

一种适合 3G 网络传输的 H. 264 实时视频容错编码

杜建超 吴成柯 王养利 肖 嵩

(西安电子科技大学 ISN 国家重点实验室, 西安 710071)

摘 要 为了在 3G 网络上传输 H. 264 实时视频流, 提出了一种鲁棒性容错编码策略。该系统首先将视频序列分成两个子序列分别进行编码, 并在不同路径中传输, 以降低因信道突发错误而造成连续丢包的概率, 对丢失的数据则采用高精度的双向加权运动补偿进行预测。为了使出错子流快速恢复正常解码状态, 采用多参考帧策略, 使用正确解码子流的参考帧作为出错子流的参考帧来进行预测编码, 以消除误差扩散。实验结果表明, 该方法在无线网络环境中不仅能有效地抵抗误码和丢包对传输图像质量的影响, 而且能保证视频流的正常播放和改善图像质量。与相关方法比较, 该方法提高重建图像质量超过 0.6dB。

关键词 H. 264 视频编码 容错 3G 网络 多参考帧

中图法分类号: TN919.8 **文献标识码:** A **文章编号:** 1006-8961(2005)11-1355-05

An Error Resilient Coding Scheme for Transmitting H. 264 Real-time Video Streaming Over 3G Networks

DU Jian-chao, WU Cheng-ke, WANG Yang-li, XIAO Song

(National key laboratory of ISN, XiDian University, Xi'an 710071)

Abstract To deliver H. 264 real-time video streaming over 3G networks, an error resilient coding scheme is presented in this paper. The system first divides original sequence into two sub-sequences to encode separately and then transmits the resulting two bitstreams over different channels so that the effect of burst channel error on image quality can be decreased. The frames which are not correctly decoded are predicted by bidirectional weighted interpolation of the frames from the other sub-stream. In order to restore the decoding process of the sub-stream suffering transmission error, the reference picture of the other sub-streams correctly decoded is selected as reference for the error sub-stream so that the error propagation can be eliminated. The simulation results show that the proposed method can efficiently combat bit error and packet loss in wireless channels and improve the image quality. The performance of the proposed method gains more than 0.6dB compared with the related approaches.

Keywords H. 264, video coding, error resilience, 3G network, multiple references

1 引 言

与所有以前的视频编码标准, 比如 ITU-T 的 H. 263 和 ISO/IEC 的 MPEG-4 相比, 由于 H. 264/AVC^[1]极大地提高了编码效率, 其对于同样的编码

质量, 可减少 1/3 到 1/2 的码率^[2], 因此, 以多媒体通信为目标的 3G 系统支持 H. 264 作为其视频编码标准^[3]。虽然 3G 网络在移动环境下的带宽可达 384kbps, 在静止环境下的带宽可以达到 2Mbps, 但是由于信道衰减、建筑物遮挡、终端移动、多用户干涉等原因影响, 使得信道是时变, 且高误码的^[4], 因

基金项目: 国家自然科学基金资助项目(60272050, 60372043)

收稿日期: 2005-08-16; 改回日期: 2005-09-18

第一作者简介: 杜建超(1977 ~), 男。2000 年获西北大学数学专业硕士学位, 现为西安电子科技大学通信工程学院博士研究生。主要研究方向为视频编码、容错与传输。E-mail: dujianchao@21cn.com

此,在 3G 网络上传输视频流时,仅仅追求高的压缩效率是不够的,必须有一定的容错和错误掩盖措施。H.264 中虽然提供了一些容错工具,但是它们有各自不同的用途和目的,即在不同的场合需要选择不同的组合来使用^[4]。比如多参考帧的使用不仅能够提高编码效率,而且当与反馈信道一起使用时,还能够及时地消除误差扩散,但是这通常要求编解码两端都有较大的内存容量,才能存储较多的参考帧。

Wenger 等人提出一种基于多路径的视频冗余编码(video redundancy coding, VRC)方法^[5],该方法是先将视频序列按照奇偶序号分成两组独立编码的子序列,然后在编码时,奇帧参考奇帧,偶帧参考偶帧,这样当其中一个子序列出现传输错误而导致质量下降和误差扩散的时候,另外一个子序列不受影响,这就保证了视频流的基本质量。范晨等提出一种基于棋盘分割插值的多描述视频编码策略^[6],该策略是先将每帧图像按照棋盘模式分割,并插值形成两路相同分辨率的近似图像,然后对这两路近似图像进行独立的运动预测、补偿和编码。当反馈信道可用时,则采用主动跟踪解码器的状态来消除误差积累。但这种方法在具有突发错误特征的无线网络上有-定的缺陷,因为要使跟踪是正确的,必须基于如下的假设:对某一出错帧而言,从编码器发送该帧到收到该帧的错误报告这一时间段内,其他所有发送的帧都被正确收到。很明显这一假设不适合于突发错误。

因此,本文提出一种新的编码方案,即基于多路径和多参考帧策略,首先将视频流编码成两路子流,并在不同的路径中传输,然后使用双向加权运动补偿来预测丢失的数据。而对于出错子流,则使用多参考帧策略来恢复其解码状态。该方法不但不要求编解码端有较大的内存容量,而且在具有突发错误特征的无线信道上传输时,能够及时消除误差积累和改善视频质量。

2 相关工作分析

在实时应用中,丢失的数据包不能通过重传来纠正,因为这将引起不可接受的延迟。这必然使解码图像质量下降,甚至可能无法继续解码。VRC 编码是按照奇偶序号将视频序列分成两路独立编解码的子序列,首先产生两路视频流,并分别在不同的信道中传输。这样,当其中一路视频流受信道误码或者突发

错误影响的时候,另一路视频流还能够正常解码,这就保证了视频流的连续播放;然后,周期性地插入同步帧,使出现错误的子流恢复正常。其编码结构如图 1 所示,图中 S 是同步帧, P_i 是第 i 帧普通编码帧。由于同步帧的插入使得 VRC 码流的冗余度高达 35%^[7],因此这不利于低码率场合的应用。

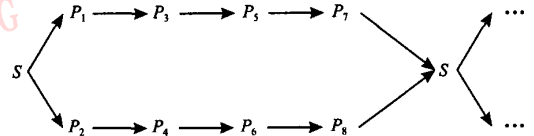


图 1 VRC 编码

Fig. 1 Video redundancy coding

基于棋盘分割的多描述视频编码与 VRC 编码类似,也是将视频序列编码成两个子流,其不同之处在于分割是基于空间域的。当反馈信道可用时,就采用主动跟踪法来消除误差积累。主动跟踪法的原理如下:解码器对出错信息进行错误掩盖的方法事先与编码端约定好,这样当编码端收到解码端的错误报告后,就可以采取同样的掩盖步骤来重复解码器的操作,以使得当前参考帧与解码端的参考帧一致,这样就可以消除误差积累。但是这样做有以下两点不足:(1)编码端缓冲区需要保留的帧数将依赖于错误报告反馈的延迟,若延迟较大,则编码端需要较大的缓存;(2)若编码端收到第 $n-d$ 帧的错误报告,且反馈延迟为 d 帧,当前编码的帧序号为 n ,则该方法假设第 $n-d+1, \dots, n-1$ 帧都被正确收到,然而这是不能保证的,特别是对于有突发错误特征的无线信道,因为编码器需要对第 $n-d, n-d+1, \dots, n-1$ 帧重复进行解码器的错误掩盖操作,以便使编码端用作参考的第 $n-1$ 帧与解码端重建的视频帧一致,如果其中某帧出现了传输错误,则编码端的操作仍然与解码端不一致,而这需要收到以后的错误报告才能得知。

3 基于多参考帧的两路视频流编码方法

3.1 系统结构

由于两路视频流编码比单路编码明显地改善了传输质量^[8],因此本文的方法基于两路视频流编码,其系统结构如图 2 所示。类似于 VRC 编码,其将视频序列按照奇偶序号分成两个子序列分别进行

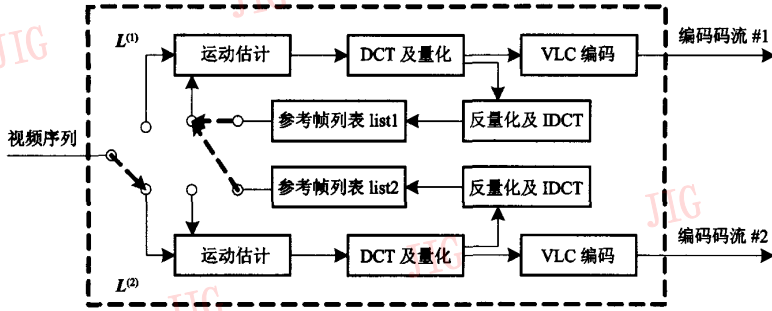


图 2 系统结构

Fig. 2 System structure

预测、补偿和编码。这样就产生了两路码流,并分别不同的信道中传输。不同于同步帧编码,本文方法是用多参考帧来恢复错误子流的解码进程,但这要求编码器的两个预测环共享参考帧列表,并且反馈信道可用。

在图 2 中, $L^{(1)}$ 和 $L^{(2)}$ 指两个独立的预测环,各自有运动估计、DCT(discrete cosine transform)变换及量化、本地重建和 VLC(variable length coding)编码等模块,它们构成完整的编码系统,只是由于输入的视频帧被交替地送入不同的预测环,因此形成了两路不同的编码子流。两个预测环各自维护一个参考帧列表,用于存储前面的解码帧,以便供后续编码帧做运动估计时进行参考。 $L^{(1)}$ 中的参考帧列表称为 $list_1$, $L^{(2)}$ 中的参考帧列表称为 $list_2$ 。正常情况下, $list_1$ 和 $list_2$ 各自服务于所属的预测环。当恢复出错子流时,正确子流预测环中的参考帧列表将被用来选择出错子流的参考帧。比如:当编码端知道 $L^{(2)}$ 产生的码流出现传输错误而导致其编解码两端的参考帧出现误匹配时,则将 $list_1$ 中的帧送入 $L^{(2)}$ 进行参考。

3.2 出错子流的状态恢复

设两个子序列分别为 $N^{(0)} = \{0, 2, 4, \dots, 2k, \dots\}$ 和 $N^{(1)} = \{1, 3, 5, \dots, 2k+1, \dots\}$ 。若编码端收到第 $n-d$ 帧的错误报告时,正在编码第 n 帧,且反馈延迟为 d 帧,则编码器会遇到如下两种情况:

(1) 若 $d\%2 = 0$, 这里 $\%$ 表示取模运算,则当前编码帧(第 n 帧)与出错帧(第 $n-d$ 帧)位于相同子序列,此时需要从另一个编码预测环的参考帧列表中选择一帧作为第 n 帧的参考帧。本文选择第 $n-1$ 帧作为参考帧。

(2) 若 $d\%2 \neq 0$, 则当前编码帧(第 n 帧)与出错帧(第 $n-d$ 帧)位于不同子序列,这时第 n 帧按

正常过程进行编码,即利用第 $n-2$ 帧作为参考帧,而第 $n+1$ 帧(位于出错子序列)则使用第 n 帧(位于正确子序列的参考帧列表)作为参考帧。

在使用参考帧选择恢复出错子流的状态后,编码器启动一个标志 $flag = 1$, 用来跟踪状态恢复是否成功。在上述第 1 种情况下,若编码端收到的下一个错误报告的帧序号 $error_frame \geq n$, 则说明第 $n-d+1, n-d+3, \dots, n-1$ 帧都被正确收到,此时状态恢复成功, $flag$ 清零,否则状态恢复不成功;在第 2 种情况下,若编码器收到的下一个错误报告的帧序号 $error_frame \geq n+1$, 就说明第 $n-d+1, n-d+3, \dots, n$ 帧都正确收到,此时状态恢复成功, $flag$ 清零,否则状态恢复不成功。

之所以状态恢复不成功,是由于两路视频流同时出现传输错误所致,这时必须通过插入 Intra Slice 的方法来进行恢复。该方法在 H.264 的容错工选项中有定义。

3.3 出错帧的误码掩盖

如果系统出现上节所述的第 1 种错误情况,则由于预测编码的原因,其第 $n-d+2, n-d+4, \dots, n-2$ 帧将会出现误差扩散,即使它们被正确收到,其解码结果也是有误差的,因此,本文使用双向加权运动补偿来预测这些帧。

设需要预测的帧为第 k 帧,用 f_k 表示,而其前后两帧 f_{k-1} 和 f_{k+1} 均在另一个子流中。 f_{k+1} 参考 f_{k-1} 的运动矢量集合为 $\{V|f_{k+1} \rightarrow f_{k-1}\}$ 。设 $M_{k+1,i}$ 为 f_{k+1} 中的第 i 个宏块,其运动矢量为 $v_{k+1,i}$, 则 f_k 中同样位置的宏块 $M_{k,i}$ 的前向运动矢量 $v_{k,i}^{(1)}$ ($v_{k,i}^{(1)} \in \{V|f_k \rightarrow f_{k-1}\}$) 按下式计算:

$$v_{k,i}^{(1)} = (d_k \times v_{k+1,i}) / d_{k+1} \quad (1)$$

其中, d_k 和 d_{k+1} 分别表示 f_{k-1} 与 f_k 、 f_{k-1} 与 f_{k+1} 之间

的时间距离。由式(1)得

$$v_{k,i}^{(2)} = v_{k,i}^{(1)} - v_{k+1,i} \quad (2)$$

$v_{k,i}^{(2)}$ 为 $M_{k,i}$ 的后向运动矢量, 即 $v_{k,i}^{(2)} \in \{V | f_k \rightarrow f_{k+1}\}$ 。

由式(1)和式(2), 通过运动补偿就可以分别得到 $M_{k,i}$ 的两个预测宏块 $M_{k,i}^{(1)}$ 和 $M_{k,i}^{(2)}$, 并由此可得到 $M_{k,i}$ 的双向加权补偿宏块 $\hat{M}_{k,i}$

$$\hat{M}_{k,i} = \frac{d_{k+1} - d_k}{d_{k+1}} M_{k,i}^{(1)} + \frac{d_k}{d_{k+1}} M_{k,i}^{(2)} \quad (3)$$

上式右边的两个预测宏块分别取不同的权值, 以对应不同时间距离对补偿宏块的不同影响。

4 实验结果

仿真实验的设置为: 编码器使用修改的 JM97 版本, 网络丢包使用丢包模版仿真软件^[9] 产生; 编

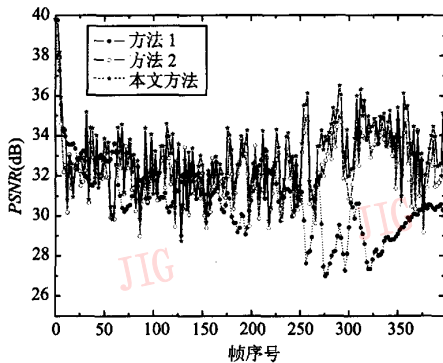
码器参数使用的级别为: ProfileIDC = 66, LevelIDC = 10。视频序列采用 QCIF 格式的 Foreman 和 News, 编码结构为 IPPP..., 帧率为 15, 仅第 1 帧为 I 帧编码。网络带宽设为 64Kbps, 误码率为 2.9e-3 和 1.7e-4 两种模式, 信道编码采用 Turbo 码。RTP (real-time transport protocol) 包最大设为 200Bytes。为验证本文算法的效果, 将 VRC 编码(方法 1)、基于跟踪的棋盘分割编码(方法 2)和本文提出的方法进行了比较, 解码端重建图像亮度的平均峰值信噪比如表 1 所示。

图 3 给出了 Foreman 序列和 News 序列在误码率为 2.9e-3 时各自重建图像的峰值信噪比。由图中可以看出, 本文的方法优于其他两种编码方法, 尤其对于图像内容变化平缓的 News 序列, 本文方法不仅具有优良的抗丢包性能, 而且重建图像的平均峰值信噪比可提高大约 0.6dB。

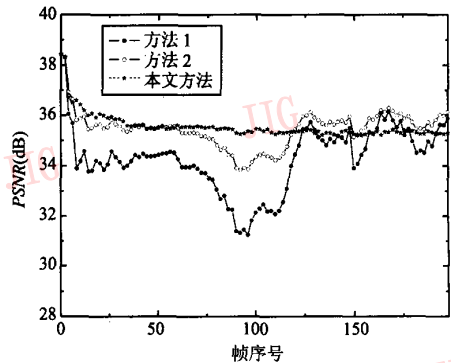
表 1 重建图像的亮度峰值信噪比

Tab.1 The luminance PSNR of reconstructed images

视频序列	编码帧数	误码率	PSNR (dB)		
			方法 1	方法 2	本文方法
Foreman	400	2.9e-3	30.81	32.29	32.84
		1.7e-4	31.18	32.89	33.46
News	200	2.9e-3	34.24	35.48	36.17
		1.7e-4	35.85	36.84	37.73



(a) Foreman 序列重建图像的峰值信噪比



(b) News 序列重建图像的峰值信噪比

图 3 误码率为 2.9e-3 时重建图像的峰值信噪比

Fig.3 The PSNR of reconstructed images with bit error rate 2.9e-3

5 结论

基于 3G 网络高误码率和突发错误的特性, 本

文提出了一种易于实现的鲁棒性编码策略。该方法不仅可通过产生两路编码子流来降低信道突发错误的影响, 并且易于对出错帧进行双向预测。在发生传输错误时, 还可使用正确解码子流的参考帧作为

出错子流的参考帧,以快速恢复出错子流的解码状态和消除误差扩散。实验证明,本文提出的方法在无线网络环境中,由于能有效地抵抗误码和丢包对图像质量的影响,从而可保证视频流的正常播放和改善图像质量。与相关方法比较,可提高重建图像质量约 0.6dB。

参考文献 (References)

- 1 Joint Video Team (JVT). Advanced video coding for generic audiovisual services [S]. H. 264 ISO/IEC 14496-10, ITU-T Recommendation, 2005.
- 2 Wiegand T, Sullivan G J, Bjontegaard G, *et al.* Overview of the H.264/AVC video coding standard [J]. *IEEE Transactions on Circuits System Video Technology*, 2003, 13(7):560 ~ 576.
- 3 3GPP. Packet Switched Conversational Multimedia Applications; Default Codecs V6.3.0 [S]. TS 26.235, 3rd Generation Partnership Project (3 GPP), 2005.
- 4 Stockhammer T, Hannuksela M M, Wiegand T. H.264/AVC in wireless environments [J]. *IEEE Transactions on Circuits System Video Technology*, 2003, 13(7):657 ~ 673.
- 5 Wenger S, Knorr G, Ott J, *et al.* Error resilience support in H.263+[J]. *IEEE Transactions on Circuits System Video Technology*, 1998, 8(7):867 ~ 877.
- 6 FAN Chen, CUI Hui-juan, TANG Kun. Chessboard-interpolation-based multiple description video coding [J]. *Science in China (E Series)*, 2003, 33(10):939 ~ 946. [范晨,崔慧娟,唐昆.基于棋盘分割插值的多描述视频编码[J].*中国科学(E辑)*,2003,33(10):939 ~ 946]
- 7 Apostolopoulos John G. Error-resilient video compression via multiple state streams [A]. In: *Proceedings of the International Workshop on Very Low Bitrate Video Coding [C]*, Kyoto, Japan, LNCS, 1999: 168 ~ 171.
- 8 Eric Setton, Yi Liang, Bernd Girod. Adaptive multiple description video streaming over multiple channels with active probing [A]. In: *Proceedings of IEEE International Conference on Multimedia & Expo [C]*. Baltimore, ML, USA, 2003:509 ~ 512
- 9 Roth G, Sjoberg R, Liebl G, *et al.* Common test conditions for RTP/IP over 3GPP/3GPP2 [S]. ITU-T SG16 Doc. VCEG-M77, ITU-T Recommendation 2001.