

采用块匹配的多视点视频编码算法的改进

黄君婷, 冯桂

(华侨大学 信息科学与工程学院, 福建 厦门 361021)

**摘要:** 利用相邻已编码宏块的模式对当前编码宏块的模式选择进行优化,减少部分细化小块的分割及搜索. 然后结合传统单视点视频编码中的简化非对称十字型多层次六边形格点搜索,对运动估计搜索算法进行改进. 实验表明:在保证其重建视频质量及码率开销的前提下,两种改进算法使不同特性的视频序列的编码时间平均减少 78%.

**关键词:** 多视点视频编码; 块匹配; 模式选择; 运动估计搜索算法; 非对称十字型多层次六边形格点搜索

**中图分类号:** TN 919. 81      **文献标志码:** A

多视点视频<sup>[1]</sup>是由多台摄像机组成的阵列从不同角度对同一场景进行拍摄,捕获到的一组视频信号. 与传统单视点视频相比,多视点视频能够更加生动地再现现实场景,给人以身临其境的感觉. 多视点视频编码(multi-view video coding, MVC)技术是未来视频通信领域的一项关键技术,主要致力于如何对多视点视频进行高效的压缩编码<sup>[2]</sup>. 多视点视频的数据量会因为摄像机数目的增加而成倍增加<sup>[3]</sup>. 因此,多视点视频编码要处理更大的数据量,在存储和传输上有更高的要求. JVT(joint video team)发布的 MVC 的软件参考模型 JMVC( joint multi-view video coding)兼具 JM(joint model)中多尺寸块分割、多参考帧、块匹配等算法,同时又根据多视点视频编码的特性对每个分割块采用了运动估计搜索算法(TZSearch)算法. 在 JMVC 中,预测结构采用了多种模式,这种方法虽然可以大幅度提高编码效率,但其计算量也在成倍地增加<sup>[4]</sup>. JMVC 中的 TZSearch 采用了光栅搜索、八点钻石搜索、星形搜索等多种搜索模型,较之全搜索在编码时间上有了很大的提高,但仍然有很大的改进空间. 本文对模式选择及 TZSearch 算法进行改进,在保证重建视频质量及码率的前提下,尽可能地减少模式选择及 TZSearch 算法所消耗的时间,从而减少多视点视频编码的时间.

1 模式复杂度

JMVC 中采用的预测模式有 SKIP, Inter 16×16, Inter 16×8, Inter 8×16, Inter 8×8, Inter 8×4, Inter 4×8, Inter 4×4 和 Intra 4, Intra 8, Intra 16. 文献[5]中,作者统计了 5 个不同特性序列的模式选择,表明选择大尺寸模式的宏块比例远大于小尺寸模式. 在平滑区域,几乎都选择了大尺寸模式,只有在运动边缘、纹理边缘和深度不连续区域,选择小尺寸模式的概率比较大<sup>[6]</sup>. 在视频编码过程中将图像分割为不重叠的块,相邻的块之间具有很大的相关性,相邻块的模式选择也具有一定的关系. 鉴于此,利用相邻已编码宏块的编码模式信息得到当前宏块的模式复杂度,针对不同复杂度的宏块分配不同的模式集. 当前宏块与其相邻已编码宏块,与当前块 E 相邻的已编码宏块有上边块 A、左边 B、左上块 C 以右上块 D,如图 1 所示.

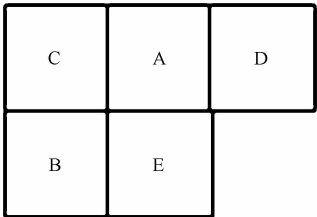


图 1 当前宏块与其相邻已编码宏块

Fig. 1 Current macroblock and the adjacent coded macroblock

定义当前宏块的预测模式复杂度(PMC),有

$$PMC = \frac{1}{4}(\omega_A + \omega_B + \omega_C + \omega_D),$$

(1)

式(1)中: $\omega_A, \omega_B, \omega_C, \omega_D$  分别表示块 A,B,C,D 编码模式的权重. 不同预测模式的权重,如表 1 所示.

表 1 预测模式的权重

Tab. 1 Weight of prediction model

模式	Inter								Intra
	SKIP	16×16	16×8	8×16	8×8	8×4	4×8	4×4	
$\omega$	0	1	2	2	3	4	4	5	6

通过计算 PMC 的值,将宏块预测模式复杂度分为  $C_1, C_2, C_3$  3 个等级, $C_1 < C_2 < C_3$ . 则有

$$\left. \begin{aligned} PMC &\leq 0.25, & MB &\in C_1, \\ 0.25 < PMC &\leq 0.5, & MB &\in C_2, \\ PMC &> 0.5, & MB &\in C_3. \end{aligned} \right\}$$

(2)

对 10 个不同特性的序列进行测试,得到不同复杂度宏块的编码模式分布,如表 2 所示. 由表 2 可知: $C_1$  中,超过 99%的宏块的编码模式为 SKIP 模式或者 Inter 16×16,选择小尺寸块模式的宏块不到 1%.  $C_2$  中,超过 94%的宏块选择 SKIP 或者 Inter 16×16,平均有 5%的宏块选择 Inter 16×8 或者 Inter 8×16,其余不到 1%的宏块选择 Inter 8×8.  $C_3$  中,选择大尺寸块模式的约有 75%,Inter 16×8,Inter 8×16,Inter 8×8 分别为 8.86%,9.94%,4.38%. 由此可知:不同复杂的宏块,其模式分布不同,复杂度越高,选择小尺寸块概率越大.

表 2 不同复杂度宏块的编码模式分布

Tab. 2 Coded model distribution in different complexity macroblock

序列	$C_1$					$C_2$					$C_3$					%
	SKIP	16×16	16×8	8×16	8×8	SKIP	16×16	16×8	8×16	8×8	SKIP	16×16	16×8	8×16	8×8	
ballroom	97.28	2.24	0.24	0.20	0.04	68.40	24.33	3.42	3.35	0.50	39.36	37.47	9.48	11.07	2.62	
exit	98.20	1.51	0.11	0.17	0.01	69.11	24.65	2.55	3.52	0.17	53.97	30.83	5.51	8.76	0.94	
vassar	99.05	0.82	0.06	0.05	0.01	73.18	20.39	3.45	2.64	0.33	49.68	29.62	9.38	9.49	1.83	
flmenco2	95.26	3.90	0.43	0.36	0.05	74.49	18.15	3.42	3.46	0.48	44.59	31.99	11.08	9.65	2.69	
objects2	98.34	1.56	0.05	0.04	0.02	80.05	18.40	0.48	0.81	0.27	30.40	25.78	9.20	16.90	17.73	
crowd	93.32	5.43	0.48	0.57	0.20	72.61	20.58	3.01	3.11	0.69	43.76	32.79	8.94	10.03	4.49	
racel	95.14	4.31	0.30	0.19	0.06	72.59	22.42	2.78	1.93	0.28	48.17	32.98	10.59	6.10	2.16	
breakdancer	96.62	3.04	0.20	0.14	0	70.58	23.95	2.57	2.77	0.13	49.61	35.85	6.93	6.92	0.69	
uli	96.50	2.97	0.24	0.26	0.04	74.09	19.91	2.66	2.99	0.36	42.57	33.00	9.49	11.28	3.65	
puppy	97.98	1.90	0.05	0.06	0.01	78.10	17.25	2.15	2.46	0.05	46.50	29.32	7.99	9.20	7.00	
平均	96.77	2.77	0.22	0.20	0.04	73.32	21.00	2.65	2.70	0.33	44.86	31.96	8.86	9.94	4.38	

2 模式分配算法

根据表 2,提出以下 3 个步骤的模式分配算法.

**步骤 1** 根据当前宏块的相邻已编码宏块的编码模式,由式(1)计算得到当前宏块的 PMC 值.

**步骤 2** 根据 PMC 的值,由式(2)得到当前宏块所属复杂度等级.

**步骤 3** 为不同复杂度等级的宏块分配模式. 具体分配原则如下: $C_1$  中的宏块只进行 SKIP 模式和 Inter 16×16;对  $C_2$  中的宏块分配的模式集为{SKIP,Inter 16×16,Inter 16×8,Inter 8×16};而对  $C_3$  中的宏块则采用原算法即遍历所有模式的算法.

3 快速搜索算法的改进

在多视点视频编码中,除了有帧内预测和帧间预测外,还存在视点间预测. 对于帧间预测,传统单视点视频编码中采用的 Simplified UMHexagonS<sup>[7]</sup>,能针对各种运动特性的视频序列取得良好的效果,且

编码速度较快. 对于视点间相关性较大的序列, TZSearch 算法中<sup>[8]</sup>采用的光栅搜索, 能够迅速又准确地匹配到最优点. 因此, 对帧间预测与视点间预测采用不同的算法, 达到提高编码速度的效果. 具体有如下 4 个改进算法步骤.

**步骤 1** 确定初始化集合, 其中包含中值预测矢量, 参考帧对应宏块的左、上、右上位置宏块的矢量以及 (0, 0) 矢量. 从中选择最优矢量的坐标作为下一步的搜索中心.

**步骤 2** 判断当前块进行的是否是帧间编码. 若是则进行 Simplified UMHexagonS, 然后结束搜索; 否则, 跳至步骤 3.

**步骤 3** 进行光栅搜索, 根据 iStartX 的值设定光栅搜索的步长. 若 |iStartX| 大于阈值, 则步长为 4, 否则步长为 3.

**步骤 4** 同 TZSearch 步骤 4, 细化搜索, 在得到的搜索中心重复地依次进行步长为 1, 2, 4, 8, 16, 32, 64 共 7 轮的八点菱形搜索或八点方形搜索, 直到最优匹配点出现在搜索中心, 此时退出循环.

4 实验结果及分析

4.1 实验说明

实验平台为 JMVC8.3, 所用 PC 配置为 Inter(R) Core(TM)2 Quad CPU2.66 GHz, 4 GB 内存, Windows 7 操作系统. 测试序列共有 10 个, 分别为 ballroom (640 × 480), breakdancer (512 × 384), crowd (640 × 480), exit (640 × 480), flamenco2 (640 × 480), objects2 (640 × 480), puppy (720 × 480), race1 (640 × 480), uli (1024 × 768), vassar (640 × 480). 其中 breakdancer 编码 90 帧, 其余 9 个序列编码 100 帧. 这 10 个序列中, ballroom, breakdancer, flamenco2, crowd, uli 都是属于大面积运动的序列, exit, objects2, puppy, vassar 都是属于背景基本静止, 运动较为平缓的序列, race1 是摄像头随运动目标移动, 且运动较为剧烈的序列. 量化参数 QP 为 37, GOP 大小为 8, 参考帧数为 2, 搜索范围为 64. 实验测试了 2 个视点, 即 0, 1 视点, 其中 0 视点为主视点, 只进行帧内预测和帧间预测, 视点 1 除了进行帧内预测与帧间预测, 还进行了视点间预测, 实验数据均来自视点 1.

4.2 模式分配算法实验结果及分析

模式分配算法与遍历模式选择算法比较, 其改变量如表 3 所示. 表 3 中:  $\eta_1$  为比特率变化率;  $\eta_2$  为时间变化率. 从表 3 中可以看出: 模式分配算法在编码时间上较原算法有很大的提高, 且峰值信噪比 (PSNR) 值减小量不超过 0.2 dB, 比特率变化较小. 根据各个序列的时间变化率可以看出, 模式分配算法对于背景基本静止、运动较平缓的序列效果较好, 对大面积运动的序列效果稍差一些.

4.3 改进的快速搜索算法实验结果及分析

改进的快速搜索算法与 TZSearch 比较, 其改变量如表 3 所示. 由表 3 可以看出: 改进的快速搜索算法在保证了 PSNR 值与比特率变化较小的前提下, 提高了编码速度. 且该算法对大面积运动的序列以及运动场景变化且运动较激烈的序列的改进效果更好.

4.4 综合改进实验结果及分析

上述两种算法的综合改进与 JMVC 中原算法比较, 其改变量如表 3 所示. 从表 3 可以看出: 两种算法综合起来对 10 个不同特性的序列都达到了大幅减少编码时间的效果 (平均提高了 78%). PSNR 值普遍下降, 但均不超过 0.11 dB, 比特率普遍变大, 不过增大的幅度不超过 2%. 该算法在减少编码时间的同时, 对 PSNR 值及比特率影响较小, 达到了很好的效果.

表 3 几种算法的比较结果

序列	模式分配与遍历模式选择算法			改进的快速搜索与 TZSearch			综合改进与 JMVC 原算法		
	$\Delta$ PSNR/dB	$\eta_1$ /%	$\eta_2$ /%	$\Delta$ PSNR/dB	$\eta_1$ /%	$\eta_2$ /%	$\Delta$ PSNR/dB	$\eta_1$ /%	$\eta_2$ /%
ballroom	-0.027 9	-0.193 9	46.529 5	-0.047 3	0.288 1	48.176 9	-0.076 0	0.006 3	76.382 0
breakdancer	-0.009 4	-0.210 3	48.835 6	-0.073 2	-0.013 0	48.103 9	-0.098 4	-0.815 4	77.276 5
crowd	-0.047 4	0.076 7	55.672 0	-0.007 3	-1.232 8	40.628 1	-0.061 4	-1.183 5	75.218 1
exit	-0.021 9	-0.844 7	64.936 6	-0.011 7	-0.392 7	39.474 9	-0.037 7	-1.324 2	80.048 2

续表

Continue table

序列	模式分配与遍历模式选择算法			改进的快速搜索与 TZSearch			综合改进与 JMVC 原算法		
	$\Delta$ PSNR/dB	$\eta_1$ /%	$\eta_2$ /%	$\Delta$ PSNR/dB	$\eta_1$ /%	$\eta_2$ /%	$\Delta$ PSNR/dB	$\eta_1$ /%	$\eta_2$ /%
flamenco2	-0.168 8	0.167 3	49.349 2	-0.103 3	-1.737 1	75.851 1	-0.049 9	-1.451 7	47.191 4
objects2	-0.007 6	-0.747 8	69.237 3	0.018 0	-1.160 7	24.102 1	-0.003 1	-1.405 8	77.798 8
puppy	-0.004 1	-0.952 4	72.877 7	-0.038 2	0.324 3	25.659 5	-0.053 4	-0.566 5	78.849 0
race1	-0.025 6	-0.807 8	59.471 4	-0.008 7	-1.067 8	55.879 8	-0.030 6	-0.888 4	82.556 1
uli	-0.030 0	0.170 3	50.334 8	-0.014 9	-0.592 0	39.049 2	-0.042 4	-0.599 4	73.719 9
vassar	-0.009 9	-1.172 0	72.687 1	-0.023 5	1.194 6	24.563 6	-0.039 8	0.278 6	78.822 0

5 结束语

分析 JMVC 中的模式选择算法与快速搜索算法 TZSearch,发现当前宏块与相邻与编码宏块间模式的相关性,建立预测模式复杂度模型.根据当前宏块的复杂度为其分配模式集,避免进行不必要的模式搜索.此外,针对不同预测方法的特点,选择不一样的搜索算法,达到提高搜索速度的效果.

实验结果表明:两种改进算法结合起来对不同特性的序列均能取得一定程度的效果,在保证重建视频的质量及编码开销的同时,提高了编码速度.

参考文献:

[1] SMOLIC A. Introduction to multi-view video coding[C]//73th MPEG Meeting. [s. l.]: MPEG Video Subgroup, 2008;23.

[2] 霍俊彦,常义林,李明,等.多视点视频编码的研究现状及其展望[J].通信学报,2010,5(31):113-121.

[3] 张兆杨,安平,张之江,等.发展 3DTV 需解决的技术及其应用趋势[J].电视技术,2010,3(46):4-6.

[4] 张兆杨,安平,张之江,等.二维和三维视频处理及立体显示技术[M].北京:科学出版社,2010;137.

[5] SHEN Li-quan,LIU Zhi,LIU Su-xing,et al. Selective disparity estimation and variable size motion estimation based on motion homogeneity for multi-view coding [J]. IEEE Transaction on Broadcasting,2009,55(4):761-766.

[6] SHEN Li-quan,LIU Zhi,ZHANG Zhao-yang,et al. Faster inter mode decision using spatial property of motion field [J]. IEEE Transactions on Multimedia,2008,10(10):1208-1214.

[7] 毕厚杰.新一代视频压缩编码标准:H.264/AVC[M].北京:人民邮电出版社,2005;41-43.

[8] 汤秀丽,戴声奎,杨中华.联合多视点视频编码中的快速搜索算法分析[J].华侨大学学报:自然科学版,2011,32(4):393-396.

Improvement of Multi-View Video Coding  
Algorithm Using Block Matching

HUANG Jun-ting, FENG Gui

(College of Information Science and Engineering, Huaqiao University, Xiamen 361021, China)

**Abstract:** Use the mode of adjacent coded macroblock to optimize the mode decision of current macroblock, and reduce the division and search of refined small block. The motion estimation search algorithm is improved in combination of simplified unsymmetrical cross multi hexagon search with traditional single view video coding. Experimental results show that under the premise of insuring reconstruction video quality and coding overhead, the two improved algorithm can reduce the encoding time 78% averagely.

**Keywords:** multi-view video coding; block matching; mode decision; motion estimation search algorithm; unsymmetrical cross multi hexagon search