

多层次十字菱形快速整像素搜索算法

曾接贤, 郑大芳, 符 祥

(南昌航空大学软件学院, 南昌 330063)

摘 要: 为提高运动估计速度, 提出一种面向 H.264 的多层次十字菱形快速整像素搜索算法。利用绝对误差和值的空间相关性, 自适应地设置阈值, 运用运动矢量场的时空相关性预测初始搜索点。根据当前最小代价值和阈值的关系, 进行提前终止判断和自适应搜索模板选择, 同时精确定位多层次十字模板。实验结果表明, 该算法在保持搜索精度的同时降低了运动估计时间。

关键词: H.264 标准; 多层次十字菱形; 整像素搜索; 时空相关性; 自适应; 运动估计

Fast Integer Pixel Search Algorithm of Multi-level Cross Diamond

ZENG Jie-xian, ZHENG Da-fang, FU Xiang

(School of Software, Nanchang Hangkong University, Nanchang 330063, China)

【Abstract】 To improve the speed of motion estimation, a fast integer pixel search algorithm for H.264 called Multi-level Cross Diamond Search(MCDS) algorithm is proposed. The algorithm uses the spatial correlation of Sum of Absolute Difference(SAD) value to set thresholds adaptively. It uses the spatial-temporal correlations of the motion vector field to predict the initial search points. And it makes early-termination judgment and selects search templates adaptively according to the relationship of the current minimum cost and the thresholds. It locates multi-cross diamond template precisely adopting the technologies of early-termination judgment and adaptive search templates selection. Experimental results show that the proposed algorithm can save a lot of motion estimation time while maintaining the search accuracy basically compared with the current algorithms.

【Key words】 H.264 standard; multi-level cross diamond; integer pixel search; spatial-temporal correlation; adaptive; motion estimation

DOI: 10.3969/j.issn.1000-3428.2013.04.067

1 概述

运动估计是视频压缩编码中的关键技术之一, 采用运动估计和运动补偿技术可消除视频信号的时间冗余并提高编码效率。H.264/AVC 是最新的国际视频压缩编码标准, 经实验证明, 其运动估计时间占整个 H.264/AVC 编码器编码时间的 60%~80%。因此, 研究高效快速的运动估计算法具有重要意义。

块匹配算法具有简单高效、额外开销小、易于超大规模集成实现等优点, 包括 H.264 在内的绝大多数视频编码标准将其作为运动估计算法。

全搜索法(FS)是最原始的块匹配运动估计算法。FS 虽然简单、可靠, 并且找到的匹配点是全局最优, 但是其计算量较大。为了解决计算量大的问题, 研究人员提出了

很多基于固定模板的快速搜索算法, 如三步法(TSS)、四步法(FSS)^[1]、基于块的梯度下降法(BBGDS)^[2]、菱形法(DS)^[3]和六边形法(HS)^[4]等。

上述算法与 FS 相比, 虽然加快了搜索速度, 但是搜索过程中容易陷入局部最优, 产生错误的运动矢量, 搜索精度和速度都有待提高。利用视频序列运动矢量十字中心偏置特性的十字菱形搜索法(CDS)^[5]、十字菱形六边形搜索法(CDHS)^[6]和新十字菱形搜索法(NCDS)^[7], 采用十字模板和各种方向模板进行搜索, 并使用提前终止判断策略, 进一步提高了搜索速度, 但是没有利用运动矢量场的时空相关性, 搜索速度仍可提高。

最近几年出现了一些基于预测矢量集的快速搜索算法, 如非对称十字型多层次六边形搜索法(UMHS)、改进的预测式区域搜索算法(EPZS)^[8]、H.264 中基于预测方向的多

基金项目: 国家自然科学基金资助项目(61165011); 江西省科技支撑计划基金资助项目(20112BBG70092); 航天科技创新基金资助项目(CASC201102); 航空科学基金资助项目(2010ZC56006)

作者简介: 曾接贤(1958—), 男, 教授, 主研方向: 图像处理, 计算机视觉; 郑大芳, 硕士研究生; 符 祥, 博士

收稿日期: 2012-05-09 **修回日期:** 2012-07-08 **E-mail:** zengjx58@163.com

模板切换运动估计^[9]和基于矢量相关性的自适应运动估计搜索算法^[10]。

这类搜索算法充分利用运动矢量场的时空相关性产生预测矢量集,通过搜索预测矢量位置产生搜索区域的中心位置,采用了混合模板搜索策略,根据运动类型自适应选择搜索模板,并加入了提前终止判断策略,提高了搜索速度。但这类搜索算法仍然复杂和费时。

本文针对上述算法计算复杂度高的问题,提出了一种面向 H.264 的多层次十字菱形快速整像素搜索算法(Multi-level Cross Diamond Search, MCDS)。其主要思想是首先进行中值预测提前终止判断;然后利用上层块模式运动矢量预测方法和时域邻近的参考帧预测方法预测初始搜索点;最后根据当前最小代价值和阈值的关系进行提前终止判断和自适应搜索模板选择。

2 算法描述

搜索策略和搜索模板决定了运动估计搜索算法的计算复杂度和搜索精度。本文提出的 MCDS 算法主要采用了如下 3 种技术:

(1) 自适应阈值技术^[11-12]

根据 SAD 值的空间相关性设置提前终止搜索的判定阈值 $T1$,再根据 $T1$ 设置 $T2$,自适应阈值比固定阈值更合理有效。

(2) 自适应搜索模板选择技术

根据当前最小代价值与阈值的关系判断运动类型,对于不同运动类型采用不同的搜索模板。

(3) 多层次十字模板搜索技术

在多层次十字模板的搜索过程中进行提前终止判断和自适应模板选择,精确定位多层次十字模板,提高搜索速度。

2.1 MCDS 的匹配准则

MCDS 算法是面向 H.264 的搜索算法,H.264 运动估计中采用的匹配准则是拉格朗日代价函数,其表示如下:

$$J(\mathbf{m}, \lambda_{\text{motion}}) = SAD(\mathbf{s}, \mathbf{c}(\mathbf{m})) + \lambda_{\text{motion}} \cdot R(\mathbf{m} - \mathbf{p}) \quad (1)$$

其中, \mathbf{m} 表示候选的运动矢量; \mathbf{p} 表示预测的运动矢量; λ_{motion} 为拉格朗日乘数因子; $SAD(\mathbf{s}, \mathbf{c}(\mathbf{m}))$ 表示原始块和候选块的绝对差值的和; $R(\mathbf{m} - \mathbf{p})$ 表示编码差分运动矢量所需要的比特数; $J(\mathbf{m}, \lambda_{\text{motion}})$ 表示拉格朗日代价值,其值越小,表明率失真越小,预测出来的图像也就越逼近原始图像。

2.2 MCDS 的初始搜索点预测

MCDS 算法利用运动矢量场的时空相关性预测初始搜索点,采用如下 4 种预测方式:

(1) 中值运动矢量预测

设当前块的左邻块运动矢量为 MV_A ,上邻块运动矢量为 MV_B ,右上邻块运动矢量为 MV_C ,则取它们的中值作为预测运动矢量。则:

$$MV_{\text{median}} = \text{median}(MV_A, MV_B, MV_C) \quad (2)$$

其中, MV_{median} 表示中值运动矢量; median 表示取中间值操作。

(2) 上层块模式运动矢量预测

H.264 支持的运动补偿块大小从 16×16 到 4×4 ,分别对应 $Mode1 \sim Mode7$ 。设当前块模式为 $Mode_{\text{curr}}$,上层块模式为 $Mode_{\text{up}}$ 。则:

$$Mode_{\text{up}} = \begin{cases} \text{unavailable} & \text{if } Mode_{\text{curr}} = Mode1 \\ Mode1 & \text{if } Mode_{\text{curr}} = Mode2, Mode3 \\ Mode2 & \text{if } Mode_{\text{curr}} = Mode4 \\ Mode4 & \text{if } Mode_{\text{curr}} = Mode5, Mode6 \\ Mode5 & \text{if } Mode_{\text{curr}} = Mode7 \end{cases} \quad (3)$$

由于运动物体具有空间完整性,因此可用上层块模式的运动矢量作为当前块的预测运动矢量。

(3) 时域邻近的参考帧预测

由于运动矢量具有连续性,因此当前块相对于不同参考帧的运动矢量也存在一定的线性关系。设当前帧对应时刻为 t ,相邻参考帧对应时刻分别为 t' 和 $t'+1$,相对于时刻 t' 和 $t'+1$ 的参考帧的运动矢量分别为 $MV_{\text{pred_NRP}}$ 和 MV_{NRP} ,则:

$$MV_{\text{pred_NRP}} = MV_{\text{NRP}} \times \frac{t - t'}{t - t' - 1} \quad (4)$$

(4) 零值矢量预测

由于视频序列中有很多静止块或小运动块,因此可以用零值矢量作为当前块的预测运动矢量,即:

$$MV_{\text{zp}} = (0, 0) \quad (5)$$

2.3 MCDS 的搜索模板

UMHS 算法由于其搜索精度较高,搜索时间较短,而被 H.264 参考软件 JM10.1 作为一种快速运动估计算法。UMHS 算法中使用了一种非对称十字型模板,主要是依据现实世界中很多运动序列的水平方向的运动多于垂直方向的运动,以及运动矢量的中心偏置特性设计的。然而采用这种非对称搜索模板的方法,在一些具有大规模规则运动的宏块上容易陷入局部最优。因此,本文构造 3 种对称的搜索模板:

(1) 小菱形模板,如图 1 所示,其搜索点数少,易收敛,可用于检测小范围的运动矢量。

(2) 大菱形模板,如图 2 所示,包含水平、垂直和对角方向上的 8 个搜索点,搜索点在各个方向上分布均匀,可用于检测中等范围的运动矢量。

(3) 多层次十字模板,如图 3 所示,由 8 个不同尺寸且尺寸大小为偶数的十字模板组成,可用于检测大范围的运动矢量。搜索多层次十字模板时,按其中的十字模板的尺寸从大到小依次搜索,搜索过程中进行提前终止判断和自适应搜索模板选择,这种搜索方式可减少不必要的搜

索点。

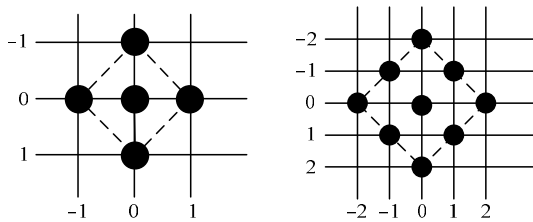


图1 小菱形模板

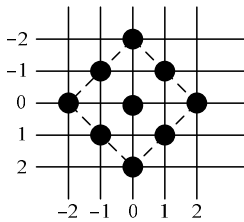


图2 大菱形模板

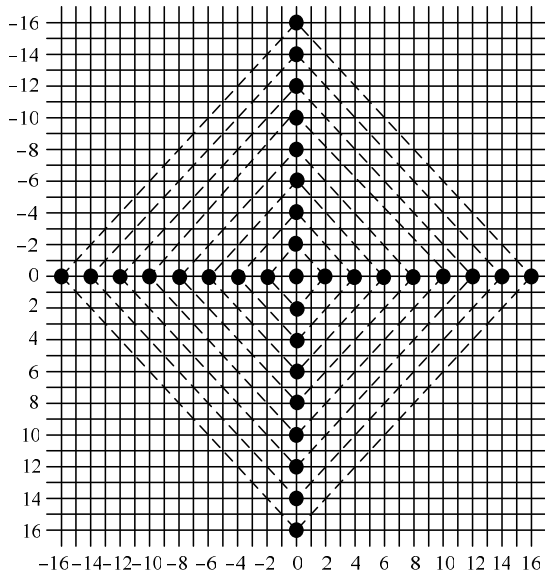


图3 多层次十字模板

2.4 MCDS 的自适应阈值

H.264 支持 7 种不同尺寸的块,在进行提前终止搜索和运动类型判定时,对于不同尺寸的块采用相同的阈值是不合理的,因此,MCDS 算法采用自适应阈值。因为视频序列的运动矢量场和 SAD 值具有时空相关性,所以 MCDS 算法利用相邻块的 SAD 值来设置对于当前块提前终止搜索的判定阈值。首先根据式(6)设置阈值 T_1 ,然后根据式(7)设置阈值 T_2 , T_1 的范围控制在[256, 500]; T_2 的范围控制在[500, 800]。

$$T_1 = \min(SAD_A, SAD_B, SAD_C) \quad (6)$$

$$T_2 = f(T_1) = T_1 + 244 \quad (7)$$

其中, SAD_A 为当前块的左邻块的 SAD 值; SAD_B 为当前块的上邻块的 SAD 值; SAD_C 为当前块的右上邻块的 SAD 值; \min 代表取最小值。

2.5 MCDS 的多层次十字模板搜索判断

实验表明,当前矢量和中值矢量、上层块模式矢量的相关性高于当前矢量和前一帧相应位置的矢量、时域邻近的参考帧预测矢量的相关性。这一结果也表明了空间预测是更好的选择,尤其是上层块模式预测在 H.264 中效果更好^[13]。

因此,MCDS 算法通过判断是否存在上层块来决定是否进入多层次十字模板搜索。如果当前块的大小是 16×16 ,则进行多层次十字模板搜索,否则搜索上层块模式矢量对

应的位置。

2.6 MCDS 的多层次十字模板搜索

搜索多层次十字模板时,按其中十字模板的尺寸从大到小依次搜索,搜索过程中进行提前终止判断和自适应搜索模板选择,即当前最小代价值小于或等于 T_1 时,结束搜索;当前最小代价值大于 T_1 小于或等于 T_2 时,进入小菱形搜索;否则继续多层次十字模板的搜索。

具体步骤如下:

Step1 将经过初始搜索点预测后的具有最小代价值的点作为搜索模板的中心,依次进行尺寸大小为 $2i(i = 8, 7, \dots, 1)$ 的十字模板搜索。在搜索每个点时,进行提前终止判断和自适应搜索模板选择,即当前最小代价值小于或等于 T_1 时,结束搜索,输出运动矢量;当前最小代价值大于 T_1 小于或等于 T_2 时,转入 Step2;否则继续搜索该尺寸的十字模板的下一个点,当一个十字模板搜索完毕时,继续搜索尺寸大小为 $2(i-1)$ 的十字模板。

Step2 小菱形搜索。将当前具有最小代价值的点作为搜索模板的中心,进行小菱形搜索,如果当前最小代价值的对应位置为搜索模板的中心,则结束搜索,输出运动矢量;否则重复执行 Step2。

3 算法总体流程

设预测矢量集 $= \{MV_{zp}, MV_{pred_NRP}\}$,则多层次十字菱形搜索算法的流程如图 4 所示。

算法各步骤描述如下:

Step1 中值预测提前终止判断。采用式(2)进行中值运动矢量预测,搜索其相应位置,并采用式(1)计算拉格朗日代价值,如果代价值小于或等于 T_1 ,则结束搜索,输出运动矢量。否则转入 Step2。

Step2 采用式(5)进行零值矢量预测,采用式(4)进行时域邻近的参考帧矢量预测,并搜索其相应位置,得出当前最小的拉格朗日代价值。如果代价值小于或等于 T_1 ,则结束搜索,输出运动矢量;如果代价值大于 T_2 ,转入 Step3;否则,转入 Step5。

Step3 通过判断是否存在上层块决定是否进入多层次十字模板搜索。其判断方法是:如果当前块的块大小为 16×16 ,转入多层次十字模板搜索;否则,采用式(3)进行上层块模式运动矢量预测,并搜索其相应位置。

Step4 大菱形搜索。将当前具有最小代价值的点作为搜索模板的中心,进行大菱形搜索,如果当前最小代价值的对应位置为搜索模板的中心,则转入 Step5;否则,重复执行 Step4。

Step5 小菱形搜索。将当前具有最小代价值的点作为搜索模板的中心,进行小菱形搜索,如果当前最小代价值的对应位置为搜索模板的中心,则结束搜索,输出运动矢量;否则重复执行 Step5。

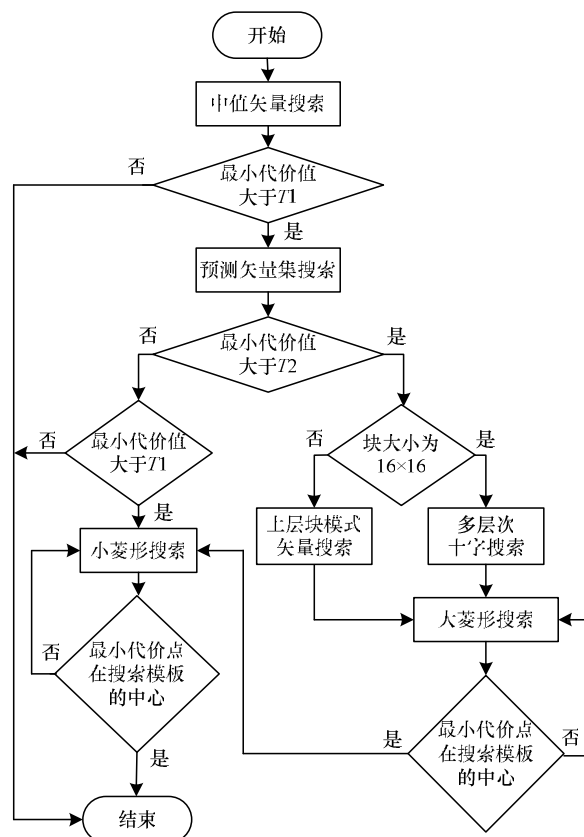


图 4 MCDS 算法总体流程

4 实验结果及分析

为了验证算法的有效性, 本文将 MCDS 算法和 FS 算法、UMHS 算法、SUMHS 算法、EPZS 算法、NCDS 算法

进行了比较, 并从重构帧的峰值信噪比和总运动估计时间分析算法的有效性。实验基于 H.264/AVC 参考模型和 JM10.1 平台进行。

实验硬件环境: Pentium(R) Dual-Core CPU 2.6 GHz, 2 GB RAM; 编码环境如下: 搜索范围是[-16, 16], 5 帧作为参考帧, 序列类型是 IPPP, YUV 的格式是 4:2:0, 帧率是 30 fps, 量化参数为 28。实验选取 6 个运动程度不同的 QCIF 格式的图像测试序列: mobile, foreman, soccer, news, mother-daughter, akiyo, 对这些序列中的前 100 帧进行视频编码。

表 1 所示是 MCDS 算法和其他算法的总运动估计时间比较表。从表 1 可以看出, MCDS 算法相对于其他算法, 总运动估计时间有了大幅度的下降。与 FS 算法相比, 总运动估计时间大约下降了 90%; 与 UMHS 算法相比, 总运动估计时间下降了 40%~70%; 与 SUMHS 算法相比, 总运动估计时间下降了 20%~55%; 与 EPZS 算法相比, 总运动估计时间下降了 40%~55%; 与 NCDS 算法相比, 总运动估计时间下降了 25%~55%。

表 2 所示是 MCDS 算法和其他算法的峰值信噪比比较表。从表 2 可以看出, MCDS 算法对不同类型的测试序列有很强的适应性。和 FS 算法相比, 峰值信噪比最高下降了 0.06 dB, 平均下降了 0.03 dB。和其他算法相比, 峰值信噪比也相差不大, 特别是相对于 SUMHS 算法, 平均提高了 0.01 dB。当测试序列为 soccer 时, MCDS 算法的峰值信噪比比 FS 的峰值信噪比还要高 0.01 dB, 文献[14]说明了 FS 算法的峰值信噪比不一定最高的原因。

表 1 MCDS 算法和其他算法的总运动估计时间比较

测试序列	总运动估计时间/s						变化量/(%)				
	FS	UMHS	SUMHS	EPZS	NCDS	MCDS	ΔFS	ΔUMHS	ΔSUMHS	ΔEPZS	ΔNCDS
mobile	110.659	26.763	18.318	18.295	12.504	9.167	-91.72	-65.75	-49.96	-49.89	-26.69
foreman	112.955	30.383	21.139	21.536	22.421	11.262	-90.03	-62.93	-46.72	-47.71	-49.77
soccer	116.410	38.605	26.784	24.363	26.361	13.354	-88.53	-65.41	-50.14	-45.19	-49.34
news	110.433	12.233	9.088	12.045	11.435	5.878	-94.68	-51.95	-35.32	-51.20	-48.60
akiyo	111.132	9.494	7.283	10.705	11.506	5.584	-94.98	-41.18	-23.33	-47.84	-51.47
mother-daughter	111.237	12.846	10.154	12.401	12.632	6.441	-94.21	-49.86	-36.57	-48.06	-49.01

表 2 MCDS 算法和其他算法的峰值信噪比比较

测试序列	峰值信噪比						变化量				
	FS	UMHS	SUMHS	EPZS	NCDS	MCDS	ΔFS	ΔUMHS	ΔSUMHS	ΔEPZS	ΔNCDS
mobile	32.91	32.89	32.88	32.89	32.91	32.88	-0.03	-0.01	0.00	-0.01	-0.03
foreman	35.77	35.74	35.71	35.74	35.75	35.75	-0.02	0.01	0.04	0.01	0.00
soccer	35.34	35.32	35.33	35.31	35.35	35.35	0.01	0.03	0.02	0.04	0.00
news	36.22	36.19	36.14	36.19	36.23	36.16	-0.06	-0.03	0.02	-0.03	-0.07
akiyo	37.68	37.66	37.65	37.66	37.69	37.64	-0.04	-0.02	-0.01	-0.02	-0.05
mother-daughter	36.77	36.77	36.74	36.76	36.75	36.75	-0.02	-0.02	0.01	-0.01	0.00

综上所述, MCDS 算法与 FS 算法、UMHS 算法、SUMHS 算法、EPZS 算法、NCDS 算法相比, 在基本保证了搜索精度的同时, 大幅减少了总运动估计时间。

5 结束语

本文提出了一种面向 H.264 的多层次十字菱形快速整像素搜索算法。利用 SAD 的空间相关性自适应设置阈值,

进行中值预测提前终止判断, 并利用上层块模式运动矢量预测方法和时域邻近的参考帧预测方法预测初始搜索点, 然后根据当前最小代价值和阈值的关系进行提前终止判断和自适应搜索模板选择, 并判断其上层块是否存在来决定是否进入多层次十字模板搜索, 最后结合大菱形和小菱形模板搜索以找到最佳匹配点。实验结果表明, 与现有算法相比, MCDS 算法保证了搜索精度, 同时大幅减少了运动估计时间。

参考文献

- [1] Po L M, Ma W C. A Novel Four-step Search Algorithm for Fast Block Motion Estimation[J]. IEEE Trans. on Circuits and Systems for Video Technology, 1996, 6(3): 313-317.
- [2] Liu L K, Feig E. A Block-based Gradient Descent Search Algorithm for Block Motion Estimation in Video Coding[J]. IEEE Trans. on Circuits and Systems for Video Technology, 1996, 6(4): 419-422.
- [3] Zhu Shan, Ma Kaikuang. A New Diamond Search Algorithm for Fast Block-matching Motion Estimation[J]. IEEE Trans. on Image Processing, 2000, 9(2): 287-290.
- [4] Zhu Ce, Lin Xiao, Chau L P. Hexagon-based Search Pattern for Fast Block Motion Estimation[J]. IEEE Trans. on Circuits and Systems for Video Technology, 2002, 12(5): 349-355.
- [5] Cheung C H, Po L M. A Novel Cross-diamond Search Algorithm for Fast Block Motion Estimation[J]. IEEE Trans. on Circuits and Systems for Video Technology, 2002, 12(12): 1168-1177.
- [6] Cheung C H, Po L M. Novel Cross-diamond-hexagonal Search Algorithms for Fast Block Motion Estimation[J]. IEEE Transactions on Multimedia, 2005, 7(1): 16-22.
- [7] Zhu Shiping, Tian Jun, Shen Xiaodong, et al. A New Cross-diamond Search Algorithm for Fast Block Motion Estimation[C]//Proc. of the 16th International Conference on Image Processing. [S. l.]: IEEE Press, 2009: 1581-1584.
- [8] Tourapis A M. Enhanced Predictive Zonal Search for Single and Multiple Frame Motion Estimation[C]//Proc. of Conference on Visual Communications and Image Processing. San Jose, USA: SPIE Press, 2002: 1069-1079.
- [9] 刘晓明, 高 燕, 李毓慧, 等. H.264 中基于预测方向的多模板切换运动估计[J]. 计算机工程, 2010, 36(15): 214-215.
- [10] 沈淦力, 王维东, 郭 斌, 等. 基于矢量相关性的自适应运动估计搜索算法[J]. 中国图象图形学报, 2008, 13(9): 1689-1693.
- [11] Qiu Xiaobin, Huang Chunqing. An Improved Algorithm of Fast Motion Estimation Based on H.264[C]//Proc. of the 5th International Conference on Computer Science and Education. [S. l.]: IEEE Press, 2010: 1717-1721.
- [12] Cui Zhaopeng, Wang Dajuan, Jiang Guang, et al. Octagonal Search Algorithm with Early Termination for Fast Motion Estimation on H.264[C]//Proc. of the 5th International Conference on Information Assurance and Security. [S. l.]: IEEE Press, 2009: 123-126.
- [13] Xu Jianfeng, Chen Zhibo, He Yun. Efficient Fast ME Predictions and Early-termination Strategy Based on H.264 Statistical Characters[C]//Proc. of the 4th International Conference on Information, Communications and Signal Processing. [S. l.]: IEEE Press, 2003: 218-222.
- [14] Chen Zujue, Cao Lie, Liu Shaoqing. A Fast Motion Estimation Optimization for H.264/AVC Encoder[C]//Proc. of International Conference on Wireless Communications, Networking and Information Security. [S. l.]: IEEE Press, 2010: 116-119.

编辑 顾逸斐

(上接第 290 页)

- [4] Weinhardt M, Luk W. Memory Access Optimization for Reconfigurable Systems[J]. Computers and Digital Techniques, 2001, 148(3): 105-114.
- [5] Lin E C, Yu Kai, Rutenbar R A, et al. A 1 000-word Vocabulary, Speaker-independent, Continuous Live-mode Speech Recognizer[C]//Proc. of the 15th International Symposium on Field Programmable Gate Arrays. New York, USA: IEEE Press, 2007: 60-68.
- [6] Choi Y K, You K, Choi J. A Real-time FPGA-based 20 000-word Speech Recognizer with Optimized DRAM Access[J]. IEEE Trans. on Circuits and Systems, 2010, 57(8): 2119-2131.
- [7] Wang Shaojun, Peng Yu, Zhao Guangquan. Accelerating On-line Training of LS-SVM with Run-time Reconfiguration[C]//Proc. of the International Conference on Field Programmable Technology. New Delhi, India: [s. n.], 2011.
- [8] 叶肇晋, 张稀楠, 马 磊. 基于 XILINX FPGA 片上嵌入式系统的用户 IP 开发[M]. 西安: 西安电子科技大学出版社, 2008.
- [9] 郭 磊. 矩阵运算的硬件加速技术研究[D]. 长沙: 国防科学技术大学, 2010.
- [10] 商 磊. 大规模线性方程组求解的并行算法及应用[D]. 西安: 西北工业大学, 2007.

编辑 顾逸斐

