

# 基于传输的网络视频编码技术研究

## Research on the Web - Video Coding Technique Based on Transmission

张利飞 毛文林 张月红 孙西全 (西安交通大学电子与信息工程学院 710049)

**摘要:**视频传输是提高网络视频 QoS 的一个关键环节。随着视频传输的广泛应用,基于传输的网络视频编码技术受到越来越多的关注。文章系统的讨论了传统视频编码在传输中的应用,深入研究了 MPEG-4 视频编码,指出了基于传输的网络视频编码技术的发展前景。

**关键词:**视频编码 网络视频 可扩展性 MPEG-4

为了实现视频在网络上的传输,各研究机构提出两种不同的解决思路。一种思路是以网络为中心,要求网络中的路由器/交换机为视频传输提供带宽、延迟、抖动、丢失率的服务质量保证。另一种思路只涉及到端系统,不对网络提出要求,通过端到端的协议实现视频传输的 QoS 保障功能。第二种思路不依赖于现有的和未来的网络实现,因而是一种比较容易实现的方法。

### 1 网络视频编码系统的组成

网络视频压缩系统的组成<sup>[5]</sup>如图 1 所示,包括信息源编码器、信道编码器、调制器、解调器、信道解码器、信息源解码器和传输信道等。

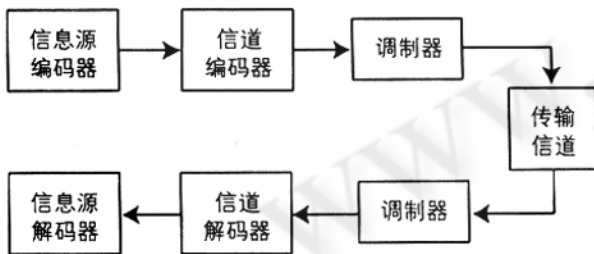


图 1 网络视频编码系统组成

### 2 传统视频编码在传输中的应用

传统视频编码在传输中有多种方式的应用。它不对网络提出要求,通过各自的特定协议或处理方式来适应不同的网络环境,主要是基于第二种思路。

#### 2.1 最小传输方式

为了解决视频在网络中传输带宽不足的问题,最简单的方法是最低传输,即按带宽最低用户的接受条件产生视频码流传送给所有用户。这种方法可以保证所有用户都获得视频,但是带宽足够的用户也只能观看质量较差的视频,带宽利用率极低。

#### 2.2 自适应编码

自适应编码的方法是根据信道的条件调整编码器的参数,生成适当的视频码流。通常采用 RTP/UDP/IP 协议,发送方将视频压缩数据划分成 RTP 数据包发送给接收方,接收方监测 RTP 包的传输延迟和丢失率等,采用 RTCP 协议将统计数据反馈给发送方,发送方根据网络带宽情况调整数据发送码率以使之相匹配。发送方根据反馈信息估计发送码率的方法主要包括基于探测的方法和基于模型的方法。

当发送方估计出发送码率后,就可以采用码率控制技术改变视频编码器的量化参数 QP,或者改变视频的帧率来生成适当的码流。自适应编码可以适应网络带宽的变化,但是它对服务器的要求太高。当用户增加时,服务器的开销太大。

#### 2.3 代码转换

为了降低自适应编码的复杂性,并且将编码和传输分离,人们提出了代码转换的概念。代码转换就是服务器保存一个质量足够好的压缩视频流,当需要降低输出码率时,服务器只需要进行部分的解码和编码就可实现。其操作是有选择性的丢弃压缩数据中不重要的部分,例如选择性丢帧、丢弃压缩数据中 DCT 参

数的高频部分或者再进行一次量化,丢弃后不会严重影响观看内容的质量。

### 2.4 码流切换

码流切换的基本思想是服务器保存同一个视频节目的不同质量的多个拷贝,这些码流的关键帧(I帧)是对齐的。当网络带宽发生变化时,服务器选择一个适当的码流传送给用户。

当码流切换和组播相结合时,这种技术常称为联播,如图2所示。它的基本思想:将同一段视频内容编码成多个不同码率的压缩视频流,将这几个视频流同时用独立的组播通道发到网络,用户根据自己的需求和网络情况选择接受某一个视频流,随着网络带宽的动态变化,用户可以在这几个视频流之间切换,以达到动态调整的目的。

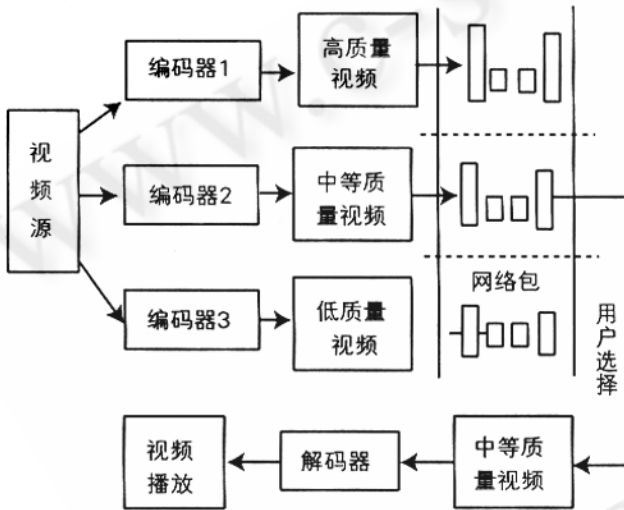


图 2 联播的码流切换原理

在码流切换方案中,服务器的计算复杂度很低,但是要有比较大的磁盘存储空间,实际上是将自适应编码中的计算复杂性转移到了存储。

### 2.5 分层可扩展性编码研究

分层可扩展性编码已经被许多现行的国际压缩标准所接纳,例如 H.263+ 和 MPEG-4。它主要有三种:时域可扩展性、空域可扩展性和质量可扩展性。其中质量可扩展性也称为 PSNR 可扩展性。

(1) 时域可扩展性。时域可扩展性编码通常是通过在码流中添加 B 帧来实现。B 帧是使用与它在时间上最近邻的前后两个 I 帧或 P 帧来预测的,而自己并不

作为任何其他帧的参考图像。因此在传输中丢弃 B 帧并不影响其他帧的质量,而仅仅降低帧率。

(2) 空域可扩展性。空域可扩展性编码是通过为视频中的每一帧都创建多分辨率的表示来实现的。当进行空域可扩展性编码时,原始视频首先通过下采样得到低分辨率的视频,编码得到基本层码流,然后编码原始视频和基本层视频的差生成增强层码流。

(3) 质量可扩展性。质量可扩展性编码的思想和空域可扩展性编码很类似,只是这里不需要对原始视频进行下采样,而是进行一次很粗的量化形成基本层码流。然后对原始视频和基本层视频的差再进行一次量化,生成增强层码流。如果有多个增强层码流,则重复上面的过程。图3是质量可扩展性编码的示意图。

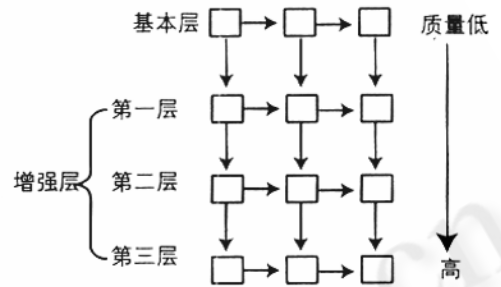


图 3 质量可扩展性编码示意图

将分层可扩展性编码与组播相结合也是解决视频在网络上传输的一种方法,称为分层组播。其基本思想是:发送方将每一层的压缩视频流发送到一个组播组中,用户根据自己的网络带宽状况接收基本层数据和多个增强层数据,然后解码获得不同质量的视频。

## 3 MPEG-4 视频编码研究

由于网络的异构性和缺乏 QoS 保证,带宽在一个很大的范围内发生变化,因此面向网络传输的视频编码的目标是将视频压缩成适合一个码率范围的码流。目前最具代表性的编码技术为精细可扩展性编码(FGS)和渐进的精细可扩展性编码(PFGS),MPEG-4 视频编码就是基于它们的。其基本思想是:将视频编码成一个可以单独解码的基本层码流和一个可以在任何地点截断的增强层码流,其中基本层码流适应最低的网络带宽,而增强层码流用来覆盖网络带宽变化的动态范围。

### 3.1 MPEG 标准的分层结构

MPEG 视频图像数据流是一个分层结构,目的是

把流中逻辑上独立的实体分开,防止语义模糊,并减轻解码过程的负担。MPEG 视频位流共包括 6 层,每一层支持一个确定的函数,或者是一个信号处理函数(DCT、运动补偿),或者是一个逻辑函数(同步、随机存取点)等。参看表 1。

表 1 MPEG 视频流语法的 6 个层次

图像序列层(随机存取单元:上下文)
图像组层(随机存取单元:视频编码)
图像层(基本编码单元)
宏块片层(重同步单元)
宏块层(运动补偿单元)
块层(DCT 单元)

每一层的开始有一个头,作为说明参数。在图像序列层中,装有视频序列参数,如图像宽度、高度、像素长宽比、帧率、位率、缓冲区尺寸等。

### 3.2 MPEG-4 的主要功能特点

(1) 基于对象的编码。基于对象的编码是 MPEG-4 的主要特征。对象的划分可以根据其独特的纹理、运动、形状、模型和高层语义为依据。这种编码是一种基于内容的数据压缩方式。以前的压缩算法只是去掉帧内和帧间的冗余,MPEG-4 则要求对图像和视频做更多的分析,甚至是理解。

(2) 基于对象的分级。分级工具主要用于因特网和无线网等窄带的视频通信、多质量视频服务及多媒体数据库预览等服务。MPEG-4 提供了两种基本的分级工具:时域分级和空域分级。时域分级是降低原视频序列的频率;空域分级是降低原视频序列的分辨率。在每类分级工具中,视频序列都可以分为两层:基本层和增强层。

(3) 支持基于内容的操作与比特流编辑。MPEG-4 无需编码就可进行基于内容的操作与比特流编辑。例如,用户可在图像或比特流中选择一个具体的对象,然后改变它的某些特性。

(4) 自然与合成数据混合编码。它提供将自然视频图像同合成数据(如文本、图形等)有效结合的方式,同时支持交互性操作。

(5) 基于内容的尺度可变性。内容尺度可变性意味着给图像中的各个对象分配优先级,比较重要的对

象用较高的空间/时间分辨率表示。它是 MPEG-4 的核心,因为一旦图像中所含对象的目录及相应的优先级确定后,其他基于内容的功能就比较容易实现了。对于甚低比特率应用来说,尺度可变性是一个关键因素,因为它提供了自适应可用资源的能力。

### 3.3 MPEG-4 视频编码的模式

MPEG-4 视频编码模式分为 VLBV 核 HBV 核两类,分别支持不同的比特率,如图 4 所示。

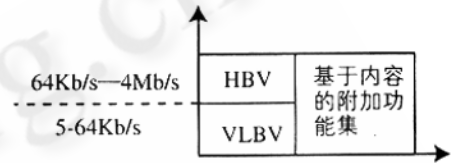


图 4 MPEG-4 的视频编码模式

VLBV 核提供了专为 5—64Kb/s 比特率视频操作与应用的算法及工具,支持较低空间分辨率(低于  $352 * 288$  像素)和较低的帧率(低于 15Hz)的图像序列。HBV 的范围为 64Kb/s—4Mb/s,采用与 VLBV 核相同的算法和工具。HBV 除了支持 VLBV 的功能外,还支持较高的空间与时间分辨率。

为了支持基于内容的图像编码,MPEG-4 视频标准引入了 VOP(Video Object Plane)视频对象平面的概念。假定输入视频序列的每一帧图像都被分割成多个任意形状的图像区域,则每个图像区域就称为一个 VOP。一个 VOP 包含图像的某一内容,如场景中的实物等。连续的属于同一对象的 VOP 序列,被称为 VO(视频对象)。这些 VOP 可以是任意形状,且形状与位置对于每帧可以是不同的。同一 VO 的 VOP 的形状、运动、纹理等信息被单独编码并传送。MPEG-4 基于内容的视频编码方法如图 5 所示。

VO 的生存期是一个镜头。MPEG-4 首先对视频序列进行镜头分切,对一个镜头中的每一帧进行物体分割,得到各个 VO,各个 VO 分别独立编码传输。解码器基本上为编码器的反过程。

## 4 结束语

视频数据与传统数据有着显著区别:数据量远远大于传统媒体数据,视频数据有质量属性。从用户角度来看,网络视频 QoS 由如下因素组成:视频质量、启

动延迟、抖动等。视频的传输是提高网络视频 QoS 最重要的环节之一。本文正是以此为基点,深入研究了基于传输的网络视频编码技术,将视频编码技术在

同传输中的应用进行了仔细分析。这些编码技术对于提高网络视频的服务质量有很大的帮助,具有很好的应用前景。

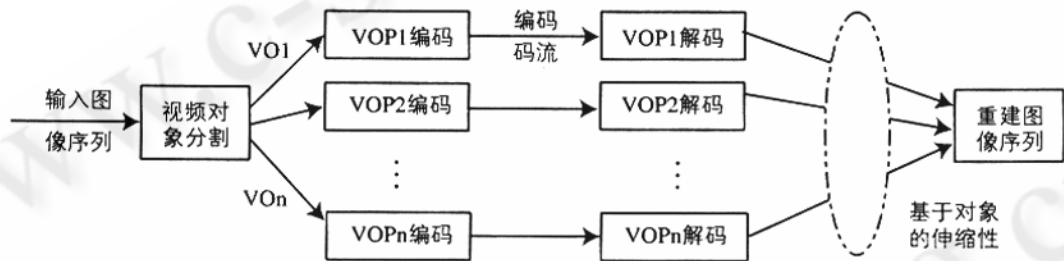


图 5 MPEG-4 基于内容的视频编码方法

### 参考文献

- 1 ISO/IEC International Standard 11172 - 2. Coding of Moving Picture and Associated Audio Information: Video. Nov. 1994.
- 2 Weiping Li, Fine granularity scalability using bit-plane coding of DCT coefficients. MPEG98/m4204, 1998.
- 3 Shipeng Li, FengWu, Ya-Qin Zhang. Study of a

new approach to improve FGS video coding efficiency. MPEG99/m5583, 1999.

- 4 Tanenbaum AS. Computer Networks. 3rd ed. Englewood Cliff, NJ: Prentice Hall, 1997.
- 5 詹青龙, 网络视频技术及应用, 西安电子科技大学出版社, 2004.
- 6 钟玉琢、向哲、沈洪, 流媒体和视频服务器, 北京清