



专题：视频技术的理论与实践

AVS2 视频编码标准技术特色及应用

马思伟¹, 罗法蕾², 黄铁军¹

(1. 北京大学信息科学技术学院, 北京 100871;

2. 中国科学院计算技术研究所, 北京 100081)

摘要: 围绕视频编码核心技术, 简单介绍 AVS 视频编码标准发展历程, 详细介绍了最新一代 AVS 视频编码标准——AVS2 (GB/T 33475.2-2016) 的核心创新技术, 主要包括: 灵活的预测划分方式、多假设帧间预测、优化的层次变换设计和自适应环路滤波处理技术以及面向场景视频应用 AVS2 提出的基于场景建模的高效预测编码技术。和第一代 AVS 视频编码标准相比, AVS2 编码效率提升一倍以上; 和同期 HEVC/H.265 国际标准相比, AVS2 在场景视频编码方面有显著优势。另外简单介绍了 AVS 标准在数字电视广播等行业中的应用情况。

关键词: AVS; 变换; 预测; 熵编码; 环路滤波

中图分类号: TN919.81

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2017245

Kernel technologies and applications of AVS2 video coding standard

MA Siwei¹, LUO Falei², HUANG Tiejun¹

1. School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China

2. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100081, China

Abstract: A brief introduction was given to the development history of AVS video coding standards. Then an overview of the key technologies adopted in AVS2 video coding standard (GB/T 33475.2-2016) was provided, including flexible prediction partition, multiple hypothesis prediction, optimized two-level transform, adaptive loop filter and background picture model based prediction coding for scene video. Compared to the AVS1, AVS2 can achieve more than 50% bits saving. And compared to HEVC/H.265, AVS2 can achieve significant coding efficiency improvement for scene video coding. Moreover, the applications of AVS standard in digital TV broadcasting was introduced briefly.

Key words: AVS, transform, prediction, entropy coding, loop filter

1 引言

1948 年香农发表的《通信中的数学理论》开启了编码理论研究之门, 至今即将迈入第 70 个年头, 其中提出的信源编码、信道编码理论对发展

高效的通信系统已产生巨大的影响。就信源编码来讲, 主要是研究文本、语音、音频、图像和视频等各种媒体信息的编码, 最近几年又出现了深度数据、点云数据和基因数据等新型数据类型的压缩, 但由于视频数据量之大最为突出, 长期以

收稿日期: 2017-07-10; 修回日期: 2017-08-08



来视频编码一直是信源编码领域的研究热点。以监控视频为例，北京市部署的监控摄像头数量达 50 万个之多，每小时产生的视频时长就相当于中央电视台建台以来的所有库存视频。因此，从传输和存储成本考虑，都需要研究压缩效率更高的视频编码技术。

由于视频压缩的重要性，几十年来学术界和工业界都对视频编码技术进行了长期而又深入的研究。国际上也一直有专门的标准组织制定视频编码标准，即 ISO/IEC MPEG 工作组和 ITU-T VCEG 工作组，制定了广为人知的 MPEG-1/MPEG-2/MPEG-4、H.261/H.263/H.264 等系列标准^[1,2]以及最新的 HEVC/H.265 标准。我国在图像和视频编码技术方面有多年的研究历史，但早期在国际标准制定中的参与度不高，标准技术几乎全部为国外若干公司所垄断，国内企业面临巨大的专利风险。2002 年，DVD 专利收费事件发生时举国震动，大批企业的主要利润纷纷外流，给我国视听产业造成很大压力。在经济全球化的大背景下，若没有掌握核心技术与标准，我国数字媒体产业的健康发展以及国家利益都得不到充分保证。

正是在这种背景下，2002 年，原信息产业部组织成立了数字音视频编解码技术标准（Audio Video coding Standard, AVS）工作组，目的是面向我国的信息产业需求，联合国内企业和科研机构，制定数字音视频的编解码标准，为数字音视频产业提供高效的编解码技术，服务于数字电视广播、无线宽带多媒体通信、互联网宽带流媒体等重大信息产业应用。至今已成功制定 AVS1/AVS+/AVS2 等系列标准并获得成功应用^[3]。特别是 2016 年底颁布的被列为国家标准的 AVS2，其性能已优于同期 H.265/HEVC 国际标准。尤其是对于场景视频编码，AVS2 标准具有显著优势，能够比 HEVC/H.265 节省 40% 左右的码率。

本文首先回顾视频编码的关键技术演变过程，然后介绍 AVS 视频编码标准的发展历程，并就 AVS2 视频编码标准（GB/T 33475.2-2016）介绍 AVS 在视频编码关键技术方面的特色创新。最后介绍 AVS 标准在数字电视广播等行业中的应用情况。

2 视频编码关键技术

视频能够进行压缩的原因在于数字化后的视频信号存在很多冗余，一般可将这些冗余归纳为空域冗余、时域冗余、统计冗余和感知冗余。空域冗余和时域冗余分别是指图像在空间和时间上存在很大相关性，如重复或非常接近的相邻像素值。统计冗余是指视频数据信息存在统计上的不均匀性。视觉感知冗余是指人眼对图像中的一些复杂细节信息感受不到，丢弃这些信息不会影响视觉感受质量。视频编码即要降低或消除这些冗余信息，对此研究者提出了一系列的编码方法，主要可以分为预测编码、变换编码和统计编码（熵编码）三大类。

早期的视频压缩技术由于在硬件上受到计算和存储资源的限制，主要是通过统计编码的办法降低数据的统计冗余，比如哈夫曼码、哥伦布码等编码方法。后来随着计算和存储资源的提升开始提出更复杂也更有效的预测、变换编码方法。预测编码的基本思想来源于 1952 年贝尔实验室 Culter 等人^[4]提出的差分脉冲编码调制系统，即用前一信号对当前信号进行预测，然后对预测差值进行编码传送。经过预测后，残差信号的能量比源信号的能量要低，从而降低了待编码信号的码率。同年 Oliver 和 Harrison^[5,6]将差分脉冲编码技术应用到视频编码中，对空域相邻像素进行差分预测编码，后来提出基于块的运动补偿预测技术使得预测编码性能获得极大的提升^[7,8]。

经过预测后的图像像素数据之间依然存在

较强的相关性,通过变换编码进行正交变换分解去相关可以达到进一步降低空间冗余的目的。变换编码于20世纪60年代后期被引入视频编码领域,早期的变换技术包括傅立叶变换、哈达码变换等。70年代初期提出了离散余弦变换(discrete cosine transform, DCT)^[9],并证明对空域相关性很高的图像信号离散余弦变换是最优的变换。由于DCT具有复杂度低、去相关性能好的优点,在视频编码中得到广泛应用,现今多采用从其演化出的类DCT整数变换^[10]。在最近的变换编码研究方面,对于随着预测距离增加相关性逐渐降低的帧内预测残差信号,离散正弦变换(discrete sine transform, DST)比DCT更有效^[11]。

70年代中期,预测编码开始与变换编码结合起来使用^[12],到80年代已形成了经典的预测/变换混合编码框架^[13]。基于预测/变换的视频编码框架如图1所示,首先通过对图像进行块划分,进行基于块的帧内预测或帧间预测降低空时域冗余,然后通过变换进行正交分解,在消除数据相关性的同时进行能量集中,可以用一部分变换系数来近似表示原来的能量信息,随后的量化过程即可达到这一个目的,量化将使部分高频系数量化为零,然后进行熵编码,从而达到压缩的目的。由于预测、变换都是基于块进行的,在低码率端会出现明显的块效应,对重构图像进行一些去块滤波处理不仅能提升主观质量,还能进一步提升编码效率。这一编码框架后来广泛应用于大家所熟知的MPEG、H.26x、AVS系列编码标准,直至今天最新一代的HEVC/H.265、AVS2标准也都是基于这一框架。此外,在这一编码框架中环路滤波也是非常重要的模块之一,它起源于早先主要用于提高主观质量的后处理滤波,自H.263+开始用于编码标准,从H.264/AVC开始广为应用。环路滤波不仅能提升主观质量,滤波后的

图像作为参考还能提升客观编码性能。在HEVC/H.265、AVS2标准的制定过程中,环路滤波技术进一步发展,除了去块效应滤波还增加了样值偏移补偿(sample adaptive offset, SAO)、自适应修正滤波(adaptive leveling filter, ALF)等滤波技术^[14]。

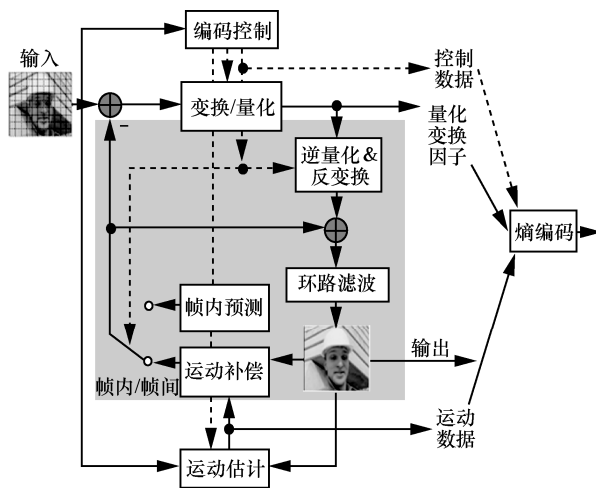


图1 基于预测/变换的视频编码框架

3 AVS标准发展历程及技术特色

2002—2017年的15年间AVS工作组一直致力于制定具有自主知识产权的视频编码标准,至今已成功地制定了面向高清数字电视、移动视频、监控视频等应用的高效视频编码标准。2012年,为了加快推进AVS在广播电视行业中的应用,国家广播电影电视总局科技司与工业和信息化部电子信息司联合发文共同成立“AVS技术应用联合推进工作组”(以下简称为AVS推进组),从此AVS标准发展进入快车道,并于2012年7月成功制定AVS+行业标准,如今AVS+已在数字电视广播等应用中取得重要进展。2011年底AVS启动了面向超高清应用的新一代视频编码标准AVS2的制定工作,2015年底标准技术制定完成,2016年5月获批为广电行业标准GY/T 299.1-2016,2016年12月正式获批为国家标准GB/T 33475.2-2016。AVS主要标准的制定发展过程见表1。



表 1 AVS 视频编码标准制定历程

技术完成时间	AVS / 类	应用及主要技术特征
2003 年 12 月	AVS1 / 基准	面向数字电视广播应用, 采用基于 8×8 像素块的帧内预测, 8×8 像素块变换编码, 去块效应滤波, 变块大小运动补偿 (8×8 像素~ 16×16 像素)
2008 年 6 月	AVS1 / 伸展	面向监控视频应用, 采用基于背景帧、核心帧的编码技术
2008 年 9 月	AVS1 / 加强	面向数字电影应用, 采用基于上下文的算术编码、加能量化等技术
2009 年 6 月	AVS1 移动	面向移动视频应用, 采用了 4×4 像素帧内预测、 8×8 像素/ 4×4 像素自适应变换等技术
2011 年 7 月	AVS1 监控	增强的监控视频编码, 采用背景图像建模编码技术
2012 年 5 月	AVS+ / 广播	面向高清数字电视广播, 采用了基于上下文的算术编码、帧级加权量化、增强场预测编码等技术
2015 年 12 月	AVS2 / 基准	面向超高清数字电视广播、场景视频等应用, 采用了自适应预测划分、多假设预测、层次变换、自适应算术编码、自适应滤波等技术

15 年来 AVS 吸引了上百家科研院所、企业公司的参与, 在视频编码制定过程中实现了视频编码核心技术的创新, 比如最新一代的 AVS2 标准已经在性能上优于 HEVC/H.265 国际标准, 甚至在场景视频编码方面实现了大幅度超越。下面简要介绍 AVS2 标准的主要特色技术, 包括灵活的预测块划分、多假设帧间预测、优化的层次变换、基于上下文的熵编码、自适应环路滤波和基于背景建模的场景编码。最后就 AVS2 与 HEVC 的性能加以比较。

3.1 灵活的预测块划分

从 20 世纪 90 年代的第一代 MPEG-1、H.261 标准, 到以 H.264/AVC、AVS1/AVS+ 为代表的第二代编码标准, 都采用基于块的预测变换编码技术, 即以 16×16 像素的宏块 (macroblock) 为基本的数据组织编码单位, 然后对宏块进行划分预测 (如 16×8 像素、 8×16 像素或 8×8 像素块等), 最后进行基于块的变换 (一般是 8×8 像素块变换), 由此可见划分的形式约束较多。而最新的 H.265/HEVC、AVS2 标准采用了基于四叉树递归划分的编码架构, 即一帧图像首先划分成连续不重叠的最大编码单元 (largest coding unit, LCU), 最大编码单元按照四叉树递归划分的方式划分成更小的编码单元 (coding unit, CU), 如图 2 所示。在 AVS2 中, 最大编码单元大小为 64×64 像素, 最小可划分成 8×8 像素的编码单元。编码单元的

最终划分模式一般通过率失真优化编码决策选择。编码单元大小确定后, 又进一步划分成预测单元进行帧内预测或帧间预测, 最后对预测残差进行变换编码。在 AVS2 中, 变换编码采用了类似编码单元划分的层次变换划分以选择最优的变换。

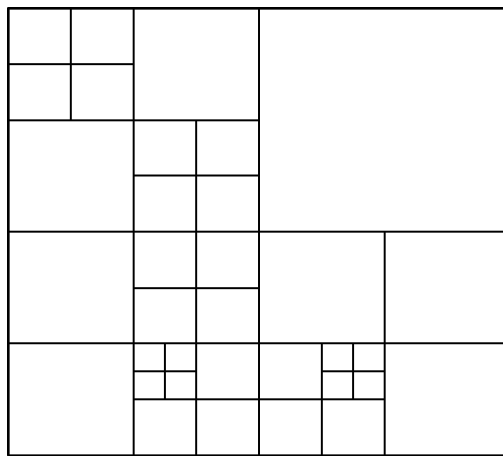


图 2 AVS2 编码单元 (CU) 四叉数划分示例
(最小块为 8×8 像素)

对于预测单元划分的形状, AVS2 也更加灵活。AVS2 的帧间预测划分方式如图 3 所示, 包括对称划分和非对称划分 (asymmetric motion partition, AMP), 其中非对称划分方式适用于 16×16 像素及以上 CU 大小。对于帧内预测, AVS2 有方形和非方形两种预测划分模式, 如图 4 所示。相比于方形划分, 非方形划分可以进一步缩短预测距离, 从而提高预测的准确度^[15]。考虑到复杂度和性能之间的

折中, 非方形划分仅适用于 32×32 像素和 16×16 像素的帧内预测编码单元。

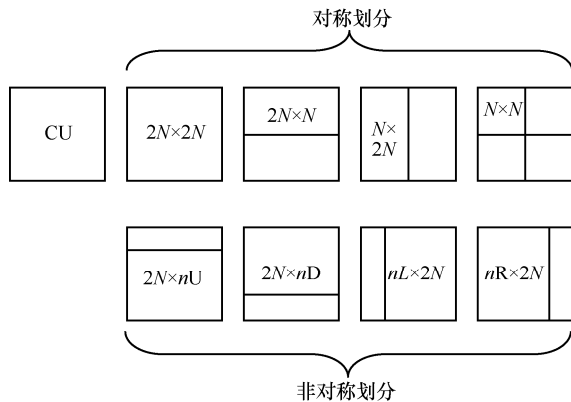


图3 AVS2 预测单元 (PU) 帧间预测划分模式

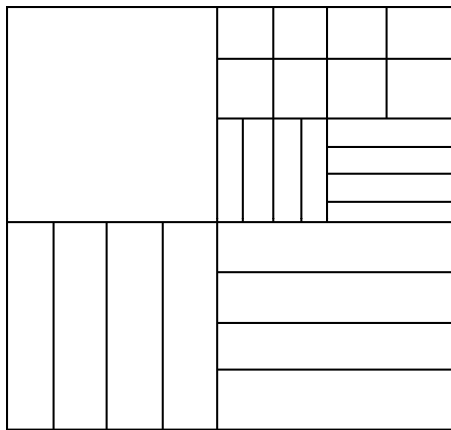


图4 AVS2 预测单元 (PU) 帧内预测划分模式

3.2 多假设帧间预测

帧间预测编码是一种有效降低时间冗余的预测方式, 从 MPEG-1 引入的前向 P 帧、双向 B 帧预测开始, 后来的帧间预测技术陆续又有多参考帧预测、层次参考帧预测^[16]以及针对静止运动或一致性运动的跳过模式、直接模式, 用于提高运动预测精度的 $1/4$ 像素精度预测, 用于提高运动矢量预测编码效率的中值预测、多运动矢量预测等编码技术。

类似 B 帧有两个方向的预测, AVS2 中新增加了一种支持前向双假设预测的图像, 称为 F 帧。按照双假设预测的来源可以分为时域双假设和空域双假设两类^[17,18]。时域双假设如图 5 所示, 当前块可以利用两个前向预测块的加权平均作为当

前块的预测, 其中只编码运动矢量 1, 运动矢量 2 由运动矢量 1 根据时间距离导出。空域双假设采用预定方向上的两个预测子的平均结果作为预测值, 如图 6 所示, 中心位置为初始预测子, 选择连线经过初始预测子的 4 种模式进行计算, 其中 $1/2$ 像素距离和 $1/4$ 像素距离位置的 4 种模式分别计算, 再加上初始预测值, 共有 9 种模式供选择。

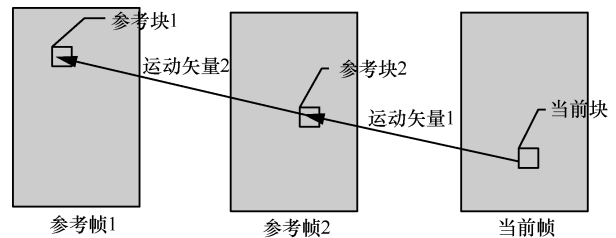


图5 时域前向双假设预测模式示意

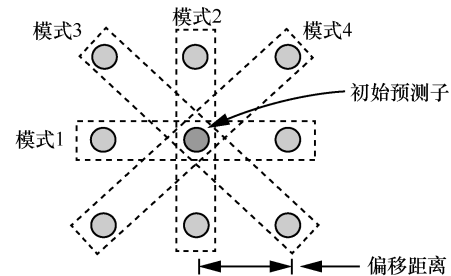


图6 空域双假设预测模式示意

3.3 优化的层次变换

类似编码单元对图像原始数据的划分组织方式, AVS2 对于残差数据采用基于变换单元的层次变换编码。AVS2 支持 4×4 像素~ 32×32 像素大小的方形进行 DCT, 并且对除了 $2N \times 2N$ 和 $N \times N$ 帧内预测以外的其他预测残差使用基于四叉树的非方形 DCT (non-square quadtree transform, NSQT)^[19]。与完全的方形变换相比, 非方形变换可以避免因跨预测块边界变换而增加高频系数, 提高了熵编码性能。为了降低变换复杂度, AVS2 中采用了一套简化的 NSQT 方案, 只允许亮度残差选择方形或非方形划分。如图 7 (a) 和图 7 (b) 所示, 分别是水平和垂直划分成 4 个非方形块的变换结构。此外, 对于 8×8 像素块大小



以上的亮度帧内预测残差, AVS2 对其 DCT 低频变换系数再多进行一次 4 点的二次变换, 如图 8 所示。二次变换使得系数分布更为集中, 从而提高帧内预测残差的压缩效率^[20]。

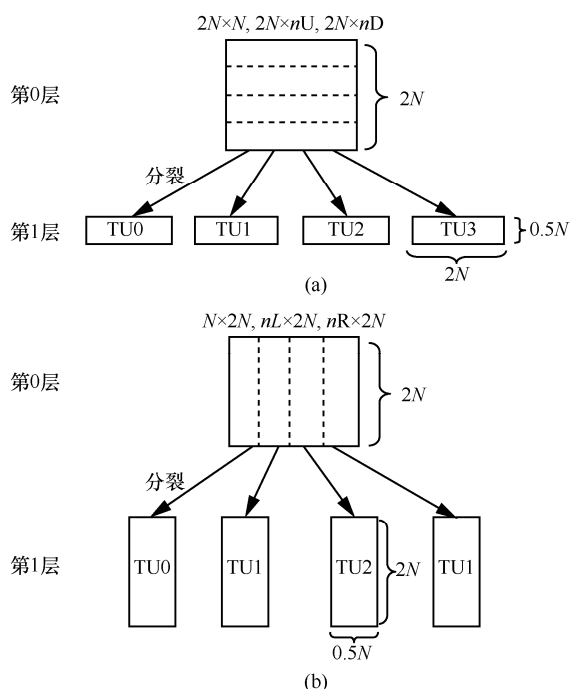


图 7 NSQT 示意

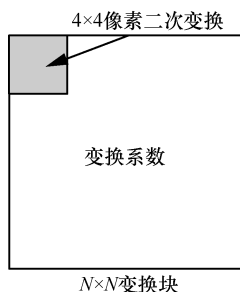


图 8 二次变换示意

3.4 基于上下文的熵编码

前面提到早期的编码标准主要采用哈夫曼码对编码系数进行熵编码, 优点是复杂度较低, 缺点是哈夫曼码对任何一个信源符号均需至少分配 1 bit 的码字, 特别是当实际信源概率分布跟预先设计的哈夫曼码概率分布不一致时会降低编码效率。算术编码通过联合编码信源符号, 理论上可以为每一个信源符号分配逼近信源概率分布的码字, 但由于复

杂度高一直在视频编码中未被应用^[21]。在 H.264/AVC 标准制定过程, 提出的基于上下文的二值算术编码很好地解决了算术编码复杂度问题^[22]。首先通过二值化实现信源符号的统一算术编码过程, 即将不同的信源符号转换成 0 或 1 的二进制串, 对不同信源符号的编码转换成仅对不同 0 或 1 的概率分布编码, 通过上下文选择不同的概率进行编码提高了信源符号编码效率, 同时通过查表实现算术编码过程中复杂的概率更新计算过程。

在 AVS2 中, 对高层语法如序列头、图像头、条带头采用变长编码 (variable length coding, VLC), 而对编码树 (coding unit tree) 级的其他语法元素采用高级熵编码 (advanced entropy coding, AEC)。无论是 VLC 还是 AEC, 首先都需要将各语法元素的值进行二值化, 得到二进制串。对于二值化, AVS2 有 3 种二值化方法: 一元码 (unary code)、截断一元码 (truncated unary code) 和指数哥伦布码 (exp-Golomb code), 可以将不同取值范围的信源符号高效地转换成二进制串进行编码。

进一步, AVS2 对变换系数编码采用两层组织编码机制, 即每个 TU 进一步划分成若干个 4x4 像素块系数组 (coefficient group, CG) 进行编码^[23]。图 9 描述了一个 8x8 像素的块编码过程, 即首先划分成 4 个 CG, 然后每个 CG 内的系数按照 Zig-Zag 扫描

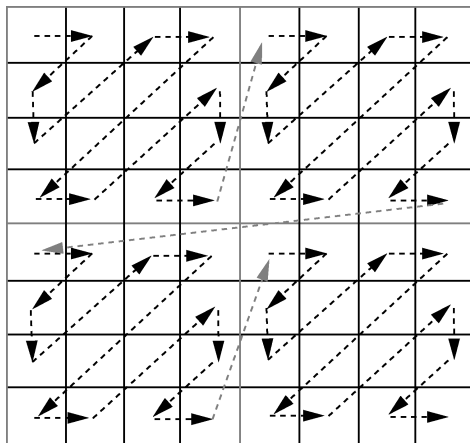


图 9 8x8 像素变换块中的 Zig-Zag 扫描顺序以及其划分的 4 个子块 CG 中的系数的扫描顺序

方式进行编码。基于CG的编码方式有两个优点：实现模块化编码，也就是不同变换块大小的包含的子块都可以使用的统一的方式进行编码；与扫描整个变换块相比，在软件和硬件实施方面，都降低了计算复杂度。同样，对于 16×16 像素的TU划分成16个子块， 32×32 像素的TU划分成64个子块，如图10所示。

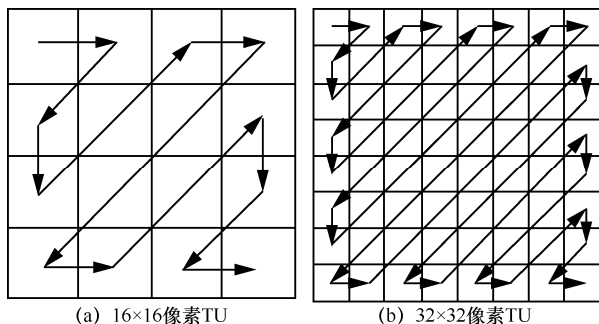


图10 两层变换系数扫描方式

3.5 环路滤波

环路滤波是视频编码器的重要模块之一。早期的滤波处理主要是作为一种视频编码后处理手段，处理后的图像即送给显示模块进行显示，环路滤波即体现在这些滤波处理处在编码环路中。AV2中的环路滤波包括3个模块，即去块效应滤波、样值偏移补偿^[24]和自适应修正滤波^[25]，分别用于滤波处理编码带来的不同类型的失真。去块效应滤波主要是用于去除由于基于块的编码在块边界所产生的块效应失真，通过滤波平滑处理可以有效提升主观质量。样值偏移补偿是指在去块效应滤波之后基于像素进行的非线性滤波操作，基于量化失真的统计特性，进行基于区间或边缘模式的失真补偿操作。区间模式是指将0~255的像素值划分成若干连续的区间，根据失真统计情况为相应区间的像素分配一个滤波偏置值，将这个偏置值叠加到相应像素位置进行滤波。边缘模式分为水平、垂直和两种对角模式4种，如图11所示。根据当前滤波后样本的值与其相邻的滤波后样本的值的关系查表得到偏置值，对当前像素进行滤波。

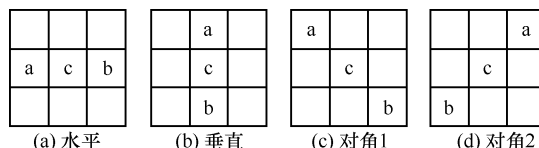


图11 样值偏移补偿的边缘模式下样本和相邻样本的位置关系

自适应修正滤波对样值偏置滤波后的重建图像进行类似维纳滤波的处理，以进一步减小编码失真，提高参考图像质量。样本补偿滤波器形状如图12所示，为 7×7 十字加 3×3 方形的中心对称滤波器，滤波器系数 $C_0 \sim C_7$ 的取值范围是 $[-64, 63]$ ，滤波器系数 C_8 的取值范围是 $[0, 127]$ 。对于亮度分量，将图像按照LCU对齐的规则均匀划分为16个区域并训练对应的16套滤波器系数，而对于色度分量(Cb/Cr)分别只训练一套滤波器系数。对于亮度分量的16套滤波器系数，编码端根据率失真性能进行自适应合并，得到最优的滤波器数目。亮度和色度分量通过帧级标志和最大编码单元级标志来分别控制滤波器是否被使用。

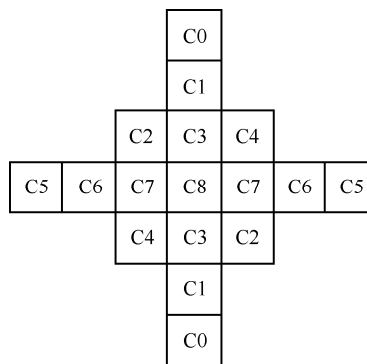


图12 样本补偿滤波器形状

3.6 场景编码

在诸如视频监控、视频会议等应用中，视频内容有其自身的特性，即背景内容通常保持不变，称此类视频为场景视频。和传统影视视频相比，场景视频的数据冗余还可以被进一步挖掘，比如可以通过建立高效的背景视频模型降低背景信息冗余。为此AVS提出了基于背景建模的视频编码方法^[26]，针对监控视频、视频会议等场景视频内容实现编码效率比同期国际标准提升约一倍左右。



AVS 提出的基于背景建模的监控视频编码流程如图 13 所示, 首先经过背景建模、背景编码、背景解码后, 生成高质量的重建背景图像; 然后, 编码器使用该重建背景完成对每幅图像进行背景预测以降低背景部分的编码代价, 进而得到更优的编码效率; 最后, 背景图像码流和视频序列码流都传输给解码过程用于解码。

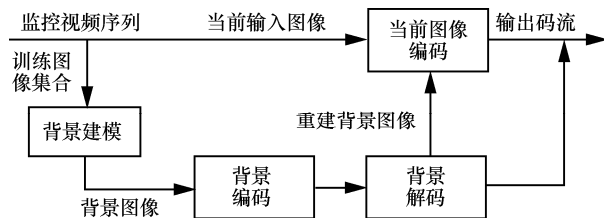


图 13 基于背景建模的监控视频编码框架

3.7 性能比较

使用 AVS2 参考软件 RD12.0 的编码性能与第一代标准 AVS1 对比结果见表 2。在随机访问 (random access, RA)、低时延 (low delay, LD) 和全帧内 (all intra, AI) 编码 3 种配置下, AVS2 分别获得了 52.9%、51.0%和 31.2%的亮度性能增益。与 HEVC 参考软件 (HM16.6) 的性能对比见表 3, 在随机访问和全帧内编码配置下, AVS2 相比 HEVC 有 0.9%和 0.1%的性能提升; 在低时延配置下, AVS2 相比 HEVC 有 2.1%的性能损失。

针对监控类视频, AVS2 与 HEVC 的性能对比见表 4, 在随机访问和低时延配置下, 基于背景建模的场景编码能分别提升 39.1%和 31.3%的编码性能。

表 2 AVS2 与 AVS1 对比性能测试

配置序列	RA			LD			AI		
	Y	U	V	Y	U	V	Y	U	V
UHD	-50.5%	-56.3%	-56.2%	-57.6%	-66.5%	-67.1%	-31.2%	-30.8%	-30.5%
1080P	-51.3%	-61.7%	-63.2%	-44.3%	-59.0%	-62.1%	-33.1%	-37.4%	-39.2%
WVGA	-52.8%	-58.1%	-58.6%	-50.5%	-58.1%	-59.9%	-30.4%	-28.2%	-29.8%
WQVGA	-52.4%	-58.9%	-59.0%	-49.4%	-59.2%	-58.7%	-26.6%	-24.8%	-26.9%
720P	-57.2%	-66.1%	-63.8%	-56.3%	-71.1%	-69.3%	-34.0%	-37.7%	-34.8%
平均	-52.9%	-60.5%	-60.5%	-51.0%	-62.1%	-62.8%	-31.2%	-32.4%	-33.0%
编码复杂度	1 210%			2 102%			1 228%		

表 3 AVS2 与 HEVC 对比性能测试

配置序列	RA			LD			AI		
	Y	U	V	Y	U	V	Y	U	V
UHD	-0.3%	5.6%	6.3%	2.7%	4.8%	9.0%	-2.2%	2.1%	2.1%
1080P	-2.3%	5.6%	3.8%	0.7%	-2.4%	-1.6%	-0.7%	2.1%	1.2%
WVGA	0.0%	6.7%	7.7%	0.9%	2.2%	5.2%	1.5%	3.9%	3.8%
WQVGA	1.1%	7.7%	8.8%	4.9%	8.1%	9.8%	2.8%	4.5%	4.9%
720P	-2.4%	-7.3%	-5.5%	1.9%	-5.1%	0.3%	-2.1%	-5.3%	-4.2%
平均	-0.9%	4.1%	4.6%	2.1%	1.3%	4.1%	-0.1%	1.4%	1.4%
编码复杂度	315%			312%			327%		

表 4 AVS2 基于背景建模的编码与 HEVC 对比性能测试

配置序列	RA			LD		
	Y	U	V	Y	U	V
Crossroad_720×576_30	-39.2%	-37.3%	-37.7%	-25.6%	-64.6%	-60.4%
Office_720×576_30	-24.0%	-24.7%	-25.0%	-15.5%	-46.1%	-44.3%
Overbridge_720×576_30	-60.7%	-56.9%	-56.8%	-38.4%	-58.3%	-56.2%
Intersection_1 600×1 200_30	-18.7%	-29.0%	-28.4%	-23.2%	-33.7%	-29.6%
Mainroad_1 600×1 200_30	-52.7%	-65.2%	-58.6%	-53.8%	-58.9%	-54.3%
平均	-39.1%	-42.6%	-41.3%	-31.3%	-52.3%	-49.0%
编码复杂度	334%			332%		

由上述结果可以看出,第二代标准 AVS2 比第一代标准 AVS1 有突出的性能提升,与同期国际标准 HEVC 相比已在场景编码方面大幅领先。

使用 AVS2 快速编码器 CAVS2 的性能与 HEVC 开源编码器 x265 对比见表 5 和表 6,和 HEVC 参考软件的性能对比见表 7。测试中使用 HEVC 通测条件^[27]中的 5 个 1080P 高清序列,其中 x265 使用 2.4 版本。CAVS2 使用的量化参数 QP 设置为 27、32、38、45,与 AVS2 通测条件保持一致,为保持两者码率和质量接近,HEVC 测试使用的 QP 为 24、28、33、38,对应 x265 使用的 QP 为 27、31、36、41,均使用单线程进行编码。此外,在 RA 配置下,相邻的 P 帧之间的 B 帧数量为 7 个,其中 x265 的超快速档次下的数量为 3 个。

可以看到在超快速~平和 5 个档次下,CAVS2

在速度与 x265 接近的情况下都具有明显性能优势,其中 LD 配置下平均增益分别为-16.54%、-4.70%、-10.43%、-8.71%、-9.90%,RA 配置下平均增益分别为-3.65%、-9.51%、-15.94%、-5.80%、-8.83%。

和 HM16.6 相比,在超快速~平和 5 个档次下,CAVS2 在 LD 配置下平均性能损失分别为:107.97%、51.97%、29.20%、17.55%、11.00%,在 RA 配置下平均性能损失分别为 87.51%、39.76%、18.11%、12.60%、5.31%。其中序列 Cactus 的 RD 曲线如图 14 所示,5 个序列的平均性能损失与编码速度曲线如图 15 所示。

4 AVS 标准应用

AVS 的出发点是实现知识产权的自主可控,让标准回归为产业发展服务的本位,支撑我国数

表 5 RA 配置下 AVS2 快速编码器 CAVS2 与 x265 性能对比

项目序列	配置	超快速(配置 0)	比较快速(配置 3)	中速(配置 5)	慢速(配置 7)	平和(配置 9)
BasketballDrive 1 920 dpi×1 080 dpi, 50 帧/s	Y BD-Rate	11.36%	-12.55%	-18.98%	-4.29%	-11.15%
	加速比(x265)	14.02	7.02	4.10	0.64	0.10
	加速比(CAVS2)	8.75	3.80	1.32	0.75	0.12
BQTerrace 1 920 dpi×1 080 dpi, 60 帧/s	Y BD-Rate	-26.95%	-4.19%	-10.93%	-3.67%	-3.59%
	加速比(x265)	19.46	11.65	5.68	0.96	0.12
	加速比(CAVS2)	10.00	5.83	2.03	0.99	0.11
Cactus 1 920 dpi×1 080 dpi, 50 帧/s	Y BD-Rate	-7.25%	-15.63%	-19.48%	-9.06%	-14.10%
	加速比(x265)	16.23	8.06	4.95	0.75	0.11
	加速比(CAVS2)	10.28	4.37	1.56	0.83	0.12
Kimino 1 920 dpi×1 080 dpi, 24 帧/s	Y BD-Rate	11.48%	-6.57%	-14.27%	-6.63%	-9.06%
	加速比(x265)	13.78	6.41	3.91	0.79	0.12
	加速比(CAVS2)	10.44	4.94	1.77	0.94	0.14
ParkScene 1 920 dpi×1 080 dpi, 24 帧/s	Y BD-Rate	-6.87%	-8.63%	-11.55%	-5.33%	-6.25%
	加速比(x265)	16.09	7.91	4.71	0.84	0.12
	加速比(CAVS2)	9.54	4.65	1.73	0.92	0.12
平均	Y BD-Rate	-3.65%	-9.51%	-15.04%	-5.80%	-8.83%
	加速比(x265)	15.92	8.21	4.67	0.80	0.11
	加速比(CAVS2)	9.80	4.72	1.68	0.89	0.12



表 6 LD 配置下 AVS2 快速编码器 CAVS2 与 x265 性能对比

项目序列	配置	超快速(配置 0)	比较快速(配置 3)	中速(配置 5)	慢速(配置 7)	平和(配置 9)
BasketballDrive 1 920 dpi×1 080 dpi, 50 帧/s	Y BD-Rate	3.30%	2.19%	-3.54%	-0.33%	-4.99%
	加速比(x265)	12.99	7.18	4.25	0.63	0.10
	加速比(CAVS2)	8.30	3.15	0.81	0.40	0.07
BQTerrace 1 920 dpi×1 080 dpi, 60 帧/s	Y BD-Rate	-60.50%	-20.96%	-24.22%	-25.57%	-20.77%
	加速比(x265)	16.64	11.02	6.00	0.86	0.11
	加速比(CAVS2)	7.79	4.43	1.43	0.58	0.08
Cactus 1 920 dpi×1 080 dpi, 50 帧/s	Y BD-Rate	-12.32%	-6.44%	-11.32%	-8.51%	-11.65%
	加速比(x265)	14.74	8.51	5.26	0.71	0.11
	加速比(CAVS2)	8.91	3.50	0.96	0.45	0.07
Kimino 1 920 dpi×1 080 dpi, 24 帧/s	Y BD-Rate	9.59%	7.33%	-0.37%	1.37%	-2.17%
	加速比(x265)	13.00	7.09	4.29	0.74	0.11
	加速比(CAVS2)	9.83	3.92	1.02	0.48	0.08
ParkScene 1 920 dpi×1 080 dpi, 24 帧/s	Y BD-Rate	-22.77%	-5.63%	-12.70%	-10.51%	-9.91%
	加速比(x265)	14.55	8.47	5.03	0.76	0.11
	加速比(CAVS2)	8.19	3.78	1.03	0.48	0.07
平均	Y BD-Rate	-16.54%	-4.70%	-10.43%	-8.71%	-9.90%
	加速比(x265)	14.38	8.45	4.97	0.74	0.11
	加速比(CAVS2)	8.60	3.76	1.05	0.48	0.07

表 7 AVS2 快速编码器 CAVS2 与 HM 性能对比

项目序列	配置	超快速(配置 0)	比较快速(配置 3)	中速(配置 5)	慢速(配置 7)	平和(配置 9)
BasketballDrive 1 920 dpi×1 080 dpi, 50 帧/s	Y BD-Rate (LD)	108.10%	49.68%	28.79%	17.64%	8.43%
	Y BD-Rate (RA)	100.70%	44.37%	20.36%	15.02%	3.12%
BQTerrace 1 920 dpi×1 080 dpi, 60 帧/s	Y BD-Rate (LD)	186.08%	86.19%	49.47%	28.78%	25.33%
	Y BD-Rate (RA)	128.49%	54.41%	27.43%	17.67%	12.66%
	Y BD-Rate (LD)	97.37%	45.54%	23.76%	14.39%	5.34%
	Y BD-Rate (RA)	80.08%	35.74%	15.01%	11.11%	1.67%
Cactus 1 920 dpi×1 080 dpi, 50 帧/s	Y BD-Rate (LD)	69.76%	38.31%	22.02%	14.18%	7.73%
	Y BD-Rate (RA)	65.26%	33.91%	13.81%	9.80%	3.92%
Kimino 1 920 dpi×1 080 dpi, 24 帧/s	Y BD-Rate (LD)	78.56%	40.13%	21.97%	12.79%	8.19%
	Y BD-Rate (RA)	63.03%	30.37%	13.93%	9.40%	5.21%
	Y BD-Rate (LD)	107.97%	51.97%	29.20%	17.55%	11.00%
	Y BD-Rate (RA)	87.51%	39.76%	18.11%	12.60%	5.31%

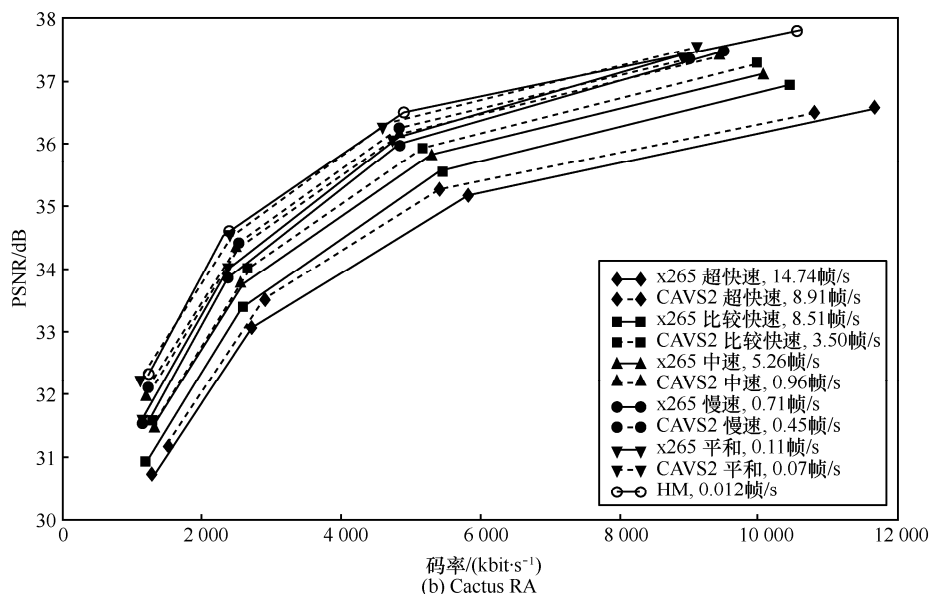
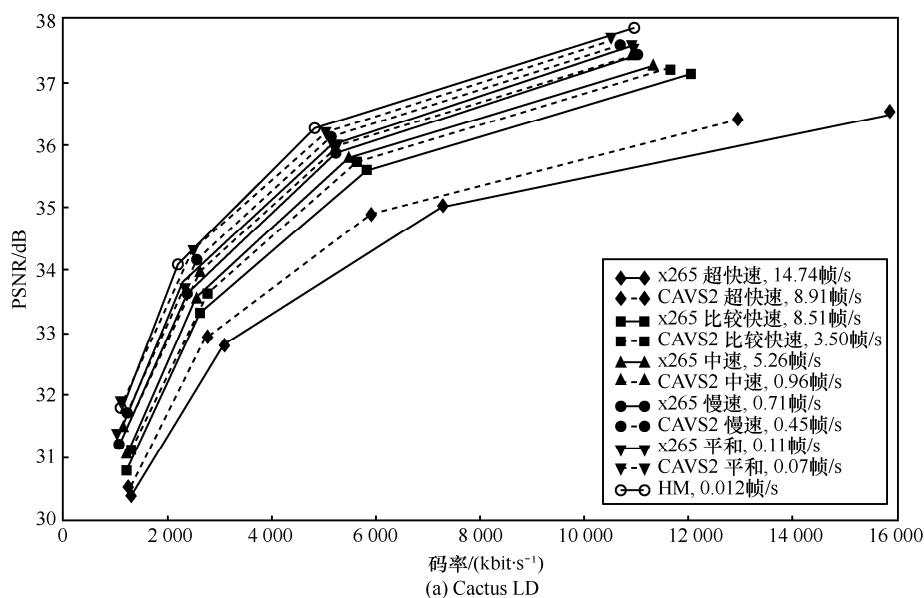


图14 AVS2快速编码器CAVS2与x265性能对比

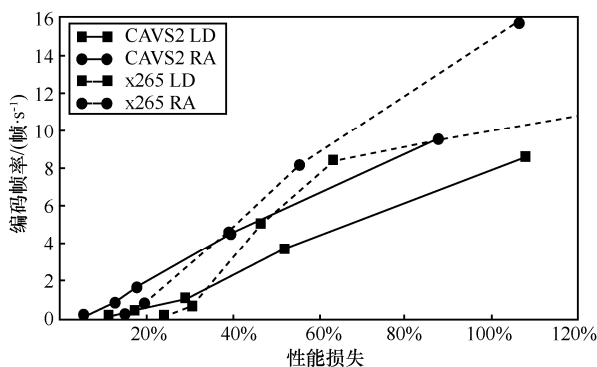


图15 AVS2快速编码器CAVS2与x265速度对比

字视音频产业的健康发展。为有效推进 AVS 标准的发展和运用, 2002 年、2004 年和 2005 年分别成立了负责组织研究制定技术标准的“AVS 工作组”、负责知识产权事务管理的“AVS 专利池管理委员会”和负责推动 AVS 产业应用的“AVS 产业联盟”, 也被形象地称为“AVS 三驾马车”。AVS 的制定和推广有赖于 3 个组织的协同工作。

AVS 标准已成为我国自主创新实施的一个典型案例, 推动了我国自主技术创新与民族产业



的发展。2006年第一代AVS标准颁布为国家标准, 2012年7月AVS+颁布为广电行业标准并实施, 进一步加快了自主创新AVS国家标准在广播电视领域的应用进程。目前中央电视台14套高清频道已全部采用AVS+标准播出, 各省卫视也已经全部切换到AVS+标准, AVS已成为支撑我国数字视听产业链“由大变强”的基础标准。2014年11月, 财政部批复《中央广播电视节目无线覆盖工程》, 投入近50亿对全国2562个发射台站进行数字化改造和全覆盖, 明确规定全国范围内地面电视节目编码方式采用AVS+编码。在海外, AVS在柬埔寨、老挝、斯里兰卡、古巴、吉尔吉斯斯坦等国的数字电视领域得到广泛应用。

2016年5月, AVS2广电行业标准的正式发布将更进一步推动我国数字视频产业的蓬勃发展, 特别是超高清数字电视广播在国内的应用。目前我国已有计划在2018年进行4K超高清试转播世界杯赛事节目, 在2019年正式开通4K频道。AVS标准为我国的广播电影电视行业创造了巨大的产业机会, 推动了我国数字视频产业的持续蓬勃发展。

5 结束语

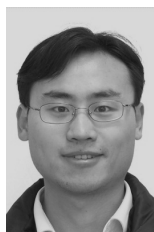
本文简单介绍了AVS视频编码标准发展历程以及最新一代AVS2标准的核心创新技术。和AVS1视频编码标准相比, AVS2编码效率提升一倍以上, 和同期HEVC/H.265国际标准相比, AVS2在场景视频编码方面有显著优势。AVS作为我国技术创新的典型已获得广泛应用, 形成了完善的AVS产业群体, 在支撑我国数字音视频产业实现“中国制造”到“中国创造”的战略转型中做出了历史性贡献。

参考文献:

- [1] 朱秀昌. 视频编码新标准——H.264[J]. 电信科学, 2002, 18(12): 26-29.
ZHU X C. A new video coding recommendation——H.264[J]. Telecommunications Science, 2002, 18(12): 26-29.
- [2] 解伟, 郭晓强, 全子一. IPTV 中的视频压缩技术研究[J]. 电信科学, 2006, 22(3): 39-42.
XIE W, GUO X Q, QUAN Z Y. Research on video compression in IPTV application[J]. Telecommunications Science, 2006, 22(3): 39-42.
- [3] 郭春辉, 周经野, 刘华东, 等. AVS1-P7 与 H.264 关键技术及性能比较[J]. 电信科学, 2006, 22(1): 47-50.
GUO C H, ZHOU J Y, LIU H D, et al. Key technologies and performance comparison of AVS1-P7 with H.264[J]. Telecommunications Science, 2006, 22(1): 47-50.
- [4] CUTLER C. Differential quantization of communication signals: US patent 2605631[S]. 1952.
- [5] OLIVER B. Efficient coding[J]. The Bell System Technical Journal, 1952(31): 724-750.
- [6] HARRISON C. Experiments with linear prediction in television[J]. The Bell System Technical Journal, 1952, 31(7): 764-783.
- [7] CANDY J, FRANKE M, HASKELL B, et al. Transmitting television as clusters of frame-to-frame differences[J]. The Bell System Technical Journal, 1971, 50(7): 1889-1917.
- [8] NETRAVALI A, ROBBINS J. Motion-compensated television coding: part I[J]. The Bell System Technical Journal, 1979, 58(3): 631-670.
- [9] AHMED N, NATARAJAN T, RAO K R. Discrete cosine transform[J]. IEEE Transactions on Computers, 1974, C(32): 90-93.
- [10] MALVAR H, HALLAPURO A, KARCZEWICZ M, et al. Low-complexity transform and quantization in H.264/AVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 598-603.
- [11] NARROSCHE M. Coding efficiency of the DCT and DST in hybrid video coding[J]. IEEE Journal of Selected Topics in Signal Processing, 2013, 7(6): 1062-1071.
- [12] TESCHER A G, COX R V. An adaptive transform coding algorithm[C]//1976 IEEE International Conference on Communications, June 14-16, 1976, Philadelphia, USA. New Jersey: IEEE Press, 1976.
- [13] KOGA T, IINUMA K, HIRANO A, et al. Motion-compensated inter-frame coding for video conferencing: NTCG5.3.1-G 5.3.5[R]. 1981.
- [14] 马思伟. AVS 视频编码标准技术回顾及最新进展[J]. 计算机研究与发展, 2015, 52(1): 27-37.
MA S W. History and recent developments of AVS video coding standards[J]. Journal of Computer Research and Development, 2015, 52(1): 27-37.
- [15] 余全合, 曹潇然, 李蔚然, 等. 短距离帧内预测技术: AVS_M3171[S]. 2013.
YU Q H, CAO X R, LI W R, et al. Short range intra prediction technique: AVS_M3171[S]. 2013.
- [16] FLIERL M, WIEGAND T, GIROD B. Rate-constrained multi-

- hypothesis prediction for motion compensated video compression[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(11): 957-969.
- [17] 凌勇, 虞露. F 帧 CE: 一种前向双假设预测图像块编码技术: AVS_M3326[S]. 2014.
- LING Y, YU L. F frame CE: a block image coding technique based on forward two hypothesis prediction: AVS_M3326[S]. 2014.
- [18] KIM I K, LEE S, PIAO Y J, et al. Directional multi-hypothesis prediction (DMH) for AVS2: AVS_M3094[S]. 2013.
- [19] 李蔚然, 袁媛, 曹潇然, 等. 非平方的二叉树变换: AVS_M3153[S]. 2013.
- LI W R, YUAN Y, CAO X R, et al. Non-square quad-tree transform: AVS_M3153[S]. 2013.
- [20] 林和源, 童怡新, 王颂文. 帧内编码二次转换改进提案: AVS_M3296[S]. 2014.
- LIN H Y, TONG Y X, WANG S W. Improvement proposal of the second time transform of intraframe coding: AVS_M3296[S]. 2014.
- [21] RISSANEN J. Generalized draft inequality and arithmetic coding[J]. IBM Journal of Research and Development, 1976(20): 198-203.
- [22] MARPE D, SCHWARZ H, WIEGAND T. Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 620-636.
- [23] WANG J, WANG X, JI T, et al. Two-level transform coefficient coding: AVS_M3035[S]. 2012.
- [24] LEE J S, KIM C, FU C, et al. Sample adaptive offset for AVS2: AVS_M3197[S]. 2013.
- [25] ZHANG X, SI J, WANG S, et al. Adaptive loop filter for AVS2: AVS_M3292[S]. 2014.
- [26] ZHANG X, HUANG T, TIAN Y, et al. Low-complexity and high-efficiency background modeling for surveillance video coding[C]//2012 Visual Communications and Image Processing (VCIP 2012), Nov 27-30, 2012, San Diego, USA. New Jersey: IEEE Press, 2012.
- [27] BOSSEN F. Common HM test conditions and software reference configurations: ITU-T SG16 Contribution JCTVC-L1100[S]. 2013.

[作者简介]



马思伟 (1979-), 男, 北京大学信息科学技术学院教授, 主要研究方向为视频处理与编码。

罗法蕾 (1991-), 男, 中国科学院计算技术研究所博士生, 主要研究方向为视频编码。

黄铁军 (1970-), 男, 北京大学信息科学技术学院教授、计算机科学技术系主任, 主要研究方向为视觉信息处理与神经形态计算。