

AVS-P2、P7 基于移动视频应用性能的比较

陈永华^{1,2}, 王毅¹, 刘东华², 杨立志²

(1. 湘潭大学 信息工程学院, 湖南 湘潭 411105;

2. 上海中科院计算技术研究所, 上海 201203)

摘要: AVS是《信息技术 先进音视频编码》系列标准的简称,是中国自主制定的音视频编码标准,主要面向高清晰度电视、高密度光存储和移动媒体等应用。它是一套包含系统、视频、音频、媒体版权管理在内的完整标准体系,其中视频标准包括两部分:面向数字电视应用领域的 AVS-P2 和面向移动应用领域的 AVS-P7。针对 AVS 两种视频标准基于移动视频应用领域上的关键技术进行比较,通过实验数据进行分析;对两种视频标准在移动视频领域的应用前景进行探讨。

关键词: 音视频编码标准;移动视频;性能比较

中图分类号: TN919.81

文献标识码: A

文章编号: 1673-629X(2008)07-0225-04

Performance Comparison between AVS-P2 and P7 for Mobile Video Application

CHEN Yong-hua^{1,2}, WANG Yi¹, LIU Dong-hua², YANG Li-zhi²

(1. Information Engineering School, Xiangtan University, Xiangtan 411105, China;

2. Shanghai Division Institute of Computing Technology, the Chinese Academy of Sciences, Shanghai 201203, China)

Abstract: AVS, the series of information technology advanced audio and video coding standard abbreviation, is a domestic audio and video coding standard of China. AVS is a set of integrity standard system which contains system, video, audio and media copyright management in the inside standard system. It includes two parts: AVS-P2 applying in DTV and AVS-P7 in mobile multimedia. In this paper, key techniques between the two video standards are compared, and the result are analyzed and proved through experiment data, and the prospect used in mobile video application is discussed in detail.

Key words: audio video coding standard; mobile video; performance comparison

0 引言

AVS中的第二部分, AVS-P2(以下简称 P2)在 2006年3月1日被正式批准为中国国家标准(GB/T 20090.2);第7部分, AVS-P7(以下简称 P7)为 AVS 系列标准中的移动视频编码标准,其目的是为手机等移动设备提供视频编解码标准^[1]。作为视频编码标准, P2、P7 在关键技术上有非常密切的联系,但是在技术细节上也各有特征。

1 移动视频应用及 P2、P7 技术特点

1.1 移动视频应用的特点

移动视频编码标准,需要满足低码率、稳定性好、

信噪比高以及算法复杂度较低等要求。移动网络带宽相对较窄,容易产生丢包,低码率对于移动网络上数据传输起着至关重要的影响,网络不畅通或丢包等出现时,编、解码应该不受影响或所受影响不大;而过低的信噪比用户是不能接受的,算法复杂度过高,也难于与性能相对较低的移动终端设备相匹配。

1.2 P2、P7 技术特点及流程

P2 本身使用帧编码和场编码,但是移动视频图像相对较小,比如 QCIF、QVGA、CIF 等,视频数据源按帧采集,完全适合用帧编码,且场编码算法过于复杂,不易于在移动设备上实现,所以通常不考虑场编码。P7 只使用帧编码。下面针对 P2 与 P7 的性能比较仅限于帧编码。

P2 与 P7 的编码核心是把数字视频压缩为原来的几十分之一甚至更小^[2],其压缩主要是:通过帧内预测消除图像内冗余,通过帧间预测消除图像间冗余,通过变换和量化消除图像内视觉冗余^[3]。P2 与 P7 均支持

收稿日期: 2007-10-22

基金项目: 上海市重大共性技术联合攻关项目(PKJ2006-L03)

作者简介: 陈永华(1974-),男,湖南宁乡人,硕士,研究方向为嵌入式系统;王毅,副教授,硕士生导师,研究方向为嵌入式系统。

采样为 4:2:0 的 YCbCr 视频格式,将各帧分成固定大小的宏块,其中亮度宏块以 16×16 个像素为单位,色度宏块以两个 8×8 个像素为单位。编码时,以宏块为单位进行预测、变换、量化,然后一方面进行熵编码,另一方面进行反量化、反变换得到重构图像作预测参考。此外还采用了环路滤波、条带等技术^[1]。而 P2、P7 编码的不同点在于:P2 采用了 P7 不具备的 B 帧预测、加权预测,而且在实现预测、亮度插值、变换、量化、熵编码和环路滤波等技术模块时,与 P7 也有较大的差异。

2 P2、P7 在移动视频应用的关键技术及性能比较

根据 P2、P7 编码流程,可以把编码分为以下几个主要模块进行比较:预测,变换、量化,熵编码,环路滤波和抗误码机制等。

2.1 预测

P2、P7 预测利用原始值和预测值之差的绝对值和(SAD)或者失真率优化(RDO)准则选择与之最相匹配的块。预测分为帧内预测和帧间预测。

2.1.1 帧内预测

帧内预测是在同一帧内利用相邻块预测当前块,选择与之最相匹配的预测模式和相应的块。一方面,P2、P7 均采用非直接预测模式,以适用不同图像特征;P7 还采用适合图像平坦区域、编码相对简单的直接预测模式,降低了码流。另一方面,就预测块大小和预测模式而言,P7 基于 4×4 块预测,支持 9 种亮度预测模式、3 种色度预测模式;P2 基于 8×8 块预测,支持 5 种亮度预测模式、4 种色度预测模式^[1]。性能比较:P2 减少了预测模式,且块数只有 P7 的 1/4,实现更为简洁、复杂度大幅度降低;P7 则基于 4×4 块,支持直接预测模式以及更复杂的预测模式,率失真略优于 P2。

2.1.2 帧间预测

帧间预测是在参考帧内搜索与之最相匹配的预测块和相应的运动矢量,对运动矢量预测编码。P7、P2 帧间预测主要不同在于宏块划分和预测方向。

P7、P2 宏块的划分均包括 16×16 、 16×8 、 8×16 、 8×8 四种类型,P7 的 8×8 块类型还可以进一步划分为 8×4 、 4×8 、 4×4 三种子块类型。子块类型的引入,进一步提高了搜索的精确度,但是也增加算法复杂度。较少的宏块划分,降低了运动矢量和模式传输的开销,也大幅度降低了编码器的复杂度;不足的是降低了一定的图像质量。

P2 与 P7 预测方向相差较大。P7 仅采用 P 帧前向帧间预测,参考前两帧;P2 不仅采用类似 P7 的 P 帧前向帧间预测,而且引入了 B 帧,采用 B 帧的前向、后

向、双向帧间预测。B 帧前向帧间预测参考前一帧;后向帧间预测参考后一帧;双向帧间预测参考前后各 1 帧,预测值兼顾前后两帧。

在移动视频应用上,采用 B 帧有两个突出的优点。一是码率低。B 帧有两种双向预测模式:直接模式和对称模式。直接模式中,当前块的前、后向运动矢量都是由后向参考图像相应位置块的运动矢量导出,无需传输运动矢量;而对称模式只传送前向运动矢量,后向运动矢量直接根据给出的前向运动矢量及前向距离索引、后向距离索引计算得出^[4]。因此都可以大幅度降低码率。二是稳定性更好。与 I 帧、P 帧不同,B 帧既不是关键帧,也不是参考帧,B 帧的丢失对后续视频数据解码不会带来任何影响,很适合于无线网络传输。

B 帧的缺点:其一是造成了微小的延时。但是,在移动视频应用上,无论采用何种视频标准,完全没有延时是不可能的,这种延时只能尽量减小。P2 减小延时的方法有很多:一是减少 B 帧的数目;二是加快移动服务端编码速度;三是在移动终端提高解码速度。上述三种方法都不难实现。从另一角度考虑,加入 B 帧后,如果以编码速度 15 帧每秒计算,B 帧仅增加延时 0.13 秒,这对用户来说并没有什么影响。其二是存储空间占用问题,解码器需要能存储两帧参考图像的缓冲区。基于这个问题,P2 允许 P 帧使用在其之前的连续两个 I/P 帧作为参考图像,在不增加缓冲区大小的前提下提高了编码效率;而且随着移动硬件设备的不断提高,内存空间也在不断提高。其三是 B 帧复杂度较高。但随着移动硬件性能的不不断提高,B 帧带来的硬件性能要求已经不难满足。

在运动矢量搜索中,搜索的精确度达到了 1/4 像素,相应的样本值通过插值来完成。P2 和 P7 插值由亮度插值和色度插值组成。对于 1/2 亮度插值,P7 采用一个水平 8 抽头滤波器($-1, 4, -12, 41, 41, -12, 4, -1$)、两个垂直 4 抽头滤波器($-1, 5, 5, -1$);P2 均采用 4 抽头滤波器($-1, 5, 5, -1$)。对于 1/4 亮度插值,P7 采用 2 抽头线性滤波器;而 P2 则主要采用 4 抽头滤波器($1, 7, 7, 1$),部分采用 2 抽头滤波器。比较起来 P2 的亮度插值综合了较多的像素个数,且利用较多中间值,因而精确度较高。对于色度插值,P2 和 P7 采用的方法相同,均采用 8 抽头滤波器,达到了较高的精确度。

此外,P2 采用加权预测,在改善编码效率、提高编码质量上,简易而有效;各宏块可以自由选择,在场景转换和照明变化时,效果尤其显著^[4]。

2.2 变换、量化

P2、P7 均采用整数离散余弦变换(ICT),只需加法

和移位即可实现,这样降低了算法复杂度也提高了编码器的效率。量化时,P2、P7 均采用 64 级量化,完全可以适应不同应用对码率和质量的要求。具体来说,P2、P7 采用的变换、量化方法如下:P2 通过式(1)、(2)完成变换,通过式(3)、(5)完成量化;P7 变换通过式(1)完成变换,通过式(4)、(5)完成量化。其中 P2 基于 8×8 块变换、量化,P7 基于 4×4 块变换、量化。其中 C_r 为变换矩阵, X 为图像数据块, $ScaleM[i][j]$ 为缩放参数, $Q_TAB[qp]$ 是与量化步长 qp 有关的量化参数, qp_const 为常数, $Abs()$ 是求绝对值函数, $Sign()$ 是求符号函数, $Z[i][j]$ 为得到的结果。P2、P7 参数均不相同。

$$Y = C_r X C_r^T \quad (1)$$

$$Z[i][j] = (Z[i][j] + (1 < < 14)) >> 15 \quad (2)$$

$$Z[i][j] = (((Abs(Y[i][j]) * ScaleM[i][j] + (1 < < 18)) >> 19) * Q_TAB[qp] + qp_const) >> 15 \quad (3)$$

$$Z[i][j] = (((Abs(Y[i][j]) * ScaleM[i][j] + (1 < < 14)) >> 15) * Q_TAB[qp] + qp_const) >> 19 \quad (4)$$

$$Z[i][j] = Z[i][j] * Sign(Z[i][j]) \quad (5)$$

由上述可以看出,P2、P7 变换、量化方法大致相同,主要区别在于基于块的大小不同。P7 基于 4×4 块,P2 基于 8×8 块。基于较大块,可以克服较小块的去相关性不足以及较小块在编码数据中引入较多的附加开销的问题,但复杂度略高。

2.3 熵编码

变换、量化后的系数矩阵,经 Zigzag 扫描形成(run, level)数对,映射成哥伦布码后按逆扫描的顺序写入二进制比特流。P2、P7 采用 k 阶哥伦布码进行熵编码压缩,通过采用自适应变长编码(CAVLC),多阶码表自适应切换,大幅度提高编码效率;同时进行熵编码压缩的还有运动矢量、预测模式等信息^[5]。P7 采用 0 阶为主、1 阶为辅的哥伦布码;而 P2 采用 2 阶为主、0 阶、1 阶和 3 阶为辅的哥伦布码^[1],相同长度的哥伦布码,阶数越高,能够表示的信息量越多,因而码率越低。

2.4 环路滤波

基于宏块的视频编码,不可避免产生块效应^[4]。P2、P7 采用一种自适应强度的环路滤波,可以很好地消除块效应,既达到改善重建图像质量

的目的,又可以提高编码效率^[4]。区别在于,P7 基于 4×4 块滤波,选取相邻边界 4 个样本点;而 P2 基于 8×8 块滤波,选取相邻边界 6 个样本点,一方面,编码块越大,块效应越小;另一方面,滤波时样本点选择越多,复杂度越高,但效果越好。

2.5 抗误码机制

P2 和 P7 支持抗误码机制,主要采取条带(Slice)划分。各条带数据独立编码,相互不影响,有利于视频应用中的错误隐藏、恢复需求等。

3 P2、P7 性能比较及分析实验

选用运动程度适中、细节变化程度分别为高、低的视频序列 mobile、foreman,视频格式分别为 CIF、QCIF 进行测试。图 1 实验中,P2 不编 B 帧,图 2 实验中,P2 采用一个 I 帧/P 帧后紧跟两个 B 帧,其它参数和 P7 相同,选择默认设置。从实验曲线图来看,不采用 B 帧时,P2 与 P7 性能基本相当,当 P2 采用 B 帧后,性能明显优于 P7。根据性能分析,前者主要是因为 P2 采用的 DCT、 8×8 块帧间预测、熵编码等比 P7 相应模块更好地降低了码率,以及通过环路滤波、加权预测等不同程度的比 P7 相应模块提高了信噪比;P7 则主要是采用了更高效的帧内预测, 4×4 块帧间预测,较大地提高了信噪比。P2 与 P7 各有所长,难分高下。而 P2 加入 B 帧后,B 帧大幅度降低了码率,性能明显提高。

4 结束语

虽然 P2 是面向数字电视应用领域的视频标准,但

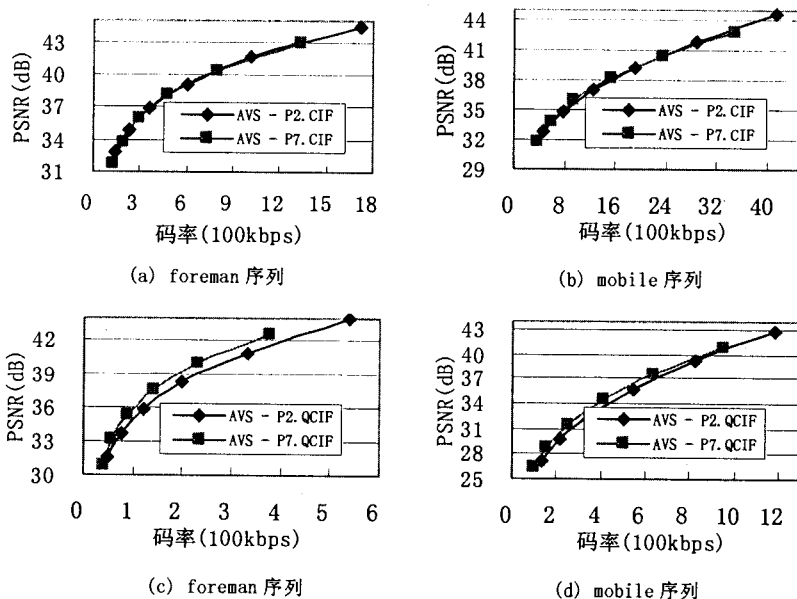


图1 不用B帧的P2和P7不同序列率失真曲线对比图

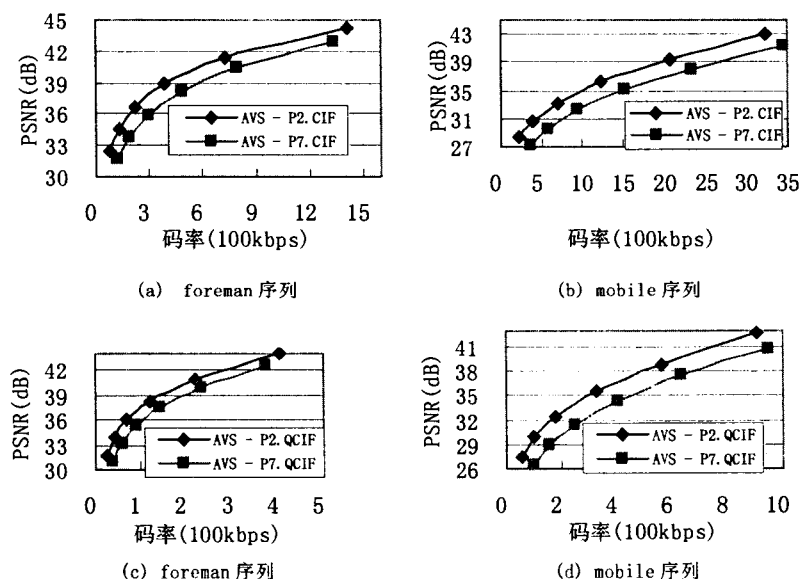


图 2 采用 B 帧的 P2 和 P7 不同序列率失真曲线对比图

是从以上的性能分析及实验结果表明, P2 不采用 B 帧时, 完全具备与 P7 相当的性能, 因此 B 帧带给 AVS 移动视频标准一个良好的选择余地; 在采用它独有的 B 帧后, 率失真曲线更好, 稳定性更佳, 虽然硬件要求略有提高, 但随着移动硬件性能的不不断提高, P2 带来的硬件性能要求也不难解决 (如在 Lenovo ET560 手机上,

图像大小 CIF, 采用 P2, 本地解码达到每秒 25 帧, 完全可满足实际应用)。2006 年, P2 已经被批准为国家标准, 无疑是对它的充分肯定和强有力的推动。同时, 随着 H.264、MPGE 系列国外标准的顺利推广和产业化逐步形成, 给国产视频标准带来了巨大的挑战。国内能否专注于统一的视频标准, 把有限的人力、物力资源投入到 AVS 视频标准中的哪一部分, 将关系到国产视频产业化发展和标准的国际化。

参考文献:

[1] 数字音视频编解码技术标准工作组. GB/T 20090.2 信息技术, 先进音视频编码,

第二部分: 视频[S]. 北京: 中国标准出版社, 2006.

- [2] 刘东华. 一种新的移动视频标准[C]//第十三届信息论学术年会. 长沙: 国防科技大学出版社, 2005.
- [3] 郭春辉. AVS1-P7 与 H.264 关键技术及性能比较[J]. 电信科学, 2006(1): 47-49.
- [4] 梁凡. AVS 视频标准的技术特点[J]. 数字电视与数字视频, 2005(7): 12-15.
- [5] 徐士麟, 喻莉, 崔春晖, 等. AVS1.0 中熵编码技术的研究[J]. 电视技术, 2005(10): 8-10.

(上接第 224 页)

来分别做为终端和座席端对所构建的模型进行了测试。

通过测试发现 eyebeam Phone 和 SIP Phone1.0 均支持两次 INVITE 来协商媒体信息, 可以随时地跟据信令信息更改目的媒体地址和端口, Avtech SoftPhone 不支持这种两次 INVITE 信令协商的模式, 它比较适合于两个终端直接通信, 但是通过服务器中转媒体流信息, 仍然可以实现双方的正常通信。因此, 构建的模型可以实现生成呼叫和清除呼叫等功能, 可以实现双方的正常通信。

3 结束语

主要以生成呼叫和清除呼叫为例, 阐述了如何通过 SIP 来实现 CSTA 的呼叫模型的方式, 对于其他的模型实现的方法类似。因此, 可以通过 SIP 来实现 CSTA 的呼叫模型, 并且根据呼叫中心实际情况的需要, 可以两端都使用基于 SIP 协议的 IP Phone, 也可以一端使用基于 SIP 协议的 IP Phone, 另一端使用基于其它协议的 IP Phone, 可以通过 IPPBX 进行信令的翻译 (H.323 和 SIP 协议的转换还没有完全实现)。例

如, 终端可以使用 SIP Phone, 而座席端仍然使用基于 H.323 协议的 IP Phone, 通过 IPPBX 的翻译, 将 SIP 信令翻译成 H.323 的模式信令, 这样就可以在不改变现有呼叫中心模式的情况下引入新的终端, 扩展了终端选择的灵活性。

参考文献:

- [1] 周海华, 边恩炯. SIP 原理与应用[M]. 北京: 机械工业出版社, 2006.
- [2] 邱克民, 田小华. H.323 呼叫模型的实现探讨[J]. 电信工程技术与标准化, 2003(8): 26-29.
- [3] Rosenberg J, Schulzrinne H, Camarillo G. SIP: Session Initiation Protocol[S]. IETF RFC 3261, 2002.
- [4] 张智江, 张云勇, 刘韵洁. SIP 协议及其应用[M]. 北京: 电子工业出版社, 2005.
- [5] Johnston A, Donovan S, Cunningham C. Session Initiation Protocol (SIP) Basic Call Flow Examples[S]. IETF RFC 3665, 2003.
- [6] ECMA. Services for Computer Supported Telecommunications Applications (CSTA) Phase II[S]. 1994.
- [7] Camarillo G. SIP 揭秘[M]. 白建军, 等译. 北京: 人民邮电出版社, 2003.