

文章编号 1004-924X(2010)11-2489-08

H. 264 快速运动估计算法的改进

李桂菊¹, 刘 刚^{1,2}, 梁静秋¹

(1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033; 2. 中国科学院 研究生院, 北京 100039)

摘要: 为了提高视频压缩效率, 对 H. 264 中采用的快速运动估计算法 UMhexagonS 进行了改进。首先, 在起点预测后加入一个结束搜索条件, 判断是否结束对当前宏块的运动估计。然后, 对于需要进一步运动估计的宏块, 从两个方面对原算法进行了改进。提出了搜索模板分割方法, 在进行十字形模板和大六边形模板搜索时, 只需要根据该方法选择少量搜索点进行运动估计。根据运动矢量的统计特点, 减少了正方形模板和扩展六边形模板的搜索点数。同时验证了所采用的搜索区域分割方法的合理性。实验结果显示, 改进算法的运动估计时间比原算法平均减少了 15.59%, 而峰值信噪比和码率基本不变, 并且能够适应各种运动类型的视频序列。得到的结果表明改进算法提高了总体编码性能。

关键词: 视频压缩; 区域分割; UMhexagonS 算法; 运动估计; 统计特点

中图分类号: TN941.1; TP301.6 **文献标识码:** A **doi:** 10.3788/OPE.20101811.2489

Improvement of fast motion estimation algorithm used in H. 264

LI Gui-ju¹, LIU Gang^{1,2}, LIANG Jing-qiu¹

(1. *Changchun Institute of Optics, Fine Mechanics and Physics,
Chinese Academy of Sciences, Changchun 130033, China;*

2. *Graduate University of Chinese Academy of Sciences, Beijing 100039, China*)

Abstract: A fast motion estimation algorithm used in H. 264, UMhexagonS, is optimized to improve the efficiency of video compression. Firstly, a stop-search technique is used to judge whether the motion estimation of current macro block should be stopped after the initial search point is predicted. Then, the original algorithm is improved in two respects for macro blocks which need further motion estimation. One improvement is that fewer search points are used in motion estimation by the proposed partition method when multi-hexagon-grid pattern or cross pattern are adopted, and the other improvement is that the search points of square search pattern and extended-hexagon search pattern are reduced according to statistical characteristics of motion vectors. Moreover, the rationality of the partition method is proved. Experimental results show that the average time of motion estimation is reduced by 15.59% compared with that of the original algorithm, whereas, the Peak Signal to Noise Ratio (PSNR) and bit rate remain almost unchanged. The proposed algorithm improves the coding performance and is suitable for various types of video sequences.

收稿日期: 2010-01-22; 修订日期: 2010-04-13.

基金项目: 中国科学院知识创新工程领域前沿项目

Key words: video compression; area partition; UMhexagonS algorithm; motion estimation; statistical characteristics

1 引言

运动估计是视频压缩领域中的一项关键技术,而块匹配方法由于其适合硬件实现,因此被广泛应用于运动估计中。在已有的块匹配方法中,全搜索算法(FS)是对给定的搜索窗口内的所有点进行匹配,因此能够得到最佳的运动估计效果,但是随着搜索半径的增大,其运算量成倍增长,导致整个编码过程时间消耗过多。

为了减少运动估计消耗时间,提高视频编码效率,近年来国内外学者提出了各种快速算法。其中,新三步法(NTSS)^[1]、钻石法(DS)^[2]、以及基于六边形的搜索法(HEXBS)^[3]等,都是利用运动矢量的中心偏置分布特性^[4]来进行搜索模板和搜索方法设计,加快了搜索速度,而且在视频序列运动比较平缓的情况下表现出良好运动估计效果。但是,这些算法没有考虑到运动矢量的相关性,在运动剧烈的情况下容易陷入局部最优而导致运动估计效果降低。而后来提出的运动矢量场自适应快速搜索算法(MVFAST)^[5],以及可预测运动矢量场自适应快速搜索算法(PMVFAST)^[6]则在先前基础上进一步考虑到运动矢量的时空相关性^[7],对运动矢量进行精确的起点预测,缩小了搜索区域范围,同时降低了搜索陷入局部最优的可能性,不过这些算法搜索模板相对单一,搜索策略不够灵活,并且运动估计效果依然有待改进。作为被 H. 264/AVC 参考编码模型 JM10.2 所采纳的快速运动估计算法之一,非对称十字型多层次六边形搜索算法(UMhexagonS)利用运动矢量的水平垂直偏置分布特性和合适的半途中止条件^[8],结合其他搜索方法的部分优点,提出了混合搜索策略,能够在比 FS 节省约 90% 的计算量下,得到非常接近 FS 的运动估计效果,并且能够适应于各种运动类型的视频序列。但是,该算法没有利用运动矢量的方向性^[9],搜索过程中会进行不必要的搜索,因此,为了加快搜索速度,需要进行改进。

由于物体运动的复杂性,混合搜索策略能更

好的利用运动矢量的各种性质,因此目前最新的各种快速算法也大都采用了混合搜索策略^[10-11]。但是这些混合算法都不同程度地牺牲了运动估计效果,导致 PSNR 的降低,以此换取搜索速度的提高。本文在研究 UMhexagonS 混合搜索策略的基础上,通过在原算法中加入运动矢量方向性判断条件和提前结束搜索条件,并且对 UMhexagonS 的搜索模板和搜索策略进行改进,减少不必要的搜索,提高了运动估计速度,同时获得了良好的 PSNR 性能。

2 基于 UMhexagonS 的改进算法

2.1 UMhexagonS 算法介绍

UMhexagonS 算法主要包含了以下 3 个部分。

(1) 精确的运动矢量起点预测。预测矢量集包含了当前宏块的中值预测矢量、时间相关块预测矢量、上层预测矢量、邻近参考帧预测矢量。运动矢量起点预测就是在上述预测矢量集中找出能够使得公式(1)给出的拉格朗日代价函数最小的运动矢量,以此作为最佳起点运动矢量(记为 MV_{PRED})。

$$m_i = \underset{m \in M}{\operatorname{argmin}} \{ D_{\text{DFD}}(S_i, m) + \lambda_{\text{MOTION}} R_{\text{MOTION}}(S_i, m, \text{ref}) \}, \quad (1)$$

式(1)中, S_i 表示当前宏块, M 为可以得到的编码模式的集合, R_{MOTION} 为编码运动矢量 m 和对应参考帧 ref 所需比特数, D_{DFD} 为失真度。

(2) 混合搜索模板的使用。对于预测得到的最佳搜索起点, UMhexagonS 算法以该点为中心,进行如图 1 所示的非对称十字形模板(记为 A)搜索,找到该模板中最小代价所在点后,以此为中心,再进行 $[-2, +2]$ 范围内的正方形模板搜索(记为 B),同样找到最佳点后,再以该点为中心进行大六边形模板(记为 C)搜索,然后用相同的方法依次对六边形模板 D、小菱形模板 E 不断搜索,最终得到的运动矢量就是当前宏块的最佳运动矢量(记为 MV_{BEST})。

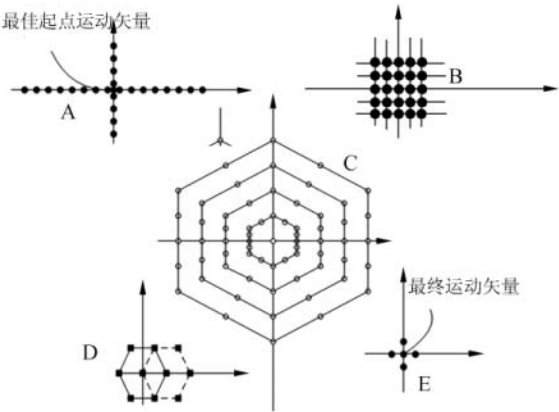


图 1 UMhexagonS 混合搜索模板
Fig. 1 Mixed search patterns of UMhexagonS

(3)采用了半路中止条件。在采用(2)中混合模板搜索的同时,针对不同的宏块分割类型和不同的参考帧,UMhexagonS 可以设置在当前模板搜索结束后,跳过部分搜索模板,直接进行后面的模板搜索。

2.2 UMhexagonS 搜索模板的优化

由上述 UMhexagonS 搜索过程可以看到,混合搜索模板需要搜索的候选点数过多,这就制约了 UMhexagonS 的速度。如果能够采用一定的限制条件缩小搜索范围,那么可以有效地减少搜索点数。由于 UMhexagonS 采用了精确的起点预测,因此可以利用起点运动矢量方向对最佳运动矢量方向所落入的角度范围进行预测,只要预测精确度足够高,那么只需要在该角度范围内搜索就可以找到最佳搜索点,有效地节省搜索点数。

考虑到文献[4]中统计得到的运动矢量在水平和垂直方向的偏置特性,本文提出的改进算法将运动矢量方向划分为偏向正 x 轴方向(标记为 $area_1$),偏向正 y 轴方向(标记为 $area_2$),偏向负 y 轴(标记为 $area_3$),偏向负 x 轴方向(标记为 $area_4$)共 4 个角度范围,如图 2 所示,因此整个运动物体所处平面为:

$$area = area_1 \cup area_2 \cup area_3 \cup area_4, \quad (2)$$

如果将当前宏块运动估计得到的最佳运动矢量记为 $curMV(cur_mvx, cur_mvy)$,则该运动矢量方向的计算方法为

$$\theta = \arctan \frac{cur_mvy}{cur_mvx}, \quad (3)$$

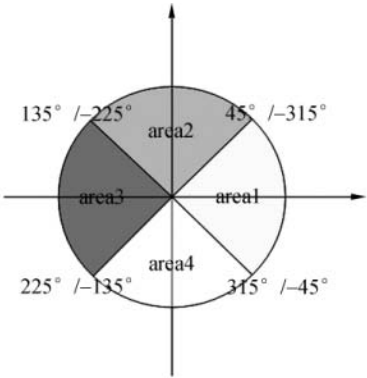


图 2 运动矢量所属 4 个区域
Fig. 2 Four areas of motion vectors

判断该运动矢量落入区域的方法如下:

$$\begin{cases} \text{if } (-45^\circ < \theta < 45^\circ) \\ \quad curMV \in area_1 \\ \text{elseif } (45^\circ < \theta < 135^\circ) \\ \quad curMV \in area_2 \\ \text{elseif } (135^\circ < \theta < 225^\circ) \\ \quad curMV \in area_3 \\ \text{elseif } (225^\circ < \theta < 315^\circ) \\ \quad curMV \in area_4 \end{cases}, \quad (4)$$

为了验证该划分方法的合理性,本文记 MV_{PRED} 和 MV_{BEST} 落入相同区域 $area_1, area_2, area_3, area_4$ 的概率分别为 p_1, p_2, p_3, p_4 ,并令 N_i 表示 MV_{PRED} 和 MV_{BEST} 落入某一相同区域的视频序列宏块数, N_t 表示所统计的宏块总数。那么

$$p_i = N_i / N_t \quad (i = 1, 2, 3, 4), \quad (5)$$

因此 MV_{PRED} 和 MV_{BEST} 落入相同区域的总概率 p_t 计算方式如下:

$$p_t = \sum_{i=1}^4 p_i, \quad (6)$$

这里选择 QCIF 和 CIF 格式的 8 种代表不同运动类型的标准测试序列,在 JM10.2 下采用 baseline 方式分别编码前 100 帧来测试 p_t 的大小,得到的结果如表 1 所示。

表 1 测试得到的 p_t 值

Tab. 1 Test values of p_t

QCIF 测试序列	p_t	CIF 测试序列	p_t
News	0.95	Akiyo	0.98
Foreman	0.96	Foreman	0.97
Mobile	0.96	Bus	0.96
coastguard	0.94	football	0.94
平均概率	0.95	平均概率	0.96

从上表可以看出,采用本文划分方法后, MV_{PRED} 和 MV_{BEST} 落入相同区域的概率对于 QCIF 图像而言,平均达到了 0.95,而对于 CIF 图像,平均概率达到 0.96,因此采用该划分方式准确度很高。

由以上分析可以看出,采用原 UMhexagonS 算法在各个模板给出的所有候选点进行搜索是不必要的。因此,对应于图 2 给出的 4 种不同的运动矢量所属区域,改进算法对于 UMhexagonS 算法的非对称十字模板、大六边形模板进行如图 3(a)、图 3(c)所示的区域划。

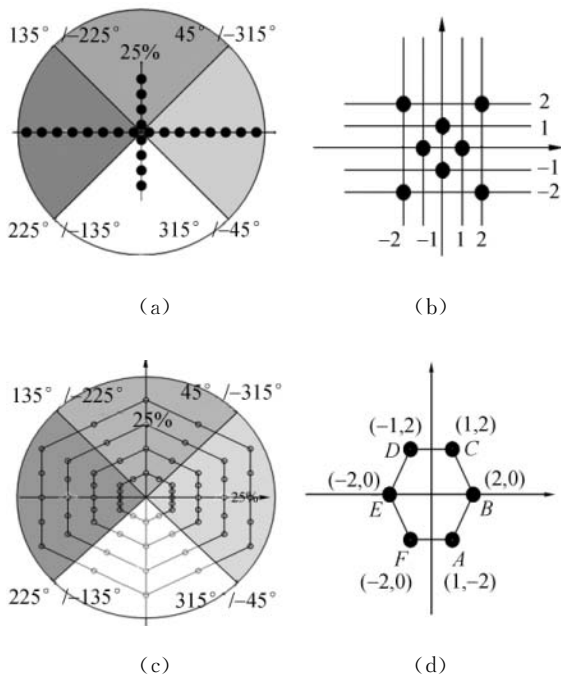


图 3 改进算法的优化模板

Fig. 3 Optimized patterns of improved algorithm

以搜索范围 $\text{search_range} = 16$ 为例,从图 3(a)可以看到,原来算法中需要对整个平面内所有 24 个候选点进行搜索,而优化模板只需要搜索其中 1/4 区域的候选点,也就是 4 个(curMV 落入 area_2 或者 area_4 时)或者 8 个点(curMV 落入 area_1 或者 area_3 时),节约了大量的搜索点,将此优化方法记为 OP1。而对于图 3(c)而言,原来需要搜索 64 个点,优化模板只需要搜索 12 或者 20 个点,将此优化方法记为 OP2。

同时对于图 3(d)中的六边形模板,当 curMV 落入 area_1 时,搜索图 3(d)中 A, B, C 等 3 个点,当 curMV 落入 area_2 时,搜索其中 C, D 等 2 个点,当 curMV 落入 area_3 时,搜索 D, E, F 等 3 个点,当 curMV 落入 area_4 时,搜索 F, A 等 2 个点。相比于原模板对 6 个点进行搜索,改进算法最多搜索 3 个点,将此优化方法记为 OP3。

其次,文献[15]通过统计说明了正方形模板搜索过程中该模板每个点成为最佳点的概率是不同的,因此应该根据统计概率对其进行优化。考虑到该模板在 UMhexagonS 中和小菱形模板一样,都是用于在小范围内搜索,此时运动矢量的方向性不强,所以本文所采用的优化模板如图 3(b)所示,修改后的模板只需要对 8 个点进行搜索,并且依然保持了各个方向上搜索点数的均匀。将此优化方法记为 OP4。

2.3 搜索结束条件

因为加入了精确的起点预测方法,UMhexagonS 对当前宏块进行运动估计时很有可能在起点矢量预测的时候已经找到了最佳运动矢量,但是根据其算法流程,UMhexagonS 至少还需要采用小菱形模板在最佳点周围进行不断搜索,这样就造成了时间浪费,因此改进算法在原算法中起点预测过程完毕后加入一个结束搜索条件,当满足该条件时,直接保存该预测得到的最佳运动矢量为最终运动矢量即可。改进算法的结束搜索条件类似原算法中的半途中止条件,采用如下式(7)进行判断:

$$SAD_{\text{cur_mv}} < SAD_{\text{Pred}}(1 + \beta), \quad (7)$$

其中 $SAD_{\text{cur_mv}}$ 为预测矢量集中得到的最小代价, SAD_{Pred} 的值通过如下方式得到:当参考帧大于 0 时,该值为将预测矢量集中的邻近参考帧预测矢量作为最佳运动矢量所得到的代价。当参考帧为 0 且宏块分割类型为 16×16 时,该值为将预测矢量集中的中值预测矢量作为最佳运动矢量所得到的代价。当参考帧为 0 且宏块分割类型不是 16×16 时,该值为将预测矢量集中的上层预测

表 2 改进算法与原算法的运动估计时间(MET)比较

Tab. 2 Comparison of METs of improved algorithm and UMHexagonS

测试序列	原算法 MET(s)	改进算法 MET(s)	节省率 (%)
news_qcif	64.936	58.582	9.785
foreman_qcif	88.293	77.112	12.664
mobile_qcif	95.255	78.510	17.579
coastguard_qcif	108.711	84.015	22.717
akiyo_cif	320.378	303.593	5.239
foreman_cif	423.628	367.875	13.161
bus_cif	538.372	438.170	18.612
football_cif	623.640	468.118	24.938
平均值			15.59

表 3 改进算法与原算法 Y 分量峰值信噪比(PSNR)比较

Tab. 3 Comparison of PSNRs of improved algorithm and UMHexagonS

测试序列	原算法 PSNR (dB)	改进算法 PSNR(dB)	PSNR 变 化(dB)
news_qcif	36.70	36.72	+0.02
foreman_qcif	36.47	36.48	+0.01
mobile_qcif	33.35	33.36	+0.01
coastguard_qcif	34.04	34.05	+0.01
akiyo_cif	39.77	39.78	+0.01
foreman_cif	36.96	36.95	-0.01
bus_cif	34.77	34.77	0.00
football_cif	36.48	36.47	-0.01

表 4 改进算法与原算法码率比较

Tab. 4 Comparison of bit-rates of improved algorithm and UMHexagonS

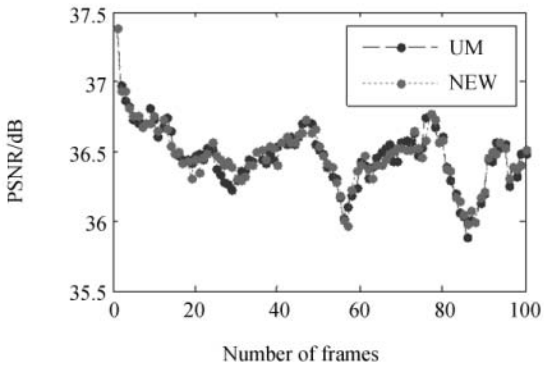
测试序列	原算法码 率(kbit/s)	改进算法码 率(kbit/s)	码率变 化率(%)
news_qcif	75.05	75.48	+0.572
foreman_qcif	135.11	135.08	-0.000
mobile_qcif	425.87	426.23	+0.085
coastguard_qcif	248.70	248.96	+0.104
akiyo_cif	96.26	96.10	-0.167
foreman_cif	403.50	405.37	+0.463
bus_cif	1 143.77	1 165.34	+1.89
football_cif	1 511.27	1 531.02	+1.31

从实验结果可以看出,对于 QCIF 大小的图像,改进算法的 PSNR 值平均提高了 0.01 dB,同时码率增加率最大不超过 1%,运动估计时间

平均减少了 15.69%;而对于 CIF 大小的图像,改进算法的 PSNR 值最大下降 0.01 dB;对于运动较剧烈的 bus 和运动剧烈的 football 测试序列,码率最大增加不超过 2%;其他序列不超过 0.5%,而平均运动估计时间也减少了 15.49%。

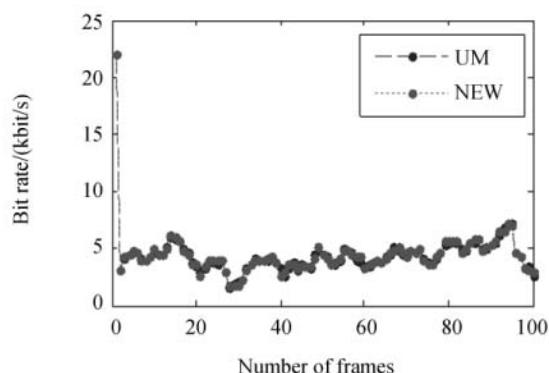
其中,对于运动缓慢的 news 和 akiyo 而言,改进算法速度提高比较少,分别只有 9.785%和 5.239%,这是因为此类视频中存在很多静止和运动缓慢的宏块,这些宏块没有明显的方向性,因此算法中采用的运动矢量方向判断作用比较小,主要是提前结束搜索条件限制了搜索点数。而随着运动剧烈程度的提高,运动矢量方向性就会越明显,改进算法所采用的运动矢量方向判断条件的作用也越大,从表中可以看出,改进算法相比于原算法,运动估计速度的提升也逐渐增大,尤其是对于运动变化剧烈的 football 和 coastguard 而言,改进算法速度提高都在 20%以上。因此,本文适合各种运动类型的视频序列,并且能够在 PSNR 和码率接近 UMHexagonS 的情况下,大幅减少运动估计时间。

图 5(b)和图 5(c)分别是 UMHexagonS 算法(图中 UM)和改进算法(图中 NEW)在 foreman_qcif 测试序列前 100 帧下的 PSNR 和码率比较,而图 5(c)则是每一个编码帧中所有宏块在每一个参考帧下的平均搜索次数。其中,图 5(c)中搜索次数虽然降低了近 50%,但是由于搜索过程中增加了很多条件判断,影响了搜索速度,因此,总的编码时间减少只有 12%。



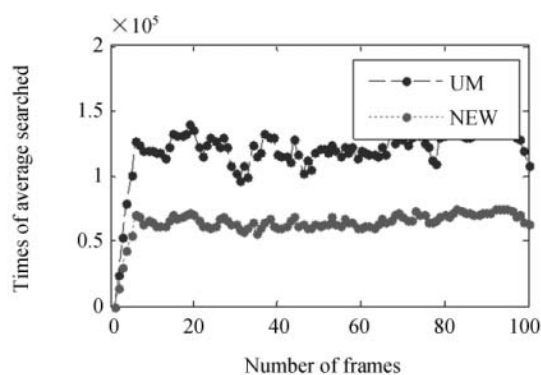
(a) 平均 PSNR 比较

(a) Comparison of average PSNR



(b) 码率比较

(b) comparison of bit rate



(c) 平均搜索次数比较

(c) Comparison of average search time

图 5 改进算法和 UMhexagonS 对于 foreman 序列的搜索性能对比

Fig. 5 Performance comparison of improved algorithm and UMhexagonS on foreman sequences

需要指出的是,对于 CIF 格式图像,在表 3 结果中可以看到,相比于 QCIF 格式图像在原算法基础上 PSNR 的提高,CIF 格式图像的 PSNR 有所减少。这主要是由于结束搜索条件还存在一定的误判造成的,对于 CIF 格式的图像,其宏块数更多,因此误判的概率有所增加,反映在表 4 中,这样的误判就造成了 CIF 格式图像的码率增加率也相比 QCIF 也有所上升。所以,如果能够找到更好的结束搜索条件,运动估计效果会进一步提高。同时,采用更好的区域分割方法,在增加表 1 中 p_i 值的情况下,减少搜索区域,也可以进一步加快运动估计速度。

4 结 论

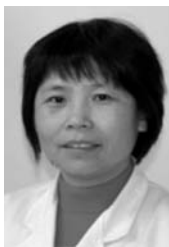
为了提高视频压缩系统速度,本文提出了针对 H. 264/AVC 中所采纳的快速运动估计算法 UMhexagonS 的改进算法。在对 UMhexagonS 进行分析的基础上,通过引入运动矢量方向性判断准则,并且加入提前结束搜索条件,优化了原算法中的混合搜索模板及其搜索策略。实验结果表明,改进算法在 PSNR 和码率与原 UMhexagonS 算法相近的情况下,运动估计时间的平均减少幅度达 15.59%。因此,改进算法比原算法更适合视频图像的快速压缩处理。本文的下一步工作就是应用该改进算法对移植到 DSP 上的 H. 264 压缩代码进行优化,进一步提高压缩速度。

参考文献:

- [1] LI R, ZEN G B, LIOU M L. A new three-step search algorithm for block motion estimation[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1994,(4):438-442.
- [2] ZHU S, MA K K. A new diamond search algorithm for fast block-matching motion estimation[J]. *IEEE Trans. Image Processing*, 2000, 2(9): 2872-2890.
- [3] ZHU C, LIN X, CHAU L P. Hexagon-based search pattern for fast block motion estimation[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2002,12(5):349-355.
- [4] 梁燕,刘文耀. 基于起点预测的自适应交叉-准菱形运动估计算法[J]. *光学精密工程*, 2005,13(2): 237-246.
- [5] LIANG Y, LIU W Y. Adaptive cross-quasi-diamond search algorithm based on initial search point estimation[J]. *Opt. Precision Eng.*, 2005,13(2): 237-246. (in Chinese)
- [6] HOSUR P I, MA K K. Motion vector field adaptive fast motion estimation[C]. *Int. Conf. Information, Communications and Signal Processing (ICICS 1999)*, 1999:1-4.
- [7] TOURAPIS A M, AU O C, LIOU M L. Predictive motion vector field adaptive search technique (PMVFAST) enhancing block-based motion estimation[C]. *Proc. SPIE Conf. Visual Communication and Image Processing*, 2001:883-892.

- [7] 闫敬文,余见,屈小波,等. 优化预测运动矢量的快速运动估计方法[J]. 光学 精密工程,2007,15(10):1622-1627.
YAN J W, YU J, QU X B, *et al.*. Novel fast motion estimation algorithm based on optimizing predictive motion vector [J]. *Opt. Precision Eng.*, 2007,15(10):1622-1627. (in Chinese)
- [8] CHEN Z B, XU J F, HE Y, *et al.*. Fast integer-pel and fractional-pel motion estimation for H. 264/AVC [J]. *Visual Communication and Image Representation*, 2006,17(2):264-290.
- [9] JIA H J, ZHANG L. Directional diamond search pattern for fast block motion estimation[J]. *Electronics Letters*, 2003,39(22):1581-1583.
- [10] 祝世平,申晓东. 十字交叉六边形块运动的估计搜索[J]. 光学 精密工程,2009,12(12):3069-3076.
ZHU SH P, SHEN X D. Cross-hexagon search algorithm for fast block-matching motion estimation [J]. *Opt. Precision Eng.*, 2009,12(12):3069-3076 (in Chinese)
- [11] 廖泰敏,郭宗明. 十字菱形六角形优化搜索算法[J]. 中国图象图形学报,2009,14(8):1530-1533.
LIAO T M, GUO Z M. Enhanced cross-diamond-hexagonal search algorithms [J]. *Image and Graphics*, 2009,14(8):1530-1533. (in Chinese)
- [12] EZHILARASAN M, THAMBIDURAI P. Simplified block matching algorithm for fast motion estimation in video compression[J]. *Journal of Computer Science*, 2008,4(4):282-289.
- [13] UKALKAR A R, BAWANE N G. Analysis of low complexity motion estimation algorithms for H. 264 video compression standard[C]. *Proceedings of the International MultiConference of Engineers and Computer Scientists*, 2009:18-20.
- [14] 向东,陈加忠,周敬利. H. 264 中搜索范围的自适应调整算法[J]. 华中科技大学学报(自然科学版), 2008, 36(5):46-48.
XIANG D, CHEN J ZH, ZHOU J L. An adaptive adjustment algorithm of midium search range for H. 264[J]. *Hua Zhong Univ. of Sci. & Tech. (Natural Science Edition)*, 2008,36(5):46-48. (in Chinese)
- [15] 丁燕,宋雪桦,闫述,等. 基于快速运动估计 UM-hexagonS 算法的改进[J]. 数据采集与处理,2009, 24(5):660-663.
DING Y, SONG X H, YAN SH, *et al.*. Improvements of UMHExagonS algorithm based on fast motion estimation [J]. *Data Acquisition & Processing*, 2009,24(5):660-663. (in Chinese)

作者简介:



李桂菊(1964—),女,吉林人,研究员,1985 年于吉林大学获得学士学位,1988 年于中科院长春光学精密机械与物理研究所获得硕士学位,主要从事数字图像处理及 DSP 应用方面的研究。E-mail:lgjciom666@yahoo.com.cn



梁静秋(1962—),女,吉林长春人,博士,研究员,博士生导师,主要从事微光机电系统及微结构光学研究。E-mail:liangjq@ciomp.ac.cn



刘刚(1986—),男,陕西西安人,硕士研究生,主要研究方向为视频压缩,嵌入式开发等。E-mail:liu2006gang@sina.com