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/© 2014 Elsevier Inc. All rights reserved.

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-