

文章编号: 1000-5013(2011)04-0393-04

# 联合多视点视频编码中的快速搜索算法分析

汤秀丽, 戴声奎, 杨中华

(华侨大学 信息科学与工程学院, 福建 泉州 362021)

**摘要:** 对比分析联合多视点视频编码中的全搜索算法和 TZSearch(TZS)快速算法,描述 TZS 算法的原理及搜索过程并对其进行分析,并根据实验统计结果设置了使搜索提前终止的阈值.实验表明,在编码比特率轻微增加和峰值信噪比略微下降的条件下,设置了阈值的 TZS 算法比原有的联合多视点视频编码中的 TZS 算法在编码时间上大幅减少.对于运动比较平缓的序列适合采用阈值为 2 的 TZS2 算法,而对于运动比较剧烈的序列适合采用阈值为 3 的 TZS3 算法.

**关键词:** 联合多视点视频编码; 快速搜索算法; 终止搜索; 阈值; 运动估计

**中图分类号:** TN 919.81 **文献标志码:** A

在当前的联合多视点视频编码(JMVC)<sup>[1]</sup>参考代码中,视差估计和运动估计是共用的,采用与运动估计相似的基于块的算法,选择相邻视点对应时刻的帧为参考帧.在最新视频编码标准 H.264/AVC 中,由于允许多参考帧和可变块尺寸编码,运动和视差估计搜索时就要对每一种分块模式和每一个参考帧遍历搜索,计算量非常大.据统计,编码单路视频的时候,运动估计大约占总编码时间的 50%(1 个参考帧)至 80%(5 个参考帧)<sup>[2-3]</sup>,是限制编码速度的主要瓶颈.因此,联合多视点视频编码中的运动和视差估计虽然提高了压缩效率,但却以巨大的处理时间为代价.显然,这在需要对视频数据进行实时传播的场合是不切实际的.本文研究和分析联合多视点视频编码中 TZSearch 快速算法.

## 1 TZSearch 算法原理与分析

全搜索算法(PelBlockSearch, PBS)和混合的快速搜索算法(TZSearch, TZS)都是联合多视点视频编码参考代码中提供的搜索算法. TZSearch 的搜索过程有如下几个步骤.

(1) 建立预测运动矢量集合. 包括通过中值预测得到的运动矢量,参考帧对应宏块的左、上、右上位置宏块的运动矢量、(0,0)位置运动矢量.

(2) 以预测运动矢量集中具有最小匹配误差的的点为搜索中心,依次进行步长为 1,2,4,8,16,32,64(若搜索范围为 64)的钻石形搜索(xTZ8PointDiamondSearch),如图 1 所示.

(3) 当步骤(2)中搜索得到的最优点距当前点的距离 cStrukt. uiBestDistanced 大于 iRaster(光栅扫描的步长)时,将进行 iRaster 个像素步长的光栅扫描.

(4) 取步骤(2),(3)中的最佳点作为细化的起点,根据情况进行 xTZ8PointDiamondSearch 或 xTZ2PointSearch(对最佳点的边缘点进行扫描),直到求得 cStrukt. uiBestDistance 等

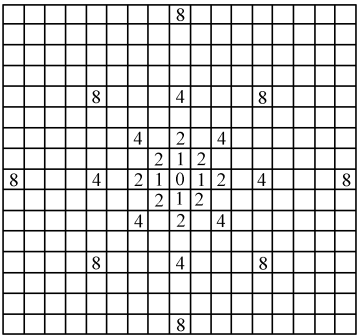


图 1 步长 8 以内的钻石形搜索

Fig. 1 Diamond search of stepsize which is smaller than 8

于 0 为止.

对以上两种搜索算法进行实验对比,共编码 248 帧,其结果如表 1 所示.表 1 中: $\Delta R_{\text{PSN}}$  为峰值信噪比的变化; $\Delta R_{\text{ATE}}$  为序列总码率的变化; $\Delta t$  为序列编码时间的变化;量化参数  $P_Q$  为 37.所做测试使用的是多视点视频编码的经典测试序列, MERL(Mitsubishi Electric Research Laboratories)的 ballroom, exit 和 vassar.这 3 组视频均是由相距 20 cm 的一字形排列的摄像机获得,采集帧率为  $25 \text{ 帧} \cdot \text{s}^{-1}$ ,图像大小为  $640 \text{ px} \times 480 \text{ px}$ ,采样格式为  $4:2:0$ ,其平均峰值信噪比( $R_{\text{PSN}}$ )<sup>[4]</sup> 为

$$R_{\text{PSN}} = \frac{R_{\text{PSN},Y} \cdot 4 + R_{\text{PSN},U} + R_{\text{PSN},V}}{6}.$$

其中: $R_{\text{PSN},Y}, R_{\text{PSN},U}, R_{\text{PSN},V}$  分别为  $Y, U, V$  分量的峰值信噪比.

由表 1 可知,相比于 PBS 算法, TZS 算法可以在失真代价很小的情况下取得一定编码时间的降低,但效果还不是很理想,需做进一步的优化和改进.

表 1 PBS 算法与 TZS 算法的实验对比

Tab. 1 Experiment contrast of PBS and TZS algorithm

视频序列	算法	$R_{\text{PSN}}/\text{dB}$	$\Delta R_{\text{PSN}}/\text{dB}$	$R_{\text{ATE}}/\text{kbit} \cdot \text{s}^{-1}$	$\Delta R_{\text{ATE}}/\%$	$t/\text{s}$	$\Delta t/\%$
exit	PBS	36.027 4	0.001	88.098 4	0.072	20 550.5	-14.43
	TZS	36.028 3		88.162 1		17 585.2	
vassar	PBS	34.487 3	0.001	62.728 2	-0.125	16 860.1	-21.55
	TZS	34.488 3		62.650 0		13 226.0	

2 TZSearch 统计分析和阈值确定

由 TZSearch 算法的步骤(2)可知,若搜索范围为 64,则要进行 7 次钻石形搜索,此时的计算量最大.若在每次搜索的过程中找到新的最优点,则令  $\text{cStrukt. uiBestRound} = 0$ ;否则,  $\text{cStrukt. uiBestRound}$  自加 1.因此,  $\text{cStrukt. uiBestRound}$  只有 8 种值,分别是 0,1,2,3,4,5,6,7,代表找到最优点后使用钻石形搜索的次数.

统计  $\text{cStrukt. uiBestRound}$  分别为这 8 个值的概率情况,如图 2 所示.图 2 中: $D$  为  $\text{cStrukt. uiBestRound}$  值; $P$  为概率值.由图 2 可知, $D=7$  的概率最大,即原点处就是最优点; $D=6$  次之,即在步长为 1 的钻石形搜索中找到了最优点; $D=5$  再次之,在步长为 2 的钻石形搜索中找到了最优点.说明在当前搜索点的周围找到最优点的概率最大,也符合运动矢量的中心偏置特性.因此,对 TZSearch 算法的步骤(2)设置阈值使其提前终止搜索,以达到降低编码复杂度、减少编码时间的目的.阈值  $T$  设置是当步骤(2)中的  $D>T$ (理论上取值范围为  $1\sim6$ ),则停止搜索,继续执行步骤(3).

根据实验得到的统计结果,可将阈值  $T$  的取值范围缩小为 1,2 或 3.这是因为根据运动矢量的中心偏置特性和上述实验统计结果可知,在当前搜索点进行步长小于等于 4 的钻石形搜索就可以比较早地找到最优点.

3 实验结果和性能评价

实验平台为 Intel Pentium M1.6 G,1 GM 内存,Microsoft XP 操作系统.量化参数  $P_Q$  分别采用 32,37 和 42,搜索范围为 64,画面组(GOP)大小为 8.对视频序列 exit,vassar,soccer,puppy 各编码 88 帧,以及日本电信服务提供商 KDDI 提供的 objects1,flamenco1 序列各编码 200 帧.soccer 与 puppy 的图像大小为  $720 \text{ px} \times 480 \text{ px}$ ;objects1 与 flamenco1 的图像大小为  $320 \text{ px} \times 240 \text{ px}$ ;TZS 是联合多视点

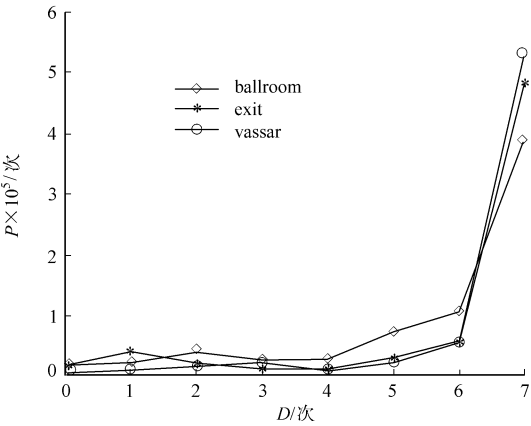


图 2 cStrukt. uiBestRound 分布规律  
Fig. 2 Distribution of cStrukt. uiBestRound

视频编码中的 TZSearch 快速算法;TZS1,TZS2,TZS3 分别是阈值为 1,2,3 的 TZSearch 快速算法.表 2 为不同算法的编码时间比较,图 3 为各个序列不同算法的失真率(RD)测试结果.表 2 中: $t_{TZS}$ , $t_{TZS1}$ , $t_{TZS2}$ 和  $t_{TZS3}$  分别为 TZS,TZS1,TZS2,TZS3 算法的编码时间.

表 2 不同算法的编码时间比较

Tab. 2 Encoding time contrast of different algorithms

视频序列	$P_Q$	$t_{TZS}/s$	$t_{TZS1}/s$	$\Delta t/\%$	$t_{TZS2}/s$	$\Delta t/\%$	$t_{TZS3}/s$	$\Delta t/\%$
exit	32	6 313.69	2 892.61	54.19	3 506.89	44.46	4 169.22	33.97
	37	5 749.38	2 406.41	58.14	2 977.89	48.21	3 603.86	37.32
	42	5 169.81	1 952.48	62.23	2 466.84	52.28	3 038.59	41.22
vassar	32	4 821.66	2 206.75	54.23	2 729.45	43.39	3 247.45	32.65
	37	4 313.53	1 879.23	56.43	2 332.00	45.94	2 815.38	34.73
	42	3 768.64	1 571.28	58.31	1 944.39	48.41	2 396.55	36.41
objects1	32	4 822.55	2 309.06	52.12	2 910.63	39.65	3 481.26	27.82
	37	4 396.83	1 961.59	55.39	2 491.11	43.34	3 015.59	31.41
	42	3 932.03	1 597.41	59.37	2 077.78	47.16	2 573.61	34.55
flamencol	32	3 934.56	1 926.16	51.05	2 437.61	38.05	2 953.59	24.93
	37	3 516.24	1 461.48	58.44	1 909.39	45.70	2 407.63	31.53
	42	3 184.72	1 136.88	64.30	1 515.80	52.40	1 985.25	37.66
soccer	32	6 733.05	2 964.31	55.97	3 765.83	44.07	4 569.84	32.13
	37	5 760.80	2 420.19	57.99	3 033.66	47.34	3 715.56	35.50
	42	4 706.42	1 988.47	57.75	2 475.94	47.39	3 018.50	35.86
puppy	32	7 101.05	3 542.33	50.12	4 314.89	39.24	5 038.64	29.04
	37	6 464.22	2 944.50	54.45	3 672.41	43.19	4 388.45	32.11
	42	5 643.06	2 299.14	59.26	2 970.63	47.36	3 666.73	35.02

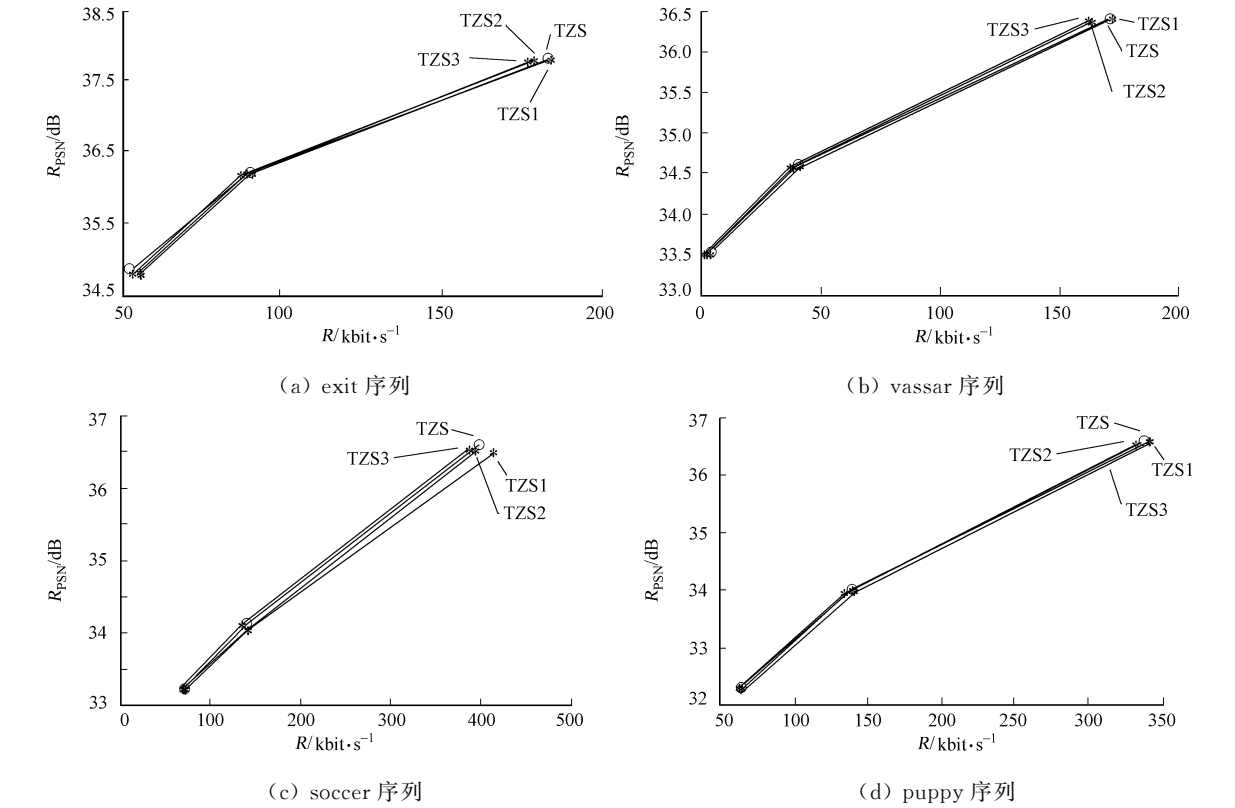


图 3 4 种编码序列的失真率曲线

Fig. 3 RD curves of four kinds encode sequences

从表 2 的结果可知,设置了阈值的 TZS1,TZS2 和 TZS3 算法比 TZS 算法在编码时间上有大幅减少,而总体峰值信噪比下降非常小,均未超过 0.08 dB,码率增加不超过 5%.

对于 exit 序列,使用 TZS3 算法比使用 TZS1 算法和 TZS2 算法在峰值信噪比的降低和编码比特率的增加上有所改善,且 TZS3 算法能保证 33.97%~41.22%的编码时间降低.因此,exit 序列适合用 TZS3 算法.对于 vassar 序列,TZS2 和 TZS3 两种算法的失真率曲线几乎重合;而在编码时间的降低,TZS2 算法(43.39%~48.41%)优于 TZS3 算法(32.65%~36.41%).因此,vassar 序列适合用 TZS2 算法.同理,对于 objects1,soccer 序列适合用 TZS3 算法,flamencol,puppy 序列适合用 TZS2 算法.

exit 序列的内容是个出口,人流量比较大,序列运动比较剧烈;objects1,soccer<sup>[5]</sup>序列也属于运动比较剧烈的序列;vassar 序列的图像中包含有大面积的静止背景,时间相关性要远远强于视点间相关性;flamencol,puppy 序列属于运动比较缓慢的序列.

综上所述,对于运动比较剧烈的序列适合采用 TZS3 算法,即采用步长稍微大的钻石形搜索(步长小于等于 4);对于运动比较平缓的序列适合采用 TZS2 算法,即采用步长稍微小的钻石形搜索(步长小于等于 2).此外,在高码率条件下,采用阈值的算法与原算法差别不大.

4 结 束 语

目前,TZS 算法使用的是一种粗犷的复杂度控制算法,其阈值集合的最大个数 6,而理论上阈值越多控制越精细.因此,今后可以沿着保证编码效率、提高复杂度控制精细度等方向进一步研究.

参 考 文 献:

[1] 杨劼,姜光.基于视觉几何的多视点视频编码研究[D].西安:西安电子科技大学,2010.  
[2] XU Xiao-zhong,HE Yun.Improvements on fast motion estimation strategy for H.264/AVC[J].IEEE Transactions on Circuits and Systems for Video Technology,2008,18(3):285-293.  
[3] CHEN Zhi-bo,XU Jian-feng,HE Yun,et al.Fast interger-pel and fractional-pel motion estimation for H.264/AVC [J].Journal of Visual Communication and Image Representation,2006,17(2):264-290.  
[4] 朱劲,罗琳.基于 JMVC 参考模型的双目视觉压缩编码研究[D].南京:东南大学,2010.  
[5] 林志鹏,罗琳.基于双目视觉的立体视频压缩关键技术研究[D].南京:东南大学,2009.

An Analysis of TZSearch Algorithm in  
Joint Multiview Video Coding

TANG Xiu-li, DAI Sheng-kui, YANG Zhong-hua

(College of Information Science and Engineering, Huaqiao University, Quanzhou 362021, China)

**Abstract:** Comparative Analysis of full search algorithm and TZSearch (TZS) fast algorithm in joint multiview video coding, the principle and search process of TZS algorithm was described and analyzed. According to the experimental results, some threshold for early termination search was set. Experimental results show that under the conditions of a slight increase in the encoding bit rate and a slight decrease in the peak signal to noise ratio, TZS algorithm which is setted threshold has significant reduction in encoding time compared to the original TZS algorithm in the joint multiview video coding. For the sequence of gentle exercise are suitable for using the TZS2 algorithm with threshold value of 2, and for the sequence of more intense exercise are suitable for using the TZS3 algorithm with threshold value of 3.

**Keywords:** joint multiview video coding; TZSearch algorithm; termination search; threshold; motion estimation

(责任编辑:钱筠 英文审校:吴逢铁)