

H.264/AVC 帧间预测模式的快速分类方法

裴世保, 李厚强, 俞能海

(中国科学技术大学电子工程与信息科学系, 合肥 230026)

摘要: H.264/AVC 采用 RDO 技术进行帧间模式选择, 提高了编码效率, 但增加了计算复杂度。该文提出了快速分类算法, 利用相邻宏块率失真特征以及预测模式之间的时空相关性, 缩小了模式选择的范围, 降低了帧间预测的复杂度。结果表明, 该方法可减少 25%~46% 的序列编码时间, 保持了 H.264/AVC 的编码性能与视觉效果。

关键词: H.264/AVC 视频编码标准; 帧间预测; 模式选择; 率失真优化

Fast Classification Method for H.264/AVC Inter-frame Prediction Mode

PEI Shibao, LI Houqiang, YU Nenghai

(Department of Electronic Engineering and Information Science, University of Sciences and Technology of China, Hefei 230026)

【Abstract】 H.264/AVC employs RDO technique in inter-frame mode selection to enhance coding efficiency, but it increases computational complexity drastically. The proposed fast algorithm exploits rate distortion characteristic and spatio-temporal correlation among neighboring macroblocks to reduce the range of candidate modes, which can efficiently save the inter prediction complexity. Experimental results show that the approach reduces the coding time by 25%~46% with similar visual quality and bit rate of H.264/AVC.

【Key words】 H.264/AVC video coding standard; inter-frame prediction; mode decision; rate distortion optimization(RDO)

新一代视频编码标准 H.264/AVC 在压缩效率、图像质量、网络适应性等方面具有优越性, 适用于交互、非交互应用环境^[1]。在提高 H.264/AVC 性能的同时, 编码器的计算复杂度大幅度地增加, 大约是 MPEG-2 的 9 倍^[2], “编码器的计算复杂度”制约着 H.264/AVC 的发展。

导致编码器复杂度大幅度增加的原因包括: 运动估计, 变换和量化, 环路滤波, 熵编码, 模式选择等。模式选择过程是增加编码器复杂度的一个非常重要的因素。H.264/AVC 为了提高编码效率, 采用 RDO 技术进行模式选择, 对于每个宏块(16×16)分别计算每种模式的率失真代价(RDcost), 选择具有最小代价的模式作为最优模式。帧间模式代价函数的计算涉及运动估计、变换/反变换、量化/反量化等复杂过程, 增加了模式选择过程中的计算量。如果能根据某些特征, 预先判断出最优模式或者预测出可能的模式范围, 就可以避免不必要的代价计算, 减少模式选择时间, 降低编码器的复杂度。研究帧间预测模式的快速分类方法具有重要的实用价值。

1 帧间模式选择的分析

1.1 帧间预测模式介绍

H.264/AVC 的每个亮度宏块可以划分为 4 种尺寸的子块(16×16、16×8、8×16、8×8)进行运动补偿计算, 对应的预测模式分别记为 MODE_16×16、MODE_16×8、MODE_8×16、MODE_8×8。如果选择了 MODE_8×8, 每个 8×8 的子块可以进一步划分为微子块(8×8、8×4、4×8、4×4), 对应的预测模式分别记为 INTER_8×8、INTER_8×4、INTER_4×8、INTER_4×4。

除了以上介绍的模式, 对于 P 帧还有一种重要的 SKIP 模式, 主要用于编码图像中静止或具有全局运动的部分。由于这种模式不需要对残差信号和运动矢量进行编码, 因此编

码一个宏块只需很少的比特。本文主要研究 P 帧的快速模式分类算法。

1.2 采用 RDO 技术的模式选择过程

H.264/AVC 支持 16×16~4×4 共 7 种不同尺寸和形状的运动补偿块, 同时还允许在帧间编码方式下, 采用帧内预测模式^[3]。帧间模式选择过程就是从所有帧间、帧内模式中, 选择 RDcost 最小的模式, 从而用最少的码率获得最优的图像质量, 提高压缩效率。

代价函数公式为

$$RDcost = SSD + \lambda * Rate$$

其中, SSD(sum of squared difference)表示当前宏块与重建宏块差值的平方和, 称为失真量(distortion), 反映了运动补偿的效果; λ 表示拉格朗日系数, $\lambda = 0.85 * 2^{(QP-12)/3}$; Rate 表示运动矢量、预测模式及运动补偿后的残差等信号经过变换、量化以及熵编码后的码率^[4,5]。

由于 H.264/AVC 支持的模式数较多, RDcost 计算复杂, 因此模式选择过程的复杂度显著增加。

2 快速模式选择算法分析

2.1 相邻宏块之间的时空相关性

视频序列是由多帧具有一定顺序的静止图像构成的, 帧内和帧间图像的宏块之间具有较强的相关性。为了验证相邻宏块预测模式的相关性, 本文选择图 1 所示的相邻宏块位置关系(C 表示当前宏块, 1、2 分别表示其左边和上边的宏块,

基金项目: 国家自然科学基金资助重点项目“高效视频编码中的关键技术研究”(60333020)

作者简介: 裴世保(1972-), 男, 硕士, 主研方向: 视频处理与通信; 李厚强, 副教授; 俞能海, 教授

收稿日期: 2006-08-22 **E-mail:** psb@ustc.edu.cn

3表示前一帧中对应于C的宏块),分别测试了相邻宏块中有2个以上(包括2个)的SKIP、MODE_8×8模式时,当前宏块为SKIP、MODE_8×8的百分比,相邻宏块的相关性统计如表1所示(用百分比表示)。

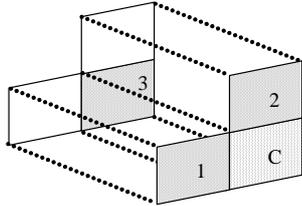


图1 相邻宏块的位置关系

表1 相邻宏块的相关性统计

	QP=24		QP=28		QP=32		QP=36		QP=40	
	SKIP	8×8	SKIP	M8	SKIP	M8	SKIP	M8	SKIP	M8
Bus	29.09	69.34	36.74	68.35	48.27	63.72	61.33	56.96	68.20	52.36
Football	54.86	77.91	65.30	75.29	73.13	70.17	80.91	61.54	85.40	45.50
Foreman	53.62	60.60	59.83	51.01	66.39	40.60	70.31	30.93	74.81	28.26
Mobile	10.34	72.45	23.71	67.30	42.46	57.69	61.66	50.10	72.76	39.90
News	87.94	62.43	88.95	58.44	90.30	59.95	91.94	56.10	93.35	49.77
Paris	86.14	75.22	87.89	70.64	88.62	66.17	89.09	57.65	89.84	40.44
Silent	87.51	64.34	88.33	56.46	89.07	45.81	90.41	31.66	91.89	25.42
Stefan	69.52	71.00	68.30	70.12	68.11	65.74	70.78	57.05	74.69	51.48

相邻宏块的模式之间确实具有较强的相关性,这种相同性质的模式聚集在一起的现象,可以判断模式、预测模式的大致范围。在相邻宏块中,预测模式、率失真特性之间均具有较强的相关性。

2.2 SKIP 优先的快速算法思想

代价函数 RDcost 主要包括码率(rate)和失真(distortion)两项分量,与进行运动补偿的块大小和形状密切相关。对于SKIP模式,由于运动矢量是零或较小值,因此编码后的码率项很小。而对于8×8以及更小的划分方法,一般适用于编码图像中运动剧烈的部分,其特征是运动矢量较大、编码后的码率较大。对于这类宏块,如果仍用SKIP模式进行编码,失真量比用最优模式编码得到的失真量增加许多倍。

对于每个宏块,首先采用SKIP模式进行编码,然后根据编码后的率失真特征以及相邻宏块的模式相关性判断是否是SKIP模式。如果不是SKIP模式,则利用获得的失真量以及模式相关性来判断它属于某类模式,通过缩小模式选择的范围来减小模式选择的复杂度。对于部分特征不明显、难以判断是否为SKIP模式、难以进行分类处理的宏块,采用H.264/AVC原来的模式选择方案,确定最优模式。

这样设计主要是基于2个原因:(1)SKIP模式的比例较高^[6],如果当前宏块是SKIP模式,就省去很多其它模式的RDcost计算,大幅度地减小了模式选择的复杂度。(2)即使当前宏块不是SKIP模式,仍可以利用SKIP模式编码后获得的失真量来帮助判断它属于某类模式。

如何判断SKIP模式是算法性能的关键。假设当前宏块是SKIP模式,那么根据模式聚集的特点,其相邻宏块也很可能是SKIP模式。同时,实验发现相邻SKIP模式宏块的失真量非常接近。如果用SKIP模式编码当前宏块,获得的失真量D接近相邻SKIP模式宏块的平均失真量AD(比值R近似

1.05),可以初步判断当前宏块是SKIP模式。

如果 $1.05 < R < 2$ 且相邻宏块的模式属于大的划分类型,那么可以判断当前宏块的范围M16 {MODE_16×16,MODE_16×8,MODE_8×16}。如果比值R大于阈值T,同时,相邻宏块的模式属于小的划分类型,那么可以判断当前宏块的模式范围M8 {MODE_8×8,MODE_8×4,MODE_4×8,MODE_4×4}。为了确定合适的阈值T,采用H.264/AVC原模式选择算法编码一帧或多帧图像,统计MODE_8×8模式宏块用SKIP模式编码后的失真量与实际失真量之比,选取最大比值与最小比值的平均值作为阈值T。

实验结果显示,该算法简单、有效,在基本保持H.264/AVC原来编码性能的情况下,平均可减少25%~46%的编码时间(QP:24~40)。算法的流程如图2所示。

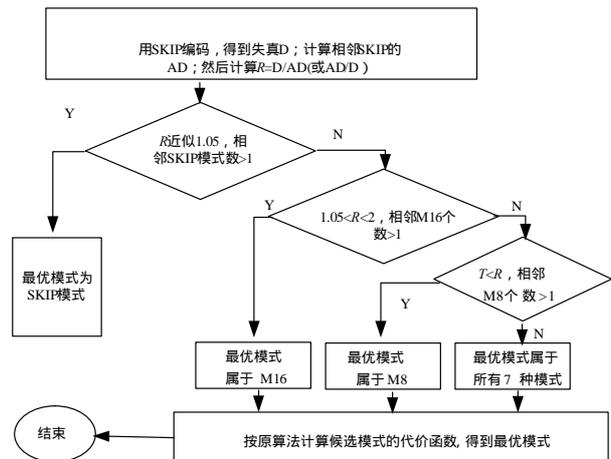


图2 快速算法流程

3 实验结果及分析

本文将该快速算法应用于H.264/AVC JM61e参考软件^[7]中,在IPPP...、RDO模式下,分别测试了不同种类的QCIF格式的视频序列,其中Bus 150帧、Football 250帧、Foreman 300帧、Mobile 300帧、News 300帧、Paris 300帧、Silent 300帧、Stefan 300帧。实验结果如表2所示,表中正数表示快速算法比H.264/AVC原算法的增加量,负数表示快速算法比原算法的减少量,BITS、TIME两列用百分比表示。实验结果显示,采用该算法平均可减少25%~46%的序列编码时间(QP:24~40),基本保持了H.264/AVC的编码性能。当QP较大时,快速算法的性能更为优越,适用于各种性质的序列,由此说明该快速算法对SKIP模式的判断是准确的。

表2 快速算法与H.264/AVC原算法的比较

参数	QP=24			QP=28			QP=32			QP=36			QP=40		
	PSNR	BITS	TIME												
Bus	-0.01	-0.02	-12.00	-0.02	-0.25	-16.64	-0.02	-0.37	-24.02	-0.05	-0.11	-30.84	-0.04	-0.25	-33.54
Football	-0.01	-0.02	-11.77	-0.02	-0.08	-14.30	-0.02	0.03	-18.18	-0.06	0.12	-21.66	-0.07	-0.41	-25.49
Foreman	-0.05	-0.13	-24.73	-0.05	-0.55	-30.56	-0.04	-0.35	-35.51	-0.10	-0.70	-40.25	-0.12	0.18	-45.35
Mobile	-0.02	-0.04	-5.92	-0.05	-0.49	-28.74	-0.10	-1.16	-46.63	-0.08	-0.06	-53.71	-0.05	-0.37	-57.70
News	-0.04	0.03	-47.64	-0.07	0.07	-50.22	-0.03	-0.42	-50.84	-0.05	0.16	-52.13	0.02	0.19	-51.72
Paris	-0.03	0.36	-44.95	-0.09	0.42	-49.06	-0.09	0.57	-52.26	-0.07	0.78	-55.15	-0.15	0.15	-57.11
Silent	-0.05	-0.04	-48.43	-0.06	-0.05	-52.08	-0.02	-0.78	-54.63	-0.07	-0.98	-56.57	0.01	-0.13	-59.49
Stefan	-0.01	-0.10	-7.76	-0.02	-0.10	-13.90	-0.03	-0.29	-21.77	-0.03	-0.41	-30.76	-0.10	-0.24	-38.94
平均值	-0.03	0.0	-25.4	-0.05	-0.13	-31.94	-0.04	-0.35	-37.98	-0.06	-0.15	-42.63	-0.06	-0.11	-46.17

(下转第212页)