基于感兴趣区域的 H. 264 视频加密算法

于俊清 刘 青 何云峰
(华中科技大学计算机科学与技术学院 武汉 430074)

摘 要 为了解决现有的基于 H. 264 的加密算法无法满足安全性和加密效率之间较好折衷的问题,提出两种基于感兴趣区域的 H. 264 加密算法,将感兴趣区域的提取和基于熵编码的 H. 264 加密算法相结合,只对提取出来的感兴趣区域进行加密. 设计了一种基于人脸检测的加密算法,并通过修改模式选择算法,去掉了使得非人脸区域参考人脸区域的帧间宏块预测类型,解决了由帧间预测引起的人脸加密区域变形的问题. 将运动对象的检测与提取和 H. 264 的编码过程结合,设计了一种基于运动对象检测的 H. 264 加密算法. 实验结果表明,这两种加密算法能够在不降低算法安全性的情况下获得安全性和加密效率之间的较好折衷,可以满足实时应用的需求.

关键词 视频加密; H. 264 编码; 熵编码; 感兴趣区域
中图法分类号 TP391 **DOI 号**: 10. 3724/ SP. J. 1016. 2010. 00945

H 264 Video Encryption Algorithm Based on Region of Interest

YU Jun-Qing LIU Qing HE Yun-Feng
(College of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074)

Abstract In order to resolve the problem that the existed H. 264-based video encryption algorithms cannot obtain the tradeoff between security and encryption efficiency, two encryption algorithms are proposed based on the regions of interest in H. 264 video. These algorithms combine extraction of region of interest with entropy encoding algorithm and only encrypt the extracted region of interest macroblock. A face-detection encryption algorithm is put forward, in which the mode selection algorithm is modified so that the inter macroblock prediction type of face region's referring to non-face region is eliminated. Another algorithm extracts motive object and encrypts macroblocks which belong to motive object during encoding procedure. Experimental results demonstrate that the two proposed algorithms can obtain a good tradeoff between security and efficiency and can meet the need of real-time applications.

Keywords video encryption; H. 264 encoding; entropy encoding; region of interest

1 引 言

随着计算机技术和网络技术的迅猛发展,网络多媒体服务成为网络服务中新的应用领域. 但是由于网络的开放性和匿名性,使得网络上传输的多媒体内容很容易遭受攻击. 而传统的视频加密算法不能满足不同网络多媒体服务的要求,如何设计满足不同应用需求的视频加密算法,以保护视频内容的安全性十分重要. H. 264 是当前最流行的视频编码标准之一,与传统的视频编码标准相比, H. 264 能在同等视频质量的情况下获得更低的码率[1],但以

收稿日期:2009-07-10;最终修改稿收到日期:2009-10-09. 本课题得到国家自然科学基金(60703049)、武汉市青年科技晨光计划(200850731353)和华中科技大学自由创新基金(M2009019)资助.于俊清,男,1975 年生,博士,副教授,主要研究方向为数字媒体处理与检索、多核处理器编程环境研究. E-mail: yjqing @hust. edu. cn. 刘 青,女,1985 年生,硕士,主要研究方向为多媒体安全、视频编码. 何云峰,男,1977 年生,博士研究生,主要研究领域为数据视频索引与检索.

大幅增加的计算复杂度为代价.由于视频本身的一些特性(如数据量大、实时性要求高、存储格式多样等),用传统的算法对视频流进行加密效率较低,而效率相对较高的选择加密算法安全性不够,如何设计加密算法以获得安全性和加密效率之间的较好折衷是基于 H.264 加密算法研究中十分重要的研究内容.本文重点讨论如何结合基于 H.264 熵编码的加密算法,设计基于感兴趣区域(Region of Interest,ROI)的视频加密算法,从而满足保护视频隐私的需要.

2 相关工作

许多研究者和研究机构都提出了一些基于 H.264 的视频加密算法.韩国成均馆大学的 Ahn 等人提出加密帧内预测模式的方法[2],帧内 4×4 块有 9 种预测模式,用 3 比特流密钥对它进行加密;帧内 16×16 块有 4 种预测模式,而帧内 16×16 块的预测模式和编码块模式(Coded Block Pattern,CBP)一起编码,为了保证格式兼容性,实际上只能用 1 比特流密钥对它进行加密.虽然这种方法能够达到一定的加密效果,但是由于密钥空间不大,因此安全性不高.韩国成均馆大学的 Kwon 等人提出了一种使用运动向量和片再定位的视频加密算法[3],该算法置乱不同宏块的运动向量差,置乱表是事先写好的.由于 *I* 帧没有运动向量差(Motion Vector Difference,MVD),因此随着编码过程的进行,置乱 MVD 的效果更加明显,但是对于帧号靠前的图像,加密效果并不明显,这是由运动向量加密算法的本质决定的.由于该算法没有加密视频的纹理信息,算法不能抵抗已知明文的攻击.北京工业大学蔡勉等人提出一种基于熵编码的加密算法[4],它考虑到熵编码的上下文相关性,如果加密基于上下文的自适应可变长编码(Context-Adaptive Variable-Length Coding,CAVLC)过程中的某个码字索引可能使加密后的索引在码表中不存在,这样就会导致解码失败.文中通过分析码表的分布规律,提出了一种保持格式兼容的加密算法,将加密算法嵌入到非零系数的数目(TotalCoeffs),4×4 块最后一个非零系数前零的个数(TotalZeros)以及色度直流 2×2 系数的 TotalZeros 和每个非零系数前连续零的个数(Runbefore)中,但是这种加密方法无法完全满足 TotalCoeffs、拖尾 1 的个数(TrailingOnes)、除拖尾系数之外的非零系数的幅值(Levels)、TotalZeros、Runbefore 这些变量之间的约束关系,因此没有格式兼容性.合肥工业大学的齐美彬等人提出基于检测与跟踪的人脸图像保护加密算法[5],该算法结合自适应的皮肤颜色模型和传统贝叶斯分类器的方法来检测人脸,并用卡尔曼滤波器对检测到的人脸区域进行跟踪,人脸检测、跟踪过程和编码的预测、变换、量化过程同步,并在熵编码过程之前对检测到的人脸部分进行加密,该方法计算复杂度较高,不能满足实时性的要求.

基于 H.264 的视频加密算法不是设计加密算法本身,而是基于已有的加密算法在 H.264 的编码过程中选择合适的加密位置和加密强度来设计一种适用于视频的加密算法.H.264 视频加密的关键是如何选择加密位置,使得加密算法具有格式兼容性,并且能获得安全性和复杂度之间较好的折衷.由于对感兴趣区域进行加密能够保护视频的重要信息,并且减小加密算法对编码的计算负载,而人脸和运动对象是感兴趣区域的典型代表,因此本文将人脸检测和运动对象检测与基于熵编码的加密算法相结合,实现了基于感兴趣区域的 H.264 加密算法.对输入视频进行人脸检测,如果检测到人脸,则使用基于人脸检测的加密算法进行加密,当前帧加密完成后对下一帧进行人脸检测,更新人脸位置,进行下一帧的编码、加密过程;如果未检测到人脸,则进行正常编码,在编码的过程中利用生成的运动向量检测运动对象,如果检测到运动对象,则使用基于运动对象的加密算法;否则,根据配置文件选择使用基于 CAVLC 的加密算法或基于上下文的自适应二进制算术编码(Context-Adaptive Binary Arithmetic Coding,CABAC)的加密算法进行加密,编码完成后将结果输出成码流.

3 基于熵编码的 H 264 加密算法

传统的基于 H.264 的加密算法都是选择对 DCT 系数、运动向量差的符号、帧内预测模式等进行加密.这些算法要么密钥空间太小,要么无法满足格式兼容性,加密后的码流标准解码器无法解码,并且这些算法中对基于 H.264 熵编码的加密研究较少.如果选择加密在熵编码之前进行,压缩性能会降低;如果选择加密在熵编码之后进行,将会失去格式依赖性.为了更好地达到压缩的目的,用 H.264 编

码后的数据,无论是残差还是其它信息最后都要进行熵编码,因此本文选择在熵编码过程中嵌入加密算法,结合 H. 264 的两种熵编码方法 CABAC 和 CAVLC,改进了两种已有的基于熵编码的加密算法.

3.1　基于 CABAC 的加密算法

包先雨等人提出的基于 CABAC 的数字视频加密算法[6],针对旁路编码和正规编码两种模式,将基于 CABAC 的加密算法分为旁路编码模式加密和正规编码模式加密. 它的思想是分别对用 CABAC 编码的残差系数、帧内预测模式以及运动向量差进行加密,因为这些数据包含了视频的细节信息以及运动信息,对它们进行加密能获得较好的安全性. 加密算法嵌入在 CABAC 编码过程的第一步,即语法元素的二值化过程.

但是,这种算法对 k 阶指数哥伦布(k-th order Exp-Golomb, EGk)码字用等长的 $2l(x)+k+1$ 大小的流密钥进行加密的过程不能保持格式兼容性. 因为对于残差系数来说 $k=0$,这时前缀和后缀的长度相同. 解码器通过读入连续的 1 的个数来判断后缀的长度,进而算出残差系数的值. 如果加密前缀,很可能导致前缀长度判断错误,进而后缀长度也判断错误. 这样读该残差系数的时候就会多读或少读比特流,导致把本不是该残差系数的比特流当成该残差系数的比特流来对待或者下一个语法元素解析的时候指针错误,最后导致整个码流的解析错误. 对于运动向量差来说,由于 $k=3$,后缀的长度比前缀的长度大 3,解码的时候通过将读入的连续 1 的个数加 3 得到后缀的长度 n. 再读入 n 个比特,对该比特进行解析得到该语法元素对应的值. 如果前缀的长度解析失败,必然会造成整个码流的错位.

本文与原算法不同之处在于只对 EGk 后缀加密,流密钥使用 FLEX 流密钥[7]. 前缀和后缀的长度相差 k,中间由一个 0 隔开,数值越大编码后码流的长度越长. 前缀的长度 $l(x)=\left\lfloor \log_2\left(\frac{x}{2^k}+1\right)\right\rfloor$,其中,$x$ 为要编码的残差系数或运动矢量差的值,前缀每一位值都为 1. 后缀的长度为 $l(x)+k$,表示的值为 $x+2^k(1-2^{l(x)})$. 由于前缀不变,这种算法使得加密后的残差系数幅值的改变在一定的范围之内.

对残差系数来说,无论是舍位一元(Truncated Unary, TU)前缀加密还是 EGk 后缀加密,加密算法本身都能保证加密后的残差和加密前的残差相比保持在一定范围之内,因此可以对所有残差进行加密,而不会产生对码率较大的影响. 为了获得更好的安全性,改进算法对所有残差系数的符号进行加密,而不仅仅加密残差系数绝对值为 1 的系数符号位.

3.2　基于 CAVLC 的加密算法

法国电信北京研发中心提出的一种加密算法[8]仅仅改变运动向量差的符号,这种算法不能提供足够的安全性,因为可以根据运动向量差的绝对值恢复出低分辨率的图像[9]. 本文在该算法的基础上提出对运动向量差采用指数哥伦布加密算法进行加密,并且除加密 AC 系数的符号外,还对 AC 系数的 RunBefore 表格进行置乱.

由于编码器只对非零系数进行编码,并且每个非零系数都对应一个 RunBefore 值,该值表示该非零系数前连续 0 的个数,因此对 RunBefore 表格进行置乱可以改变亮度 4 ×4 变换块系数中 0 的分布情况,再结合对系数符号、DC 系数幅值、帧内预测模式以及运动向量差的加密,可以达到较好的加密效果. 这里只对帧内宏块残差进行加密,因为编码后的码流中帧内宏块的数据占据了码流的极大部分,如果帧内宏块加密强度足够的话,后面以帧内宏块为参考的帧间宏块必定失真很大,这时只需要对帧间宏块的运动向量差进行加密就能达到足够的安全级别. 并且对 DC 系数幅值的加密算法将会较大程度地影响压缩效率,如果加密所有残差的 DC 系数,将会使得加密算法对码率的影响较大,不能满足实时的应用需求.

文中提到用流密钥对幅值进行加密,但是并未提到流密钥的长度. 这就涉及到 DC 系数和 AC 系数可能达到的最大值的问题. 根据计算可以得出量化后 DC 系数的最大可能值为 1632,(1,1) 上的最大可能值为 1469. 因此,为了保证加密后的幅值在范围之内,最多可用 10 比特的流密钥对幅值进行加密. 流密钥的长度将极大地影响最终的压缩效率和图像的可理解性,流密钥长度越长,加密效果越好,但是压缩效率越差. 因此为了在压缩率和安全性之间进行折衷,通过实验数据得出,流密钥的长度为 8 是比较合适的. 加密时要注意加密后的幅值不能为 0,否则也会出现无法解码的情况.

另外,对 DC 系数的加密过程只能适用于帧内 4 ×4 块的 DC 系数. 对于帧内 16 ×16 宏块和色度直流系数在进行 DCT 变换后还要进行哈达码变换,不能对哈达码变换后的系数进行加密. 因为改变的 DC 系数可能影响后面 AC 系数码表的选择,如果码表选择错误,解析时长度就会发生错误,最后导致解码失败.

4 基于感兴趣区域的加密算法

感兴趣区域是视频中重要性较高的部分,通常包含更多的敏感信息.对视频中的感兴趣区域进行加密能够在减少加密算法对码率、编码效率的影响的情况下保护视频的重要信息.本文提出了两种基于感兴趣区域的加密算法,它们分别将人脸检测和运动对象检测与基于熵编码的 H.264 加密算法相结合.

4.1 基于人脸检测的加密算法

人脸是视频中出现频率较高的对象,也是观众注意力集中的地方.文献[5]中提出了一种基于人脸检测和跟踪的基于 H.264 的视频加密算法.本文将提出一种不同的基于人脸检测的加密算法.该算法与文献[5]的不同之处在于通过修改模式选择算法[10]避免了对非认证用户加密效果扩散导致加密区域变形的问题.

为了加密视频中的 ROI,文献[11]提出结合灵活宏块排序(Flexible Macroblock Ordering,FMO)类型 6,将整帧分成两个片组,分别代表前景和背景.前景片组对应需要加密的感兴趣区域.在 H.264 编码过程进行加密的流程如图 1 所示.很明显,加密参数不参与预测,因此不会影响以当前帧为参考帧的其它帧的编码,也就是说编码端的错误不会扩散.

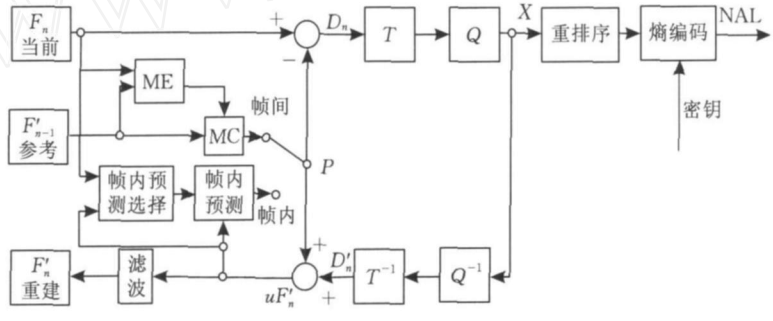

图 1 编码端加密流程图

解码端解密框图如图 2 所示.对拥有密钥的认证用户,熵解码过程恢复出的数据和在编码端熵编码前的数据相同,并且解密过程在预测循环之前,因此重构出来的图像也和编码端的重构图像相同,加解密过程完全可逆.由于加密算法本身的格式兼容性,对非认证用户仍然可以正常解码加密视频.但是由于熵解码的数据未经过解密,参与预测循环的数据也是未经过解密的数据,会导致加密数据发生漂移.这种漂移是由帧间预测产生的.由于该算法采用 FMO,人脸和非人脸区域属于不同的片组,而 H.264 标准规定帧内预测不能跨片进行,帧内预测错误不会发生漂移.但是帧间预测没有片边界,运动估计的时候不会判断参考帧的当前搜索像素属于哪个片,这样通过运动估计得到的运动向量可能使得属于非人脸的背景片组的像素以属于人脸的前景片组为参考.而人脸部分是加密的,背景部分就会被加密的前景部分影响,重构出来的图像会发生扭曲的现象,不再是一个矩形区域.

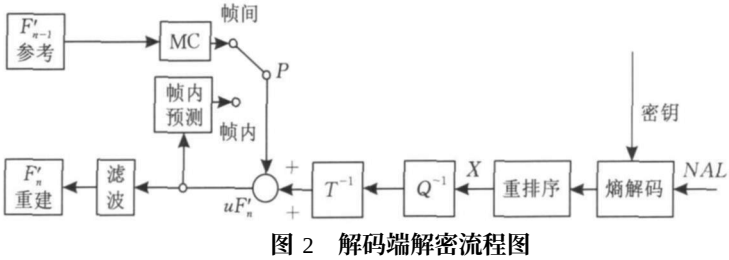

图 2 解码端解密流程图

文献中提出了这个问题,但是文中说强制某些帧间预测的宏块变为帧内预测.实验结果显示,这种方式将会不必要地增大码流.本文提出的修改模式选择的过程为对 16×16、16×8、8×16、8×8、P-SKIP模式进行运动估计,根据每一种模式估计出的运动向量计算参考块左上、右上、左下、右下 4 个角的坐标,并计算出这 4 个角所属的宏块号.判断这 4 个宏块中是否有任何一个宏块属于前景片,而当前编码宏块属于背景片,如果有则将当前模式置为不可用.后面选择最佳模式时无论是否启用 RDO 都不用判断这种模式.

本文使用 FMO 类型 2,用 AdaBoost 算法[12]检

测人脸,对检测的人脸区域进行加密,加密算法使用第 3. 2 节中介绍的基于 CAVLC 的加密算法.

4.2　基于运动对象检测的加密算法

运动对象也是视频中的感兴趣区域之一. 要对运动对象进行加密首先要提取运动对象. 现有的基于 H. 264 的运动对象提取都是基于压缩域进行的. 而基于 H. 264 的加密必须和编码过程结合起来. 本文提出的算法在编码过程中提取运动对象,并对运动对象进行加密,它利用了编码过程中产生的运动向量. 这种思想在前人的文章中并未提出. 由于运动对象提取本身是一个比较复杂的方向,本文简化了刘志等人在文献[13]中提出的基于 H. 264 压缩域的实时运动对象分割算法,并对提取的运动对象进行加密. 运动对象提取过程和编码过程同步.

由于 H. 264 运动向量搜索算法的最小单位是 4 ×4 块,运动向量的精度达到四分之一精度. 根据模式选择的结果,对以 16 ×16、16 ×8、8 ×16、8 ×8 等为单位进行运动估计的宏块,将运动估计的结果(该块的运动向量)赋给该块包含的所有 4 ×4 子块. 由于 H. 264 采用了可变块大小的运动估计算法,在用运动矢量场进行运动对象分割的时候以 4 ×4 块为基本单位. 运动对象提取分为运动矢量场的归一化和积累、全局运动补偿以及运动对象提取三步. 其中运动矢量场的归一化和积累的目的是获得更加平滑的稠密运动向量场,全局运动补偿的目的是去掉由摄像机运动引起的全局运动,剩下由运动对象产生的局部运动,最后用全局运动估计得到的参数计算得到残差运动向量场.

残差越大说明其属于运动对象可能性越大,残差越小说明其属于背景的可能性越大. 对所有残差求一个平均值,大于这个平均值的 T%的被认为是运动对象,此为阈值 1. 由于最后加密的单位是 16 × 16 宏块,而这里得到的是每个 4 ×4 块被判断为运动对象还是非运动对象. 因此必须设置一个阈值,如果一个 16 ×16 宏块中被判断为前景的 4 ×4 宏块的个数大于这个阈值,那么这个 16 ×16 被判断为运动前景,此为阈值 2. 如果阈值设置过大,可能提取的运动对象不完全,如果阈值设置过小,则会造成将不属于运动对象的背景区域错划成运动对象.

由于 I 帧包含大量信息,并且 I 帧不包含运动向量信息,因此本算法对 I 帧的所有宏块加密. 仅仅加密 I 帧不能提供足够的安全性抵抗已知明文攻击. 未加密的 I 块和 B 帧、P 帧的运动向量可以用来恢复部分原始视频的视觉信息. 本算法对 P 帧运动对象所属的宏块进行加密,由于 P 帧以 I 帧作为参考, P 帧被影响的宏块不仅仅是加密的宏块,未加密的宏块由于帧间预测也会被影响.

运动对象加密算法流程图如图 3 所示,加密算法使用 3. 1 节中介绍的基于 CABAC 的加密算法. 算法首先判断当前编码帧是否为 I 帧,如果是则对所有 I 帧的宏块加密、熵编码、写码流;如果不是则通过当前的帧号判断是否已经记录了 3 帧的运动向量,如果是则用上文的方法提取运动对象,否则记录运动向量等信息,这些信息包括用于提取运动对象的运动向量差以及延迟写码流的所有宏块信息,每 3 帧写一次码流直到一个图像组(Group Of Pictures, GOP)编码结束. 提取运动对象的过程和编码过程可以并行,采用的加密算法对编码效率几乎没有影响,因此该算法基本不影响编码效率.

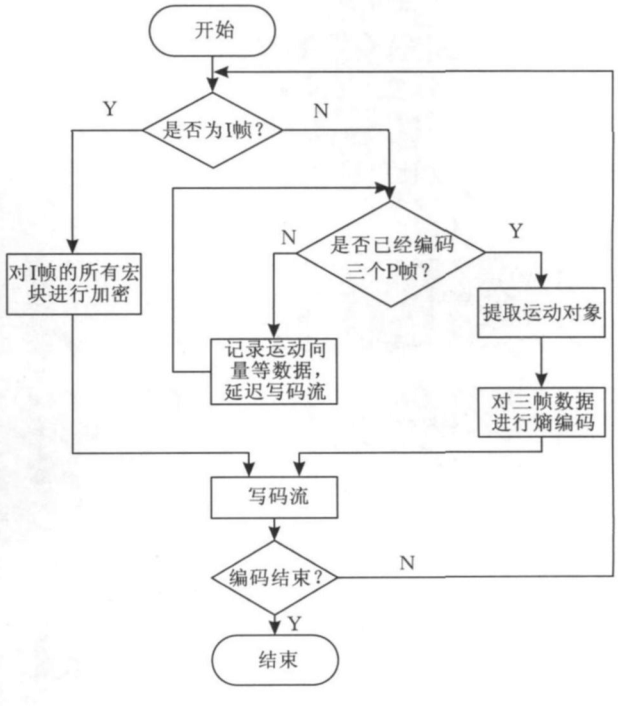

图 3　运动对象加密算法流程图

5　实验结果与分析

5.1　基于人脸检测的加密算法实验结果

使用基本档次进行编码,配置文件的参数为编码 10 帧,每 10 帧一个 GOP,包含 1 个 I 帧以及 9 个 P 帧,量化参数 QP 为 30,CAVLC 编码,用人脸检测的结果初始化 FMO 参数.

图 4 给出了采用本文的算法,加密前和加密后 PSNR 值的变化. 横轴表示 18 段视频序列,纵轴表示 PSNR 值. 可以看出加密后的 PSNR 值平均降低

了 25dB ,质量退化比较明显. 因此 ,该算法具有较好的感知安全性.

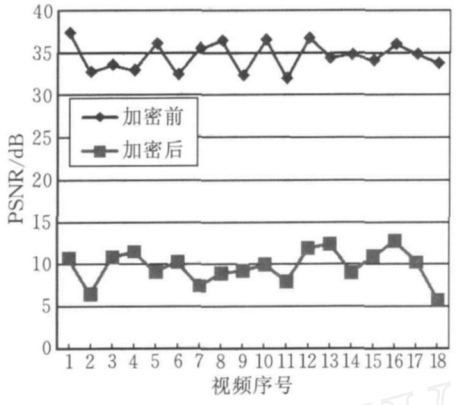

图 4 加密前后 PSNR 值比较图

图 5 给出了不修改模式选择算法加密 suzie 序列 10 帧的效果. (a) ～(j) 为加密后的第 1 帧到第 10 帧. 由图中可以看出 ,加密的人脸区域发生了扭曲的现象. 使用文献[5]的算法修改模式选择后的加密效果如图 6 所示 ,(a) ～(j) 为加密后的第 1 帧到第 10 帧. 可以看出直接将某些帧间宏块修改改为帧内宏块的做法使得加密的人脸矩形区域的边界清晰可见.

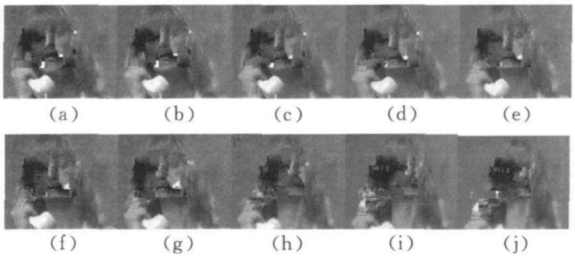

图 5 加密人脸区域扭曲效果图

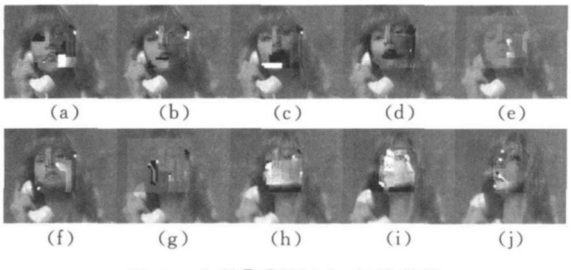

图 6 文献[5]算法加密效果图

使用本文提出的算法修改模式选择后的加密效果如图 7 所示 ,(a) ～(j) 为加密后的第 1 帧到第 10 帧. 可以看出本文提出的算法使得人脸矩形区域的边界有轻微的模糊现象. 这可能是因为人脸区域和非人脸区域边界像素的判断问题. 算法通过判断参考块四角像素所属的宏块是否和当前编码宏块属于同一个片来决定是否使用当前模式作为候选模式. 但可能参考宏块中四角的像素属于前景片 ,但由于像素处于前景片和背景片的边界 ,这个像素值(二分之一精度或四分之一精度) 是由属于前景片的像素和属于背景片的像素共同插值得到的 ,造成人脸区域边界的轻微模糊.

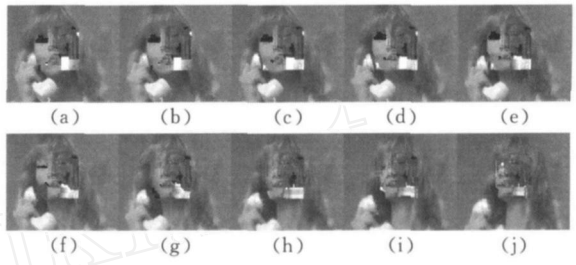

图 7 本文算法加密效果图

本文算法和不修改模式选择算法以及文献中算法对码率影响的比较如图 8 所示. 可以看到本文算法基本不改变原始的码率 ,但是文献[5]中的算法对码率的影响却在两倍以上. 这可能是因为直接将某些帧间宏块模式修改为帧内模式时并没有考虑率失真优化. 这种算法去掉了最佳模式是帧间模式 ,且它使得非前景片参考了前景片或相反的这种帧间预测模式. 但是除了这种模式之外 ,可能还有另一种帧间模式使得率失真达到次优 ,却不满足非前景片参考了前景片或相反 ,而它被文献中的方法剔除掉了 ,导致了码率大大地增加. 从时间上来看完全加密算法的时间比不加密的时间有略微的增长 ,这个增长是由加密过程产生的. 而本文的人脸加密算法和文献[5]中提到的人脸加密算法相比不加密的算法时间上还有下降 ,文献中的算法时间下降更加明显 ,这主要是由于模式选择过程中某些模式的判断提前结束造成的. 由于模式选择过程占整个 H. 264 编码时间的较大部分 ,因此文献算法在时间上有比较明显的降低. 但是文献[5]算法在时间上获得 8. 8 %左右的减少却是以增大 153 %的码率为代价的. 而本文的算法和完全加密的算法相比在时间上有所降低 ,码率的增加也不大. 实验在 CPU 为 P4 2. 4GHz ,内存 512MB 的微机上进行.

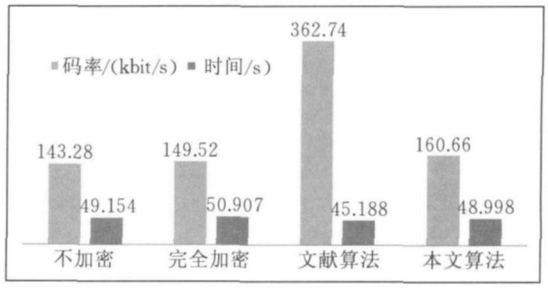

图 8 本文算法和其它算法性能比较图

加密算法采用的是基于 CAVLC 的加密算法,本算法的安全性和基于 CAVLC 的加密算法类似. 由于置乱的前景对象和背景对象有不同的特征,前景不能通过背景预测得到,因此不能使用错误隐藏攻击恢复视频内容.

5.2　基于运动对象检测的加密算法实验结果

本文对运动序列进行测试,使用主档次进行编码,配置文件的参数为编码 10 帧,每 10 帧一个 GOP,包含 1 个 *I* 帧以及 9 个 *P* 帧,量化参数 *QP* 为 30,CABAC 编码. 阈值 1 设置为 22. 对于阈值 2,图 9 给出了不同阈值的分割结果. 可以看出阈值为 1 时分割出的运动对象包括不是运动对象但有轻微运动的背景;阈值为 15 时分割出的运动对象只是运动对象的一部分,并不完全;阈值为 6 时分割出的运动对象比较合适.

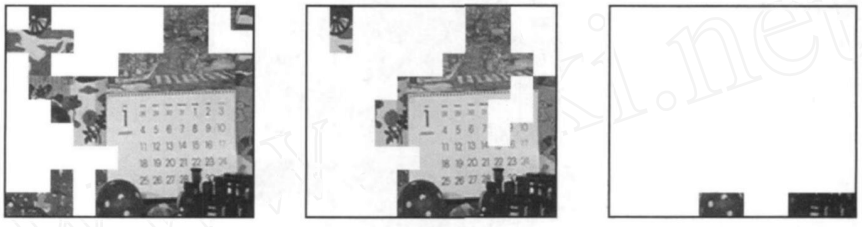

图 9　运动对象分割结果图

加密效果如图 10 所示,(a)和(b)分别为第 1 帧加密前后的对比图,(c)和(d)分别是第 10 帧加密前后的对比图. 实验结果表明,无论对于 *I* 帧还是 *P* 帧该算法都能达到较好的加密效果.

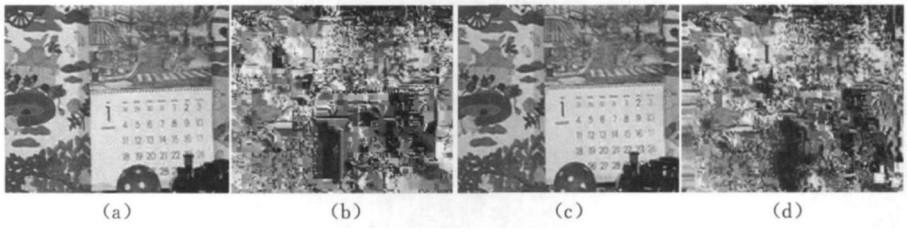

图 10　加密效果图示意图

图 11 给出了加密前和加密后 PSNR 值的比较. 横轴表示视频序列编号,纵轴表示 PSNR 的值. 可以看出加密后的 PSNR 值平均下降了 23dB. 因此,该算法具有较好的感知安全性.

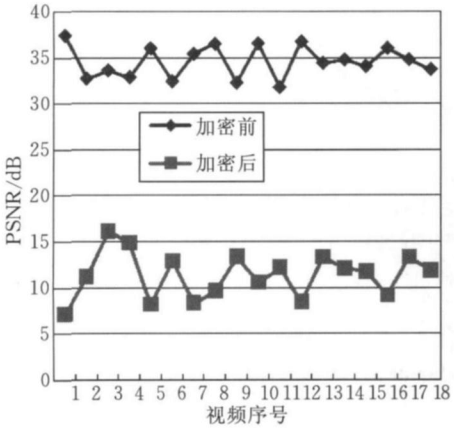

图 11　加密前后 PSNR 值比较图

运动对象的加密算法对码率的影响与其他算法的比较见图 12 所示. 可以看出,由于加密的数据量变少,该方法与对 *I* 帧、*P* 帧的所有宏块进行加密的算法相比码率有所下降,比仅加密 *I* 帧的算法码率有所增加,所增加的码率除了是由对运动对象加密导致的以外,还包括帧间预测时错误扩散产生的. 提取的运动对象的大小越大,增加的码率越大.

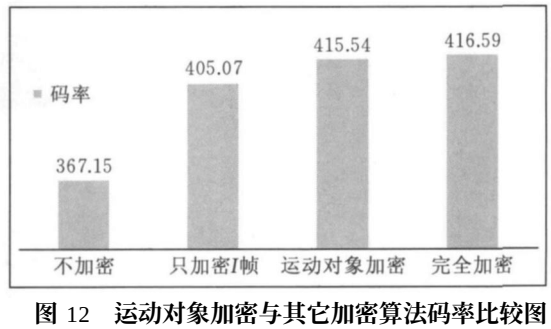

图 12　运动对象加密与其它加密算法码率比较图

基于运动对象的加密算法是建立在基于 CABAC 的加密算法的基础上的,它的安全性和基于 CABAC 加密算法的安全性类似. 由于被加密的视频对象被隐藏了,敌手除了破解算法本身外,还需要破解哪些宏块是被加密宏块. 因此,这种算法能够在不降低安全性的情况下,减小加密算法对码率的影响,满足实时性的要求.

6 结 论

由于视频中的感兴趣区域是视频需要重点保护的对象,为了减少视频加密算法对码率的影响,获得安全性和加密效率之间较好的折衷,本文提出一种基于感兴趣区域的 H.264 加密算法.该算法能够在不降低安全性的情况下节省一定的码率,达到安全性和加密效率之间的较好折衷,并可以满足实时应用的需求,为进一步研究基于 H.264 的加密算法提供了一种切实可行的研究思路.今后我们将继续深入研究基于感兴趣区域的加密算法中的几个关键问题,这主要包括:(1) 设计基于 H.264 扩展 SVC 的感兴趣区域的加密算法;(2) 结合其他视频处理方法分析基于熵编码的加密算法的安全性;(3) 在使用全局运动估计的运动向量差的基础上结合其他方法以更加精确地提取运动对象;(4) 设法利用不同视频序列的特性自动确定提取运动对象时所需的阈值;(5) 对大量视频进行统计,分析在 H.264 的编码过程中各参数对可理解程度的影响,根据分析结果针对不同重要性的参数使用不同安全级别的加密算法进行处理.

参 考 文 献

[1] Wiegand T, Sullivan G J, Bjontegaard G, Luthra A. Overview of the H.264/AVC video coding standard. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 560-576

[2] Ahn Jinhaeng, Shim Hiuk Jae, Jeon Byeungwoo et al. Digital video scrambling method using intra prediction mode// Proceedings of the 5th Pacific Rim Conference on Multimedia. Tokyo, 2005: 386-393

[3] Kwon Sang Gu, Woong Il Choi, Jeon Byeungwoo. Digital video scrambling using motion vector and slice relocation// Proceedings of the 2nd International Conference. Toronto, 2005: 207-214

[4] Cai Mian, Jia Jia, Yan Lei. An H.264 video encryption algorithm based on entropy coding// Proceedings of the 3rd International Conference on Intelligent Information Hiding and Multimedia Signal Processing. Kaohsiung, 2007: 41-44

[5] Qi Meibing, Chen Xiaorui, Jiang Jianguo et al. Face protection of H.264 video based on detecting and tracking// Proceedings of the 8th International Conference on Electronic Measurement and Instruments. Xi'an, 2007: 2172-2177

[6] Bao Xian-Yu, Jiang Jian-Guo, Yuan Wei et al. Study of CABAC-based digital video encryption in the H.264/AVC standard. Journal of Communications, 2007, 28(6): 24-27 (in Chinese)
(包先雨, 蒋建国, 袁炜等. H.264/AVC 标准中基于 CABAC 的数字视频加密研究. 通信学报, 2007, 28(6): 24-27)

[7] Biryukov Alex. The design of a stream cipher LEX// Proceedings of the 13th International Workshop. Montreal, 2007: 67-75

[8] Wang Yajun, Cai Mian, Tang Feng. Design of a new selective video encryption scheme based on H.264// Proceedings of the International Conference on Computational Intelligence and Security, 2007: 883-887

[9] Liu Zheng, Li Xue. Motion vector encryption in multimedia streaming// Proceedings of the 10th International Multimedia Modelling Conference. Brisbane, 2004: 64-71

[10] Lim K P, Sullivan G, Wiegand T. Text description of joint model reference encoding methods and decoding concealment methods, Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-X101// Proceedings of the 24th Meeting. Geneva, 2007: 1-57

[11] Frederic Dufaux, Touradj Ebrahimi. H.264/AVC video scrambling for privacy protection// Proceedings of the 15th IEEE International Conference on Image Processing. San Diego, 2008: 1688-1691

[12] Paul Viola. Rapid object detection using a boosted cascade of simple features// Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2001: 1511-1518

[13] Liu Zhi, Xu Jiang-Feng, Lu Yu et al. A real-time H.264 compressed domain based moving object segmentation algorithm. Journal of Image and Graphics, 2006, 11(11): 1614-1618 (in Chinese)
(刘志, 徐剑峰, 陆宇等. 基于 H.264 压缩域的实时运动对象分割算法. 中国图象图形学报, 2006, 11(11): 1614-1618)

YU Jun-Qing, born in 1975, Ph.D., associate professor. His research interests include digital media processing and retrieval, multi-core programming environment.

LIU Qing, born in 1985, M.S.. Her research interests include multimedia security and video encoding.

HE Yun-Feng, born in 1977, Ph.D. candidate. His research interests include digital video index and retrieval.

Background

With the fast development of internet and multimedia technology, the massive volume of digital video data is growing in an explosive way. All sorts of video sharing application such as video-on-demand, digital television, Web TV, video conference and online video chat continue to bring out. For the purpose of protecting video content from intercepting, digital video encryptions appear to be a rising research highlight in multimedia security domain. Based on the widely used H. 264 standard, it is of great importance to research on video encryption algorithm for protecting video content.

Based on several previous papers on H. 264-based video encryption algorithms, modified algorithms are proposed to resolve the shortcomings in the platform of JM8. 6 open source project. Depending on the analysis of the reason why two previous algorithms on entropy coding-based video encryption is not format-compliant, improved algorithms are put forward, one of which is perceptual encryption algorithm which degrades video quality to a certain degree. Another CAVLC-based algorithm which encrypt intra prediction mode, motion vector difference and intra macroblock residual differently is raised. Putting the idea of region of interest to the selection of encryption position, face detection-based and moving object-based detection encryption algorithms are implemented. Mode selection algorithm is modified and inter macroblock prediction type which make non-face region's reference macroblock locates in face region is forced not available, resolving the problem of distortion of encrypted face region resulted from inter prediction. Experimental results show that the proposed algorithms can get enough security level for real-time transmission application and the bitrate and coding efficient increment is not obvious. However, security of the proposed encryption algorithm need to be further studied; accuracy of moving object extraction needs to be improved. These are our research focus in the future work.

This work is financially supported by the National Natural Science Foundation of China under Grant No. 60703049 and the "Chen Guang" Foundation for Young Scientists of Wuhan under Grant No. 200850731353. In recent years, the group concentrated on the research on media processing and retrieval on multi-core architecture. Several papers have been published in international conference and journals. A testing platform prototype has been developed.