

# 互联网视音频技术现状与展望

罗云川 姜秀华 刘平

**【摘要】**流媒体协议、封装格式、视音频压缩编码标准是目前互联网视音频主要技术。对比主要的几种视频编码标准和音频编码标准，以及视音频平台采用的技术标准，预测未来视音频技术的发展方向。根据全国文化资源共享工程中出现的部分节点采用私有方案、部分节点的方案偏旧等问题，建议视频编码标准采用H.264、音频编码标准采用AAC、文件封装格式采用FLV/MP4、点播采用的协议采用HTTP、直播采用的协议采用RTMP。

**【关键词】**互联网 视频 音频 标准

互联网视频从一个模糊的概念发展到高清甚至3D，并随着近2年PAD、智能手机等各种智能终端的兴起，更加刺激了互联网视频的高速发展。互联网视频已经从单一的PC端向包括电视屏幕在内各种终端延伸。根据CNNIC发布的数据，2012年中国网络视频用户增长了14.3%，达到3.72亿人。而欧美国家收看网络视频的比例已经达到80%以上<sup>[1]</sup>。

在当前这个互联网视频蓬勃发展的年代，互联网视音频技术显得尤为关键。本文将就互联网视音频技术中的流媒体协议，数据封装，以及音视频编解码技术进行深入研究。

## 一、互联网视音频技术

网络视音频技术主要包含以下几点：流媒体协议技术，封装技术，视频压缩编码技术以及音频压缩编码技术。

### （一）视音频数据格式

互联网上的视频文件通过视频播放器进行播放，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<sup>[2]</sup>（见如图1）。

解协议的作用，就是将流媒体协议的数据，解析为标准的封装格式数据。输入的数据采用的传输



图1 互联网视频文件播放步骤

协议种类很多，例如HTTP，RTMP，或是MMS等等<sup>[3]</sup>，采用这些协议的数据中通常包含一些控制视音频传输和播放的信令数据，这些数据如果送去解码，必然会造成系统出错。解协议操作即去除数据中的协议信息，并将数据处理输出成标准的封装格式。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TS，FLV，AVI，MP4，MKV，RMVB等等<sup>[4]</sup>，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。因此，解码视音频之前必须先解封装，才能得到分离开的视频/音频压缩编码数据。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<sup>[5]</sup>，视频的压缩编码标准则包含H.264，MPEG2，VC-1等等<sup>[6]</sup>。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。

（二）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。不同种类的流媒体系统通常使用不同的流媒

表1 主要流媒体协议一览表<sup>[3]</sup>

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

体协议。当前网络上主要的流媒体协议如表1所示。

由表1可见，RTSP+RTP这样的国际标准并未广泛用于互联网视频服务。因为互联网网络环境的不稳定性，像RTP这种采用UDP作为传输层协议的流媒体协议容易产生丢包等现象<sup>[7]</sup>，从而影响到传输的视频的质量。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。

（三）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，流行的封装格式如表2所示。

由表2可见，多数的封装格式都支持流媒体（除了AVI）播放，因此都可以应用于互联网视频系统。RMVB和TS因为客户端播放器不太普

及而并未广泛由于互联网视频网站。而MP4，FLV，MKV这样的格式是如今互联网视频网站采用的最多的封装格式。

（四）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

表3 主要视频编码一览表<sup>[6]</sup>

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（新）
MPEG4	MPEG	2001	各个领域（旧）
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

表2 主要封装格式一览表<sup>[4]</sup>

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

视频的数据量。视频编码是互联网视音频技术中最重要的技术之一。视频码流的数据量占了互联网视音频总数据量的绝大部分。高效率的视频编码在同等码率下，可以获得更高的视频质量。

由表3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。实验结果表明，同等码率下，不同视频压缩编码标准的视频质量有如下关系：

HEVC > VP9 > H.264 > VP8 > MPEG4 > H.263 > MPEG2。

（五）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

表4 主要音频编码一览表<sup>[5]</sup>

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

频编码标准，而导致音频数据量有所增加，也不会对互联网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码率下，可以获得更高的音质。

由表4可见，近年来并未推出全新的音频编码方案，现有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。实验结果表明，在低码率的情况下（48kbps），不同音频编码标准的音频质量有如下关系：

AAC+ > MP3PRO > AAC > RealAudio > WMA > MP3

AAC+可以获得最好的主观质量，而MP3编码的质量相对较低。

二、视音频平台对比

（一）全国文化共享工程参数对比

全国文化信息资源共享工程的很多节点都使用了互联网视音频技术，力求为网民带来中华民族几千年来积淀的各种类型的文化信息资源精华以及贴近大众生活的现代社会文化信息资源。通

表5 全国文化信息资源共享工程参数[点播]

名称	协议	封装	视频编码	音频编码	播放器
北京	MMS	WMV	WMV	WMA	WMP
上海	MMS	WMV	WMV	WMA	WMP
浙江	私有（泰初）				
陕西	私有（Vicworl）				
内蒙古	RTMP	FLV	VP6	MP3	Flash
湖北	MMS	WMV	WMV	WMA	WMP
河北	MMS	WMV	WMV	WMA	WMP
广东	HTTP	FLV	VP6	MP3	Flash
桂林	MMS	WMV	WMV	WMA	WMP
江苏	私有（泰初）				

过调查发现，很多节点都提供了具有地方特色的戏剧影视作品，视音频服务以点播为主，部分节点的视音频参数如表5所示。

（二）主流互联网视音频平台参数对比

调查显示，部分点击率较高的互联网视

音频平台的视音频参数包括一些电视台（如CNTV），以及视频分享网站（如优酷，土豆）。现有的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

表6 现有互联网视音频平台参数对比[直播]

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
CUTV	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

表7 现有互联网视音频平台参数对比[点播]

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

台，社交视频网站较为常见。主流互联网视音频平台直播服务的参数对比和主流互联网视音频平台直播服务的参数对比分别见表6和表7。

三、网络视音频技术的发展方向

（一）流媒体协议发展方向：点播：HTTP；直播：RTMP，RTMFP

点播服务使用HTTP的好处主要是减轻了服务器的压力，即当多媒体文件下载完成后就可以断开连接，从而节省出服务器资源再为其他客户端服务。此外，视频服务商可以直接使用普通的Web服务器而不必花费额外的资金购买流媒体服务器，从而节省了一笔经费。从调查数据可以看出：优酷网，土豆网，乐视网等大部分互联网视音频服务商都采用了该种方式提供多媒体服务。

直播服务使用RTMP或者RTMFP的好处主要是客户端不需要另外安装插件。一般的流媒体协议都需要相应的客户端播放器的支持。比如说MMS协议需要客户端安装Windows Media Player。而RTMP或RTMFP支持使用Flash播放器作为客户端，而Flash播放器在全世界PC上的普及度已经达到99%。因而，如果视音频服务网站采用RTMP或RTMFP传输流媒体，普通用户打开网页就可以收看直播，而不必做一些安装播放器的操作，这大大提高了用户的方便程度。此外，RTMFP是基于P2P的流媒体协议，在未来有着很大的发展潜力。从调查数据可以看出：很多知名网站都采用了RTMP。

（二）封装格式发展方向：FLV，MP4，MKV  
封装格式相比压缩编码以及流媒体协议来



说,不能算作一个非常关键的技术。而且不同封装格式之间的差别很小。因而很多推出比较早的封装格式(例如AVI, RMVB)和刚推出的封装格式(例如MKV)都占据的一定的市场份额。

就网络视音频来说,封装格式可以分为两种:支持流媒体的格式和不支持流媒体的格式。支持流媒体的格式(例如:FLV, MP4, MKV)可以用于点播服务,其特点是可以“边下边播”。不支持流媒体的格式(例如:AVI)则必须在下载完整文件后才能开始播放。从调查数据可以看出:FLV, MP4占据了很大的比例。此外, MKV作为一个先进的开源封装格式在未来有很大的发展前途。

### (三) 视频压缩编码技术发展方向: HEVC

视频数据占据了互联网视音频数据总量的绝大部分比例,因而是互联网视音频中最关键的技术。面对视频应用不断向高清晰度、高帧率、高压压缩率方向发展的趋势,当前主流的视频压缩标准协议H.264的局限性不断凸显。同时,面向更高清晰度、更高帧率、更高压缩率视频应用的HEVC协议标准应运而生。

H.264从2003年5月草稿发布以来,凭借其相对于以往的视频压缩标准在压缩效率以及网络适应性方面的明显优势,逐步成为视频应用领域的主流标准。根据 MeFeedia的数据,至2011年底,80%的视频使用H.264编码,并且随着支持H.264解码的设备不断增多,这一占有率还将进一步增长。

但是,随着数字视频应用产业链的快速发展,视频应用向高清晰度、高帧率、高压压缩率

几个方向发展的趋势愈加明显,原有的H.264在很多方面已经不能应对这些变化。2010年1月, VCEG和MPEG联合成立JCT-VC(Joint Collaborative Team on Video Coding),统一制定下一代编码标准:HEVC。第一版的HEVC/H.265视频压缩标准在2013年4月13日被接受为ITU-T(国际电信联盟)的正式标准。HEVC/H.265目标是在H.264标准2~4倍的复杂度基础上,将压缩效率提升一倍以上。

就调查的数据来看, H.264无疑是应用最广泛的视频压缩编码标准。它拥有成熟的软件支持和硬件支持。由于HEVC尚未成熟,在未来的2-3年内, H.264依然是最主流的视频压缩编码标准。但是随着时间的推移, HEVC会逐渐取代H.264成为主流的视频编码标准。

### (四) 音频压缩编码技术发展方向: AAC

音频数据占据了互联网视音频数据总量的小部分比例,因此音频压缩编码技术改进的需求并没有视频就压缩编码技术改进的需求那么大。

AAC将是MP3格式的接棒者。然而,由于MP3的空前流通,在目前来说,其他格式并未能威胁其地位。MP3不仅有广泛的用户端软体支持,也有很多的硬件支持,比如MP3播放器、DVD等。

由调查的数据可知,网络视音频平台已经普遍采用AAC作为其音频压缩编码的标准。鉴于其优秀的性能,在不久的将来, AAC应该也会在硬件等其他领域超过MP3,成为市场上最主流的音频压缩编码格式。

#### 四、全国文化信息资源共享工程改进方案

##### (一) 存在的问题

在当前全国文化信息资源共享工程的视音频方案中,部分节点的视音频参数有以下不足:

##### 1.部分节点采用私有方案

部分节点使用的是私有的流媒体解决方案(例如:泰初, Vicworl)。一方面,在收看这些节点视频的时候,必须安装相应的播放器,给用户的使用带来了诸多不便。另一方面,这些播放器对浏览器的支持还不完善,不兼容火狐、Chrome等浏览器,导致很多用户无法收看视频。

##### 2.部分节点的方案偏旧

由于很多节点使用的是微软的流媒体解决方案,然而微软的解决方案目前已经偏旧,当用户希望在网页上播放微软流媒体系统的视频时,必须通过Windows Media Player的ActiveX控件来播放视频。而像火狐、Chrome等浏览器并不支持此类ActiveX控件,这会导致很多用户无法播放视频。

##### (二) 改进方案建议

##### 1.视频编码标准采用H.264

H.264是当前主流视频编码标准,将近十年的发展已经让它的软件和硬件设备十分成熟,因此,在稳定性各方面很有保障。H.264能够在低码率情况下提供高质量的视频图像,而且有较强的网络适应能力,适合互联网视音频的传输环境。

##### 2.音频编码标准采用AAC

AAC是当前主流音频编码标准,虽然MP3依然如日中天,但是随着近年来支持它的设备日渐

增多,在不久的将来会逐渐取代MP3成为最普遍的音频编码标准。

##### 3.文件封装格式采用FLV/MP4

FLV/MP4是当前主流的封装格式,大部分网络视音频服务平台都采用了这一封装格式。

##### 4.点播采用的协议采用HTTP

HTTP点播实现简单而且节约服务器资源,并可以使用Flash播放器直接收看HTTP传输的视频而无需下载额外的播放器。HTTP是实现视频点播的最佳选择。

##### 5.直播采用的协议采用RTMP

RTMP可以使用Flash播放器直接收看而无需安装额外插件。而且RTMP是基于TCP协议进行传输的,因此,不会出现丢包情况,有较好的传输质量。RTMP是实现直播的最好选择。

#### 参考文献

- [1] 中广互联.在线视频平台——正在崛起的视频新产业 [EB/OL].[2014-03-30].  
<http://www.sarft.net/a/41884.aspx>.
- [2] Lei Xiaohua, Jiang Xiuhua, Wang Caihong. Design and Implementation of streaming processing software based on FFMPEG [C].2013.
- [3] Wikipedia. Comparison of streaming media systems [EB/OL].[2014-03-25]. [http://en.wikipedia.org/wiki/Comparison\\_of\\_streaming\\_media\\_systems](http://en.wikipedia.org/wiki/Comparison_of_streaming_media_systems).
- [4] Wikipedia. Comparison of container formats [EB/OL].[2014-03-31]. [http://en.wikipedia.org/wiki/Comparison\\_of\\_container\\_formats](http://en.wikipedia.org/wiki/Comparison_of_container_formats).
- [5] Wikipedia. Comparison of audio formats[EB/OL].[2014-02-28]. [http://en.wikipedia.org/wiki/Comparison\\_of\\_audio\\_formats](http://en.wikipedia.org/wiki/Comparison_of_audio_formats).
- [6] Wikipedia. Comparison of video codecs[EB/OL].[2014-03-20]. [http://en.wikipedia.org/wiki/Comparison\\_of\\_video\\_codecs](http://en.wikipedia.org/wiki/Comparison_of_video_codecs).
- [7] 苏佳,姜秀华. IPTV视频质量评价介绍[J].电视技术, 2011, 35(6): 78-81.

(作者单位:文化部全国公共文化发展中心,中国传媒大学)

责任编辑 冯佳