

视频编码中一种新的用于几何块划分的变换结构

袁媛¹, 郑萧桢², 何芸¹

(1. 清华大学 清华信息科学与技术国家实验室, 北京 100084;
2. 华为技术有限公司 海思半导体和元件事业部研发部, 广东 深圳 518129)

摘要: 针对几何块划分技术提出了一种任意形状的变换结构. 虽然几何块划分技术比现有视频编码标准中采用的块划分方式更为灵活, 但其所对应的变换结构并未得到合理设计, 从而导致变换效率下降. 因此, 提出对两个任意形状的残差块分别进行任意形状的变换, 并通过翻转的方法保留像素间的相关性. 实验结果表明, 将任意形状的变化结构应用于几何块划分技术可以节约 1.16%~3.05% 比特率.

关键词: 几何块划分; 变换块; 任意形状的变换

中图分类号: TP 37

文献标志码: A

文章编号: 1007-2861(2013)03-0240-05

A New Transform Structure for Geometry Motion Partitioning in Video Coding

YUAN Yuan¹, ZHENG Xiao-zhen², HE Yun¹

(1. Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, Beijing 100084, China;
2. R&D Department of Hisilicon Semiconductor and Component Business Department, Huawei Technologies, Shenzhen 518129, Guangdong, China)

Abstract: This paper proposes a transform structure with arbitrary shape blocks for geometry block partitioning. A square coding block is split into two arbitrary shaped prediction blocks by geometry partitioning, which is more flexible than the partitioning method used in current video coding standards. In the current design, using a square transform block for combined square residual block is not reasonable and will affect transform efficiency. To further improve the transform structure, we propose that the two arbitrary shaped residual blocks are transformed separately with arbitrary shaped transform blocks. A turnover method is used to retain the correlation of residuals. Experimental results show that a bit rate saving of 1.16%~3.05% can be achieved.

Key words: geometry block partitioning; transform block; transform with arbitrary shaped block

随着视频应用的不断发展, 许多高效的视频编码标准陆续出台, 其压缩效率的提升主要依赖于预测、变换、熵编码、环路滤波等技术的不断改进, 其中帧间预测中的块划分结构的改进是提升压缩效率的关键因素, 也成为近年来视频编码领域的热点研究问题. 预测技术中的帧间预测是将已经编码的图像作为当前编码块的参考图像, 其目的是充分利用图像像素时域的相关性以消除像素间的冗余信息. 由于图像中存在各种复杂的运动形式, 因此需要将编码块依据运动的边界划分成更小的预测块分别进行预测, 以提高预测的准

确性, 这就是本研究所提出的块划分技术.

在早期的视频编码标准 H.261^[1], MPEG-1^[2] 和 MPEG-2^[3] 中, 预测块大小被限定为 16×16. H.263 标准^[4] 引入了大小为 8×8 的预测块. H.264 标准^[5] 中则使用了更加灵活的划分方式, 即采用了 7 种不同大小的预测块: 16×16, 16×8, 8×16, 8×8, 8×4, 4×8 和 4×4. 然而以上标准中使用的块划分结构都仅仅采用了垂直或者水平方向的直线作为预测块的划分曲线, 也就是说仅使用了方形或者矩形的预测块, 而实际视频序列中物体的边界是任意形状的, 不是仅局限于垂直或水

收稿日期: 2013-05-19

基金项目: 国家重点基础研究发展计划(973 计划)资助项目(2009CB320903); 国家科技重大专项资助项目(2010ZX03004-003)

通信作者: 何芸(1955—), 女, 教授, 研究方向为视频编码、图像处理等. E-mail: hey@mail.tsinghua.edu.cn

平方向的. 因此, 为了进一步提高预测精度, 已有研究者提出了更加灵活的几何块划分技术^[6-12], 即使用一条任意位置和任意方向的直线对编码块进行划分.

已有的几何块划分技术将一个方形的编码块划分成两个任意形状的预测块, 经过帧间预测之后便可得到两个任意形状的残差块, 再将两个任意形状的残差块拼接成一个方形的残差块, 然后使用一个方形的变换块对其进行变换. 然而, 预测块的边界处残差通常是不连续的, 甚至存在较大的跳变, 如果变换块内部存在较大的残差的跳变, 变换后就会产生较多的高频分量, 会严重影响变换效率. 为了解决这一问题, 本研究针对已有几何块划分技术提出了一种新的变换结构, 即对两个残差块分别进行形状自适应的变换, 以进一步提高几何块划分技术的编码效率.

1 几何块划分技术

几何块划分技术是现有视频编码标准中水平垂直划分方式的一种扩展, 它使用一条任意形状和任意方向的直线对编码块进行划分. 假设编码块中存在一条任意形状的物体边界如图 1(a) 所示, 那么图 1(b) 则给出了一种几何块划分技术所使用的划分直线, 该直线能够比较准确地拟合物体边界的位置和方向, 同时其标识方式也相对简单.

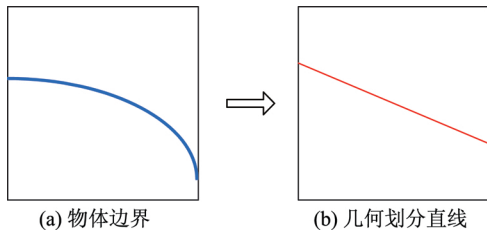


图 1 几何块划分的曲线拟合

Fig. 1 Curves fitting of geometry block partitioning

对于直线的标识方法有很多种, 包括使用直线的两个端点进行标识^[6]、使用极坐标 ρ 和 θ 进行标识^[8]、利用直线端点距顶点的距离进行标识^[10-12]等. 图 2 给出了 3 种不同的标识方法. 无论使用哪种标识方法, 在获取直线信息之后, 就可以根据直线判断每个像素点是属于哪个预测块的, 并生成一个二值图 $Mask(x, y)$. 得到二值图 $Mask(x, y)$ 的方法是将像素点坐标 (x, y) 带入直线方程 $f(x, y)$, 若 $f(x, y) > 0$, 则标识为 1, 若 $f(x, y) \leq 0$, 则标识为 0, 可以由下式表示:

$$Mask(x, y) = \begin{cases} 1, & \text{if } f(x, y) = ax + by + c > 0, \\ 0, & \text{if } f(x, y) = ax + by + c \leq 0. \end{cases} \quad (1)$$

得到每个像素点的标识位之后就可以对编码块进行精确划分. 使用两个任意形状的预测块分别进行运

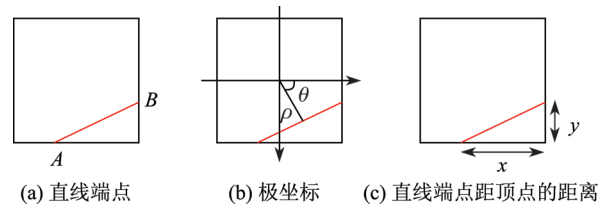


图 2 几何块划分的 3 种标识方法

Fig. 2 Three kinds of signaling methods of geometry block partitioning

动搜索得到其运动矢量, 经过预测之后可得到两个任意形状的残差块. 然而如何进行下一步的变换编码也是影响编码效率的关键. 利用已有的方形变换块, 可以使用如下两种方法对得到的残差块进行编码: ① 将两个残差块拼成一个方形的残差块再进行方形变换; ② 将两个任意形状的残差块分别补齐成两个方形的残差块再依次进行编码 (见图 3).

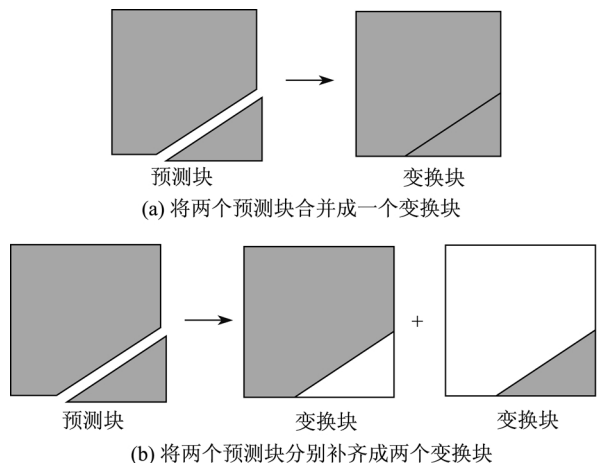


图 3 几何块划分的变换方法

Fig. 3 Transform methods for geometry block partitioning using square blocks

第一种方法的变换块内部存在预测边界, 而在大部分情况下预测边界两侧的残差存在像素值的跳变, 因此会造成变换后高频系数的增加, 影响变换效率. 第二种方法期望在一定程度上缓解残差的跳变, 但是选择何种像素值对这些空白部分进行填充是最关键的问题. 如果选用最简单的 0 或 1, 则在同样存在残差跳变的情况下, 根据现有的残差值进行计算, 会带来一定的计算复杂度. 同时, 对两个方形变换块进行编码的数据量比原来增加了一倍, 熵编码等部分需要处理的数据也会增多, 同样会影响最终的编码效率. 本研究期望在不引入更多数据的条件下解决残差跳变的问题, 因此提出了在基于内容的编码中使用形状自适应变换技术.

2 形状自适应变换技术

大多数视频编码标准采用基于块的编码思想,也就是将编码图像分割成若干个方形的编码块逐个进行编码. 而MPEG-4-2 标准^[13]则采用了基于内容的编码思想,将视频图像分割成若干个编码目标逐个进行编码,再使用混合编码框架中的技术进行预测、变换、熵编码等操作. 由于每个编码目标都是任意形状的,因此为了提高变换效率,本研究采用了一种形状自适应离散余弦变换 (shape adaptive discrete cosine transform, SADCT) 技术^[14-15]提高编码效率.

在基于内容的编码过程中,仍然以块为单位进行编码,只是需要编码的部分具有任意形状,因此只涉及编码块中的一部分区域. 假设得到一个任意形状的残差块 (见图 4(a)). 首先,将所有的残差上移至编码块的最上方 (见图 4(b)). 然后,逐列进行不同长度的 DCT 变换,得到列变换系数,列变换长度分别为 1, 3, 4, 5, 3 (见图 4(c)),再将列变换系数全部平移至块的最左边 (见图 4(d)). 最后,进行长度不同的行变换 (见图 4(e)),得到最终的变换系数. 此时,得到的变换系数已经全部集中在块的左上角,与方形变换是一致的,因此使用 zig-zag 顺序对变换系数进行扫描,输入熵编码模块即可. 对于基于内容的视频编码,由于已对编码目标的轮廓进行编码并传输至解码端,且编码目标的每行每列有多少个像素点是已知的,因此并不需要记录像素点的平移过程.

上述任意形状的 DCT 变换可有效完成对任意形状残差块的变换,并具有较高的变换效率,可应用于图像去噪、去块效应、去振铃效应等方面^[16-17],并且可进一步应用于几何块划分技术的变换结构中.

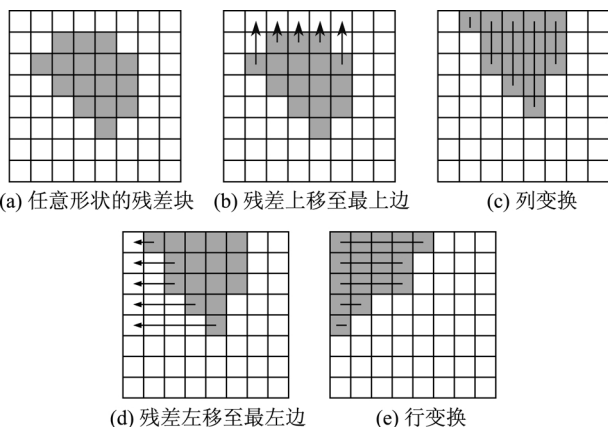


图 4 SADCT 技术
Fig. 4 Technology of SADCT

3 几何块划分中任意形状的变换技术

为了提高几何块划分中残差块的变换效率,本研

究将上述 SADCT 技术应用于几何块划分中. 将几何块划分中的两个预测块分别看作任意形状的残差块,并分别使用任意形状的 DCT. 对每个残差块重复进行以下编码过程: ①将残差块翻转至左上角; ②计算残差块每行像素点的个数,逐行进行行变换,得到行变换系数; ③计算行变换系数每列的像素点个数,逐列进行列变换,得到最终的变换系数; ④采用 zig-zag 顺序对变换系数进行扫描,输入之后的熵编码模块.

在基于内容的 SADCT 中,由于需要对像素点进行移位,因此在一定程度上减弱了相邻像素点之间的相关性,进而影响了变换效率. 而在几何块划分中,由于使用了一条直线将编码块从中间分割开,因此每个残差块都至少包含一个直角,则残差块可能是直角三角形或者直角四边形. 为了有效保留相邻像素间的相关性,本研究采用翻转残差块的方法来保留像素间的相对位置. 对于残差块的翻转可以按照对称轴分为 3 种情况: 水平翻转、垂直翻转和对角翻转 (见图 5).

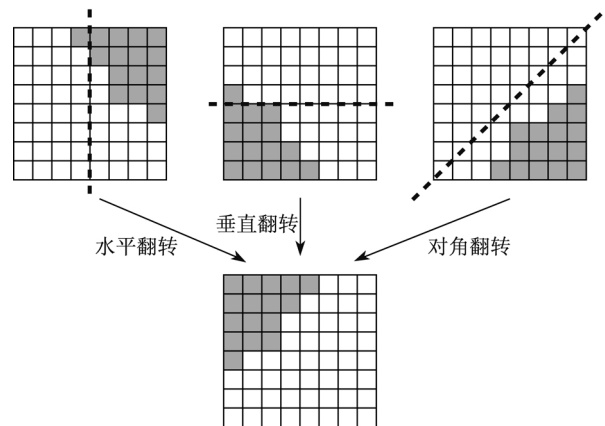


图 5 残差块的翻转

Fig. 5 Overturn of a residual block

经过翻转之后,残差全部集中在块的左上角,翻转过程并没有改变相邻像素的相对位置,从而保留了像素间的空域相关性,因此在此基础上进行变换可以获得更好的变换效率. 同时,对两个残差块分别进行任意形状的变换,能够有效解决变换块内部包含预测边界的问题,从而有效提升变换效率.

经过行变换和列变换得到变换系数后,使用如图 6 所示的 zig-zag 扫描顺序将二维的变换系数转变成一个一维向量输入熵编码模块进行编码. 虽然需要对两个变换块进行熵编码,但是由于变换系数的总数并没有增加,因此没有增加熵编码模块的编码复杂度.

4 实验结果

将所提出的几何块划分技术的任意形状变换结构应用于 kta2.7-qualcomm 平台^[18],即可得到本研究

的实验结果. 该平台是高通公司在最初制定下一代国际视频编码标准 H.265/HEVC 阶段提出的, 主要是在 H.264/AVC 的参考软件 JM16.2 的基础上增加了一些新的编码工具,

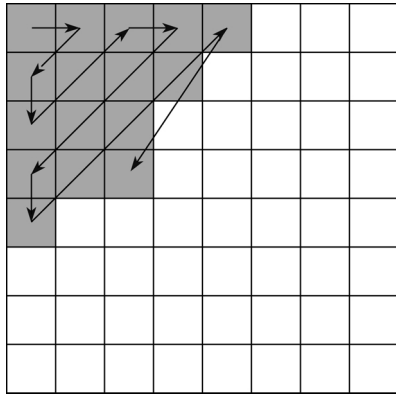
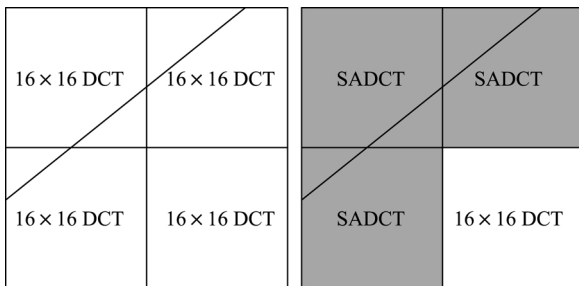


图 6 Zig-zag 扫描
Fig. 6 Zig-zag scan

对编码块的大小进行了扩展, 新增加了大小为 32×32 和 64×64 的编码块. 变换块仍然采用 16×16 以及 8×8 两种, 并且使用了 IPPP 的编码结构, 采用了 AMVP, CABAC 等高级编码技术.

该平台集成了本研究所需要的几何块划分技术, 将其应用在 16×16 , 32×32 和 64×64 的编码块中, 同时使用极坐标 ρ, θ 对划分直线的位置进行标识. θ 的变化步长设定为 11.25° , 也就是说 θ 共有 $360/11.25=32$ 种可能的取值. ρ 以一个像素为变化单位, 例如对于 16×16 的编码块, ρ 有 8 种可能的取值 (0, 1, ..., 7). 因此, 对于 16×16 的编码块进行几何块划分共有 $32 \times 8=256$ 种可能的划分方式; 而对于 32×32 和 64×64 的编码块, 则分别有 512 和 1 024 种可能的划分方式.

该平台对于几何块划分得到的残差块进行如下变换: 以 32×32 的编码块为例, 如果使用几何块划分, 则将得到的残差拼成大小为 32×32 的残差块, 然后使用 4 个 16×16 的变换块进行变换编码, 如图 7(a) 所示.



(a) 全部使用方形的变换 (b) 在包含预测边界的部分使用 SADCT, 其余部分仍然使用方形变换

图 7 32×32 编码块的几何块划分的两种变换结构

Fig. 7 Two transform structures for 32×32 coding block

本研究提出的变换结构如图 7(b) 所示. 由于实验平台中最大的变换块大小为 16×16 , 因此仍然将 32×32 的编码块划分成 4 个 16×16 的区域. 对于包含预测边界的区域使用 SADCT 技术, 其余部分仍然使用大小为 16×16 的变换, 这样可有效解决变换块内部存在预测边界的问题, 从而进一步提高变换效率.

实验结果如表 1 所示, 表中给出了使用任意形状变换结构之后比特率 (bit ratio, BR) 的变化和相应峰值信噪比 (peak signal to noise ratio, PSNR) 值的变化情况. 测试使用了 416×240 以及 832×480 两种分辨率的 8 个序列. 从实验结果可以看出, 使用任意形状的变换结构能够节约 1.16%~3.05% 比特率, 对应的 PSNR 值提升了 0.06~0.13 dB.

表 1 实验结果

Table 1 Experimental results

测试序列	Δ BR/%	Δ PSNR/dB
BasketballPass	-2.59	0.12
BlowingBubbles	-1.93	0.08
BQSquare	-3.15	0.13
RaceHorses	-1.32	0.07
BasketballDrill	-2.36	0.08
BQMall	-3.05	0.13
PartyScene	-2.03	0.10
RaceHorses	-1.16	0.06

具体分析每个序列的测试结果可以看出, 本研究提出的任意形状的变换结构对于运动比较剧烈的序列 (如 BasketballPass) 以及线条纹理比较多的序列 (如 BQSquare, BQMall) 体现出了较好的编码性能. 这是因为对于运动剧烈或者线条纹理多的序列, 使用几何块划分技术能够得到更好的预测结果, 因而应对更多的编码块选择使用几何块划分模式而非传统的方形或矩形预测模式. 因此, 本研究提出的新的变换结构也能够获得更高的编码效率.

5 结束语

本研究对几何块划分技术的变换结构进行了分析, 提出了使用 SADCT 对几何块划分中两个任意形状的残差块进行变换编码, 从而有效解决了变换块内部存在预测边界的问题, 并可有效提高变换效率. 实验结果表明, 所提出的 SADCT 能够节约 1.16%~3.05% 比特率, 从而有效提升了编码器的压缩率.

相比水平和垂直的划分技术, 几何块划分技术的优点是其划分方式非常灵活, 能够准确地将运动物体和背景分隔开, 因此可以得到较小的残差. 虽然本研究提出的 SADCT 能够在一定程度上提高编码效率,

但是编码复杂度则相应提高,由分析可知一个 64×64 编码块可能的划分方式有 1 024 种,编码过程中就需要遍历这 1 024 种划分方式,得到编码性能最好的一种,相应的编码时间将达到原来的 3~4 倍.因此,如何进一步降低几何块划分技术的实现复杂度将是下一步的研究重点.

参考文献:

- [1] ITU-T. Video codec for audiovisual services at $p \times 64$ Kbit/s [EB/OL]. [2013-05-19]. <http://www.itu.int/rec/T-REC-H.261-199012-S/en>.
- [2] ISO/IEC JTC1. Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s-part 2: video [EB/OL]. [2013-05-19]. <http://webstore.ansi.org/RecordDetail.aspx?sku=ISO/IEC%2011172-2/Cor3:2003#.UcDk9qV6qZE>.
- [3] ITU-T, ISO/IEC JTC1. Generic coding of moving pictures and associated audio information-part 2: video [EB/OL]. [2013-05-19]. <http://www.itu.int/rec/T-REC-H.262-199507-S/en>.
- [4] ITU-T. Video coding for low bit rate communication version 1 [EB/OL]. [2013-05-19]. <http://www.itu.int/rec/T-REC-H.263-199603-S/en>.
- [5] ITU-T, ISO/IEC JTC1. Advanced video coding for generic audiovisual services [EB/OL]. [2013-05-19]. <http://www.itu.int/rec/T-REC-H.264-200305-S/en>.
- [6] KONDO S, SASAI H. A motion compensation technique using sliced blocks in hybrid video coding [C]// IEEE International Conference on Image Processing. 2005: II-305-8.
- [7] DIVORRA Ò, YIN P, GOMILA C. Geometry-adaptive block partitioning [EB/OL]. [2013-05-19]. http://wftp3.itu.int/av-arch/video-site/0704_San/.
- [8] DIVORRA Ò, YIN P, DAI C, et al. Geometry-adaptive block partitioning for video coding [C]// IEEE International Conference on Acoustics, Speech and Signal Processing. 2007: I-657.
- [9] GUO L, YIN P, FRANCOIS E. TE 3: simplified geometry block partitioning [EB/OL]. [2013-05-19]. <http://phenix.int-evry.fr/jct/>.
- [10] FRANCOIS E, BORDES P, GUO L, et al. CE2: simplified geometry block partitioning [EB/OL]. [2013-05-19]. <http://phenix.int-evry.fr/jct/>.
- [11] ZHENG X Z, YU H P. Flexible macroblock partition for inter-frame coding [EB/OL]. [2013-05-19]. <http://phenix.int-evry.fr/jct/>.
- [12] ZHENG X Z, YU H P. CE2: Huawei and Hisilicon report on flexible motion partitioning [EB/OL]. [2013-05-19]. <http://phenix.int-evry.fr/jct/>.
- [13] ISO/IEC JTC1. Information technology-coding of audio/visual objects, part 2: visual [EB/OL]. [2013-05-19]. <http://webstore.ansi.org/RecordDetail.aspx?sku=ISO/IEC%2014496-2:2004#.UcDltKV6qZA>.
- [14] SIKORA T, MAKAI B. Shape-adaptive DCT for generic coding of video [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1995, 5(1):59-62.
- [15] XIE T, CHENG J W, HE Y, et al. A new approach to arbitrary shaped DCT [C]// IEEE International Conference on Picture Coding Symposium. 1997: 97-101.
- [16] FOI A, KATKOVNIK V, EGIAZARIAN K. Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images [J]. IEEE Transactions on Image Processing, 2007, 16(5):1395-1411.
- [17] FOI A, DABOV K, KATKOVNIK V, et al. Shape-adaptive DCT for denoising and image reconstruction [C]// International Society for Optics and Photonics in Electronic Imaging. 2006: 60640N.
- [18] KARCZEWICZ M, CHEN P S, JOSHI R, et al. Video coding technology proposal by Qualcomm Inc. [EB/OL]. [2013-05-19]. <http://phenix.int-evry.fr/jct/>.