

针对 H. 264/AVC 的时域错误隐藏算法

周启亚¹⁾, 杨高波²⁾, 刘志成³⁾, 彭勇³⁾, 周庞荣¹⁾

¹⁾(湖南铁路科技职业技术学院信息技术系, 株洲 412000) ²⁾(湖南大学计算机与通信学院, 长沙 410082)

³⁾(湖南铁道职业技术学院信息工程系, 株洲 412001)

摘要: H. 264/AVC 采用了新的编码特征工具, 充分利用它们能有效地改进错误隐藏的性能。树型结构的可变分块大小运动估计, 使得与错误宏块直接相邻的分块数最多可达 16 个。本文提出一种基于平面拟合的时域错误隐藏算法。通过把与受损宏块直接相邻的每个分块的运动矢量定义为 3 维空间的一个点, 采用平面拟合的方法表征小范围内相邻运动矢量的变化趋势, 拟合得到受损宏块的运动矢量实现时域错误隐藏。对多种视频序列的仿真实验结果表明, 相对于典型的错误隐藏算法, 主观视觉质量和客观质量评价都有一定的提高。

关键词: 平面拟合; 错误隐藏; 时域; H. 264/AVC

中图分类号: TP301.6 文献标志码: A 文章编号: 1006-8961(2010)09-1338-07

Temporal error concealment algorithm for H. 264/AVC video stream

ZHOU Qiya¹⁾, YANG Gaobo²⁾, LIU Zhicheng³⁾, PENG Yong³⁾, ZHOU Pangrong¹⁾

¹⁾(Department of Information Technology, Hunan Vocational College of Railway Technology Zhuzhou 412000)

²⁾(School of Computer & Communication, Hunan University, Changsha 410082)

³⁾(Department of Information Engineering, Hunan Railway Professional Technology College, Zhuzhou 412001)

Abstract: A new unique encoding tools are adopted in H. 264/AVC. Fully utilization of these tools in error concealment will greatly improve the concealment performance. The tree-structure macroblock division mode with variable sizes is used in H. 264/AVC so that the amount of the divided blocks directly adjacent to the lost macro-block may be up to 16. The motion vector of each divided block surrounding the lost macro-block is projected to one point in 3D space. Then the plane fitting based MV recovery algorithm is proposed, in which one plane is constituted to present the MV changing tendency in small regions. Experimental results demonstrate the proposed algorithm is superior to other common used algorithm in both subjective and objective quality review.

Keywords: plane fitting; error concealment; temporal domain; H. 264/AVC

0 引言

随着网络与通信技术的发展, 多媒体应用越来越广泛。然而由于受带宽的限制, 视频、图像等多媒体数据需以压缩的形式来传输, 而压缩数据通常采用了上下文相关编码, 致使它们对传输通道的错误非常敏感, 1 bit 数据的错误可能会导致整帧的数据遭到损坏, 从而使得解码后的多媒体数据视觉质量难以容忍, 因此, 相关的错误处理技术是必不可少的。

通常有多种途径来避免错误的发生及由此导致的重建图像质量的下降, 其一就是改善传输通道的质量, 杜绝发生错误的根源, 或建立多传输通道, 以便在发生错误时进行重传, 显然该方法成本太高; 其二是在编码端实现的错误抑制技术, 这样就可以抵抗外界轻度的干扰造成码流灾难性的毁坏。在 H. 264 标准中采用的参数集结构 (SPS, PPS), ASO, FMO, 数据分割等, 均能较好地抑制错误的发生与扩散, 使得 H. 264 具有良好的网络亲和性。但该类技术的缺陷是增加了编码复杂度, 使编码的效率大大降低, 不利

收稿日期: 2009-01-18; 改回日期: 2009-04-29

第一作者简介: 周启亚 (1980—), 男, 讲师。2008 年于湖南大学获工学硕士学位。主要研究方向为视频编码、视频信号处理。

E-mail: zhoukia@126.com。

于实时应用,同时也会增加部分码流;其三就是在解码端实现的错误隐藏技术,该类方法在错误不可避免发生时,根据人眼视觉特性对发生的错误进行掩盖,显然不会增加编码器的负担,也无需花费额外传输通道成本,实现起来方便,可操作性强。

实现错误隐藏主要是利用像素空时域相关性来进行的。依据所利用的信息不同,可以分为两类,即空域错误隐藏和时域错误隐藏。文献[1]提出了一种简单的空域加权错误隐藏算法。对于受损宏块中每个像素点用与其在同一水平位置和同一垂直位置的4个边界点来进行加权插值,边界像素的加权系数与其到受损像素之间的距离成反比。考虑到图像本身的纹理特征,Xu等人利用Prewitt方向算子来求得受损块相邻像素中梯度最大的一个方向,然后按照这个方向进行插值^[2]。马宇峰等人^[3]提出一种基于最小像素跨度准则的空域错误隐藏方法,利用丢失块边界的像素,在无须进行边缘估计的情况下能够有效地恢复出丢失块中可能存在的边缘。时域错误隐藏利用的是视频帧间的时域相关性,较有效的方式是从参考帧中根据运动矢量选择用于运动补偿的宏块来替代受损宏块。在用于运动补偿的运动矢量也丢失的情况下,问题关键在于估计错误宏块的运动矢量。常用的受损宏块运动矢量估计方法有置零(TR)、取周围直接相邻宏块运动矢量的中值、均值或加权均值等,它们有各自的应用场合,适合于具有不同运动特点的视频序列^[4-5]。基于最小边界匹配失真的运动矢量估计算法(BMA)也得到了广泛应用,其基本思想是使所恢复的宏块与周围宏块的边界像素失真最小,即把满足像素最大空间平滑性作为运动矢量估计的依据^[6-7]。Yan等人在此基础上提出了一种空时域边界匹配准则用于恢复运动矢量,其综合利用了视频信号的空域连续性和时域连续性^[8]。

H.264/AVC作为新一代视频编码标准,因其较高的压缩效率和良好的网络亲和性,正逐渐地取代其他编码标准而得到广泛应用。目前针对H.264/AVC码流的错误隐藏也成为研究热点之一。Zheng等人先提出利用与受损宏块每个顶点直接相邻的3个 4×4 子块的运动矢量来构建一个低阶平面,根据该低阶平面来计算顶点像素的运动矢量值,再用平滑插值的方法来分别求得受损宏块内部其他每个像素的运动矢量值,从而依据每个像素的运动矢量来恢复受损宏块^[9]。后来他们又提出利用与丢失宏

块相邻的块运动矢量来构成一个多项式模型,以此来描述小范围运动矢量的变化趋势,进而预测得到受损宏块的运动矢量^[10]。H.264/AVC的编码特点决定了对其码流的错误隐藏算法具有新的特征。本文在分析了H.264/AVC错误隐藏特征的基础上,提出一种基于平面拟合的运动矢量恢复算法。

1 H.264/AVC 错误隐藏的新特征

1) 灵活的宏块组织形式(FMO)有利于错误隐藏。H.264/AVC中一帧可以分为多个片组(Slice),宏块可以通过不同的映射方式映射到不同的片组,从而形成不同式样的片组结构,如交错型、散乱型等。对于大小 $M \times N$ 像素的帧,宏块的个数为 $(M/16) \times (N/16)$ 。若把帧划分为 i 个片组,把映射记为 F ,则宏块到片组的映射过程可描述为

$$\text{MacroBlock}_k \xrightarrow{F} \text{Slice}(j)$$

其中, $k \in \{0, \dots, (M/16) \times (N/16) - 1\}$, $j \in \{0, \dots, i - 1\}$ 。若把帧划分成2个片组,即 $i = 2$,映射 F 取 $f = (k / (M/16)) \% i$ 就可得到图1(a)所示的交错型片组;当 F 取 $f = k \% i$ 则可得图1(b)所示的散乱型片组。宏块以不同的形式组织,对错误隐藏时可利用的信息量有不同的影响,例如当图1中标记为白色的宏块丢失时,图1(a)中可用于隐藏的相邻宏块只有上下2个(标记为黑线框,不计角点宏块),而图1(b)中有上下左右4个。显然以这种交替形式来组织宏块,在发生错误时有更多空间相邻宏块的像素信息用于恢复操作,更有利于错误隐藏。

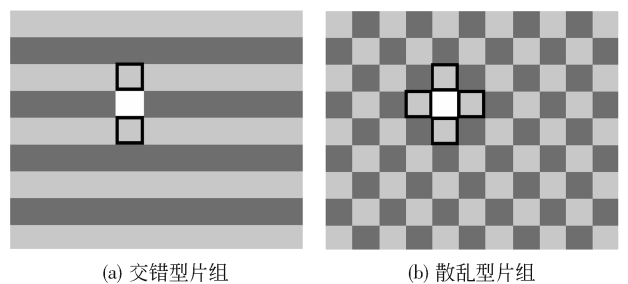


图1 片组结构

Fig. 1 The slice structure

2) 可变块大小的运动估计与补偿使得用于错误隐藏的运动信息更加丰富。H.264/AVC在宏块级支持 16×16 , 16×8 , 8×16 , 8×8 等4种划分模式,其中对于 8×8 的块模式,可以进一步划分为

$8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4$ 等 4 种子块模式。图 2 给出了一帧采用前向预测编码时的宏块划分示意图。每个划分对应一个运动矢量,宏块划分得越细,所拥有的运动矢量越多,一个宏块最多可拥有 16 个运动矢量。同时,块划分得越细,各块运动矢量的相关性越强。在进行损坏宏块运动矢量估计时,利用这种强相关性恢复出的运动矢量越准确。H. 264/AVC 中这种精致的树状结构宏块划分模式为准确的错误隐藏提供了条件,后面的算法中将用到这一运动矢量的相关性原理。

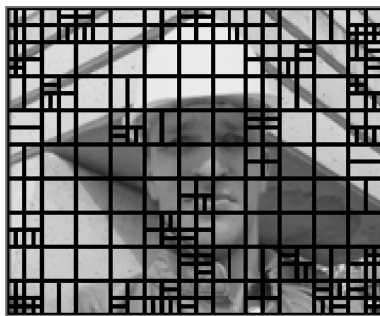


图 2 P 帧可变大小分块示意图

Fig. 2 The macroblock coded mode map with variable size of P frame

3) 多参考帧的使用是 H. 264/AVC 编码框架中的特色内容之一。为了获取更精确的匹配块, H. 264/AVC 从缓冲的多帧中进行搜索,找到与当前块匹配误差最小的块作为参考块,如图 3 所示。显然,可以把多参考帧匹配策略运用于错误隐藏,这样使得用时域相邻像素来恢复错误宏块时有更多的选择,进一步提高隐藏精度。

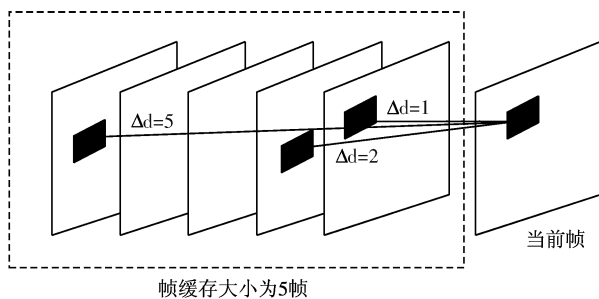


图 3 多参考帧运动搜索

Fig. 3 The motion estimation using multiple reference frame

2 基于平面拟合的运动矢量恢复算法

同一物体局部小区域通常有着相同的运动属

性,即保持运动一致性。而如前所述,H. 264/AVC 采用可变大小的分块尺寸来进行运动估计与补偿,最小分块为 4×4 像素。这样,分块越细,小范围内的块运动矢量也就具有相似性。利用 H. 264/AVC 这种块运动矢量的相关性,提出一种基于平面拟合的方法来估计受损宏块的运动矢量,本质上是通过一个平面来表征受损宏块周围直接相邻划分运动矢量的变化趋势,进而通过此拟合的平面来求运动矢量。

假设第 m 个宏块受损,与损坏宏块上下左右直接相邻的 4 个宏块均正确。以损坏宏块的中心点为坐标原点,根据相邻宏块的划分模式,各个划分的中心点在 2 维平面的坐标很容易求得。以图 2 中的第 13 个宏块发生错误为例(上起第 2 行,左起第 3 列),其周围直接相邻宏块的分割模式如图 4 所示。左、下、右 3 个宏块均采用 16×16 模式,那么其中心坐标相应为 $(-16, 0), (0, -16), (16, 0)$ 。而上边宏块采用 $P8 \times 8$ 模式,其中宏块中的第 1 个和第 3 个块进一步划分为 2 个 8×4 和 2 个 4×8 子块。那么与错误宏块直接相邻的上边划分有 3 个,中心坐标依次为 $(-6, 12), (-2, 12), (4, 12)$ 。又每个划分拥有一个运动矢量,若把运动矢量的值视为 Z 值,可以把每个划分对应的运动矢量描述为 3 维空间中的一点。上面的受损宏块其直接相邻的每个划分运动矢量在 3 维空间的分布如图 5 所示,实现时水平和垂直分量单独考虑,这些与受损宏块直接相邻的划分运动矢量集合记为 P 。在 3 维空间中对 P 的元素用一个平面来进行拟合,拟合后的平面反映

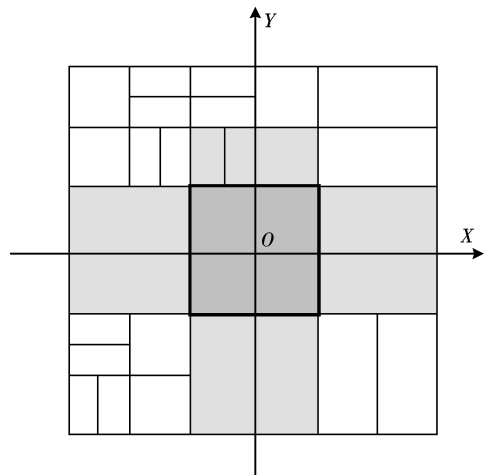


图 4 损坏宏块周围直接相邻宏块的分割模式

Fig. 4 The divided mode map of macroblocks directly surrounding the lost one

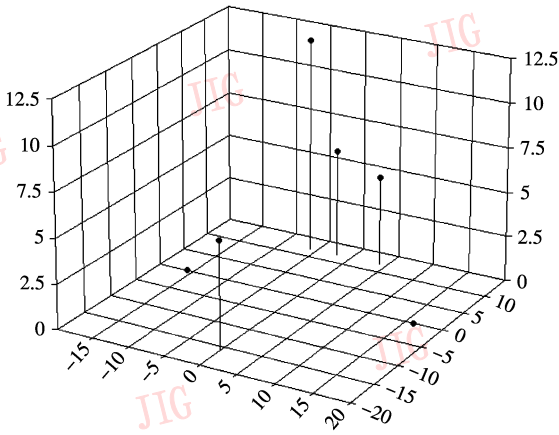


图 5 损坏宏块周围直接相邻划分的运动矢量分布
Fig. 5 The distribution of the motion vectors directly surrounding the lost macroblock

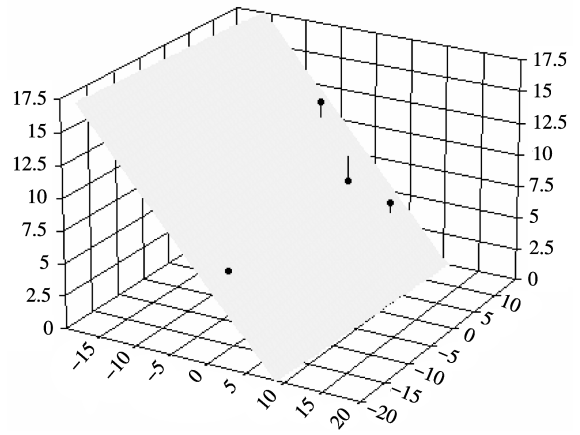


图 6 平面拟合示意图
Fig. 6 The illustration of plane fitting

的就是小范围内相邻运动矢量的大致变化情况。令待求的平面方程为

$$z(x,y) = a + bx + cy \quad (1)$$

式中, (x,y) 为每个划分的坐标, z 为相应的运动矢量, a,b,c 为待定系数。 a,b,c 通过使得样点值与拟合值之差的平方和最小求得。

$$S = \sum [P_i - (a + bx + cy)]^2, P_i \in \mathbf{P} \quad (2)$$

注意到式(2)中的 P_i, x, y 已知, 故 a, b, c 是 S 的函数。由无约束最优化问题的一阶必要条件知, 当 S 取得最小值时, S 对 a, b, c 的偏导数 $\frac{\partial S}{\partial a}, \frac{\partial S}{\partial b}, \frac{\partial S}{\partial c}$ 必为 0, 即

$$\begin{cases} 2 \sum [P_i - (a + bx + cy)] = 0 \\ 2 \sum [P_i - (a + bx + cy)]x = 0 \\ 2 \sum [P_i - (a + bx + cy)]y = 0 \end{cases} \quad (3)$$

将式(3)展开整理得:

$$\begin{cases} (\sum 1)a + (\sum x)b + (\sum y)c = \sum P_i \\ (\sum x)a + (\sum x^2)b + (\sum xy)c = \sum xP_i \\ (\sum y)a + (\sum xy)b + (\sum y^2)c = \sum yP_i \end{cases} \quad (4)$$

解式(4)即可求得 a, b, c , 其中累加的项数为集合 \mathbf{P} 中元素的个数, 即与受损宏块直接相邻的划分运动矢量个数。求得了拟合平面, 坐标 $(0,0)$ 处的 Z 值, 此时 $Z(0,0) = a$, 即为待估计的受损宏块运动矢量值, 图 5 中的运动矢量分布采用平面拟合后的效果如图 6 所示。

3 实验结果及分析

实验平台为 Intel Pentium 4 3.00GHz CPU, 512MB DDR 内存的 PC 机, 程序运行环境 VC++ 6.0, 在 JM8.6 标准测试模型中实现算法。编码端设置帧的结构为 IPPP..., 对 Foreman, mobile, coast-guard 3 个序列进行编码生成 H.264 码流。所选序列均为 QCIF 格式。为了测试算法在不同码率情况下的隐藏性能, 分别设置 QP 的值为 16, 20, 24, 28 来进行试验。表 1 列出了在各种不同码率、不同宏块丢失率的情况下本文算法(PF)以 PSNR 为依据的隐藏性能评价结果, 并与几种典型算法对发生相同错误时的隐藏性能进行了比较。注意, 表中 PSNR 值只是对亮度分量进行计算得到的。显然, 客观上, 本文算法的性能整体上比直接替代算法(TR)和运动矢量均值补偿隐藏算法(AVG)要好。图 7 给出了它们在 Foreman 序列上执行后主观质量比较, 图 7(a)中黑色的方块表示一帧中受损的宏块, 图 7(b)、图 7(c)和图 7(d)分别为采用 TR, AVG 和本文算法隐藏后的效果, 注意观察人头上的帽子前沿和身后墙体空隙处受损的那 3 个宏块, 可以明显地看出用本文算法恢复得很真实, 而用其他方法恢复后的结果出现明显的不连续。

实验表明:

1) 理论上算法中集合 \mathbf{P} 的元素越多, 平面拟合后来求得的运动矢量结果越准确, 错误隐藏的质量越好。实验中参与拟合的样点不能少于 4 个, 实际

表 1 本文算法与几种典型隐藏算法的隐藏性能比较

Tab.1 The objective performance comparison between the proposed algorithm and other common used error concealment algorithm

视频序列	QP	原始 PSNR/dB	算法	不同宏块丢失率(MLR)下的 PSNR/dB		
				1%	5%	8%
Foreman	16	44.787 9	TR 算法	37.066 8	34.018 1	32.845 3
			AVG 算法	44.055 7	41.787 8	41.651 3
			本文算法	44.055 7	42.967 1	42.795 8
	20	41.665 7	TR 算法	36.372 3	33.681 0	32.600 9
			AVG 算法	41.665 7	40.439 2	39.607 3
			本文算法	41.665 7	40.822 1	40.260 3
	24	39.036 7	TR 算法	35.424 2	33.173 8	32.206 9
			AVG 算法	37.967 0	37.191 3	36.877 6
			本文算法	38.947 0	36.218 6	36.027 3
28	36.395 8	TR 算法	34.105 7	32.382 8	31.595 4	
		AVG 算法	35.219 9	34.046 3	33.337 8	
		本文算法	36.375 8	35.398 4	35.236 6	
mobile	16	44.113 5	TR 算法	42.807 5	38.312 6	36.456 5
			AVG 算法	42.807 5	40.956 9	34.464 8
			本文算法	43.626 4	41.447 1	32.249 5
	20	40.499 2	TR 算法	39.896 0	37.111 5	35.651 1
			AVG 算法	39.896 0	39.070 5	37.396 3
			本文算法	40.279 6	39.355 2	37.531 4
	24	36.836 4	TR 算法	36.572 7	35.106 2	34.128 0
			AVG 算法	36.572 7	36.134 1	35.205 8
			本文算法	36.750 3	36.291 0	35.539 1
28	33.502 6	TR 算法	33.393 1	32.705 6	32.127 4	
		AVG 算法	33.393 1	33.240 1	32.747 6	
		本文算法	33.393 1	33.058 4	32.544 6	
coast-guard	16	44.203 1	TR 算法	43.383 1	41.452 8	40.293 5
			AVG 算法	44.155 5	43.339 0	42.178 9
			本文算法	44.155 5	43.509 6	42.292 2
	20	40.705 0	TR 算法	40.321 1	39.280 0	38.576 5
			AVG 算法	40.689 0	40.157 2	39.526 7
			本文算法	40.689 0	40.407 2	39.742 1
	24	37.469 1	TR 算法	37.286 5	36.814 0	36.429 3
			AVG 算法	37.408 1	37.283 6	36.980 2
			本文算法	37.469 1	37.386 9	36.950 6
28	34.573 4	TR 算法	34.477 7	34.253 9	34.081 7	
		AVG 算法	34.477 7	34.290 5	34.146 5	
		本文算法	34.477 7	34.245 9	34.074 0	

上这一条件正因为 H.264/AVC 具有可变尺寸的树型结构宏块划分模式这一编码特点而很容易满足,前面已提到 H.264/AVC 中与受损宏块直接相邻的

划分个数最多可达 16 个。另外,实现时对于 $Z=0$ 的点不参与拟合。

2) 以 PSNR 值作为算法客观质量评价时,可以

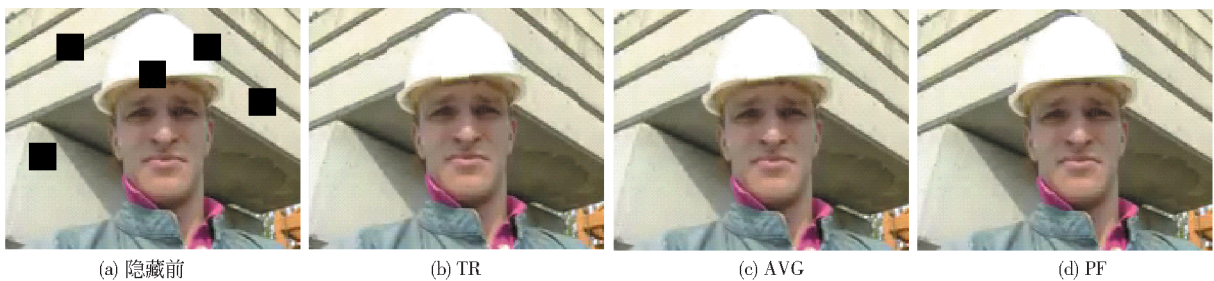


图7 本文算法与几种典型隐藏算法执行后主观质量比较

Fig. 7 Comparison of the subjective quality review between the proposed algorithm and other common used error concealment algorithm

选用整个 GOP 解码后帧的 PSNR 均值,也可以仅以隐藏后得到的当前帧 PSNR 值,表 1 中的数据是用后者。不过视频压缩由于采用时域预测编码,解码时当前帧的错误隐藏效果会直接影响到后续帧的解码质量,也即会影响到序列解码后的 PSNR 值。实际应用中 GOP 的长度不宜过长,适当地设置 IDR 帧能有效避免时域错误的扩散。

3) 对于采用多参考帧搜索的情况,当受损宏块直接相邻的每个划分所用的参考帧不一致时,可以认为错误宏块的参考帧与它们其中某个划分相同,遍历这些候选的参考帧,根据最小边界匹配失真的原则来确定错误宏块最终所用的参考帧。

4) 当受损宏块的运动特性与其相邻宏块不发生剧变时,用该算法能获得非常满意的恢复效果。对于只有局部、小幅运动的视频序列,由于可能有较多宏块被编码为帧内编码模式,这些宏块就没有相应的运动矢量,此时用直接时域替代的效果最好。对于那些具有快速、复杂运动特点的视频序列,用直接时域替代、均值运动补偿、平面拟合算法都是不够理想的。本文算法较适用于具有中等运动速度的视频序列。

5) 本文算法把一个宏块作为整体来恢复的,而有可能受损宏块在编码时也被划分为更多的子块,此时再作为整体来处理显然不大适合。如何确定受损宏块的划分模式,再对每个划分来恢复是下一步值得研究的问题。

6) 为了获取更好的错误隐藏效果,可以把前面的操作视为错误隐藏过程的第一步,接着可以采用另外一些方法来进一步改进重建图像的质量。例如可建立自适应的马尔可夫随机场图像先验模型,用最大后验估计来获取更好的效果。

4 结 论

在分析 H.264/AVC 错误隐藏新特征的基础上,提出一种基于平面拟合的运动矢量恢复算法,算法的创新点在于在 3 维空间中用一个平面来表征小范围内相邻运动矢量的整体变化情况,从而用拟合得到的平面来求得运动矢量。实验结果表明该算法无论在主观质量评价还是客观质量评价较典型隐藏算法都具有优越性。

参考文献 (References)

- [1] Cao Jihua, Li Fengting. Error concealment techniques in MPEG-2 video decoders[J]. Journal of Tsinghua University (Science & Technology), 2004, 44(7): 921-924. [曹继华, 李凤亭. MPEG-2 视频解码器中的错误隐藏技术[J]. 清华大学学报: 自然科学版, 2004, 44(7): 921-924.]
- [2] Xu Y, Zhou Y. H.264 video communication based refined error concealment schemes [J]. IEEE Transactions on Consumer Electronics, 2004, 50(4): 1135-1141.
- [3] Ma Yufeng, Cai Anni. A new spatial error concealment method based on minimum pixel span criteria [J]. Journal on Communications, 2005, 26(1): 1-5. [马宇峰, 蔡安妮. 基于最小像素跨度准则的空域错误隐藏方法[J]. 通信学报, 2005, 26(1): 1-5.]
- [4] Peng Qiang, Yang Tianwu, Zhu Changqian. Block-based temporal error concealment for video packet using motion vector extraction [C]//Proceeding of ICCAS' 02. Piscataway, NJ, USA: IEEE Press, 2002: 10-14.
- [5] Shahram S, Faouzi K, Rabab W. Concealment method for video communications in an error-prone environment[J]. IEEE Journal on Selected Areas in Communications, 2000, 18(6): 1122-1128.

- [6] Lam W M, Reibman A R, Liu B. Recovery of lost or erroneously received motion vectors [C]//Proceedings of ICASSP. Piscataway, NJ, USA: IEEE Press, 1993, 5: 417-420.
- [7] Al-Mualla M, Canagarajah N, Bull D R. Temporal error concealment using motion field interpolation [J]. Electronics Letters, 1999, 35(3): 215-217.
- [8] Chen Y, Hu Y, Oscar C A, et al. Video error concealment using spatio-temporal boundary matching and partial differential equation [J]. IEEE Transactions on Multimedia, 2008, 10(1): 2-15.
- [9] Zheng J H, Chau L P. Error concealment algorithm for H. 26L using first-order plane estimation [J]. IEEE Transactions on Multimedia, 2004, 6(6): 801-805.
- [10] Zheng J H, Chau L P. Efficient motion vector recovery algorithm for H. 264 based on a polynomial model [J]. IEEE Transactions on Multimedia, 2005, 7(3): 507-513.