

流媒体技术及其开发方法

逢 栋,姜昌金

(东南大学 自动化研究所,江苏 南京 210096)

摘要:随着 Internet 技术的迅速发展,各种各样的网络多媒体应用层出不穷,其中以流媒体技术为背景的应用发展最为迅速,其目的在于缩短客户对多媒体数据回放的等待时间和为用户提供较好的服务质量,深受教育和视频监控行业的青睐。文中主要论述了流媒体技术、原理及其在网络上传送时所使用的 RTP/RTCP 协议,并与传统的 TCP/IP 传输协议进行了详细的比较,进一步阐明了 RTP/RTCP 协议所具有的优势。同时介绍了微软公司最新推出的基于 COM 技术的流媒体应用开发工具 DirectShow,利用它可以大大缩短流媒体应用的开发周期。文中结合一个简单的例子说明了它一般的使用方法。

关键词:流媒体;RTP/RTCP;DirectShow;Filter;COM

中图分类号:TN919.8

文献标识码:A

文章编号:1005-3751(2006)02-0145-03

Streaming - media Technique and Methods of Its Development

PANG Dong, JIANG Chang-jin

(Research Institute of Automation, Southeast University, Nanjing 210096, China)

Abstract: With the rapid development of Internet, there is more and more various applications based on multimedia on Internet, especially for applications of streaming-media technique. The main purpose is to reduce the time of playback and supply the better quality of service, causes greates interest in education and video surveillance. Describes the streaming-media technique, principles and the RTP/RTCP protocols when it is transported on the network, at the same time, compares with traditional TCP/IP protocols to make clear the advantages of RTP/RTCP. Simultaneously a toolkit developed by Microsoft Corporation called DirectShow which is based on COM technique is introduced, it can simplify the process of development. A simple example is to explain the method that how to use it.

Key words: streaming - media; RTP/RTCP; DirectShow; filter; COM

0 概述

流媒体指在 Internet/Intranet 中使用流式传输技术的连续时基媒体,如:音频、视频或多媒体文件。流式媒体在播放前并不下载整个文件,只将开始部分内容放入内存缓冲区,流式媒体的数据流随时传送随时播放,只是在开始时有些延迟。流媒体实现的关键技术就是流式传输,实现流式传输有两种方法:实时流式传输(Real time streaming)和顺序流式传输(Progressive streaming),实时流式传输是指保证媒体信号带宽与网络连接匹配以使媒体可被实时观看到^[1]。顺序流式传输是顺序下载,在下载文件的同时用户可观看在线媒体,在给定的时刻,用户只能观看已下载的那部分而不能随意地跳到还未下载的前头部分,顺序流式传输不像实时流式传输在传输期间根据用户的连接速度作调整。由于标准的 HTTP 服务器可发送这种形式的文件,也不需要其他特殊协议的支持,故经常被称作 HTTP 流式传输。由于该文件在播放前观看的部分是无

损下载的,这可以保证媒体文件播放的高质量。而实时流式传输要使用专门的流媒体服务器(如:Windows Media Server)和专门的传输协议(如:RTP/RTCP)。实时流式传输总是实时传送,特别适合现场事件,也支持随机访问,用户可快进或后退以观看前面或后面的内容。但是实时流式传输底层使用的是面向非连接的 UDP 传输协议,当网络拥挤或出现问题时,由于丢失的信息被忽略掉而使媒体播放的质量难以得到保证^[2]。

1 流媒体传输协议及其所具有的优势

流媒体面向的是多媒体应用,而多媒体应用的一个显著特点是数据量大并且许多应用对实时性都有比较高的要求。传统的 TCP/IP 协议最初是为提供非实时数据业务而设计的,IP 协议负责主机间的数据传送,其本身不进行检错和纠错,因此经常发生数据包丢失或失序现象。为保证数据的可靠传输,人们将 TCP 协议用于 IP 数据的传输,从而提高了接收端的检错和纠错能力,但不可避免地引起传输延时和耗用网络带宽。因此传统的 TCP/IP 协议传输实时流媒体的能力较差。当然在有足够大的缓冲区和充足的网络带宽的情况下,在 TCP 协议上实现接近实时的传输也是可能的。如果网络状况不佳,那么用 TCP 协议

收稿日期:2005-05-13

作者简介:逢 栋(1977—),男,山东青岛人,硕士研究生,主要从事多媒体技术在工业控制中的应用;姜昌金,副教授,研究方向为工业过程控制。

传输实时流媒体数据几乎是不可能的。其主要原因在于以下 4 个方面:TCP 的重发机制,TCP 的拥塞控制机制、报文头结构、启动速度慢。

1.1 TCP 的重发机制

在 TCP/IP 协议中,当发送方发现数据丢失时,他将会重发丢失的数据包,从 TCP/IP 的快速重发机制可以知道这需要额外的 3 个帧延迟,而这种重发机制对于实时性要求很高的流媒体数据传输来说会造成灾难性的后果。因为这样很容易造成网络的拥塞并且接收方不得不等待丢失数据的到来,从而造成音频和视频出现严重的不连续。

1.2 TCP 的拥塞控制机制

TCP 的拥塞控制机制在探测到有数据包丢失时,它就会减小其拥塞窗口。另一方面,音频、视频在特定的编码方式下,产生的编码数量是不可能突然改变的。例如标准的 MPEG1 视频需要,加上一些额外的控制信息,它不能在低于这个带宽要求的网络上传输。正确的拥塞控制应该是变换视频信息的编码方式,调节视频信息的帧频,或图像的大小。在 RTP 协议中,发送方能够根据 RTCP 返回的信息进行调节。

1.3 报文头结构

TCP 协议的报文头比 UDP 和 RTP 协议的报文头要大(TCP 的报文头为 40 字节,而 RTP 的报文头仅为 12 字节),并且 TCP 协议不能提供时间戳(Time Stamp)和编解码信息等这对于流媒体传输至关重要的内容。

1.4 启动速度慢

即便在网络状况良好,没有丢包的情况下,由于 TCP 的启动需要建立连接,因而在初始化的过程中需要较多的时间。而在实时流媒体应用中,应当尽量减少延迟。

可见,TCP/IP 不适合传输实时流媒体数据。为了解决以上所面临的问题,IETF 在 RFC1889 中发布了支持实时流媒体通信的 RTP/RTCP 传输层协议。RTP 协议位于 UDP 协议之上,因为 UDP 是面向非连接的协议,所以传输数据没有 TCP 可靠,并且无法保证实时业务的服务质量,这需要通过 RTCP 协议来保证。TCP 协议在流媒体应用中通常用来传送可靠性要求较高的控制报文。实际应用中 RTP 协议的实现经常被融合到应用程序之中,RTP 协议中没有连接的概念,其报文的发送既可建立在面向连接的底层协议上(如 TCP),也可建立在面向非连接的底层协议上(如 UDP),因此 RTP 协议不具有报文的分段功能,其数据包的分段是由其底层的协议来完成的。RTP 协议由两部分组成:数据报文部分(RTP 报文)和控制报文部分(RTCP 报文)^[3]。

RTP 报文由报文头和数据部分组成。报文头长度为 12 字节,包含 V(版本号),X(扩展位),CC(CSRC 个数),PT(负载类型),Sequence Number(顺序号),TimeStamp(时间戳),SSRC(时间戳),CSRC(贡献源)等信息。

①PR(负载类型):7 位,标识所传输数据的编码类型,接收方使用此字段对于接收到的数据进行解码。

②TimeStamp(时间戳):32 位,此字段反映了 RTP 报文第一个字节的采样时间。为了同步和计算网络抖动,时钟必须单调线性增长。如果定时产生发送数据,则相邻 RTP 报文的时间戳之差就是采样间隔时间。

③SSRC(同步源):32 位,标识数据的来源。

④CSRC(贡献源):32 位,标识混合报文的各个来源。由于 CC 字段的长度为 4 位,因此一个 RTP 混合报文中最多能够标识 15 个 CSRC。

RTCP 协议是 RTP 协议中的控制功能协议,它单独运行于底层协议之上。RTCP 报文是由接受方向发送方发送的报文,它负责监视网络服务质量、通信带宽以及网上传送的信息,并将其通知给发送端。RTCP 的主要功能如下:

- 对数据的传输质量提供反馈。

所有的接收方把它最近的接收情况报告给所有的发送方。这些信息包括所收到数据包的最大序列号、丢失的包数、乱序的包数以及用于估计传输延时的时间戳信息。这些信息反映了当前的网络状况,接受方在接受到这些信息之后,自动调整它们的发送策略。

- 提供不同媒体之间的同步。

例如在视频会议系统中,RTP 源有视频/音频两种媒体需要传输,这些媒体之间的同步需要依靠 RTCP 中包含的时钟信息和相关的 RTP 时间戳信息来进行同步。

另外,RTCP 报文中的源描述(SEDs)提供了会话参与者的详细描述(如:姓名、地址等),这样为在会话中实时地显示参与者的有关信息提供了方便。根据功能的不同 RTCP 报文可分为 RR(接收者报告),SR(发送者报告),SEDs(源描述),BYE(离开声明)和 APP(应用)5 类。

通过以上对 RTP 和 RTCP 报文的分析,可以知道 RTP 协议确实提供了一种适合实时流媒体信息传输的机制。

2 流媒体应用的开发工具与开发方法

随着数字音视频处理的不断发展,基于数字编解码的多媒体应用已越来越深入人们的日常生活之中。但是流媒体的处理向来以其复杂性和技术性而闻名,怎样使流媒体的处理变得简单而富有成效,逐渐成为人们关注的焦点问题。为了向开发人员提供功能更为完备的处理软件,Microsoft 公司推出了新一代的多媒体开发软件包——DirectShow。

它主要提供播放本地或 Internet 服务器上的多媒体格式的文件或数据,支持基于 VFW 和 WDM 驱动的音视频捕获。

2.1 DirectShow 体系结构简介

DirectShow 是一种通过自定义或内置的过滤器(filter)来控制和处理多媒体数据的体系结构,其核心是被称为过滤器的 COM 组件。该体系定义了如何控制和处理过滤器内部和相互间的多媒体数据流,每个过滤器都拥有输入针或输出针,由它负责过滤器之间的连接。输入针接收数据

进入过滤器;输出针向其他过滤器提供数据。典型的转换过滤器,如:解压缩过滤器,提供一个输入针和一个输出针,而视频捕获过滤器仅拥有一个输出针。应用所需的过滤器放在过滤器图表中(filter graph),过滤器图表管理器(filter graph manager)组件负责检查这些过滤器的连接和控制过滤器之间的数据流动。为了使过滤器图表正常工作,过滤器必须通过输入针和输出针以正确的顺序相连,数据流也必须以正确的次序启动和停止。过滤器图表管理器负责连接过滤器和控制媒体流,它能够自动搜寻可以递交指定媒体类型的过滤器配置,并且建立与之相应的过滤器图表。当搜寻递交配置时,过滤器图表管理器首先读取注册表以决定可用的过滤器类型;接着,过滤器图表管理器不断地试图连接接受该数据类型的过滤器,直到数据到达递交过滤器。每个过滤器均注册有一个优先值,在都可以处理数据时,优先级最高的过滤器首先被使用。DirectShow 系统与应用程序之间的关系如图 1 所示^[4]。

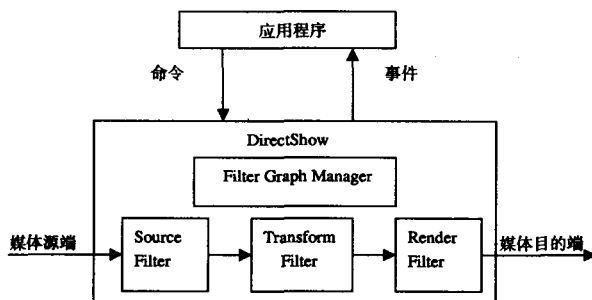


图 1 应用程序与 DirectShow 的关系图

2.2 基于 DirectShow 的转换过滤器设计实例

过滤器是一种进程内的 COM 组件,由于 COM 规范本身比较复杂,而且 DirectShow 系统自身也包含了很多实现的细节^[5]。为此,DirectShow SDK 提供了一套开发 Filter 的基类源代码,基于这些基类开发 Filter 将大大简化开发过程。该例中要开发的过滤器其功能是将显示器上播放的图像进行左右的反转,其开发步骤如下:

(1)首先选取开发所需要的基类,因为要开发的 Filter 是接收数据,经过一定的处理后再输出数据,所以选择了 CTransInPlaceFilter 作为基类^[6],并且不需要重新实现自

己的 Pin 类。

(2)以 CTransInPlaceFilter 为基类建立我们自己的类 class CVideoConverter : public CTransInPlaceFilter, 在 CVideoConverter 中重载基类中的成员函数 CheckInputType (const CMediaType * mtIn)用来判断从输入针进入的媒体类型;在 CompleteConnect (PIN_ DIRECTION direction, IPin * pReceivePin)中由媒体类型选择相应的反转处理函数;在 Transform (IMediaSample * pSample)中完成图像的反转算法。

(3)调用 CoCreateInstance 方法建立过滤器图表实例 CoCreateInstance (CLSID-FilterGraph, NULL, CLSCTX_INPROC, IID_ IgraphBuilder, (void * *) m_ pConverter)。m_ pConverter 是指向 IfilterGraph 型的指针,调用其中的 AddFilter 方法即可将过滤器加入到过滤器图表当中,最后调用 Render 方法通过 DirectShow 提供的智能连接机制即可自动完成 Filter 之间的连接。

3 结束语

文中主要介绍了流媒体技术的有关概念,并讲述了 DirectShow 的主要组成部和工作原理,以一个例子简要介绍了开发过滤器组件的一般方法。对于从事视频监控和多媒体软件开发的编程人员具有一定的参考价值。

参考文献:

- [1] 藏劲松. 流媒体技术及其在远程教学中的应用[J]. 微机发展, 2003, 13(1): 55-57.
- [2] 景海丽, 裘雪红. DirectShow 技术在网络编程中的应用[J]. 微机发展, 2003, 13(11): 96-98.
- [3] Schulzrinne H, Casner S, Frederick R, et al. RTP: A Transport Protocol for Real-Time Applications. [S]. RFC 1889. 1996.
- [4] 陆其明. DirectShow 开发指南[M]. 北京:清华大学出版社, 2003.
- [5] 潘爱民. COM 原理与应用[M]. 北京:清华大学出版社, 2004.
- [6] Microsoft Corporation. Microsoft DirectX Media SDK [CP/DK]. <http://www.microsoft.com>, 2003.

(上接第 144 页)

- [3] Chen Mao, Dorer K. User Manual RoboCup Soccer Server for Soccer Server Version 7.07 and later[Z]. In RoboCup Soccer Server User Manual, 2003.
- [4] Quinlan J R. C4.5: Programs for Machine Learning[M]. San Mateo, CA: Morgan Kaufmann, 1993.
- [5] Mutakami K. Potential Field Approach to Short Term Action Planning in RoboCup F180 League[J]. Computer Science, 2003, 2019: 345-350.
- [6] Meyer J, Adolph R. Decision-Making and Tactical Behavior With Potential Fields[A]. In RoboCup-2002 Artificial Intelligence[C]. [s.l.]: Springer, 2002. 304-311.
- [7] Zhang Bo, Cai Qingsheng, Chen Xiaoping, et al. An Agent Team for RoboCup Simulator League[A]. Proceedings of the 3rd World Congress on Intelligent Control and Automation[C]. Hefei, P. R. China: [s.n.], 2000. 189-193.
- [8] Kraus S. Negotiation and Cooperation in Multi-Agent Environments[J]. Artificial Intelligence, 1997(9): 79-97.
- [9] Kitano H, Asada M, Osawa E, et al. A challenge problem[J]. AI Magazine, 1997, 18(1): 73-85.
- [10] 李实, 陈江, 孙增圻. 清华机器人足球队的结构设计与实现[J]. 清华大学学报(自然科学版), 2001, 41(7): 94-97.