改进的帧内帧间模式选择快速算法

朱蕾琦 张其善 杨东凯 胡薇薇

(北京航空航天大学 电子信息工程学院, 北京 100191)

摘 要: 为提高 I帧和 P帧的编码效率, H. 264视频编码标准采用多模式帧内预测和多子块模式帧间预测,并选择率失真最优的模式进行编码,导致编码复杂度很高.为此,提出一种基于小波变换的帧内模式快速选择结合帧间预测下帧内模式预判断的算法.利用小波变换后的低频子图做帧内模式选择,由此预判断宏块中每个4×4子块可能的预测模式,使原来9次的代价函数计算减少到平均 3. 75次;使用去均值的差值平方和,预判断宏块是否选择帧间预测下的帧内编码模式,减少 P帧编码时间.实验表明,改进算法在峰值信噪比降低不超过 0. 0644 dB的情况下,编码速度平均提高 19%左右.

关 键 词: H. 264; 帧内预测; 帧间预测; 预测模式; 小波变换; 平均绝对差

中图分类号: TN 919. 8

文献标识码: A 文 章 编 号: 1001-5965(2008) 12-1411-04

Fast mode selection for intra and inter prediction

Zhu Leiqi Zhang Qishan Yang Dongkai Hu Weiwei

(School of Electronics and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

Abstract: H. 264 video coding standard is high complex and time-consuming owing to its multimode in intra prediction and multi-size sub-block modes in inter prediction to improve the efficiency of I frame and P frame. A new algorithm with fast mode selection based on wavelet transforms in intra prediction and combined with pre-estimation of intra mode in inter prediction was proposed. Firstly, the method utilized the optimal modes selected in low frequent part transformed by wavelet to estimate the possible modes in sub-block of original macro-block. This method reduces the calculate times from 9 to 3. 75 averagely. Secondly, pre-estimation of intra mode in inter prediction was implemented using the parameter of mean removed mean absolute difference, and it reduced the P frame coding time. Simulation experiment results from video sequences show the method effective with the encoding time of sequences lower about 19% averagely while peak signal-to-noise ratio lower 0. 0664 dB at most.

Key words: H. 264; intra prediction; inter prediction; prediction mode; wavelet transforms; mean absolute difference

作为迄今为止最新的国际视频编码标准, H. 264采用了多种新技术,与以前的 H. 263+及 MPEG4相比,在保证相同编码质量的前提下, H. 264的输出码率降低了 50%[1-2].然而,新技术的引用在提高编码效率的同时也增加了计算复杂度.如帧内预测多模式选择,需要计算每个宏块在不同模式下亮度和色度分量的代价函数,从而选择率失真最优的编码模式.同时,在帧间预测模式选择时,为了提高编码效率引入了帧间预测的帧内编码模式,而帧内预测模式选择的计算量是帧间预测的 5倍[3],由此大大增加了编码复杂度.因此,如何提高帧内、帧间预测编码速度成为 H. 264研究的一个方向.文献[4]提出一种在帧内预测模式下重新安排宏块中 16个 4×4子块编码顺序的改进算法;文献[5]利用平均边缘误差和对帧间预测下帧内模式进行预判断,但会增加额外的

收稿日期: 2008-01-07

基金项目: 国家自然科学基金资助项目(60602046)

作者简介: 朱蕾琦(1982-),女,湖北武汉人,博士生, zhu leiqi@gmail.com.

计算; 文献[6]利用改进的率失真优化函数针对不同的帧内预测模式提出快速算法, 但是对于 PC 机上的编码时间并没有实际的提高; 文献[7]分别从亮度块和色度块的模式选择计算方法上提出改进, 但算法没有综合考虑亮度块和色度块; 同时文献[8]也提出一种分两步选择的快速算法.

基于帧内帧间模式选择的复杂性, 本文从两个方面提出一种新的改进算法. 首先对 I帧进行二级小波变换, 然后对其低频子图中每个 4×4块做帧内模式选择, 根据得到的最优模式预判断原宏块可能的预测模式, 以减少模式选择的时间. 其次对于 P帧宏块是否采用帧内编码模式, 本文将一种预判断的算法引入到改进算法中, 以提高编码速度. 经试验证明, 本文所采用的方法实现简单, 并且可行、有效.

1 H. 264的帧内帧间预测技术

在 H. 264帧内预测中, 亮度块分 4×4子块预测(intra4×4)和 16×16宏块预测(intra16×16)两种[9]. 其中 intra4×4有 9种预测模式, 包括 1种直流模式 DC(Direct Current)即 mode 2和 8种方向预测模式, 即 mode 0, mode 1, mode 3~mode 8, 如图 1所示. intra16×16有 4种预测模式, 分别是垂直、水平、直流和平面. 帧间预测采用了多子块模式选择的方式以提高编码效率. 另外, 由于场景的快速变换, 利用宏块本身的特性进行帧内编码相比帧间预测匹配能得到更好的编码质量, 因此在 H. 264的帧间预测中引入了帧内编码模式[10]. 由于人眼对亮度信息比较敏感, 同时色度分量模式选择与亮度分量的模式有很强的相关性, 所以本文中帧内预测模式选择只针对亮度分量, 色度分量本文不做详细阐述.

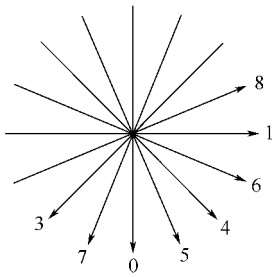


图 1 intra4×4亮度预测模式方向

2 基于小波变换的帧内模式选择

帧内预测模式选择时, 为了均衡编码质量和编码所用比特数, 每一个宏块都需要计算多种模式下的代价函数, 以得到最优的率失真 RDO (Rate-Distortion Optimization)模式. 帧内预测 RDO 的计算如下式:

$$C_{\text{mode}} = D_{\text{SSD}}(S_{\text{cur}}, S_{\text{rec}}, M) + \lambda_{\text{mode}} \times R(S_{\text{cur}}, S_{\text{rec}}, M) \quad (1)$$

其中, C_{mode}为该模式下的代价函数; S_{cur}和 S_{rec}分别为原始信号和重构信号; M 为编码所采用模式; λ_{mode}为拉格朗日乘数. 另外, D_{SSD}为原始信号和重构信号之间的差值平方和; R 为编码宏块头信息、编码模式以及所有 DCT(Discrete Cosine Transform)变换系数所需要的比特数. 代价函数最小的模式为帧内预测模式.

2.1 帧内模式选择快速算法思想

基于小波变换的帧内预测基本思想是: 图像经过小波变换后会分解为 1个低频子图和 3个高频子图, 其中低频子图包含了大量原始图像的特征信息, 而高频子图则包含细节信息. 若能首先对低频子图进行帧内的模式预测, 然后根据其选择的模式, 就能预判断出原始图像中宏块可能的预测模式, 从而减少模式选择的时间.

2.2 视频图像的小波变换

视频图像小波变换是通过一组低通滤波器 H 和高通滤波器 G 对图像进行二维离散小波分解. 根据多分辨分析, 任一尺度空间的剩余系数和小波系数都可以由上一尺度空间的剩余系数与共轭滤波器组 (H^*, G^*)进行卷积并 2倍下采样得到, 即可以在任意尺度空间上对图像的行和列进行离散小波变换, 最终实现图像的小波分解. 每一级的小波变换将原始图像按不同频带和分辨率分解成 4个子带图像: 低频子图 W_{LL}, 高频垂直方向子图 W_{LH}, 高频水平方向子图 W_{HL}和高频对角线方向子图 W_{HH}. 图像小波分解的框图如图 2所示.

其中, $W_{\text{LL},\,i+1}$是需要进行小波变换的原始图像或变换后的低频分量, $W_{\text{LL},\,i}$是对 $W_{\text{LL},\,i+1}$小波变换后的低频分量, $W_{\text{LH},\,i}$, $W_{\text{HL},\,i}$, $W_{\text{HH},\,i}$是对 $W_{\text{LL},\,i+1}$小波变换后的 3个高频分量.

选择合适的小波函数对图像进行分解, 需要满足两个条件: 其一, 低频子图应尽可能多地保留原图特征信息; 其二, 小波分解的系数应尽可能地小, 以减少计算量. 文献[11]通过一系列性能比较得出, 利用 biorthogonal7/9能够满足以上两个条件, 所以本文采用该小波函数对视频序列进行分解.

2.3 帧内模式选择快速算法步骤

原始宏块中每个 4×4子块可能的预测模式, 由对应低频子图选择的预测模式、两个临近方向的预测模式及 DC模式组成参考模式. 但是对于

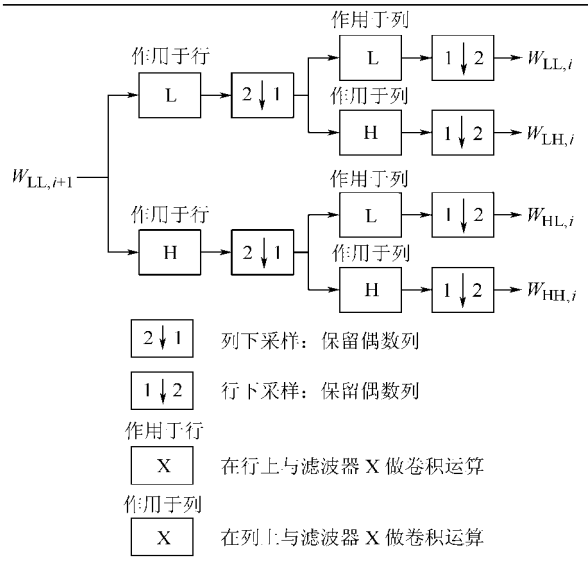


图 2　二维小波变换分解算法结构图

左边、上面和右上不全有相邻块的宏块,若其对应的低频子图选择的预测模式为 DC模式,原宏块中 16个 4×4子块不一定全为 DC 模式,此时就需要参考相邻块的预测模式.具体的算法步骤如下.

1) 对 I帧做 2级小波变换,取出原始图像的低频子图作为预测对象,其中低频子图中每个 4×4块记为 lbw4×4,对应原图像中的一个宏块.

2) 对低频子图的每个 lbw4×4块做 intra4×4的模式选择,保留率失真代价函数最小的模式.

3) 若当前 lbw4×4块左边、上面和右上不全有相邻块,且参考模式为 mode2,则转入步骤 4),反之若 3个方向全有邻块,则对原宏块中 16个 4×4子块做符合如下规则的 intra4×4模式选择:

①若 2)中选择的是 mode2,则对宏块做 intra16×16的 DC 变换;②若 2)中选择的是 mode0, mode1, mode4, mode5, mode6或 mode7,则取该模式及其临近两个方向的模式,以及 mode2作为参考模式.③若 2)中选择的是 mode3或 mode8,则取该模式及其临近 mode7或 mode1,以及 mode2作为参考模式.

4)对于低频子图中左边、上面和右上不全有相邻块的 lbw4×4块,且步骤 2)中选择的模式为 mode2,按如下规则对原图宏块 16个 4×4子块进行 intra4×4模式选择 (图 3):①若该 lbw4×4块上面和右上没有相邻块,则取其左边邻块的预测模式为参考模式;②若该 lbw4×4块左边没有相邻块,则分别取其上面和右上邻块的预测模式为参考模式;③若该 lbw4×4块右上没有相邻块,则分别取其左边和上面邻块的预测模式为参考模式;④若该 lbw4×4邻块的预测模式也均为 mode2,则上述①②③种情况中均计算 9种预测模式,以取得 RDO最优的预测模式.

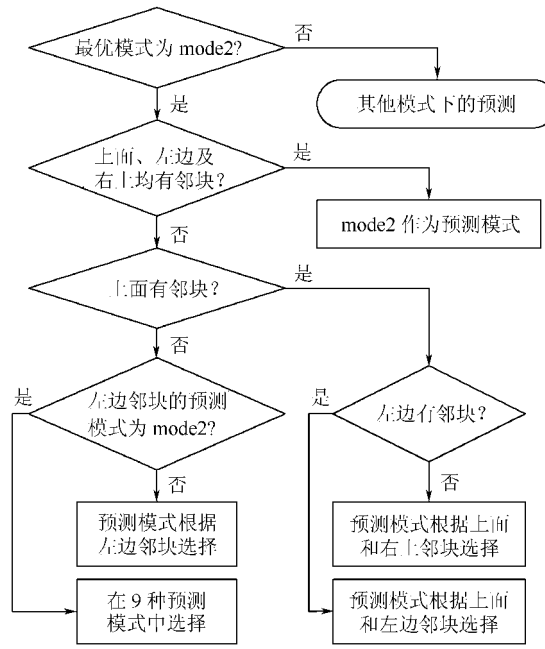


图 3　lbw4×4预测模式为 mode2时,
原图宏块的模式选择算法框图

2.4　帧内模式选择快速算法分析

由改进算法的实现过程可知,对于左边、上面和右上都有邻块的 lbw4×4块,其对应宏块中每个 4×4子块所需要计算的 RDO次数平均为 (6×4+2×3)/8=3.75,大大少于原始算法中 9次.即使在最坏的情况下 (对于低频子图中左边、上面和右上不全有相邻块的 lbw4×4块,在步骤 2)中选择的模式为 mode2,且其相邻块的模式也为 mode2),一帧图像中需要计算 intra4×4的 RDO次数为 56×16×9+340×16×4+396×9=33388,而相对于原始算法,一帧图像 intra4×4的 RDO计算次数为 396×16×9=57024,同时还需要计算 396×4=1584次 intra16×16的 RDO.由上述比较可看出,改进的算法在最坏的情况下也能减少约一半的计算量,显著的提高了编码效率.

3　帧间预测快速判断算法

为了提前判断帧间预测是否采用帧内编码模式,本文采用的是类似于文献 [1]中提到的算法,该算法是利用一种去除均值的平均绝对差 (mean removedMAD)作为参数,提前判断在 P帧的某个宏块中是否使用帧内预测模式.

MAD(Mean Absolute Difference)是常用的衡量宏块间相似性的一个参数,大部分的运动估计算法都使用该参数作为宏块匹配准则.然而 MAD

有时却并不能把编码质量与编码码流大小统一起来. 由此文献[1]提出一种去除均值的 MAD, 即 mMAD. 该参数在计算 MAD 之前, 先去除两个宏块的均值, 如式(2). 由此得到的 mMAD 能更清晰地反映出两个宏块之间的相似性.

$$D_{\mathrm{mMAD}} = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|(x_{ij}-m_x)-(y_{ij}-m_y)\right| \tag{2}$$

其中 $N=16$ 为宏块的大小; m_x 和 m_y 分别为两个宏块亮度的均值; 而 x_{ij} 和 y_{ij} 分别是两个宏块中每个像素的亮度值.

由于 mMAD 可以更好地提供衡量编码质量和编码码流大小的信息, 所以在 P 帧模式选择时采用下式提前判断是否使用帧内预测模式.

$$\left.\begin{array}{ll}\text{if} & D_{\mathrm{mMAD}}-E_x>T \quad \text{intra coding}\\ \text{else} & \text{inter coding}\end{array}\right\} \tag{3}$$

其中 $E_x=\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|x_{ij}-m_x\right|$; T 是由统计得到的阈值. 若 $D_{\mathrm{mMAD}}-E_x$ 大于给定阈值, 则该宏块做帧内预测, 反之则做帧间的多子块预测.

对于不同特征的序列, 阈值的选取也可能不同. 如运动缓慢的序列, 大部分的宏块都可以由帧间预测模式得到较好的预测效果, 所以此时就可以设定一个较小的阈值, 以得到更好的编码质量和较小的编码码流.

4 实验结果及分析

为了验证本算法的有效性, 实验对 3 个 CIF (Common Intermediate Format)格式典型序列 foreman、bus和 news分别测试了改进的帧内预测模式选择算法和整个改进算法, 并在编码时间和亮度分量 Y 以及两个色度分量 U、V 的 PSNR (Peak Signal-to-Noise Ratio)值方面与原算法进行了比较. 实验参数设置如下: ①允许 RDO; ②I帧的 QP (Quantization Parameter)值设为 30; ③熵编码采用 CAVLC (Context Adaptive Variable Length Coding); ④P帧的参考帧数为 1; ⑤GOP (Group of Pictures)采用 IPPP; ⑥允许采用哈德曼变换. 实验程序所运行的环境: CPU 为 Intel Pentium Processor 1.60 GHz, 内存为 768 MB, 操作系统为 Windows XP Professional, 采用的测试序列编码的帧数为 120帧. 表 1给出了编码前后的实验结果及两种方案对比的数据. 其中 OA 表示原算法, NA1表示只采用了基于小波变换的帧内预测模式选择快速算法, NA2表示在 NA1的基础上加入了帧间预测帧内编码模式预判断的改进算法, 上述改进算法与原始算法的比较结果分别为 R1和 R2

表 1 优化前后数据对比

测试序列		编码时间/s		PSNR/dB		
		所有帧	I帧	Y	U	V
foreman	OA	16.2884	6.3416	36.0136	40.4920	43.4588
	NA1	14.2700	4.2964	35.9983	40.4793	43.4161
	NA2	13.2282	4.2964	35.9931	40.4897	43.4208
	R1	12.39%	32.25%	-0.0153	-0.0127	-0.0427
	R2	18.79%	32.25%	-0.0205	-0.0023	-0.0380
bus	OA	20.4800	6.2260	33.5118	39.2938	40.2953
	NA1	18.1245	4.7273	33.4516	39.2491	40.2771
	NA2	16.6245	4.7273	33.4474	39.2548	40.2811
	R1	10.91%	24.07%	-0.0602	-0.0447	-0.0182
	R2	18.83%	24.07%	-0.0644	-0.0390	-0.0142
news	OA	11.4481	6.3952	37.0199	39.5908	40.5120
	NA1	9.6464	4.6236	36.9566	39.5853	40.4590
	NA2	9.2600	4.6236	36.9558	39.5863	40.4655
	R1	15.74%	27.70%	-0.0633	-0.0055	-0.0530
	R2	19.11%	27.70%	-0.0641	-0.0045	-0.0465

从实验结果可以看出, 本文提出的改进算法使 I帧的编码时间减少了 24.07% ~ 32.25%, 整个序列的编码速度提高了 18.79% ~ 19.11%. 由于 news序列运动平缓, 且其纹理简单, 所以该序列的时间相关性很强, 帧间预测的模式选择相对耗时较少, 此时帧内预测模式选择的时间就会对整个编码时间有重要的影响. 从表 1可以看出, 只采用基于小波变换的帧内预测模式选择快速算法就可以使 news序列的编码时间提高 15.74%. 另外, 从表 1还可以看出, 测试序列的编码质量几乎保持了不变. 改进算法 1中 Y 分量的 PSNR 仅下降了 0.0153 dB, 而改进算法 2中 Y 分量的 PSNR 最多也只下降了 0.0644 dB, PSNR 微小的改变完全不会影响人眼主观的视觉评价, 因此视频的编码质量几乎没有变化.

5 结束语

基于帧内帧间模式选择的复杂性, 本文分两方面提出了一种新的改进算法. 针对帧内模式选择, 首先对 I帧小波变换后的低频子图做帧内预测的模式选择, 然后根据其模式确定原宏块可能的预测模式, 从而使原来 9次的 RDO 计算减少到平均 3.75次, 复杂度明显降低, 提高了编码性能. 对于帧间预测的模式选择, 通过计算宏块的去除均值 MAD, 提前判断出该宏块是否需要进行帧内编码模式, 进一步提高了编码速度. 通过实验可以看出, 本算法在保证码流结构和编码质量不变的情况下, 较大的提高编码速度, 不失为一种行之有效的改进算法.

(下转第 1427页)

参考文献 (References)

[1] Rossoni P, Panetta P V. Developments in nanosatellite structural subsystem design at NASA-GSFC[R]. SSC-99-V-3, 1999

[2] 朱毅麟. 微型航天器的公用平台和关键技术[J]. 中国空间科学技, 1997, 33(6): 25-31
Zhu Yilin. Utility platform and critical technologies for miniature spacecraft[J]. Chinese Space Science and Technology, 1997, 33(6): 25-31(in Chinese)

[3] 邓建国, 罗隆福. 基于 BP神经网络的大型水轮发电机定子铁心的温度预测[J]. 湖南大学学报, 2002(1): 72-75
Deng Jianguo, Luo Longfu. BP neural network method for temperature rise prediction of large hydrogenerator[J]. Journal of Hunan University, 2002(1): 72-75(in Chinese)

[4] 闵桂荣, 郭舜. 航天器热控制[M]. 北京: 科学出版社, 1998: 420-428
Min Guirong, Guo Shun. Aerocraft thermal control[M]. Beijing: Science Press, 1998: 420-428(in Chinese)

[5] 李运泽, 宁献文, 王晓明, 等. 纳卫星热系统的在轨动态特性分析与建模[J]. 空间科学学报, 2007, 27(3): 233-238
Li Yunze, Ning Xianwen, Wang Xiaoming, et al. Analysis and modeling for the in-orbit dynamics of nanosatellite's thermal system[J]. Chin J Space Science, 2007, 27(3): 233-238(in Chinese)

[6] Satish Kumar. 神经网络[M]. 北京: 清华大学出版社, 2006: 120-125
Satish Kumar. Neural network[M]. Beijing: Tshinghua University Press, 2006: 120-125(in Chinese)

[7] 许东, 吴铮. 基于 MATLAB 6.X的系统分析与设计——神经网络[M]. 陕西: 西安电子科技大学出版社, 2003: 159-163
Xu Dong, Wu Zheng. System analysis and design based on MATLAB 6.X— neural network[M]. Shanxi: XiDian University Press, 2003: 159-163(in Chinese)

[8] 肖业伦. 航空航天器运动的建模——飞行动力学的理论基础[M]. 北京: 北京航空航天大学出版社, 2003: 233-240
Xiao Yelun. Aerocraft movement modeling— theoretical basis of flight dynamics[M]. Beijing: Beijing University of Aeronautics and Astronautics Press, 2003: 233-240(in Chinese)

(责任编辑: 文丽芳)

(上接第 1414页)

参考文献 (References)

[1] Turaga D S, Chen T. Estimation and mode decision for spatially correlated motion sequences[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(10): 1098-1107

[2] Dong H Y, Zhang Q S. Fast inter coding mode selection algorithm for H.264[J]. Chinese Journal of Aeronautics, 2006, 19(3): 119-224

[3] JVT-J033 Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), Fast mode decision for H.264[S]

[4] Lee W, Jung Y, Lee S, et al. High speed intra prediction scheme for H.264/AVC[J]. IEEE Transactions on Consumer Electronics, 2007, 53(4): 1577-1582

[5] Choi I, Lee J, Jeon B. Fast coding mode selection with rate-distortion optimization for MPEG-4 part-10 AVC/H.264[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(12): 1557-1561

[6] Tseng C H, Wang H M, Yang J F. Enhanced intra-4x4 mode decision for H.264/AVC coders[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(8): 1027-1032

[7] 孟庆磊, 姚春莲, 宋建斌, 等. 一种面向 H.264/AVC的快速帧内预测选择算法[J]. 北京航空航天大学学报, 2007, 33(2): 119-223
Meng Qinglei, Yao Chunlian, Song Jianbin, et al. Fast selective algorithm of intra prediction for H.264/AVC[J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(2): 119-223(in Chinese)

[8] Kim C S, Li Q, Kuo C C J. Fast intra-prediction mode selection for H.264 codec[J]. Proceedings of SPIE-The International Society for Optical Engineering, 2003, 5241: 99-110

[9] Richardson Iain E G. H.264 and MPEG-4 video compression: video coding for next generation multimedia[M]. England: John Wiley & Sons Ltd, 2003: 177-184

[10] ITU-T I ISO/IEC, ITU-T Recommendation H.264: Advanced video coding for generic audiovisual services[S]

[11] Zan J W, Ahmad M O, Swamy M N S. Comparison of wavelets for multiresolution motion estimation[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(3): 439-446

(责任编辑: 娄　嘉)