

基于 H. 264 的立体视频右图像 整帧丢失错误隐藏算法

庞林娟^{1),2),4)} 郁梅^{1),2)} 蒋刚毅^{2),3)} 叶锡恩²⁾ 蒋志迪²⁾ 刘尉悦^{2),3)}

¹⁾ (中国科学院计算技术研究所, 北京 100080) ²⁾ (宁波大学信息学院, 宁波 315211)

³⁾ (北京大学视听信息处理国家重点实验室, 北京 100871) ⁴⁾ (中国科学院研究生院, 北京 100080)

摘要 当 H. 264 编码立体视频流在 Internet 上传输时, 由于信道错误所引起的数据丢失常常会造成整帧图像的丢失。为了恢复丢失的整帧立体视频右图像, 提出了一种基于 H. 264 的立体视频右图像整帧丢失错误隐藏算法, 该算法依据立体视频编码特点进行相关性分析, 首先确定丢失帧中每个宏块的预测方式, 然后采用运动补偿或视差补偿对其进行恢复, 进而重现丢失帧。实验结果表明, 该算法能够在较低的计算复杂度下, 获得较高质量的立体视频图像。

关键词 立体视频 整帧丢失 错误隐藏 预测方式时域相关性

中图分类号: TN919.3 **文献标识码**: A **文章编号**: 1006-8961(2006)11-1584-04

An Error Concealment Algorithm based on H. 264 for Entire Right Frame Loss in Stereoscopic Video Sequence

PANG Lin-juan^{1),2),4)}, YU Mei^{1),2)}, JIANG Gang-yi^{2),3)}, YE Xi-en²⁾, JIANG Zhi-di²⁾, LIU Wei-yue^{2),3)}

¹⁾ (Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080)

²⁾ (Faculty of Information Science and Engineer, Ningbo University, Ningbo 315211)

³⁾ (National Key Laboratory of Machine Perception, Peking University, Beijing 100871)

⁴⁾ (Graduation School of Chinese Academic of Sciences, Beijing 100080)

Abstract When the compressed stereoscopic video stream generated by H. 264 coder is transmitted over the Internet, it is sensitive to packet loss. Due to the codec structure based on H. 264, a packet loss always causes the loss of an entire video frame. In such a case, a novel error concealment algorithm dealing with the loss of an entire right frame is proposed in this paper. Firstly it performs relativity analysis according to the coding scheme of stereoscopic video sequence, then obtains the prediction mode for every MB and conceals it. The experimental results show that the proposed algorithm can achieve better concealed quality with low complexity, so it is an efficient error concealment algorithm.

Keywords stereoscopic video, loss of a whole right frame, error concealment, temporal relativity of prediction mode

1 引言

由于信号传输干扰、噪声和信道拥塞等原因, 导致信道不可避免会有误码或丢包等现象发生, 对于 H. 264 编码立体视频流来说, 可能会造成整帧图像的丢失, 从而不仅影响当前帧的质量, 而且由于采用帧间编码, 这种影响还会连续传递到以后的解码帧, 从

而导致整个立体视频序列质量的严重恶化。在这种情况下, 为了能够得到质量可以接受的视频图像, 在解码端采用错误掩盖技术是一种行之有效的方法。然而, 目前绝大多数错误掩盖技术都只适用于单通道视频中的若干宏块或片 (slice) 出错的情况^[1,2], 极少数已有的整帧丢失错误掩盖算法^[3,4] 由于没有考虑到立体视频序列的特点, 故也不能很好地适用于立体视频流。鉴于此, 本文提出了一种基于 H. 264 的立体视

基金项目: 教育部科研基金项目(2060590); 浙江省自然科学基金(RC01057, Y105577); 浙江省科技攻关项目(2004C31105)

收稿日期: 2006-06-15; **改回日期**: 2006-08-03

第一作者简介: 庞林娟(1981~), 女, 2004年获河北工业大学学士学位, 现为中国科学院计算技术研究所在读硕士研究生。现主要研究方向为立体视频网络传输、错误隐藏。E-mail: panglj04@mails.gucas.ac.cn

频右图像整帧丢失的错误隐藏算法。它首先根据立体视频的编码特点进行相关性分析,然后根据分析结果,利用右通道图像预测方式的时域相关性来确定丢失帧中每个宏块的预测方式以恢复丢失帧。

2 立体视频编码与相关性分析

利用 H. 264/AVC 的时域分级特性^[5]来编码立体视频信号。时域分级包含基本层和增强层两层,分别对应立体视频的左右通道。目前常用的立体视频编码方案是 MCP(motion compensation prediction)和 DCP(disparity compensation prediction)相结合的方法^[6]。编码时,首先对左通道采用 H. 264 编码,对右通道中每帧则分别进行 MCP 和 DCP,然后选择误差较小者作为预测结果。此编码方案不仅可以去除各通道内部相邻帧之间的时域冗余信息和立体图像对之间的空域冗余信息,而且 MCP 和 DCP 相结合还可以对由于运动引起的“遮挡”和“暴露”区域进行很好的处理。

2.1 视差矢量的时域相关性

视差矢量信息是进行立体图像分析的基础,它不仅体现了立体图像对中左右图像的空间相关性,并且还具有很强的时域相关性^[7],可以相邻图像中对应块的视差矢量差值的分布情况为标准来分析这一相关性。由于本文算法只是针对平行相机系统,故只考虑视差矢量的水平分量 d_x 。

令 $[-S, +S]$ 为视差估计的水平搜索范围, $d_x^{(n-1)}, d_x^{(n)}$ 分别为相邻图像中相同位置宏块视差矢量的水平分量, $k = d_x^{(n)} - d_x^{(n-1)}$ 为这两个视差分量的差值。可通过计算相邻帧中满足 $k=0$ 的宏块所占的比例来评价视差矢量的时域相关性,评价函数为

$$L = p_0/p \tag{1}$$

其中, p 是一帧中的总宏块数, p_0 是一帧中满足 $k=0$ 的宏块数。分别对 crowd 和 puppy 立体视频序列进行的实验结果如图 1、图 2 所示,图中 crowd 序列相邻两帧至少 70.7% 的宏块具有相同的视差矢量, puppy 序列相邻两帧至少 73% 的宏块有相同的视差矢量,可见立体视频序列视差矢量具有很强的时域相关性。

2.2 右图像预测方式的时域相关性

由于编码立体视频的右通道图像采用了 MCP 和 DCP 两种预测方式,故本文提出的预测方式相关性是针对右通道图像而言。相关性函数定义如下:

$$R = p_{\text{same}}/n_{\text{total}} \tag{2}$$

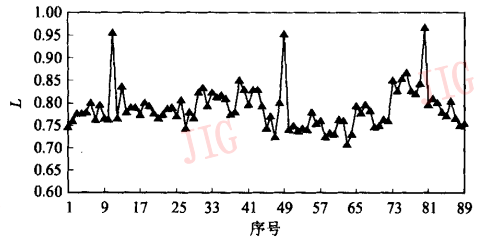


图 1 crowd 序列视差矢量时域相关性实验结果

Fig. 1 Disparity vector's temporal relativity for crowd

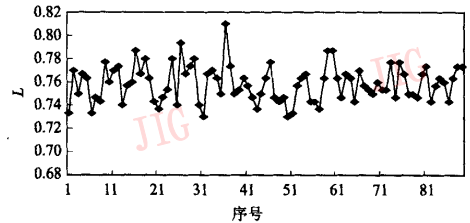


图 2 puppy 序列视差矢量时域相关性实验结果

Fig. 2 Disparity vector's temporal relativity for puppy

其中, p_{same} 指相邻帧中预测方式相同的宏块数(预测方式相同是指对应宏块同时采用 MCP 或 DCP), n_{total} 是一帧中总宏块数。实验以 H. 264 JM86^[8] 为平台,最大参考帧数为 2。在 JM86 中,参考队列按参考图像序号 $N_{\text{reference}}$ 降序排列,当宏块的 $N_{\text{reference}}$ 为 0 时,表示采用 DCP 预测方式;当宏块的 $N_{\text{reference}}$ 为 1 时,表示采用 MCP 预测。预测方式确定是通过统计右视点相邻两帧具有相同参考图像序号的宏块个数,即 p_{same} 值,再利用式(2)评价立体视频序列右通道的预测方式的时域相关性。对 crowd 和 puppy 序列分别进行实验,实验结果如图 3、图 4 所示, puppy 序列相邻帧至少 94.5% 的宏块具有相同的预测方式, crowd 序列相邻帧至少 82.5% 的宏块具有相同的预测方式,可见立体视频右通道图像预测方式具有很强的时域相关性。

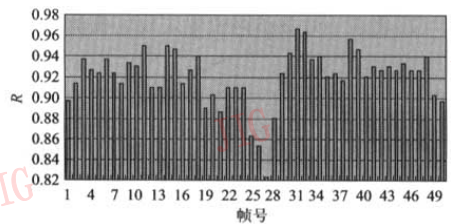


图 3 crowd 序列右通道预测方式时域相关性实验

Fig. 3 Prediction mode's temporal relativity for crowd

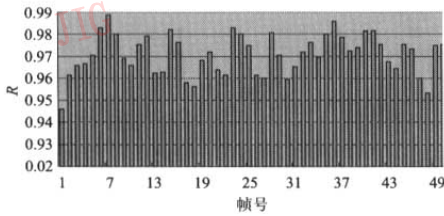


图 4 puppy 序列右通道预测方式时域相关性实验
Fig. 4 Prediction mode's temporal relativity for puppy

3 本文提出的错误隐藏算法

受损宏块的运动矢量可通过线性外推和加权平均的方式获得(如图 5 所示)。其中, F_n 为当前丢失的右通道帧, F_{n-2} 和 F_{n-1} 表示右通道正确接收的前两帧($n, n-1$ 和 $n-2$ 分别表示帧序号), B (斜线标示的方块)表示 F_n 中当前受损宏块。设定搜索范围为 16×16 (如图 5 所示), 在 F_{n-1} 中有 9 个相关宏块 ($B_1 \sim B_9$) 可以利用, 其中 B_5 是 F_{n-1} 中与 B 处于相同位置的宏块。首先判断每个相关宏块的预测方式, 而对于进行 MCP 的相关宏块则根据其自身运动矢量将它们投影到 F_n , 然后以投影宏块与受损宏块的重叠区域(黑色区)的面积(图 5 中的 A_5)作为权值, 同时利用式(3)计算当前受损宏块的运动矢量 $V_E^{(n)}$ 。其中, α_i 表示第 $n-1$ 帧 F_{n-1} 中第 i 个宏块运动矢量 $V_i^{(n-1)}$ 的权值, 即重叠区域的面积。 $V_i^{(n-1)}$ 是 F_{n-1} 中宏块 B_i 的运动矢量。在没有任何重叠区域时, 令 $V_E^{(n)} = 0$ 。

$$V_E^{(n)} = \frac{\sum_{i=1}^M \alpha_i \times V_i^{(n-1)}}{\sum_{i=1}^M \alpha_i} \quad (3)$$

基于前述的相关性分析和运动矢量的获取过程, 本文错误隐藏算法主要步骤如下所示:

- (1) 获取右通道前一帧的相关信息, 包括预测方式和运动矢量等;
- (2) 将前一时刻相同位置宏块的预测方式作为

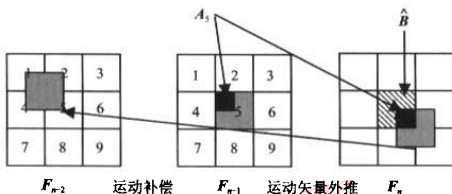


图 5 外推运动矢量的获取过程
Fig. 5 Process of extrapolating motion vector

当前受损宏块的预测方式;

- (3) 对于 MCP 宏块, 先利用图 5 所示方法获取该宏块的运动矢量, 再通过运动补偿恢复该块;
- (4) 对于 DCP 宏块, 先根据视差矢量的时域相关性, 取前一时刻右图像相同位置宏块的视差矢量作为其视差矢量, 然后利用视差补偿对其进行恢复;
- (5) 重复步骤(2)至步骤(4), 恢复整个丢失帧。

4 实验结果

为了评价本算法的效率, 对 puppy 和 crowd 序列做了模拟实验。实验中, 编码器产生的码流格式为 IPPP, 编码 $QP=28$ 。对于每个测试序列, 每次实验先设定其中的一帧右图像丢失, 再用本文进行恢复。为了比较不同算法的优劣, 还设计了如下实验: 利用 BAD (boundary absolute difference)^[8] 作为当前受损宏块用 DCP 和 MCP 进行恢复的判断标准, 其他部分不变。BAD 是 H. 264 进行时域错误隐藏时为丢失块选择最佳运动矢量所采用的判定方法, 称此为基于 BAD 的错误隐藏算法。客观恢复效果如图 6, 图 7 所示, 图 8 给出了 crowd 右通道第 33 帧用本文算法和基于 BAD 的算法恢复的结果图像。结果表明, 本文算法与基于 BAD 的错误隐藏算法相比, crowd 序列的 PSNR 平均提高了 0.73dB, puppy 序列的峰值信噪比 (PSNR) 平均提高了 1.0dB。同时, 以恢复一帧的平均时间衡量两种方法的计算复杂度 (P4 3.0G Hz), 结果如表 1 所示。 t_1 为正确解码一帧的平均时间, t_2 为本文方法恢复一帧的平均时间, t_3 为基于 BAD 的方法恢复一帧的平均时间。表 1 显示本文算法的复杂度明显低于基于 BAD 的算法。最后, 以 puppy 序列为例分析实验结果。设 $C = N_{mc}/N_{dc}$, 其中 N_{mc} 为一帧中进行 MCP 的宏块数, N_{dc} 为一帧中进行 DCP 的宏块数。图 9 表明本文算法在恢复一帧时, 对预测方式的判断更接近于正确

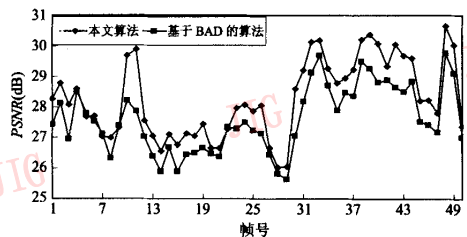


图 6 crowd 两种方法对比实验结果
Fig. 6 Comparison result of two methods for crowd

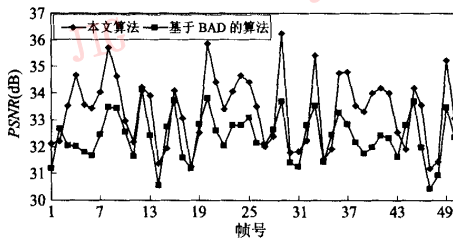


图 7 puppy 两种方法对比实验结果

Fig. 7 Comparison result of two methods for puppy



(a) 正确解码帧(35.17dB)



(b) 本文算法的恢复帧(29.73dB)



(c) 基于 BAD 的错误隐藏算法的恢复帧(28.09dB)

图 8 不同算法的恢复图像对比

Fig. 8 The comparison of different algorithms' restored image

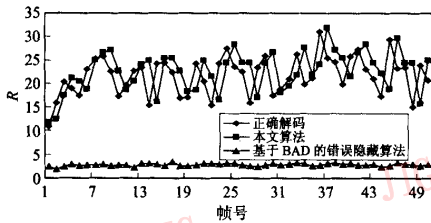


图 9 puppy 序列右通道图像预测方式判断对比试验

Fig. 9 The comparison experiment of judging prediction mode for right view image in puppy

5 结论

本文针对立体视频右通道整帧丢失的情况,提出了一种错误隐藏算法,即先以右通道图像预测方式的时域相关性为判断准则,将前一时刻右图像相同位置宏块的预测方式作为当前受损宏块的预测方式,然后根据选定的预测方式求取合适的运动矢量或视差矢量来恢复当前宏块,重复此步骤即可恢复丢失的整帧右图像。实验结果表明,本文提出的算法能够使恢复帧获得很好的主客观质量。

参考文献 (References)

1 Pyun J, Lee J, Jeong J, et al. Robust error concealment for visual communication in burst-packet-loss networks[J]. IEEE Transactions

表 1 计算复杂度对比实验

Tab.1 Comparison experiment of computing complexity

	t_1 (ms)	t_2 (ms)	t_3 (ms)
crowd	208.72	205.94	284.00
puppy	666.10	690.96	1087.44

解码,故本文方法的恢复效果更好。另外,在基于 BAD 方法中,由于对每个宏块都要计算运动矢量和选取视差矢量,然后才能根据 BAD 在两者间进行选择,故计算复杂度比较大。

on Consumer Electronics, 2003, 49(4):1013~1019.

2 Zhang Rong-fu, Zhou Yuan-hua. An error concealment method for video communications considering both frequency and spatial domains [J]. Journal of Shanghai Jiaotong University, 2004, 38(4):606~609. [张荣福,周源华. 视频通信中频域和空域相结合的错误掩盖技术[J]. 上海交通大学学报, 2004, 38(4):606~609.]

3 Chen Yu, Yu Kim, Li Jiang, et al. An error concealment algorithm for entire frame loss in video transmission [A]. In: Proceedings of Picture Coding Symposium (PCS' 2004) [C], San Francisco, CA, USA., 2004.

4 Liu Li, Yang Gao-bo, Sun Tao. A minimum motion vector span criteria based error concealment algorithm for entire frame losses in H.264 [A]. In: Proceedings of Picture Coding Symposium (PCS' 2006) [C], Beijing, China, 2006.

5 Wiegand T, Sullivan G. Draft ITU-T recommendation and final draft international standard of joint video specification [S]. ITU-T Rec. H.264/ISO/IEC 14 496-10 AVC, in Joint Video Team (JVT) of ISO//IEC MPEG and ITU-T VCEG 2003.

6 Li Shi-ping, Jiang Gang-yi, Yu Mei. Approaches to H.264-based stereoscopic video coding [J]. Computer Engineering and Application, 2005, 41(1):77~79. [李世平,蒋刚毅,郁梅. 基于 H.264 的立体视频编码方法[J]. 计算机工程与应用, 2005, 41(1):77~79.]

7 Zhu Zhong-jie, Jiang Gang-yi, Yu Mei, et al. Fast disparity correspondence algorithm for stereo video coding based on MPEG-2 [J]. Journal of Circuits and Systems, 2003, 8(1):32~35. [朱仲杰,蒋刚毅,郁梅等. 一种基于 MPEG2 的立体视频编码中的视差匹配快速算法[J]. 电路与系统学报, 2003, 8(1):32~35.]

8 JVT reference software[CP/OL]. <http://iphome.hhi.de/suehring/tml/>