

# 视频流关键技术的研究进展

卓 力, 沈兰荪, 朱 青

(北京工业大学信号与信息处理研究室, 北京 100022)

**摘 要:** 视频流是在因特网上进行视频信息传送的主流方式. 为了在因特网上传输高质量的视频流, 需要采取相应的传输机制. 本文从视频流传输框架出发, 系统讨论了当前视频流关键技术的研究进展, 分析了各种技术的特点, 并指出进一步发展的前景.

**关键词:** 因特网; 流媒体; 视频流; 阻塞控制; 差错控制

**中图分类号:** TN915. 01      **文献标识码:** A      **文章编号:** 0372-2112 (2002) 08 1213-06

## Research Advances in Key Technology of Video Streaming

ZHUO Li, SHEN Liansun, ZHU Qing

(Signal and Information Processing Lab, Beijing Polytechnic University, Beijing 100022, China)

**Abstract:** Video streaming is the mainstream manners in delivery of video content over the Internet. To deliver high quality video streaming over the Internet, suitable transmission mechanism must be adopted. Starting from introduction to video streaming delivery framework, research advances in key technology of video streaming are summarized. The characteristics of techniques are analyzed and prospect is also pointed out.

**Key words:** internet; streaming media; video streaming; congestion control; error control

### 1 引言

目前, 因特网已经发展成为一个跨区域、跨国界、跨文化的信息传输平台, 网络多媒体化成为因特网的发展趋势. 在网络上传输多媒体信息时, 可以采用下载 (Download) 和流式 (Streaming) 传输两种方案. 音/视频数据文件一般都比较, 由于网络带宽的限制, 下载需要很长的时间, 需要的存储容量很大. 为了解决这些问题, 一种遵守特定网络协议的流式媒体 (Streaming media) 技术应运而生.

流式媒体技术是一种新的传送时间连续的数据和音/视频媒体的技术. 流式传输时, 用户不必等到整个文件全部下载完毕, 只需经过几秒或十数秒的启动延时即可进行观看. 当音/视频等媒体在客户端播放时, 文件的剩余部分继续下载, 用户不需要存储文件就可以边下载边连续、不中断地播放. 这样流式传输不仅使启动延时大大缩短, 还不需要太大的缓存容量<sup>[1-5]</sup>.

在流式媒体中, 视频流 (Video streaming) 技术是最为重要的. 本文从视频流的传输框架出发, 讨论了视频流的关键技术及应用前景.

### 2 视频流传输系统框架及其特点

根据因特网上视频流业务的特点可以将视频流分为两种

传输模式<sup>[1-3]</sup>: 直播 (Live streaming) 和点播 (On demand streaming). 点播与直播之间最大的区别在于视频编码器是否实时地对视频信息进行编码. 点播是将编码后的视频码流存储起来, 编码离线进行, 而直播需要编码器实时地对视频信息进行编码. 点播可以支持快进/快退/停止/回放等 VCR (Video Cassette Recording) 功能, 而直播一般难以具有这样的功能. 点播和直播可以采用单播 (Unicast) 和组播 (Multicast) 两种传输机制, 两种机制的视频流传输框架如图 1 所示.

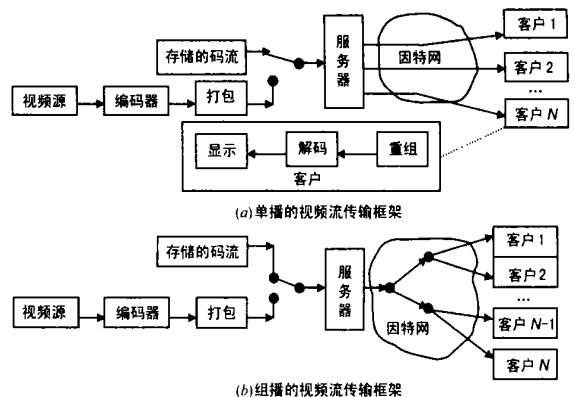


图 1 视频流传输框架

由图 1 可以看出,单播是在客户端与服务器之间建立一个单独的数据通道,服务器送出的数据包只能传送给一个用户.组播是适用于会议电视的一种传输模式,服务器将连续的数据流同时发送给多个用户,多个用户共享同一信息.显然,点播更适合采用单播机制进行传送,服务器应不同用户的点播要求将不同的视频内容传输给用户,而直播则适合采用组播机制.组播减少了网络上传输信息包的总量,大大提高了网络的利用效率,降低了成本.但传输的稳定性和灵活性还有待于进一步提高.单播虽然会造成服务器负载过重、网络利用率低等问题,却是方式灵活、适用性好的一种传播机制.

设计视频流系统时,应考虑以下特点<sup>[6-10]</sup>:

(1) 当前的因特网是为了传输数据而设计的,并不适合传输连续的流媒体.因特网是一个异构、时变、缺乏 QoS 控制的网络,虽然传输效率高,但带宽波动大,易发生丢包现象.网络提供尽力而为服务,很难满足用户个性化的 QoS (Quality of Service) 需要.

(2) 视频流基于服务器-客户模型传输.与传统的传输模型不同,视频编码器通过服务器与信道相连,视频信息预先以某一速率进行编码,编码器无法根据信道的状况调整编码速率.

(3) 在视频流系统中,每个用户的连接速率、终端的处理速度与显示能力、QoS 需求等存在很大的差异,解码器不一定象传统的方式那样将接收到的信息全部解码,而是根据实际情况动态调整.

(4) 视频流服务器同时为多个客户服务,服务器只能针对每个客户进行简单的处理.如果现有的用户用完了服务器的带宽,则会存在新用户难以接入等问题.

为了能在因特网上为用户提供连续、流畅的高质量视频服务,必须针对用户的异构性、因特网和视频业务的特点研究相应的机制,满足不同用户对 QoS 的个性化需求.这些研究涉及视频流的编解码技术、QoS 控制机制、网络传输协议的设计、用户端的误码恢复/隐藏以及缓存器的设计等.

### 3 视频流编解码技术

视频流技术的特点要求视频流的压缩技术不仅是高效的,还必须是码率可调整的,即视频数据只压缩一次,但可以以多个帧率、空间分辨率或视频质量进行解码,从而仅用一个码率可调整的码流就可以支持所有的用户.

针对视频流发展的需求,2000 年 MPEG-4 标准增补了视频流应用框架,提出了可精细扩展编码方法(FGS, Fine Granularity Scalability).经过多次实验,DCT 系数的比特平面编码因为实现复杂度低、性能优等特点被 MPEG-4 选定为 FGS 编码机制.

MPEG-4 FGS 采用的是一种混合分层/嵌入式编码策略.压缩后的码流包括一个基本层和一个增强层,基本层采用非可扩展的基于 DCT 变换的运动补偿混合编码方法,对原始图像与基本层重建图像之间的残差图像采用 DCT 系数的比特平面编码方法来得到增强层的码流.增强层码流可以在任意处截断,从而获得对码流速率的连续控制.这就是被称作“可

精细扩展”的原因<sup>[11]</sup>.

FGS 编码方法实现简单,可以在编码速率、显示分辨率、内容、解码复杂度、抗误码等方面提供灵活的自适应和可扩展性,具有很强的带宽自适应能力和抗误码性能.但还存在以下问题:

(1) 编码效率低于不可扩展的编码方法.这是由于非可扩展编码与 FGS 编码的出发点不同,前者的目的是在给定的编码速率下使编码的失真最小,而后者的目的是为了可精细地调整编码速率.为了避免因为接收不到增强层的信息而造成的误码扩散问题,FGS 预测时采用的是一种“开环”增强层的预测方式,预测时没有利用高质量的增强层信息.

(2) FGS 方法在编码时确定帧尺寸、帧率、解压缩的质量,在图像质量和编码效率之间达到均衡,并没有考虑实际可用的传输带宽、网络丢包率、用户爱好以及接收端的部分解码等问题,造成接收端得到的视频质量并非最优.

目前 MPEG-4 FGS 支持质量(SNR)可扩展,而不支持时间可扩展,无法进行时间分辨率的调整.为此,Mihaela 等<sup>[12]</sup>提出了一种混合时间 SNR 的 FGS 编码方案,采用一个 FGS 增强层就可以调整质量和时间分辨率,扩展了增强层的码率范围.Wu 等<sup>[13]</sup>提出的渐进可精细扩展方法 PFGS (Progressive Fine Granularity Scalability) 针对 FGS 方法编码效率不高的缺点,利用多个增强层进行预测以减少预测误差,提高了编码效率,但同时运算复杂度也会相应地提高.近来,人们在把对 FGS 的研究集中在提高增强层视频的视觉质量方面,包括频率加权、对感兴趣区有选择地增强以提高视觉质量、提高增强层的抗误码能力等.

### 4 VCR 功能的支持

在实际应用中,用户希望能够实现对多媒体信息的有效、快速的浏览,这其中的关键是能为用户提供 VCR 功能,如快进、快退、停止、随机存取等.要实现这些功能需要更多的网络带宽以及解码复杂度的提高.

随着新 MPEG 国际标准的不断推出,许多视频流应用都采用 MPEG 格式对视频信息进行压缩.因此,不少学者对 MPEG 码流支持 VCR 功能进行了深入的研究.MPEG 标准采用的是基于 I/P/B 帧结构的运动补偿预测编码机制,如果要对 P 帧解码,则需要首先对前面的 I/P 帧进行解码,而要对 B 帧解码,则前后的 P 帧都要先解码.这种 I/P/B 帧结构便于实现快进等前向播放功能,但要实现后向播放功能则难度很大.

一种实现后向播放功能的方法是将整个 GOP 解码后存储在一个大的缓存器中,然后再反向播放.但这需要用户端有极大的缓存器来存储解码帧,因此虽然实现简单但在实际应用中并不可行.Chen 等<sup>[14]</sup>提出的方法是在客户端将 P 帧转换成 I 帧,切断 I 帧和 P 帧之间的相关性.在进行帧类型转换和帧顺序重排后,采用交换运动矢量的办法进行新 I/B 码流的反向播放.但将 P 帧转换成 I 帧时会造成解码复杂度高、存储空间大等问题.

Wee 等<sup>[15]</sup>提出的方法将 I/P/B 格式的码流分成两部分: I/P 帧部分和 B 帧部分.采用码间转换的方法将 I/P 帧转换成反

向帧顺序的 FP 码流, 然后根据原有 FP 帧的前向运动矢量估计新 FP 码流的反向运动矢量, 降低了转换过程中的运算复杂度. 对于 B 帧, 采用一种用于反向播放的运动矢量交换机制, 但是这种码间转换过程所需的计算量仍然很大, 还会因为运动矢量估计的误差造成误码扩散.

以上这些方法都没有完全解决因为支持 VCR 功能而造成的网络流量增加和解码复杂度增大等问题. Omoigui 等<sup>[16]</sup>提出的解决快进问题的方法是存储多个具有不同时间分辨率的压缩码流, 然后根据用户的要求发送某一个具有适当的时间分辨率的码流. 这种方法不会造成网络流量过大的问题, 但是由于存储的码流数目有限, 快进的速度受到一定的限制. Lin 等<sup>[17]</sup>在服务器上采用双向码流结构来解决反向播放的问题, 并基于这种结构, 提出了一种服务器端的帧选择机制来减小所需的网络带宽和解码器复杂度. 采用误差补偿机制来减少由于码流切换而造成的误码扩散问题, 最终实现用 MPEG 4 视频流系统提供全部的 VCR 功能.

## 5 QoS 控制机制

QoS 控制机制有基于网络的和基于终端系统(end system)两种. 基于网络的方法是由网络中的路由器、交换机等提供 QoS 支持, 比如路由器发生阻塞时不再是随机丢包, 而是根据服务的优先级或包中信息的重要程度有选择地丢包等<sup>[18]</sup>. IETF 先后制定了支持一定 QoS 的因特网服务模型, 如 IntServ、DiffServ 等. 虽然采用这些服务模型可以提供一定的 QoS 保证, 但由于网络带宽波动问题依然存在, 还需要调度控制、排队等其他复杂的管理措施.

通常基于网络的方法成本比较高, 目前的因特网还无法在很大范围内支持这类方法. 基于终端系统的方法是由服务器和客户端采取 QoS 控制措施来提高视频质量, 而不需要网络的参与, 可以适应于现有的和未来的网络. QoS 控制机制可以分为两类: 阻塞控制(Congestion control)和差错控制(Error control)<sup>[19]</sup>. 下面进一步详细讨论.

## 6 阻塞控制技术

因特网的网络带宽是时变的. 视频流传输速率高于网络带宽时会发生阻塞, 造成突发的丢包和延时过大. 但如果视频流传输速率低于网络可用带宽, 就无法有效地利用网络资源. 因此, 阻塞控制技术的关键在于准确地估计网络带宽, 通过使视频流的传输速率与网络带宽匹配来防止阻塞的发生.

现有的 TCP 协议通过重传来保证数据的可靠传输, 不适合视频流的传输需求. 但目前因特网上的主要流量是基于 TCP 协议的, 采用的阻塞控制技术必须具有 TCP 友好的特性, 即一个新的视频流加入时不应该影响其他 TCP 流的正常传输<sup>[7-20]</sup>. 阻塞控制机制包括对网络可用带宽的估计、码率匹配两个方面.

### 6.1 网络可用带宽的估计

目前采用的网络带宽估计方法包括基于码率的方法和基于窗口的方法. 基于码率的方法根据网络的反馈信息来控制传输速率, 往往依据 AIMD(Additive Increase Multiplicative De-

crease) 准则或采用基于 TCP 模型的公式来估计网络带宽. 1988 年 Jacobson<sup>[21]</sup>提出的 AIMD 算法被 TCP 阻塞控制机制采用成为一种常用的带宽估计方法, 这是一种“试探”(probe)的算法: 网络没有丢包时则加性增加传输速率, 一旦网络发生了丢包, 则乘性降低传输速率. AIMD 算法具有一定的鲁棒性, 对阻塞反应灵敏, 但即使是网络带宽不发生变化, 也会由于周期性的信道状况检测而造成传输速率的波动, 造成视频传输质量的下降.

Raze 等<sup>[22]</sup>基于 AIMD 算法提出了 RAP 算法, 这是一种端到端的基于码率自适应的阻塞控制机制, 利用每个包的 ACK 估计回程时间和丢包率, 可以用于实时的视频流传输, 并具有 TCP 友好性能. Rohi 等<sup>[23]</sup>提出的 LIMD/H 算法利用过去的丢包信息区分丢包发生的原因, 并对丢包区别处理, 既对阻塞反应灵敏, 又能平缓调整传输速率.

基于 TCP 模型的公式法是通过 TCP 流量的大量分析, 推导出利用回程时间、丢包率来估计网络带宽的公式<sup>[2]</sup>. 这种方法可以平缓调整传输速率, 但过分依赖对回程时间和丢包率的估计, 对网络动态变化的反应比较慢.

基于窗口的方法通过调节拥塞窗口的尺寸大小来控制传输速率. 与 TCP 采用的控制机制类似, 根据接收端的 ACK 信息, 依据 AIMD 准则调整阻塞控制窗口的大小, 控制网络的流量, 避免阻塞的发生<sup>[2]</sup>.

以上这些阻塞控制算法都深受现有的 TCP 阻塞控制机制的影响, 过分强调具有 TCP 友好的特性, 不能有效利用网络的资源. 因此如何根据因特网的特点和采用的协议设计高效、快速的阻塞控制算法仍是值得研究的问题.

### 6.2 码率匹配

码率匹配是使视频码流的传输速率与网络可用带宽适配. 目前采用的技术主要包括码率转换、动态码率切换和可扩展编码等<sup>[1,4,7]</sup>.

码率转换方法是指编码器预先以某一固定速率对视频内容进行编码, 服务器通过有选择地丢帧、丢 DCT 高频系数以及重新量化等方式来调整编码速率. 采用这种方法只能调整有限的码率范围, 而且算法复杂度高, 需要服务器参与复杂的处理.

动态码率切换技术是指对同一视频内容采用多个码率进行编码, 服务器根据客户的信道状况反馈信息选择最匹配的编码速率进行传输. 这种方法需要用到多个速率的码流, 占用的存储资源大. 客户端解码时同步困难, 码率调整的范围有限.

可扩展编码方法被认为是可用于视频流以解决因特网异构特性的编码方法, 服务器根据网络带宽的具体状况决定是否传输增强层、传几个增强层, 服务器参与的处理比较少<sup>[4]</sup>. 在解码端, 增强层要么根本不可用, 要么作为一个整体被解码, 因此采用可扩展编码方法获得的视频质量是不连续的, 呈阶梯状变化. 如果只有一个或两个增强层可用, 那么码率调整的程度有限. 如果有多个增强层可用, 则编码开销增大, 编码效率降低.

前面讨论的 FGS 编码方法是可扩展编码方法中的一种,

但不是采用多个增强层来获得中间质量,而是可以获得对码率的连续调整,解码质量与解码比特数成正比。服务器只需要根据网络带宽对增强层码流截断,参与的处理很少,因而在视频流应用中得到了越来越多的应用<sup>[6,8]</sup>。

## 7 差错控制

差错控制的目的是为了解决丢包问题,这包括应用层和传输层的差错控制。应用层的差错控制包括从视频压缩角度考虑的抗丢包能力以及客户端的丢包检测与恢复,传输层的差错控制包括打包算法的设计以及 FEC 等。

### 7.1 信源编码的抗丢包能力

目前 H.263、MPEG-4 标准中所采用的多种抗误码技术如重同步标记、数据分割以及数据恢复等针对无线信道的误码控制,对因特网的丢包恢复并不适合。具有抗丢包能力的视频编码方法往往利用码流中的冗余信息来有效地抗丢包,这涉及多描述编码和最优模式选择。

多描述编码的基本思想是对同一视频内容采用多种方式进行描述,每一种描述都可以获得可接受的视频质量,多个描述方式结合起来可以使视频质量得到增强。这种方法虽然压缩效率低,但传输时不需要采取很强的保护措施就可以有效地抗丢包<sup>[23]</sup>。多描述编码的效率和计算复杂度等问题还有待于进一步的解决。

Zhang 等<sup>[24]</sup>研究了在传统的率失真框架下如何自动选择 Intra 编码宏块的数目和位置,在给定了丢包率和编码速率的情况下可以获得最小的全局失真。这种方法是在给定的条件下通过编码模式选择使重建视频与原始视频的失真达到最小,而没有充分考虑网络的具体状况和客户端的误码恢复能力。Wu 等<sup>[25]</sup>提出的端到端的全局率失真最优模式选择方法则综合考虑了信源端的量化、打包、信道特性、接收端的误码隐藏能力等因素,获得了更好的性能。

Cote 等<sup>[26]</sup>把率失真最优问题应用到可扩展视频编码中,基于可扩展编码和优先级传输原理提出了一种视频通信框架。这种方法综合考虑了信道状况、信道编码的误码恢复能力、解码器的差错恢复/隐藏能力等各种因素,为每一层的每个编码块选择编码模式,可以在给定的码率下使接收端重建的失真最小。

### 7.2 客户端的误码恢复与隐藏

误码恢复与隐藏技术是在发生了丢包的情况下,由接收端采取的一种“后处理”技术。视频信息不同于一般数据(或文本)信息,它具有极强的空间和时间相关性。因此可以充分利用这一特性,寻找一些相关的数据来代替误码数据,使误码不被人眼觉察出来。Wang 等<sup>[27]</sup>对接收端的误码恢复与隐藏技术作了很好的总结,但最大平滑恢复、凸集投影等方法比较适合于 ATM 网络和无线环境,并不适合因特网上的视频流应用。

时间/空间插值法是两种常用的误码恢复技术,时间插值法适合于 INTER 编码模式下的误码恢复,比如由前一重建帧代替发生了误码的当前帧;由前一重建帧中相应位置处的块代替受损块;由前一重建帧中由运动矢量确定的块代替受损

块等都属于这类方法<sup>[28]</sup>。空间插值法适合于 INTRA 模式下的误码恢复,利用受损块周围的像素插值来恢复误码。时间、空间插值还可以结合起来使用,以最大限度地提高重建视频的质量。

### 7.3 打包机制

设计打包机制时要综合考虑传输效率、抗丢包能力等因素,打包过程可以提供一种重同步的方法,增强系统的抗丢包能力。Zhu 等<sup>[29]</sup>在打包前将相邻几个图像块的信息作适当交织,在发生了误码的情况下比较容易恢复出原始的视频信号。Le 等<sup>[30]</sup>研究了 MPEG-4 码流的打包方法,每个包采用固定尺寸,一个宏块可以打成两个数据包,这样数据包之间就会具有一定的相关性,因此这种方法的抗丢包能力差。Turletti 等<sup>[31]</sup>提出了一个宏块一个数据包的打包算法,与 Le 的方法相比增强了抗误码能力,但传输效率低。Zhu 等<sup>[32]</sup>采用一个 GOB 一个包的打包算法,提高了传输效率。Wu 等<sup>[33]</sup>则利用 MPEG-4 的 VOP 特性设计了一个混合层的打包算法,在传输效率高的同时也提高了抗误码能力。

Michael 等<sup>[9]</sup>则对基于“GOB”或“Slice”结构的多种打包算法进行了深入的研究,并给出了比较结果。这些打包算法包括一个数据包可以包括 1 个、2 个甚至多个 GOB,一个帧的数据可以放在一个数据包中,奇数行的 GOB 交织后放在一个数据包中、偶数行的 GOB 交织后放在一个数据包中等。实验表明,数据包的尺寸小可以有效地抗丢包,但传输效率低。

### 7.4 FEC

FEC 通过给压缩后的视频码流加上一定的冗余信息来有效抗误码,是视频流传输时经常采取的差错控制机制,主要包括信道 FEC 编码、基于信源编码的 FEC 以及信源/信道联合编码<sup>[8, 28, 34, 35]</sup>等。

因特网上的信道编码主要采用块编码的形式,如 RS (Reed-Solomon) 码。先将视频码流的每个分段打成  $K$  个包,通过对这  $K$  个包中数据块的线性无关组合产生  $N(N \geq 1)$  个冗余块,这样共得到  $M = K + N$  个包块。用户只要正确接收  $M$  个包中的任意  $K$  个包(但必须至少接收  $K$  个包,否则接收到的数据包全部作废)就可以完全恢复一个分段。因此,客户端可以根据自己的处理能力选择一定数目的包进行解码,从而适应客户端异构性的特点。但用户必须至少接收到  $K$  个包后才能开始解码,因此增加了解码延时。基于信源编码的 FEC 这样加冗余信息:第  $N$  个包中不仅包括码流信息还包括第  $N - 1$  个包的冗余信息,如果  $N - 1$  包丢失,则根据第  $N$  个包的冗余信息恢复丢失的数据。

以上这两种都是基于信源/信道分离理论的编码技术,无法根据信道的状况来调节编码速率,不适用于因特网这类时变的传输信道。信源/信道联合编码则在给定网络可用带宽的情况下,通过最优地分配信源和信道编码速率,使视频传输失真达到最小<sup>[35]</sup>。随着可扩展编码技术的发展,人们开始研究根据视频信息的重要性程度、客户定义的优先级以及网络的丢包率高低等来决定所采取的不等错误保护措施<sup>[8-10, 35]</sup>,不等的错误保护与可扩展编码相结合可以看作是一种可以有效抗丢包的信源/信道联合编码方案。Rohit 等<sup>[23]</sup>将 FEC 与多描

述编码相结合, Alexander 等<sup>[10]</sup>则将 FEC 与基于 SPIHT 算法的可扩展编码方法相结合, 在因特网上进行视频流传输均取得了很好的效果. FGS 方法与不等的错误保护结合这是当前视频流传输技术的一个研究热点<sup>[6, 8, 12]</sup>. 研究结果表明, 这种方法具有很强的抗丢包能力. 与其他方法相比, 可以将 PSNR 提高 5dB 左右.

## 8 缓存器的设计

视频流传输的实现需要缓存. 因特网以包传输为基础进行异步传输, 对于一个数据量很大的视频文件来说, 在传输中会被分解为许多包. 由于网络的带宽是动态变化的, 各个包选择的路由不尽相同, 到达客户端的延时也不一样. 为此, 需要使用缓存器来弥补延时波动的影响, 保证视频播放的连续, 流畅地进行. Hayder 等设计了一种传输-解码缓存模型, 与基本层的重传相结合, 可以保证解码的流畅进行.

## 9 视频流传输中采用的协议

在因特网上传输视频流时需要相应的传输协议, 所涉及的协议包括: 网络层的 IP 协议, 传输层的 TCP/UDP, RTP/RTCP 协议, 以及会话层的 RTSP, SIP 协议等<sup>[30]</sup>. IP 提供了在因特网上传送 UDP/TCP 数据包的公共平台, UDP/TCP 是用于传送 RTP/RTCP/RTSP/SIP 数据包的底层传输协议, 这些协议结合起来可以提供因特网上的视频流服务.

TCP 可以通过重传丢失的数据包而提供可靠的传输, 利用阻塞控制来防止网络阻塞. 由于 TCP 反复重传会引入过大的延时, 因此在因特网上传输视频流时往往采用 UDP 协议. 但 UDP 不能保证数据包的可靠传输, 需要利用上层的 RTP 协议来检测是否丢包.

RTP/RTCP 运行在 TCP/UDP 协议之上. RTP 是用于因特网上的多媒体数据的传输协议, 可以为实时应用提供时间信息和流同步. RTP 本身并不能为顺序传送数据包提供可靠的传输机制, 也不提供流量控制或拥塞控制. RTCP 是监视 RTP 包传送的控制协议, 可以给发送端提供 QOS 反馈, 和 RTP 配合使用, 能以有效的反馈和最小的开销使传输效率最佳化.

RTSP 在位于 RTP/RTCP 之上, 是控制流媒体在因特网上传输的协议. 它可以提供 VCR 功能, 还可以建立、控制服务器和客户端之间的连续视频/音频流. SIP 是会话控制协议, 可以建立或终止与一个或多个用户的会话. 与 RTSP 不同的是, SIP 支持用户的移动.

## 10 结束语

以宽带为基础, 视频流不仅可以用于视频点播、数字图书馆等多媒体业务, 还能用于一些对实时性要求很高的场合, 如实况转播新闻、球赛、重要会议等. 另外, 视频流技术还可以用于远程监控、安全监督以及互动视频节目等方面.

作为多媒体与网络领域的交叉学科, 流式媒体的应用与研究得到了迅速的发展. 它衍生出了适合流式传输的网络通信技术, 多媒体数据采集技术、数据压缩技术和存储技术等基础技术. 流式媒体已经发展成为一个产业. 可以预见, 流式媒

体将成为未来因特网应用的主流, 并将推动因特网整体架构的革新.

### 参考文献:

- [1] Jian Lu. Signal processing for internet video streaming: A review [A]. Proceedings of SPIE Image and Video Communications and Processing [C]. San Jose, CA USA: 2000. 1- 14.
- [2] Dapeng Li, et al. Streaming video over the internet: Approaches and directions [J]. IEEE Trans on Circuits and Systems For Video Technology, 2001, 11(1): 1- 20.
- [3] Gregory J Conklin, et al. Video coding for streaming media delivery on the internet [J]. IEEE Trans Circuits and Systems For Video Technology, 2001, 11(3): 269- 281.
- [4] Reza Rejaje, et al. Layered quality adaptation for internet video streaming [J]. IEEE Journal Selected Areas in Communications, 2000, 18(12): 2530- 2543.
- [5] Jane Hunter, et al. A review of video streaming over the internet [DB/OL], <http://archive.dst.c.edu.au>. 2001- 05.
- [6] Hayder Radha, et al. Scalable internet video using MPEG-4 [J]. Signal Processing Image Communication, 1999, 15: 95- 126.
- [7] Reza Rejaje, et al. Architectural consideration for playback of quality adaptive video over the internet [DB/OL], <http://www.citeseer.nj.nec.com>. 2001- 06.
- [8] U horn, et al. Robust internet video transmission based on scalable coding and unequal error protection [J]. IEEE Trans Image Processing, 1999, 15: 77- 94.
- [9] Michael Gallant. Rate-distortion optimized layered coding with unequal error protection for robust internet video [J]. IEEE Trans Circuits and Systems for Video Technology, 2001, 11(3): 357- 372.
- [10] Alexander E Mohr. Unequal loss protection: Graceful degradation of image quality over packet erasure channels through forward error correction [J]. IEEE Journal Selected Areas in Communications, 2000, 18(6): 819- 828.
- [11] Weiping Li. Overview of fine granularity scalability in MPEG-4 video standard [J]. IEEE Trans Circuits and Systems For Video Technology, 2001, 11(3): 301- 317.
- [12] Mihaela van der Schaar. A hybrid temporal-SNR fine granular scalability for internet video [J]. IEEE Trans Circuits and Systems For Video Technology, 2001, 11(3): 318- 331.
- [13] Wu Feng, et al. A framework for efficient progressive fine granularity scalable video coding [J]. IEEE Trans Circuits and Systems For Video Technology, 2001, 11(3): 332- 344.
- [14] Chen M S, et al. Download and stream conversion: Supporting interactive playout of videos in a client station [A]. IEEE Conf Multimedia Computing and Systems [C]. Washington: 1995. 73- 80.
- [15] Wee S J, et al. Compressed domain reverse play of MPEG video streams [A]. Proc SPIE Conf Multimedia Systems and applications [C]. Boston, MA USA: 1999. 237- 248.
- [16] Omoigui N, et al. Time compression: System concerns, usage, and benefits [A]. Proc ACM SIGHI Conf [C]. New York USA: 1999. 136- 143.
- [17] Chir wen Lin, et al. MPEG video streaming with VCR functionality

- [ J ]. IEEE Trans Circuits and Systems for Video Technology. 2001, 11 ( 3 ): 415- 425.
- [ 18 ] Huar Rong Shao, et al. User and content aware object based video streaming over the Internet [ A ]. Visual Communications and Image Processing [ C ]. Sadinia Italy: 2000. 23- 30.
- [ 19 ] V Jacobson. Improving QoS in IP networks [ DB/OL ]. <http://www.citeseer.nj.nec.com>. 2002- 06.
- [ 20 ] Jong Won Kim. TCP Friendly Internet video streaming employing variable frame rate encoding and interpolation [ J ]. IEEE Trans Circuits and Systems for Video Technology. 2000, 10( 7 ): 1164- 1177.
- [ 21 ] V Jacobson. Congestion avoidance and control [ A ]. Proc SIGCOMM [ C ]. Stanford, CA, USA: 1988. 314- 319.
- [ 22 ] Reza Rejaje, et al. RAP: An End to End Rate based Congestion Control mechanism for real time Streams in the Internet [ DB/OL ]. <http://www.citeseer.nj.nec.com>. 2001- 06.
- [ 23 ] Rohit Puri, et al. Forward error correction (FEC) codes based multiple description coding for internet video streaming and multicast [ J ]. Signal Processing: Image Communication 2001, 16: 745- 762.
- [ 24 ] Rui Zhang, et al. Video coding with optimal inter/ intr mode switching for packet loss resilience [ J ]. IEEE Journal of Selected Areas in Communications. 2000, 18( 6 ): 966- 976.
- [ 25 ] Dapeng wu, et al. An end to end approach for optimal mode selection in internet video communication: theory and application [ J ]. IEEE Journal Selected Areas in Communications. 2000, 18( 6 ): 977- 995.
- [ 26 ] Guy Cote, et al. Optimal mode selection and synchronization for robust video communications over error prone networks [ J ]. IEEE Journal Selected Areas in Communications. 2000, 18( 6 ): 952- 965.
- [ 27 ] Wang Yao, et al. Error control and concealment for video communication: A Review [ J ]. Proceedings of The IEEE. 1998, 86( 5 ): 974- 997.
- [ 28 ] Shahram Shirani, et al. Error concealment methods, A Comparative Study [ A ]. IEEE Canadian Conference on Electrical and Computer Engineering [ C ]. Edmonton, Canada: 1999. 835- 840.
- [ 29 ] Q F Zhu, et al. Coding and cell loss recovery for DCF based packet video [ J ]. IEEE Trans Circuits Systems for Video Technology. 1993, 3: 248- 258.
- [ 30 ] F Le Lenec, et al. Error resilient video transmission over the Internet [ A ]. Proc. SPIE Visual Communications and Image Processing [ C ]. Newport Beach, CA USA: 1999. 101- 113.
- [ 31 ] IETF. RTP payload format for H. 261 video streams [ S ]. RFC 2032. 1996.
- [ 32 ] IETF. RTP payload format for H. 263 video streams [ S ]. RFC 2190. 1997.
- [ 33 ] Dapeng Wu, et al. On end to end architecture for transporting MPEG-4 video over the Internet [ J ]. IEEE Trans Circuits and Systems for Video Technology. 2000, 10( 6 ): 923- 941.
- [ 34 ] Dapeng Wu, et al. Transporting real time video over the internet: Challenges and Approaches [ J ]. Proceedings of The IEEE. 2000, 88: 1- 19.
- [ 35 ] Geoffrey Davis, John Danskin. Joint source and channel coding for Internet image transmission over lossy packet networks [ EB/OL ]. <http://www.citeseer.nj.nec.com>. 2001- 06.
- [ 36 ] Andrew Campbell, et al. A review of QoS architectures [ EB/OL ]. <http://www.citeseer.nj.nec.com>. 2001- 06.

### 作者简介:



卓力女, 1971年出生于江苏徐州, 1992年和1998年分别获得电子科技大学和东南大学学士和硕士学位, 现为北京工业大学讲师、博士生, 近年来发表论文多篇, 主要研究方向为视频编码、无线IP视频传输等。



沈兰荪男, 1938年生于江苏苏州, 北京工业大学教授、博士生导师, 主要研究兴趣为图像与视频信号的压缩编码、处理与传输等, 著有“图像编码与异步传输”、“视频编码与低速率传输”等多部专著, 发表学术论文二百余篇。



朱青女, 1965年生于浙江义乌, 1994年和1997年分别获得日本早稻田大学硕士和博士学位, 现为北京工业大学研究员, 主要研究领域为多媒体通信及国际文字符号的标准化等, 发表学术论文40多篇。