

文章编号: 1005-0523(2002)03-0058-04

网上语音数据传输的 Delphi 实现技术

蒋先刚, 涂晓斌, 陈海雷

(华东交通大学 基础科学学院, 江西 南昌 330013)

摘要: 介绍并分析语音数据在网上的传递技术和应用, 主要介绍用 Delphi 在网上开发语音传输程序的关键技术和技巧, 详细介绍音频压缩管理器控件的应用。

关键词: 音频压缩管理器; 嵌套字控件; 数据流

中图分类号: TP391.6

文献标识码: A

0 引言

语音信息的交流为网上教学提供一个更生动的媒体, 在公司局域网上对设计产品进行语音形式的讨论和聊天, 从网上实时选择电台广播和收听 MP3 音乐, 通过很少的编码就可设计出自己的语音复读机, 这些涉及到语音数据处理的程序设计问题。语音信息作为人们交流的媒体, 表现出亲切、可信和交互性强的特点。在 Internet 日益发展的今天, 人们正借助于网络浏览信息、交换数据和商业交易。而目前的网上信息的交换基本是基于文字形式的, 图片形式信息的交换量的规格和实时性仍受网上带宽的限制, 而语音媒体在 Internet 上的应用呈现着诸多的优越性和信息量巨大的技术性问题。从程序员的角度看, 文字、图片、动态图象及语音都是数据流, 只要合理调用开发环境中发布流和接受流的控件就可上传递和获取这些多媒体信息。而网上信息的传递受流量的制约, 网络程序跟工业监控程序一样表现出一定的实时性。如我们选择的被传递的语音数据格式为双声道的 PCM 数据流, 其数据流为 16KB/S, 如果网络比较拥挤, 每个用户只能拥有 4KB/S 的流量, 则我们听到的实时点播的音乐或者电话是断断续续的。因此比较实用的 Internet 电话

程序必须具有良好的兼顾各种语音质量的语音数据压缩程序段。语音数据的压缩比是 Internet 电话实现的关键参数。

1 音频压缩管理器控件的应用

Windows 提供了一定的音频处理和音频压缩的 API 函数。Codecs 是音频压缩的解码编码器, 它提供对媒体格式进行转换的功能, Windows 带有其相应的动态库。其相应的调用 Codec 功能的接口函数库 ACM(Audio Compression Manager)提供给用户对媒体格式进行转换的快捷工具。在 Delphi 开发音频压缩的程序中, 其相应的 API 声明单元为 MSACM.pas。语音压缩的程序中, 音频压缩管理器控件组包括 TACMConvertor、TACMIn 和 TACMOut 这三个控件。程序员可在相应的网站获得这些控件。TACMConvertor 控件可以转换两个不同媒体格式的音频数据, 该控件可以用来指定 ACM 流的输入输出格式, 用户既可在设计过程中设定媒体的输入输出格式, 也可在程序运行时由对话框调整这些参数。TACMIn 控件用于从麦克风接受音频数据, 一般用 PCM 格式的音频流来记录其数据。TACMOut 控件用于声音的回放, 回放也一般用 PCM 格式的音频流, 这个控件的音频数据缓冲区大小 NumBuffers 属性用于指定播放前使用

收稿日期: 2001-12-17

基金项目: 获 2001 年国家留学基金资助, 编号 21836011

作者简介: 蒋先刚(1958-), 男, 湖南永州人, 华东交通大学教授。

的缓冲区数,这对于流量不稳定的网络的音频传输显得非常重要,它可以缓冲连接的波动数据而输出稳定的不间断的声音.

处理麦克风音频数据的控件 TACMIn 一般获取 PCM 格式的数据,而回放音频数据的控件 TACMOut 也一般采用 PCM 格式播放,而 PCM 格式的双声道音频流达 16KB/S,如果只用这两个控件,这样相对于传统的 Windows API 音频传递与播放程序设计来说,它提供的新的特性就只是音频数据的缓冲功能而使获取和播放的声音不间断,但网上传递的音频数据仍偏大,因此要想在 Internet 上实现电话和音频广播功能,就必须对音频数据在上网前进行压缩,而从网上接受到这个压缩的音频数据后将其解压而变成声卡可播放的音频数据.实现音频数据的压缩和解压功能的控件是 TACMConverter.它提供将转换的音频数据的输入输出格式的选择,这些音频包括 PCM、Microsoft G. 723. 1、Lernout & Hauspie、VivoActive 和 Voxware 等 10 多种格式.我们选择的压缩和解压的流程为:麦克风音频数据 PCM 8 位单声道(8KB/S) > TACMConverter > GSM 8 位单声道(2KB/S) > Internet > TACMConverter > 喇叭音频数据 PCM 8 位单声道(8KB/S).这样在网上进行的音频广播和打电话时,只需 2KB/S 的流量就可传送和听到不间断的音频信号了.

2 网上音频数据传输的软件技术

2.1 用用户数据报文协议控件实现语音数据的传

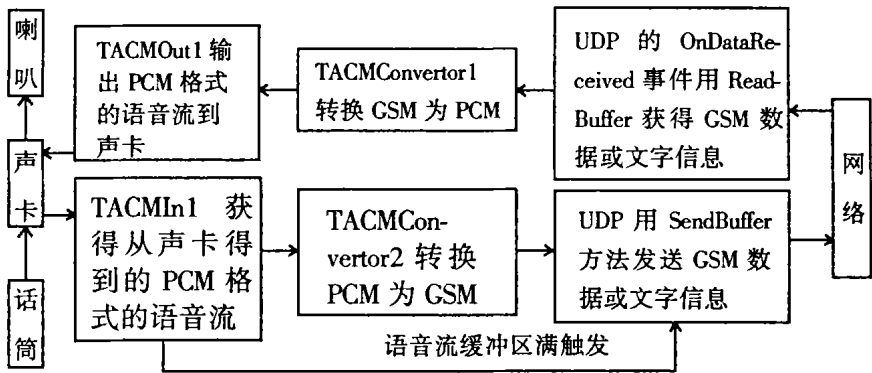


图 1 用用户数据报文控件进行音频数据传输的程序框图

2.2 用嵌套字控件进行音频数据的传输

用嵌套字控件进行音频数据传输的程序框图如图 1 所示,在程序表单中引入 TACMIn1 用于麦克风音频数据的获得,TACMOut1 用于从网上获取的

输

用户数据报文协议 UDP(User Datagram Protocol)与 TCP/IP 协议不同,它作为一个无连接的协议,只是简单地将数据放在端口中,UDP 不在发送和接受方之间建立一个连接.这种数据传输方式的安全性较差些,但程序运行过程的鲁棒性要强许多.我们使用 TNMUDP 控件来传输音频数据或文字信息,发送报文的任务由方法 SendBuffer 或者 SendStream 完成,它负责完成由 TACMIn1 控件获取的并经 TACMConverter2 控件转换的压缩音频数据的发送.而在它接受到一个 UDP 报文时,将触发 OnDataReceived 事件,在此事件函数中,我们调用 ReadBuffer 或者 ReadStream 来读取报文的数据,并将这些数据经 TACMConverter1 解压后由 TACMOut 控件加以播放.考虑到保留文字信息交流的必要性,在程序表单中我们仍然放入发送文字信息和显示文字信息的文字编辑控件 TMemo,其文章信息的发送与获得仍用 TNMUDP 用来发送与获得音频数据的方法和事件.如果发送信息和接受信息的计算机 IP 同为一机的话,这时的程序起着语音复读机的作用.用用户数据报文控件进行音频数据传输的程序框图如图 1 所示.

音频数据的播放,而 TACMConverter2 用于上网前将音频数据进行压缩,而 TACMConverter1 用于将网上获得的音频数据的解压.程序表单放入了客户嵌套字控件 ClientSocket1 和服务端嵌套字控件 Server-

Socket1. 每一个语音程序即包括客户程序段又包括服务器程序段,这两个控件同时具有发送和接受数据流的方法和事件. 如果服务器的客户地址为本机的话,这时的程序起着语音复读机的作用. 作为服务器端时,它可具有多个客户端,即这时可作为音频数据的广播站,可让多个与你相连的用户听到你的声音,这在网上多媒体教学中可作为教师的指导语音播放;作为用户端,你需发送你的服务器的 IP 地址并连接之,传递你的语音数据到对方. 作为更加准确的信息传递,程序保留了文字信息交流的功能,客户嵌套字控件和服务器嵌套字控件能够同时接受到语音数据和文字信息,程序靠分析数据流的长度和类型而分别去进行音频数据的解压或者文字数据的显示. 当麦克风音频数据缓冲区满时,其触发事件就将收集的音频数据发送出去. 其程序段如下:

```
Procedure TAudioTalkMain.ACMin1BufferFull (Sender: TOb-
ject: Data; Pointer:
Size: Integer); //麦克风音频数据缓冲区满的事件函数
var
NewSize : Integer;
```

```
begin
Move(Data^,ACMConvector2.BufferIn^,Size);
NewSize := ACMConvector2.Convert;
//将 PCM 转换为 GSM 格式
TransmitAudioData ( ACMConvector2.BufferOut, NewSize );//
调用数据发送程序 end;
Procedure TAudioTalkMain.TransmitAudioData ( Buffer :
Pointer; Size: Integer);
Var //音频数据发送程序段
I: Integer;
begin
if IsServer=True Then
begin //程序处在服务器状态就向连接的多个客户广播
其音频数据
for I:=0 to ServerSocket1.Socket.ActiveConnections -1 do
ServerSocket1.Socket.Connections [I].SendBuf ( Buffer ^,
Size);
end;
if IsServer=False then//程序处在客户状态就向其服务
器发送其音频数据
ClientSocket1.Socket.SendBuf (Buffer^,Size);
end;
```

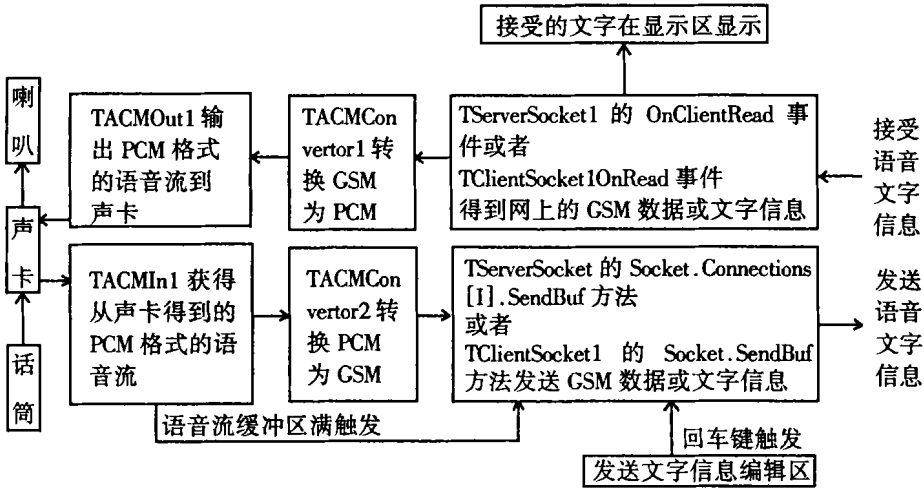


图 2 用嵌套字控件进行音频数据传输的程序框图

服务器嵌套字控件和客户嵌套字都具有获取网上数据的能力,其程序段的源码是一样的,下面给出服务器嵌套字控件获取音频数据或者文字信息并播放或者显示的程序段:

```
procedure TAudioTalkMain.ServerSocket1ClientRead (Sender:
TObject;
Socket: TCustomWinSocket);
Var//获取并播放或显示音频数据或文字信息的程序段
NewData: TPointer;
NewDataSize: Integer;
```

```
begin
if ReceiveMemo.Lines.count>200 then
ReceiveMemo.Lines.Clear;
//显示区接受的内容过多就清空
Status('接受到' + IntToStr(Socket.ReceiveLength) + '字
节');
if Socket.ReceiveLength<=1 000 then
//判断是文字信息还是语音数据
begin//文字信息在显示区显示
ReceiveMemo.lines.add (socket.RemoteAddress + ',' +
```

```
socket.ReceiveText);
End else begin
with Socket do begin //处理音频数据程序段
NewDataSize := ReceiveLength;
//实际获得的音频数据的大小
GetMem(NewData, NewDataSize);
try
ReceiveBuf(NewData^, NewDataSize);
if not ACMOut1.Active then begin
ACMOut1.Open(ACMConvector1.FormatIn);
end;
ACMOut1.Play(NewData^, NewDataSize);
//播放音频声音
finally
FreeMem(NewData);
end;end;
end;
end;
```

3 结束语

在构建语音 Web 应用系统时,着重点放在它与构建一般的文字型的网络系统的不同处,它传送的

语音数据必须是经压缩的,并且每个客户端必须有一下载的能进行语音数据获取、发送和压缩的用户端程序,而服务器端则负责客户的登录、端口分配和用户程序的分发.网络电话的实用性直接取决于语音数据的压缩比,通过合理选择语音压缩管理器控件的各项参数及 Delphi 提供的合适的 Internet 控件,可高效设计出鲁棒性良好的 Internet 网上电话的程序.通过程序实际的运行检验,说明传递压缩后的语音数据与传递一般文字信息一样可靠、快捷和清晰,这在 Internet 上提供了一个更直接、信赖的传递媒介.

参考文献:

[1] 林志清,陈福民.基于网络多媒体技术的电视新闻制作系统[J].计算机应用,2000,(8):23~24.
[2] Paul Perry,陈向群,等.多媒体开发指南[M].北京:清华大学出版社,1995.
[3] 乔林,等.Delphi 5.0 程序设计-Internet 应用务实篇[M].北京:中国铁道出版社,2000.
[4] 任旭钧,[美]Steve Teixeira Xavier Pacheco. Delphi 5 开发人员指南[M].北京:机械工业出版社,2000.

Research on Audio Transferring Technology on Internet

JIANG Xian-gang, TU Xiao-bin, CHEN Hai-lei

(School of Natural Science, East China Jiaotong University, Nanchang 330013, China)

Abstract: This paper introduces and analyses the technologies and applications of transferring audio information on Internet, core skills of developing audio transferring program by Delphi and the application of the ACM component.

Key words: audio compression manager ;twinsocket ; data stream