视频信息隐藏的置乱策略与方法

陈真勇 唐 龙 唐泽圣
(清华大学计算机科学与技术系,北京 100084)

摘 要 随着网络和多媒体技术的飞速发展,视频数据的安全问题越来越突出。深入探讨了视频数据在空域、频域以及运动矢量上的置乱策略与方法。对如何提高置乱方法的安全强度,以及如何降低置乱对编码效率与视频图像质量的影响进行了详细分析,提出了置乱应遵循的原则。在此基础上,改进现有的算法,并通过实验进行了验证。实验结果表明,基于频域与运动矢量相结合的视频置乱策略与方法具有更强的安全性。与传统数据加密方法不同,置乱方法能够与视频编解码进行紧密的结合。

关键词 视频信息隐藏 MPEG-4 置乱
中图法分类号:TP391 **文献标识码**:A **文章编号**:1006-8961(2005)10-1242-06

Scrambling Strategy and Methods for Video Data Hiding

CHEN Zhen-yong, TANG Long, TANG Ze-sheng
(*Department of Computer Science and Technology, Tsinghua University, Beijing* 100084)

Abstract With considerable development of the network and multimedia technology, the security problem of the video data becomes more urgent. In this paper, the scrambling strategy and methods to hide the video data are discussed deeply in the spatial and frequency domains, followed by the security of these methods and the principles of the scrambling. The means to increase the security of the scrambling methods and reduce the influence of compression efficiency and video quality due to the scrambling are analyzed in details. Based on these, the existing methods are improved in the experiments. The experimental results show that the combination of spatial and frequency domain methods enhances the security of video data. Differing from the traditional encryption methods, these scrambling methods being proposed can be embedded into the video encoding and decoding processes compactly.

Keywords video data hiding, MPEG-4, scrambling

1 引 言

在当今信息时代与全球化背景下,对等主体(国家、企业、研究机构等)之间的竞争日益激烈,信息的安全问题越来越受到重视。视频信息作为多媒体信息的核心,在网络、通讯以及视频技术本身(如 MPEG-X 与 H.26X 系列视频编解码国际标准等)蓬勃发展推动下,它的发布与交流更加普及,例如视频出版物(video publications)、视频点播(video-on-demand, VOD)、可视电话和视频会议(videoconferences)等等,因此,人们迫切需要解决其在商业、军事以及其他私密背景应用中的安全性问题。

传统的密码学密钥加密方法是用来解决视频数据安全性问题的途径之一[1~3],但其不能与视频编解码过程进行很好地结合。为了得到合适的安全强度,密钥加密方法容易导致大量的额外开销,因为即使是编码后的视频压缩数据,也是非常庞大的。而在视频实时发布应用场合,额外开销带来的问题将会变得更加突出。为了避免这些问题,从另外一个

基金项目:国家自然科学基金重点项目(60133020)
收稿日期:2004-11-28;**改回日期**:2005-03-09
第一作者简介:陈真勇(1974~),男,博士后。主要研究方向为信息隐藏、数字水印、图形图像与可视化技术。
E-mail:chzhyong98@mails.tsinghua.edu.cn

途径来研究解决视频数据安全性的方法,即紧密结合视频编解码过程的视频置乱方法及其恢复,分别从空域、频域与运动向量角度来探讨提高视频数据安全强度,同时降低对编码效率和视频图像质量影响的置乱策略方法。

数字视频信号具有庞大的数据量,为了节省存储容量和传输带宽,通常以压缩方式进行存储或传输。目前视频压缩、解压缩的国际标准为 MPEG-X 和 H.26X系列,它们是以帧内块变换(DCT 或小波变换)和帧间运动预测补偿为基础的混合编码方案[4],图 1 是其编码过程。因此,视频数据的置乱方法可以从空域、频域与运动矢量这 3 个方面来进行探讨。

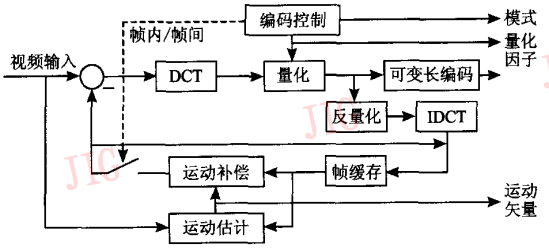

图 1　视频混合编码方案图解

Fig. 1　Video hybrid coding scheme

2　空域置乱

帧内块变换是针对一定大小的块单元进行变换,通常为 8×8 块,为了避免对原始视频数据在空间上的统计特性造成较大的影响,空域置乱以同样大小的块为单位。一种简单的处理方法是在编码前,对各帧原始数据(YUV 格式)在本帧范围内置乱,然后进行压缩编码,经过存储或传输,使用时在解码之后对置乱加以恢复,即如图 2 所示的外置置乱方式。这种方式操作简单,但对视频数据的时域统计特性影响较大,使帧间运动预测补偿编码效率显著降低,尤其是对常用的基于 16×16 宏块的运动补偿方式,因为置乱使原始宏块破碎,即使是基于 8×8块或更小的运动补偿方式,也会从总体上增大运动向量,且大量的块将超出运动搜索范围。

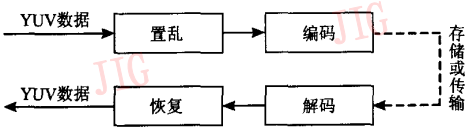

图 2　外置置乱方式

Fig. 2　Additive scrambling mode

如果在运动估计前将参考帧恢复,将能够消除置乱对帧间统计特性造成的影响,这就要求置乱恢复必须内嵌到视频编解码过程中。以 I 帧为例,如图 3 所示,在编码过程中,置乱恢复被内置于参考帧的重构与运动估计环节之间,相应地在解码过程中,运动补偿也要以恢复后的 I 帧为参考。通常情况下,对 I 帧置乱会直接导致依赖它的 P/B 帧也不可识别,这样已经能够满足诸如付费视频点播这类低安全级别要求的应用。但是,当 P/B 帧上剩余能量(主要是 P/B 帧的帧内编码块)多时,P/B 帧上剩余能量的积累会逐渐变得清晰,从而泄漏部分图像信息,这种情况在 I 帧出现频率低时尤其如此。因此,为了得到更高的安全强度,对 P/B 帧的置乱是必要的,但内嵌情况也会变得复杂一些,此时置乱只能针对 P/B 帧的帧内编码块,否则会影响 P/B 帧内运动补偿块的编码,并且这些 P/B 帧在作为其他帧的运动补偿参考之前需要恢复其帧内编码块的置乱。

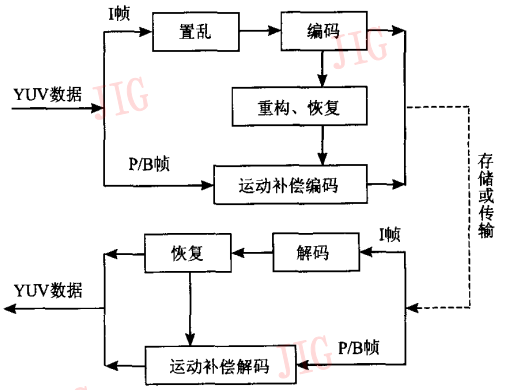

图 3　内嵌 I 帧置乱方式

Fig. 3　Embedded I-frame scrambling mode

3　频域置乱

前面描述的帧内空域置乱是以变换块(8×8 大小)为最小单元的,其块内纹理是可见的,这可能被非法攻击者所利用。因此,一些研究人员从频域角度来探讨视频数据的置乱,即在视频编码过程中,在对亮度 Y 与色度信号 UV 进行 DCT 变换之后,打乱这些系数的顺序,使得帧图像在正常解码情况下不可识别,即使块内纹理也是如此。Tang[5] 在 8×8 块内对 DCT 系数置乱,虽然达到了置乱效果,但却使

码率增加了 55% 左右，这是由于 DCT 系数的 zig-zag 扫描顺序被打乱，其统计特性受到较大影响造成的。另外，仅在块内作置乱有可能泄露大尺度上的图像信息。Zeng[6] 对此进行了改进，在 slice 范围内，将亮度信号的 DCT 系数按不同频带分开置乱，因为相同频带的数据其统计特性更加接近，色度信号也是如此，这在较大程度上改善了码率增加的问题，使码率增加降低到 23% 左右，但是这种改进在一帧图像由多个 slice 组成的情况下，仍有可能泄露垂直方向上的大尺度图像信息。高科[7] 将其应用于网络视频加密系统的设计。

为进一步减小对 DCT 系数统计特性的影响，以及避免泄露大尺度上的图像信息，在 Zeng[6] 的基础上，将 DCT 系数按相同位置分组，在组内进行置乱，通常相同位置系数的统计特性更加接近，对压缩码率的影响会更小；置乱的范围从 slice 扩大到整个帧，这样同时增大了非法攻击者破译的难度。仍以 I 帧的置乱为例，图 4 为其内嵌到视频编码过程中的示意图，置乱的恢复为其反过程。DCT 相同位置系数分组如图 5 所示，以 $q_k^{(l)}(k=0,1,\cdots,63;l=0,1,\cdots,n-1)$ 表示第 $l+1$ 个 DCT 块的第 $k+1$ 个量化后的系数，n 为帧内 DCT 块的总数，对于亮度信号 Y，$n=4m$，m 为宏块数，对于色度信号 UV，$n=2m$。$q_k^{(l)}$ 按在块内的相同位置分成 64 组进行置乱，每组 n 个。

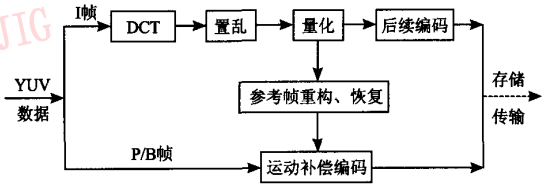

图 4 I 帧频域置乱方式

Fig. 4 I-frame scrambling mode in frequency domain

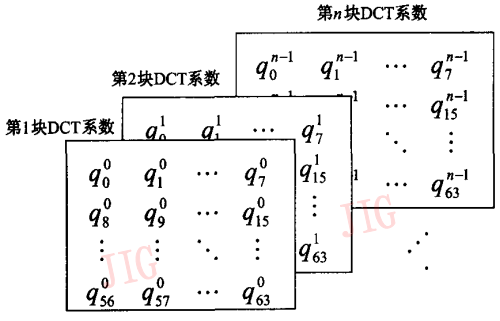

图 5 DCT 系数置乱图解

Fig. 5 DCT coefficient scrambling scheme

与 I 帧相比，P/B 帧的频域置乱稍微复杂一些，除帧内编码块的置乱同 I 帧外，运动补偿编码块的运动预测误差的 DCT 系数也可以采用相同的方法进行置乱。

4 运动向量置乱

实验结果表明，视频数据在空域或频域被置乱后，物体和场景的运动信息仍会是可见的。因此，运动信息常被作为视频分析的主要关键特性，对于高安全级别要求的应用，对运动信息的置乱是必须的。Zeng[6] 对运动向量的置乱，是将运动向量（包括水平 X 和垂直 Y 两个方向的分量）的符号标志位（Sign）有选择性地进行反转，如 1 变为 0 或 0 变为 1，以此来达到置乱运动向量的目的。

然而仅有选择性地改变运动向量的符号标志位，对运动信息的安全保密还是不够的，因为在运动物体或场景区域内，块的运动向量通常具有相同的符号标志位，这一点很有可能会被非法攻击者所利用，所以有必要进一步对运动向量的整体进行置乱。置乱 P/B 帧运动向量的 X、Y 两个分量可以分开，也可以作为整体，但必须以宏块为单位，因为在压缩码流中，对运动向量的编码模式标识是以宏块为单位的。以 MPEG4 为例，对于采用前向预测的 P 帧，在基于 8×8 块的运动补偿方式处于激活状态时，运动向量主要有 INTER16（1 个运动向量）、INTER8（4 个运动向量）两种模式，置乱时应分开来处理。对于采用双向预测的 B 帧，运动向量主要有直接（direct mode）、前向、后向和双向 4 种模式，前 3 种只有 1 个运动向量，双向模式有 2 个运动向量，同 P 帧，它们也各自分开处理，将前 3 种模式分开的目的是可以减小对运动向量统计特性的影响，因为运动矢量采用预测编码，而前向和后向运动矢量的正负值往往是相反的。

5 实验与分析

5.1 视频置乱实验

实验采用伪随机方法，以密钥为种子来生成乱序表，根据乱序表对要置乱的对象——空域 8×8 块单元、频域 DCT 系数以及运动向量来进行置乱。具体步骤是：用密钥作为种子，生成 $[0,n-1]$ 范围内的伪随机整数序列，n 为要置乱对象的数目，当出现

相同的整数值时只保留第 1 次出现的整数,也就是说去掉第 2 次以后出现的整数,直到序列覆盖[0, $n-1$]范围内的所有整数,这样就得到需要的乱序表。置乱恢复时仅需密钥,依照相同方法得到乱序表,再根据乱序表恢复被置乱的对象。

实验使用的视频编解码标准是基于国际标准化组织 ISO 提供的 MPEG4 源码 Momusys[8]。使用的测试视频序列 Vectra[9] 的格式为 CIF(352×288),4:2:0 采样,共 142 帧,记录了一辆外观为白色的汽车的运动过程,其 YUV 文件大小为21 593 088字节。实验的帧编码模式为每 9 帧有 1 个 I 帧、2 个 P 帧、6 个 B 帧,即 I-B-B-P-B-B-P-B-B-I 模式,正常编码后的码流文件大小为491 852字节,压缩比为 43.90,以色度信号 U 为例(后面同此),142 帧图像的平均 PSNR 为 39.08dB。实验对前面所提出的几种空域、频域和运动向量置乱方法,以及它们的结合作了实现,实验结果如表 1、图 6 ~ 图 14 所示。其中,表 1 列出了编码后的码流文件大小、码率增加幅度和 PSNR。

表 1　实验结果

Tab. 1　Experimental result

置乱策略	编码文件大小(字节)	码率增加幅度(%)	PSNR(dB)
外置所有帧空域置乱	1 580 146	221.26	30.50
内置 I 帧空域置乱	570 960	16.08	36.99
I 帧频域置乱	575 935	17.10	39.08
P 帧运动向量置乱	499 081	1.47	39.08
B 帧运动向量置乱	504 885	2.65	39.08
内置 I 帧空域置乱和 P/B 帧运动向量置乱	591 222	20.20	36.99
I 帧频域置乱和 P/B 帧运动向量置乱	596 197	21.21	39.08

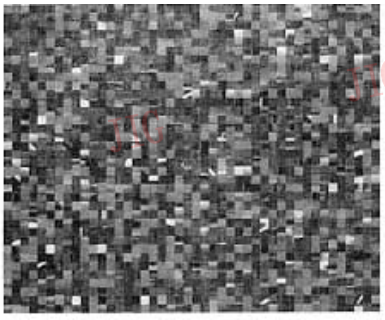

图 6　I 帧空域置乱

Fig. 6　I-frame scrambling in spacial domain

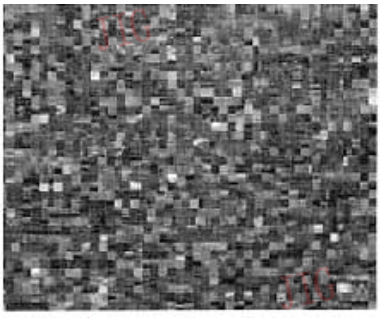

图 7　I 帧频域置乱

Fig. 7　I-frame scrambling in frequency domain

图 8　P 帧运动向量置乱

Fig. 8　P-frame motion vector scrambling

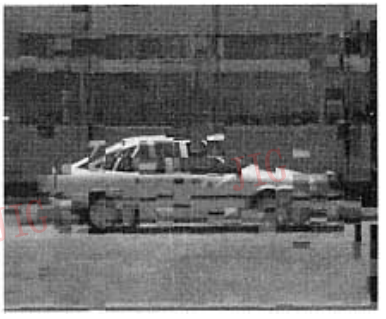

图 9　B 帧运动向量置乱

Fig. 9　B-frame motion vector scrambling

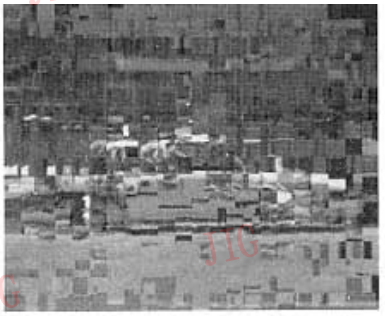

图 10 P/B 帧运动向量置乱后的 B 帧

Fig. 10 P/B frame motion vector scrambling

图 11 正常压缩解压缩图像

Fig. 11 Normal decoded image

图 12 外置空域置乱恢复图像

Fig. 12 Unscrambled image of additive mode

图 13 内置空域置乱恢复图像

Fig. 13 Unscrambled image of embedded mode

图 14 频域置乱的恢复图像

Fig. 14 Unscrambled image of frequency mode

从表 1 可以看出,外置空域置乱虽然简单,但压缩比却降低为不足原有压缩比的 1/3,且对图像质量造成较大的影响,PSNR 降为 30.50dB,这在图 12 中也可以看出。空域置乱内置后,情况得到比较大的改善,I 帧置乱码率增加幅度为 16.08%,此时对图像质量仍有影响,PSNR 为 36.99dB,但人眼几乎觉察不出置乱恢复图像(图 13)与正常编码解图像(图 11)的差别。表 1 表明,频域和运动向量置乱并不影响视频图像的质量,因为 PSNR 没有发生变化。I 帧频域置乱码率的增加幅度与内置 I 帧空域置乱接近,为 17.10%,但它没有给视频图像带来额外的质量下降,更重要的是其块内纹理是不可见的(见图 7),而空域置乱时视频图像的块内纹理是可见的(见图 6),这验证了频域置乱的优点。另外,运动向量置乱仅使码率小幅上升,P 帧为 1.47%,B 帧为 2.65%,在 I 帧频域与 P/B 运动向量均置乱的情况下,码率增加幅度也仅为 21.21%,这些结果好于 Zeng[6] 置乱方法的实验结果。

此外,前面提到的仅在空域或频域置乱时,物体或场景的运动仍是可见的,这一点在实验中得到了证实。

表 2 码流增加率实验结果比较

Tab. 2 Comparison of bit overhead

单位:%

	Zeng 方法[6]	本方法
I 帧频域置乱	19.8	17.1
频域 + 运动矢量置乱	23.6	21.2

5.2 视频置乱的安全性分析

对于 352 × 288 大小的帧,忽略帧内具有完全相同的块,非法攻击者要恢复 I 帧空域置乱需要尝试 $n!$ ($n = 44 \times 36$)种可能性,每种可能性对应的计算量为解码 I 帧的计算量;而频域置乱的可能性为 $\prod_{i=1}^{64} n_i! \times \prod_{j=1}^{64} l_j!$,$n_i$、$l_j$ 分别为非零的 Y 信号与 UV 信号 DCT 系数的个数,$n_i \leqslant 44 \times 36$,$l_j \leqslant 2 \times 22 \times 18$,忽略 DCT 系数为零及非零 DCT 系数相同的情况。对于 P 帧运动矢量置乱,假设 INTER16 和 INTER8 两

种模式具有非零运动向量的宏块个数分别为 r、s 个($r+s\leqslant 22\times 18$),则可能性为 $r!\times s!$;B 帧运动矢量置乱的可能性与 P 帧类似。由此可以看出非法攻击需要付出的计算代价是非常巨大的,这说明用置乱方法加密视频信息具有比较高的安全性。这种安全性除了依赖乱序算法本身的鲁棒性以外,还有赖于置乱后视频图像的视觉特性,如果置乱的视觉特性仍可能被利用,则非法攻击的计算代价就会降低。因此,综上所述,可以得出这样的结论:当 I/P/B 帧均在频域置乱,且 P/B 帧运动向量置乱时,置乱策略具有最好的安全性。

6　结　论

为了加强安全性,置乱应以尽量模糊可能被非法攻击者利用的视频信息特性为原则之一;置乱在一定程度上改变了视频数据的统计特性,因此,其策略方法则应以尽量小地影响统计特性为另一原则。基于这些原则,对已有的频域与运动矢量置乱算法进行改进,并得到实验验证。实验结果还说明,基于运动矢量与频域相结合的视频置乱策略与方法具有更强的安全性。对安全性的分析表明,基于置乱的视频信息隐藏技术能够为视频数据提供较高的安全性。此外,它可以与视频编解码过程很好地进行结合,避免传统加密方法的大量额外开销。在视频信息的发布与交流(如视频会议)越来越普及的今天,该方法具有很好的应用前景。

参考文献(References)

1　Agi I, Gong L. An empirical study of secure MPEG video transmissions[A]. In: The Internet Society Symposium on Network and Distirbuted System Security [C], San Diego, CA, 1996:137 ~ 144.

2　Yang Xiao-wen, Sheng Zhi-fan, Zheng Zhi-hang. Design of conditional access system for compressed video data [J]. TV Engineering, 1996,(1):1 ~ 6.[杨晓文, 盛志凡, 郑志航. 视频压缩数据有条件接收系统设计[J]. 电视技术, 1996,(1):1 ~ 6.]

3　Liu Bao-feng, Zhang Wen-jun. Research on security of video scrambling transmission [J]. Communications Technology, 2003,(7):88 ~ 90.[刘宝峰, 张文军. 视频流传输安全性的研究[J]. 通信技术, 2003,(7):88 ~ 90.]

4　Hartung Frunk, Girod Bernd. Watermarking of uncompressed and compressed video[J]. Signal Processing, 1998, **66**(3): 283 ~ 301.

5　Tang L. Methods for encrypting and decrypting MPEG video data efficiently[A]. In: Proceedings of the 4th ACM International Multimedia Conference[C], Boston, MA, 1996:219 ~ 229.

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

7　Gao Ke, Xu Wen-bo, Zheng Zhe-xing. The design of network video encryption system based on frequency domain selective scrambling algorithm [J]. Computer Engineering and Applications, 2003,(32):179 ~ 181.[高科, 须文波, 郑哲星. 基于频度域选择置乱算法的网络视频加密系统的设计[J]. 计算机工程与应用, 2003,(32):179 ~ 181.]

8　ISO. Source code of MPEG4 Encoder and Decoder of Publicly Available Standards [EB/OL]. http://www.iso.org/iso/en/ittf/PubliclyAvailableStandards/14496-5_Compressed_directories/Visual/Natural.zip, 2004-03-20.

9　Lagendijk R L. Test Video Sequence of Vectra[EB/OL]. http://www-it.et.tudelft.nl/~inald/vcdemo/otherdownloads/vectra.zip, 2004-06-06.