DCLSA:一种适用于 H.264/AVC 的 DCT 系数分层置乱算法

包先雨 蒋建国 李 援
(合肥工业大学计算机与信息学院,合肥 230009)

摘 要 基于目前的 DCT 系数加密算法在安全性、压缩比和信噪比上都各自存在缺点,提出了一种新的适用于 H.264/AVC 的 DCT 系数分层置乱算法(DCLSA)。该算法针对 H.264/AVC 中 4×4 DCT 变换的特点,首先将同一宏块中每个 4×4 块 DCT 系数进行块间分层,构建系数分层模型,然后可根据安全性要求对不同层系数进行随机置乱,实现加密编码。通过性能比较和具体实验效果分析,此算法具有更高的安全性、更优的压缩比和较好的信噪比,适合于 H.264/AVC 的网络安全应用。

关键词 H.264/AVC DCT 系数置乱 视频加密 密钥同步
中图法分类号:TP309.7 TP37 **文献标识码**:A **文章编号**:1006-8961(2008)04-0618-06

DCLSA: A DCT Coefficients Layered Scrambling Algorithm for H.264/AVC

BAO Xian-yu, JIANG Jian-guo, LI Yuan
(*Department of Computer and Information, Hefei University of Technology, Hefei* 230009)

Abstract This paper presents a novel DCT coefficients layered scrambling algorithm (DCLSA) for H.264/AVC based on the deficiencies of current DCT coefficients encryption algorithms in respect of security, compact ratio and signal-to-noise ratio. According to the characteristics of the 4×4 DCT transform of H.264/AVC, the algorithm first divides the coefficients of every 4×4 block of the same macro-block into several layers, and build a coefficients layered model. Then different layer is scrambled, respectively, by security requirements to achieve secure video coding. DCLSA has shown significant advantages on security, compact ratio and signal-to-noise ratio through performance comparisons with other algorithms and concrete experimental results, thus making it especially suitable for secure network applications.

Keywords H.264/AVC, DCT coefficients scrambling, video encryption, key synchronization

1 引 言

随着视频压缩编码技术和网络传输技术的快速发展,视频及视频相关服务(如数字电视、在线影视、视频聊天、视频会议等)的安全性正逐渐成为研究的热点。由于视频具有码流结构特殊、数据量大和实时性要求高等特点,所以需要针对其编码性质设计特殊的加密算法。

目前的视频加密算法根据压缩和变换方式的不同,可分为基于早期标准(如 H.263, MPEG-2 和 MPEG-4 等)的加密方式和基于 H.264/AVC[1] 的加密方式;根据编码流程又大致可分为压缩前信源数据加密、预测后预测模式加密、变换/量化后 DCT 系数加密、熵编码过程中加密和压缩编码后数据加密等。本文主要研究 H.264/AVC 标准中变换/量化后 DCT 系数加密。

基金项目:国家自然科学基金项目(60474035);安徽省"十五"二期科技攻关重大计划项目(040020382)
收稿日期:2006-08-14;**改回日期**:2006-11-28
第一作者简介:包先雨(1981~),男。现为合肥工业大学计算机与信息学院博士研究生。现主要研究兴趣为多媒体安全与版权保护。E-mail:baoxianyu@163.com

2　相关工作分析

2.1　基于早期标准的 DCT 系数加密算法

早期标准对预测残差采用了基于 8×8 DCT 变换。针对该变换域,目前存在很多 DCT 系数加密算法。Zigzag-permutation[2]是最早的 DCT 系数加密算法,属变换后块内系数置乱。它的优点是理论置乱空间可达 64!,安全性很高,但由于系数统计特性完全被破坏,熵编码后的压缩比降低约 50%。Tosun 和 Feng 对其做了改进,将块内 64 个系数按频带划分为基本层(0~3)、中间层(4~17)和增强层(18~63),然后根据安全性要求选择不同层进行置乱[3]。该算法虽然可以获得相对较高的压缩比,而且支持多重安全级别,但理论最大置乱空间为 4!×14!×46!<64!,安全性降低。此外,由于块内 DCT 系数能量大小具有 Zigzag 排序的特点,该类置乱算法都无法防止 FBA (frequency-based attack) 攻击[4]。

VEA[5]是对所有 DCT 系数符号位进行加密。它通过一个长为 m 的二进制序列 b_i 作为密钥 k,明文为 $M_S=s_1s_2\cdots s_ms_{m+1}\cdots s_{2m}$,其中,$s_i$ 为系数符号位,加密后的密文 $C_S=E_k(M_S)=(b_1\oplus s_1)(b_2\oplus s_2)\cdots(b_m\oplus s_m)(b_1\oplus s_{m+1})\cdots(b_m\oplus s_{2m})$。此算法简单、易实现,并且密钥搜索空间 2^m 可通过改变 m 的长度来调节;缺点是随着加密数据量增大,密钥重复次数多,难以抵抗已知明文攻击。

Zeng 和 Lei 提出了两种 DCT 系数加密的一般性方法[6]:(1)将 I 帧分成多个片段(Segment),然后在每个片段内进行块间置乱;(2)块的旋转。这些方法均具有较高的安全性,例如,若一个片段包含 n 个块,其中有 s 个全零块和 t 个相似块,则其置乱空间为 $n!/(s+t)!$;若每块有 8 种旋转方式,其块旋转空间为 $(8\times n)!/s!$。其缺点是对信噪比和压缩比都有较大的影响,而且不能防止块间相关性攻击[9]。

2.2　基于 H.264/AVC 的 DCT 系数加密算法

曹弈提出了 3 种基于 4×4 DCT 的 H.264/AVC 视频加密算法[7],这些算法最大的特点是加密过程取在量化之后熵编码之前,实现简单,密钥开销较小,但在安全性、压缩比和信噪比上都各自存在缺点。下面分别对这三种算法进行分析:

算法 1　DCT 系数符号的翻盘

此算法对所有 4×4 块中的 DCT 系数符号位进行加密,实现简单且不影响压缩比,但它没有考虑到系数熵编码的特点,只是以较多的计算资源换取了更高的安全性。文献[8]提出了仅加密拖尾系数(±1)符号位的方法,极大地减少了计算开销。另外,符号加密没有改变系数的统计特性,安全性较低,一般只用于视频透明加扰。

算法 2　4×4 块内系数的洗牌(B_Shuffling)

将同一宏块中每个 4×4 块作为基本单元进行随机置乱,而保持块内系数相对位置不变。去掉 n 个全零块和相似块,则一个 16×16 宏块的加密复杂度为 16!/n!,安全性较高。其缺点是:(1)由于每个 4×4 块内部信息完全保留,攻击者可以利用块间相关性(如边缘连续性,色彩和纹理相似性等)进行破译;(2)洗牌过程改变了相邻块系数的统计特性,并且系数的位置变动明显,因而对信噪比和压缩比都具有较大的影响。

算法 3　块内非零高低频系数之间洗牌(C_Scrambling)

考虑到实际编码时存在全零块和量化后系数许多为零等情况,假设非零系数占总系数的 1/3,则每个 4×4 块系数置乱空间为(16/3)! ≈5!,整个宏块置乱空间为$(5!)^{16}$,安全性较高;若将同一个 8×8 或 16×16 块中非零系数进行置乱,则安全性更高。与文献[2]、[3]算法相比较,其优点是洗牌过程只改变了块内非零系数的统计特性,对压缩比影响相对较小,但由于 4×4 块内非零系数数量很少,因此也更容易使用 FBA 攻击来进行解密。

3　建议的 DCLSA 算法

根据对以上算法的分析,现有的 DCT 系数加密算法在安全性、压缩比和信噪比上都各自存在缺点。为此,本文利用 H.264/AVC 中 4×4 DCT 变换的特点,通过构建 DCT 系数分层模型,提出了下面的 DCLSA 算法,同时还建议了一种简单的密钥分发与同步方法。

3.1　DCT 系数分层模型

在模型建立之前,首先将每个宏块以 4×4 为单位分成 16 个块:$\boldsymbol{Z}_0,\boldsymbol{Z}_1,\cdots,\boldsymbol{Z}_{15}$,每块包含 16 个 DCT 系数,即 $\boldsymbol{Z}_i=\{C_{i,0},C_{i,1},\cdots,C_{i,15}\}$ $(0\leqslant i\leqslant15)$。如 2.2 小节所述,现有的加密算法如 B_Shuffling 即是对 $\boldsymbol{Z}_0,\boldsymbol{Z}_1,\cdots,\boldsymbol{Z}_{15}$ 这 16 个块进行块间随机置乱,而块内系数相对位置不变;C_Scrambling 即是对每个块 $\boldsymbol{Z}_i$ 内非零系数进行随机置乱。由于这些算法在安全性、压缩比和信噪比方面都各自存在不足,这里按如下步骤构建了一个新的 DCT 系数分层模型:

(1) 将同一宏块中每个 4×4 块 $\boldsymbol{Z}_i$ $(0\leqslant i\leqslant15)$

包含的 2 维系数矩阵映射成 1 维系数矩阵,映射函数为 Zigzag 排序,如图 1 所示。

(2) 一个宏块按步骤 1 映射 16 次后,可得到对应的 16 个 1 维系数矩阵,再将这些系数矩阵组合成一个 2 维 16×16 系数矩阵,如图 2 所示。

(3) 将此 2 维系数矩阵按行顺序(即按系数能量 Zigzag 排序)分成 16 层(Layer 0, Layer 1, …, Layer 15),建立如图 2 的分层模型,其中 Layer i = $\{C_{0,i}, C_{1,i}, \cdots, C_{15,i}\}$ $(0 \leqslant i \leqslant 15)$。

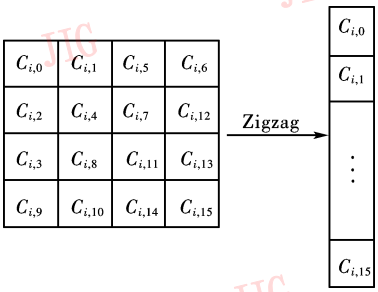


图 1 $\boldsymbol{Z}_i$ $(0 \leqslant i \leqslant 15)$ 系数矩阵映射

Fig. 1 Coefficients matrix mapping of $\boldsymbol{Z}_i$ $(0 \leqslant i \leqslant 15)$

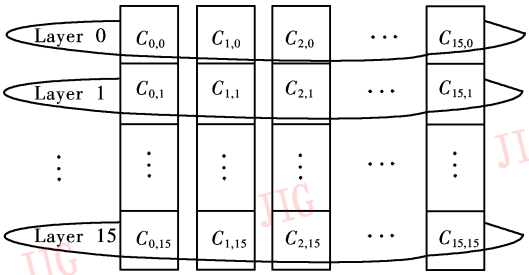


图 2 DCT 系数分层模型

Fig. 2 DCT coefficients layered model

3.2 DCLSA 算法实现

从系数矩阵映射过程可以看出,$C_{i,0}$ 代表宏块中 $\boldsymbol{Z}_i$ $(0 \leqslant i \leqslant 15)$ 块的直流(DC)系数,其余为交流(AC)系数。即 Layer 0 包含了宏块中 16 个 4×4 块的 DC 系数,本文将其定义为 DC 系数层;Layer i $(1 \leqslant i \leqslant 15)$ 包含了宏块中 16 个能量相当的 AC 系数,本文将其定义为 AC 系数层 i $(1 \leqslant i \leqslant 15)$。

图 3 是一通用型分层置乱算法,可以根据安全性要求选择其中的一层或多层系数分别进行随机置乱。值得一提的是,若仅加密 DC 系数层或仅加密所有 AC 系数层,则无法防止逼近攻击,如图 4 所示,甚至可以使用模式识别技术近似恢复原有视频图像,因而建议对 DC 系数层和 AC 系数层都进行置乱。同时考虑到每个 4×4 块右下角系数经量化后很多为零值,假设非零系数占总系数的 1/3,则图 2 中 Layer 4 之后绝大部分系数均为零,所以在后续实验中只对前 5 层(Layer 0, Layer 1, …, Layer 4)系数分别进行了随机置乱,置乱层次越多,安全性越高。又由于每层系数在各自的 4×4 块中所占能量相当,使用了文献[10]中的 Controller 来控制系数的置乱程度,以增强置乱效果。

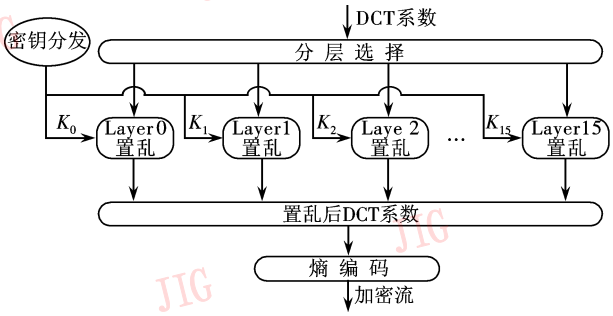


图 3 通用型分层置乱算法

Fig. 3 Generic layered scrambling algorithm

(a) 设置所有DC系数为0 (b) 设置所有AC系数为0

图 4 逼近攻击(去色度信息)

Fig. 4 Approximation attacks (removing chroma)

3.3 密钥分发与同步

如图 5 所示,密钥分发以帧为基本单位,每个密钥序列的更新单位为图组。为了节省密钥开销和减小计算资源消耗,采用密钥公用方法:DC 系数层(Layer 0)使用一个密钥,其他 AC 系数层 i $(1 \leqslant i \leqslant 15)$ 公用一个密钥。若图像大小为 $H \times W$,则密钥序列的长度为 $H \times W \times 2/(16 \times 16)$。每帧开始加密时,密钥序列指针复位,逐层置乱,直到该帧结束。

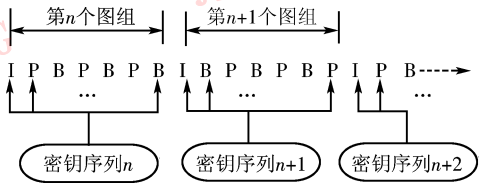


图 5 密钥分发与更新

Fig. 5 Key distribution and refresh

对于新用户开始接入系统时的密钥同步,建议

在 H.264/AVC 中的序列参数集和图像参数集之间填充 32 位二进制序列 $p_0p_1\cdots p_{31}$。虽然增加了些码长,但对一个图组来说是忽略不计的。又因为两种参数集所在的 NALU(网络适配层单元)起始码相同(H.264/AVC 规定为 0x000001),只要保证填充序列 $p_0p_1\cdots p_{31}$ 内部不出现与起始码相同的二进制串,就不会引起解码出错。新用户接入系统时,首先从码流中解析出填充序列,然后使用该序列快速迭代计算出当前图组的密钥序列,最终实现密钥重同步。

该方法的优点在于:当网络传输中出现丢包、丢帧等情况时,通过每帧开始加密时将密钥序列指针复位,不会影响下一帧的正常解密;在码流中插入填充信息,不仅可实现密钥重同步,而且兼容标准码流;密钥存储空间小,如 $H\times W$ 图像的密钥存储数量为 $H\times W\times 2/(16\times 16)$。

4　性能分析

主要从安全性、信噪比和压缩比等角度对 DCLSA 算法性能进行了分析。为了进一步证明本算法的有效性和实用性,还将其与文献[7]中 B_Shuffling和 C_Scrambling 算法进行了对比。

4.1　安全性分析

通过构建 DCT 系数分层模型,将每层系数分别进行随机置乱,属块间置乱算法。以下将主要分析 DCLSA 算法在穷举攻击、FBA 攻击和块间相关性攻击时的安全性。

与 B_Shuffling 和 C_Scrambling 算法相比较,DCLSA 算法对 FBA 攻击和块间相关性攻击具有较高的安全性,如表 1 所示。这是由于算法采用了分层置乱思想,将同一个宏块中每个 4×4 块能量相当的 DCT 系数进行了随机置乱,因而(1)置乱过程几乎不影响置乱后块内系数能量大小的 Zigzag 排序,可以有效地防止 FAB 攻击;(2)置乱过程改变了每个 4×4 块内信息,这就保证了 DCLSA 算法对块间相关性攻击的安全性。

表 1　不同加密算法抗攻击能力比较

Tab. 1　Anti-attack comparison of different algorithm

加密算法	穷举攻击	FBA 攻击	块间相关性攻击
DCLSA	√	√	√
B_Shuffling	√	√	×
C_Scrambling	√	×	√

DCLSA 算法还具有较大的明文搜索空间。以一个宏块为例,从 2.2 节分析可知,B_Shuffling 算法的明文搜索空间为 $M_1=16!/n!$,其中,n 为宏块中全零块和相似块的数量;C_Scrambling 算法的实际搜索空间为 $M_2=(5!)^{16}$。DCLSA 算法建议对同一宏块中前 5 层系数分别进行随机置乱,考虑到每层系数能量相当,假设每层系数有 1/3 相等,则一个宏块的明文搜索空间为 $M_3=(16!/5!)^5$。因此,这三种算法的明文搜索空间满足关系:$M_1<M_2<M_3$。即 DCLSA 算法的明文搜索空间最大,安全性最高。

另外,算法的实际穷举空间应为 $\min\{|\text{明文搜索空间}|,|\text{密钥空间}|\}$,其中,$|\cdot|$为集合的势。因为 DCLSA 算法采取 DC 系数层和 AC 系数层各公用一个密钥,所以其密钥空间远小于明文搜索空间。以 352×288 视频序列为例,每个图组使用的密钥流长度为 $352\times288\times2/(16\times16)=792$ 位,因此采用穷举攻击来进行解密是很困难的。又由于密钥序列在每个图组的实时变化,算法还可以防止已知/选择明文攻击。

4.2　信噪比与压缩比分析

从信噪比的影响来看,DCLSA 算法改变了相邻块系数的统计性质,对信噪比有些影响,而且这种影响是随机的,即可能优化信噪比,也可能使其变差,如表 2 所示。与块间置乱算法 B_Shuffling 相比较,DCLSA 算法可以获得更优的信噪比。

从压缩比的影响来看,C_Scrambling 算法只是破坏了块内非零系数的统计性质,与以往的块内系数置乱[2]和分段置乱算法[3]相比较,可获取较好的压缩比;B_Shuffling 算法虽然没改变块内系数的统计性质,但系数位置出现了巨大变动,对压缩比有一定的影响,如表 2 所示。通过比较,本文建议的 DCLSA 算法对压缩比影响最小,这是由于分层置乱几乎不影响块内系数能量大小的 Zigzag 排序,而且系数位置的变动相对较小。

4.3　实验结果

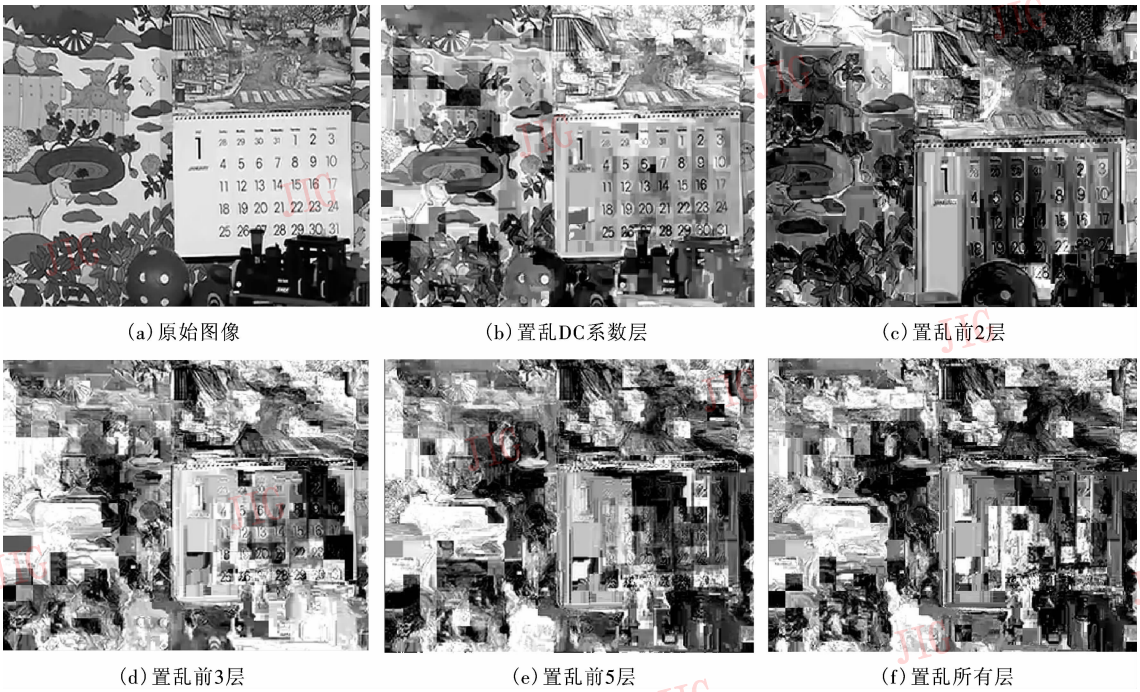
采用标准的 Mobile(352×288,去色度信息)序列测试,加密后图像如图 6 所示。从加密效果可以看出,当仅对 DC 系数层(Layer 0)进行置乱时,图像视觉安全性较差;随着置乱层次的增加,图像变得越来越混乱。直至加密前 5 层后,图像基本难以辨别,而且其视觉效果与置乱所有层的安全性相当,这是由于第 5 层之后的系数绝大部分为零值。因此为了节省计算资

表 2 加密前后信噪比和压缩比比较

Tab. 2 Comparison of compact ratio and SNR before and after Encryption

视频序列 QCIF	信噪比(dB)			压缩比			
	加密前	DCLSA	B_Shuffling	加密前	DCLSA	B_Shuffling	C_Scrambling
Foreman	34.26	34.26	34.19	89.33	89.29	90.94	86.47
News	34.83	34.84	34.81	155.50	153.81	154.41	150.36
Hall_monitor	35.52	35.52	35.51	185.27	184.69	185.53	180.83
Salesman	33.78	33.75	33.76	195.57	193.32	192.79	187.35
Silent	34.03	33.99	33.98	129.10	128.00	127.70	123.25
Container	34.64	34.67	34.66	303.11	302.05	296.40	297.51

注:测试条件为: JM82,100frames,iInterval = 9,bInterval = 1



(a)原始图像 (b)置乱DC系数层 (c)置乱前2层

(d)置乱前3层 (e)置乱前5层 (f)置乱所有层

图 6 分层置乱的实验效果图

Fig. 6 Test results of layered scrambling

源,建议对前 5 层系数进行随机置乱,同时也可根据安全性要求选择不同层系数进行置乱,实现加密编码。

5 结 论

针对 H.264/AVC 中 4×4 DCT 变换的特点,通过构建 DCT 系数分层模型,提出了一种新的 DCT 系数分层置乱算法 DCLSA。该算法的特点是:

(1) 较高的安全性:在保证加密视觉效果的条件下,算法对 FBA 攻击和块间相关性攻击具有较高的安全性,另外还支持多重安全级别;

(2) 更优的压缩比与信噪比:与 B_Shuffling 和 C_Scrambling 算法相比较,DCLSA 可以获得更优的压缩比和较好的信噪比;

(3) 网络实现的可行性:当网络传输中出现丢包、丢帧情况时,不会影响下一帧的正常解密;在参数集之间插入同步信息,可实现新用户初始接入系统时的密钥重同步。

本文构建的 DCT 系数分层模型实质上是宏块级系数能量大小的可分级模型,以此为基础,如何在 H.264/AVC 视频中嵌入数字水印是下一步工作的重点;另一个可能的方向是研究具有可控加扰程度的 H.264/AVC 透明加扰算法。

参考文献(References)

1 ITU-T Rec. H. 264 | ISO/IEC 14496-10:2005 (E). Advanced Video Coding for Generic Audiovisual Services [S].

2 Tang L. Methods for encrypting and decrypting MPEG video data efficiently [A]. In: Proceedings of the ACM Multimedia'96 [C], Boston, USA, 1996: 219 ~ 229.

3 Tosun A S, Feng W C. Efficient Multi-layer coding and encryption of MPEG video streams [A]. In: Proceedings of IEEE International Conference on Multimedia and Expo [C], New York, USA, 2000, **1**: 119 ~ 122.

4 Qiao L, Nahrstedt K. Comparison of MPEG encryption algorithms [J]. International Journal on Computers and Graphics, 1998, **22**(4): 437 ~ 448.

5 Shi C G, Bhargava B. An efficient MPEG video encryption Arithm [A]. In: Proceedings of the 6th ACM International Multimedia Conference[C], Brisol, United Kingdom, 1998, **9**:381 ~ 386.

6 Zeng W J, Lei S. Efficient frequency domain selective scrambling of digital video [J]. IEEE Transactions on Multimedia, 2003, **5**(1): 118 ~ 129.

7 Cao Yi, Zhang Rong, Liu Zheng-kai. Research on DCT-based video encryption under H.264 [J]. Journal of Image and Graphics, 2005, **10**(8): 1047 ~ 1051. [曹弈, 张荣, 刘政凯. H.264 标准中基于 DCT 的视频加密研究 [J]. 中国图象图形学报, 2005, **10**(8): 1047 ~ 1051.]

8 Bao Xian-yu, Jiang Jian-guo, Li Yuan. A new encryption scheme for H.264 real-time video transmission [J]. Acta Electronica Sinica, 2006, **34**(11):2099 ~ 2102. [包先雨, 蒋建国, 李援. 一种适合于 H.264 实时视频传输的新型加密方案 [J]. 电子学报, 2006, **34**(11):2099 ~ 2102.]

9 Mao Y N, Wu M. A joint signal processing and cryptographic approach to multimedia encryption[J]. IEEE Transactions on Image Processing, 2006, **15**(7): 2061 ~ 2075.

10 Liu Z, Li X. Motion Vector Encryption in Multimedia Streaming [A]. In: Proceedings of the 10th International Multimedia Modeling Conference(MMM'04)[C], Melbourne, Australia, 2004: 64 ~ 71.