

基于 oRTP 库的无线语音视频直播系统的设计

付 鹏

(南京邮电大学 通信与信息工程学院,江苏 南京 210003)

摘 要:随着 3G 和 4G 网络的全面覆盖,通信网络的带宽大大增加,另外,智能终端(Android, iPhone 等)快速普及,使得无线语音视频直播的广泛应用成为可能。设计了一种基于开源 oRTP 库的无线语音视频直播系统。给出了系统的总体结构,在描述了 iOS 系统实现语音和视频采集的方法后,重点介绍了基于 oRTP 库的语音和视频传输部分的设计思路和程序实现,并给出了关键部分的实现代码,简要介绍了客户端如何实现直播语音和视频。通过测试,手机端处于 WiFi 和 3G 环境下,该系统具有较好的 QoS 性能,和良好的可靠性。实验证明,无线语音直播应用已具备应用和推广的初步条件。

关键词:oRTP 库;iOS 系统;实时传输协议;无线语音视频直播

中图分类号:TP277

文献标识码:A

文章编号:1673-629X(2013)02-0188-04

doi:10.3969/j.issn.1673-629X.2013.02.048

Design of Wireless Voice and Video Broadcast System Based on oRTP Library

FU Peng

(College of Telecommunications & Information Engineering, Nanjing University of Posts
& Telecommunications, Nanjing 210003, China)

Abstract: With comprehensive coverage of the 3G and 4G networks, the bandwidth of communication networks has greatly increased, in addition, the intelligent terminals (Android, iPhone, etc) spread rapidly, which makes extensive wireless voice and video broadcast applications possible. A design of wireless voice and video broadcast system based on the open source oRTP library is introduced in this paper. The overall structure of the system is given, after describing the method iOS system uses to complete the collection of the voice and video, focusing on the design ideas and procedures, and the critical section and the codes of key parts are given, then how the client can achieve voice and video broadcasting is briefly introduced. Testing result shows that the system can achieve better QoS performance and reliability in both WiFi and 3G environment. Experiments show that the wireless voice and video broadcast system gets the initial conditions of the application and promotion.

Key words: oRTP library; iOS; RTP; wireless voice and video broadcasting

0 引 言

随着 WiFi, 3G 网络的全面覆盖, 以及移动智能终端的普及, 使得无线语音视频直播系统的广泛应用成为可能。业界广泛讨论的视频监控系统^[1], 一般是基于固定的摄像头设备, 且常常只进行图像监控, 而无线语音视频直播系统则对声音、视频进行同步直播, 且具有高移动性的特点。未来无线视频直播还会在突发事件服务、医疗救助和灾难救援等领域发挥典型作用。同时, 随着移动互联网的发展, 视频直播业务作为一项不可或缺的重要应用, 还会在丰富人们业余生活、促进人们情感沟通等领域发挥显著作用^[2]。

笔者设计了一种基于 Apple 公司 iOS 系统移动终端(iPad, iPhone4 等) 为语音和视频采集端, 并以 WiFi 和 3G 网络为传输网络的语音和视频直播系统。客户端通过与服务器建立连接, 实现语音视频的单向直播。

1 系统总体结构

文中实现的无线语音视频直播系统由语音图像采集和编码, 发送端和服务器接收端, 客户请求端组成, 系统工作的流程框图如图 1 所示。

2 无线语音视频直播系统的设计

2.1 iOS 语音视频采集

项目基于美国苹果公司 iOS 系统终端, 以 iPhone4 为例, 它的主频为 1024MHz, RAM 为 512MB, 摄像头为 500 万像素。以视频采集为例, iOS4 系统的 SDK 中为

收稿日期: 2012-05-30; 修回日期: 2012-09-04

基金项目: 国家自然科学基金面上项目(61171092)

作者简介: 付 鹏(1988-), 男, 安徽合肥人, 硕士生, 研究方向为多媒体通信和下一代网络。

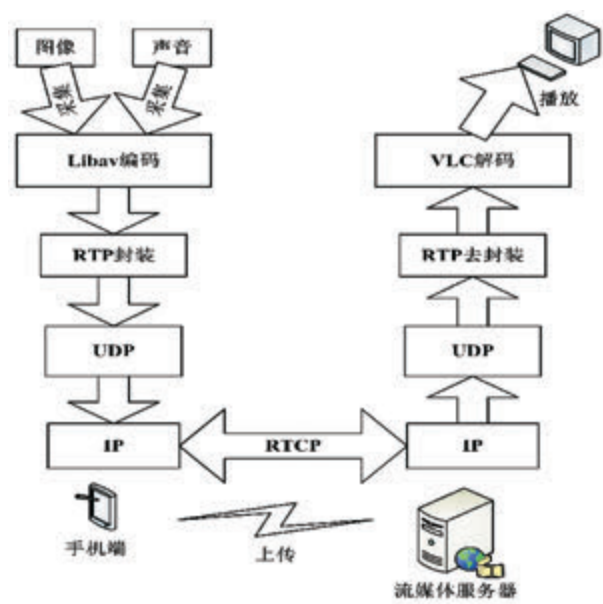


图1 移动语音视频直播系统的组成框图

人们提供了 AVCaptureSession 类,使用这个类,就可以轻松访问摄像头设备来实时采集视频。该系统中,视频的帧率为 12 ~ 20fps。其流程图如图 2 所示:

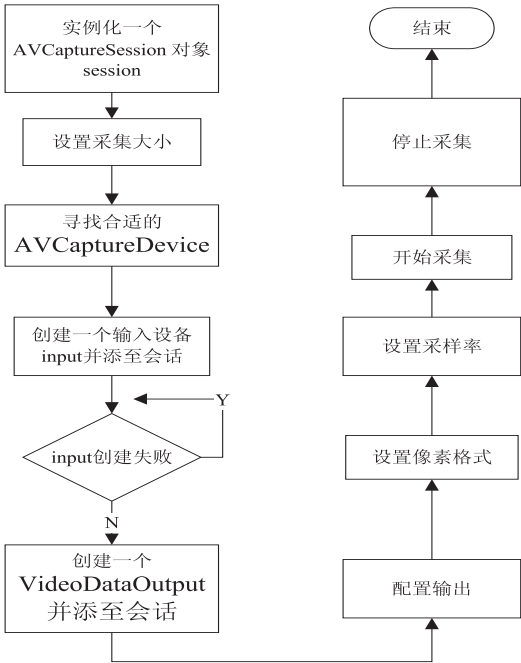


图2 视频采集流程图

2.2 语音视频编码

利用 Libav 库^[3](旧称 ffmpeg)进行语音和视频的编码。Libav 是一套可以用来记录、转换数字音频、视频,并能将其转化为流的开源计算机程序。

系统中视频编码采用 H. 264^[4],H. 264 是新一代视频编码标准。其编解码流程主要包括五个部分:帧内预测和帧间预测、变换和反变换、量化和反量化、环路滤波、熵编码。H. 264 采用视频编码层 (VCL) 和网络抽象层 (NAL) 分层的结构设计,以网络所要求的方

式对数据进行打包和发送,码流结构对网络的适应性强,增加差错恢复能力,使 H. 264 能够很好适应 IP 和无线网络的应用^[5]。

音频编码采用 Nellymoser^[6](又称 Asao),是由 Nellymoser 公司开发的音频编码标准。

2.3 语音视频流传输系统的设计

2.3.1 RTP/RTCP 协议概述

在进行流媒体的实时传输时,TCP 的重传机制不能满足实时性的要求,且 TCP 在检测到丢包后采取的发送速率减半的策略不适合流媒体要求的传输速率平稳性的要求。另一种传输协议 UDP 不含任何控制机制,容易抢占带宽造成网络瘫痪。

通常进行流媒体传输时选择 RTP (Real-time Transfer Protocol)^[7],RTP 负责对流媒体数据按照 RTP 数据包格式进行封装,并利用与它绑定的传输层协议进行数据包的实时传输。RTP 为 Internet 上端到端的实时传输提供时间信息和流同步,但并不保证服务质量,服务质量由 RTCP(Real-time Transfer Control Protocol)来提供。

RTP 包头第一个 4 字节定义了版本号 (V),填充位 (P),扩展位 (X),CSRC 计数器 (CC),标记位 (M),载荷类型 (PT),序列号 (sequence number);第二个 4 字节表示时间戳 (timestamp),时间戳为音、视频同步提供重要信息;第三个 4 字节表示同步源标识 (SSRC),它标识 RTP 会话的来源,在一次会话时随机选取的。前 12 个字节是 RTP 固定的,紧接其次的是贡献源列表 (CSRC List)^[8]。

2.3.2 音、视频数据上传至服务器

该部分的工作是在编码得到音频帧、视频帧后,将其上传至服务器,在正式上传数据之前,首先经历手机端与服务器端的预连接。预连接阶段,手机端向服务器发送预连接信息,服务器端收到预连接信息后,对信息进行记录处理,然后回送“ready”信息,手机端收到“ready”后,则正式发送流媒体数据。

1) 预连接。

手机端用 TCP socket 向服务器发送预连接信息,包含音频、视频的 SSRC,和该会话的 URL。预连接信息的格式为: SSRC; SSRC; URL (例: 2364833500; 4563535269; rtmp://192.168.1.250/live/njupt)。服务器收到预连接信息后,判断音、视频 SSRC 是否已经存在于建立的列表中,若不存在,则添加。然后回送手机端一个“ready”信号,手机端收到“ready”信号后,则正式发送流媒体数据。

2) 流媒体数据的发送与接收。

文中利用开源的 oRTP 库^[9],oRTP 库是用 C 开发的 RTP 库。该库能够在 Linux (和其他类 Unix 系统)

和 Windows 上运行。它遵循 RFC3550,并且拥有容易使用的 API。oRTP 库具有支持多种 RTP 格式、发送接收实时调度、单线程支持多路媒体流、实现自适应的缓冲区算法、并且实现了 RTCP 等等优于其他 RTP 封装库的特点。

手机端发送部分的流程图如图 3 所示:

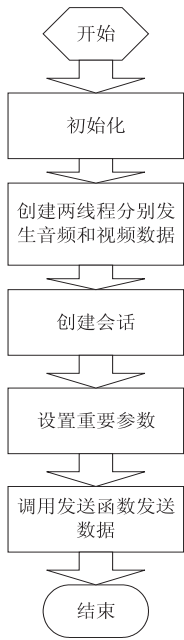


图 3 手机端发送的程序流程图

以音频为例,手机发送端部分关键语句如下:

```
//初始化
ortp_init();
ortp_scheduler_init();

//创建音频会话
RtpSession * audio_sess;
audio_sess = rtp_session_new ( RTP_SESSION_SEN-
LY);/* 只发送 */

//重要参数的设置
rtp_session_set_payload_type(audio_sess,0);
rtp_session_set_scheduling_mode ( audio_sess,1);/* 设置调
度管理功能 */
rtp_session_set_blocking_mode ( audio_sess,1);/* 设置成阻
塞模式 */
rtp_session_set_remote_addr( audio_sess,SERVER_ADDR,des
_audio_port);

//调用发送函数
sendBytes=rtp_session_send_with_ts( audio_sess,sendBuffer,
len,m_curTimeStamp);
if ( sendBytes > 0)
m_curTimeStamp += m_timeStampInc; /* 增加时间戳 */
服务器端创建两个线程,分别接收音频、视频数
据。其流程图如图 4 所示。
```

以接收音频为例,部分代码如下:

```
//初始化(同上,省略)
```

```
//创建音频会话
RtpSession * audio_sess;
audio_sess = rtp_session_new ( RTP_SESSION_RECVON-
LY);/* 只接收 */

//重要参数的设置(负载类型、调度模式、阻塞模式设置)
rtp_session_set_local_addr( audio_sess,"0.0.0.0",audio_
portbase);
rtp_session_set_connected_mode(audio_sess,TRUE);

//接收数据部分
while(1) {
int have_more=1, totalBytes=0, rcvBytes=0;
while (have_more) { /* 一次读不完会继续循环直至读完
*/
rcvBytes = rtp_session_rcv_with_ts( audio_sess, buffer1+tot-
alBytes,
READ_AUDIO_BYTES_PER_TIME,m_curTimeStamp_audio,
&have_more);totalBytes+= rcvBytes;
}
m_curTimeStamp_audio+=m_timeStampInc;
if( totalBytes <= 0 )
continue;
...
}
```

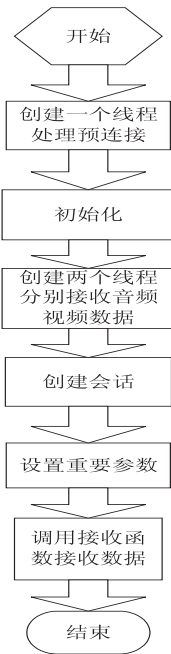


图 4 服务器端的接收数据的程序流程图

2.4 客户端解码实时收看

客户端使用开源的 VLC 库^[10]实现音频、视频流的解码和播放。VLC 作为一个开源、跨平台的视频播放器,支持 Libav 支持的所有编码器和所有文件格式,有极强的播放能力。

3 系统测试与分析

为分析系统的性能,对数据包的延时以及丢包情

况进行测试。分别让 iPhone4 采集端处于内网 WiFi (手机端和服务端处于同一内网) 和外网 3G (手机端通过 3G 网络,服务器部署于内网) 两种环境下,观看端均位于内网情况下,进行 5 组实验。在测算数据包延时的实验中,每组实验统计 100 个数据包的延迟数据,得到 5 组实验的平均延迟;在测算丢包的实验中,每组实验追踪 1000 个数据包,统计其丢失包数目并计算丢包率。结果如表 1 所示。

表 1 数据包延迟表

实验组别 环境	1	2	3	4	5
内网 WiFi 延迟 秒数/丢包率	0.96/ 0.2%	0.90/ 0.1%	0.98/ 0.2%	1.04/ 0.3%	1.08/ 0.2%
外网 3G 延迟 秒数/丢包率	3.78/ 0.8%	3.88/ 1.0%	3.84/ 1.1%	3.92/ 1.2%	3.98/ 0.9%

实验表明,系统较稳定,内网 WiFi 下的性能优于外网 3G,在两种环境下,基本能够进行语音和视频的直播。

4 结束语

文中给出了一种基于 iOS 系统和开源 oRTP 库设计多媒体数据上传的具体实施方案,在描述了移动视频直播系统的总体结构和模块后,具体给出了音视频上传部分的程序流程图以及部分实现代码。实验表明,该系统无论置于 WiFi 还是 3G 环境均具有良好的稳定性,在系统延时、画面和声音质量等方面均有不错的表现。但此方案还存在不足之处,例如如何进一步提高在 3G 环境下的直播性能,则是今后需要考虑的问题^[11,12]。

(上接第 187 页)

则参数选择就符合标准。

今后工作中应考虑的问题:编制程序,统计观测系统的炮检距及方位角的分布是否均匀,是否满足叠加次数,比较所选三维观测系统的优劣程度。目前尚无这些技术评价标准,需专门研究,确定相关技术标准。这些问题解决了,软件实际应用能力及准确度都将得到更大提高。

参考文献:

[1] 陆基孟. 三维地震勘探方法原理及资料解释[M]. 北京:石油工业出版社,1991.

[2] 赵靖舟,张春林. 油气成藏地质学的内涵及其在油气地质学中的定位[J]. 石油与天然气地质,2007(2):139-142.

[3] Horvath P S. History proves 3-D seismic surveys can be effective[J]. World Oil,1985,201(1):67-70.

[4] Parra J O. The transversely isotropic poroelastic wave equation including the Biot and the squirt mechanisms:theory and ap-

参考文献:

[1] 徐世强,赵 霁,牛泽民. 基于 GPRS 的远程视频监控系统的设计与实现[J]. 计算机技术与发展,2011,21(12):150-153.

[2] Krikelis A. Considerations for a new generation of mobile multimedia communication systems[J]. Multimedia Application, 2000(4-6):80-82.

[3] Wikipedia[EB/OL]. [2012-03-01]. <http://zh.wikipedia.org/wiki/Libav>.

[4] 毕厚杰. 新一代视频压缩编码标准-H. 264/AVC[M]. 北京:人民邮电出版社,2005.

[5] 谭雪理. 基于 RTP 的 H. 264 视频实时传输系统的研究与实现[D]. 南京:南京邮电大学,2011.

[6] Wikipedia[EB/OL]. [2012-01-12]. [http://en.wikipedia.org/wiki/Asao_\(codec\)](http://en.wikipedia.org/wiki/Asao_(codec)).

[7] Schulzrinne H,Casner S,Frederick R,et al. RTP:A Transport Protocol for Real-time Application[S]. Internet Draft, RFC3550,1996.

[8] Koistinen T. Protocol overview:RTP and RTCP,Nokia Telecommunications[EB/OL]. [2011-11-12]. <http://www.netlab.tkk.fi/opetus/s38130/k99/presentations/4.pdf>.

[9] ORTP. a Real-time Transport Protocol (RTP,RFC3550) Library[EB/OL]. [2011-11-12]. <http://www.linphone.org/eng/documentation/dev/ortp.html>.

[10] VLC media player:VideoLan[EB/OL]. [2011-11-20]. <http://www.videolan.org/vlc/>.

[11] 范 鹏. 多媒体信息实时传输及其 QoS 的研究与实现[D]. 郑州:中国人民解放军信息工程大学,2004.

[12] 李 思. 基于 RTP 的视频传输控制机制的研究[D]. 西安:西安电子科技大学,2009.

plication geophysics[J]. Geophysics,1997,62(1):309-318.

[5] Shatilo A P,Sondergeld C,Rai C S. Ultrasonic attenuation in Glenn Pool rocks, northeastern Oklahoma [J]. Geophysics, 1998,63(2):465-478.

[6] 陈振华,余永权,张 瑞. 模糊模式识别的几种基本模型研究[J]. 计算机技术与发展,2010,20(9):32-35.

[7] 周俊杰,李士祥,贾运巧. 复杂山区煤田三维地震数据采集技术研究[J]. 地震地质,2007(1):105-113.

[8] 张念鲁. VB 网络编程及应用实训[M]. 北京:高等教育出版社,2003.

[9] 赵 波,王大翊,荣 霓,等. 一种面向开发过程的软件可靠性预测方法[J]. 计算机技术与发展,2011,21(9):14-18.

[10] 王永庆. 人工智能原理与方法[M]. 西安:西安交通大学出版社,2006.

[11] 张明学. 地震勘探原理与解释[M]. 北京:石油工业出版社,2010.

[12] 马在田. 三维地震勘探方法[M]. 北京:石油工业出版社,1989.

基于oRTP库的无线语音视频直播系统的设计

作者: [付鹏](#)
作者单位: [南京邮电大学 通信与信息工程学院, 江苏 南京 210003](#)
刊名: [计算机技术与发展](#)
英文刊名: [Computer Technology and Development](#)
年, 卷(期): 2013(2)

本文链接: http://d.g.wanfangdata.com.cn/Periodical_wjtz201302050.aspx