

Internet 视频通信差错隐藏技术研究*

Review of Error Concealment Technologies in Internet Video Communications

胡 飞 朱光喜 朱耀庭
(华中科技大学电信系 武汉430074)

Abstract This paper reviews Internet-based video communications and relevant error concealment techniques. By analyzing the essential characters of Internet and video communications and listing the main causes that packet loss is inevitable in Internet, the necessity of error resilient coding for video communications over Internet is demonstrated. Then error detect technologies, error concealment technologies at decoder end and those by coder and decoder interacting proposed recently are classifiedly reviewed.

Keywords Internet, Video communication, Error concealment, Error control

1 引言

TCP/IP 协议使得 Internet 能够建立于结构简单、成本较低的通信设施基础之上,具有广泛包容多种通信链路、多种网络技术和多种终端设备的能力,能够提供更加灵活的服务模式,可以在同一通信网络上提供多种应用,因此具有带宽利用率高、支持多种业务、操作灵活、易于管理、成本低等特点。然而目前 Internet 只提供尽力服务(Best-effort)的服务模式,不提供服务质量(QoS)保证。因此在 Internet 中,时延及时延抖动以及数据包丢失等传输差错不可避免,是 Internet 通信协议及其应用必须解决的重要问题之一^[1]。

在 Internet 迅速发展和普及的同时,近年来视频通信技术的研究和应用也取得了巨大发展^[1~3]。随着网络带宽增长,通讯成本下降,视频通信逐渐成为 Internet 的一类重要应用,如计算机支持的协同工作(CSCW)、远程教育、实时监控、远程诊断、可视电话、视频会议、视频点播(VOD)、视频流化播放、新闻视频浏览等,这些应用具有普遍的应用需求和广泛的应用前景。

然而 Internet 的 Best-effort 服务模式与视频通信对传输网络的 QoS 要求存在本质上的矛盾,成为阻碍视频通信应用进一步发展的重要因素。这一矛盾也是视频通信领域的一个重要研究内容,目前国内外研究者已经提出大量解决或缓解该矛盾的方法,如信道可靠性控制^[4~6]、容错编码^[7~9]、差错隐藏等。

其中信道可靠性控制只单纯地从提高信道传输质量出发,没有考虑到传输内容的特性和要求;容错编码通过对视频信息编码时增加一定的容错信息,降低传输差错对视频信息的解码和回放质量的影响。冗余性的容错信息要占用一定的带宽,显然降低了网络资源的利用率。另外由于不同的应用要求和网络状态的变化,容错编码依然无法完全消除传输差错的影响。这两类方法虽然具有一定的独立性,但在网络资源利用效率和差错控制效果上很难实现全局优化。而差错隐藏通过对多媒体流中残留或预留的冗余信息进行进一步挖掘实现差错控制。这类方法应用灵活,可以在解码端独立进行,也可通过编/解码交互进行,是目前 Internet 视频通信中广泛使

用的差错控制方法。

本文分析了 Internet 视频通信的特点,并对近年来出现的 Internet 视频通信差错隐藏技术进行了分类概述,并对差错检测技术进行了描述。

2 Internet 中的传输差错

目前 Internet 只提供 Best-effort 服务模式,不提供 QoS 保证。传输差错是 Internet 通信协议及其应用必须解决的重要问题之一。对于 Internet 视频通信而言,传输差错主要表现为数据包丢失。导致 Internet 传输中数据包丢失的主要原因有:

1)应用的流量突发性。Internet 共享带宽的工作模式使突发流量的出现不可避免,而有限的路由器缓冲区不可能完全吸收流量的突发性,因此数据包会溢出交换节点的缓冲区,导致数据包丢失。

2)拥塞和丢失是 Internet 设计思想的一部分^[1]。在尽力服务模式下,网络不向应用确定地分配带宽等资源,而是必须由应用程序本身去动态地发掘可用资源。TCP 就是通过故意地连续增加流量直至发生丢失再降低流量的方式去发现最多可用资源。

3)存储转发机制引入了排队时延和时延抖动的问题。对于具有严格时间要求的视频流,过大的时延导致数据包在回放时刻后到达而无效,最终表现为数据包丢失。

4)对于近年来出现的群组通信(Multicast)应用,问题变得更为复杂。在群组通信应用中,数据可能由一个发送者发往成千上万个接收者,从而经过各种各样的链路和网络,因而出现数据包丢失的概率要大得多。

5)必须承认,网络技术在飞速发展,目前我们正快速进入光通信时代,巨大容量的通信链路正成为现实。或许链路容量的增加将消除主干网络上绝大多数的数据包丢失现象,但由于 Internet 网的链路异构性和网络技术异构性,如“最后一公里”的接入部分、无线通信链路等问题依然存在,有时丢失主要原因不仅仅是链路容量问题,还包括天气等其它影响因素,因此无法避免数据包丢失。

6)通信链路带宽不断增长的同时,应用的带宽需求也会

*)本课题得到国家教育部重点科学技术项目基金(项目编号2000175)、国家科技部技术创新基金(项目编号00C26224210641)资助。胡 飞 博士研究生,主要研究方向为包交换网络多媒体通信差错控制技术。朱光喜 教授,博士生导师,主要研究领域为多媒体信号处理及信息网络技术。朱耀庭 教授,博士生导师,主要研究领域为通信与电子系统技术。

相应地不断增加。例如,目前只有低比特率视频通信能够通过 Internet 传输,而且只能提供勉强可以接受的视频质量。随着高清晰度视频技术的发展与人们对视频质量的要求不断提高,这类应用必然要求更高的传输速率。因此,传输网络提供的带宽与应用的需求之间的矛盾将继续存在,致使数据包丢失不可避免。

7) 由于 Internet 网络端设备的异构性,发送/接收设备的速度、CPU 计算能力、存储容量等存在不匹配问题,在应用层依然会出现数据包丢失。这种情况对于如手持设备、移动设备等尤为突出。

总之,传输差错是 Internet 通信中不可避免的现象,是由 Internet 建立在简单、经济的架构上、提供灵活的服务模式、允许同时提供多种应用、具有较高的网络资源利用率等设计思想和技术本身决定的,是上述一系列优点的副产物,也是网络协议和网络应用必须考虑的重要问题之一。

3 视频通信^[2-4]

视频信息包含信息量大,内容丰富,是重要的一类多媒体信息。与电子邮件、文件传输等其它网络应用相比,视频通信具有独特的特征:1) 视频信息所要求的传输带宽较高,从数十 Kbps 到数十 Mbps 不等;2) 视频信息具有周期性(如要求每秒 10~30 帧的传输速率),传输持续时间长(几分钟到数小时不等);3) 视频流具有严格的实时传输要求,当数据包的到达由于传输时延超过回放时刻,该包无效,等同于丢失;4) 视频信息往往经过压缩处理,压缩后的视频流具有变比特(VBR)性,其流量突发性强;5) 压缩后的视频信息由于存在视频帧之间的依赖性,对数据包丢失敏感,即一帧的数据包丢失不仅影响该帧的正确解压,还将影响其后的视频帧的正确解压。图1描述了数据包丢失对采用 H. 263 协议^[10]压缩的视频流的影响。从图中可以看出,当数据包丢失率达到 5% 时,已经可以明显感觉到视频回放质量的失真;当数据包丢失率达到 10% 时,视频回放质量已经令人无法接受;6) 另一方面,由于人的生理特征,视频通信能够容忍一定的包丢失率。如图 1 中包丢失率为 3% 时,几乎看不出数据包丢失对视频回放质量的影响。



上左:无数据包丢失;上右:3%的包丢失;

下左:5%的包丢失;下右:10%的包丢失

图1 传输差错对采用 H. 263 协议压缩的视频流的影响

4 差错检测

实施任何差错控制技术的前提是检测到是否发生了传输差错以及对差错发生位置准确定位。因此差错控制的第一步是差错检测。差错检测技术可以分为两类:传输层差错检测和视频解码端差错检测。

4.1 传输层差错检测

在基于数据包的视频传输系统中,传输层差错检测最常用的一种方法是利用数据包的头部信息。视频发送端对视频编码器输出的压缩视频流打包时,在每个数据包前加上头部信息和载荷信息。其中头部信息包含着一个连续的数据包序号。当数据包丢失发生时,连续到达传输层解码器的数据包的序号将不再保持连续,接收端传输层解码器从而可以判断数据包丢失的发生。这种差错检测方法也叫做基于失序的差错检测方法,普遍应用于基于包交换网络的视频传输系统中。如广泛用于视频会议的复用协议 H. 223^[11]就采用了这种差错检测方法。

传输层采用的另一种差错检测方法是 FEC 技术。这种方法对视频编码器输出的压缩视频流分段进行纠错编码,即为每个数据段提供纠错字段,传输层解码器根据纠错字段判断差错的发生及其发生位置。FEC 方法通常工作在比特层,即发现或纠正比特差错。如 H. 223 协议就采用 FEC 对复用的数据包的头部信息和载荷信息进行差错检测。H. 263^[10]协议也为每 493 比特的视频流采用 18 比特的 FEC 字段。

4.2 解码端差错检测

解码端差错检测一般工作在应用层。视频解码端进行的差错检测利用了视频信号的自然特性。一种方法是利用相邻行的像素之间的差值来对用脉码调制(PCM)或差分脉码调制(DPCM)编码的视频流中的差错进行检测^[12]。当差值大于某个设定的门限时,可以认为出现了传输差错。另一种方法^[13]是通过检测一个块的边界像素与其相邻的四个块的边界像素 DCT 系数之间的差值来进行差错检测:解码器通过取当前块与其相邻块在四个方向上一个像素宽度的差值形成四个向量,然后对这些向量进行一维 DCT 变换。如果这些相邻的块是平滑变化而且视频传输过程中没有差错发生,那么变换所得的结果应该足够小。如果变换所得向量包含了较大的系数,说明有传输差错发生,而且差错发生位置可以根据系数位置进行确定。

解码端差错检测还可以工作在频域^[14]。这种方法在每个块的扫描行尾部插入一个同步码字。当解码器取得这个同步码字时,解码器检查已经解码的块的数量并与同步码字相比较,如果结果不一致则意味着差错的发生。

当视频编码器采用 VLC 技术编码时,视频流中任何一个比特的错误将导致其后的所有比特无法正确解压,这种现象称为失同步。这种特性也可用来进行差错检测。当视频解码器发现了不属于其码表的码字时,可以判断出传输差错的发生。另外,视频流所采用的语法也可以用来实现差错检测,比如,如果解码器解码结果中包含了值为 0 的量化步长,或者解码所得的 DCT 系数多于最多可能的数目(例如,8×8 的块 DCT 变换所得的系数最多为 64 个),都可以作为判断差错发生的依据。

基于数据包头部信息和 FEC 的两种方法工作在传输层,差错检测可靠性高,容易实现,但需要消耗一定的额外带宽。工作于视频解码端的差错检测则不需要分配额外的比特,而是利用视频信号的平滑特性进行差错检测。但这类方法可靠性低,容易发生误判,而且实现起来非常复杂,可能消耗计算机 CPU 资源并产生一定的时延。使用同步码字的方法则是上述两种方法的折衷:通过在编码的过程中保留较低的冗余信息,使得解码端差错检测更加容易。显然,这些方法并不互相排斥,而是可以根据需要和使用场合相互结合使用。

5 解码端差错隐藏

尽管经过信道编码保护,视频流经过有丢失信道到达接收端时仍可能有残余丢失。解码端差错隐藏是指在解码端进行的对丢失信息的恢复或估计。如果采用基于块的复合视频编码方法,要正确解压一个损坏了的MB,需要对三种信息进行估计:1)纹理信息,包括原始图像的块或差分块的像素值或DCT系数;2)运动信息,主要指MB的运动向量MV;3)MB的编码模式。

众所周知,自然场景的图像或视频中低频分量是主要成分,即除了边缘信息,在时间和空间上相邻的像素值的变化往往是非常平滑的。因此几乎所有用来恢复受损图像纹理信息的技术都利用图像或视频的平滑特性,普遍采用时间或空间上的插值方法对丢失的纹理信息进行估计。图像或视频的平滑特性还反映在运动向量MV的平滑变化上,因此插值方法还可以用来恢复MV。至于编码模式信息的恢复则普遍采用启发式方法。

5.1 纹理信息恢复

运动补偿时间预测:恢复受损的MB的一个简单但非常有效的方法是根据MB的运动信息用前一解压帧中对应的MB代替当前受损的MB。这种方法的恢复性能关键在于MV是可知的。如果MV也丢失了,那么在进行基于运动补偿时间预测的MB恢复之前必须首先对MV进行估计。

空间插值:另一种纹理信息恢复方法是空间插值:用正确收到的相邻的块插值估计受损块的像素值。通常一幅图像中同一行的所有块或MB被装载在一个数据包里传输,往往同时丢失。因此一般用受损块上面或下面相邻的块来进行插值。由于相邻块中大部分像素值是不同于丢失的块的像素值的,插值一般只利用相邻块的边界信息^[15]。

除了直接对像素值进行插值估计,一个更简单的方法是估计受损块的直流分量(DC),并且只用直流分量解压受损块。DC可以通过对周围相邻块的DC取平均进行估计,也可以利用相邻块的边界信息进行估计^[16]。交叉打包机制有助于空间插值差错隐藏:相邻的块分别被装载在不同的数据包里,一个数据包的丢失不会损坏所有其它可利用的相邻信息。

最大平滑效果时空插值:空间插值方法存在的一个问题是如何确定合适的插值模板,而且该方法没有对可能正确收到的部分DCT系数加以利用。一种改进的方法是^[17],同时在空间上利用同一帧内相邻块的像素和在时间上利用前/后一帧内相关的块,对当前受损块丢失的信息进行平滑恢复。在恢复所得的DCT系数与正确收到的部分系数吻合的前提下进行尽可能的平滑。上述方法可以抽象为非受限优化问题。该问题在不同的丢失模式、采用不同的空间、时间甚至频域的插值模板的情况下,对应不同的求解结果。

凸集投影空间插值:凸集投影(POCS, Projection onto Convex Sets)方法^[18]的思想是,将每个未知量各自的限制条件形成一个凸集,所有凸集的交集为问题的最优解。求解通过将过程解反复迭代投影到各个凸集上实现。由于迭代过程需要大量时间,这种方法在视频通信的差错恢复上应用受到限制。

5.2 编码模式和运动向量的恢复

相对于纹理信息,编码模式和运动向量对于视频信息的恢复起着更重要的作用。前文所述的许多差错隐藏方法实现的前提是假设编码模式和运动向量能够正确得到,因此对编

码模式和运动向量的保护更加重要。在H.263和MPEG4标准中,编码器可以选择采用数据划分技术,将编码模式和运动向量独立打包,并在传输中为包含编码模式和运动向量的数据包提供更强的保护。

然而模式信息和运动向量的丢失依然无法完全避免。针对编码模式丢失的一种简单有效的差错隐藏方法是认为该MB采用I模式编码,并采用空间插值方法对该MB进行估计。另一种更好的方法是在相邻块的编码模式可知的前提下,根据相邻块编码模式的统计分布,估计受损MB最有可能的编码模式^[19]。

对运动向量MV的估计也有多种方法^[16,20],包括:1)假定丢失的MV为零向量。这种方法极其简单,而且当视频信息中的运动强度很小时,差错隐藏的效果很好。2)使用前一帧对应块的MV代替当前受损块的MV。3)使用空间相邻块MV的均值代替当前受损块的MV。4)使用空间相邻块MV的中值代替当前受损块的MV。通常情况下,受损MB水平方向相邻的MB的MV也因为受损而不可知,因此3)和4)一般在受损MB上下相邻MB的基础上进行。除了上述方法,还有文[21]提出对MB中不同的像素区进行不同的MV估计,取得了更好的效果。

6 编/解码交互差错隐藏

前文所述的各项差错隐藏技术都是在解码端独立进行的。如果通信链路提供信道允许解码端向编码器提供反馈信息,则可以进行交互式差错隐藏,即编码器根据解码端返回信息对编码参数或编码方法作相应调整,使得解码端的差错隐藏实现更加容易、更加充分,进一步减少或消除传输差错的影响。

基于重传的ARQ就是一种简单的交互差错控制方式,但这种方式由于引入较大的处理时延,在交互视频应用中效果不很理想。相对于传输差错对当前视频帧的影响,差错扩展对视频回放质量的影响更大^[6]。因此,当传输差错出现时,限制差错扩展范围对改善视频回放质量更加重要。

基于反馈信息的参考帧选择:通过反馈信道,编码器可以知道在解码端发生错误的视频帧的信息。根据该信息,编码器在对下一帧进行预测编码时,就可以选择以前的、但解码端已经正确收到的帧作为参考帧,而不采用最近的、在解码端发生错误的视频帧作为参考帧。这种方法称为参考帧选择(RPS: Reference Picture Selection),要求视频编码端和解码端都缓存一定数量的已编码或解码的视频帧。相对于I模式刷新,如果预测距离不是太远并且不发生场景变换,RPS编码效率显然要高。

尽管利用了反馈信息,但RPS不会引入处理时延,因为编码器对当前帧的编码无需等待已编码帧的信息,而是直接选择反馈信息通知出错的视频帧之前的帧作为参考帧即可。但反馈信道时延会影响RPS的工作性能。比如,由于链路时延,解码器端第n帧出错的反馈信息直到编码器对第n+d帧编码时才到达。由于解码器使用发生错误的帧作为解压时的参考帧,因此视频帧n+1到n+d-1都会包含错误。编码器得到反馈信息后,通过选择帧n-1作为对帧n+d编码的参考帧,阻止了传输差错在帧n+d以后的扩展。显然,产生并传输反馈信息所需的时间越长,差错影响的范围就越大;并且由于预测距离增加,编码效率也会相应下降。

基于反馈信息的错误追踪^[22]:RPS以早期的完好帧作为

参考帧对当前帧进行编码。更精确地,编码器可以通过反馈信息追踪第 n 帧中受损区域是如何影响第 $n+1$ 到 $n+d-1$ 帧中各个块的解压的,即知道差错扩展的范围。这样,编码器就可以:1)对帧 $n+d$ 中将用 $n+d-1$ 帧中受损块预测编码的块,采用 I 模式编码。2)避免使用 $n+d-1$ 帧受损的块,而使用未受损的块作预测编码的参考块。3)对帧 $n+1$ 到 $n+d-1$ 使用与解码器相同的错误隐藏技术,这样编码器对第 n 帧编码时使用的参考帧与解码器端解码时使用的参考帧相同。前两种方法只要求编码器追踪受损区域的位置,第三种方法要求编码器对帧 $n+1$ 到 $n+d-1$ 进行与解码器相同的操作,复杂度更高一些。但上述方法都能够阻止传输差错在帧 $n+d$ 以后的扩展。

总结 尽管基于 Internet 的视频通信有着普遍的应用需求和广泛的应用前景,然而 Internet 服务模式与视频通信 QoS 要求之间存在的矛盾阻碍着视频通信应用的进一步发展。对于基于 Internet 的视频通信而言,差错隐藏是目前广泛使用的有效的差错控制方法。

本文对 Internet 视频通信的特点以及近年来国内外出现的主要视频差错隐藏技术进行了概述。然而随着通信技术和信息处理技术的不断发展,新的应用需求将不断涌现;同时差错隐藏作为视频通信领域的一个研究热点,研究成果也不断取得新的进展。要了解最新研究成果,还必须及时跟踪国际最新动态。但总的来讲,Internet 视频通信技术将不断完善,最终得以实现高质量、高可靠性视频通信的普遍应用。

参 考 文 献

- Comer D E. Computer Networks and Internets. Prentice Hall International, Inc., 1997
- Steinmetz R, Nahrstedt K. Multimedia Computing, Communications and Applications. Prentice Hall International Inc., 1995
- Tekalp A M. Digital Video Processing. Prentice Hall International Inc., 1996
- 胡飞,朱耀庭,朱光喜. Internet 视频点播差错控制. 计算机研究与发展, 录用待刊
- Hu Fei, Zhu Guangxi, Zhu Yaoting. Enhanced ARQ-based Packet Loss Recovery for Real-time Communication. International Conference on Info-tech & Info-net (ICII'2001), Beijing, Oct. 2001
- Girod B, Faerber N. Feedback-based error control for mobile video transmission. Proc. IEEE, Special Issue on Video for Mobile Multimedia, 1999, 87(10): 1707~1723
- 胡飞,朱耀庭,朱光喜. 复合嵌入式零树编码低码率容错视频传输. 计算机研究与发展, 录用待刊
- Zhang R, Regunathan S L, Rose K. Video coding with optimal Inter/Intra-mode switching for packet loss resilience. IEEE JSAC, June 2000
- Chung D-M, Wang Y. Multiple description image coding using signal decomposition and reconstruction based on lapped orthogonal transforms. IEEE Trans. on Circuits and Systems for Video Technology, Sept. 1999. 895~908
- Video coding for low bitrate communication. ITU-T Recommendation H. 263 Version 2, Jan. 1998
- Multiplexing protocol for low bitrate multimedia communication. ITU-T Recommendation H. 223, 1995
- Rose K M, Heiman A. Enhancement of one-dimensional variable-length DPM images corrupted by transmission errors. IEEE Trans. Commun., 1998, 37: 373~379
- Mitchel O R, Tabatabai A J. Channel error recovery for transform image coding. IEEE Trans. Commun., Dec. 1981. 1754~1762
- Lam W, Reibman A. An error concealment algorithm for images subject to channel errors. IEEE Trans. Image Processing, May 1995. 533~542
- Aign S. Error concealment for MPEG-2 video. Signal Recovery Techniques for Image and Video Compression and Transmission. Kluwer Academic Publishers, 1998
- Hong M C, Kondi L, et al. Video Error Concealment Techniques. Signal Processing: Image Communications, special issue on Error Resilient Video, 1999, 14(6-8): 437~492
- Zhu Q F, Wang Y, Shaw L. Coding and cell loss recovery for DCT-based packet video. IEEE Trans. CAS for Video Tech. June 1993. 248~285
- Yu G, Liu M M, Marcellin M W. POCs-based error concealment for packet video using multiframe overlap information. IEEE Trans. CAS for Video Tech. Aug. 1998. 422~434
- Sun H, Challapali K, Zdepski J. Error concealment in digital simulcast AD-HDTV decoder. IEEE Trans. Consumer Electronics. Aug. 1992. 108~117
- Lam W M, Reibman A R, Liu B. Recovery of lost or erroneously received motion vectors. Proc. ICASSP, 93, Minneapolis, Apr. 1993. 417~420
- Zhu Q, Wang Y. Error concealment in visual communications. Compressed Video over Networks. Marcel Dekker Inc., 2000
- Stenbach E, Farber N. Standard compatible extension of H. 263 for robust video transmission in mobile environments. IEEE Trans. Circuits. Syst. Video Tech. Dec. 1997. 872~881
- 李霖, 周兴铭. 移动数据库技术: 机遇和挑战. 计算机科学, 1997, 4
- Imielinski T, Badrinath B R. Data Management for Mobile Computing. Rutgers University, 1992
- Imielinski T, Badrinath B R. Mobile Wireless Computing: Challenges in Data Management. CACM, 1994, 37(10)
- Spyrou C, et al. Wireless Computational Models: Mobile Agents to the Rescue. IWOS'99, 1999
- STEMM M, KATZ R H. Measuring and Reducing Energy Consumption of Network Interface in Hand-Held Devices. IEICE TRANS. COMMUN, 1997, E80-B(8)
- 赵宇鹏, 徐明. 一种基于时隙的邮件消息发布算法. 待发表

(上接第123页)

- Bukhres O, Morton S. Mobile Computing Architecture for a Battlefield Environment. Computer Science, Purdue University, 1997
- Imielinski T, Rutgers University. Energy Efficient Data Filtering and Communication in Mobile Wireless Computing. 1994
- 李霖, 周兴铭. 非对称网络环境中的数据广播的启发式多盘调度算法. 计算机学报, 1999, 22(1): 1~7
- Imielinski T, et al. Data on Air: Organization and Access. [Tech. Report]. Rutgers University, 1994
- JIN JING GTE Laboratories Incorporated, et al. Client-Server Computing in Mobile Environments. ACM Computing Surveys, 1999, 31(2)