

一种基于 H.264 的快速帧间预测模式选择算法^①

曾庚卓^{1,2} 马跃^{1,2} 张伟^{1,3} (1.中国科学院沈阳计算技术研究所 辽宁 沈阳 110171;
2.中国科学院研究生院 北京 100049; 3.中国科学技术大学 安徽 合肥 230027)

摘要: 提出了一种基于 H.264 的快速帧间预测模式选择算法, 该算法针对 H.264 中帧间预测模式采用的对所有候选模式进行全搜索的方法, 提出了一种基于图像运动剧烈程度分级的模式选择算法, 达到快速选择预测模式的效果。实验证明, 本文提出的模式选择算法与 H.264 原算法相比在基本保持原算法率失真特性良好的同时较大幅度的提高了编码效率。

关键词: H.264; 运动估计; 帧间预测; 模式选择

A Fast Inter Prediction Mode Selection Algorithm of H.264

ZENG Geng-Zhuo^{1,2}, MA Yue^{1,2}, ZHANG Wei^{1,3}

(1. Shenyang Institute of Computing Technology, Chinese Academy of Sciences, Shenyang 110171, China;
2. Graduate University of Chinese Academy of Sciences, Beijing 100049, China; 3. Department of Computer Science and Technology, University of Science and Technology of China, Hefei 230027, China)

Abstract: This paper presents a fast, inter-frame prediction mode selection algorithm based on H.264. For H.264, the algorithm method of inter prediction mode searches for all the candidate models. This paper proposes an algorithm based on an image intensity classification model selection system to quickly select the best prediction model. Experiments show that the algorithm proposed in this paper increases coding efficiency under the premise of maintaining the basic rate-distortion characteristics.

Keywords: H.264; motion estimation; inter prediction; mode selection

1 引言

最新的视频编解码标准 H.264/AVC 由 ISO/IEC 的 MPEG 组织和 ITU-T 的 VCEG 组织联合制定的, 它同时也是 MPEG4 第 10 部分的主要内容。与其他的视频编解码标准相比, H.264 标准的目的在于更加有效地提高视频编码效率和它对网络的适配性, 其编码算法可广泛应用于数字电视、无线视频通信和视频会议中^[1]。在相同的条件下, 与 H.263+ 和 MPEG-4 相比, H.264/AVC 能减少 50% 的码流, 但是这种编码效率的提升却是以增加计算复杂度和存储能力为代价获得

的^[2]。这种高计算复杂度将是对 H.264 在嵌入式领域和实时应用场合应用的一个瓶颈。

帧间预测是视频压缩的关键技术之一, 通过利用已经重建的参考帧对当前参考帧进行预测编码, 可以对当前块与参考块的运动矢量以及相应的运动补偿残差进行编码, 有效地消除视频编码中的时间冗余, 降低编码码率, 提高编码效率^[3]。在 H.264 编码标准中, 帧间预测编码是该标准的重要编码方式, 而帧间预测编码的模式选择算法是 H.264 编码计算复杂度极高的主要原因之一。

^① 收稿时间:2010-01-29;收到修改稿时间:2010-03-10

2 H.264帧间编码

帧间预测编码是从已被编码的重建的一帧或多帧中产生预测模板。这个模板是通过在参考帧中进行采样点的平移得到的。在H.264中仍然采用了从H.261起就被广泛运用的基于块的运动补偿。和早期的标准不同的是H.264支持更多更细的块大小划分以及亚像素运动矢量估计^[4]。

H.264的帧间预测使用了7种基于不同块大小和块形状的模式,包括16×16, 16×8, 8×16, 8×8, 8×4, 4×8和4×4(如图1所示),其中模式1到模式4一般称作宏块级的模式,而模式4所包含的模式(模式4到模式7)则称作亚宏块级的模式。多模式的运用极大地提高了编码效率和R-D性能。

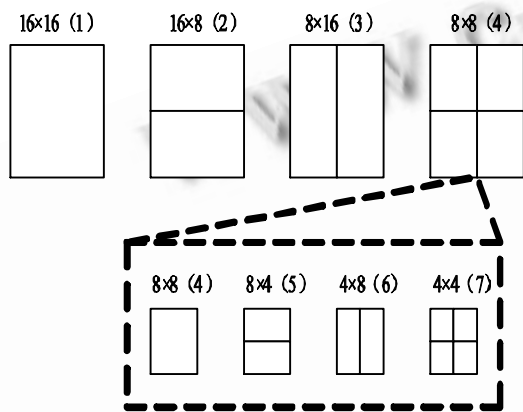


图1 帧间预测模式(共7种)

在图1所示的7种模式中,主要依据图像的细节变化程度采用不同的预测模式。如在视频会议中,视频图像背景区域变化很小,可采用宏块级的大模式,而在足球、网球比赛等图像背景丰富且图像变化较剧烈的视频中,则一般采用亚宏块级的小模式作为帧间预测的最佳模式。

H.264在进行运动估计时,首先对当前块采用所有的模式进行运动估计,然后用率失真优化(RDO)方法比较采用不同模式编码时所需码率的大小,码率小的模式将作为最终的模式。该过程是一个全搜索的过程,它盲目地搜索了所有候选模式,从而造成了极大的计算复杂度。虽然通过这种方法能够得到最优的效果,但是在很多硬件设备中不能支持这种庞大的计算能力要求。据统计,多模式的块匹配运动估计占据了整个编码约80%的计算量,巨大的运算量使得H.264的编码复杂度约是H.263的4~5倍。因此,帧间模

式选择快速算法的研究是目前H.264研究中的重要课题。

目前,已有一些针对H.264模式快速选择的研究。如某算法分析了视频序列中的目标运动和纹理区域特点,提出了快速多块运动估计的算法,该算法先进行8×8模式搜索,并通过4个8×8向量的一致性特点来选出16×16的模式。该算法通过快速选出一组序列中出现最多的16×16模式,节省了大量运算量,而码率和编码质量几乎不变。还有一些算法测试相邻宏块运动向量的相关性,进行大模式和小模式的分割,同时结合快速的运动搜索算法,比传统的全搜索方法速度上有较大的提高^[5]。

本文针对H.264帧间预测模式中采用的穷尽模式选择算法,提出了一种基于图像运动剧烈程度分级的模式选择算法,该算法通过对各预测模式采用阈值分级来区分图像的运动剧烈程度,从而采用不同的搜索模式以加快帧间预测模式的选择。实验证明,本文提出的快速模式选择算法与H.264原算法相比在基本保持原算法率失真特性良好的前提下较大幅度的提高了编码效率。

3 快速帧间模式选择算法

在通常的视频序列中,相邻图像之间普遍存在着时间相关性,即相邻图像的大部分区域的变化是缓慢的。在H.264/AVC中,因为图像的纹理不同,故采用不同的分块模式。纹理特征可以用宏块间的残差表示。对运动补偿后,残差小的采用大的分块模式,残差大的采用小的分块模式,纹理特征可以从帧间宏块间的SAD获得^[6]。

$$SAD(i, j) = \sum_{m=0}^M \sum_{n=0}^N |f_k(m, n) - f_{k-1}(m+i, n+j)| \quad (1)$$

公式(1)表示大小为M×N的宏块当前帧和前一帧的残差值。

在H.264标准中,通常使用率失真优化方法从所有候选模式中选择最佳的模式,该方法公式如下:

$$J_{mode} = SAD + I_{mode} \times R \quad (2)$$

其中SAD表示当前帧与重构帧的差值;mode表示候选模式中的最佳模式;I_{mode}指拉格朗日系数。R表示编码宏块头、运动矢量、DCT系数的比特率。该公式不仅可以判断宏块级的模式,同时也适合亚宏块级模式的判断。

3.1 改进的阈值判断准则

众所周知，对于图像中平滑的区域，其能量主要集中在低频分量上，所以高频分量可以用来衡量图像的细节程度^[7]，即该图像的运动剧烈程度。对于一个 $M \times N$ 块，AC 系数总能量可以用式(3)表示：

$$E_{AC} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [(F_{(i,j)})^2 - (F_{(0,0)})^2] \quad (3)$$

其中， $F_{(i,j)}$ 表示大小为 $M \times N$ 的宏块各像素的 DCT 系数， $F_{(0,0)}$ 表示该宏块的 DC 系数。

当一个 $M \times N$ 块像素值 $Y_{(i,j)}$ ，其上下左右的像素值与其相差为 255(灰度值用 8 位表示，最大值为 255，最小为 0)，表明此时块的细节最丰富。根据能量守恒定律，变换后的总能量等于变换前的总能量，所以对于大小为 $M \times N$ 的块，其 AC 系数的最大能量为：

$$E_{ACmax} = \frac{M \times N}{2} [(F_{max})^2 - (F_{min})^2] - (M \times N) \left(\frac{F_{max} + F_{min}}{2} \right)^2 \quad (4)$$

其中 $F_{max} = 255$ ， $F_{min} = 0$ 。

把 E_{ACmax} 作为一个标准化因子，对 E_{AC} 和 E_{ACmax} 取自然对数，得到：

$$T = \frac{\ln E_{AC}}{\ln E_{ACmax}} \quad (5)$$

这样， T 被限制在 $[0,1]$ 区间，可以将 T 作为新的阈值，根据 T 值的大小作为图像运动剧烈程度的判别标准进行分级，以衡量图像的细节程度。

根据图像的运动剧烈程度，将图像分为准静止块、运动平缓块和剧烈运动块，定义变量 **MotionActivity** 为图像运动剧烈程度标识，分别定义宏 **LOW**、**MEDIUM** 和 **HIGH** 表示准静止块、运动平缓块和剧烈运动块。

表 1 阈值分级模型

If	$T \leq 0.46$	
	MotionActivity=LOW;	(准静止块)
If	$0.46 < T < 0.78$	
	MotionActivity=MEDIUM;	(运动平缓块)
If	$T \geq 0.78$	
	MotionActivity=HIGH;	(剧烈运动块)

表 1 根据对大量图像的统计分析，当阈值 $T \leq 0.46$ 时，图像变化甚微，可归类到准静止块；当 $0.46 < T < 0.78$ 时，图像有较小的细节变化，可归类到运动平缓块；当 $T \geq 0.78$ 时，图像具有较大的细节变化，属于高细节块，因此可归类到剧烈运动块。因此，定义阈值分级模型如表 1 所示。

3.2 改进的帧间模式选择算法

基于表 1 的阈值分级模型，将图像进行分类并根据图像所在的类别采用不同的搜索模板和搜索方式。新的快速帧间模式选择算法步骤如下：

Step1. 根据公式(v)计算 16×16 宏块的 T 值，若该宏块 **MotionActivity** = **LOW**，则该宏块为准静止块，则该宏块的帧间预测模式即为模式 1(如图 1，以下同)；若该宏块 **MotionActivity** = **MEDIUM**，则表示该宏块运动较为平缓，则计算该 16×16 宏块分成的 4 个 8×8 宏块(如图 2(a)所示)；若该宏块 **MotionActivity** = **HIGH**，则表示该宏块运动较为剧烈，则计算该 16×16 宏块分成的 16 个 4×4 宏块(如图 2(b)所示)，转入 Step4；

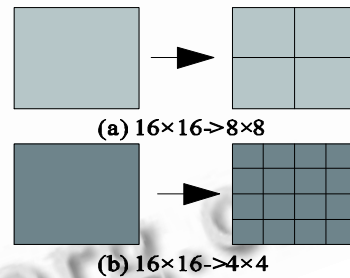


图 2 Step1 示意图

Step2. 对于 Step1 中的 8×8 宏块，若该宏块 **MotionActivity** = **LOW**，则该宏块的帧间预测模式即为模式 4；若该宏块 **MotionActivity** = **MEDIUM**，则可依据模式 2 和模式 3 扩大该宏块的尺寸为 16×8 和 8×16 (如图 3(a)所示)，转入 Step3；若该宏块 **MotionActivity** = **HIGH**，则计算该 8×8 宏块分成的 4 个 4×4 宏块(如图 3(b)所示)，转入 Step4；

Step3. 对于 Step2 中的 8×8 、 16×8 和 8×16 宏块，分别依据公式(iv)计算这些宏块的 T 值，选择 T 值最小的模式作为帧间预测的模式；(注：由于 8×8 宏块已确定属于运动平缓块，因此在这个步骤中的三个宏块的最差也是属于运动平缓块，如由该 8×8 宏块向上合并的 16×8 和 8×16 宏块存在准静止块，

则能找到更优的匹配模式);

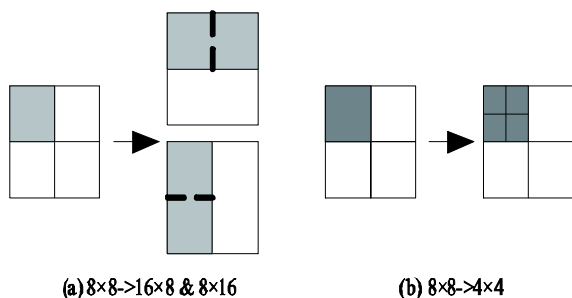


图 3 Step2 示意图

Step4. 对于 4×4 宏块, 如果该宏块仍存在 $MotionActivity = HIGH$, 则表示已经细分到最小匹配模式的宏块, 则该宏块的帧间预测模式即为模式 7; 若该宏块 $MotionActivity = LOW$ 或者 $MotionActivity = MEDIUM$, 则表示细分到 4×4 的宏块属于低细节宏块, 则可依据模式 5 和模式 6 扩大该宏块的尺寸为 8×4 和 4×8 (如图 4 所示), 转入 Step5;

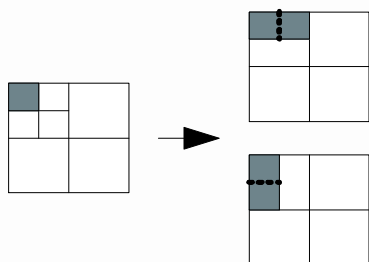


图 4 Step4 示意图

Step5. 对于 Step4 中的 4×4 、 8×4 和 4×8 宏块, 分别依据公式(iv)计算这些宏块的 T 值, 选择 T 值最小的模式作为帧间预测的模式。

本文提出的帧间预测模式选择算法流程图如图 5 所示。

4 实验结果及分析

为了验证本文所提出算法的性能, 在相同的条件下, 对 H.264 原模式选择算法和本文算法在 JM10.1[8] 代码中进行了计算机仿真实验, 实验选取 stefan、foreman、mother-daughter 为测试序列, 每个序列选取前 100 帧。其中 stefan 为网球运动, 属于剧烈运动, 为 cif 格式, 而 foreman 和 mother-

daughter 均为 qcif 格式且为小运动序列。实验对编码时间 ($\Delta t/\%$)、比特率 ($\Delta kbits/\%$)、PSNR ($\Delta PSNR/dB$) 指标进行了验证和比较, 实验结果如表 2 所示。

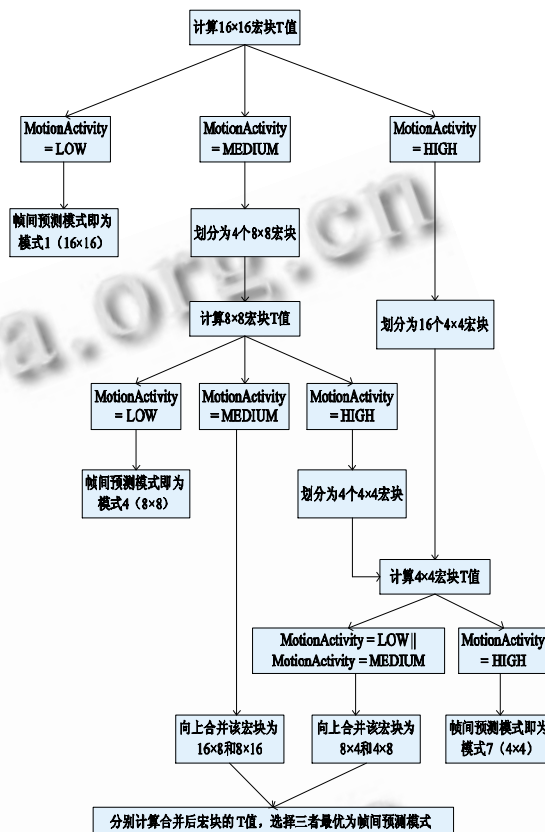


图 5 本文提出的帧间预测模式选择算法示意图

表 2 H.264 原算法和改进算法参数比较

序列	格式	编码时间	比特率	PSNR
M-d	qcif	-52.5 8	+2.19	-0.04
Foreman	qcif	-51.2 6	+2.76	-0.05
Stefan	cif	-42.5 3	+3.51	-0.12

由表 2 可知, 本文提出的模式选择算法相对原 H.264 模式全搜索算法在编码时间上平均减少了 48.79%, 而平均比特率增加仅为 2.82%, 平均峰值信噪比下降仅为 0.07dB, 说明该算法极大地提高了编码的效率, 同时该算法对图像质量的影响很小。

5 结论

本文对 H.264 帧间预测的模式选择算法进行了探讨,并针对 H.264 中帧间预测模式选择采用的全搜索方式,提出了一种快速帧间模式选择算法。该算法在对运动估计准确性和图像质量的影响都不大的同时,大大减少了编码时间,降低了计算复杂度,提高了编码器的编码效率。

参考文献

- 1 宋彬,常义林,李春林. H.264 帧间预测模式的快速选择算法.电子学报,2007,35(4):697-700.
- 2 Puri A, Chen XM, Luthra A. Video coding using the H.264/MPEG-4 AVC compression standard. Signal Processing: Image Communication 19, 2004:793-849.
- 3 周巍,史浩山,周欣. H.264 帧间预测快速算法.计算机辅助设计与图形学学报,2008,20(6):770-774.
- 4 Shenolikar PC, Narote SP. Different Approaches for Motion Estimation. International Conference on Control, Automation, Communication and Energy Conservation. 2009:1-4.
- 5 林巍晓,方向忠,黄修超等. 一种快速的 H.264 帧间模式选择算法.上海交通大学学报,2006,40(1):1-6.
- 6 Lin ZH, Zou YB. A Novel Fast Motion Estimation Algorithm Based on Starting Search Point Prediction. 2009 International Conference on Measuring Technology and Mechatronics Automation: 746-749.
- 7 Bai MS, Tian YP, Tian XD. Fast inter-mode decision algorithm based on UMHexagonS. Computer Applications, 2007,27(9):2150-2152.
- 8 Tourapis AM, Leontaris A et al. H.264/MPEG-4 AVC Reference Software Manual. Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6) 30th Meeting: Geneva, CH, 29 Jan -3 Feb, 2009.