

基于 TCP/IP 协议的远程多媒体 视频会议系统

葛广英 (山东聊城师院通信工程系 252059)

摘要: 文章介绍了基于 TCP/IP 协议的远程多媒体视频会议系统的构成、工作原理和相关技术, 并对数字视频编解码技术和网络编程进行了详细探讨。

关键词: 多媒体 视频会议系统 TCP/IP 协议 编解码技术

结合我国目前计算机通信网络普及率和传输率低而有有线电视 CCTV 的电缆已遍及乡镇、普及率高的特点, 我们研制开发了符合我国国情的基于 TCP/IP 协议的多媒体视频会议系统, 它不仅使与会者节省了时间和人力物力, 又可使与会者看到和听到会议参加者, 共同面对面地商讨和研究问题, 使每一位与会者有身临其境的感觉, 有利于双方的交流。本文介绍了它的构成、工作原理以及主要功能, 并对视音频编解码技术和网络编程技术等相关技术进行了详细探讨。

1 视频会议系统结构

多媒体视频会议系统的结构如图 1 所示, 它主要由远程视频会议终端——分会场、主会场、控制中心、控制管理软件以及多媒体信息传输网络等组成。

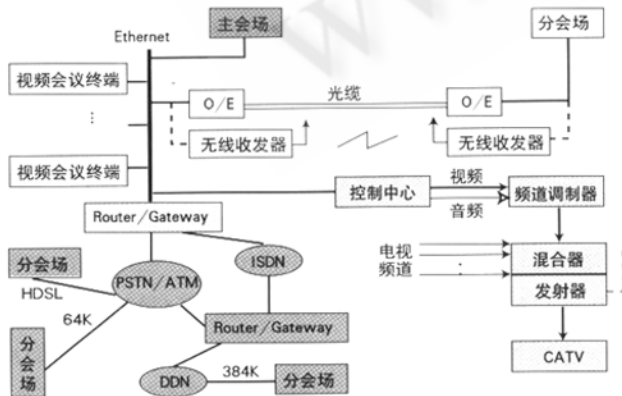


图 1 多媒体远程视频会议系统结构框图



主会场或控制中心主要由多媒体控制主机、视音频信息接收解压卡和相应的各种控制、处理、记录、显示和输出设备, 主要完成网络上视音频压缩数据的接收、解压、显示、输出以及系统监控管理等功能。管理人员通过界面友好、操作简单方便的计算机控制台, 实现(1)实时处理和记录由视频会议终端现场传来的视频、音频、文字以及各种数据信息; (2)对网内的远程视频会议终端现场发出各种操作和控制指令, 以完成在现场所能完成的各种操作; (3)根据会议要求选择主会场或某一分会场的信息作为会议的主信息通过网络向其他分会场传送。而对于哪些无网络接口的分会场或单位, 可将会议信息送入电视频道调制器, 经调制变换成某一电视频道的高频调制信号, 和其他频道的电视信号一起进入前端混合器, 经天线发射器发射出去或进入有线电视网 CATV。

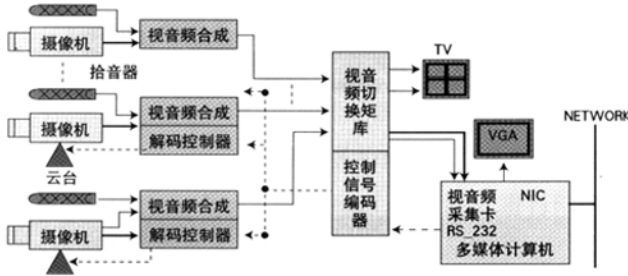


图 2 远程视频会议终端系统结构框图

远程视频会议终端——分会场主要由多媒体计算机、视音频信息采集压缩编码卡、网络适配器(NIC)或Modem以及摄像机、语音采集设备,编解码控制部分和驱动执行机构等组成,如图2所示。

分会场完成视频会议现场视频图像和音频声音的实时捕获、实时压缩编码处理和显示输出,以及控制摄像机的各种动作,将符合国际标准的压缩码流经网络接口设备送到通信网络上,通过一种或多种类型的通信网,将这些信息传送到主会场、控制中心或其他分会场。此外,远程视频会议终端还要接收控制中心或主会场发出的各种指令,根据指令控制摄像机的动作,采集对应的视音频信息,经压缩后通过网络进行上传。

2 数字视频图像的编解码技术

视频图像的数字压缩编解码,实质是在一定质量(信噪比要求或主观评价得分)条件下,以最少比特数来传送图像,有效地减小图像的数据量,降低视频图像实时性对网络带宽的要求,因此必须对其进行压缩编码。目前已经形成的数字视频编解码标准主要有 MPEG—I、H.261 和 H.263 等。

2.1 MPEG—I 图像编解码标准

MPEG—I 图像编解码标准即运动图像的编码方法主要有两类:一类是无损压缩编码,如哈夫曼可变长度特征编码、LZW 编码以及游程编码 LC 等,这类编码技术不丢失任何信息,但压缩比较低;另一类是有损编码,其方法包括:预测、变换、编码。目前运动图像的视音频信息编码技术采用的是预测—离散余弦变换法(DCT),即以静态图像编码为基础,加上运动估值与运动补偿等,所形成的运动图像编码技术,在 120 倍的压缩比下,由数据流恢复的图像质量仍很高。在 MPEG—I 中利用预测法——消除

前后两帧之间的时域冗余,利用离散余弦变换法(DCT)——消除图像空间域的冗余,再经变字长编码(即对出现概念大的像素以短码,对出现概念小的像素以长码),对数据进一步压缩,适合于在 LAN 网上传输运动图像。

本系统采用美国 Digital 公司的 21230-AA 芯片来实现 MPEG—I 的实时编解码。21230 芯片按照 MPEG—I 和 H.261 标准,实现实时视频编解码。利用 21230 芯片和外围电路构成 PCI 总线的计算机运动图像压缩编解码卡,和 CPU 协同工作完成 MPEG—I 的编解码功能。21230 芯片可以和计算机 PCI 接口实现无缝连接,完成所有基于像素点如 DCT 变换、量化、运动估计和补偿等的操作,另外还有音频压缩单元 Analog Devices 2181 实现对音频部分的编解码。

2.2 H.261 图像编解码标准

H.261 图像编解码标准是国际电话电报咨询委员会(CCITT)制定的 $P \times 64$ Kbps 运动图像编码协议,是一个网络视频图像传输的编解码标准。H.261 采用帧内正交编码(DCT)、帧间差分编码和变字长编码。帧内正交编码采用了离散余弦变换,消除了空域冗余;帧间差分编码,采用了运动补偿、帧间预测方式,消除了时域冗余,最后再经变字长编码。H.261 可使数据速率压缩至 $P \times 64$ Kbps ($P=1 \sim 20$),在 32~384Kbps 时图像可达 CIF 或 QCIF 格式 15 帧/秒的图像质量。

H.261 编码大幅度地降低了数据流的瞬时变化,与 MPEG—I 相比,图像质量略逊于 MPEG—I,但编码后的数据流速率更低,适合于在 ISDN 网或 DDN、PSTN 网上传输运动的图像。

2.3 H.263 图像编解码标准

为了在普通电话线传输运动图像信息,CCITT 制定了比特率更低的视频编码方案——H.263 标准,该标准基于 H.261,考虑到更低比特率的特点,增加了四个通用的编码选项:①不限制运动矢量模式;②以语义算法编码模式替代哈夫曼编码(变字长编码),减小视频的码率;③采用高级预测模式减小方块效应;④ PB 帧格式。在考虑这四种编码选项后,可将图像数据最低编码到 20Kbps,通过电话线上以 22.8Kbps 的 V.34 Modem 传输,图像质量达到 176×144 或 128×96 分辨率下 5~15 帧/秒的水平。

上述 MPEG—I、H.261、H.263 三种编码方式都是针对低成本应用的编码方案,采用去除时域冗余和空域冗余的方法,尽可能的降低码流速率,在低带宽的信道上,实现运动图像的传输。

3 多媒体网络通信技术

3.1 多媒体网络通信协议

目前, 由于 Internet 网具有广泛的应用基础, 为视频会议系统将来在因特网上的拓展奠定基础, 本系统采用 TCP/IP 协议实现异种网的互联。TCP/IP 协议为不同硬件体系结构和不同操作系统计算机所组成的网络提供了互相通信的能力, 具有广泛的通用性。在 TCP/IP 协议组中, 与 ISO 的 OSI 七层参考模型中传输层相对应的是 TCP 和 UDP(User Datagram Protocol, 用户数据报协议)两大协议, 二者都是建立在网络层 IP 协议基础上的高层协议。TCP 提供面向端到端的连接服务, 是有序且可靠的, 而 UDP 提供的是无连接不可靠的数据服务, 无需对方作出确认, 不进行重传, 处理过程比较简单, 传输延迟小, 适合实时性要求很高的多媒体通信。

3.2 Winsock 编程实现多媒体信息网络传输

本系统中采用 Client/Server(客户机/服务器)模型, 建立可靠的网络传输, 编程采用 Winsock 技术实现网上传输。Winsock 是 Microsoft 公司推出的网络编程的标准, 它运行于 Win9X 平台上, 是一个基于 Socket 模型的 Windows API 集, 用它可实现 Windows 环境下 TCP/IP 网络传输, 实现在 Internet 传输数据和交换信息, 是开发网络应用程序常用的技术。Winsock 技术基于 C/S 模式, 客户端和服务端是通过套接字 Socket 进行通信的, 对套接字的读写就完成了信息的传输。

Winsock 编程有多种类型, 包括面向连接的、无连接的、阻塞方式、非阻塞方式等, 本系统采用 UDP 的非阻塞方式进行连接, 即将客户端 Socket 的 ClientType 特性设为 ctNonBlocking(非阻塞方式), 服务器端 Socket 的 ServerType 特性设为 stNonBlocking。当客户端 Socket 试图进行读或写时, 服务器端的 Socket 就会得到通知, 产生 OnClientRead 事件或 OnClientWrite 事件; 当服务器端的 Socket 试图进行读或写时, 客户端 Socket 就会得到通知, 产生 OnRead 事件或 OnWrite 事件。执行 MPEG-1 编码的计算机作为网络客户端, 提供视频压缩数据; 执行 MPEG-1 解码和显示的计算机作为网络服务器端, 得到视频压缩数据。利用 Delphi 5.0 提供的 TserverSocket 服务器套接字和 TclientSocket 客户端套接字实现网络传输。

公司的 Delphi 5.0 作为开发工具进行的二次开发。Delphi 5.0 以其使用方便、功能强大深受广大程序员的喜爱, 它提供了一套 Internet 组件, 使开发网络通信程序更加方便、快捷。在这套组件中, 包括 TNMUDP 组件, 这个组件是专门用于 UDP 通信的。TNMUDP 不区分服务器端和客户端, 只区分发送端和接收端, TNMUDP 既可以放在发送端, 也可以放在接收端, 两个 TNMUDP 也可以放在同一个应用程序中。本系统中的视频图像数据、音频数据以及摄像机动作的控制数据就是通过它进行传输的。

在图 1 所示中, 各个多媒体远程视频会议终端现场的计算机由视音频采集卡与摄像机相连, 负责采集并以 MPEG——I、H.261 和 H.263 方式压缩图像声音数据, 然后以 UDP 数据报的形式向控制中心、主会场或其他分会场传送。控制中心负责接收各个会议现场发送来的 UDP 数据报, 并将 UDP 数据报组装还原成图像声音数据, 最后解压显示图像和播放声音。

4.1 远程视频会议终端——分会场软件的设计

远程视频会议终端的软件主要由视音频数据采集压缩模块和视音频数据传输模块组成。现场视频会议终端捕获的图像和声音, 经压缩后通过 Sender 控件向网上传送, 通过 MsComm 控件向串口 COM 写数据, 发送控制指令, 其程序流程如图 3 所示。

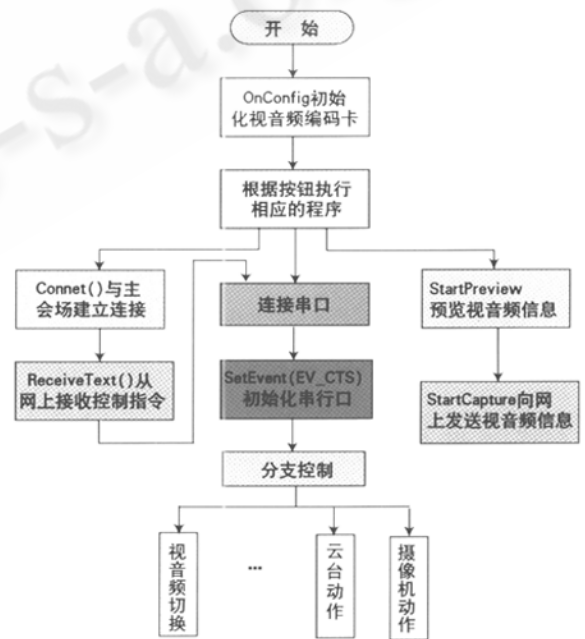


图 3 远程视频会议终端程序流程

4 控制管理软件的设计

本系统的软件是在 Windows 98 环境下, 选用 Inprise

4.2 主会场或控制中心软件的设计

主会场或控制中心的控制软件负责接收 UDP 数据报并还原成完整的图像数据,然后解压显示图像。要使控制中心的服务器监听客户端的连接请求,首先要设置 TServerSocket 的 Port 特性来指定端口号。指定了端口号后,可以调用 TServerSocket 的 Listen() 来开始监听。程序中将 Active 特性设为 True,则服务器程序启动的时候就会自动进入监听状态。服务器进入监听状态后,当客户端提出连接请求时,服务器端自动接收请求,然后建立连接并触发 OnClientConnect 事件,程序可以在此事件的处理函数中读取图像数据。为了使控制中心能及时接收所有的数据包,该事件处理函数应尽可能简单、迅速地完成。另外,图像的解压显示功能应在另一个线程中完成。程序流程图如图 4 所示。

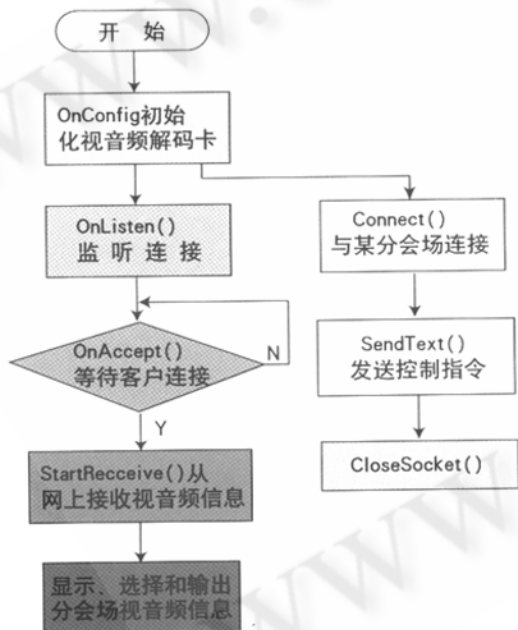


图 4 视频会议控制中心程序流程

如果要断开连接,只要在服务器端调用 Close() 或者把 Active 特性设为 False。服务器端对所有客户的连接都将断开,并退出监听状态。如果在客户端调用 Close(), 将只断开该客户与服务器的连接,不影响其他客户端的连接。

5 多媒体视频会议系统的主要功能。

多媒体视频会议系统以 LAN/WAN、ISDN、DDN、ATM、PSTN 等多种通信网络为基础,采用网络化的管理手段,使数据、报文、图表、视频、音频以及控制数据等多种媒体信息在网络上传输和交互使用,共享资源。其主要功能为:

5.1 远程视频会议终端

(1) 采用 MPEG—I 或 H.261、H.263 编解码标准,对视音频图像进行实时编码和回放,在 9.6kbps~2Mbps 的速率下,图像分辨率达 CIF、QCIF 格式 1~30 帧/秒。

(2) 在网络传输中采用 TCP/IP 协议, TCP/IP 传输数据通信控制信息, UDP/IP 传输视音频信息。

(3) 可手动或自动切换 1—64 个摄像点,视音频可同步切换或异步切换,可实时将分会场信息显示在监视器上,并通过网络上传;

(4) 通过鼠标操作,实现对现场摄像机的上、下、左、右等动作的控制,以及对摄像机镜头的光圈大小、变焦的远近、聚焦的虚实等进行调整;

5.2 视频会议控制中心或主会场

(1) 网络管理,适应各种通信网络和传输链路,可根据需要将控制系统设在主会场或广播电视控制中心;

(2) 通过网络可实时接收远程视频会议端——分会场传送的视音频信息和其他数据,通过 IP 地址和端口号区分不同的分会场;

(3) 可实时显示与分会场的连接状态,向分会场发出的各种控制指令,控制摄像机的各种动作;

(4) 控制中心根据会议要求将主会场或某分会场的信息转播给各分会场;

(5) 多媒体信息尤其是视音频信息,用硬盘或可写光盘等设备进行记录,并回放检索;

(6) 管理与控制更加自动化、智能化,各种信息的交流更加合理、方便,人机界面更加友好。■

参考文献

- 1 钟玉琢, 多媒体技术(高级), 1999, 7.
- 2 李智渊等, 高速网络技术及其应用, 1997, 12.
- 3 黎连业, 计算机网络基础和工程, 1999, 2.
- 4 徐新华, Delphi 5 高级编程, 2000, 5.