

互联网中基于MPEG-4流媒体播放的研究

高科, 刘心松, 詹骥

(电子科技大学计算机科学与工程学院 成都 610054)

【摘要】在分析RTP/RTCP协议的基础上分析了传输MPEG-4编码视音频数据的RTP分组负载格式。研究并提出了一种MPEG-4视音频流播放方案。介绍了其线程模型、总体实现框架等。并采用视频流同步音频流的方案使得播放器的播放质量有了较高的保证。实用表明,采用该方案实现的客户端具有稳定性好,占用传输带宽低等优点。

关键词 流媒体; 实时传输协议; 线程同步; 多媒体压缩编码标准(MPEG-4)
中图分类号 TP393.09 文献标识码 A

Study on MPEG-4 Streaming Player over Internet

GAO Ke, LIU Xin-song, ZHAN Ji

(School of Computer Science and Engineering, Univ. of Elec. Sci. & Tech. of China Chengdu 610054)

Abstract Based on the analysis of RTP/RTCP protocol, The RTP data format used for MPEG-4 media is analyzed. An implementation scheme of client using MPEG-4 media data is then proposed. Adopting the scheme that video streams are synchronized to audio stream produces a high playback quality of the media player. The thread model and main architecture of the client and are introduced. Applications show that the clients implemented by the scheme have good stability.

Key words streaming media; real-time transport protocol; thread synchronization; moving picture experts group

随着互联网应用的日益广泛发展,流媒体技术的运用已经深入到了许多领域。流媒体技术是指以流式传输信息的网络信息传输。多媒体压缩编码标准(Moving Picture Experts Group, MPEG-4)^[1],压缩标准是由国际标准化组织下属的运动图像专家小组制定的新一代超低压缩比规范。如何将MPEG-4编码的影视信息运用于互联网中,使网络用户可以通过互联网享受到实时的,可交互多媒体服务成为了一个亟待解决的问题。本文研究了流媒体播放协议并提出了一种基于互联网的MPEG-4流媒体播放方案。

1 RTP/RTCP/RTSP协议与MPEG-4编码

实时传输协议(Real-Time Transport Protocol, RTP)由Internet工程任务组制定^[2]。RTP协议提供了一种通过网络采用单播或组播的形式传输视频,音频等实时信息流的方法。RTP协议的每个分组都包含序列号以及时间戳。因此,它可以为实时数据传输提供时序重组,帧丢失检测,多路复用等服务。RTP固定首部的格式如图1a所示。字段V指明了RTP包的版本号,当前为2。负载格式PT一般不使用。当数据需要加密时标志位P置1。当一个RTP包中包含多个视频对象面(Video Object Plane, VOP)时,标志位M置1。CC字段指明了当一个数据流由多个数据流混合而成时的参与源数目。序列号指明当前分组的序号,并且应呈线性递增。时间戳用于确定分组中数据的正确回放时间。时间戳必须单调线性增加,接收方根据当前会话的时间戳来确定当前的时序,以便选择正确的回放时间。当需要把多个源数据流混合成

V	P	X	CC	M	PT	序列号
						时间戳
						同步源标识SSRC
						参与源标识CSRC

a. RTP固定首部格式

V	P	RC	PTYPE	LENGTH
数据区(Data Area)				
⋮				

b. RTCP分组格式

图1 RTP/RTCP数据报文格式

收稿日期: 2003-06-10

基金项目: 四川省科技攻关计划资助项目(02C0C0006-018)

作者简介: 高科(1980-),男,硕士,主要从事多媒体数据传输、分布式并行服务器方面的研究。

一个数据流时,同步源标识(Synchronization Source, SSRC)指出数据流的源站,该项随机选择且唯一。参与源标识符(Contributing Source, CSRC)表示对包内载荷起作用的源。实时流控制协议(Real-Time Control Protocol, RTCP)是RTP协议的伴生协议,用于拥塞控制和流控制。RTCP包格式如图1b所示。字段V指明了当前的版本号,当数据需要加密时标志位P置1。RC字段指明了首部后面的报告计数。PTYPE指定了分组报文类型。LENGTH字段指定了分组的总长度。数据区即为各种报告信息。实时流协议(Real-Time Streaming Protocol, RTSP)定义了如何有效地通过IP网络传送多媒体数据^[3]。RTSP协议提供了选择底层数据传输协议的功能,如使用UDP或TCP协议。在本播放系统中,把RTSP协议实现在TCP协议上,以提高传输的可靠性。MPEG-4编码按宏块进行,采用预测编码、形状编码和离散余弦变换的混合编码方法。预测编码通过运动估计和运动补偿来实现。为了适应任意形状的VOP(视频对象平面),MPEG-4引入了多边形匹配技术和图像填充技术。MPEG-4流媒体信息的RTP数据组包时,RTP包和VOP并无绝对关系,一个VOP可以放入多个RTP包中。当VOP占据空间较小时,多个VOP亦可放入一个RTP包。RTP包分为两部分。开始部分是RTP固定首部信息,它提供了通用的、关于协议服务、消息格式等信息;后一部分是与MPEG-4传输具体相关的部分,鉴于MPEG-4编码的视频流本身已有很强的纠错能力,因此传输MPEG-4编码信息流的RTP包并不像传输MPEG-2编码信息流的RTP包那样在一个RTP通用头部后面提供一个用于纠错的特殊媒体头部。而是在其固定头首部后面紧跟视频数据区。

2 播放系统的研究

2.1 播放系统实现方案的总体框架

基于以上对RTP/RTCP/RTSP协议以及MPEG-4编码的分析,本文研究并提出了一种在互联网中基于RTP/RTCP/RTSP协议的MPEG-4实时流媒体播放系统实现方案。其总体框架如图2所示。



图2 流媒体播放系统的总体框架

在播放过程中,播放器共有5个端口与服务器交互。它们分别是:(1) 视频流RTP端;(2) RTCP端;(3) 音频流RTP端;(4) RTCP端;(5) 播放器与服务器进行交互的RTSP端口。播放器通过RTP端口获得视频音频信息,通过RTCP端口获得会话期间

底层网络状况。通过RTSP端口,播放器可以实现与服务器的交互功能,如开始播放,暂停播放,拖动时间进度条。在整个播放过程中,CPlayerSession类是播放器的核心,负责维护整个会话,其主要伪码如下:

```

class CPlayerSession {
public:
    CPlayerMedia *m_my_media; //媒体类指针,指向音频类或视频类,两个类指针形成一个链表
    int play_all_media( double start_time = 0.0); //播放媒体流
    int m_current_state; //当前播放器的工作状态
    int m_current_time; //当前播放时间
    .....
private:
    CAudioSync *m_audio_sync; //音频播放同步类,存有音频播放缓存
    CVideoSync *m_video_sync; //视频播放同步类,存有视频播放缓存
    CMsgQueue *m_master_msg_queue; //同步播放线程的消息队列
    .....
}
  
```

2.2 播放系统的数据流控制技术

播放过程中,播放器共有5个线程同时运行,以保证播放过程视频流与音频流的同步。其线程模型如图3所示。接收线程负责接收从服务器端传入的RTP/RTCP数据包,每个接收线程各自维护一个双向循环链表。当有RTP数据包传入时,接收线程将其置入链表头部。当解码线程需要解时,从循环链表尾部取出RTP包数

据, 组帧, 然后解码。同步播放线程是播放器的核心, 控制着会话期间播放器播放状态的转变, 同时也控制接收线程, 解码线程的工作时间, 以保证视频信息和音频信息的同步。在此重点介绍视频信息的传播过程, 如图4所示。服务器把经过MPEG-4编码的视频信息打成RTP包后发送至客户端, 播放器在接收RTP包的同时也接收RTCP包。播放器在播放的同时进行分组过滤, 丢包率检测等工作。再将这些信息经过RTCP分析器分析汇总后传至RTCP包生成器, 重新生成RTCP数据包。传至服务器, 供服务器分析、监测会话期间底层网络状况以及提供端点之外的带外通信。

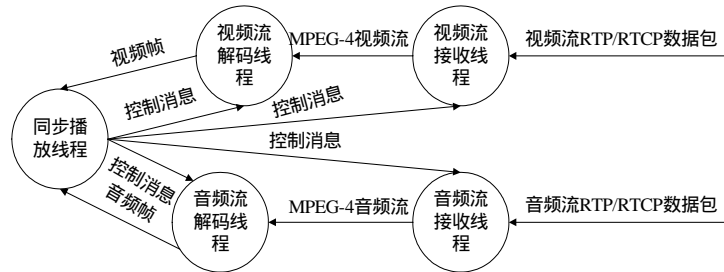


图3 线程模型

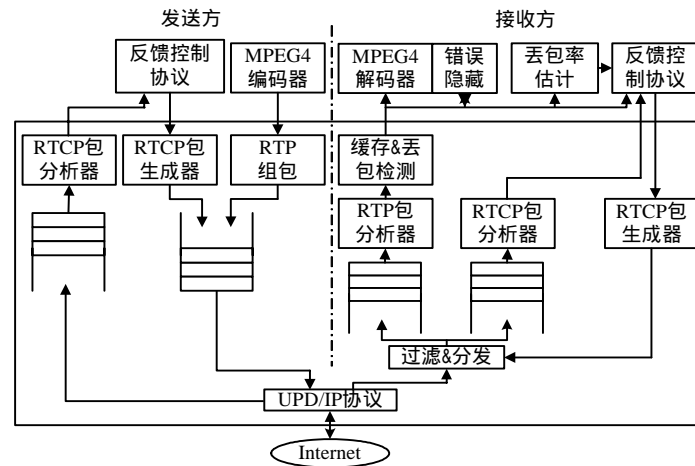


图4 视频流传输示意图

3.3 播放系统的视音频同步算法

在整个播放过程中, 视频流和音频流的同步播放是播放器实现过程中的重点和难点。图5所示描绘出了视频流与音频流的同步过程。在整个播放过程中, 同步播放线程始终维护着一个消息队列, 其中存放着自定义的消息。当用户暂停播放或拖动时间进度条时, 这些消息被发送给接收和解码线程, 使得同步播放线程能够控制着播放器的工作状态。

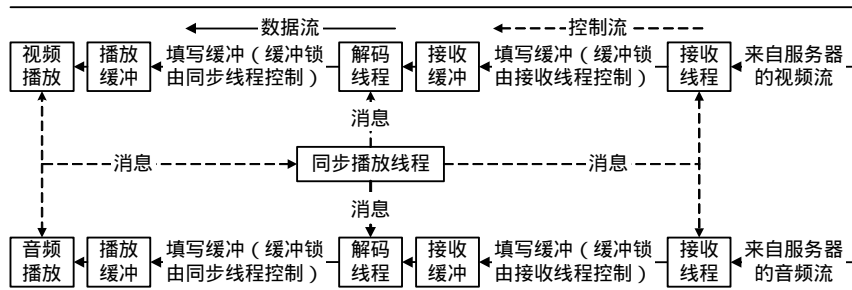


图5 同步算法示意图

考虑音频流的码率较视频流码率小, 在网络状况差的情况下具有更高的稳定性。因此, 整个播放过程中, 选用音频流作为时间基准。视频流参照音频流的时间戳同步。

2.3.1 音频流的播放

初始会话时,客户端将收到的第一个RTP音频包的时间戳中的时间作为基准时间 $m_current_time$ 。若RTP包含有一个以上的音频帧,则同一RTP包中音频帧的时间戳通过 $m_current_time += ((m_samplesperframe(每帧的样本数) * 1000) / m_freq(采样率))$ 计算得到。无论音频包和视频包,只要是属于同一帧的信息,则它们的时间戳相同。会话初始化后,若接收线程接收到两秒钟长度的数据,则释放接收缓冲锁,解码线程开始解码。当解码器对一帧数据解码后,解码线程向同步播放线程发送消息,同步播放线程对播放缓冲加锁。同步播放线程调用DirectX中DirecxSound的标准接口实现声音的播放。

2.3.2 视频流的播放

对于视频流的播放,接收、解码过程部分与音频流类似,在此不再赘述。不同的是同步部分,因MPEG-4编码的视频流对网络带宽占用较大,容易导致丢包。为此,视频帧的播放时间应与 $m_current_time$ 一致。有两种情况须考虑:(1)若待解码的视频帧的时间戳比 $m_current_time$ 小200 ms时,说明视频播放已落后于音频播放,同步播放线程调用CPlayerMedia的跳帧成员函数,跳过当前视频帧,以追上音频播放;(2)若待解码的视频帧的时间戳比 $m_current_time$ 大200 ms时(虽然此种情况出现的概率很小,但也应考虑),说明视频已超越音频的播放,此时同步播放线程将暂停视频播放,以待音频播放与视频播放同步。最后,同步播放线程调用DirectX中DirecxShow的标准接口实现视频的播放。

2.4 播放系统的控制流管理

控制流是指用户控制客户端播放状态的信息流。通过控制流,用户可以实现暂停、播放、停止、拖动时间进度条等功能。控制流的描述采用标准的会话描述协议,该协议工作于RTSP协议之上。服务器和播放器的交互采用RTSP协议。播放器通过由同步播放线程维持的一个有限状态机来实现播放状态的转变。当用户点击播放器上的控制按钮时,播放器接收到用户的控制信息,从而导致播放器播放状态的转变。图6所示指明了该有限状态机的各个状态,及其转变条件。同步播放线程中的 $m_current_state$ 变量标识了播放器当前的播放状态。当用户拖动时间进度条时,播放器暂停,然后重新设置播放时间 $m_current_time$,最后继续播放。若视音频不同步,则播放器也进入暂停状态,待视音频同步后继续播放。

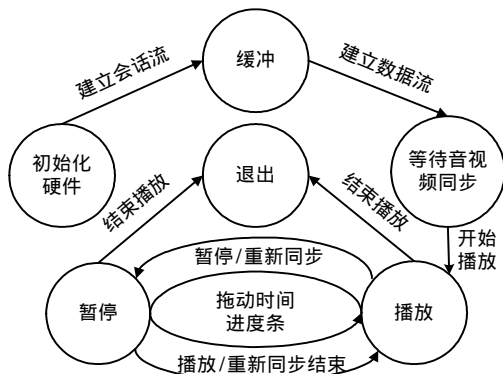


图6 客户端的状态转变

3 结束语

本文着重分析了基于RTP/RTCP协议采用MPEG-4编码的视音频流传输,并提出了一种在互联网流媒体传输过程中的客户端实现方案。实用表明采用本方案实现的播放器具有传输带宽低,传输成本低等优点,具有良好的运用前景。

参 考 文 献

- [1] 钟玉琢, 王 琪, 贺玉文. 基于对象的多媒体数据压缩编码国际标准——MPEG-4及其校验模型[M]. 北京: 科学出版社, 2000.
- [2] Schulzrinne H, Casner S, Frederick R, et al. RTP: a transport protocol for real-time application. RFC 1889[S]. Internet Engineering Task Force, 1996.
- [3] Schulzrinne H, Lanphier R, Rao A. RTSP: real time streaming protocol. RFC 2326[S]. Internet Engineering Task Force, 1998.

编 辑 孙晓丹