

容错视频编码与传输技术研究

姜 东 李 炜 李 波 范圣印 陈 立

(北京航空航天大学计算机学院数字媒体实验室, 北京 100083)

摘 要 随着无线网络和多媒体技术应用的广泛和深入,在不可靠信道上传输视频的需求日益增长,视频编码和传输中的错误控制问题已经引起了广泛关注。本文全面回顾了过去 10 多年来的容错编码与传输技术,包括基于编码端的容错、基于解码端的差错掩盖以及信源信道联合编码。此外还对最新的视频编码标准 H. 264/AVC 的容错工具进行了介绍并展望了容错视频编码与传输技术的发展方向。

关键词 容错视频编码 视频传输 差错掩盖 信源信道联合编码 H. 264/AVC

中图分类号: TP37 **文献标识码:** A **文章编号:** 1006-8961(2006)06-0778-09

Error Resilient Video Coding and Video Transmission Technology

JIANG Dong, LI Wei, LI Bo, FAN Sheng-yin, CHEN Li

(School of Computer, Beijing University of Aeronautics and Astronautics, Beijing 100083)

Abstract With the development of wireless network and multimedia technology, error control technology in video coding and video transmission over lossy channels has been widely studied. This paper reviewed the error-resilient coding and transmission technology, including encoder based error resilience, decoder based error concealment, network based error control and joint source and channel coding. In error resilience, redundant information is added into the bitstream to confine the error propagation. Error concealment estimates or recovers the losing data making use of the character that most nature scenes change smoothly between neighbor blocks. Network based error control, joint source and channel coding utilize the channel feedback information to guide the encoder to make an optimum decision. Additionally, error resilient tools in the latest video coding standard H. 264/AVC and the preview of the future error resilient technology are also considered. Experiments indicate that the proposed error resilient methods are effective in video transmission over lossy channels.

Keywords error resilient coding, video transmission, error concealment, joint source and channel coding, H. 264/AVC

1 引 言

视频领域的两大国际组织 ISO-MPEG 与 ITU 已经相继推出了一系列视频压缩标准如 MPEG-1、MPEG-2 和 MPEG-4 以及 H. 261、H. 263、H. 263 + 等,这些标准不断地提高了视频编码效率。与此同时,各种基于互联网的应用,特别是实时多媒体应用的快速增长使得网络基础设施常常处于超负荷工作状态。而且伴随着网络通信设施的改进,又会衍生出更新更高的应用需求^[1]。人们越来越认识到,无

论如何提高视频的压缩比,海量视频数据与有限的信道带宽始终是一对矛盾。

当前基于 DCT 的变换编码和基于 VLC 的熵编码极大地提高了信源编码的效率,但同时也使得视频流对于传输错误十分敏感。在不可靠信道如电话网和无线网络的情况下(图 1 所示的传输体系),信道出现突发噪声、随机噪声的可能性大大增加,导致视频传输中数据包的丢失^[2]。如果不加控制措施,数据包的丢失会引起错误的扩散,从而导致视频质量的严重下降。因此,研究视频编码和传输中的容错技术,对于解决由于信道误码引起接收端视频质

基金项目:“863”计划(2002AA11901006);国防基础科研和武器装备预研项目

收稿日期:2004-10-10;改回日期:2005-07-11

第一作者简介:姜东(1976~),男,北京航空航天大学计算机学院博士研究生。主要研究方向为多媒体通信、视频容错编码。已在国内外杂志、会议上发表论文近 10 篇。E-mail:duff99@163.com

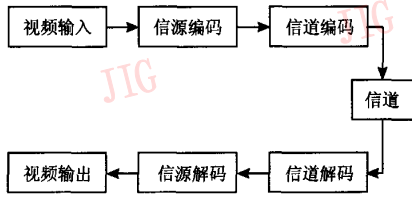


图1 视频传输系统框图

Fig. 1 Diagram of video transmission system

量下降以及由此带来的误码扩散问题具有十分重要的意义。

当前流行的视频编码标准 MPEG-4、H. 263 都各自提供了不同的容错工具^[3,4], 联合视频组 JVT 最新制定的 H. 264/MPEG-4 Part 10 AVC 标准更是把视频流的容错性与视频编码效率摆在了同等重要的地位加以考虑^[5-7]。

2 基于编码端的容错

基于编码端容错的主要做法是以添加适当的冗余信息来防止错误的扩散, 减小错误造成的影响范围。但这会降低编码的效率, 而在遇到传输错误时却能使解码端获得较好的视频质量。容错编码器的设计目标是以最小的冗余得到最大的容错效果。

容错编码有多种添加冗余的做法, 有些可以阻止错误传播, 有些可以方便解码端更好地进行差错掩盖, 有些则在发生错误的情况下力求还能够提供基本的服务。视频编码标准 MPEG-4、H. 263 定义了一系列错误恢复策略, 包括添加再同步标志 (RM)、数据分割 (OP)、可逆变长编码 (RVLC)、帧内编码更新、头信息扩展码和独立分段预测等^[3,4]。此外, 添加区别保护的层编码与多描述编码也是重要的研究方向。

2.1 再同步标记、数据恢复与数据分割

当压缩码流由于传输错误而失去同步信息时, 如果没有补救措施, 解码器将丢弃大量数据从而引起接收端质量的迅速下降。插入再同步标记是一种简单有效的容错方法, 它能将错误限制在两个同步标记之间从而不会引起错误的扩散。

可逆变长编码就是作为一种数据恢复工具, 其思想是在设计 VLC 码的时候做到前向、反向都能解码。研究表明, 对于视频而言, RVLC 能获得接近 VLC 的编码效率和良好的码字结构^[1,8]。RVLC 或

许在一定程度上降低了编码效率但可以充分地提高视频流对错误的鲁棒性。

通常, 一个宏块由运动向量信息和 DCT 数据组成, 另外还有一些控制信息。Ghanbri 提出把数据分割成不同的部分, 以便于在网络上进行不同优先级的传输^[9]。使用数据分割使得运动向量信息的数据对传输错误的鲁棒性得以加强, 同时由于运动边界标记 (MBM) 的插入, 在数据丢失时不必丢弃所有的运动向量信息。文献 [4] 表明在 BER (bit error rate) 为 $10E-3$ 时, 采取数据分割后视频的平均 PSNR 值有近 2dB 的提高。

RM、DP、RVLC 都是 MPEG-4 推荐的容错工具, 这 3 种方法结合使用可取得更好的容错效果。图 2 显示了在突发错误持续时间为 1ms, BER 为 10^{-2} 的信道下 3 种容错工具的不同效果。

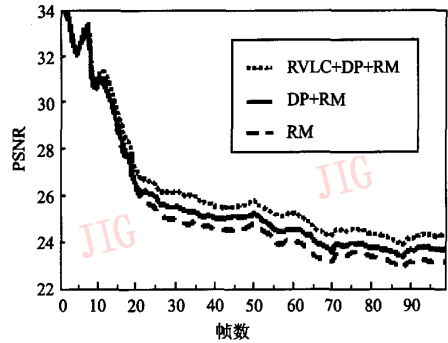


图2 突发错误下 RM、DP、RVLC 的性能对比

Fig. 2 Comparison of the performance of RM、DP、RVLC under burst error

2.2 头信息扩展码

视频数据中有些信息如编码模式、时间戳等头信息对于解码器来说具有特别重要的意义, 如果传输过程中此类信息丢失, 解码器只有丢弃该帧的全部数据。头信息扩展码 (HEC) 是 MPEG-4 为重点保护头信息而提出的一种方法, 其思想是重复备份视频包内的头信息, 这样即便原来的头信息出错也不必丢弃整帧数据。

2.3 容错预测编码

容错预测编码的思想是及时地引入帧内编码或帧内预测的帧或宏块, 这样便可有效的阻止错误的传播^[3]。容错性预测编码的主要做法有帧内编码更新和独立分段预测。

帧内编码更新 (Intra-MB Update): 周期性地插

入帧内编码的帧或宏块是防止错误扩散的有效手段并且同时具有很强的灵活性。文献[10]研究了选择帧内编码的宏块个数的策略,其思想是根据网络的反馈信息动态决定何时插入帧内编码的宏块。文献[2]表明,采用率失真优化策略来决定插入帧内编码宏块的个数及位置可以获得非常好的效果。

独立分段预测(ISP):ISP的思想是把一幅图切成更小的片,在进行运动估计、运动补偿时都限制在该片内进行,也可达到限制错误扩散的目的。H.264/AVC中基于Slice的编码思想其实即是来源于ISP。

2.4 分层编码与多描述编码

为了解决多点播送中主机的带宽异构问题,分层编码(LC)技术已成为一个研究热点并被MPEG-4采用。LC将编码数据分割为一个基本层和若干增强层^[11],每层对应于分辨率从低到高的不同信息。基本层提供了视频序列的基本信息,增强层提供了视频序列更高的分辨率和细节;基本层可以单独传输和解码,而增强层则必须与基本层一起传输和解码。但每个增强层可依次使视频质量更高,如图3所示。

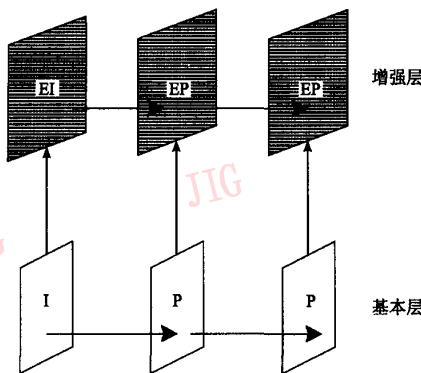


图3 分层编码示例图

Fig. 3 Layered coding

LC作为可扩展编码的主流得到了研究者广泛的关注,当前LC的做法有时域可扩展、空域可扩展、SNR可扩展以及频域可扩展^[12,13]。LC作为一种灵活的编码方式,可以比较有效地解决带宽受限时的视频传输问题。但会大大增加系统的复杂度。此外,LC需要与特殊的传输方案配合使用才能发挥出优势,最常使用的是非平等错误保护(UEP)。UEP的思想是对基本层和增强层视频流采用不同优先级的传输控制策略。文献[14]提出了一种使用UEP的基于率失真优化的分层编码方案,对基本层采用基于前向纠错技术(FEC)来保证相对较高的可靠性。

与分层编码类似,多描述编码(MDC)也是生成多股码流,每一股为一个描述,但生成的每个描述都是相关的并且具有同等的重要性。每一个描述都能独立解码并可获得基本的重建图像,多个描述一起则可使视频质量更好。但从另一个角度说,每个描述的相关性也决定了其冗余性。多路径传输(MPT)作为MDC的传输方案,得到了网络传输研究者的关注^[15-18]。MPT研究如何在物理层上对信道进行比特级别的划分并对传输进行流量和路由控制的技术。但长期以来,MDC与MPT的研究者都是独立进行研究的。文献[19]提出了一种联合设计MDC与MPT的方案以在提高视频流容错性的同时还能保证端到端的可用带宽,如图4所示。

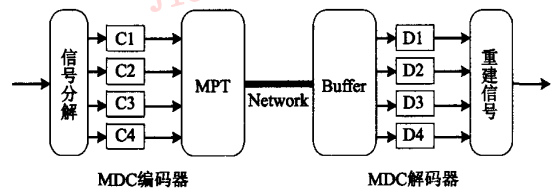


图4 多描述编码及传输框图

Fig. 4 Multiple description coding and transmission architecture

与LC相比,MDC的优势是不需要网络专门提供一个可靠的反馈信道来获得接收端的反馈信息。举例说,在一个丢包率较高的网络上,多次重传意味着多次传输冗余信息,此时,MDC的效率会更高。综合比较,LC采用分层的、不相关的分解方法,而MDC采用非层次结构的、相关的分解。

3 基于解码端的差错掩盖

基于解码端的差错掩盖(error concealment)利用正确接收到的信息来估计或恢复由于传输问题丢失的数据。该方法的主要依据是:自然图像通常主要由低频信息构成,如果没有剧烈运动或不是位于物体的边缘,空间上相邻像素或时间上前后相邻帧对应位置像素色度值的变化是平滑的;另外,人类的视觉系统对于低频信息也比高频信息更敏感。基于解码端的差错掩盖主要是利用了上述特点。

在解码端使用差错掩盖技术,首先要进行错误检测。对于基于块的编码方式,要恢复一个丢失的宏块,首先需知道该宏块的编码模式。另外,无论是intra-块还是inter-块,纹理信息都是必须要恢复的。

对于 inter-块,还需要恢复运动向量(MV)。

3.1 错误检测

错误检测分为基于信道的错误检测和基于解码端的错误检测^[20]。增加包头信息是在一种传输层进行的常用的错误检测方法。在基于包的传输中,编码器的输出数据被打成数据包,每个数据包又包括包头和有效载荷。包头包含有包序号并按照数据包传输的顺序进行编号。在接收端,包序号便可用来进行丢包检测。前向纠错技术是另一种错误检测的方法。编码器输出的码流分段使用错误校验编码,在解码端,错误校验解码进行错误检测并尽可能修正位错误。

3.2 编码模式的恢复

通常,对重要信息如编码模式需添加特别的保护措施,即便这样,该信息有时也可能丢失或受损,估计受损 MB 编码模式的方法是统计与该 MB 相邻的 MB 的编码模式,从中挑选一个最相似的模式。

3.3 运动向量的恢复

对于运动向量 MV 的估计,有以下 5 种方法:在运动较小的情况下认定 MV 为 0;使用前一帧对应块的 MV;使用空间相邻块的 MV;使用空间相邻块的 MV 的中值;重新估计 MV。

当一个宏块受损时,由于打包的原因,其水平方向的相邻宏块一般也受损,因此只能从其上一行或下一行的对应宏块估计,如图 5(b)所示。文献[21]、[22]表明重新估计 MV 的方法可以得到最好的重建质量。文献[23]、[24]认为与给整个 MB 使用一个 MV 相比,在同一 MB 内的不同像素区域使用不同的 MV 会得到更好的重建结果。

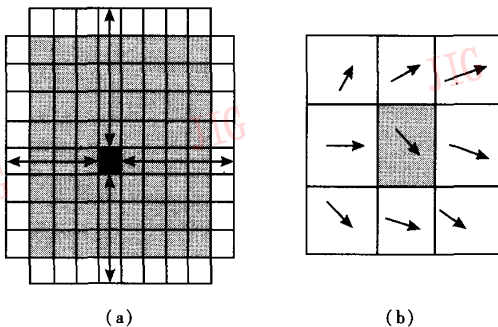


图 5 运动向量的恢复

Fig. 5 Recovery of motion vector

3.4 纹理信息的恢复

纹理信息反应了图像的细节特征,它是构成图

像的基本成份,因此对于纹理信息恢复的好坏直接决定了整个宏块的恢复质量。其恢复方法主要是利用人眼对高频信息不敏感的特点,包括空域插值、时域恢复、频域恢复及其结合。

4 基于传输层的差错控制

传统的传输协议其设计初衷是用来传输非实时数据的,如电子邮件与文件传输等,这导致在目前的“尽力而为”的网络上,传统协议不能很好地处理视频的传输。基于传输层的差错控制技术主要有自动重传、FEC 及 RTP/RTCP 协议等。

4.1 自动重传 (ARQ)

面对传输过程中产生的错误,ARQ 是传统的做法,其思想是重新发送丢失的数据。ARQ 技术由于需要重新发送分组,从而带来大量的冗余数据。网络丢包率太高时,重传不仅不能解决丢包问题,由于循环重传而会导致网络性能的急剧下降,从而带来更大的传输延迟。有选择性地重传可以在一定程度上降低延迟,可以作为一般应用下的视频传输方案。

4.2 前向纠错码 (FEC)

FEC 是经典的信道编码方法,其原理是在原有的数据流中添加冗余信息^[24]。即第 N 个包中不仅包括码流信息还包括第 $N-1$ 包的冗余信息,如果第 $N-1$ 包丢失,则根据第 N 包中的冗余信息恢复丢失的数据。当错误率保持在一定限度时,接收端能利用冗余信息重新构造出丢失的数据。与 ARQ 相比,FEC 具有较小的延迟,因而更适合视频流的实时传输。

4.3 实时传输协议 RTP/RTCP

基于多媒体实时传输的特殊需求,IETF(互联网工程任务组)定义了实时传输多媒体数据的标准协议 RTP/RTCP,RTP/RTCP 运行在 UDP 协议层之上,是非常薄的一层协议。在 RTP/RTCP 体系下,编码器生成的视频流依次经过 RTP 层、UDP 层进入 IP 网络(如图 6 所示)。在接收端,经历逆过程然后送交视频解码器进行解码。接收端的 QoS 监视器根据接收状况判断当前的网络状况(如丢包率和延迟等),反馈控制模块对这些信息进行分析后将该信息反馈至发送端;发送端的位率控制模块根据此信息判断当前的网络状态并自适应调整编码器的输出位率,使之与当前的可用带宽相匹配。

除此之外,还有几个专门面向多媒体的传输协

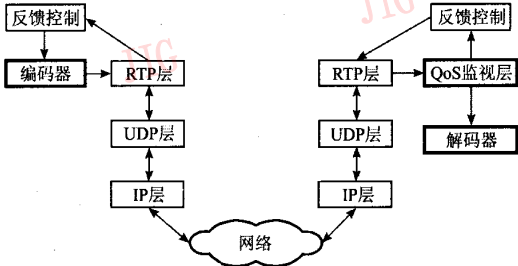


图 6 RTP/RTCP 传输体系结构
Fig. 6 RTP/RTCP transmission architecture

议。资源预留协议 (RSVP) 对异构的接收端有较好的适应性并且与现有的网络协议 IPv4 及下一代协议 IPv6 具有较好的兼容性,但 RSVP 不能提供一种机制来防止用户提出超越本身需求的带宽;用户身份验证也是其薄弱环节,因此很难在现有网络上使用,被认为是未来的网络服务协议。此外还有 ITU 专门为视频会议制定的标准 H. 323,它适用于对等网络、有交互性的多媒体传输。

5 H. 264/AVC 的容错工具

最新的视频编码标准 H. 264/AVC 是联合视频组 (Joint Video Team) 制定的,于 2003 年 3 月正式成为国际标准^[6]。H. 264/AVC 采用了一系列最新的技术,如块大小自适应的运动估计、整数块变换、改进的循环滤波以及高效率的熵编码等,这些技术使得 H. 264/AVC 标准能在重建图像质量相同的前提下比 H. 263 或 MPEG-4 降低大约 40% ~ 50% 的码流量,或者可以在编码码率相同的情况下重建图像质量平均有 3dB 的提高^[25]。

需要特别指出的是,H. 264/AVC 明确地把视频流的容错性、对带宽的适应性等需求作为与编码效率同等重要的因素加以考虑^[6]。H. 264/AVC 的容错手段^[26]有:SP/SI 帧编码、多参考帧、FMO、ASO、RS、Intra-MB 及数据分割等。

5.1 SP/SI 帧

现有的视频编码标准 H. 263, MPEG-2, MPEG-4 主要定义了 3 种类型的帧。I-帧主要消除空域冗余,P-帧消除时域冗余,B-帧使用双向参考帧并提供更高的压缩比。

P-帧、B-帧在消除时域冗余的同时为压缩码流引入了时域相关性,这就使得压缩码流对于传输过程中的错误非常敏感。为此,H. 264/AVC 定义了

SP-帧,SI-帧,其主要特点就是为同一帧使用不同的参考帧编码,但其重建结果相同。实验结果^[27]表明,SP-帧用近似于 P 帧的编码效率达到了接近 I 帧的效果。

SP/SI 帧最先提出是为了解决码流切换时所遇到的问题,但它在容错编码与传输中也有很好的应用前景。如图 7(a) 所示,假设图中打叉的 P 帧在传输过程中丢失,根据传统的编码方法,则后续的所有 P 帧都将因参考帧丢失而无法解码。SP/SI 帧的思想是在码流中周期性的插入一些 SP/SI 帧,如图中 S 帧(主 SP 帧),如果没有传输错误发生,则传输顺序为 S、P、P、P、S、P、P。若图中 P 帧出错,其后的 S 帧将无法解码,此时根据网络反馈的信息,发送端在得知出错信息后发送 SP 帧(辅 SP 帧),该 SP 帧的参考帧为第 1 个 S 帧,因而可以正常解码。但是该辅 SP 帧要与图中其对齐的主 SP 帧具有相同的重建结果才能被后续帧用作参考帧。

如图 7(b) 所示,假设 S_1 使用 P_1 为参考帧, S_2 使用 S 做参考帧,但 S_1 、 S_2 的重建结果要相同。

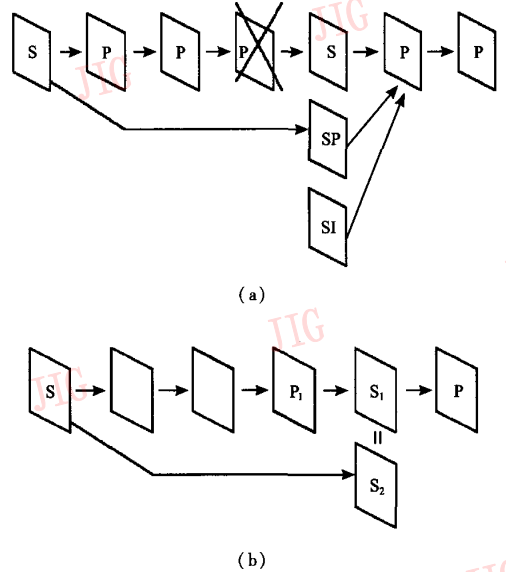


图 7 SP/SI 帧用于容错
Fig. 7 SP/SI used for error resilience

5.2 多参考帧

基于多参考帧的运动补偿既可以大幅提高运动估计的性能,也解决了使用单一参考帧时的遮挡问题,同时由于使用了多个参考帧也使得视频流具有更好的容错性能,如图 8 所示。基于多参考帧的运动补偿带来的问题是编码与解码的计算量增大,同

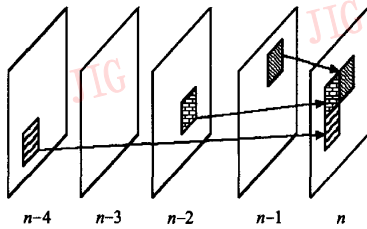
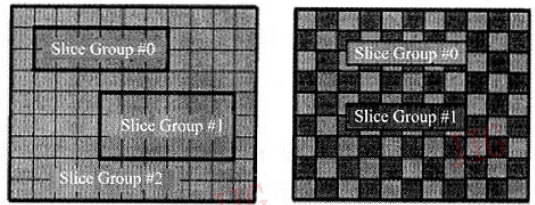


图 8 基于多参考帧的运动估计

Fig. 8 Multiple reference frame based motion estimation



(a) (b)

图 10 使用 FMO 后 Slices 的划分

Fig. 10 Partition of a picture with FMO

时也因参考帧的增多而使编码时需要更多的存储资源。

5.3 灵活的宏块排列

在传统的视频编码标准中一幅图一般都是被分成大小固定的宏块,并且在这些标准中解码过程的定义都是以宏块为基本单位,而 Slices 则是按扫描顺序排列的宏块序列。在这种情况下,如果不使用灵活的宏块排列(FMO),一幅图一般被切分成如图 9 所示的一个或几个 Slices。

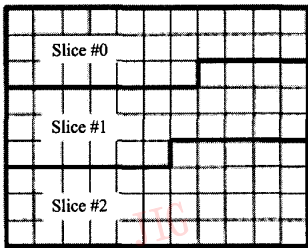


图 9 不使用 FMO 时一幅图的划分

Fig. 9 Partition of a picture without FMO

在 H. 264/AVC 中, Slice 是自包含的,即每个 Slice 都包含完整的参数集并存放在压缩码流中,在解码端每个 Slice 可单独解码而无需其他 Slices 的信息。通过引入 Slice 组的概念, FMO 改变了以往对于图的划分方法。每个 Slice 组就是一组宏块的集合,它通过 Slice 头信息中的参数集定义了每个宏块到 Slice 组的映射表。映射表记录了每个宏块在图中的位置以及该宏块属于哪个 Slice 组。

FMO 的思想是,在一幅图内,宏块的扫描模式可以有多种,例如可以是交叉排列的 Slices,其中宏块可以随意安排;可以有一个或多个“前景” Slice 组,其余的部分作为一个 Slice 组,图 10(a)所示,这种安排方式其实可以作为感兴趣区域的编码应用;也可以是像棋盘状的排列方式,图 10(b)所示,整幅

图被分成两个 Slice 组,白色的为 Slice Group#0,阴影部分为 Slice Group#1,这样传输时即便有一个包丢失,通过插值的方式进行差错掩盖仍然可以得到接受的解码图像。实验结果表明,对于 CIF 格式图像,在丢包率为 10% 的网络中,通过 FMO 方法并经过差错掩盖后,可使专业判读员感觉不到视频质量有降低^[28]。但 FMO 带来的缺点是它打乱了宏块的顺序从而使帧内预测的有效性大为降低。另外,额外的宏块重排的开销也使得处理延迟增大。

5.4 任意的 Slices 排列

任意的 Slices 排列 (arbitrary slices ordering, ASO) 使得每个 Slice 都能独立解码,从而可以在网络发送和接收时不必以传统的扫描顺序进行,而可以按任意的 Slice 顺序发送。ASO 带来灵活性的同时也如 FMO 一样会增大端到端的传输延迟。

5.5 冗余 Slices

为了提高视频流传输的鲁棒性, H. 264/AVC 允许编码器发送某一 Slice 的冗余编码数据 (RS) 以处理传输中出现的丢失。一般来讲, RS 与该 Slice 的原始编码方式应该有所不同,这样在 primary slice 丢失后,利用 RS 同样能恢复出重建值。

H. 264/AVC 的其他容错工具还有 Intra-宏块更新、数据分割等。Intra-宏块更新在以往的标准中已经被采纳,主要用于与信源信道联合编码结合使用,它通过获得网络的反馈信息动态决定 Intra-宏块更新的频率和位置。数据分割在以前的编码标准也已使用,但 H. 264/AVC 在码流中明确定义了数据分割的方法并把数据分割为 DPA、DPB 和 DPC 3 部分,同时也为这 3 类数据分配了不同的存储缓冲区。文献[29]的实验结果表明面对不同丢包率的网络,信道自适应的 Intra-宏块更新、FMO 以及多参考帧在容错效果方面具有明显的优势。

6 关于容错视频编码与传输技术的思考

在不可靠信道上进行视频传输特别是视频的实时传输是一项挑战性很强的工作,要减小传输错误对解码端视频带来的影响,必须在视频编码、网络传输以及解码端进行错误控制。研究表明,这些容错工具在面对传输错误时还是非常有效的,对于解码端的视频回放无论是客观评价还是主观评价,质量都有提高。对于如何选择合适的容错工具以达到最好的容错效果,一方面要看具体的应用对于哪些约束比较敏感,比如有些应用对添加冗余后数据量的增加敏感,而有些应用对于传输延迟敏感等等;另一方面,把不同的容错工具组合使用,特别是通过实时探测信道状况,进而实现编码器、解码器交互的差错控制,可以得到更好的效果。

当前关于容错视频编码与传输技术的发展方向及研究热点^[26,29]如下:

(1) 研究基于优化率失真的信源信道联合编码传统的信源编码和信道编码都是独立进行的,并没有考虑两者共同对视频传输带来的影响。对于基于块的编码,一个宏块的编码模式有 Intra/Inter

两种, Intra 宏块的编码效率较低,但遇到网络错误时不受其他帧(宏块)的影响,即网络容错性能好; Inter 宏块通过运动估计可大大提高编码效率,但要正确解码时必须依靠其参考帧,因此对于有丢失的网络信道失真较大。

通常,信源失真是在编码的量化阶段引入的,信道失真则是由传输错误造成的^[30]。信源信道联合编码(joint source channel coding)的思想是根据全局的率失真模型(global rate-distortion model),参考信道的反馈信息,动态地调整编码端 Intra/Inter 宏块的比例,从而获得信源与信道联合失真最小条件下的编码决策准则,使得接收端的效果达到最佳。

(2) 研究网络自适应的多维 FMO 容错编码算法

FMO 是一种允许编码器以传统扫描顺序以外的其他顺序将宏块划分到条带中的宏块组织方式。图 11 表示了几种常用的 FMO 模式^[31]: 图 11(a)为默认模式: 一帧全部宏块被划分到同一个条带; 图 11(b)为固定模式: 固定数目的宏块,一般为行宏块数的整数倍被划分到一个条带; 图 11(c)为交织模式: 偶数行的宏块被分到一个条带,奇数行宏块被分到另一个条带; 图 11(d)为分散模式: 棋盘式的宏块划分。

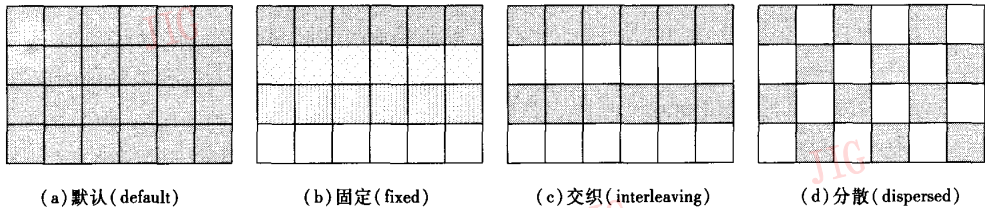


图 11 使用 FMO 前、后 Slices 的划分

Fig. 11 Comparison of a picture with/without FMO

不同的宏块组织方式可导致不同的编码效率和容错能力。现有的帧内预测规定预测不应超出条带范围。如果将一些相邻的宏块分配到同一条带内,则可以利用空间相关性进行帧内预测,可获得较高的编码效率;如果按 FMO 的思想重新排列宏块时,其空间相关性被打破,因此编码效率会有所降低。

一旦发生传输丢包的情况,前者的宏块组织方式可能导致解码器接收到的视频出现连续丢失区域,难以准确恢复;而后的宏块组织方式,未接收到的丢失区域不连续,对差错掩盖极为有利。因此针对不同的网络选用何种宏块组织方法以达到编码

效率和容错性能的最佳平衡应该是 FMO 考虑的重点。在重排空间相邻的宏块顺序的同时,重新排列时间相邻的两帧内宏块的顺序,也即采用多维 FMO 会带来更好的容错效果。

(3) 研究基于信道反馈的多参考帧编码

使用多参考帧编码可以显著提高预测的匹配度,减小残差,因而可以提高编码效率。此外,使用多参考帧也具有较好的容错性能。实验结果表明,使用多参考帧进行运动估计可以有效地控制错误的扩散^[29]。如果再建立一个反馈信道,接收端便可把丢失的数据包序号返回给编码器,编码器则从缓冲

区内找出正确的参考帧重新发送至接收端。要完成这个过程,假设网络的往返时间(RTT)为 T ,在 T 时间内传输的帧数为 D ,那么编码端需要缓存的总帧数为 D ,这也意味着接收端将带来 D -帧的延迟。

文献[31]认为,当反馈信息从接收端传输至编码端时,其延迟已大大超过了相邻两帧之间的时间间隔,即便缓冲区内有多帧最近解码的参考帧,也有可能因延迟过大而丧失作为参考帧的条件,因此建议增加缓存帧的数量。缓存帧具体数量以及参考帧的选择准则也成为有意义的研究课题。

(4) 研究基于解码端的差错掩盖及后处理算法

传统的差错掩盖算法对 Intra 编码帧仅仅采用空间掩盖算法,而对 Inter 编码帧仅仅采用时间掩盖算法。实验结果表明,当视频序列内容变化平缓时,邻近帧的时间相关性十分明显,对 Intra 编码帧采用时间掩盖算法效果更好;而当 Inter 编码帧出现场景切换的内容时,原有的时间相关性被打破,此时采用空间掩盖算法更为适用。总之,无论是 Intra 还是 Inter 编码帧,自适应地选择时间或空间 EC 算法可以达到更好的效果。

因此,采取场景变换检测、运动活跃性检测和运动矢量恢复模块分级执行的策略,以决定是采用时间相关性信息还是空间相关性信息来进行差错掩盖,是降低计算复杂度的有效途径。若出现场景切换,则只可以采用空域掩盖算法;否则,执行视频序列内容运动活跃性检测,若活跃性较低,则直接拷贝前帧对应位置宏块来替换丢失位置宏块;否则,采用经典时域 EC 算法恢复运动矢量,重建丢失宏块。

(5) 研究设计适应异构网络的应用层视频传输协议

端到端的网络带宽评估是自适应视频流传输调整发送码率的基础,其评估分为两个环节,一是开始传输之前的初始带宽评估,二是在传输过程中跟踪带宽的动态变化。网络状况发生了变化,编码器也必须调整参数,以适应网络带宽的变化。

网络发送缓冲区是编码器和网络发送模块之间的一块缓冲区域,其大小直接影响到端到端传输的时延、抖动及数据的丢失控制。如果缓冲区设置得太大,端到端时延会增大,影响传输的实时性;相反,如果缓冲区设置得太小,缓冲区溢出的风险将增加。因此研究网络发送缓冲区控制技术十分重要。采用基于网络发送缓冲区充满度的位率控制机制,为缓冲区设置二级阈值,适时地通过缓冲区的充满度为

编码器提供预警信号,从而实现一个稳升快降的位率控制算法。

因此,综合考虑不同终端用户网络的异构性,设计能够自适应网络的位率控制机制,研究基于 RTP/RTCP/UDP 协议的、应用级的容错视频传输协议也是当前的具体应用中十分典型和有意义的研究课题。

参考文献 (References)

- 1 Wang Yao, Wenger Stephan, Wen Jiangtao, *et al.* Error resilient video coding techniques—Real-time video communications over unreliable networks [J]. IEEE Signal Processing Magazine, 2000, 17(4): 61 ~ 82.
- 2 Cote G, Shirani Shahram, Kossentini Faouzi. Optimal mode selection and synchronization for robust video communications over error-prone networks [J]. IEEE Journal on Selected Areas in Communications, 2000, 18(6): 952 ~ 965.
- 3 Moccagatta Iole, Soudagar Salma, Liang Jie, *et al.* Error-resilient coding in JPEG-2000 and MPEG-4 [J]. IEEE Journal on Selected Areas in Communications, 2000, 18(6): 899 ~ 914.
- 4 Talluri Raj. Error-resilient video coding in the ISO MPEG-4 standard [J]. IEEE Communications Magazine, 1998, 36(6): 112 ~ 119.
- 5 Dogan S, Sadka A H, Kondoz A M. Error-resilient techniques for video transmission over wireless channels [J]. The Arabian Journal for Science and Engineering, 1999, 24(2): 101 ~ 114.
- 6 Wiegand Thomas, Sullivan Gary J, Bjontegaard Gisle, *et al.* Overview of the H.264/AVC video coding standard [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 560 ~ 576.
- 7 Karczewicz M, Kurceren R A. Proposal for SP-Frames [S]. In ITU-T Video Coding Experts Group Meeting, Eibsee, Germany, 2001, Doc. VCEG-L-27
- 8 Wen G, Vilasenor J. A class of reversible variable length codes for robust image and video coding [A]. In: Proceedings of IEEE International Conference on Image Processing [C], Santa Barbara, CA, 1997: 65 ~ 68.
- 9 Ghanbri M. Two-layer coding of video signals for VBR networks [J]. IEEE Journal on Selected Areas in Communications, 1989, 7(5): 771 ~ 781.
- 10 Haskell P, Messerschmitt D. Resynchronization of motion compensated video affected by ATM cell loss [A]. In: Proceedings of International Conference on Acoustics, Speech and Signal Processing [C], San Francisco, CA, 1992: 545 ~ 548.
- 11 Gallant Michael, Kossentini Faouzi. Rate-distortion optimal joint source/channel coding for robust and efficient low bit rate packet video communications [A]. In: Proceedings of IEEE International Conference on Image Processing [C], Vancouver, BC, Canada, 2000: 355 ~ 358.
- 12 Wu Feng, Li Shipeng. Efficient and universal scalable video coding

- [A]. In: Proceedings of IEEE International Conference on Image Processing[C], Rechester, New York, 2002: II-37 ~ II-40.
- 13 He Yunwen, Zhao Xuejun, Zhong Yuzhuo, *et al.* Improved fine granular scalable coding with inter-layer prediction [A]. In: Proceedings of the Data Compression Conference [C], Snowbird, Utah, 2002: 172 ~ 181.
- 14 Gallant Michael, Kossentini Faouzi. Rate-distortion optimized layered coding with unequal error protection for robust internet video [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 11(3): 357 ~ 372.
- 15 Wang Yao, Panwar Shivendra, Lin Shunan, *et al.* Wireless video transport using path diversity: multiple description VS. layered coding [A]. In: Proceedings of IEEE International Conference on Image Processing [C], Rechester, New York, 2002: I-21 ~ I-24.
- 16 Begen Ali C, Altunbasak Yucel, Ergun Ozlem. Real-time multiple description and layered encoded video streaming with optimal deverse routing [A]. In: Proceedings of the 8th IEEE International Symposium on Computers and Communications [C], Kemer-Antalya, Turkey, 2003: 887 ~ 892.
- 17 Apostolopoulos J G. Reliable video communication over lossy packet networks using multiple state encoding and path diversity [A]. In: Proceedings of SPIE Conference on Visual Communications and Image Processing [C], San Jose, 2001: 392 ~ 409.
- 18 Lin S, Mao S, Wang Y, *et al.* A reference picture selection scheme for video transmission over ad-hoc networks using multiple paths [A]. In: Proceedings of IEEE International Conference on Multimedia and Expo [C], Tokyo, Japan, 2001: 96 ~ 99.
- 19 Gogate Nitin, Chung Doo Man, Panwar Shivendra S, *et al.* Supporting image and video applitions in a multihop radio environment using path diversity and multiple description coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(9): 777 ~ 792.
- 20 Wang Yao, Zhu Qin-fan. Error control and concealment for video communication-a review [J]. Proceedings of the IEEE, 1998, 86(5): 974 ~ 997.
- 21 Hong M C, Kondi L, Scwab H, *et al.* Video error concealment techniques [J]. Signal Processing: Image communications, 1999, 14(6-8): 437 ~ 492.
- 22 Zeng Wenjun, Liu Bede. Geometric-structure-based error concealment with novel applications in block-based low bit rate coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(4): 648 ~ 665.
- 23 Lakshman T V, Ortega Antonio, Reibman Amy. VBR video: tradeoffs and potentials [A]. In: Proceedings of the IEEE [C], 1998, 86(5): 952 ~ 973.
- 24 Seung Han Jung, Kim Rin-Chul, Lee Sang-Uk. A hierarchical synchronization technique based on the EREC for robust transmission of H.263 bit stream [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2000, 10(3): 433 ~ 438.
- 25 Bo Hong. Introduction to H.264 [EB/OL]. <http://www.h2631.com/h264/www.utdallas.edu/bhong.h264.pdf>, 2002-11-22.
- 26 Sullivan Gary J, Topiwala Pankaj, Luthra Ajay. The H.264/AVC advanced video coding standard: overview and introduction to the fidelity range extension [A]. In: Proceedings of the SPIE Conference on Applications of Digital Image Processing [C], Denver, CO, 2004: 53 ~ 74.
- 27 Karczewicz Marta, Kruceren Ragip. The SP and SI frames design for H.264/AVC [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 637 ~ 644.
- 28 Wenger Stephen. H.264/AVC Over IP [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 645 ~ 656.
- 29 Stockhammer Thomas, Wiegand Thomas, Oelbaum Tobias, *et al.* Video coding and transport layer technique for H.264/AVC-Based Transmission over Packet-lossy Networks [A]. In: Proceedings of IEEE International Conference on Image Processing [C], Barcelona, Spain, 2003: 481 ~ 484.
- 30 He Zhi-hai, Cai Jian-fei, Chen Chang-wen. Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(6): 511 ~ 523.
- 31 AVS M1322-2004 A few topics on reference frame [S]. Tsingdao, China