

文章编号:1001-9081(2008)07-1898-03

一种用于 H. 264 的快速多块类型运动估计算法

颜洪奎^{1,2}, 沈燕飞¹, 朱珍民¹, 肖建华²

(1. 中国科学院 计算技术研究所, 北京 100080; 2. 湘潭大学 信息工程学院, 湖南 湘潭 411105)

(yanhkui@hotmail.com)

摘要:针对 H. 264 在多块运动估计中全搜索算法的计算复杂度高的问题,提出了一种基于图像的空间频率特性来预测运动估计的宏块划分的算法。实验结果表明,该算法在保证图像质量和码率的情况下显著地减少了运动估计计算复杂度的 24%~53%。

关键词:H. 264 帧间模式;多块类型运动估计;空间频率特性

中图分类号:TP319.4 **文献标志码:**A

Rapid algorithm of multi-block motion estimation for H. 264

YAN Hong-kui^{1,2}, SHEN Yan-fei¹, ZHU Zhen-min¹, XIAO Jian-hua²

(1. Institute of Computer Technology, Chinese Academy of Sciences, Beijing 100080, China;

2. College of Information Engineering, Xiangtan University, Xiangtan Hunan 411105, China)

Abstract: The full search algorithm of H. 264 multi-block motion estimation is highly time-consuming. The paper proposed a new multi-block algorithm by speculating motion estimation based on the characteristic of the image's spatial frequency. The results of the experiments show that our algorithm evidently reduces the complexity of motion estimation computing under the condition of keeping good image quality and high bit rate.

Key words: H. 264 inter mode; multi-block motion estimation; spatial frequency characteristic

0 引言

H. 264/AVC 视频编码国际标准是由 ITU-T 和 ISO/IEC MPEG 共同制定的^[1],与以往的标准相比,在相同的编码质量下 H. 264/AVC 可以节约大约 50% 的码率^[2]。但是其编码效率的提高是以增加编码算法复杂度为代价的,其复杂度大约是 H. 263 的 4~5 倍^[3]。

除了用整数 4×4 DCT 代替传统的 8×8 DCT 之外, H. 264 的基本编码算法和 H. 263/MPEG 是很相似的。新增的技术包括有为了更高编码效率的 1 帧的帧内预测模式,多模式的块匹配运动估计和多参考帧选择等。其中多块类型的块匹配运动估计技术是提高 H. 264/AVC 编码效率的主要原因。块匹配运动估计(Block-Matching Motion Estimation, BMME)是基于运动补偿/DCT 混合编码标准的重要组成部分,运动估计与补偿技术可以有效地去除视频序列相邻帧间存在的大量时间冗余,极大地提高了编码效率^[4]。

多块类型的全搜索(Full Search, FS)算法是 BMME 最直接的实现方法,FS 算法是通过如图 1 中所示的每一种块类型进行搜索窗内的所有点进行匹配,从而达到最佳匹配。虽然这种全搜索处理(七种块类型的检查)提供了最优的编码效果,但是它所带来的计算复杂度的增加也是惊人的。在对 H. 264/AVC 各编码模块计算量的统计结果中,在使用七种块类型的 BMME 比只用 16×16 一种块类型的 BMME 能节省 15% 的码率^[5],但是它占据了整个编码大约 80% 的计算量^[6]。因此研究快速算法成为目前这一领域的重要研究课题。

在本文中,我们提出了一种全新的快速块类型选择算法,它集中在 16×16 (mode 1), 16×8 (mode 2) 和 8×16 (mode 3)

三种模式,通过计算图像区域的空间频率特性(Spatial Frequency Characteristic, SFC)来预测块的划分模式,忽略了没有性能表现的块类型的搜索。在保证图像质量和码率情况下有效地减少了计算开销的 24%~53%,这对于实时实现是很有意义的。

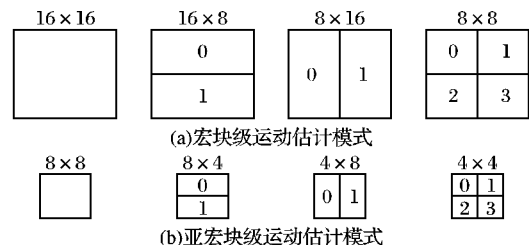


图 1 宏块及子宏块的划分

1 帧间编码模式及模式选择方法

准确的预测能有效地减少原始图像和预测图像之间的残差数据。正是基于这个理由, H. 264 为了获得更好的压缩性能而采用了多块类型的运动估计/运动补偿。但是对于一个宏块(MB)来说,它可能包含了多于一个的对象而且这些对象的运动方向也是不尽相同的。如果只使用一个运动矢量将会只有一部分的宏块能得到很好的运动补偿,而在不匹配宏块中,运动补偿后的残差能量将会很大。

正是基于上述的原因,如图 1 中所示, H. 264/AVC 在帧间预测中将一个宏块(MB)分割为 7 种大小不同的块模式。多种预测模式为宏块预测提供了更多的选择,且每种块模式能进行亚像素和多参考帧的运动估计和补偿,从而可显著提高预测的准确性。除了 7 种帧间预测模式外,在 P 帧还支持

收稿日期:2008-01-18;修回日期:2008-03-17。

作者简介:颜洪奎(1978-),男,湖南湘潭人,硕士,主要研究方向:多媒体、DSP 视频处理;朱珍民(1962-),男,湖南慈利人,教授,主要研究方向:嵌入式、多媒体、普适计算、动力学系统、实时数值仿真、并行算法;沈燕飞(1976-),男,江苏泰州人,硕士,主要研究方向:音视频编解码;肖建华(1963-),男,湖南邵阳人,教授,博士,主要研究方向:优化算法、数据挖掘。

skip 模式和帧内预测模式。帧内预测模式包括 4×4 (Intra 4) 和 16×16 (Intra 16) 两种亮度尺寸。 4×4 (Intra 4) 块支持九种预测方向; 16×16 (Intra 16) 支持四种预测方向。

在 H.264/AVC 的参考模型中,运动估计和模式类型是以 RDO^[7] 为判决依据。首先对所有帧间块模式进行运动估计,然后比较所有模式(包括帧内预测模式)的 RDO, RDO 最小的模式为最佳模式。运动估计通过式(1)确定:

$$J_{\text{motion}}[mv, ref | \lambda_{\text{motion}}] = SAD[s, r(mv, ref)] + \lambda_{\text{motion}} [R(mv - pred) + R(ref)] \quad (1)$$

其中: $mv = (m_x, m_y)^T$ 为运动矢量, λ_{motion} 为拉格朗日系数。 s 是当前块像素值; $pred$ 为预测矢量; ref 为选择的参考帧; $r(ref, mv)$ 为参考块的像素值; SAD 为当前块与参考块像素差值的绝对值和,它的计算公式如下:

$$SAD(s, c(m)) = \sum_{x=1, y=1}^{B, B} |s[x, y] - c[x - m_x, y - m_y]| \quad (2)$$

其中: $B = 16, 8$ 或 4 。 R 包括运动矢量与其预测值的差值编码比特和参考帧的编码比特。

模式判决由式(3)确定:

$$J_{\text{mode}}(s, c | Q, \lambda_{\text{mode}}) = SSD(s, c | Q) + \lambda_{\text{mode}} R(s, c | Q) \quad (3)$$

其中: c 为重建块; R 包括对宏块编码的所有比特; SSD 为当前块和重建块的差值的平方和。相对于式(1),得到重建块需要对残差进行变换和量化与反量化反变换,其过程较复杂。

很显然,这种全搜索的多块类型运动估计/补偿方法能有效地提高预测的准确性,但是它同时也给编码器带来了沉重的计算负担,因此为能实际应用有必要研究快速算法。

2 快速算法的分析与设计实现

2.1 图像的纹理特征与宏块的划分模式

图像采用的编码模式与其内容是有很大关系。一般地,静止区域和运动简单的区域(如背景)采用匹配块较大的帧间编码模式,运动复杂的区域(如图像中的运动物体)采用匹配块较小的帧间编码模式^[8]。如图2中所示的一个静态背景上有一个运动的对象。对于当前帧A,当前宏块的两个 8×16 子块有相对一致的纹理特征,而对于当前帧B中的当前宏块则是两个 16×8 子块有相对一致的纹理特征,从而为了获得更好的运动补偿效果,对于当前帧A中的当前宏块应当划分成两个 8×16 的子块,而对于当前帧B的当前宏块应当划分成两个 16×8 子块,这样的残差将会是最小的。

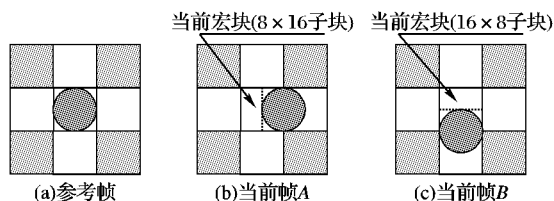


图2 静态背景上的运动物体示意图

在视频序列中,如果一个图像区域的细节或是它所包含的纹理是固定的(如有一致的变化率),则这样的图像区域被认为是一致的,一个典型的高一致性区域的例子是在卡通中的一致颜色区域。而这种图像区域的一致性程度可以通过来计算它的空间频率特性来衡量^[9]。图像的空间频率特性对应为图像在空间上的波形频率,平缓的低细节区域表现为低频成分的高度集中,而高频成分则对应的是图像的变化剧烈或是细节较多的区域。

计算图像的空间频率特性有多种方法,一种是通过图像进行 DCT 变换后对 DCT 系数值取平均来衡量该图像区域的一致性程度。计算图像区域的空间频率特性的另一种方法是计算它的均方差(MSE)值来衡量。因为对图像进行 DCT 变换是很耗时的,所以本文采用的是计算图像区域的 MSE 值来衡量图像的一致性的^[9]。

2.2 快速多块尺寸运动估计

仿真结果显示,在通常的情况下,有 70% 的宏块将会选择模式一(16×16)作为它们最终的块类型,而且模式 4 至模式 7 占全部模式的 10% 以下^[5]。因此本文提出的算法在保证图像质量和码率的情况下关闭了使用少的块模式 4 到模式 7,而主要集中在模式 1、模式 2 和模式 3 这三种模式的选择。算法流程如下。

1) 定义阈值变量等,并做初始化,对初始化只进行一次,后面的宏块需要动态修正这些值。

定义 4 个变量如下:

T : 宏块的 MSE 值。

$MSE2$: 模式 2 中两个 8×16 块中小的 MSE 值。

$MSE3$: 模式 3 中两个 16×8 块中小的 MSE 值。

$BestMode$: 最佳的宏块划分模式。

初始化设置 $T = 0$, $Best_{\text{mode}} = 1$, $MSE2$ 和 $MSE3$ 均为零。

2) 通过计算当前宏块的 MSE 值以及 $MSE2$ 、 $MSE3$ 值得来得到宏块和块的一致性程度,并对这种一致性程度进行比较和判断。 M_{\min} 为 $MSE2$ 和 $MSE3$ 中的较小值,也即一致性程度更高的块的 MSE 值。

$M_{\min} = \min(MSE2, MSE3)$;

如果 M_{\min} 小于 T , 则转第 3 步, 否则转第 4 步。

3) 通过上一步计算得到了宏块和块的一致性程度,从而设置宏块划分模式为一致性程度更高的模式。

如果 $MSE2$ 小于 $MSE3$, 则 $BestMode = 2$; 否则 $BestMode = 3$ 。

4) 模式二和模式三的一致性程度都低于模式一, 所以将宏块的划分模式设为模式一。

$BestMode = 1$ 。

5) 必须对所有的宏块进行上述计算。

跳转到第 2 步, 直到处理完所有的宏块。

为了进一步提高算法的效率,本算法采用了一种独立的模块形式来进行处理,处理模块的结构如图3所示。图中,外部视频源的视频序列被分为两路,一路输入到缓冲模块中先缓冲,另一路输入到宏块划分模块将每帧的数据划分成 16×16 宏块大小;紧接在宏块划分模块之后的是宏块的空间频率特性计算模块,它将计算出宏块的空间频率特性;块模式选择模块则依据空间频率特性计算模块计算所得的空间频率特性来进行块类型的选择;最后编码模块依据块模式选择模块所得到的块类型对由缓冲模块来的视频序列进行编码。

3 实验结果

本文的实验采用了 JVT 的 JM7.3 作为实验平台,主要的实验参数如表 1 所示。

为了客观全面地验证算法的有效性,本文对不同运动程度和纹理特点视频序列分别做了测试,选择的测试序列为 CIF (352 × 288) 格式的“Akiyo”、“Foreman”和“Mobile”。本实验进行了两组测试工作,一组是采取全搜索算法 FS (Full Search) 对全部预测模式进行编码,另一组是采用本文提出的 SFC 预测的快速运动估计算法,分别对测试序列的图像质量 (SNR(Y))、码率 (Bitrate) 和帧间运动估计的时间进行比较。

表 2 为测试序列“Foreman”使用 FS 算法和本文提出的 SFC 预测算法的信噪比 (SNR(Y)) 和比特率 (BitRate) 的对比数据,从表中可以看出,快速 SFC 预测算法对于不同的量化值 QP 在图像质量和比特率上都只有很细微影响。

表 3 为所有测试序列在不同量化值 QP 下 FS 算法和 SFC 预测算法的性能比较。对于图像细节少、背景静止且运动量

小的测试序列“Akiyo”,SFC 预测算法的图像质量、码率和 FS 算法的保持一致的同时也取得了良好的降低计算复杂度的效果。而对于图像的细节信息量大,且运动剧烈的测试序列“Mobile”,SFC 预测算法在图像质量和比特率上有一定的下降,从而可以看出,本文提出的 SFC 预测算法很适合于视频会议等纹理一致(即图像细节少,运动量小)的场合。

参数	值
FrameToBeEncoded	100
GOP structure	IPPP
UseHadamard	Used
FrameRate(Hz)	30
NumberReferenceFrames	2
InterSearchA × B	All are enable
SymbolMode	CABAC
PartitionMode	No DP
QP	26, 28, 30, 32
UseConstrainedIntraPred	0(= not constraint)
MVResolution	1/4-pel
SearchRange	Global
RateControl	Disable

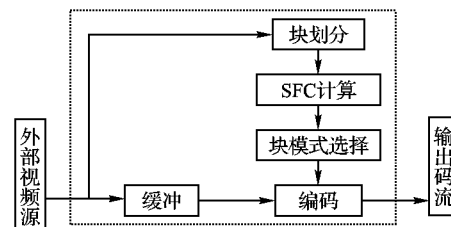


图 3 独立处理模块方框图

QP	FS 算法		SFC 预测算法	
	SNR(Y)/dB	码率/bps	SNR(Y)/dB	码率/bps
26	38.04	468.70	38.03	475.42
28	36.95	360.32	36.94	368.04
30	35.73	273.90	35.73	279.27
32	34.58	211.73	34.53	215.88

表 3 测试序列在不同 QP 值时的图像质量和码率对比

测试序列	QP	FS 算法			SFC 预测算法			优化对比		
		SNR(Y)/dB	码率/bps	时间/ms	SNR(Y)/dB	码率/bps	时间/ms	SNR(Y)/dB	码率提高百分比/%	减少时间百分比/%
Akiyo	28	40.05	151.85	26.86	40.05	152.45	12.50	0.00	0.0%	-53.45%
	32	39.59	76.54	25.84	39.58	77.61	12.82	-0.01	1.4%	-50.40%
Foreman	28	36.95	360.32	51.94	36.94	368.04	26.48	-0.01	2.1%	-49.01%
	32	34.58	211.73	50.53	34.53	215.88	26.06	-0.05	2.3%	-48.43%
Mobile	28	34.70	696.45	80.43	34.65	748.68	55.88	-0.05	7.5%	-30.52%
	32	31.27	470.21	81.55	31.20	511.11	60.06	-0.07	8.7%	-24.34%

图 4 为测试序列“Foreman”在使用 SFC 预测算法和全搜索运动估计算法对比上模式划分的正确率对比图。

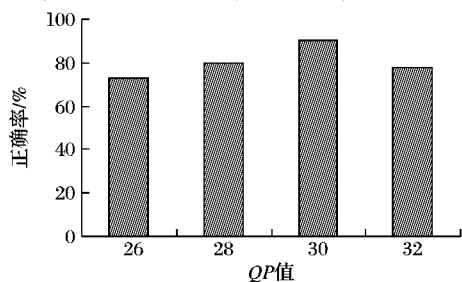


图 4 测试序列“Foreman”的正确率

4 结语

运动估计中的块划分模式选择是 H. 264/AVC 提高压缩效率和保证图像质量的重要技术之一。然而采用率失真优化虽然能做到最佳的块划分模式,却需要非常大的计算复杂度。本文对此提出了基于图像空间频率特性的运动搜索预处理的快速块划分模式选择算法。从实验结果可以看出本文算法在块划分模式的选择上进行了有效的提前预测,在明显减少计算复杂度的情况下只以图像质量和码率方面的微小牺牲作为代价。而且如果将本文所提出的算法结合优秀的快速运动搜索算法的话,块划分模式选择的计算速度将会进一步地加快。

参考文献:

- [1] WIEGAND T. Joint video team (JVT) of ISO/IEC MPEG and ITU-T VCEG[CD]. Joint Committee Draft, JVT-C167, 2002.
- [2] WIEGAND T, SULLIVAN G J, BJONTEGAARD G, et al. Over-

- view of the H. 264/AVC video coding stand[J]. IEEE Transactions on Circuits and Systems Video Technology, 2003, 13(7): 560 - 576.
- [3] LAPPALAINEN V, HALLAPURO A, HAMALAINEN T D. Optimization of emerging H. 26L video encoder[C]// IEEE Workshop on Signal Processing Systems. Washington, DC: IEEE Press, 2001: 406 - 415.
- [4] 沈兰荪,卓力,田栋,等. 视频编码与低速率视频传输[M]. 北京: 电子工业出版社, 2001.
- [5] CHANG A, AU O C, YEUNG Y M. A novel approach to fast multi-block motion estimation for H. 264 video coding[C]// 2003 International Conference on Multimedia and Expo 2003. Washington, DC: IEEE Press, 2003, 1(6/9).
- [6] ZHANG JIAN-NING, HE YU-WEN, YANG SHI-QIANG, et al. Performance and complexity joint optimization for H. 264 video coding[C]// Proceedings of the 2003 International Symposium on Circuits and Systems 2003. Washington, DC: IEEE Press, 2003, 2(25/28): II-888 - II-891.
- [7] WIEGAND T, SCHWARZ H, JOCH A, et al. Rate-constrained coder control and comparison of video coding standards[J]. IEEE Transaction on Circuits and System for Video Technology, 2003, 13(7): 688 - 703.
- [8] CHAN M H, YU Y B, CONSTANTINIDES A G. Variable size block matching motion compensation with application to video coding[J]. Communications, Speech and Vision, IEE Proceedings I, 1990, 137(4): 205 - 212.
- [9] BURAZEROVIC D, VERVOORT G J M. VIDEO ENCODING, US PCT/IB04/50145[P]. 2006 - 07 - 27.