A Secure 3D-SPIHT Codec

Shiguo Lian, Jinsheng Sun, Zhiquan Wang
Department of Automation
Nanjing University of Science and Technology
Nanjing, P.R China 210094
E-mail: *sg_lian@163.com*

Abstract

Due to the properties of large-volume and real-time requirement, video is difficult to be encrypted directly by traditional cryptography. Video encryption algorithms taking compression process into account are preferred. In this paper, an encryption scheme combining with 3D-SPIHT encoding is proposed, in which, wavelet coefficients in each data cube are confused based on octree confusion and their signs are encrypted by a chaotic stream cipher. What's more the relative positions of data cubes are confused. The encryption scheme changes the compression ratio slightly, is of low-cost and supports direct bit-rate control. These advantages make it suitable for multimedia applications with real-time operation requirement such as multimedia network, wireless multimedia, mobile multimedia and so on.

Key words *video encryption, chaotic encryption, 3D-SPIHT, octree encoding, information security*

1. Introduction

Multimedia encryption technology has become more and more important with the rapid increasing of multimedia information on the Internet. Compared with text data, multimedia data is often of large-volume and requires real-time processing, which makes it difficult to encrypt multimedia data directly with traditional cryptography. Therefore, multimedia encryption algorithms combining encryption process with compression process are preferred. In order to satisfy application requirements such as real-time operation, position seek, data cut or paste, bit-rate control and so on, multimedia encryption algorithms should be secure, fast, and keep compression ratio and data format nearly unchanged.

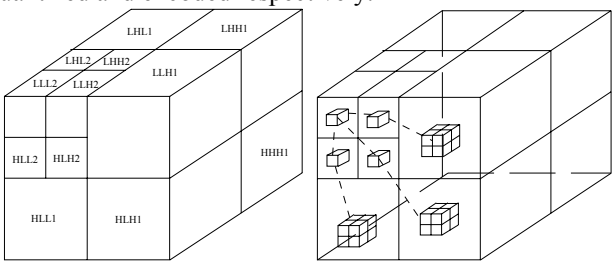
In the past decade, some encryption algorithms have been proposed for MPEG encoded videos, such as encrypting DCT coefficients [1-3], encrypting motion vectors [4] and encrypting data format information [5]. They can satisfy the requirements proposed above in different degree. Recently, wavelet transformation has been widely used in multimedia compression, for its advantages of good spatiotemporal property and easy to produce progressive bit-stream. In order to protect multimedia data compressed by wavelet transformation, some multimedia encryption algorithms based on wavelet transformation have been proposed. For example, some algorithms encrypt images by confusing wavelet coefficients [6,7], and some algorithms encrypt only tree-structure information [8]. However, the algorithms based on wavelet coefficient confusion change compression ratio greatly and the ones based on tree-structure encryption are not suitable for direct bit-rate control.

Wavelet zerotree encoding has been widely used in image encoding. And for the properties of high compression ratio and progressive coding stream, it has been extended to three-dimension space and used in large-volume data compression such as video, medical image and multispectral image. A. Bilgin [9] proposed a 3D-wavelet encoding scheme based on octree structure, Pearlman extended SPIHT [10] to 3D-SPIHT [11], and G. Menegaz [12] proposed a layered wavelet zerotree encoding scheme. These encoding schemes inherit the properties of 2D encoding schemes, such as high compression ratio and progressive coding stream, and can realize lossy compression or lossless compression. In this paper, we propose a secure 3D-SPIHT encoding scheme that combines encryption process with compression process. It is secure, cost-effective and easy to realize direct bit-rate control. What's more, it keeps compression ratio nearly unchanged. The rest of the paper is arranged as follows. Section 2 gives a brief introduction to 3D-SPIHT encoding. The secure encoding scheme is described in Section 3. The cryptosystem's security is analyzed in Section 4. In Section 5, various experiments are done to test its properties. Finally, some conclusions are drawn in Section 6.

2. 3D-SPIHT Encoding

In SPIHT, the relationship among coefficients lying in different frequency bands is based on quad-tree structure, while the one is based on octree structure in 3D-SPIHT. Given an image sequence to be encoded, 7L+1 subband image cubes are produced after L-level 3D wavelet transformation. A 2-level 3D wavelet transformation is shown in figure 1(a), and figure 1(b) shows a complete wavelet coefficient tree. Where, the coefficients lying in the lowest frequency subband (LLL2) are the roots of octrees, each of which has seven child-nodes that lie respectively in the subands (LLH2, LHL2, LHH2, HLL2, HLH2, HHL2 and HHH2) in seven directions. And each child-node has also its 8 children-nodes that lie in more refined subbands

in the according directions. Thus, except the coefficients in the lowest subband (LLL2) and in the highest subbands (LLH1, LHL1, LHH1, HLL1, HLH1, HHL1 and HHH1), each coefficient has 8 children-nodes. A complete octree is shown in figure 3. In 3D-SPIHT encoding, in order to get higher encoding speed, image sequence is often divided into smaller sized data cubes, such as $16 \times 16 \times 16$ or $32 \times 32 \times 32$. Then each data cube is wavelet transformed, quantized and encoded respectively.



(a) 3D Wavelet Decomposition (b) Zerotree Structure

Figure 1. 3D-SPIHT Encoding

3. The Proposed Secure Encoding Scheme

The secure encoding scheme is composed of three parts: intra-cube encryption, inter-cube encryption and key distribution. Where, intra-cube is composed of two steps: coefficient confusion based on octree confusion and coefficient sign encryption. Inter-cube encryption means to confuse the relative positions of data cubes. And a key distribution scheme is proposed to strengthen the cryptosystem. They are presented in details as below respectively.

3.1 Intra-Cube Encryption

The intra-cube encryption algorithm combines encryption process with compression process, which is shown in figure 2. The encryption process (dashed frame) composed of coefficient confusion and sign encryption is applied after wavelet transformation and quantization, and before zerotree encoding. Where, coefficient confusion is divided into two steps: coefficient confusion in an octree and coefficient confusion among different octrees. Each of them will be explained in details.

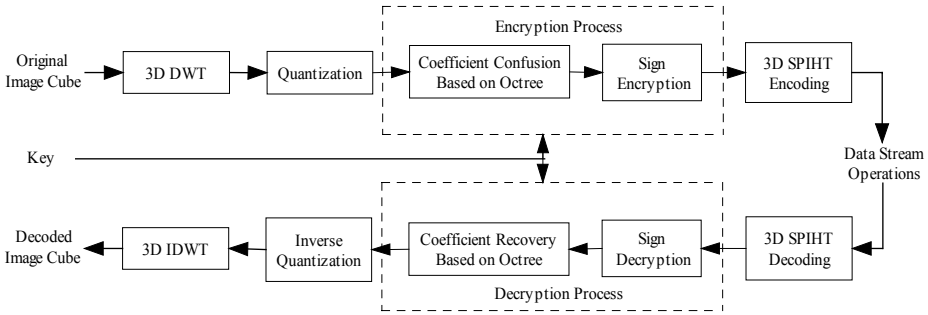


Figure 2. A Secure 3D-SPIHT Encoding Scheme

Step 1. Coefficient confusion in an octree

Considering the octree structure of wavelet coefficients, this step confuses children-nodes belonging to the same parent-node, and keeps the parent-child relationship between different coefficients unchanged, which is shown in Fig. 3. And this kind of confusion is applied along an octree from the root to the highest leaves. Thus, the coefficients in the octree are all confused. For a data cube, if L-level wavelet transformation is applied to it, the generated octree is of L-height. Then all the coefficients from 1-height to L-height are confused in this step.

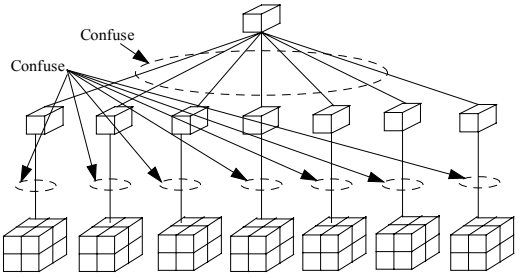


Figure 3. Octree Confusion

Step 2. Confusion among different octrees

Generally, more than one octrees are produced in 3D-SPIHT encoding. Permuting the positions of different octrees can change the coefficient distribution, thus make the decoded images too chaotic to be understood. In 3D-SPIHT encoding, the roots of the octrees lie in the lowest frequency subband. So the confusion process permutes the coefficients in the lowest frequency subband and keeps each octree structure unchanged. For example, in Fig. 1(b), if an $N \times N \times N$-sized data cube is transformed by L-level wavelet transformation, then $\left(\frac{N}{2^L}\right)^3$ L-height octrees are produced. Thus, the confusion process is done among $\left(\frac{N}{2^L}\right)^3$ coefficients, and the confusion space is $\left(\frac{N}{2^L}\right)^3!$.

Step 3. Coefficient Sign Encryption

In wavelet transformation, changes of coefficient signs affect decoded images' understandability. In order to realize high speed and direct bit-rate control, we propose to encrypt coefficients' signs by a chaotic stream cipher here. If a coefficient is lower than zero, its sign is represented as '0', otherwise as '1'. Thus, the signs of all the wavelet coefficients consist of a sign sequence that is modulated (bitwise/XOR-operation) by binary pseudo-random sequence generated from a chaotic sequence generator. Here, the chaotic sequence generator is based-on Logistic map

$$x(k+1) = \mu x(k)\left[1 - x(k)\right].$$

Where, we take $\mu = 4$ and $k \in \{0,1,\cdots,n-1\}$. Initial state x(0) is the key, and the chaotic sequence $x(0), x(1), \cdots, x(n)$ is produced through iterated chaotic map. If $x(i) = 0.b(0)b(1)b(2)b(3)\cdots$, then binary chaotic sequence $b(0), b(1), \cdots, b(m)$ may be constructed by extracting the first m bits of x(i). Thus, n states can generate a binary sequence with m × n bits. The parameter m should be determined according to computer resolution and security requirement. The bigger m is, the higher the key-sensitivity is, while the higher the computer resolution is required. Through various experiments, we recommend 7<m<17.

3.2 Inter-Cube Encryption

The previous steps give the methods to encrypt a data cube. In this step, the relative positions among different data cubes are confused to make the decoded images more chaotic. In 3D-SPIHT encoding, image sequence is often divided into smaller sized data cubes and then encoded cube by cube. For an image sequence of size F × W × B (F is the number of frames, W is the width of a frame and B is the height of a frame), if the data cube is of size F × F × F, then the number of data cubes is $\frac{W}{F} \times \frac{B}{F}$. Thus, the confusion space of the proposed inter-cube confusion is $\left(\frac{WB}{F^2}\right)!$. For example, if F=16, W=240 and B=352, then the number of data cubes in this image sequence is 15 × 22=330, and the confusion space is 330!.

3.3 Key Distribution Scheme

In 3D-SPIHT encoding, each data cube is encoded respectively. So we propose to encrypt each data cube with different key. For an image sequence of size F × W × B (F is the number of frames, W is the width of a frame and B is the height of a frame), if the data cube is of size F × F × F, then the number of data cubes is $\frac{W}{F} \times \frac{B}{F}$. Thus, we propose a key distribution scheme based on the spatiotemporal chaos named One-way Coupled Map Lattice (OCOML) that has been widely used in data encryption because it can generate pseudo-random sequences with high security [13]. The model is shown below.

$$\begin{cases} k(i+1, j) = (1-\varepsilon) f(k(i, j)) + \frac{1}{2}\varepsilon\{f[k(i, j-1)]\} \\ f(x) = 4x(1-x) \end{cases}$$

Where, $0 < \varepsilon < 1$, $1 \le i \le \frac{W}{F}$ and $1 \le j \le \frac{B}{F}$. For k(i,j) ranges between 0 and 1, it can be represented as binary fraction $k(i, j) = 0.a(0)a(1)a(2)\cdots$, then binary number $K(i, j) = a(0)a(1)\cdots a(m)$ may be constructed by extracting the first m bits of k(i,j). And K(i,j) is just the key of the (i,j)-th data cube.

4. Security Analysis

In 3D-SPIHT encoding, the octree of wavelet coefficients has a property that only the root has 7 child-nodes in 1-height layer and other nodes always have 8 children-nodes shown in Fig. 3. In an H-height (H>1) octree, there are $7 \cdot 8^{H-1}$ leaves and $(8^{H-1} - 8)$ internal nodes. And in h-height ($h \le H$) layer, there are $7 \cdot 8^{h-1}$ nodes. Denote C(h) as the number of confused octrees in h-height. The confusion space of an octree is

$$C_H = \prod_{h=1}^{H} C(h) = (7!) \cdot \prod_{h=2}^{H} (8!)^{7 \cdot 8^{h-2}} = (7!) \cdot (8!)^{\left(8^{H-1}-1\right)}.$$

Supposing T octrees are produced during the encoding process, the encryption space of intra-cube encryption is

$$P = (C_H)^T \cdot (T!) \cdot P_S = \left((7!) \cdot (8!)^{\left(8^{H-1}-1\right)}\right)^T \cdot (T!) \cdot 2^M.$$

Where $P_S = 2^M$ is the space of sign encryption, and M is the number of sign-encrypted coefficients. For example, for a $16 \times 16 \times 16$ sized data cube, 8 octrees of 3-height are produced after 3-level wavelet transformation. Thus, the encryption space of coefficient confusion, sign encryption (encrypting only the coefficients in the lowest frequency subbands) is 4.71×10^{2374}. That's only the brute-force space of a data cube. For brute-force attackers, repeated encoding and decoding operations are required, which makes it difficult to walk through the brute-force space.

Using the proposed key distribution scheme, each data cube is assigned a key different from others, which further increases the difficulty of brute-force attack. What's more, it also increases the difficult for some statistics attackers and differential attackers to destroy the cryptosystem, since the correlation between adjacent data cubes cannot be used. Furthermore, the key distribution and sign encryption are realized by chaotic stream ciphers whose security against known-plaintext attacks and select-plaintext attacks keep the secure encoding scheme secure against these attacks too.

5. Experimental Results

5.1 Direct Bit-rate Control

In this experiment, we test the PSNRs (Peak Signal-to-Noise Ratios) of original and encrypted image sequence Bus under various bit-rates (Fig. 5). The cube size is $32 \times 32 \times 32$ and the wavelet transform is 3-level. The curves in Fig. 5 show that, compared with the original image sequence, the image blur caused by the proposed encryption scheme is not more than 0.4dB, which can be accepted in most applications. This implies that the secure encoding scheme supports direct bit-rate control.

5.2 Encryption Results

The encryption results of video Bus and Flower are shown in Fig. 6. The cube size is $16 \times 16 \times 16$ and the wavelet transform is 3-level. The encrypted pictures (c) and (d) are so confused that they cannot be understood at all, which shows that the encryption scheme is of high security.

Experiments are also done to test the relationship between encryption speed and transform level, and the one between encryption speed and cube size. Encryption speed is measured by encryption time ratio (Etr) that is the time ratio between encryption process and encoding process. Experimental results show that, the encryption time ratios are not bigger than 10% when cube size is not bigger than $32 \times 32 \times 32$ and transform level is not bigger than 4. It shows that the algorithm is of low cost, which is easy to satisfy real-time requirement.

6. Conclusions

A secure 3D-SPIHT encoding scheme combining encryption process with compression process is proposed in this paper. Theoretical analysis and experimental results show that, the encryption scheme is secure, fast and supports direct bit-rate control. These properties make it suitable for multimedia applications with real-time and high compression requirements, such as secure encoding of video, image sequence, medical image and multispectral image, especially for multimedia network or mobile multimedia applications.

Acknowledgement

This work was supported by the National Natural Science Foundation of China through the grant numbers 60174005 and 70271072.

References

[1] Ali Saman Tosum and Wu-chi Feng, "Efficient multi-layer coding and encryption of MPEG video streams," IEEE International Conference on Multimedia and Expo (I) 2000, ICME 2000, 30 July-2 Aug. 2000, Page(s): 119 -122 vol.1.

[2] L. Qiao, K. Nahrstedt, and I. Tam, "Is MPEG encryption by using random list instead of zigzag order secure," IEEE International Symposium on Consumer Electronics, December 1997. Singapore.

[3] Changgui Shi, Bharat Bhargava, "A fast MPEG video encryption algorithm," In Proceedings of the 6th ACM International Multimedia Conference, Bristol, UK, pages 81-88, September, 1998.

[4] Jiu-Cheng Yen and Jiu-In Guo, "A new MPEG encryption system and its VLSI architecture," IEEE Workshop on Signal Processing Systems, Taipei, 1999, pp.430-437.

[5] L. Tang, "Methods for encrypting and decrypting MPEG video data efficiently," In Proceedings of the Fourth ACM International Multimedia Conference (ACM Multimedia'96), pages 219-230, Boston, MA, November 1996.

[6] Takeyuki Uehara, "Combined encryption and source coding," http://www.uow.edu.au/~tu01/ CESC.html.

[7] Wenjun Zeng and Shawmin Lei, "Efficient frequency domain selective scrambling of digital video," IEEE Trans, Multimedia 2002.

[8] Howard Cheng and Xiaobo Li, "Partial Encryption of Compressed Images and Videos," *IEEE Transactions on Signal Processing*, v 48, n 8, Aug, 2000, p 2439-2451.

[9] Amir Said, "A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," IEEE Transactions on Circuits and Systems for Video Technology, Vol.6, June 1996.

[10] A. Bilgin, G. Zweig, and M. W. Marcelin, "Three-dimensional compression with integer wavelet transform," Applied Optics, vol. 39, no. 11, pp. 1799-1814, April 2000.

[11] Z. Xiong, X. Wu, D. Y. Yun, and W. A. Pearlman, "Progressive coding of medical volumetric data using three-dimensional integer wavelet packet transform," Proceedings of SPIE Conference on Visual Communications, Vol. 3653, pp. 327-335, Jannuary 1999.

[12] G. Menegaz, L. Grewe, and J.-P. Thiran, "Multirate Coding of 3D Medical Data," Proceedings of IEEE International Conference on Image Processing (ICIP), Vancouver, Cannada, Vol. III, pp. 656-659, September 11-13, 2000.

[13] Wang Shihong, Kuang Jinyu, Li Jinghua, et al, "Chaos-based communications in a large community," Phys Rev, 2002, 66(6): 1-4.

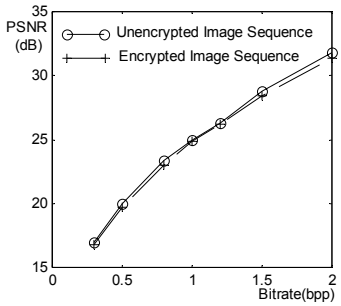
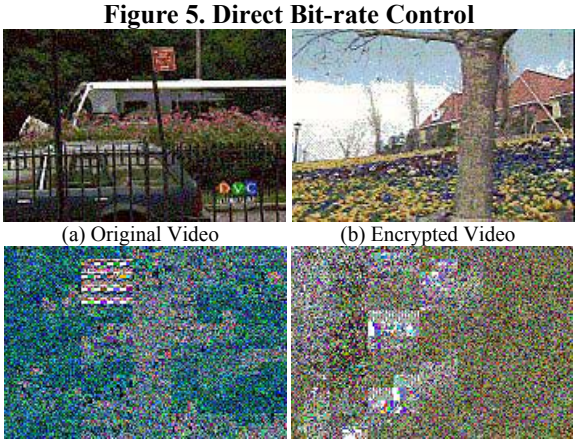


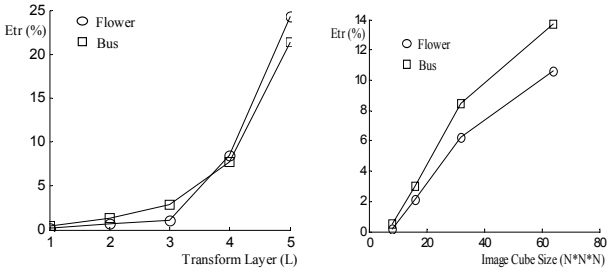
Figure 5. Direct Bit-rate Control



(a) Original Video (b) Encrypted Video

(c) Original Video (d) Encrypted Video

Figure 6. Encryption Results of Video Sequence



(a)Relationship Between Encryption Speed and Wavelet Transform Layer (b)Relationship Between Encryption Speed and Data Cube Size

Figure 7. Encryption Speed Test