

基于 H.264 的快速帧间模式选择算法

孙力帆, 普杰信

(河南科技大学电子信息工程学院, 洛阳 471003)

摘要: 针对 H.264 视频编码器中计算量大的帧间模式选择过程, 提出一种快速帧间模式选择算法。根据编码块的边缘特征信息和模式相关性, 预测当前宏块可能的候选模式类型。统计并分析帧间编码的最佳模式分布, 通过设定基于统计数据的自适应提前中止门限, 对预测模式进行筛选, 以获取最匹配块模式。实验结果表明该算法能在保证重构图像质量与编码性能的情况下, 减少 51%~68% 的编码时间, 降低编码器的计算复杂度。

关键词: H.264 标准; 模式选择; 边缘信息; 帧间预测

Fast Inter-frame Mode Selection Algorithm Based on H.264 Standard

SUN Li-fan, PU Jie-xin

(College of Electronic Information Engineering, Henan University of Science & Technology, Luoyang 471003)

【Abstract】 Aiming at the inter-frame mode selection process with large calculation of H.264 video encoder, this paper proposes a fast inter-frame mode selection algorithm. According to the edge feature information and mode correlation of encoding block, the candidate mode type of current macro-block is forecasted. The best mode distribution of inter-frame encoder is counted and analyzed. Scale of all the modes is restricted by setting self-adaptive threshold for early termination to get most matching block mode. Experimental results show that the algorithm can decrease the encoding time by 51% to 68% and reduce the computational complexity of encoder while maintaining quality and coding performance of reconfiguration image.

【Key words】 H.264 standard; mode selection; edge information; inter-frame prediction

1 概述

H.264/AVC^[1] 是最新的国际视频编码标准, 由于采用先进的视频编码技术, 因此与以往标准相比, H.264 极大提高了编码效率和图像质量, 但增加了计算复杂度, 导致其难以满足实时视频通信的要求。

近年来, 为降低帧间模式选择的计算复杂度, 研究者提出了很多快速帧间模式选择算法。例如, 文献[2]通过判决是否采用小块尺寸的运动估计来跳过不必要的编码模式。文献[3]通过运动估计获取残差宏块, 对其进行纹理分析, 并根据分析得出的宏块运动特征信息确定帧间宏块最佳编码模式。文献[4]的快速预测模式选择算法利用图像序列的时空相关性, 根据量化因子自适应更新模式选择参数, 提高了帧间预测速度。虽然上述算法可以减少待编码的模式量, 但计算量仍然很大, 且存在一定局限性。它们没有充分考虑块模式的相关性和边缘特征信息对整个编码过程的影响。如果借助此类信息并分析宏块特性, 以指导整个编码过程, 就能避免搜索对编码性能影响不大的模式。因此, 本文提出一种低复杂度的新型快速帧间模式选择算法。

2 H.264 帧间模式选择算法

H.264 在帧间预测编码时, 每个 16×16 宏块可划分成形状不等的区域作为运动描述区域, 如图 1 所示。其中每个宏块可划分成 16×16, 16×8, 8×16 和 8×8 的子块进行宏块级模式编码。对 8×8 块模式编码时, 该子块还将进一步独立地分成 8×8, 8×4, 4×8, 4×4 的亚宏块, 每种划分对应一种编码模式, 能对运动区域进行更精确的分割, 使之更符合图像中背景和运动对象的实际情况。

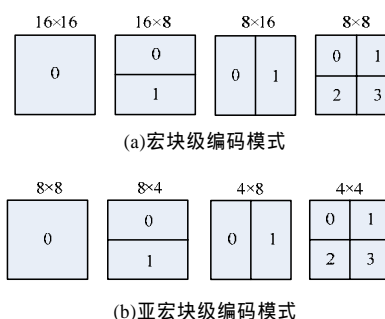


图 1 帧间块模式划分

由于 H.264 针对每个宏块可以提供多种编码模式, 因此在编码端通过采用基于拉格朗日率失真优化判决(Lagrange rate-Distortion Optimization, RDO)的代价函数, 对待编码宏块的所有可供选择的编码模式分别计算各自的率失真代价 $RDCost$, 并选取代价函数最小的模式作为最佳编码模式, 以达到最优的匹配。

$$J_{mode} = SSD_{mode} + \lambda_{mode} \cdot R \quad (1)$$

但因为上述模式选择算法要对每种模式重复执行相同的计算, 即运动估计与补偿和整数 DCT 变换、量化、反量化/整数 DCT 反变换、熵编码, 所以将导致 H.264 的编码运算量

基金项目: 国家自然科学基金资助项目(60475021); 河南省科技攻关计划基金资助项目(0524220019)

作者简介: 孙力帆(1982-), 男, 硕士研究生, 主研方向: 视频编码, 多媒体通信; 普杰信, 教授、博士

收稿日期: 2009-08-24 **E-mail:** octavian2005@126.com

成倍增加。

3 快速帧间模式选择算法

3.1 帧间块模式特点及相关性

各种分块模式在视频图像中不是均匀分布的，因此，H.264 采用多种块尺寸的帧间宏块划分模式，目的是更准确地表示视频场景的运动和纹理信息，从而降低以前标准采用单一宏块划分模式带来的误差。一般地，大尺寸块模式适合于图像平坦和静止区域，如视频会议、可视电话图像中常见的头肩部序列、背景区域位置随时间变化很小而且具有很强空间相关性的图像。这些区域往往位于图像的背景部分或运动对象的内部，当该对象运动时，其内部的各个部分具有相同或相似的活动特性。小尺寸块模式适用于图像纹理较复杂或运动性较剧烈的区域，如包含运动物体边缘的边界区域。如果一个区域内部各个块都具有相同或相似的纹理特征，那么该区域具有同样的空间特性，是具有相同性质的同类区域。

对于不同特点的视频序列，一些模式出现的概率很小，如果能提前预测出此类区域，并对预测模式进行筛选就可以跳过对大量不必要块模式的复杂率失真计算，从而减少模式选择时间。

3.2 基于块边缘特征信息及模式相关性的帧间模式判决

大量研究表明，视频对象的边界能表现出很强的边缘特征，因此，本文根据局部边缘特征信息并使用 Sobel 算子来计算边缘强度 EV ，以预测这些同类区域。对于一个像素 $P_{i,j}$ ，在其水平和垂直方向上定义其边缘强度向量为

$$dx_{i,j} = P_{i-1,j+1} + 2 \times P_{i,j+1} + P_{i+1,j+1} - P_{i-1,j-1} - 2 \times P_{i,j-1} - P_{i+1,j-1} \quad (2)$$

$$dy_{i,j} = P_{i+1,j+1} + 2 \times P_{i+1,j} + P_{i+1,j-1} - P_{i-1,j+1} - 2 \times P_{i-1,j} - P_{i-1,j-1} \quad (3)$$

则每个 4×4 块的边缘强度为

$$EV_{4 \times 4} = \sum_{i,j=0}^3 (|dx_{i,j}| + |dy_{i,j}|) \quad (4)$$

因为 4×4 、 8×8 、 16×16 块之间具有相关性，所以能很容易地通过累加 4×4 块的边缘强度求得 8×8 、 16×16 块的边缘强度。其中， 16×16 块的边缘强度可由 16 个 4×4 块的边缘强度累加计算而得，即

$$EV_{16 \times 16} = \sum_{i=0}^{15} EV_{4 \times 4}(i) \quad (5)$$

8×8 块的边缘强度可由 4 个 4×4 块的边缘强度累加计算而得，即

$$EV_{8 \times 8}(0) = \sum_{i=0}^3 EV_{4 \times 4}(i) \quad (6)$$

$$EV_{8 \times 8}(1) = \sum_{i=4}^7 EV_{4 \times 4}(i) \quad (7)$$

$$EV_{8 \times 8}(2) = \sum_{i=8}^{11} EV_{4 \times 4}(i) \quad (8)$$

$$EV_{8 \times 8}(3) = \sum_{i=12}^{15} EV_{4 \times 4}(i) \quad (9)$$

如果当前块的 EV 小于门限值 TH ，则将其归入具有相同性质的同类块，否则归入非同类块，并对其进行更小的块模式划分。可以使用边缘强度来确定 16×16 、 8×8 、 4×4 是否为当前块的最佳编码模式。

如果 $EV_{16 \times 16}$ 大于 $TH_{16 \times 16}$ ，且每个 $EV_{8 \times 8}$ 都小于 $TH_{16 \times 16}$ ，则使用边缘强度差值 ED 来确定块最佳编码模式为 16×8 或 8×16 。如果 $EV_{16 \times 16}$ 大于 $TH_{16 \times 16}$ 且 $EV_{8 \times 8}$ 大于 $TH_{16 \times 16}$ ，但属

于该 8×8 块的所有 4 个 $EV_{4 \times 4}$ 都小于 $TH_{4 \times 4}$ ，则使用边缘强度差值来确定块最佳编码模式为 8×4 或 4×8 。

因为边缘强度是用来反映一个块内所存在的边缘特征信息，所以如果 2 个相邻块的边缘强度差值小，则 2 个相邻块可能包含同一个对象，即具有相同纹理特征与空间相关性，那么 2 个块属于同类块。 16×8 、 8×16 、 8×4 、 4×8 的边缘强度差值如下：

$$ED_{16 \times 8} = |EV_{8 \times 8}(0) - EV_{8 \times 8}(1)| + |EV_{8 \times 8}(2) - EV_{8 \times 8}(3)| \quad (10)$$

$$ED_{8 \times 16} = |EV_{8 \times 8}(0) - EV_{8 \times 8}(2)| + |EV_{8 \times 8}(1) - EV_{8 \times 8}(3)| \quad (11)$$

$$ED_{8 \times 4}(x) = |EV_{4 \times 4}(4x) - EV_{4 \times 4}(4x+1)| + |EV_{4 \times 4}(4x+2) - EV_{4 \times 4}(4x+3)| \quad (12)$$

$$ED_{4 \times 8}(x) = |EV_{4 \times 4}(4x) - EV_{4 \times 4}(4x+2)| + |EV_{4 \times 4}(4x+1) - EV_{4 \times 4}(4x+3)| \quad (13)$$

其中， $ED_{16 \times 8}$ 与 $ED_{8 \times 4}(x)$ 分别表示其在水平方向上的边缘强度差值； $ED_{8 \times 16}$ 与 $ED_{4 \times 8}(x)$ 分别表示其在垂直方向上的边缘强度差值。如果当前块在水平方向上的边缘强度差值小于其在垂直方向上的边缘差值，则块最佳编码模式为 16×8 或 8×4 ，否则块最佳编码模式为 8×16 或 4×8 。

3.3 基于统计数据的动态调整门限

如果一个宏块内包含许多具有相同性质的同类区域，则该宏块很可能使用大块模式作为最佳编码模式，否则对其进行进一步块模式划分并采用边缘强度与边缘强度差值来确定最终编码模式。本文针对不同视频序列宏块采用的帧间最佳编码模式进行统计，结果如表 1 所示。

表 1 典型序列的帧间最佳编码模式统计 (%)

QP	视频序列	16×16 块	16×8 块	8×16 块	8×8 块
24	Foreman	20	10	14	56
	Tennis	17	12	9	62
	Coast guard	44	10	11	35
	Stefan	33	16	12	39
32	Foreman	46	15	20	19
	Tennis	52	17	13	18
	Coast guard	52	18	18	12
	Stefan	38	14	14	34
40	Foreman	77	11	8	4
	Tennis	73	13	11	3
	Coast guard	76	10	12	2
	Stefan	60	19	13	8

由表 1 可知，量化参数 QP 与块编码模式的选择有很强的相关性。 QP 越大，选用的块模式就越大，即当前块最佳编码模式为 16×16 的概率越大。 QP 越小，则选用的块模式越小，即当前块最佳编码模式为 8×8 及其各种子模式 (8×4 、 4×8 、 4×4) 的概率越大。因此，本文模式搜索策略中采用的提前中止门限不是一个固定常数，而是根据 QP 的变化进行调整，并在搜索过程中对其条件进行判断以确定是否需要后续模式搜索，从而避免不必要的率失真计算，达到降低模式选择复杂度的目的。经过大量实验数据统计，设定门限 $TH_{16 \times 16}$ 为 $1200 \times QP$ ， $TH_{16 \times 16}$ 为 $750 \times QP$ ， $TH_{8 \times 8}$ 为 $600 \times QP$ ， $TH_{4 \times 4}$ 为 $150 \times QP$ 时，算法性能基本达到最优，适用于多数视频序列。

3.4 算法流程

算法流程如图 2 所示。

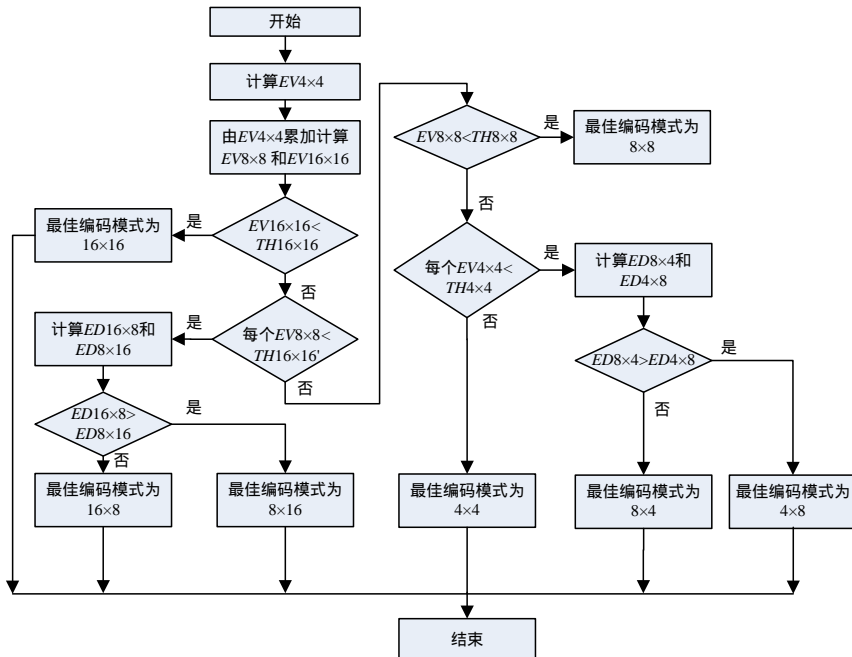


图2 算法流程

4 实验结果与性能分析

本文的快速算法在标准测试模型 JM8.6 上实现，并选取 Foreman, News, Container, Mobile 4 种不同运动程度和纹理特征的 QCIF 序列的前 100 帧进行测试。主要实验参数为：使用哈达玛变换和 RDO，QP=28, 32, 36，熵编码为 CAVLC，GOP 编码结构为“IPPP...”。表 2 给出了本文算法与 JM8.6 模式选择算法的编码性能。

表 2 本文算法与 JM8.6 模式选择算法的编码性能比较

视频序列	QP	时间/(%)	PSNR/dB	码率/(%)
Foreman	28	51.5	-0.05	1.43
	32	53.2	-0.10	1.19
	36	57.5	-0.12	0.62
News	28	60.8	-0.06	0.03
	32	62.3	-0.05	0.05
Container	28	63.2	-0.03	-0.18
	32	66.4	-0.02	0.37
Mobile	28	68.7	-0.03	0.01
	28	52.8	-0.04	0.44
Mobile	32	54.6	-0.04	0.08
	36	57.4	-0.03	-1.19

在不同 QP 下，本文算法比原算法编码时间平均减少 51%~68%，尤其对于 News, Container 这些纹理复杂度一般、空间细节较少且活动性较小的序列来说，其改进效果明显大于 Foreman, Mobile 等运动较剧烈、纹理细节较丰富的序列。在低码率，即 QP 较大的情况下，当前待编码宏块会偏向于选择宏块级编码模式。因此，可以很快找到较优的预测模式，避免不必要的小块模式搜索，从而降低编码时间，使本文快速算法发挥作用。由于在本算法提前中止搜索策略中，采用的门限自适应地根据序列的运动特征及块边缘信息对预测模式进行筛选，排除了某些较小概率的模式，因此在一定程度

上导致了 PSNR 的降低和码率的增加，但码率最多增加 1.43%，PSNR 减少不超过 0.12 dB，小于人眼视觉系统可感知的阈值(0.2 dB)，所以，基本不会影响视频图像的主观视觉效果。

图 3 给出了序列 News, Mobile 分别使用本文算法与 JM8.6 模式选择算法的率失真性能曲线，以分析本算法对综合性能的影响。对于 News 这种运动缓慢、具有较多平坦区域的序列，2 条曲线基本重合。对于运动剧烈且具有复杂背景的 Mobile 序列，本文算法显示了良好的编码性能。

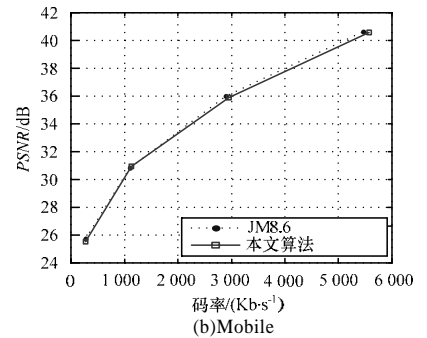
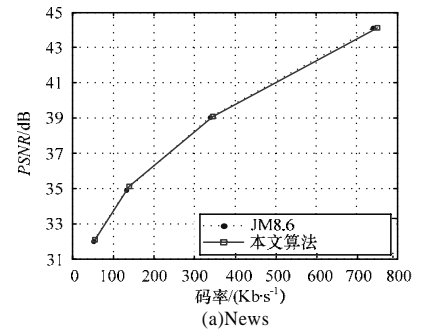


图 3 2 种序列的率失真曲线

5 结束语

基于率失真优化框架下的 H.264 编码器存在复杂度太大、不易于实时应用的问题，因此，本文提出一种低复杂度的新型快速帧间模式选择算法。将块模式相关性与边缘特征信息相结合，利用模式的预判决和模式的提前中止搜索等技术对帧间块编码模式进行预测，以快速选出最优模式。

参考文献

- [1] Wiegand T. JVT-G050-2003 Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification[S]. 2003.
- [2] Li Peijun, Lin Minglong. Fast Inter Mode Selection Algorithm for Motion Estimation MPEG-4 AVC/JVT/H.264[C]//Proc. of 2006 IEEE International Conference on Image Processing. [S. l.]: IEEE Press, 2006: 1365-1368.
- [3] Wang Xiangwen, Sun Jun. Fast Mode Decision for H.264 Video Encoder Based on MB Motion Characteristic[C]//Proc. of 2007 IEEE International Conference on Multimedia and Expo.. [S. l.]: IEEE Press, 2007: 372-375.
- [4] 宋 彬, 常义林, 李春林. H.264 帧间预测模式的快速选择算法[J]. 电子学报, 2007, 35(4): 697-700.

编辑 陈 晖