音视频编解码技术

# 一、前言

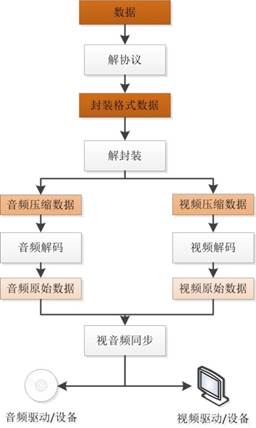
电影文件有不同的格式，用不同的后缀表示：avi，rmvb，mp4，flv，mkv等等。在这里需要注意的是，这些格式代表的是**封装格式**。何为封装格式？就是把视频数据和音频数据打包成一个文件的规范。仅仅靠看文件的后缀，很难能看出具体使用了什么**视音频编码标准**。总的来说，不同的封装格式之间差距不大，各有优劣。

注：有些封装格式支持的视音频编码标准十分广泛，应该算比较优秀的封装格式，比如MKV；而有些封装格式支持的视音频编码标准很少，应该属于落后的封装格式，比如RMVB。

# 二、视频播放器的原理

视音频技术主要包含以下几点：封装技术，视频压缩编码技术以及音频压缩编码技术。如果考虑到网络传输的话，还包括流媒体协议技术。

视频播放器播放一个互联网上的视频文件，需要经过以下几个步骤：解协议，解封装，解码视音频，视音频同步。如果播放本地文件则不需要解协议，为以下几个步骤：解封装，解码视音频，视音频同步。他们的过程如图所示。



**解协议**的作用，就是将流媒体协议的数据，解析为标准的相应的封装格式数据。视音频在网络上传播的时候，常常采用各种流媒体协议，例如HTTP，RTMP，或是MMS等等。这些协议在传输视音频数据的同时，也会传输一些信令数据。这些信令数据包括对播放的控制（播放，暂停，停止），或者对网络状态的描述等。解协议的过程中会去除掉信令数据而只保留视音频数据。例如，采用RTMP协议传输的数据，经过解协议操作后，输出FLV格式的数据。

**解封装**的作用，就是将输入的封装格式的数据，分离成为音频流压缩编码数据和视频流压缩编码数据。封装格式种类很多，例如MP4，MKV，RMVB，TS，FLV，AVI等等，它的作用就是将已经压缩编码的视频数据和音频数据按照一定的格式放到一起。例如，FLV格式的数据，经过解封装操作后，输出H.264编码的视频码流和AAC编码的音频码流。

**解码**的作用，就是将视频/音频压缩编码数据，解码成为非压缩的视频/音频原始数据。音频的压缩编码标准包含AAC，MP3，AC-3等等，视频的压缩编码标准则包含H.264，MPEG2，VC-1等等。解码是整个系统中最重要也是最复杂的一个环节。通过解码，压缩编码的视频数据输出成为非压缩的颜色数据，例如YUV420P，RGB等等；压缩编码的音频数据输出成为非压缩的音频抽样数据，例如PCM数据。

**视音频同步**的作用，就是根据解封装模块处理过程中获取到的参数信息，同步解码出来的视频和音频数据，并将视频音频数据送至系统的显卡和声卡播放出来。

# 三、流媒体协议

流媒体协议是服务器与客户端之间通信遵循的规定。当前网络上主要的流媒体协议如表所示。

主要流媒体协议一览

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 名称 | 推出机构 | 传输层协议 | 客户端 | 目前使用领域 |
| RTSP+RTP | IETF | TCP+UDP | VLC, WMP | IPTV |
| RTMP | Adobe Inc. | TCP | Flash | 互联网直播 |
| RTMFP | Adobe Inc. | UDP | Flash | 互联网直播 |
| MMS | Microsoft Inc. | TCP/UDP | WMP | 互联网直播+点播 |
| HTTP | WWW+IETF | TCP | Flash | 互联网点播 |

因为互联网网络环境的不稳定性，RTSP+RTP较少用于互联网视音频传输。互联网视频服务通常采用TCP作为其流媒体的传输层协议，因而像RTMP，MMS，HTTP这类的协议广泛用于互联网视音频服务之中。这类协议不会发生丢包，因而保证了视频的质量，但是传输的效率会相对低一些。

# 四、封装格式

封装格式的主要作用是把视频码流和音频码流按照一定的格式存储在一个文件中。现如今流行的封装格式如下表所示：

主要封装格式一览

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 名称 | 推出机构 | 流媒体 | 支持的视频编码 | 支持的音频编码 | 目前使用领域 |
| AVI | Microsoft Inc. | 不支持 | 几乎所有格式 | 几乎所有格式 | BT下载影视 |
| MP4 | MPEG | 支持 | MPEG-2, MPEG-4, H.264, H.263等 | AAC, MPEG-1 Layers I, II, III, AC-3等 | 互联网视频网站 |
| TS | MPEG | 支持 | MPEG-1, MPEG-2, MPEG-4, H.264 | MPEG-1 Layers I, II, III, AAC, | IPTV，数字电视 |
| FLV | Adobe Inc. | 支持 | Sorenson, VP6, H.264 | MP3, ADPCM, Linear PCM, AAC等 | 互联网视频网站 |
| MKV | CoreCodec Inc. | 支持 | 几乎所有格式 | 几乎所有格式 | 互联网视频网站 |
| RMVB | Real Networks Inc. | 支持 | RealVideo 8, 9, 10 | AAC, Cook Codec, RealAudio Lossless | BT下载影视 |

由表可见，除了AVI之外，其他封装格式都支持流媒体，即可以“边下边播”。有些格式更“万能”一些，支持的视音频编码标准多一些，比如MKV。而有些格式则支持的相对比较少，比如说RMVB。

# 五、视频编码

视频编码的主要作用是将视频像素数据（RGB，YUV等）压缩成为视频码流，从而降低视频的数据量。如果视频不经过压缩编码的话，体积通常是非常大的，一部电影可能就要上百G的空间。视频编码是视音频技术中最重要的技术之一。视频码流的数据量占了视音频总数据量的绝大部分。高效率的视频编码在同等的码率下，可以获得更高的视频质量。

主要视频编码一览

|  |  |  |  |
| --- | --- | --- | --- |
| 名称 | 推出机构 | 推出时间 | 目前使用领域 |
| HEVC(H.265) | MPEG/ITU-T | 2013 | 研发中 |
| H.264 | MPEG/ITU-T | 2003 | 各个领域 |
| MPEG4 | MPEG | 2001 | 不温不火 |
| MPEG2 | MPEG | 1994 | 数字电视 |
| VP9 | Google | 2013 | 研发中 |
| VP8 | Google | 2008 | 不普及 |
| VC-1 | Microsoft Inc. | 2006 | 微软平台 |

由表可见，有两种视频编码方案是最新推出的：VP9和HEVC。目前这两种方案都处于研发阶段，还没有到达实用的程度。当前使用最多的视频编码方案就是H.264。

## 1.主流编码标准

H.264仅仅是一个编码标准，而不是一个具体的编码器，H.264只是给编码器的实现提供参照用的。实际中使用最多的就是x264了，性能强悍（超过了很多商业编码器），而且开源

## 2.下一代编码标准

下一代的编解码标准就要数HEVC和VP9了。VP9是Google继VP8之后推出的新一代标准。VP9和HEVC相比，要稍微逊色一些。

# 六、音频编码

音频编码的主要作用是将音频采样数据（PCM等）压缩成为音频码流，从而降低音频的数据量。音频编码也是互联网视音频技术中一个重要的技术。但是一般情况下音频的数据量要远小于视频的数据量，因而即使使用稍微落后的音频编码标准，而导致音频数据量有所增加，也不会对视音频的总数据量产生太大的影响。高效率的音频编码在同等的码率下，可以获得更高的音质。

主要音频编码一览

|  |  |  |  |
| --- | --- | --- | --- |
| 名称 | 推出机构 | 推出时间 | 目前使用领域 |
| AAC | MPEG | 1997 | 各个领域（新） |
| AC-3 | Dolby Inc. | 1992 | 电影 |
| MP3 | MPEG | 1993 | 各个领域（旧） |
| WMA | Microsoft Inc. | 1999 | 微软平台 |

由表可见，近年来并未推出全新的音频编码方案，可见音频编码技术已经基本可以满足人们的需要。音频编码技术近期绝大部分的改动都是在MP3的继任者——AAC的基础上完成的。

# 七、现有的网络视音频平台对比

现有的网络视音频服务主要包括两种方式：点播和直播。点播意即根据用户的需要播放相应的视频节目，这是互联网视音频服务最主要的方式。绝大部分视频网站都提供了点播服务。直播意即互联网视音频平台直接将视频内容实时发送给用户，目前还处于发展阶段。直播在网络电视台，社交视频网站较为常见。

## 1．直播平台参数对比

主流互联网视音频平台直播服务的参数对比如表所示：

现有网络视音频平台参数对比

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 名称 | 协议 | 封装 | 视频编码 | 音频编码 | 播放器 |
| CNTV | 私有 | | | | |
| 华数TV | RTMP | FLV | H.264 | AAC | Flash |
| 六间房 | RTMP | FLV | H.264 | AAC | Flash |
| 中国教育电视台 | RTMP | FLV | H.264 | AAC | Flash |
| 北广传媒移动电视 | RTMP | FLV | H.264 | AAC | Flash |
| 上海IPTV | RTSP+RTP | TS | H.264 | MP2 | 机顶盒 |

可以看出，直播服务普遍采用了RTMP作为流媒体协议，FLV作为封装格式，H.264作为视频编码格式，AAC作为音频编码格式。采用RTMP作为直播协议的好处在于其被Flash播放器支持。而Flash播放器如今已经安装在全球99%的电脑上，并且与浏览器结合的很好。因此这种流媒体直播平台可以实现“无插件直播”，极大的简化了客户端的操作。封装格式，视频编码，音频编码方面，无一例外的使用了FLV + H.264 + AAC的组合。FLV是RTMP使用的封装格式，H.264是当今实际应用中编码效率最高的视频编码标准，AAC则是当今实际应用中编码效率最高的音频编码标准。视频播放器方面，都使用了Flash播放器。

## 2．点播平台参数对比

主流网络视音频平台点播服务的参数对比如表所示：

现有互联网视音频平台参数对比

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 名称 | 协议 | 封装 | 视频编码 | 音频编码 | 播放器 |
| CNTV | HTTP | MP4 | H.264 | AAC | Flash |
| CNTV（部分） | RTMP | FLV | H.264 | AAC | Flash |
| 华数TV | HTTP | MP4 | H.264 | AAC | Flash |
| 优酷网 | HTTP | FLV | H.264 | AAC | Flash |
| 土豆网 | HTTP | F4V | H.264 | AAC | Flash |
| 56网 | HTTP | FLV | H.264 | AAC | Flash |
| 音悦台 | HTTP | MP4 | H.264 | AAC | Flash |
| 乐视网 | HTTP | FLV | H.264 | AAC | Flash |
| 新浪视频 | HTTP | FLV | H.264 | AAC | Flash |

可以看出，点播服务普遍采用了HTTP作为流媒体协议，H.264作为视频编码格式，AAC作为音频编码格式。采用HTTP作为点播协议有以下两点优势：一方面，HTTP是基于TCP协议的应用层协议，媒体传输过程中不会出现丢包等现象，从而保证了视频的质量；另一方面，HTTP被绝大部分的Web服务器支持，因而流媒体服务机构不必投资购买额外的流媒体服务器，从而节约了开支。点播服务采用的封装格式有多种：MP4，FLV，F4V等，它们之间的区别不是很大。视频编码标准和音频编码标准是H.264和AAC。这两种标准分别是当今实际应用中编码效率最高的视频标准和音频标准。视频播放器方面，无一例外的都使用了Flash播放器。