

基于运动时空相关度的时域误码掩盖算法^①

黄治华^② * ** 易本顺 * 喻占武 *

(* 武汉大学电子信息学院 武汉 430074)

(** 武汉邮电科学研究院 武汉 430079)

摘 要 针对视频通信中时域误码掩盖中运动矢量恢复存在偏差的问题,提出了一种依据相似三角形特点的运动矢量估算方法。同时,利用受损块与前续块、邻域块的运动时空相关性,提出了一种基于运动时空相关度的时域掩盖算法:对于受损块,采用提出的相似三角形运动矢量估算方法和拉格朗日插值法分别估算其运动时域相关和运动空域相关的运动矢量,然后评估受损块的运动时空相关度,据此对估算的运动时域和运动空域相关的运动矢量进行加权平均以得到最终掩盖的运动矢量,从而实现受损块的时域掩盖。仿真结果表明,对不同宏块丢失率或不同视频序列,该算法均可比传统算法恢复出更高质量的图像。

关键词 视频传输,视频编码,误码掩盖,运动矢量(MV),边界匹配,H.264

0 引言

在视频通信中,当接受端接收的码流中存在误码或者丢包时,解码器重建的视频信号会存在失真。尤其是在基于运动补偿、预测编码及变长编码技术的视频编码标准中,解码错误会在时域和空域扩散,严重降低主观视觉效果。此时在解码端引入误码掩盖技术,利用视频信号的时空相关性,对受损的信号进行恢复,可大幅提高受损图像的重建质量。

当帧间编码宏块受损时,一般使用时域误码掩盖算法,其关键是运动矢量(motion vector, MV)的恢复。基本的运动矢量恢复算法包括:①运动矢量置零,即时域替换(temporal replacement, TR)算法^[1];②使用前一帧对应块的运动矢量(collocated MV, CMV);③使用邻域块运动矢量的平均值(average MV, AMV)或中值(median MV, MMV)。文献[2]提出的边界匹配算法(boundary matching algorithm, BMA)以边界失真最小为准则,从多个候选运动矢量中选择最佳运动矢量,该算法已被 H.264 参考软件 JM 所采用。文献[3]利用运动跟踪技术,将前一帧对应块邻域内的运动矢量依照其运动趋势外推,评估受损区域与外推预测区域之间的相关性,从而估算出

受损的运动矢量。文献[4]根据邻域块的运动矢量构造拉格朗日插值多项式,以此多项式恢复受损的运动矢量。文献[5]通过建立基于运动矢量变化趋势的二次多项式模型来恢复运动矢量。文献[6]提出的运动域插值算法根据邻域块的运动矢量加权插值得到每个像素的运动矢量。文献[7]运用错误跟踪的思想恢复受损的运动矢量。文献[8]针对丢失区域的特点自适应采用 BMA 算法和拉格朗日插值法。上述算法中,TR、CMV 和运动跟踪等算法利用运动的时域相关性,保证了前后帧之间运动的连续性,这对于运动平缓或具有直线运动特征的区域掩盖效果较好,但当图像中场景变化较大或物体运动剧烈时,这些算法所恢复的运动时域相关的运动矢量与真实的运动存在较大误差;AMV、MMV、BMA、拉格朗日插值等算法利用了运动的空域相关性,所估算的运动空域相关的运动矢量可使受损块与邻域块具有相同的运动趋势,但这些算法需要依赖周围邻域块的运动信息,当传输信道误码率较高,宏块大面积丢失时,由于周围有效信息减少甚至根本没有,这些算法将会完全失效。

为此,本文提出一种依据相似三角形特点的运动矢量估算方法,该方法根据运动的时域相关性,将受损块、前续块(即前一帧对应块)、邻域块的运动矢

① 863 计划(2007AA12Z214)和湖北省软件产业专项资金(2007014)资助项目。

② 男,1977 年生,博士生,工程师;研究方向:多媒体信息与信号处理,视频压缩通信;联系人, E-mail: huangzhihua_wl@yahoo.com.cn (收稿日期:2009-03-13)

量构造为相似三角形,依据相似三角形的特点,估算受损的运动矢量,该方法不仅对运动平缓的区域掩盖有效,而且对图像中物体的缩放及转动、镜头的拉升等掩盖也有效。同时,本文还综合地考虑受损块运动的时空相关度,给出了一种运动时空相关性的度量方法,据此对估算的运动时域和运动空域相关的运动矢量进行加权平均以得到最终掩盖的运动矢量,从而实现受损数据的误码掩盖。

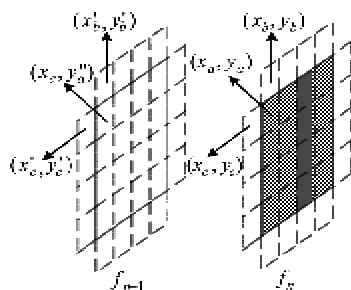
1 相似三角形运动矢量估算法

大量试验表明,镜头中的物体大都进行有规律的运动,如镜头的拉升、物体的有规律转动等,相邻帧之间对应块的运动变化具有相似性,即使物体做不规则运动,只要运动适中,运动变化的相似性条件仍然满足。本文基于这种合理假设,根据前续块及邻域块的运动信息构造相似三角形,并依据相似三角形的特点恢复受损的运动矢量。在H.264树状结构运动估计中,4×4块为运动估计的最小尺寸,其

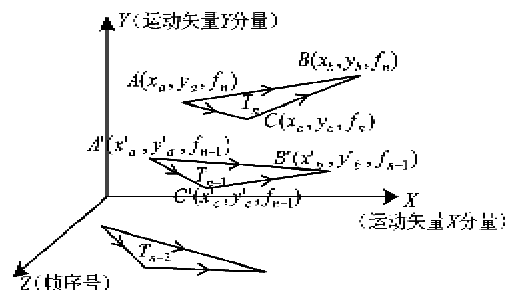
估计的精度最高,因此本文以4×4块为单位恢复受损的运动矢量。

以分别代表运动矢量的水平分量、垂直分量、帧序号的X轴、Y轴、Z轴建立三维坐标系,将视频序列的运动矢量对应到此坐标系中。对于受损的4×4块,假定其在当前帧 f_n 中的运动矢量为 $MV(x_a, y_a)$,则此运动矢量在三维空间中有一坐标点 $A(x_a, y_a, f_n)$ 与之对应。同理,在 f_n 中已完成解码且与受损块横向和纵向距离最近的两个邻域块的运动矢量为 $MV(x_b, y_b)$ 和 $MV(x_c, y_c)$,它们在坐标系中的对应点分别为 $B(x_b, y_b, f_n)$ 和 $C(x_c, y_c, f_n)$ 。受损块及其两个邻域块在前一帧 f_{n-1} 对应块的运动矢量在三维空间中的坐标分别为 $A'(x'_a, y'_a, f_{n-1})$ 、 $B'(x'_b, y'_b, f_{n-1})$ 和 $C'(x'_c, y'_c, f_{n-1})$ 。

当B、C不重合,且A'、B'和C'也两两不重合时,则在平面 $Z = f_n$ 上,A、B和C三点相连构成三角形 T_n ,在平面 $Z = f_{n-1}$ 上,A'、B'和C'三点构成三角形 T_{n-1} ,运动矢量的空间位置关系如图1所示。



(a) 4×4块位置关系



(b) 运动矢量空间坐标

图1 运动矢量空间位置关系图

由于相邻帧之间对应块的运动变化具有相似性,可知三角形 T_n 与 T_{n-1} 相似。同时,在上述6个运动矢量中,前续块及已解码的邻域块的运动矢量均已知,因此,根据相似三角形的特点,可得如下关于 x_a 和 y_a 的方程组:

$$\begin{cases} (x_a - x_b)^2 + (y_a - y_b)^2 = r^2 \\ (x_a - x_c)^2 + (y_a - y_c)^2 = s^2 \end{cases} \quad (1)$$

式中:

$$r^2 = \frac{[(x_b - x_c)^2 + (y_b - y_c)^2]}{[(x'_b - x'_c)^2 + (y'_b - y'_c)^2]} \times [(x'_a - x'_b)^2 + (y'_a - y'_b)^2]$$

$$s^2 = \frac{[(x_b - x_c)^2 + (y_b - y_c)^2]}{[(x'_b - x'_c)^2 + (y'_b - y'_c)^2]}$$

$$\times [(x'_a - x'_c)^2 + (y'_a - y'_c)^2]$$

求解上述二元二次方程组,可得 x_a 和 y_a 。当 $y_b \neq y_c$ 时, x_a 和 y_a 的计算公式为

$$\begin{cases} x_a = \frac{x_b - \omega_1 \omega_2 \pm \sqrt{(1 + \omega_1^2)r^2 - (\omega_2 + x_b \omega_1)^2}}{1 + \omega_1^2} \\ y_a = \omega_1 x_a + \omega_2 + y_b \end{cases} \quad (2)$$

式中:

$$\omega_1 = \frac{x_b - x_c}{y_c - y_b}$$

$$\omega_2 = \frac{r^2 - s^2 + x_c^2 - x_b^2 + y_c^2 - y_b^2}{2(y_c - y_b)} - y_b$$

当 $y_b = y_c$ 时, x_a 和 y_a 的计算公式为

$$\begin{cases} x_a = \frac{r^2 - s^2 + x_c^2 - x_b^2}{2(x_c - x_b)} \\ y_a = y_b \pm \sqrt{r^2 - (x_a - x_b)^2} \end{cases} \quad (3)$$

由式(2)和式(3)可见,式(1)存在两组解,而且这两组解在空间的顶点坐标关于边 BC 对称。但在实际的视频序列中,相邻帧对应块的运动方向往往不会发生改变,因此,所求的顶点 A 除了满足相似三角形的要求外,还应作如下限制: A 与边 BC 的位置关系和 A' 与边 $B'C'$ 的位置关系相同,位置关系的判定表达式为

$$S = \left((x_a - x_b) - \frac{(x_c - x_b)(y_a - y_b)}{(y_c - y_b)} \right) \times \left[(x_{a'} - x_{b'}) - \frac{(x_{c'} - x_{b'}) \times (y_{a'} - y_{b'})}{(y_{c'} - y_{b'})} \right] \quad (4)$$

将所求的两组 x_a 和 y_a 分别代入式(4)中,只有使得 $S > 0$ 的一组 x_a 和 y_a 才为有效的运动矢量。

在上述运动矢量的估算中,利用了当前帧中已解码且与受损块距离最近的邻域块的运动信息。对于受损宏块,若其 4×4 块按照从左到右、从上到下的掩盖次序,则在估算顺序靠后的 4×4 块的运动矢量时,会使用本宏块中顺序靠前的 4×4 块运动矢量,而顺序靠前块也是经过掩盖得到的,这样就必然会将顺序靠前块因误码掩盖而引入的误差传播到顺序靠后块的运动矢量中,从而导致掩盖误差的传播。因此,掩盖应遵循先边缘后中心的原则进行,次序如图2所示。

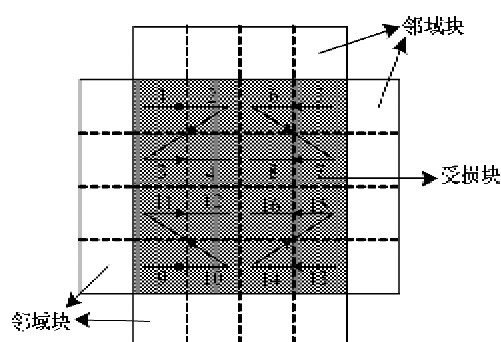


图2 4×4 块掩盖次序图

图中,深色块为受损宏块,浅色块为正确接收的邻域块,块中数字标识了受损块掩盖的顺序。

当 B 、 C 重合,或者 A' 、 B' 和 C' 两点或三点重合时,因不能构成三角形将无法使用上述相似三角形运动矢量估算法,此时,可采用 CMV 算法,即恢复的运动矢量为 $x_a = x_{a'}, y_a = y_{a'}$ 。

2 基于运动时空相关度的时域掩盖算法

基于运动时空相关度的时域掩盖算法充分利用了前续块和邻域块的运动信息,前续块、邻域块及受损块的位置关系如图3所示。

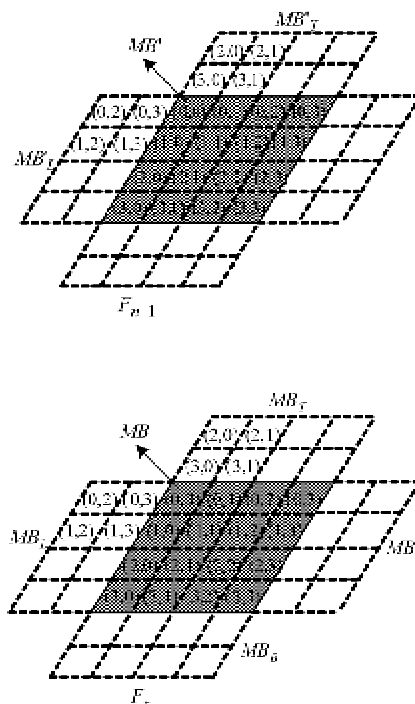


图3 前续块、邻域块与受损块位置关系图

图中, F_n 表示帧 n , 受损宏块及其邻域宏块分别为 MB 和 MB_{dir} , 其中 dir 标识了邻域宏块的位置, $dir \in \{T, B, L, R\}$; MB 、 MB_{dir} 的 4×4 块的运动矢量记为 $MV(m, n)$ 和 $MV_{dir}(m, n)$; 4×4 块在前一帧 F_{n-1} 对应块的运动矢量记为 $MV'(m, n)$ 和 $MV'_{dir}(m, n)$, 其中 m 和 n 标识了 4×4 块在宏块中的位置, $m, n \in \{0, 1, 2, 3\}$ 。

对于受损块 $MB(m, n)$, 首先根据上节提出的相似三角形运动矢量估算法和文献[4]提出的拉格朗日插值法估算其运动时域相关和运动空域相关的运动矢量, 分别记为 $MVT(m, n)$ 和 $MVS(m, n)$ 。

然后, 评估受损块的运动时空相关度。由于相邻的 4×4 块具有相似的运动趋势, 因此, 本文以 8×8 块(记为 $block0 \sim block3$)为单位给出运动时空相关度的评估方法, 即 4×4 块与其所属的 8×8 块具有相同的运动时空相关度。

运动的时域相关度反映了相邻帧之间图像运动的相似程度, 可通过相邻帧对应块的运动矢量的相

关性来衡量。对于受损的 8×8 块,从周围正确接收的宏块中选取与其直接相邻的 4×4 邻域块,并将这些邻域块的运动矢量构成集合 TS_n ,同时,将所选邻域块在前续帧中对应块的运动矢量构成集合 TS_{n-1} 。 TS_n 与 TS_{n-1} 中对应元素的相关性即为受损块的运动时域相关度,其数学表达式为

$$\alpha_T = \frac{[\text{cov}(TS_n, TS_{n-1})]^2}{D(TS_n) \times D(TS_{n-1})} \quad (5)$$

式中: $D(\cdot)$ 为统计方差, $\text{cov}(\cdot)$ 为协方差。例如,对于 $block0$, 与其相邻的 4×4 块为 $MB_T(3,0)$ 、 $MB_T(3,1)$ 、 $MB_L(0,3)$ 和 $MB_L(1,3)$, 由此可得 $TS_n = \{MV_T(3,0), MV_T(3,1), MV_L(0,3), MV_L(1,3)\}$, $TS_{n-1} = \{MV'_T(3,0), MV'_T(3,1), MV'_L(0,3), MV'_L(1,3)\}$, 根据式(5)计算可得 $block0$ 的运动时域相关度。

运动的空域相关度表现为同一帧中相邻区域图像运动的相似程度,通过计算受损块邻域内 4×4 块的运动矢量的相关性可得。计算运动的空域相关度,除了需要 TS_n 外,还需沿着邻域块的选择方向外推,选取与邻域块相邻的块,并以其运动矢量构造集合 SS_n 。 TS_n 与 SS_n 中对应元素的相关性即为受损块的运动空域相关度,其计算公式为

$$\alpha_S = \frac{[\text{cov}(TS_n, SS_n)]^2}{D(TS_n) \times D(SS_n)} \quad (6)$$

对于 $block0$, 沿着邻域块选择的方向外推可得与邻域块相邻的 4×4 块为 $MB_T(2,0)$ 、 $MB_T(2,1)$ 、 $MB_L(0,2)$ 和 $MB_L(1,2)$, 由此可得 $SS_n = \{MV_T(2,0), MV_T(2,1), MV_L(0,2), MV_L(1,2)\}$, 根据式(6)可得 $block0$ 的运动空域相关度。

得到 α_T 和 α_S 后,对估算的 $MVT(m, n)$ 和 $MVS(m, n)$ 赋予相应的加权因子,从而得到受损块掩盖的运动矢量 $MV(m, n)$, $MV(m, n)$ 的计算公式为

$$MV(m, n) = \frac{\alpha_T \times MVT(m, n) + \alpha_S \times MVS(m, n)}{\alpha_T + \alpha_S} \quad (7)$$

同理,可得 $block1$ 、 $block2$ 和 $block3$ 的运动时空相关度,从而可恢复出所有 4×4 块的运动矢量。

3 仿真结果

本文选取 Foreman、Coastguard 和 Suzie 作为测试序列,图像格式为 QCIF,序列长度为 120 帧。在仿真试验中,采用 H.264 的参考软件 JM10.2 作为测试

平台,帧结构为“IPPP...”,在两个 I 帧之间有 39 个 P 帧。I 帧和 P 帧的量化参数为 28,参考帧数为 1,使用全部 7 种不同大小的块进行运动估计。I 帧没有宏块丢失,P 帧的宏块丢失率分别为 5%、10%、20% 及 30%。

采用 4 种算法进行性能比较:本文提出的基于运动时空相关度的算法、TR 算法、BMA 算法及文献[4]提出的拉格朗日插值法(简称为 LGRG 算法)。所比较的均为 Y 分量的峰值信噪比 Y_{PSNR} ,其结果如表 1 至表 3 所示。

表 1 Foreman 序列在不同宏块丢失率下的峰值信噪比

宏块丢失率 %	峰值信噪比(PSNR) dB				
	受损图像	TR	BMA	LGRG	本文算法
0	36.89	36.89	36.89	36.89	36.89
5	17.31	31.03	32.05	33.85	33.97
10	12.72	25.53	27.91	27.97	28.73
20	10.77	23.05	24.38	25.85	26.83
30	9.46	22.53	22.80	24.49	26.02

表 2 Coastguard 序列在不同宏块丢失率下的峰值信噪比

宏块丢失率 %	峰值信噪比(PSNR) dB				
	受损图像	TR	BMA	LGRG	本文算法
0	34.70	34.70	34.70	34.70	34.70
5	19.63	29.85	30.78	30.99	31.21
10	16.21	26.68	26.97	27.59	27.97
20	13.99	24.44	25.28	25.71	26.48
30	12.56	22.90	23.94	24.49	25.15

表 3 Suzie 序列在不同宏块丢失率下的峰值信噪比

宏块丢失率 %	峰值信噪比(PSNR) dB				
	受损图像	TR	BMA	LGRG	本文算法
0	37.16	37.16	37.16	37.16	37.16
5	21.22	33.17	35.31	35.45	36.21
10	16.82	29.76	32.82	33.26	33.91
20	13.60	24.94	26.78	28.58	29.57
30	11.88	23.69	25.52	27.75	28.71

三个序列具有不同运动特性,Foreman 序列不但存在背景的抖动和变化,而且前景也在发生变化;Coastguard 序列具有中等速率的运动;Suzie 序列不仅具有中等速率的运动,而且运动方向随机。从表 1 至表 3 可见,在测试中对于不同运动类型的序列和不同的宏块丢失率,本文算法均优于 TR、BMA 和

LGRG 算法,是一种鲁棒性很强的算法。Foreman 序列中人头在有规律地转动,TR 算法不精确的掩盖导致误差扩散,使得掩盖后的图像比实际头部的转动有所延迟,因此重建图像的 PSNR 不尽人意。而本文提出的相似三角形运动矢量估算方法,通过前续块和邻域块的运动矢量构造相似三角形模型,以此模型跟踪邻域内运动矢量的相对变化趋势,从而恢复的运动矢量更准确,掩盖的效果更好。Foreman 序列在 20% 的宏块丢失率的条件下,本文算法相比 TR 算法增益 3.8dB,相比 BMA 算法增益 2.5dB,相比

LGRG 算法增益 1dB。随着宏块丢失率的升高,本文算法的优势更加明显,这是因为 BMA 算法和 LGRG 算法需要利用邻域块信息估算受损的运动矢量,当宏块丢失率较高时,周围可用信息减少,这些算法的性能明显下降,而本文算法除了利用邻域信息外,还充分利用了前续块的运动信息,因此掩盖的效果较好。

下面,以具有中等速率运动的 Suzie 序列第 85 帧为例,从主观视觉角度对本文算法和 TR、BMA 及 LGRG 算法进行比较,重建的图像如图 4 所示。



图 4 Suzie 在 30% 的丢包率下,不同算法重建的第 85 帧图像效果

从图中可见,TR 算法效果最差,在人物的左脸颊及话筒处存在大量的方块效应;BMA 算法重建的图像在人物的手腕、脖子及话筒处有明显的割裂痕迹;LGRG 算法重建的图像在人物的鼻梁及左脸颊处存在明显的切线。这些区域正是正确接收数据和丢失数据的交界处,而本文算法重建的图像消除了这些不足,使图像看上去更平滑,表现出更优异的性能。

4 结论

本文根据运动的时域相关性,提出了一种依据相似三角形特点的运动矢量估算方法,同时还提出了一种基于运动时空相关度的时域误码掩盖算法。

在这种算法中,综合利用了运动的时域和空域相关性,提高了 H.264 压缩视频码流在误码信道上的错误抵抗能力。本文算法完全基于解码端,不需要反馈信道,不改变 H.264 的码流结构,因此具有广泛的实用性。经仿真实验表明,与 TR 算法、BMA 算法和 LGRG 算法相比,本文算法对不同的宏块丢失率、不同类型的视频序列均能恢复出更高质量的图像,具有更强的错误鲁棒性。下一步的工作主要针对误码掩盖算法与编码端的鲁棒编码方法的有效结合,进行更深入的研究。

参考文献

- [1] Wang Y, Zhu Q F. Error control and concealment for video communication: a review. *Proceedings of the IEEE*, 1998, 86 (5):974-997

- [2] Wang Y K, Hannuksela M M, Varsa V, et al. The error concealment feature in the H.26L test model. In: Proceedings of the 2002 International Conference on Image Processing, Rochester, USA, 2002. 729-732
- [3] Pyun J Y, Lee J S, Jeong J W, et al. Robust error concealment for visual communications in burst-packet-loss networks. *IEEE Transactions on Consumer Electronics*, 2003, 49(4): 1013-1019
- [4] Zheng J H, Chau L P. A motion vector recovery algorithm for digital video using Lagrange interpolation. *IEEE Transactions on Multimedia*, 2003, 49(4):383-389
- [5] Zheng J H, Chau L P. Efficient motion vector recovery algorithm for H.264 based on a polynomial model. *IEEE Transactions on Multimedia*, 2005, 7(3):507-513
- [6] Chen B N, Lin Y. Temporal error concealment using selective motion filed interpolation. *Electronics letters*, 2006, 42(24): 1390-1391
- [7] Kung W Y, Kim C S, Jay C C. Spatial and temporal error concealment techniques for video transmission over noisy channels. *IEEE Transactions on Circuits and Systems for Video Technology*, 2006, 16(7):789-802
- [8] Wu J, Liu X G, Yoo K Y. A temporal error concealment method for H.264/AVC using motion vector recovery. *IEEE Transactions on Consumer Electronics*, 2008, 54(4): 1880-1885

An efficient temporal error concealment algorithm based on temporal and spatial correlation of motion

Huang Zhihua^{* **}, Yi Benshun^{*}, Yu Zhanwu^{*}

(^{*} School of Electronic Information, Wuhan University, Wuhan 430074)

(^{**} Wuhan Research Institute of Post & Telecommunication, Wuhan 430079)

Abstract

To overcome the inaccurate estimation of motion vectors in temporal error concealment in video communications, a motion vector recovery method based on the characteristics of similar triangle is given. Furthermore, a novel temporal error concealment algorithm is proposed, which efficiently exploits the temporal and spatial correlation of motion. For a lost macro block, the temporal related motion vector and the spatial related motion vector are first estimated by the similar triangle method and the Lagrange interpolation method respectively, then the temporal and spatial correlation of motion is calculated, finally, the recovery motion vector is obtained by using a weighted average of the temporal and spatial related motion vectors according to the magnitude of correlation. The simulation results show that compared with the traditional algorithms, the proposed algorithm can recover images with high quality in different rates of lost block or different video sequences.

Key words: video transmission, video coding, error concealment, motion vector(MV), boundary matching, H.264