Contents Protection for Media Streaming Service

Goo-Rak Kwon*

1. Introduction

With the advance of multimedia technology, multimedia sharing among multiple devices has become the main issue. This allows users to expect the peer-to-peer distribution of unprotected and protected contents over public network. Many audio and video (A/V) processing software including DVD players, CD rippers, MP3 encoders, and A/V players have been posted for free on the Web allowing users to build their own A/V record collections from their own CD and DVD. Inevitably, this situation has caused an incredible piracy activity and some Web sites have begun to provide copyrighted A/V data for free. In order to protect the contents from illegal attacks, digital rights management (DRM) is required as shown in Fig. 1.

The DRM system generally provides two essential functions: management of digital rights by identifying, describing, and setting the rules of the content usage, and digital management of right by securing the contents and enforcing usage rules. The basic principle of the DRM model [3]-[5] is to separate and identify three core entities: Users, Content, and Rights. Users can be any type of users from a rights holder to an end-consumer. Content is any type of contents at any level of aggregation. A right is an expression of permissions, constraints, and obligations between Users and Content. This model provides the greatest flexibility when assigning rights to any combination or layering of Users and Content. Fig. 2 shows the example of contents distribution service using DRM. In Fig. 2, a user A requests image, audio, or video in the network and goes through payment. Billing system informs payment approval, and then CP delivers encrypted contents to only an authorized user through the network.

Various encryption techniques for DRM have been researched. These techniques are classified into two approaches: scrambling and watermarking. Scrambling [6]-[8] that is generally based on old and proven cryptographic tools, efficiently ensures confidentiality, au-

※ 교신저자(Corresponding Author) : 권구락, 주소 : 광주광역시 동구 서석동 조선대학교 전자정보공과대학 정보통신공학과 8층, 전화: 062-230-6268, FAX: 062-230-6268, E-mail : grkwon@chosun.ac.kr
* 조선대학교 정보통신공학과

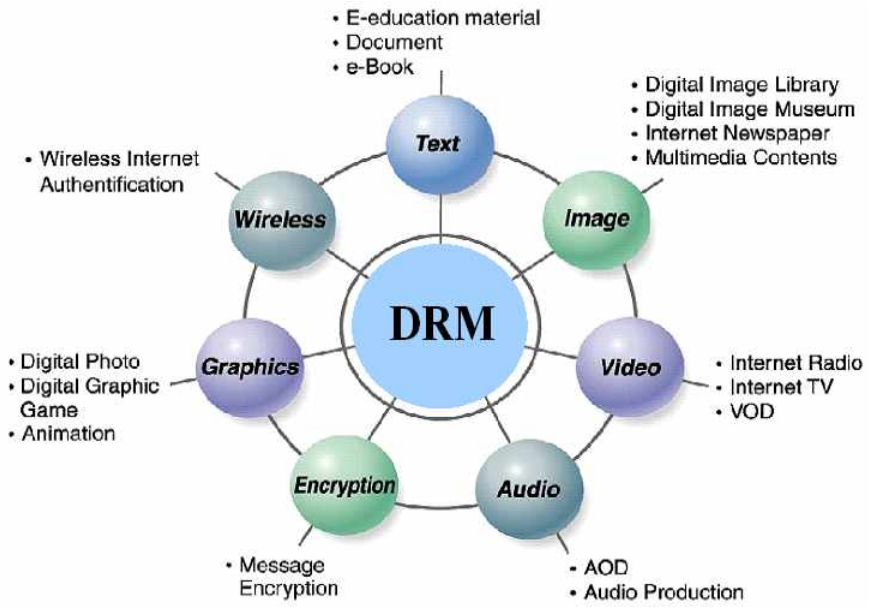


Fig. 1. Application of DRM technique.

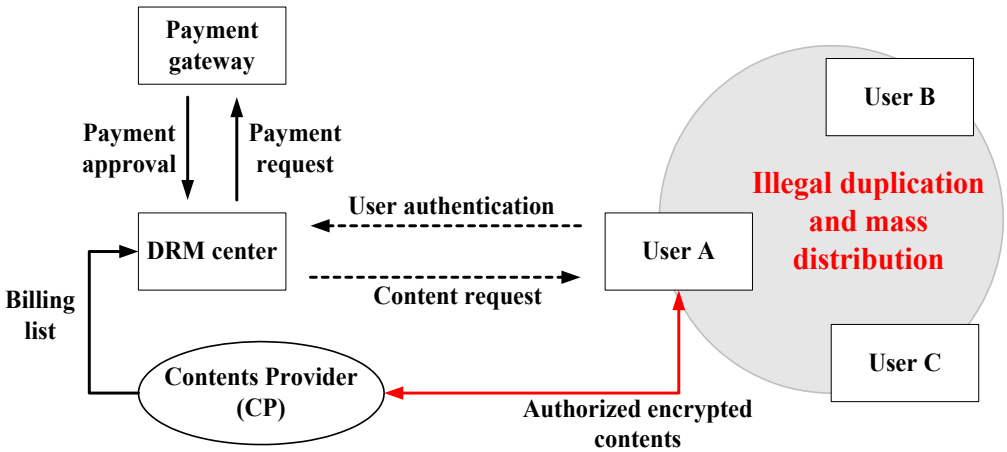


Fig. 2. Example of the contents distribution service using DRM.

thenticity, and integrity of messages. However, it does not protect against unauthorized copying after the message has been successfully transmitted and descrambled [3]. This kind of protection can be handled by watermarking [12][13], which is a more recent topic that has attracted a large amount of research and is perceived as a complementary aid in encryption. A digital watermark is a piece of information inserted and hidden in the media content [6]. This information is imperceptible to a human observer but can be easily detected by a computer. Moreover, the main advantage of this technique is to provide the nonseparability of the hidden information and the content.

A watermarking system consists of an embedding algorithm and a detecting function. The embedding algorithm inserts a message into a

media and the detecting function is then used to verify the authenticity of the media by detecting the message. The most important properties of a watermarking scheme include robustness, fidelity, tamper resistance, and data payload [8].

The rest of the paper is organized as follows. Section 0 briefly describes the Background such as DRM architecture, MP3, and MP4. In Section 0, the joint and partial encryption techniques are presented using MPEG-1 audio layer III and MP4 video coder.

2. Background

This section describes the common components in DRM systems, the working process of DRM models and the role taken by client-side applications in DRM. Each DRM vendor supports different DRM implementation, names and ways to specify the content usage rules, however, the basic DRM process is the same, which usually involves four parties: the content provider, the distributor, the clearinghouse, and the consumer. Usually a DRM system is integrated with an e-commerce system that handles financial payments and triggers the function of the clearinghouse, however, due to space considerations we does not describe the details. Fig. 3 displays the common components of a DRM system based on most existing commercial systems. Following the explanation of these common elements, a typical model used by current DRM implementations is presented.

- Content Provider such as a music record label or a movie studio holds the digital rights of the content and wants to protect these rights.
- Distributor provides distribution channels, such as an online shop or a web retailer. The distributor receives the digital content

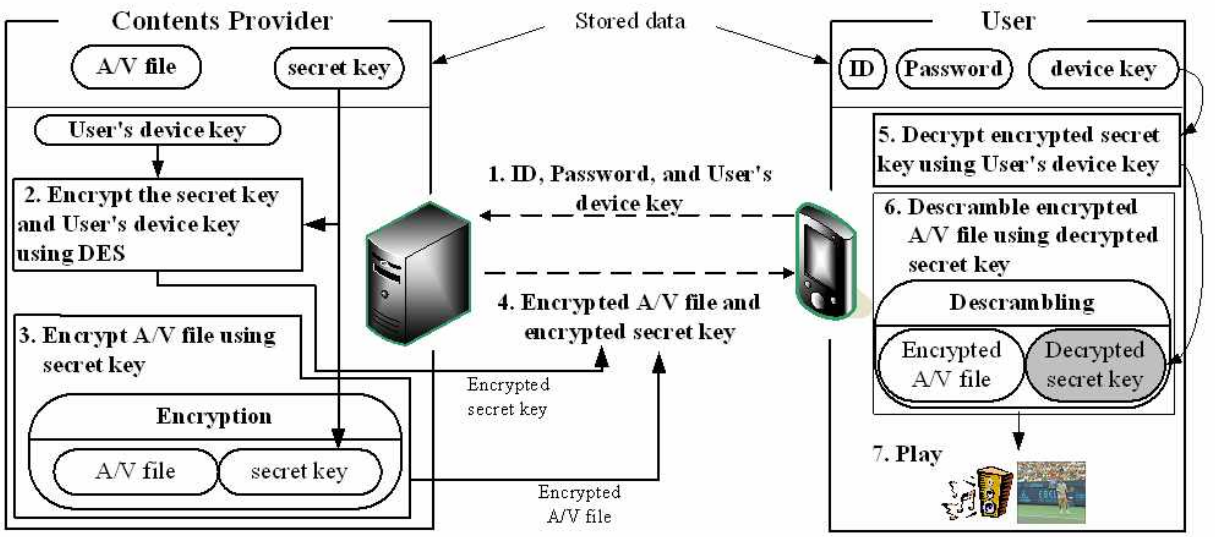


Fig. 3. Architecture of the proposed DRM system.

from the content provider and creates a web catalogue presenting the content and rights metadata for the content promotion.

- Consumer uses the system to consume the digital content by retrieving downloadable or streaming content through the distribution channel and then paying for the digital license. The player/viewer application used by the consumer takes charge of initiating license request to the clearinghouse and enforcing the content usage rights.
- Clearinghouse handles the financial transaction for issuing the digital license to the consumer and pays royalty fees to the content provider and distribution fees to the distributor accordingly. The clearinghouse is also responsible for logging license consumptions for every consumer.

A typical DRM model used by current DRM implementations works as follows: Firstly, the content provider encodes the digital content into the format supported by the DRM system. Different DRM systems provided by different DRM vendors may support different content formats. The digital content is then encrypted and packaged for the preparation of distribution. The content provider may use watermarking technology to embed digital codes into the digital content that can identify the ownership of the content and the usage rules. Next, the protected content is transferred to the appropriate content distribution server, web server or steaming server, for on-line distribution.

The digital license containing content decryption keys and usage rules is sent to the clearinghouse. The usage rules specify how the content should be used, such as copy permit, pay-per-view, a one-week rental, etc. At the other end of the process, the consumer downloads the digital content from the web server or requests streaming content from the streaming server. To be able to consume the protected content, the user has to request a valid license from the clearinghouse. After receiving the license request, the clearinghouse verifies the user's identity for example by having the user present a valid digital certificate, charges his account based on the content usage rules, and generates transaction reports to the content provider. Finally, the license is delivered to the consumer's device after the consumer has paid through the e-commerce system, and the protected content can be decrypted and used according to the usage rights in the license.

In DRM system, consumers can login along received digital content to other people through super-distribution, which lets vendors market their digital content to a vast amount of potential customers without direct involvement. Although digital content can be freely distributed, to utilize the content, the recipient has to contact the clearinghouse and provide whatever information or payment required for the license.

3. Joint and Partial Encryption Technique in A/V Codec

Before explaining the A/V scrambling technique in detail, we first introduce the concept of the joint and partial scrambling. The joint scrambling method can provide levels of security for contents encryption. The security level can be determined by the number of independent encryption methods combined. The partial encryption method obtains the higher security by simply encrypting only significant parts of the compressed data. Next, we provide a detailed explanation on the A/V encryption technique.

In order to protect the contents against eavesdropping and moreover against illegal mass distribution after descrambling, we propose an A/V watermarking and scrambling (WS) method based on our previous research results [9].

Fig. 4 shows the block diagram of the WS method for the audio data. Fig. 4(a) shows the MP3 encoder with watermark embedding and scrambling. The original PCM data is decomposed into 32 subbands using the polyphase quadrature filterbank. Each decomposed signal is transformed using MDCT to increase the resolution of frequency. The watermark is embedded into MDCT coefficients. The watermarked MDCT coefficients are synchronously scrambled using both the watermark sequence and the secret key and quantized into water-

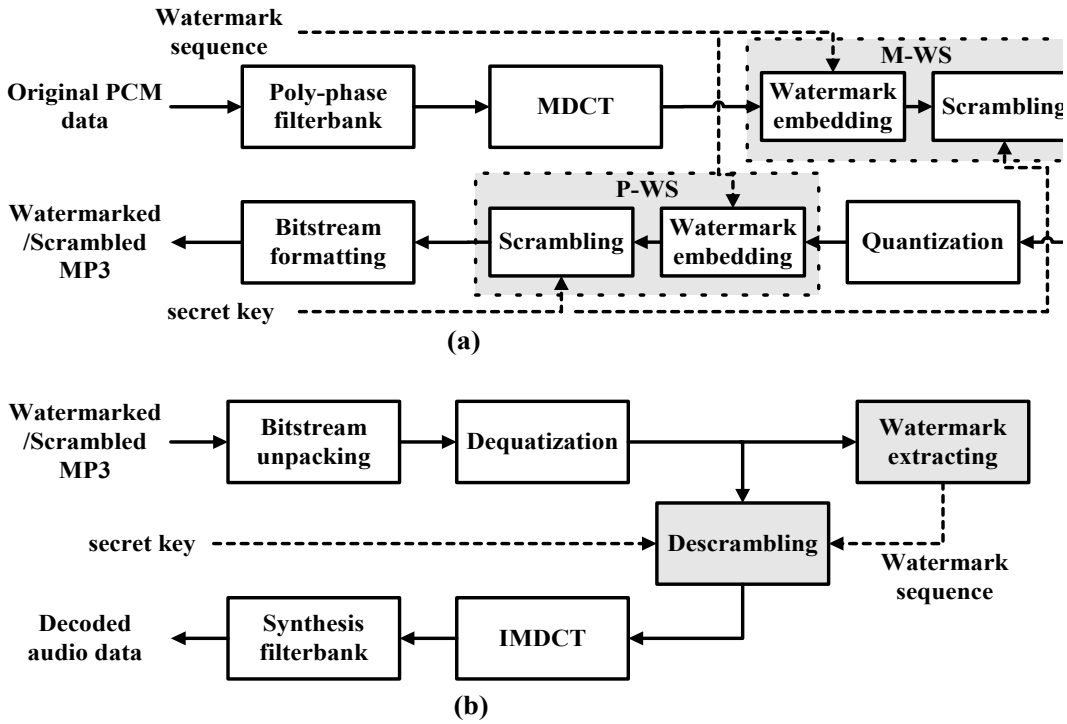


Fig. 4. Proposed audio encryption using watermarking and scrambling. (a) Encoder with watermark embedding and scrambling. (b) Decoder with watermark extracting and descrambling.

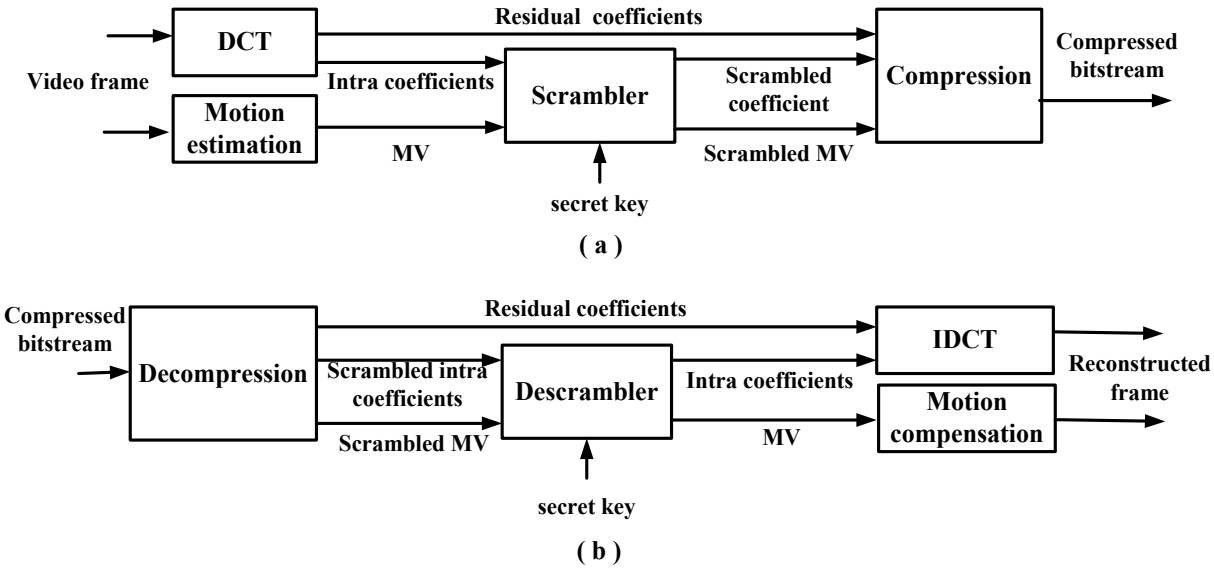


Fig. 5. Proposed video encryption using scrambling. (a) Scrambler. (b) Descrambler.

marked/scrambled MP3 bitstream. Fig. 4(b) shows the MP3 decoder with watermark extraction and descrambling. The watermarked/scarambled MP3 bitstream is quantized. The watermark sequence is extracted from the dequantized signal. Then, the dequantized signal is descrambled using the extracted watermark sequence and the secret key. The descrambled signal is inversely transformed using inverse MDCT and passed through the synthesis filterbank to produce the restored audio data.

Fig. 5 shows the block diagram of the scrambling/descrambling algorithms for the video data. At the encoder, in the case of intraframe, the original frame is transformed and encrypted using the secret key. Then, the scrambled DCT coefficients are quanitzed and fed to the entropy encoder. In the case of interframe, MVs obtained by motion estimation are scrambled using the secret key and fed to the entropy encoder while DCT coefficients are quantized and fed to an entropy encoder without scrambling. At the decoder, the compressed video bitstream is decompressed by the entropy decoder and the dequantizer. Prior to IDCT and motion compensation, the decompressed DCT coefficients of interframe and MVs of intraframe are descrambled using the secret key provided only

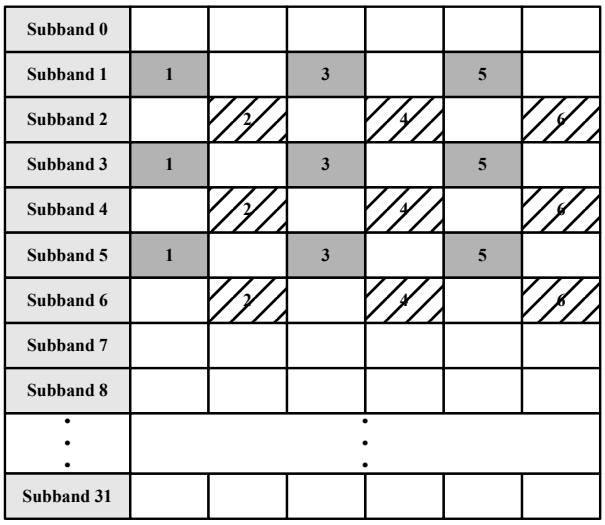


Fig. 6. Scrambling region in MDCT subbands.

for the authorized devices.

3.1. Audio encryption technique

3.1.1. Magnitude-based watermarking and scrambling (M-WS)

The WS technique encrypts MDCT coefficients in the MP3 audio. The MP3 bitstream is a concatenation of a sequence of channels. Each channel corresponds to two granules, where each granule is defined as precisely 576 consecutive samples. Therefore, each channel has 1152 samples consisting of 32 subbands each of which has 36 MDCT coefficients. Fig. 5 shows an example of the MP3 audio channel. Each subband of the frame is divided into 6 blocks. Each block has 6 MDCT coefficients. In the watermarking technique, the watermark sequence is embedded into MDCT coefficients of subband 0. The binary representation of the watermark sequence, W, is given by

$$W = b_1 b_2 b_3 \cdots b_N \quad (1)$$

where b_i represents the ith digit and N is the number of bits in W. Let P denote the number of zeros in a row after a decimal point of an MDCT coefficient. For example, P=2 for 2.003. In the watermarking technique, the first MDCT coefficient of each block except the 6th block is watermarked using value P of the MDCT coefficients and bi's of W. The watermarked MDCT coefficients are replaced by original MDCT coefficients if the following conditions are satisfied:

- bi = 1 and P of each MDCT coefficient is odd.
- bi = 0 and P of each MDCT coefficient is even.

On the other hands, the watermarked MDCT coefficients is equal to $5\times10^{(-p+1)}$ .As the remainder is equal to 5, P has the robustness against the quantization error after the watermarked MDCT coefficients are quantized and dequantized. In the decoder, W can be easily extracted by checking whether P of the MDCT coefficient is odd or not.

After embedding the watermark, scrambling is performed using the watermark sequence. Since the energy is concentrated on [subband 1 - subband 6], scrambling is performed for these bands. Fig. 6 shows the flowchart of scrambling. If bi = 0, subbands 2, 4, and 6 are selected for scrambling. Otherwise, subbands 1, 3, and 5 are selected. The secret key determines which block of the selected bands is scrambled. Fig. 3 shows the blocks selected by the secret key. Scrambling can be performed by simply amplifying the selected coefficients.

3.2. Video encryption technique

3.2.1. Segment-based DCT coefficient scrambling (S-DCTCS) for Intraframes

To scramble intraframes in MPEG-4, the S-DCTCS is applied to each MB in intraframes. The conventional approaches for scrambling change the sign bit of each coefficient in the 8×8 DCT based frames (Cox, 2000; Tang, 1996).

This sign scrambling efficiently provides an impact on distortion. However, in this method, high computational complexity is required to encrypt all coefficients for each MB. In addition, the scrambling can be easily cracked since the scrambling scheme is very simple. To solve these problems, we propose the S-DCTCS method that scrambles the sign bits of DC and AC coefficients.

In this algorithm, in order to reduce computational complexity, we first divide a frame into segments consisting of several MBs. Instead of encrypting all the coefficients, we scramble the largest DC coefficient among the DC coefficients of MBs in the segment and the largest AC coefficient of each MB. Changing the sign bits of DC coefficients and AC coefficients heavily depends on the number of MBs in Table I. As a result, Fig. 10 shows the effectiveness on efficient distortion for the sign-scrambling of each DC and AC coefficients in intraframes. Each scrambling mode is defined as shown in Table I. The mode in Table I is encrypted with the device key at a contents provider and transmitted to a user. Note that the scrambled mode set is considered as the secret key in the DRM system in Fig.. Therefore, a user who does not have the same key for decryption cannot access the original contents without visual quality degradation.

3.2.2 MV scrambling (MVS) for Interframes

Since encoded DCT coefficients of interframes (i.e, P or B frames) represent residual errors, DCT coefficients have typically small values. Thus, it is not good enough to apply the S-DCTCS for interframes. That is, for un-

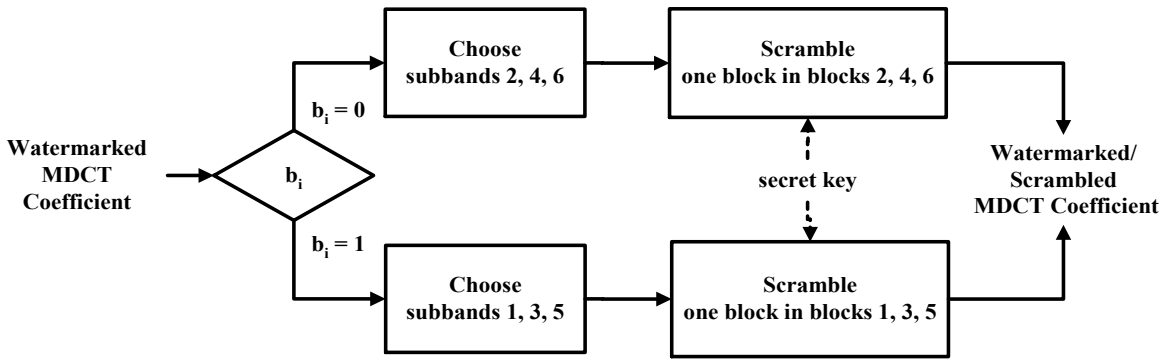


Fig. 7. Scrambling process of M-WS.

Table I. Video encryption mode set based as Secret Key

The number of MBs	Intraframe	Interframe
0	$DC' = -DC$	$MV' = MV + \Delta\alpha$
1	$AC' = -AC$	$MV' = -MV$
⋮	⋮	⋮
k	DCT sign encryption	MV sign/phase angle encryption

(a) (b)

Fig. 8. Picture after encrypting DCT coefficients. (a) DC sign scrambling. (b) AC sign scrambling.

authorized users, the S-DCTCS cannot guarantee original visual quality caused by video encryption [12]. In order to solve this problem, we propose two video encryption methods for scrambling MVs in interframes of MPEG-4 video: phase angle and sign scrambling.

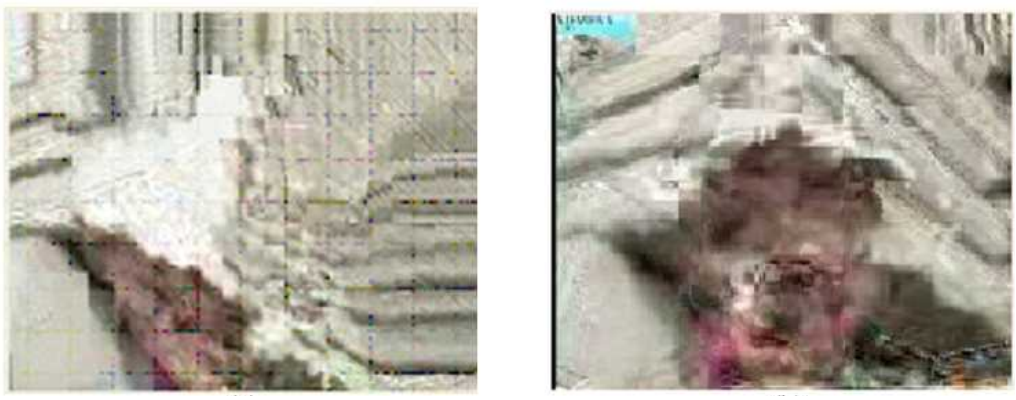
The phase angle, θ, is calculated as

$$\theta = \arctan\left[\frac{MV_v}{MV_h}\right], \tag{2}$$

where MV_v and MV_h, respectively, are the vertical and horizontal components of the MV. θ can be changed by adding weighting factors to each component. If the vertical and horizontal components of an MV are changed, the phase angle of the MV is modified as:

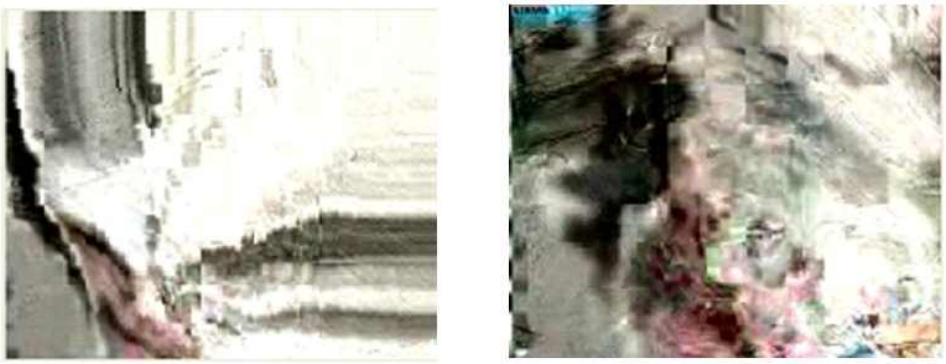
$$\begin{aligned} MV_v' &= MV_v + \Delta v, \\ MV_h' &= MV_h + \Delta h, \end{aligned} \tag{3}$$

where MV_v' and MV_h' represent the modified MV. The MV sign encryption scheme is very simple to implement. The direction of the MV is reversed as follows:



(a) (b)

Fig. 9. Picture after encrypting MVs. (a) MV phase angle encryption. (b) MV sign encryption.



(a) (b)

Fig. 10. Hybrid encryption. (a) DC sign plus MV phase angle encryption. (b) AC sign plus MV sign encryption.

$$MV_v' = -MV_v,$$
$$MV_h' = -MV_h. \quad (4)$$

Fig. 8 shows the result of MV encryption. Results of MV phase and MV sign scrambling are shown in Fig. 9 (a) and (b), respectively. Fig. 9 shows the results of differential encrypting methods. Fig. 10 (a) and (b) show the results of DC sign and MV phase angle scrambling and AC sign and MV sign encryption. These methods provide a very good compromise between compression ratio and coding efficiency since they produce a little bit overhead while scrambling DCT coefficients and MVs.

4. Conclusions

A new encryption method for contents protection of MP3 and MP4 Codec is shown. The magnitude and phase information of MDCT coefficients is scrambled for audio encryption. DCT coefficients and MVs in MP4 coder are used for video scrambling. With watermarking-scrambling technique, only users who have the secret key can access the watermarked-scrambled MP3/MP4 contents and moreover protect illegal copies of descrambled contents.

In addition, the minimal cost encryption scheme for securing the copyrighted MP3/MP4 A/V data is presented. The scrambling and watermarking techniques achieve a very good compromise between several desirable properties such as speed, security, and file size.

References

[1] A. Matsunaga, K. Koga, and M. Ohkawa, "An analog speech scrambling system using the FFT technique with high level security," IEEE Trans. on Selected Areas in Communications. 7(4), pp. 540-547, 1989.

[2] A. Piva, F. Bartolini, and M. Barni, "Managing copyright in open networks," IEEE Internet Computing. 6(3). pp. 18-26, 2002.

[3] B. Schneier, "Applied cryptography," 2nd ed., New York, Wiley and Sons, 1996.

[4] W. Zeng and S. Lei, "Efficient frequency domain selective scrambling of digital video,"

IEEE Trans. on Multimedia, 5, pp. 118-129, 2003.

[5] C. Neubauer, J. Herre, and K. Brandenburg," Continuous steganographic data transmission using uncompressed audio, Proc. Information HidingWorkshop, Portland, OR, pp. 208-217, 1998.

[6] C. Yeh and C. Kuo, "Digital watermarking through quasi m-arrays," Proc. IEEE Workshop on Signal Processing Systems, Taipei, Taiwan, pp. 456-461, 1999.

[7] D. L. Gall, "MPEG: a video compression standard for multimedia applications," Communications of the ACM, 34(4), pp. 46-58, 1991.

[8] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding: a survey," Proceedings of the IEEE, 87(7), pp. 1062-1078, 1999.

[9] G. R. Kwon, T. Y. Lee, K. H. Kim, J. D. Jin, and S. J. Ko, "Multimedia digital right management using selective scrambling for mobile handset," LNAI, 3802, pp. 1098-1103, 2005.

[10] P. Bassia, I. Pitas, and N. Nikolaidis, "Robust audio watermarking in the time domain," IEEE Trans. on Multimedia, 3(2), pp. 232-241, 2001.

[11] S.E Borujeni, "Speech encryption based on fast Fourier transform permutation," Proc. of The 7th IEEE International Conference on Electronics, Circuits and Systems, pp. 290-293, 2000.

[12] S. Farkash., S. Raz., and D. Malah, "Analog speech scrambling via the Gabor representation," Proc. of the 17th Convention of Electrical and Electronics Engineers in Israel, pp. 365-368, 1991.

[13] W. Li, H. Xinping, and H. Leung, "Performance evaluation of digital audio watermarking algorithms," Aerospace and Electronic Systems, IEEE Trans. on., 40. pp. 12-26, 2004.

권 구 락

• 2008년~현재 조선대학교 정보통신공학과 조교수
• 2006년~2007년 ㈜달리텍 대표이사
• 2002년~2007년 고려대학교 메카트로닉스 공학박사
• 관심분야: 멀티미디어 신호처리, 표준압축코덱, 정보보안