

面向监控应用的 AVS 视频编码标准技术

文/张贤国 张莉 梁路宏 马思伟 北京大学数字媒体技术研究所

摘要：本文详细介绍了AVS视频编码标准中的安防监控技术及其发展方向，并对AVS视频编码标准的框架、涉及到的技术和编解码工具等进行了简要描述，同时与H.264/AVC标准进行了比较，实验证明，AVS视频编码标准伸展档次在保证低复杂度、容错性和编码质量的同时，显著提高了安防监控视频的编码效率。因此，AVS监控标准将在安防监控中具有广阔的应用前景。

关键词：AVS 视频编解码 安防监控 伸展档次 背景建模

一、概述

AVS视频编码标准由成立于2002年的中国音视频标准（AVS）工作组制定，致力于制定中国自主知识产权的数字音视频的压缩、解码、处理和描述的技术标准，从而提高数字音视频编解码设备和系统的技术含量和经济价值。发展到现在，AVS工作组不仅正式发布了一系列编解码标准，也成为国际上最主要的音视频编码标准之一。由于不同视频应用对视频编码标准的复杂度、容错性、编码性能等需求各不相同，AVS视频编码标准定义了包括基准、移动、加强、伸展、立体在内的不同档次来满足不同视频应用需求，其中面向安防监控视频应用的伸展档次于2008年6月制定完成。

伸展档次主要面向标准的安防监控应用的视频压缩需求，在充分考虑了安防监控视频的一些特征如存在随机噪声大、场景切换少等的基础上，

在很大程度上适应了安防监控应用的特殊需求，解决了如视频编码位流存储时间长，存储代价大，对编码位流容错需求较大，对重建视频主客观压缩质量损失容忍度小，对编解码算法复杂度容忍度低，对事件检测和查询的友好性需求高等问题。

二、AVS视频编解码技术介绍

如图1所示，AVS视频编码器的模块结构是基于传统的预测/变换混合编码框架的。每个划分的宏块将按如下过程完成编码：首先对该宏块进行帧内或帧间预测进行，预测后得到

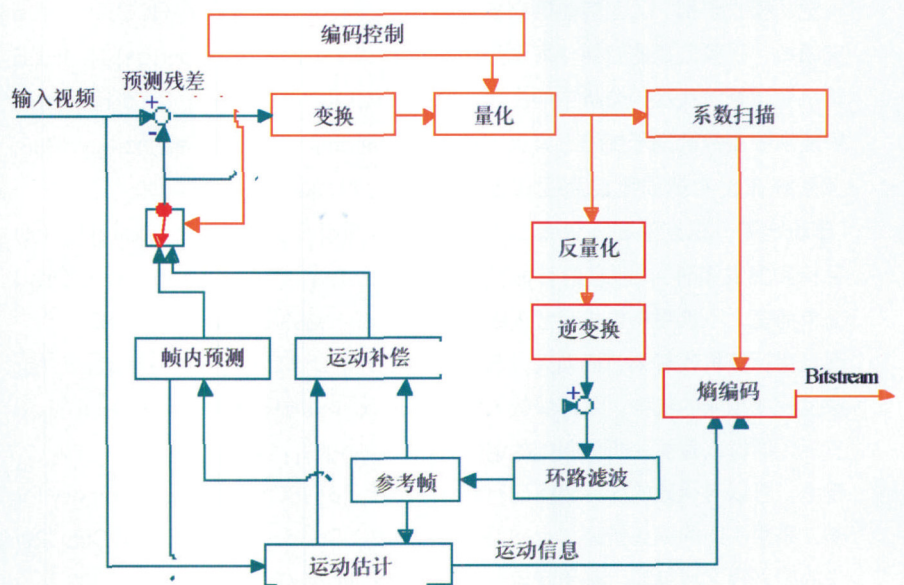


图1 AVS 视频编码器结构



的预测残差经过变换和量化后将得到量化系数矩阵，然后通过指定的系数扫描方式对系数重排列，并交由熵编码器编码后写入视频位流；在熵编码器编码量化系数的同时，量化系数还将经过逆变换与反量化得到重建的残差数据，该数据与预测样本叠加后将会得到重建图像，该重建图像经过滤波并写入图像缓存后将会用于后续宏块的预测过程。在此过程中，AVS的视频编码的主要优化工具可以分类为：帧内预测、可变块大小运动补偿、多参考帧帧间预测、四分之一或八分之一像素插值、加强的B帧直接模式预测、B帧对称模式预测、整数变换、量化及自适应加权量化、熵编码、隔行编码技术、抗误码技术、立体视频编码工具和监控视频编码工具。表1给出了AVS视频编解码标准的历史沿革，包括上述几类工具的发展过程。

三、AVS视频伸展档次——面向安防监控的优化技术

在安防监控领域，由于安防监控视频具有存储时间长，存储代价大、对容错需求较大、主客观压缩质量损失小、编解码算法复杂度低等特点，对视频编码标准的编码质量、压缩效率和容错性提出了更高的要求。因而，AVS视频编码标准引入了很多关键技术，并通过制定伸展档次来满足安防监控应用对视频编码标准的需求。AVS 伸展档次的技术主要包括背景帧（G帧）、背景预测帧（S帧）、灵活条带集、核心图像、受限DC模式、支持各类监控标记的语法元素、改进型的运动矢量缩放、自适应加权量化以及非参考P帧等。我们将首先分别说明这些技术所解决的问题和技术特征，并指出利用背景帧技术可以有效的支撑场景建模来实现高效的监控视频编码。

背景帧（G帧）：背景帧是帧内编码的场景图像，可以是输入序列中的一帧原始图像，也可以是非原始图像，它是背景预测帧的唯一参考帧。它除了具有帧内编码图像（I帧）的作用外，还具有显示监控视频场景特性的功能。与背景预测帧结合，可以显著提高监控视频的编码效率。

背景预测帧（S帧）：背景预测帧可以使用帧内预测和背景帧预测两种方式编码。在使用背景帧预测时，限定零矢量预测，而且只能16x16的块划分方式进行编码，即只能使用跳过模式16x16前向预测模式。背景预测帧可以实现部分帧内编码图像的功能，并且由于其简单的背景帧预测方式，可以保证背景预测帧具有较低的编解码复杂度的同时，得以显著提高整个视频的压缩效率。伸展档次在使用G/S帧时的解码输出顺序可以如图2所示。其中图像S10以G1为参考图

表1 AVS视频编解码标准的历史沿革

| 时间 | 文档 | 编码工具和信 |
|----------|-------------|------------------------------------|
| 2002年6月 | 基准档次需求 | |
| 2002年10月 | 基准档次工作草案 | 选定预测/变换混合编码框架 |
| 2003年7月 | 基准档次委员会草案 | 8x8整数变换、C2DVLC熵编码，B帧直接和对称模式 |
| 2003年10月 | 基准档次最终委员会草案 | 改进8x8变换、低复杂度滤波、插值、帧内预测等 |
| 2003年12月 | 基准档次标准草案 | AVS标准第二部分（视频）完成制定 |
| 2006年2月 | 基准档次成为国家标准 | AVS标准第二部分（视频）成为国家标准 |
| 2007年6月 | 伸展档次标准化开始 | 面向视频监控应用 |
| 2008年3月 | 加强档次标准化开始 | 面向多媒体娱乐应用 |
| 2008年6月 | 伸展档次最终委员会草案 | 背景帧、背景预测帧、核心帧、自适应加权量化，灵活条带，受限 DC模式 |
| 2008年6月 | 移动档次需求通过 | 手机应用 |
| 2008年9月 | 加强档次最终委员会草案 | 上下文二进制算术编码(CBAC)，自适应加权量化 |
| 2009年3月 | 移动档次委员会草案 | 可变块大小预测和变换，高精度插值，加权量化 |
| 2009年7月 | 移动档次最终委员会草案 | |
| 2010年6月 | 立体视频支持 | 语法修订 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|----|----|----|----|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| G | B | B | P | B | B | P | B | B | S | B | B | P |

图2 伸展档次的解码显示顺序示意图

像在上述低复杂度模式下完成编解码。

核心图像：出于安防监控应用的需求，监控视频位流对信道传输错误容忍度比较低，因而需要提高视频编码标准的抗误码性。AVS伸展档次在语法上可以指定若干图像为核心图像，由于核心图像一方面要求其本身最多只使用一帧最近的随机访问点后已正确解码的核心图像为参考图像，另一方面又要求核心图像后的第一个P或S帧只使用一帧参考图像，因而核心图像及其以后图像在解码时

与前面解码错误的图像的相关性就可以去除，实现核心图像后位流的抗误码。

有利于监控应用的语法元素：不仅上述的灵活条带集相关的语法元素支持一些基于感兴趣区域的监控应用，伸展档次在语法上可以支持单色和彩色序列编码，并且支持图像区域标记、区域事件标记、摄像机标记等，即用户可以在伸展档次下在视频流中的任意图像前添加上述标记信息，以支持各种基于事件检索的监控应用。见图3。

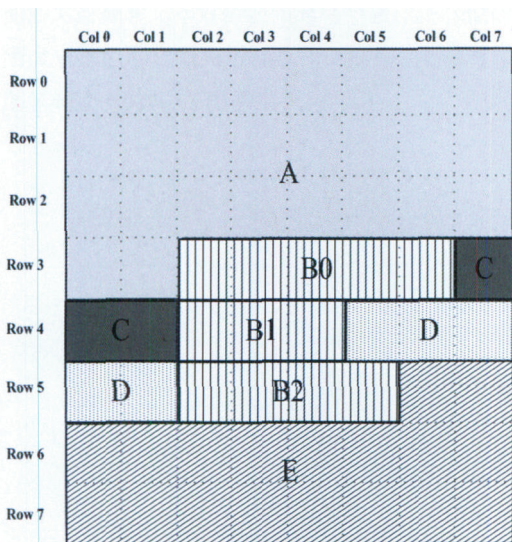


图3 灵活条带集示意图

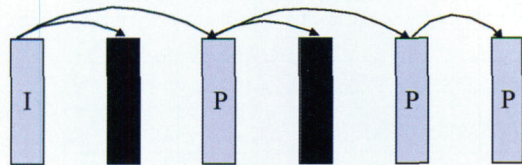


图4 具有非参考P帧的参考图像关系结构图

敏性特征，提高编码图像的主观质量。所决策的量化模式包括默认模式、考虑图像细节的模式、忽略图像细节的模式。实验证明该方法可以获得明显的主观质量提升。

改进型的运动矢量缩放：提出了基于半像素补偿的运动矢量缩放技术。由于场编码下顶场和底场图像在垂直方向上存在1行像素位置偏差，这将影响MV缩放的准确性。改进的运动矢量缩放有效的改善了场编码时的这一问题。

非参考P帧：由于在视频安防监控中，图像的清晰度比视频播放的帧率重要，当传输网络发生改变时，会首先考虑使用降低帧率的方式来满足传输带宽的限制，考虑到带宽的波动性和不同用户终端的解码能力，视频码流应具有多级帧率变化的功能。因而AVS伸展档次可以指定P帧是否可以作为参考帧，来消除部分帧之间的依赖关系，来满足视频传输中对多级帧率的需要。基于非参考P帧图像关系的结构如图4所示。

基于场景建模背景帧预测编码技术：针对监控视频场景具有场景相对固定，场景切换不频繁，噪声较大等特点，AVS伸展档次通过支持背景帧不显示输出，编码器应用可以通过场景建模得到输入视频的场景模型，并将其编码为背景帧，从而更好的完成背景帧预测，因而伸展档次可以支持如图5所示监控视频编码方式。在该编码方式下，由于背景帧可以是建模生成的场景图像，因而对于背景预测帧有更高的参考性，进而显著提高编码效率。

受限DC模式：为保证视频流不被传输错误的码流产生宏块间传递性的影响，在帧间预测帧的DC模式中，伸展档次支持增加比特标志以限制使用DC模式的当前块的相邻块可用性，以保证当前块使用帧内解码刷新时不被帧间预测的相邻块的错误累计所影响，进而保证视频在帧内刷新所应具有的差错恢复能力。

加权量化：该方法利用图像中相邻已编码宏块信息，对当前宏块进行自适应决策，在图像宏块级实现自适应量化以适应图像的内容特征，从而可以更好地符合人眼的视觉灵

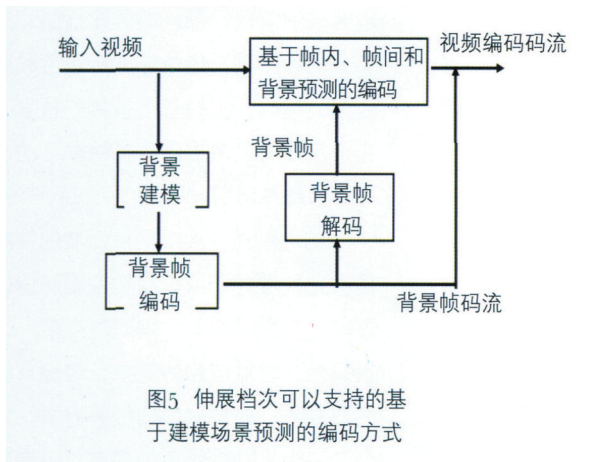


图5 伸展档次可以支持的基于建模场景预测的编码方式

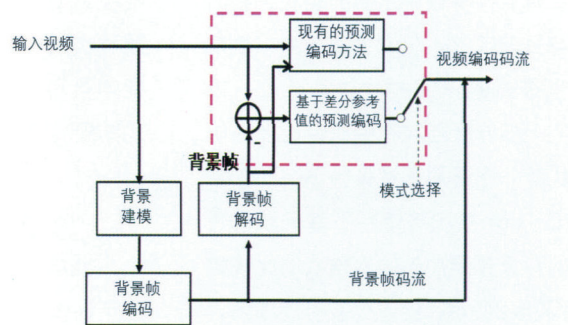


图6 AVS下一代伸展档次在审提案，基于背景建模的可选差分编码技术

AVS标准伸展档次的上述技术均已报批即将成为国家标准。此外，为了进一步解决安防监控应用中的问题，AVS下一代伸展档次的技术需求提案已经通过，正在征集更优化的编解码技术，一方面应引进新技术进一步显著提高编码效率，满足网络化条件下大量监控视频流同时实时传输的需要，明显降低存储成本；另一方面能够直接支持智能化监控所需的区域、对象和事件描述，并为监控视频的索引、检索、摘要等应用提供更便利条件。其中，以进一步提高编码效率的基于背景建模的可选差分编码技术已经经过AVS工作组的多次讨论和修改，该方法可以在监控视频的常用码率下可以比现有AVS伸展档次进一步提高5%~15%的监控视频压缩效率。该编码方法的编码结构如图6所示。

四、AVS视频伸展档次与H.264/AVC的性能对比

H.264/AVC标准也制定了包括Baseline、Main、High等在内的各个档次，但由于H.264/AVC并没有专

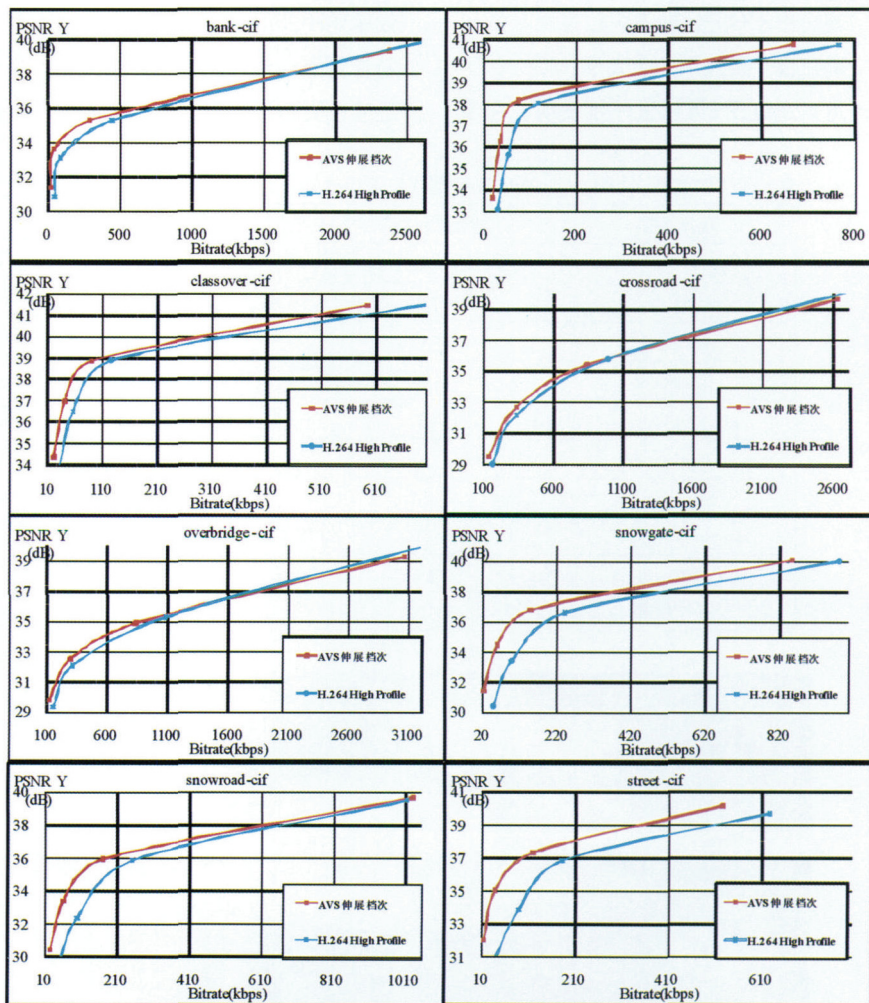


图7 AVS伸展档次与H.264/AVC High Profile率失真曲线对比

门针对监控视频制定新的档次，我们只与目前H.264/AVC性能最高、复杂度也最高的High档次进行比较，以证实AVS伸展档次在编码性能上的明显优势。在通用测试条件下，H.264/AVC High档次与使用了基于场景建模的背景预测AVS伸展档次的比较如图7和表2所示。经过比较可以看到，在图8所示的常见监控视频下，AVS伸展档次的要明显优于H.264/AVC High档次，可以实现近1个分贝的峰值信噪比提高，等效于35%的码率节省。由于CIF分辨率的监控视频的码率在极少条件下才会大于1Mbps，于是，通过图7可以进一步发现，在监

控视频编码常见码率范围内，AVS伸展档次比H.264/AVC High档次的性能增益和码率节省会比表2给出的平均结果高得更多。

五、总结

九年来，AVS在标准化过程中推出了许多极富有工业价值的技术，本文重点对涉及到安防监控领域的伸展档次技术做了深入的介绍，无论从技术本身还是从实验结果都可以看到，AVS伸展档次在典型监控场景的编码效率比H.264/AVC更高，它的抗误码特性更强，支持灵活的打包长度网络适应性更高，并且非常利于基于感兴

趣区域的编码、感兴趣区域检测、对象分割和跟踪，还能够支持单色和彩色序列编码，支持基于事件的检索，在一定程度上考虑到了多帧率下的时域可伸缩性问题。

综上所述，AVS视频标准的伸展档次充分考虑了监控视频应用对编码效率、主客观质量、编码复杂度、抗误码性、智能应用的支撑、传输环境的适应性、时域的伸缩性等需求，包含了一大批独具特色的国家自主专利技术，因而具有广阔的应用前景。

参考文献

[1]AVS工作组官方网站，<http://www.avs.org.cn>

[2]AVS标准文档、技术文档和参考软件下载地址，<ftp://124.207.250.92>

[3]法国电信，面向监控应用的背景帧和背景预测技术，AVS_M2189

[4]清华大学，AVS-S灵活条带集，AVS_M2305

[5]华为海思预研部，关于限制AVS-S的DC预测模式参考特性的建议，AVS_M2464

[6]华为海思预研部，基于半像素补偿的运动矢量缩放，AVS_M2329

[7]华为海思预研部，AVS X Profile宏块级自适应加权量化改进方案，AVS_M2331

[8]武汉大学，支持视频码流时域分级的编解码方式及其参考帧管理，AVS_M2384

[9]北京大学，基于背景建模和可选差分模式的视频编码方法，AVS_M2709

表2 AVS伸展档次与H.264/AVC High档次在监控序列下的测试比较结果

| 测试序列 | AVS伸展档次 VS. H.264/AVC High Profile | |
|------------|------------------------------------|--------|
| | 峰值信噪比提升 | 码率节省 |
| Bank | 0.71 | 39.64% |
| Campus | 0.91 | 40.34% |
| Classover | 0.68 | 30.70% |
| Crossroad | 0.30 | 8.61% |
| Overbridge | 0.36 | 12.55% |
| Snowgate | 1.55 | 47.19% |
| Snowroad | 1.29 | 41.10% |
| Street | 1.93 | 60.41% |
| 平均结果 | 0.97 | 35.07% |

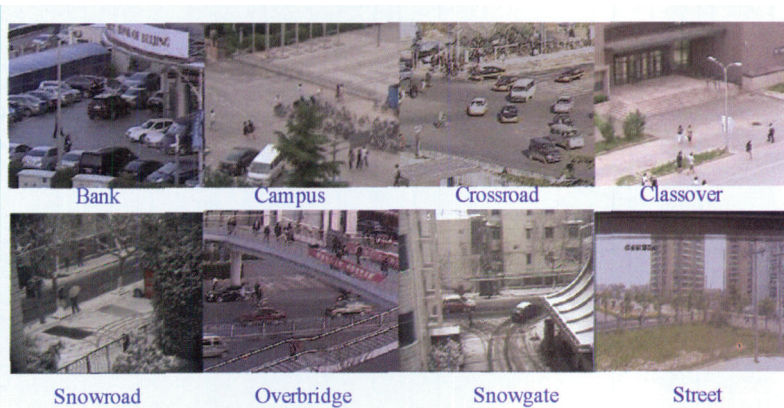


图8 AVS伸展档次与H.264/AVC High档次在比较时所使用的测试序列