schemes provide secure data exchange among peers, it implies that the ciphertext cannot be

K. Aizawa, Y. Nakamura, and S. Satoh (Eds.): PCM 2004, LNCS 3333, pp. 356–363, 2004.

altered during transmission [6]. If any one bit is received erroneously, the plaintext cannot be decrypted correctly. This is not a good property when we deliver protected contents in a broadcasting environment, where erroneous transmission may occur occasionally. There, a one-bit error may cause a totally useless content. To meet this deficiency for multimedia broadcasting, we include watermarking technique to aid encryption, because the watermarked contents can withstand some kind of *attacks*, including signal processing, geometric distortion, and transmission errors. In this paper, we combine both the cryptographic and watermarking techniques for layered content protection. On the one hand, the message for protection of multimedia contents can be perfectly decrypted by cryptography, while on the other hand, the encrypted message can be further protected by robust watermarking algorithms to resist transmission errors.

This paper is organized as follows. Sec. 2 describes the concepts and issues of layered content protection. In Sec. 3, we propose a layered protection structure with combined cryptographic and watermarking schemes. We give an application example and simulations in Sec. 4. And Sec. 5 concludes this paper.

2 Layered Protection Concepts

As discussed in Sec. 1, scalable coding is a solution to broadcast contents to devices with various playback capabilities. With the nature of layered coding, the whole media can be partitioned into blocks of data. Thus, it is straightforward to group receivers of different playback capabilities by sending different combinations of data partitions. However, the conditional access (CA) requirement is dealt in a different way in a broadcast environment. To distinguish different groups of users, a popular solution is to encrypt data by a group-shared key. Thus, the CA issue can be solved by encrypting data partitions with different keys, and a granted user has the corresponding decryption keys to the assigned data partitions.

The next issue is how to distribute the keys. Depending on the delivery infrastructure, two problems may arise. One is how to protect keys from malicious listeners. There are methods to protect keys from malicious listeners, such as the one proposed in the DVB standard [7]. The other problem is how to synchronize (in time) a key with the content. For example, to broadcast a protected content over Internet, we may send the key to users via a reliable channel (such as RTSP connection [8]), while the content goes through an unreliable channel (such as RTP sessions [9]). A reliable channel guarantees information correctness by sacrificing delivery speed, and it is likely that the key information is out-of-sync to the corresponding content.

A possible solution to eliminate synchronization problem is to transmit the key information together with the content, such as inserting it into the optional header fields of the coded stream. However, it may be destroyed by transmission errors or transcoding. Our proposed method is less sensitive to minor transmission errors. We embed the key information into the content with robust watermarking techniques. Since the key information is available at the same time

as we reconstruct the content, the (time) synchronization problem is resolved. The drawback of this approach is that if packet loss or transcoding occurs, the reconstructed content is different from the original one, and the key information may not be extracted accurately. To reduce the impact of unreliable or distorted delivery, we incorporate robust digital watermarking methods [10] to reinforce the robustness of the embedded key information.

The main steps of the layered protection is organizing secrets (keys and necessary parameters) into a watermark, robustly watermarking the base layer, and encrypting the enhancement layer. A granted user receives the base layer, extracts and derives the decryption key, decrypts the enhancement layer, and combine layers together to produce the contents. In the following sections, we will describe our proposed method in detail.

3 Proposed Method

In this section, we describe the layered decryption and decoding operations on the receiver side. Because the associated encryption and encoding operations vary depending on the scalable coding, we provide an example at the end of this section. We first describe the receiver architecture in our proposed method, then we describe the corresponding transmitter architecture in the following paragraphs.

3.1 Receiver Architecture

Scalable coding is composed of one *base layer* and several *enhancement layers* to match the network diversity for transmission. The enhancement operation is illustrated in Fig. 1. Assuming that the initial base layer B_0 has been received, the subsequent composing operations can be expressed by

$$B_i = \text{compose } (B_{i-1}, E_i), \tag{1}$$

where

$$E_i = \text{decrypt}_e \ (X_i, K_i). \tag{2}$$

In Eq. (1), B_{i-1} is the available base layer, and E_i is the enhancement layer to improve quality from B_{i-1} to B_i. During transmission, E_i is protected by a cryptic algorithm with K_i as the key, and the transmitted data is X_i in Eq. (2).

There are some secret information to be obtained prior to decrypting E_i, and the operations can be expressed as follows:

$$W_i = \text{extract } (B_{i-1}, P_{i-1}) \tag{3}$$

$$F_i = \text{decrypt}_f \ (W_i, G_i) \tag{4}$$

$$K_i = \text{key } (F_i) \tag{5}$$

$$P_i = \text{param } (F_i) \tag{6}$$

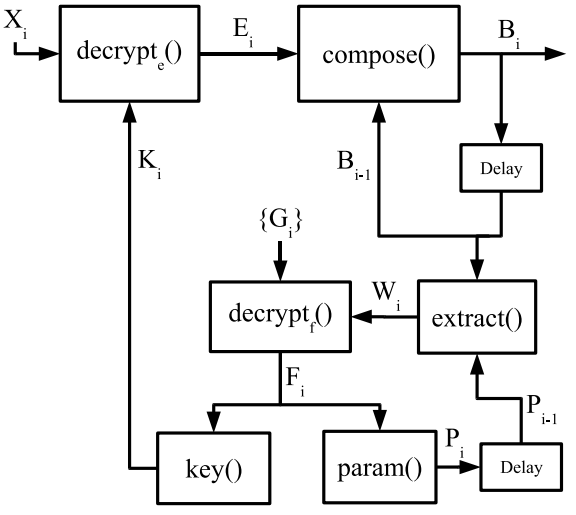


Fig. 1. Decryption and decoding of layer-protected content

W_i is the digital watermark extracted from the constructed base layer B_{i-1} with extraction parameter P_{i-1}. As described in Sec. 2, W_i represents the protected secret information. Thus, we have the secret information F_i by decrypting the watermark using user-specific key G_i. After parsing F_i, we obtain the decryption key K_i and the next watermark extraction parameter P_i.

As Fig. 1 illustrates, the decryption and composition blocks are iterative processes. There are several initial parameters required to activate these processes. We will discuss how to obtain the initial parameters in the following paragraphs.

- When the whole content is protected, namely, B_0 is encrypted, we need K_0 to decrypt X_0. In this case, K_0 should be obtained by a separate channel.
- One scenario is that B_0 is the "preview" layer; i.e., B_0 is not encrypted, we simply bypass the decrypt_e block.
- Depending on the watermarking algorithm, the extraction process may requires specific parameters. If it does, the first watermark extraction parameter P_0 should be obtained in a separate channel to activate subsequent extraction process.
- All the key-decryption keys $\{G_i\}$ should be obtained before receiving the media data, for instance, by manually or automatically update after subscription.

3.2 Transmitter Architecture

Depending on the scalable coding algorithm, the design of transmitter side varies. Fig. 2 shows one of the possible designs. The architecture is almost the inverse of the receiver architecture in Fig. 1. The watermark W_i is the encrypted version of the key K_i and the embedding parameter P_i. The B'_{i-1} is the un-watermarked

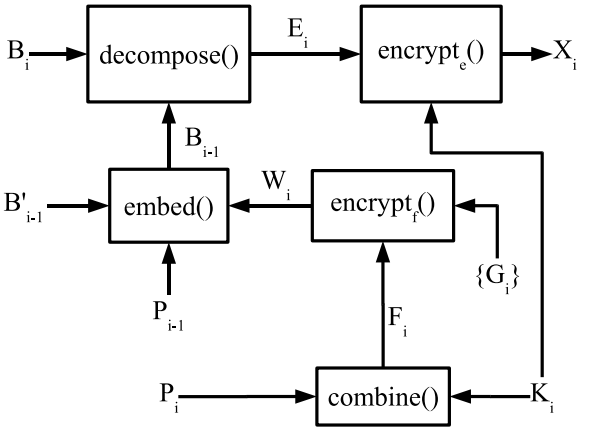


Fig. 2. Encryption and encoding of layer-protected content

base layer with lower quality. After embedding W_i into B'_{i-1}, we have the base layer B_{i-1}. The enhancement layers are generated as the differences between B_i and B_{i-1}. All the $\{K_i\}$, $\{P_i\}$, and $\{G_i\}$ are known in advance.

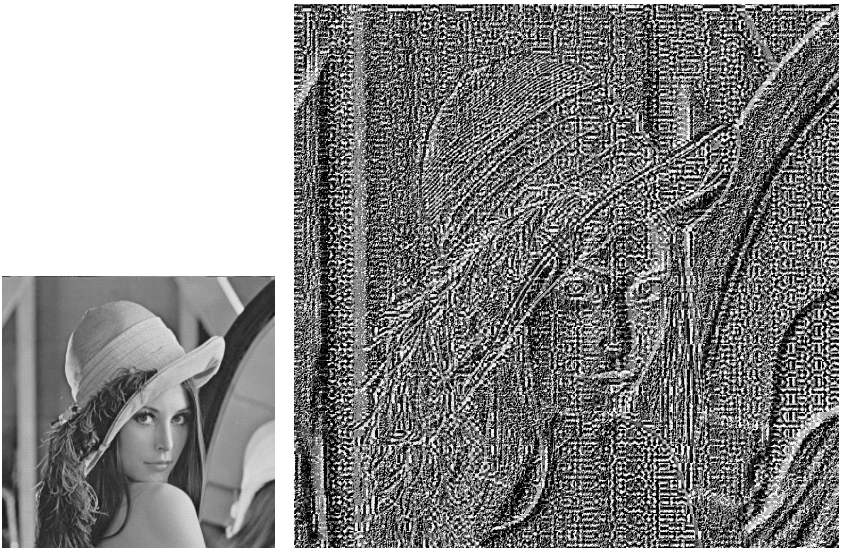
4 Simulation Results

In this paper, we use the test image `Lena` with size 1024×1024 to conduct the simulations in this section. The original `Lena` is first converted to 512×512 base layer. The DES[11] key (8 ASCII letters **"NCTU-DEE"** in Fig. 3(a)) to encrypt enhancement layer is also encrypted using DES by the user key ($\{G_i\}$ in Fig. 1) to generate the 8-byte (or 64-bit) secret. The secret is then repeated for 32 times to form the binary watermark, as shown in Fig. 3(b).

(a) (b)

Fig. 3. Plaintext encryption and watermark generation. (a) The 8-byte plain-text. (b) The converted binary watermark with size 128×128.

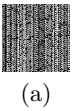
Figure 4 shows the data in transmitted base layer and the enhancement layers. Before transmission, the watermarked base layer has acceptable visual quality, with the PSNR of 39.24 dB in Fig. 4(a). We then extract the watermark from the base layer picture, derive the decryption key, decrypt the transmitted enhancement data in the next layer, and finally reconstruct the original 1024×1024 picture.



(a) (b)

Fig. 4. (a) 512×512 base layer. (b) 1024×1024 enhancement layer.

We then test the packet loss case on the base layer [12]. The packet loss rate in our simulations is set to 10%. The extracted watermark is shown in Fig. 5(a). The distortion is within the tolerance range of the extracted watermark, with the bit-correct rate of 92.74%. We then use the majority vote to produce the 8-byte secret, extracted encryption key, and decrypt the ciphertext. Finally, we can recover the original key information correctly as shown in Fig. 5(b). In addition, the 1024×1024 picture thus can be reconstructed with some defects as shown in Fig. 6.



(a) $\Longrightarrow$ **"NCTU–DEE"** (b)

Fig. 5. Watermark extraction and cipher-text decryption. (a) The extracted watermark, with the bit-correct rate of 92.74%. (b) The decrypted cipher-text, which is identical to that in Fig. 3(a).

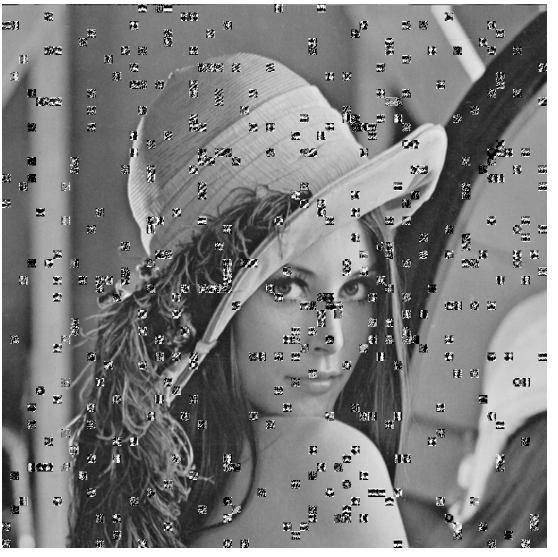


Fig. 6. The encrypted and watermarked image corrupted by transmission errors, with best-effort reconstruction.

5 Conclusion

In this paper, we proposed a structure to protect the layered (scalable) content in a broadcast environment. By combining cryptographic and robust watermarking techniques, the secret for decrypting enhancement data streams can be safely embedded in the base layer. Robust watermark enables embedding information directly in the multimedia content, and the embedded bits can be extracted even when the watermarked media experience attacks during transmission. On the other hand, cryptography provides confidentiality. But it does not tolerate any bit error. The contribution in this paper is to combine these two techniques, and offer the advantages of both for intellectual property protection.

In the proposed scheme, the encryption concept guarantees the access control, keeping away malicious eavesdroppers. Also, the embedding concept solves the key-content synchronization problem. The robust watermarking concept increases the data robustness against transmission errors and distortions. Comparing to conventional cipher-block chaining encryption, our method not only provides a way to guarantee access controls, but also synchronously transmits decryption information. Moreover, robust watermarking implicitly gives higher data integrity protection on the keys than on the contents. One simulated example demonstrates the effectiveness of the proposed structure.

In our future work, we will modify our structure with scalable video coding. We will also integrate our proposed structure with the MPEG IPMP (Intellectual Property Management and Protection) message exchange format[13,14].

References

1. Sun, X., Wu, F., Li, S., Gao, W., Zhang, Y.-Q.: Seamless switching of scalable video bitstreams for efficient streaming. IEEE Transactions on Multimedia **6** (2004) 291–303
2. Almeida, J.M., Eager, D.L., Vernon, M.K., Wright, S.J.: Minimizing delivery cost in scalable streaming content distribution systems. IEEE Transactions on Multimedia **6** (2004) 356–365
3. Wiegand, T., Sullivan, G.J., Bjntegaard, G., Luthra, A.: Overview of the H.264/AVC video coding standard. IEEE Transactions on Circuits and Systems for Video Technology **13** (2003) 560–576
4. Parviainen, R., Parnes, P.: Large scale distributed watermarking of multicast media through encryption. Proceedings of the International Federation for Information Processing, Communications and Multimedia Security Joint Working Conference IFIP TC6 and TC11 (2001) 149–158
5. Lim, Y., Xu, C., Feng, D.D.: Web-based image authentication using invisible fragile watermark. Conferences in Research and Practice in Information Technology (2002) 31–34
6. Xu, X., Dexter, S., Eskicioglu, A.M.: A hybrid scheme for encryption and watermarking. IS&T/SPIE Symposium on Electronic Imaging 2004, Security, Steganography, and Watermarking of Multimedia Contents VI Conference (2004) 723–734
7. Digital video broadcasting project (DVB): http://www.dvb.org/ (2004)
8. Real time streaming protocol: http://www.rtsp.org/ (2004)
9. Schulzrinne, H.: http://www.cs.columbia.edu/~hgs/rtp/ (2004)
10. Shieh, C.S., Huang, H.C., Wang, F.H., Pan, J.S.: Genetic watermarking based on transform domain techniques. Pattern Recognition **37** (2004) 555–565
11. Data Encryption Standard (DES): http://www.itl.nist.gov/fipspubs/fip46-2.htm (1993)
12. Chande, V., and Farvardin, N.: Progressive transmission of images over memoryless noisy channels. IEEE Journal on Selected Areas in Communications **18** (2000) 850–860
13. Avaro, O., Eleftheriadis, A., Herpel, C., Rump, N., Swaminathan, V., Zamora, J., Kim, M.: MPEG systems (1-2-4-7) FAQ, version 17.0. ISO/IEC JTC1/SC29/WG11 N4291 (2001)
14. Huang, C.C., Hang, H.M., Huang, H.C.: MPEG IPMP standards and implementation. IEEE PCM'02 (2002) 344–352