

流媒体的封装格式

1. 视频播放器原理

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

2. 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

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

RTSP+RTP经常用于IPTV领域。因为其采用UDP传输视音频，支持组播，效率较高。但其缺点是网络不好的情况下可能会丢包，影响视频观看质量。因而围绕IPTV的视频质量的研究还是挺多的。

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。

3.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：

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。

4.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。

视频编码的简单原理可以参考：视频压缩编码和音频压缩编码的基本原理

注：视频编码技术在整个视音频技术中应该是最复杂的技术。如果没有基础的话，可以先买一些书看一下原理，比如说《现代电视原理》《数字电视广播原理与应用》（本科的课本）中的部分章节。

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

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。

基于H.264标准的编码器还是很多的，究竟孰优孰劣？可参考：MSU出品的 H.264编码器比较（2011.5）

在学习视频编码的时候，可能会用到各种编码器（实际上就是一个exe文件），他们常用的编码命令可以参考：各种视频编码器的命令行格式

学习H.264最标准的源代码，就是其官方标准JM了。但是要注意，JM速度非常的慢，是无法用于实际的：H.264参考软件JM12.2RC代码详细流程

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。其基本教程网上极多，不再赘述。编码时候可参考：x264编码指南——码率控制。编码后统计值的含义：X264输出的统计值的含义（X264 Stats Output）

Google推出的VP8属于和H.264同一时代的标准。总体而言，VP8比H.264要稍微差一点。有一篇写的很好的VP8的介绍文章：深入了解 VP8。除了在技术领域，VP8和H.264在专利等方面也是打的不可开交，可参考文章：WebM(VP8) vs H.264

此外，我国还推出了自己的国产标准AVS，性能也不错，但目前比H.264还是要稍微逊色一点。不过感觉我国在视频编解码领域还算比较先进的，可参考：视频编码国家标准AVS与H.264的比较（节选）

近期又推出了AVS新一代的版本AVS+，具体的性能测试还没看过。不过据说AVS+得到了国家政策上非常强力的支持。

5. 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。

音频编码的简单原理可以参考：视频压缩编码和音频压缩编码的基本原理

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