

可伸缩视频编码在移动互联网中的应用与实现

蒋宇浩,余林琛

(中国地质大学 计算机学院,湖北 武汉 430074)

摘 要: 为了应对移动互联网中终端和网络的异构性,尝试将 H. 264/SVC 应用到移动视频点播服务中,实现基于 SVC 的码流自适应播放系统. 借助 JSVM 研究并实现了不同空间、时间,以及质量层的可伸缩视频的生成;基于 VS 模拟了可伸缩视频在传输和播放过程中码率的动态切换;通过分析子层码流证实了可伸缩视频具有码率跨度大及解码质量好等特点. 最终验证了 H. 264/SVC 在移动互联网视频服务中所具有的优势,以及应用的可行性.

关键词: 移动互联网; H. 264/SVC; 码流自适应; JSVM

中图分类号: TP37

文献标识码: A

文章编号: 1000—7180(2016)12—0134—05

The Application and Implementation of Scalable Video Coding in Mobile Internet

JIANG Yu-hao, YU Lin-chen

(School of Computer, China University of Geosciences, Wuhan 430074, China)

Abstract: In order to deal with heterogeneity of the terminal and the network in the mobile Internet, the H. 264/SVC is applied into the mobile video-on-demand service, and a bit stream self-adaptively broadcast system based on SVC is implemented. With the aid of JSVM, the generation of scalable video with different space, time and quality layers is implemented. On the basis of VS, the dynamic switching is simulated in bit rate of scalable video in the process of transmission and broadcast. By analyzing the sub-layer stream, we confirmed that scalable video has a large span on bit rate and the quality of decoded video is good, and tested the advantages of H. 264/SVC in mobile Internet video services, and the feasibility of the application.

Key words: mobile Internet; H. 264/SVC; streaming self-adaption; JSVM

1 引言

随着移动互联网技术的飞速发展,移动视频点播日益普及. 然而,由于不同的终端在硬件(CPU、屏幕、内存等)方面存在着诸多差异,而不同的网络类型又存在着信道质量(带宽、丢包率等)的差别,这就使得我们面临着一个非常复杂的异构环境. 所以,在提供视频服务时,我们需要根据实际情况实时调控视频等级.

传统的解决方案包括两种:为系统添加转码设备和存储多分等级的视频. 这两种方案分别存在算法复杂度高,灵活性不足以及容易造成存储灾难等

缺陷和不足^[1]. 2007 年,联合视频小组(JVT)发布了 H. 264 的扩展标准——H. 264/SVC (Scalable Video Coding)^[2]. SVC 凭借其优秀的压缩效率和使用灵活性,很适合移动互联网异构环境下的多媒体服务^[3].

所以本课题尝试将 SVC 应用到移动视频点播服务中,实现基于可伸缩视频编码的码流自适应播放系统. 最终,成功实现了可伸缩视频的生成,以及播放过程中码率的动态切换,从而证明了课题的可行性. 并通过分析视频的子层码流,证明了 H. 264/SVC 能够很好地应对网络和终端的异构,体现出它在移动互联网应用中所占有的优势.

2 H. 264/SVC 介绍

H. 264/SVC 作为 H. 264/AVC 的扩展标准,可以通过一次编码得到包含不同时间、空间以及质量层级的分层序列^[4].

以往的可伸缩视频编码方案编解码机制复杂度较高,效率低下,未能达到实际应用标准^[5]. H. 264/SVC 因为继承了 H. 264 的编码模式所以具有较好的压缩效率,且拥有更加全面的分层机制,能够根据需要提供时间、空间以及质量等级上的任意组合^[6]. 目前,SVC 已被普遍用于视频会议产品中^[7].

然而,对于 H. 264 的继承也导致 SVC 的编码速度缓慢,进而导致其难以进一步大规模应用. 目前 H. 264/SVC 主要处在理论研究的层面上,编码速度的提升成为该标准的瓶颈所在,也是当前国内外研究的重点^[8].

可伸缩视频一般分为一个基本层以及若干个增强层. 如图 1 所示,解码器 A 作为低等级播放端,放弃了高层视频数据,仅接收和解码了基本层的数据,获得最低质量的视频. 而解码器 B 则接收了所有层的数据,获得最高质量的视频.

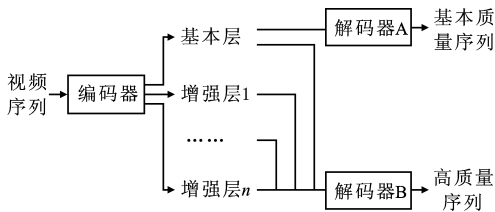


图 1 H. 264/SVC 总体概念

3 可伸缩视频生成

3.1 实现流程

生成可伸缩视频时,我们首先将原视频切割,得到视频分片;然后对分片进行解码,提取出相应的音频与视频,同时规范化视频流,转化为 $640 \times 360@25$ (非固定)的 yuv 数据;接着将规范后的 yuv 数据进行下采样,得到低空间分辨率的若干个次级 yuv 数据;然后以这些 yuv 数据为输入,以 JSVM 为工具,以设计好的配置文件为依据进行可伸缩编码,得到拥有不同时间层、空间层,以及质量层的可伸缩视频;最后,根据需要提取相应的子层码流并解码,从而得到不同等级的视频数据. 具体实现流程如图 2 所示.

3.2 配置文件

JSVM 是 H. 264/SVC 的官方参考模型,随着

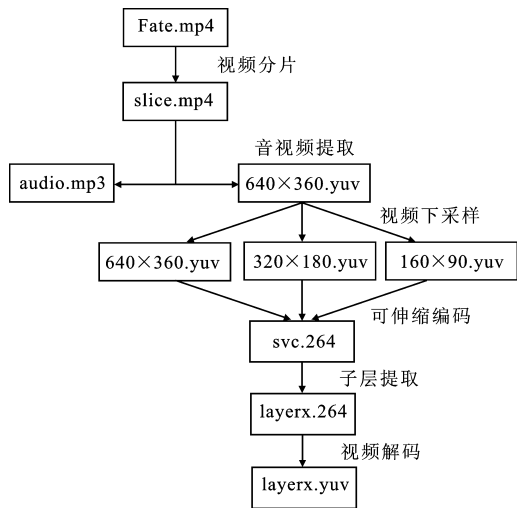


图 2 可伸缩视频实现流程图

SVC 项目的发展不断更新着,它提供了一系列的可执行文件用于 SVC 的编码以及数据分析^[9]. 在使用 JSVM 进行可伸缩编码时要用到配置文件,它能够根据具体需要,设置不同的时间层、空间层以及质量层,从而得到想要的可伸缩视频. 配置文件分为主配置文件和层配置文件. 表 1 所示为精简后的主配置文件,包含需要设计的必要参数. 下面简要介绍其设计过程.

表 1 JSVM 主配置文件

参数	值	描述
OutputFile	svc. 264	最终输出的视频
FrameRate	25. 0	输入视频最大帧率
FramesToBeEncoded	100	最大编码帧数
GOPSize	2	GOP 包含的最大帧数
BaseLayerMode	2	基本编码层模式
CgsSnrRefinement	1	质量分层模式
EncodeKeyPictures	1	关键帧编码模式
MGSCControl	2	中粒度非关键帧运动估计和运动补偿模式
SearchMode	4	搜索模式
SearchRange	32	搜索范围(Full Pel)
NumLayers	6	层配置文件数
LayerCfg	layer0. cfg	层配置文件名
.....		

这里计划生成包含 2 个时间层、3 个空间层、2 个质量层共 12 个层级的可伸缩视频. 输出文件命名为“svc. 264”;三个原始输入视频统一采用的 25f/s 帧率,所以最大帧率为 25;测试分片时长 4 s,25 f/s $\times 4\text{ s}=100\text{ f}$,所以最大编码帧数为 100 帧;若设计为 n 个时间层,则 GOP 最多包含 2^{n-1} 帧,所以

GOPSize=2;基本层编码模式统一采用包含 SEI 的 AVC 模式,赋值 2;质量分层模式采用中粒度,赋值 1;关键帧编码相应地采用中粒度模式,赋值 1;而非关键帧则采用运动估计和运动补偿相结合的模式,赋值 2;为了加快编码速度,我们选用快速搜索模式,赋值 4;与快速搜索模式相对应,搜索域设置为 32;我们总共分了 3 个空间层以及 2 个质量层,所以共有 $3 \times 2 = 6$ 个子层配置文件;后面依次列出相关配置文件名称.

3.3 视频分层结果

按照 3.1 节中的实现流程,采用 3.2 节中的配置文件,我们用 JSVM 对三个分辨率的原始视频进行可伸缩编码,最终得到了所预期的包含 12 个层级的可伸缩视频. 视频的具体分层情况如表 2 所示.

表 2 SVC 视频分层情况

层级	分辨率	帧率	码率	最小码率	DTQ
0	160×96	12.5	12.9	12.9	0, 0, 0
1	160×96	25.0	15.0	15.0	0, 1, 0
2	160×96	12.5	147.7		0, 0, 1
3	160×96	25.0	152.0		0, 1, 1
4	320×192	12.5	176.2	41.4	1, 0, 0
5	320×192	25.0	185.6	48.6	1, 1, 0
6	320×192	12.5	468.9		1, 0, 1
7	320×192	25.0	481.6		1, 1, 1
8	640×368	12.5	525.2	97.7	2, 0, 0
9	640×368	25.0	556.7	123.7	2, 1, 0
10	640×368	12.5	1 088.4		2, 0, 1
11	640×368	25.0	1 129.3		2, 1, 1

上述表格展示了 SVC 视频各个分层的详细信息,包括其所在的层级、视频分辨率(出于软件规范原因,视频分辨率显示得和样本视频有些差别,但不影响提取和解码的结果)、帧率、码率等. 参数 DTQ 显示的是该层码流详细的空间、时间,以及质量层等级(D:空间,T:时间,Q:质量).

3.4 实现流程

系统模拟了 SVC 视频传输和播放过程中的码率动态切换,分为提取、传输、解码、和播放 4 个线程. 各线程协同合作又相互制约,它们之间的结构和约束关系如图 3 所示.

图中各线程左侧为其启动条件,其中,extract_id(transfer_id、decode_id、show_id):提取(传输、解码、播放)线程当前处理的视频分片号;bufferNum:预设的缓冲分片数.

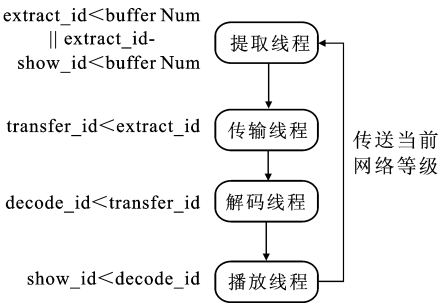


图 3 系统模拟框架图

首先提取线程提取 bufferNum 个最低级视频分片,然后挂起;当某个线程的前一线程有处理完的分片(即该线程有待处理分片)时,该线程启动;当某个线程的处理进度赶上前一进程时,该线程挂起. 当播放线程和提取线程处理进度缩短到一个 buffer-Num 之内,提取线程再次启动,根据播放端传来的当前网络等级提取下一个分片,从而实现实时的码流自适应播放. 四个线程协同工作,直到所有的分片处理完毕.

4 结果分析

4.1 视频效果评价

可伸缩视频编码的三个输入视频如下:

- 160×90.yuv(160×90@25);
- 320×180.yuv(320×180@25);
- 640×360.yuv(640×360@25).

我们将可伸缩视频的各子层码流提取并解码,然后将子层视频与相同分辨率的原始输入视频相比较,得到子层视频的峰值信噪比(PSNR),并算出子层视频的码率.

(1)子层码流传输码率

图 4 为子层视频(从 layer0 到 layer11)的传输码率分布图. 从图中可以看出,随着视频等级的上升,子层码流的传输码率逐渐上升. 码率范围涵盖从 layer0 的 14.08 kbit/s 到 layer11 的 1 131.854 kbit/s. 码率跨度较大,从而,增加了可伸缩视频的鲁棒性和灵活性,使其可以适应多样的移动终端和异构的网络环境. 也正因为如此,可伸缩编码在移动互联网应用中占有更大优势.

(2)子层码流 PSNR

图 5 展示了三种分辨率的子层码流相比于同分辨率的原始视频的 PSNR. 前四个等级的子层码流分辨率为 160×90;layer4 到 layer7 分辨率为 320×180;后面四个子层码流分辨率为 640×360(相同分

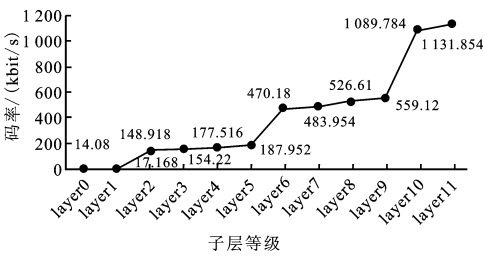


图4 子层码流传输码率分布图

分辨率的子层码流具有不同的帧率和质量等级).从图中可以看出,三种分辨率的子层码流 PSNR 呈相似变化趋势.

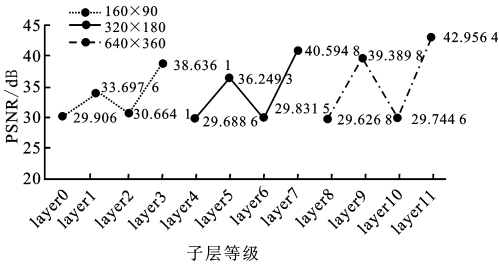


图5 三种分辨率的子层码流 PSNR

此处以 layer8 到 layer11(分辨率为 640×360)的子层码流为例进行说明. Layer8 和 layer10 帧率相同(12.5 f/s),质量等级前者为 0,后者为 1;layer9 和 layer11 帧率相同(25 f/s),质量等级前者为 0,后者为 1. 所以,帧率相同时,质量等级越高 PSNR 越高(帧率较高时该特点更为明显,如 layer9 和 layer11);质量等级相同时,帧率越高,PSNR 越高(此特点较为明显,如 layer8 和 layer9). 所以对于某一分辨率的子层码流,它的 PSNR 会随着帧率和质量等级的增高而变大. 另外,普遍认为,

- PSNR>40 dB:视频质量非常好;
- 40 dB≥PSNR>30 dB:视频质量较好;
- 30 dB≥PSNR>20 dB:视频质量较差;
- 20 dB≥PSNR:视频质量非常差.

从图中可以看出,子层码流的 PSNR 普遍在 30~40 dB 之间,甚至高于 40 dB. 说明可伸缩视频子层码流解码后还原度很高,很接近原始图像. 所以,H. 264/SVC 拥有较高的压缩性能,具有应用到移动互联网的意义与价值.

4.2 自适应码率切换

在系统模拟的过程中,根据播放端实时提供的网络等级,系统会动态地调整提供服务的分片的视频等级. 各分片播放质量如表 3 所示(播放视频并非采用的第 3 节的可伸缩视频,为了加快编码速度,测

试视频编码了 2 个时间层、2 个空间层、2 个质量层,共 8 个层级).

表3 分片播放详情

分片	网络等级	分辨率	帧率	质量等级	码率
0	0	320×180	25	0	38.4
1	0	320×180	25	0	38.3
2	4	640×360	25	1	468.9
3	2	640×360	12.5	0	291.5
4	0	320×180	25	0	38.1

图 6 显示的是播放过程中,各分片对应的网络等级以及此时提供服务的视频分片的码率. 可以看出,分片的传输码率跟随网络质量的动态变化而自我调节.

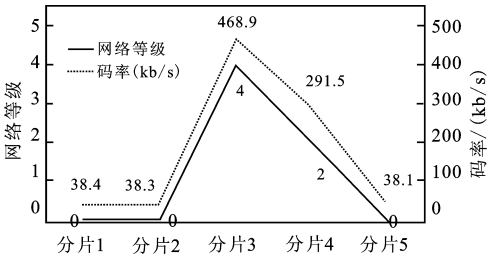


图6 分片码率随网络等级变化图

至此,初步实现了可伸缩视频编码在移动互联网中的应用.

5 结束语

为了应对移动互联网中终端的多样性以及网络的异构性. 本文尝试将 H. 264/SVC 技术应用到移动视频点播服务中,并成功实现了基于可伸缩视频编码的码流自适应播放系统. 通过对可伸缩视频的分析,证实了其码率跨度大以及解码质量好的特点,能够很好地应对移动互联网中终端和网络的异构. 通过实现可伸缩视频播放的码流自适应,证明了本课题的可行性.

参考文献:

[1] 冯进伟. H. 264/SVC 流媒体系统优化与实现[D]. 北京: 北京邮电大学, 2012.

[2] Schwarz H, Marpe D, Wiegand T. Overview of the scalable video coding extension of the H. 264/AVC standard[J]. IEEE Transactions on circuits and systems for video technology, 2007, 17(9): 1103-1120.

[3] 刘玲玲. SVC-P2P 流媒体视频质量自适应技术研究

[D]. 重庆: 重庆大学, 2014.

[4] Maarif H A Q, Gunawan T S, Priantoro A U. Complexity evaluation in scalable video coding[J]. Multi-media-An International Journal (AMIJ), 2010, 1(1): 12-25.

[5] Boyce J M, Ye Y, Chen J, et al. Overview of SHVC: Scalable Extensions of the High Efficiency Video Coding Standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26(1): 20-34.

[6] Kofler I, Kuschnig R, Hellwagner H. In-network adaptation of H. 264/SVC for HD video streaming over 802. 11 g networks[C]//Proceedings of the 21st international workshop on Network and operating systems support for digital audio and video. ACM, 2011: 9-14.

[7] 卢鑫,林茂六,金雪松,等. 新一代可伸缩视频编码标准:背景、特征、技术及其应用[J]. 电子测量与仪器学报, 2015, 29(10): 1415-1423.

[8] 任宏. H. 264 扩展标准—可伸缩视频编码(SVC)及其其中精细可伸缩编码的研究[D]. 成都: 电子科技大学, 2009.

[9] Schwarz H, Wien M, Vieron J. JSVM software manual[Z]. ISO/IEC JTC1 and ITU-T, CVS tag: JSVM_9_19_14, 2011.

作者简介:

蒋宇浩 男,(1993-),硕士研究生. 研究方向为可伸缩视频编解码研究与实现.

余林琛(通讯作者) 女,(1977-),博士,副教授. 研究方向为云计算、移动视频.

E-mail:linchenyu_cug@foxmail. com.

(上接第 133 页)

[7] Hyvarinen A. Fast and robust fixed-point algorithms for independent component analysis[J]. IEEE Transactions on Neural Networks,1999,10(3):626-634.

[8] Chenot C,Bobin J, Rapin J. Robust Sparse Source Separation[J]. IEEE Signal Processing Letters, 2015, 22(11):2172-2176.

[9] Kang Jaewook,Jung Hyoyoung, Kim Kiseon. Fast signal separation of 2-D sparse mixture via approximate message-passing[J]. IEEE Signal Processing Letters, 2015,22(11):2024-2028

[10] Ma Xiaohong, Ding Shuxue, SongJifei, et al. Underdetermined blind separation of weak sparse sources via matrix transform layer by layer in the Time-Frequency domain[C] // IEEE 7th International Conference on Awareness Science&Technology. Qinhuangdao, China:IEEE,2015:96-100.

作者简介:

王建功 男,(1982-),博士研究生. 研究方向为通信信号处理. E-mail:wangjiangong_78@163. com.

沈越泓 男,(1957-),博士,教授. 研究方向为通信与信息系统、通信信号处理.

魏红艳 女,(1982-),硕士,工程师. 研究方向为通信与信息系统、通信信号处理.