

虚拟摄像头在基于 H.323 的远程教学系统中的应用

袁 聃, 程文青, 吴 砥

(华中科技大学电子与信息工程系, 武汉 430074)

摘要: 视频的传输和显示是现代远程教学系统中的基本组成部分。由于视频数据的数据量大, 传输具有持续性的特点, 因此视频传输和显示模块的处理效率直接决定了系统的稳定性和可用性。该文提出了远程教学系统中信令和数据传输分离的两层体系结构, 以及利用内核模式解决用户模式问题的新思路, 阐述了虚拟摄像头在基于 H.323 的远程教学系统中应用的可行性, 进一步给出了该方案的具体实现方法。
关键词: 远程教学; H.323; 虚拟摄像头; DirectShow

Application of Virtual Camera in E-learning System Based on H.323

YUAN Dan, CHENG Wen-qing, WU Di

(Dept. of Electronic and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074)

【Abstract】 The transmission and display of video is an essential part in the e-learning system. Because of the great capacity of video data and durative transmission, the efficiency of the processing on the transmission and display module directly decides the stability and usability of the system. This thesis proposes a double-deck structure in which command and data are transmitted separately, as well as an idea which could solve the problems of the user mode through kernel mode. This thesis presents the feasibility of applying the virtual camera in e-learning system based on H.323, and gives a realization method furthermore.

【Key words】 e-learning; H.323; virtual camera; DirectShow

随着信息技术的发展, 教育呈现出了越来越多的形式。现代远程教学系统结合现代网络技术、多媒体技术、成熟的网络交互协议与传统教育学的优势, 改变了传统面对面的教学模式, 实现了教育资源的共享利用, 同时, 打破了时空及地域限制, 延伸至全国乃至世界, 实现教育资源的共享和优化。

现代远程教学系统多是基于 H.323 协议实现的, 笔者在远程教学系统的开发实践中, 应用了虚拟摄像头技术, 使用内核模式中的设备驱动, 虚拟了硬件设备, 对应用程序尤其是 H.323 协议提供了相当完备的支持, 从各方面改善了普通远程教学系统的性能, 使系统运作更高效, 结构更清晰, 在实际中得到了广泛的应用。

1 系统简介

1.1 虚拟摄像头简介

虚拟摄像头是没有硬件设备的摄像头, 它在操作系统中仅仅作为一个摄像头驱动程序存在, 但与一般驱动程序不同的是, 它将应用程序源虚拟成了硬件源。

由于系统对硬件的操作是调用驱动程序来实现的, 对处于用户模式的摄像头捕获程序而言, 它无法直接操作处于内核模式的驱动程序, 而是通过操作系统调用摄像头驱动程序实现, 因此对摄像头捕获程序而言, 虚拟摄像头与真实摄像头并没有区别, 可以使用同样的摄像头捕获方法去捕获虚拟摄像头中的数据, 而读出的数据即编程者指定的应用程序源提供的数据。

上述的应用程序源就是基于 H.323 协议的应用程序, 与普通远程教学系统不同的是, 本系统采用了特殊的二层结构, 分散了系统的负载, 结构更清晰。由于虚拟摄像头的支持, 本系统实现了功能上的扩展和性能上的提升。

1.2 远程教学系统两层体系结构

本文的远程教学系统构架如图 1 所示。终端之间仅通过服务器进行交互。根据交互数据性质的区别, 把交互数据分为两类: 一类是信令及音视频数据; 另一类是资源数据。基于分散负载、提高效率的考虑, 笔者提出了一种信令和数据传输分离的二层体系结构。

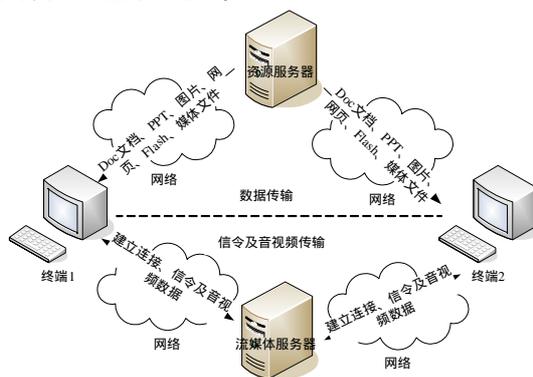


图 1 远程教学系统构架

信令和音视频传输部分采用国际通用的 H.323 协议实现, 称为 H.323 信令模块。所有终端在教学开始时使用 H.323 信令模块与服务器建立连接, 在教学过程中连接一直保持, 教师将教学命令发送给流媒体服务器, 流媒体服务器将教学命令向所有师生转发, 音视频数据也通过这种方式传输, 实现所有终端间信令和音视频的交互。

作者简介: 袁 聃(1982 -), 女, 硕士研究生, 主研方向: 多媒体网络通信; 程文青, 副教授; 吴 砥, 讲师

收稿日期: 2006-10-30 **E-mail:** Y_D_2000@sohu.com

资源数据使用资源服务器中转,教学开始前,教师将所有要用到的教学资源上传至资源服务器,教学过程中教师使用 H.323 信令模块向所有学生发出获取某项资源的命令,学生收到命令后自动从资源服务器获取指定资源。

使用这种信令和数据分离的两层体系结构,可将资源数据的传输转化成获取资源的命令的传输,减少了传输的数据量,不仅将负载从终端转移到服务器,也将数据量比较大、管理策略完全不同的两种数据:资源数据和音视频数据,分别由两个服务器负载,实现了数据量的分散、数据管理的分类集中。

1.3 虚拟摄像头与两层体系结构的结合

本系统为了发挥两层体系结构的特色,根据信令及音视频传输层和数据传输层不同的特点,对这两层采用了不同的处理方式。但是在具体实现过程中,系统性能和功能方面都受到了一定的限制。因此,本系统尝试使用虚拟摄像头打破了系统性能和功能上的限制,发挥了两层体系结构的优点,对两层体系结构提供了良好的支持,使系统得到了一定的扩展。

2 虚拟摄像头应用特点

2.1 支持两层体系结构

由于信令和音视频部分的传输要求极高的数据处理效率,因此H.323 信令模块使用了非托管的C++语言实现^[1],而托管代码在处理效率上虽不及非托管代码,但它具有跨语言集成、开发效率高等方面的优势,故资源数据传输和视频显示部分使用托管代码实现,由非托管代码向托管代码提供底层的服务支持。

上述的两层编码结构确保了系统的稳定性、安全性和简便性,但由于托管代码和非托管代码在内存管理方面的机制完全不同,因此长时间在两层之间传输大数据量的数据会使内存管理者频繁更替,除了在处理速率方面会受到影响,更易引起内存管理的混乱,导致系统的不稳定。使用非托管代码编写 H.323 信令模块从网络获取视频数据,需要使用托管代码编写的视频显示模块显示,其中必然涉及到托管代码和非托管代码之间大数据量的传输问题。

对于这个问题,笔者提出了一种新的解决思路,用户模式出现的问题,可以通过内核模式解决。解决方式如下:编写非托管代码接口,由用户模式运行的非托管代码使用 DirectShow Filter 机制向内核模式运行的虚拟摄像头写入数据;用户模式的托管代码则利用 DirectShow 提供的托管代码接口直接捕获虚拟摄像头中的数据。使用这种方式后,非托管代码和托管代码之间不再需要大数据量的传输,也不会再出现内存管理方面的问题。

2.2 支持设备重用

本系统需要课程录制功能,从摄像头获取的数据不仅发往网络,也要在通过 Windows Media Encoder 接口本地保存。Windows Media Encoder 具有很好的封装性,它能从硬件源获取数据,但不支持网络源。这就带来了一个问题:摄像头设备是独占性设备,H.323 信令模块的视频捕获程序已经占用了摄像头读取本地视频数据,Windows Media Encoder 将不能同时再次占用。

使用虚拟摄像头可以实现设备重用。方法如下:在 H.323 信令模块中开辟一块缓冲区,缓冲区大小为 1 帧视频图像的数据量大小。H.323 信令模块从摄像头捕获到视频数据后即存入此缓冲区,将缓冲区的数据写入虚拟摄像头中之后将之

发至网络。通过微观上对缓冲区进行串行操作,宏观上将一份数据并行地发向两个地方,同时完成了对摄像头视频的网络发送和本地录制,解决了设备的独占性问题,宏观上达到了设备重用的效果。

2.3 支持代码的模块性封装

在虚拟摄像头模块实现视频显示和录制部分对视频图像的处理,一方面简化了应用程序代码,另一方面使代码的模块清晰,封装性和复用性好。

3 虚拟摄像头在系统中应用的实现

3.1 DirectShow 技术与虚拟摄像头的结合

为了提高系统的稳定性,Windows操作系统对硬件操作进行了隔离,应用程序一般不能直接访问硬件。设备驱动处于内核模式,可以用DirectShow技术解决处于用户模式的应用程序向内核模式中的设备驱动写入数据这个问题^[2]。

DirectShow是微软公司提供的一套在Windows平台上进行流媒体处理的开发包。它基于模块化,每个功能模块都采取的是COM组件方式,称为Filter。DirectShow Filter工作在用户模式,而硬件工作在内核模式。为实现交互,DirectShow为采用WDM驱动程序的硬件设计了KsProxy Filter(如图2所示)。这种Filter跟其他普通的DirectShow Filter(例如,自定义Filter)处于同一个级别。它通过Stream Class控制硬件的驱动程序minidriver(在系统中对应虚拟摄像头驱动程序)。

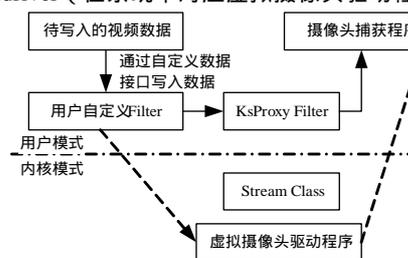


图2 远程教学系统中应用程序向硬件驱动写入数据原理

在本文系统中,DirectShow 技术不仅用于摄像头的捕获显示,也应用于向虚拟摄像头驱动中传送数据,由于DirectShow可以自定义Filter,而自定义的Filter可以留出写入数据的接口,因此,在本系统中实现由应用程序向驱动程序写入数据的原理如图2所示。准备传输数据时,调用DirectShow提供的API函数将自定义Filter和Ksproxy Filter连接起来;传输数据时,调用自定义Filter的数据接口,向自定义Filter中写入视频数据,视频数据将从自定义Filter传至Ksproxy Filter。绑定了虚拟摄像头设备的摄像头捕获程序,使用DirectShow机制,通过Ksproxy Filter从虚拟摄像头驱动中获取数据,即可得到从自定义Filter传入的数据。对于处于用户模式的摄像头捕获程序而言,即捕获了由用户模式传入内核模式的数据。

在自定义Filter中可以加入视频处理模块,如图2所示,摄像头捕获程序(远程教学系统程序)将获取经过处理后的视频图像,在系统运行过程中可随意更换视频处理形式,体现了较好的封装性、复用性和模块性独立。

3.2 Openh323 协议库与虚拟摄像头的结合

如第2节所述,本系统中虚拟摄像头有2个应用:(1)将从网络中获取的视频数据写入虚拟摄像头供虚拟摄像头捕获程序读取显示;(2)将本地视频数据读取到H.323信令模块的缓冲区中,将缓冲区中数据写入虚拟摄像头中,供Windows Media Encoder捕捉录制。

3.2.1 数据传输的准备和清理

如上所述,在数据传输时,需要通过自定义 Filter 向虚拟摄像头驱动中写入数据,因此,数据传输前,应使用 DirectShow 机制完成初始化工作:(1)初始化 DirectShow 提供的 COM 接口;(2)始化虚拟摄像头,将自定义 Filter 与虚拟摄像头驱动连接起来;(3)自定义 Filter 的初始化,主要包括自定义 Filter 中默认图片的初始化以及一些视频处理功能的初始化工作。在具体应用中根据功能需求的不同,上述第(3)步可以忽略。

在数据传输结束时需要相应的释放占用的资源。

3.2.2 视频数据在基于 H.323 的远程教学系统中的处理

openh323 系统^[3]是澳大利亚的Equivalence Pty Ltd 公司组织开发的,该协议库完全符合H.323 协议,且功能完善、技术成熟,因此,现代远程教学系统一般使用openh323 协议库提供对H.323 协议的支持。

本系统使用 openh323 协议库实现H.323 信令模块,openh323 运行时,可针对不同的数据建立不同的传输通道^[4]。下面介绍 openh323 协议库如何收发媒体数据,以及如何在 openh323 协议库中添加虚拟摄像头数据接口。

(1) Openh323 发送和接收视频数据

开始传输视频数据时, H323Channel 的子类 H323 UnidirectionalChannel 调用 Start 方法,启动了一个后台线程 H323LogicalChannelThread,这个线程用于接收和发送媒体数据。该线程调用 H323_RTPChannel 中的发送(transmit)和接收(receive)方法来传递媒体数据。

发送媒体数据时使用的是 H323_RTPChannel 类中 Transmit 方法:首先调用 H323_RTPChannel 类中的 Read 方法,从 Codec 中读出一帧媒体数据,再调用 WriteData 将该数据包封装成标准的 RTP 分组,写入 UDP 套接字传送到对端主机。

接收媒体数据时使用的是 H323_RTPChannel 类中的 Receive 方法:首先调用 H323_RTPChannel 类中的 Read BufferedData 方法从网络读取一帧媒体数据,再调用 Write 方法将该数据写入 Codec。

(2) 在 openh323 协议库中添加虚拟摄像头数据接口

如上文所述,H.323 中读取或写入数据都是通过Codec完成,Codec是编解码器^[5],与PChannel绑定。PChannel是定义了IO通道的抽象类。PVideoChannel是PChannel的子类,用于从指定IO通道中读取或写入视频数据。视频Codec绑定的即是PVideoChannel。

当发送视频数据时,如图3所示,H323_RTPChannel 类中的 Read 方法,从视频 Codec 中读出一帧视频数据,视频 Codec 通过 PVideoChannel 的 Read 方法从绑定的视频输入设备中读入视频数据至缓冲区 buf,对该缓冲区中数据进行编码,然后返回给 H323_RTPChannel 类的 Read 方法。同样,接收视频数据时,如图3所示先使用视频 Codec 对发送给 Codec 的数据进行解码,然后存储到缓冲区 buf 中,PVideoChannel 类中的 Write 方法将存储在缓冲区 buf 中的数据写入绑定的视频输出设备显示。

如上所述,PVideoChannel 与视频设备存在着直接的联系,从 IO 通道中读取或向 IO 通道中写入视频数据都要通过 PVideoChannel 类,虚拟摄像头数据接口是 IO 通道的一种,因此,需要在 PVideoChannel 类中对虚拟摄像头数据接口进行操作。

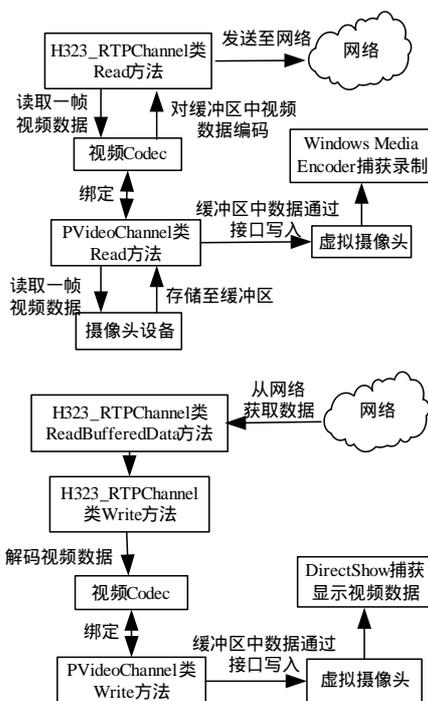


图3 发送、接收视频数据时虚拟摄像头的应用

PVideoChannel 类中的 Read 方法是从 IO 设备中读取视频数据,因此,如图3所示,应将 Read 方法 buf 缓冲区中的数据通过虚拟摄像头数据接口写入虚拟摄像头,供 Windows Media Encoder 对本地视频进行捕获录制。同理,应将 Write 方法 buf 中的数据通过虚拟摄像头数据接口写入虚拟摄像头中,供 DirectShow 捕获显示网络接收的视频数据。

4 结束语

视频交互是远程教学系统中重要的组成部分。本文基于信令和数据传输分离的两层体系结构,提出了利用虚拟摄像头显示和录制视频数据的方案,分析了利用 DirectShow 机制操作虚拟摄像头的原理,阐述了在 H.323 系统中加入虚拟摄像头模块的可行性和具体实施方法。应用虚拟摄像头,实现了对本系统两层体系结构的充分支持,也达到了视频捕获设备的重用的效果。经过验证,证实了这一方案的可行性,本系统已在实践中投入使用。但本系统仍存在着一定的不足,目前仅实现了对虚拟摄像头的支持,并未扩展到其他虚拟设备,虚拟摄像头的图像处理功能也并未进一步扩展。随着远程教学的进一步发展,在后续的研究中将进一步考虑:(1)其他虚拟设备在本系统中的应用。(2)虚拟摄像头模块中视频处理部分的扩展功能。这样将进一步扩展本远程教学系统的功能,使本系统更加完善。

参考文献

- 1 ITU-T-1998 Recommendation H.323 Packet-based Multimedia Communication Systems[S]. 1998.
- 2 陆其明. DirectShow 开发指南[M]. 北京: 清华大学出版社, 2003.
- 3 Openh323 Project[EB/OL]. (2006-06). <http://www.openh323.org/>.
- 4 卢政. 如何成功的运用 OPENH323 来开发商业的 H.323 协议栈[EB/OL]. (2002-11). <http://www.zdnet.com.cn/i/developer/story/200211/39098573/Openh323-RTH323.pdf>.
- 5 Equivalence Pty Ltd. PWLib Source Code[Z]. (2006-06). <http://www.openh323.org/code.html>.