

一种稳健的图割立体匹配方法

马东东,杨 靛,刘向增

(西安微电子技术研究所,陕西 西安 710054)

摘 要: 为了提高传统的基于图割的立体匹配算法的鲁棒性,提出一种基于图像增强的图割立体匹配方法,传统算法是在图像灰度值的基础上构建能量函数,该方法加入了图像的梯度值来构建能量函数,然后将基于二值化标号函数的 α 扩展算法和 KV 最大流算法结合起来求解能量函数. 由于梯度对于图像中的噪声、局部光照具有鲁棒性,因此该算法在一定程度上增加了传统 GC 算法的鲁棒性,最后通过仿真和实验验证了算法的有效性.

关键词: 梯度;鲁棒性;能量函数;图割;立体匹配

中图分类号: TP391

文献标识码: A

文章编号: 1000—7180(2016)04—0027—05

A Robust Method Based on Graph Cuts for Stereo Correspondence

MA Dong-dong, YANG Liang, LIU Xiang-zeng

(Xi'an Microelectronics Technology Institute, Xi'an 710054, China)

Abstract: To improve the robustness of the algorithm based on graph cuts, a new method based on graph cuts is introduced in this paper. We formulated the energy functions based on gradient. Then the method was applied to minimize the energy function. As for the robustness of the gradient for the image noise and local illumination, the method increase the robustness of the traditional algorithm based on graph cuts. At last, the effectiveness of the algorithm is verified by simulation and implementation.

Key words: gradient; robustness; energy function; graph cuts; stereo correspondence

1 引言

立体匹配技术作为视觉导航和三维重建等领域的关键技术,得到众多学者的广泛研究,其基本原理是从两幅或多幅图像中获得同一景物的立体像对,匹配出相应的像点,从而能将二维上的距离信息转化为空间的深度信息.

立体匹配算法有基于局部约束和基于全局的两种匹配算法^[1]. 局部匹配算法主要包括匹配代价的选取、匹配代价的统计、视差的优化这几个过程;全局匹配算法核心环节包括能量函数的构造和能量函数的优化,能量函数的优化常用的方法有梯度递减法、模拟退火法和基于图割算法. 梯度递减法几乎能用于所有的连续变量的能量函数,但会出现局部极小值. 模拟退火法几乎能用于所有的离散变量的能

量函数,虽能找到全局最小值,但需要很长的收敛时间. 基于图割的优化算法是一种很有效的能量函数的优化算法,在特定的条件下可以得到全局的最小值.

传统的立体匹配是基于图像的灰度值进行匹配的,图像的灰度易受到噪声、光照、采样等外界条件的影响,对此本文提出了一种增强型的匹配算法,即 IGGC(Intensity and Gradient Graph Cuts)算法,将图像的梯度值(G)和灰度值(I)结合起来定义了图像的像素值,记为图像的 IG 值,在图像的 IG 值的基础上构建能量函数,利用图割算法求解能量函数,得到图像的视差图. 综上所述,IGGC 算法分为两个步骤:一是构建能量函数模型;二是求解模型.

2 模型的构建

IGGC 算法是基于图像 IG 值,将视差问题转化为标号问题,然后构建标号函数的能量函数。

2.1 IG 值的求取

图像的灰度值(I),直接表示了该像素点的亮度大小,是描述彩色世界的主要参数,对于一个三通道的彩色图像,

$$I = rI_r + gI_g + bI_b \quad (r + g + b) = 1 \quad (1)$$

式中, r, g, b 分别为各颜色通道的权重, I_r, I_g, I_b 为各颜色通道的强度。

图像的梯度,就是图像灰度值变化显著的地方,梯度能从一定程度上体现图像的轮廓和边缘信息。梯度包括两个组成部分,一是梯度的幅值,另外一个梯度的相角,为了简化算法,在这里不考虑梯度的相角,用 $f(x, y)$ 表示一幅灰度图像,则梯度定义为

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix}, \text{则梯度用式(2)表示,为了简化计算,}$$

用式(3)近似表示:

$$G = \sqrt{G_x^2 + G_y^2} \quad (2)$$

$$G = |G_x| + |G_y| \quad (3)$$

将图像的灰度值和梯度值结合起来作为图像的像素的强度值,记为图像的 IG 值,图像的 IG 值的求取如下式:

$$IG = \alpha G + (1 - \alpha)I \quad (4)$$

式中, α 为梯度模的权重,调节参数 α 可以改变算法的抗噪性。

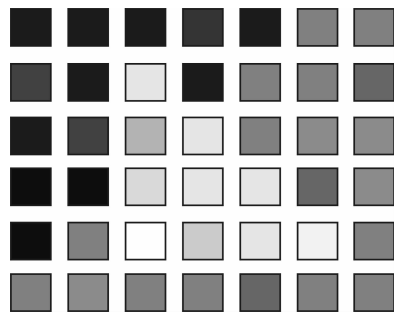
2.2 标号问题

图像处理领域的很多问题都可以看作是标号问题,如图 1 所示,设 P 为一个像素集,对于 P 中的任意一个像素 $p \in P$,都有固定的灰度值 I_p ,如图 1(a),设 f 为标号函数,对应的 $f_p = f(p) \in L$ 为像素 p 的标号,其中 L 为标号集,在图 1(b)中,将 $f_p \in \{0, 1, 2\}$ 分配每一个像素 $p \in P$,在图像分割问题中,标号代表着分割区域的编号,在图像恢复问题中,标号代表着图像的原始灰度,在立体匹配问题中,标号代表着匹配后的视差。

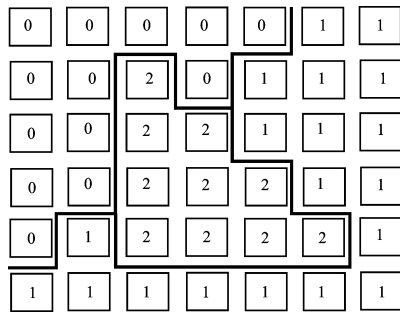
将标号问题应用到立体匹配问题中,对于立体匹配中的两幅图像分别记为图 L 和图 R,设 P 为图 L 的像素集,匹配的目的就是为每一个像素 $p \in P$ 分配一个标号 f_p ,这个标号代表像素 p 的视差,使得图 L 中的像素 p 与图 R 中的像素 $p + f_p$ 相匹配。

2.3 构建能量函数

为了求解标号函数 f ,需要做出两个假设:



(a) 灰度图



(b) 标号图

图 1 图像与标号示意图

(1) 图 L 中的像素 p 与图 R 的像素 $p + f_p$ 有相同的像素强度;

(2) 视差图是光滑的,像素 p 和其邻域中的像素 q 有相似的视差值。

对于这两个假设,可以构建如下的能量函数

$$E(f) = E_{\text{data}} + E_{\text{smooth}} \quad (5)$$

$$E_{\text{data}} = \sum_{p \in P} D_p(f), E_{\text{smooth}} = \sum_{(p, q) \in N} V_{\{p, q\}}(f_p, f_q)$$

式中, $D_p(f)$ 称为数据项,作为 f 与观测数据一致性的惩罚, $D_p(f)$ 的求取如下式所示:

$$D_p(f) = (IG(p) - IG(p + f_p))^2 \quad (6)$$

式中的 IG 为上文求取的图像的 IG 值

$V(f_p, f_q)$ 为光滑项,对应于第二个假设,体现的是区域内部的连续性和边界的不连续性,经典的光滑项为用 Potts 模型,如式(7)所示

$$V(f_p, f_q) = \lambda \cdot T \quad (f_p \neq f_q) \quad (7)$$

当参数为真时 T 为 1,否则为 0。

为了增强算法的鲁棒性,在光滑项中引入了截断值,构造后的光滑项如下式

$$V(f_p, f_q) = \min(|f_p - f_q|, R) \quad R \text{ 为截断值} \quad (8)$$

在图像的匹配中,会出现遮挡问题,其中一幅图中的某一点在另一幅图中没有相应的像点,因此,能量函数需加入遮挡惩罚项。结合文献[1],本文构造

的遮挡项的能量如下式:

$$E_{\text{occ}} = \sum_{p \in M} \lambda_{\text{occ}} \quad (9)$$

式中, M 为遮挡像素集, 由在单幅图像上的可视像素组成, λ_{occ} 单个像素状态为遮挡时的代价, 本文设为 100.

综上所述, 最终构造的能量函数为:

$$E(f) = E_{\text{data}} + E_{\text{smooth}} + E_{\text{occ}} \quad (10)$$

3 模型的求解

能量函数 $E(f)$ 的求解是 NP 难问题, 若采用穷举法, 计算量过大, 状态空间为 $L^{H \times W}$, 其中, L 为视差范围, H 为图像的高度, W 为图像的宽度, 因此, 对于能量函数的求解有很多优化算法, 如梯度递减算法、模拟退火算法和基于图割的算法, 本文采用基于图割的优化算法, 基于图割的优化算法是利用图论的知识, 将能量函数构造成图, 能量函数的最小值对应于所构建的图中的最小割, 再利用最大流最小割定理, 将求解最小割的问题转化为求解最大流的问题, 求解出的最大流对应的就是能量函数的最小值, 最小能量函数 $E_{\min}(f)$ 所对应的标号函数 $f = \arg(E_{\min}(f))$ 为我们所求的最终视差标号. 综上所述, 利用图割算法优化能量函数主要分为两个步骤: 一是利用能量函数构建合适的图; 二是将图中的最小割转化为最大流, 利用最大流算法来求解图中的最小割. 在求解能量函数过程中, 本文将 α 扩展算法和文献[2]中的最大流算法结合起来, 应用在本文的 IGGC 算法中, 这是本文的创新点.

3.1 图的构建

采用条件迭代算法(ICM算法), ICM 中的标准移动方法常会出现局部最小值, 文献[3]提出了一种 α 扩展算法, 在特定的条件下可以得到全局最小值. 算法描述如下:

- (1) 以任意标号 f 开始
- (2) 设置标志 $\text{flag} = 0$, 对于每一个标号 $\alpha \in L$
 - (2.1) 通过 α 扩展, 在 f' 中寻找 $\hat{f} = \arg(E_{\min}(f'))$
 - (2.2) 如果 $E(f') < E(f)$, 使 $f = f'$, 并设置 $\text{flag} = 1$;
- (3) 如果 $\text{flag} = 1$ 转到(2);
- (4) 返回 f ;

假设现有一个标号函数 f 和某一特定的标号 α , 经过一次 α 扩展, 标号函数 f 变为 f' , f' 中的标号不变或变为 α , 即 $f' = f$ 或 $f' = \alpha$, 由此可见, 一次 α 扩展的过程可以看做将标号函数编码为一个二进制矢

量 $\mathbf{x} = \{x_p \mid p \in P\}$, 若 $f' = f$, x_p 为 0; 若 $f' = \alpha$, x_p 为 1. 在文献[4]将这种被二值化矢量定义的标号函数记为 f^x , 这样, 就可以将一个非二值化的问题转化为一个二值化的问题, 提升了算法的效率.

构造图 $G = (V, \xi)$, 图的源点为 s , 汇点为 t , 其中, $V = \{v_1, \dots, v_n, s, t\}$, ξ 为图中的边, 在图 G 上的每一个割都有一定的费用, 任意一个割都可以形成二值矢量 $\mathbf{x}_1, \dots, \mathbf{x}_n$ 与图中节点(除过源点和汇点)的对应关系, 若 $v_i \in S$, 则 $x_i = 0$, 若 $v_i \in T$, 则 $x_i = 1$, 二值能量函数 $E(x_1, \dots, x_n)$ 与图中割的费用是相等的.

上文构建的能量函数, 二值化后的能量函数如下:

$$E(x_1, \dots, x_n) = \sum_i E^i(x_i) + \sum_{i < j} E^{i,j}(x_i, x_j) \quad (11)$$

上式在特定的条件下会出现全局最小值, 为了寻找限制条件, 文献[4]给出了求解无约束非线性优化问题 $\min f(x)$ 的整体最优解的定理

定理 1 如果函数 $f(x)$ 是 R^n 上的可微凸函数, 若 $\nabla f(x^*) = 0$, 则 x^* 是 $\min f(x)$ 的整体最优解.

当 $E(x_1, \dots, x_n)$ 满足式就可以满足可微凸函数的条件, 能量函数就可以出现全局最优解^[5].

$$E^{i,j}(0,0) + E^{i,j}(1,1) \leq E^{i,j}(0,1) + E^{i,j}(1,0) \quad (12)$$

对于一个依赖变量 x_i 的项 E^i , 分为以下两种情况:

(1) 当 $E^i(0) < E^i(1)$, 增加边 (s, v_i) , 边的权值为 $E^i(1) - E^i(0)$ 如图 2(a) 所示;

(2) 当 $E^i(0) > E^i(1)$, 增加边 (v_i, t) , 边的权值为 $E^i(0) - E^i(1)$ 如图 2(b) 所示.

对于两个变量 x_i 和 x_j 的函数 $E^{i,j}$, 设 $E^{i,j}(0,0) = A$, $E^{i,j}(0,1) = B$, $E^{i,j}(1,0) = C$, $E^{i,j}(1,1) = D$, $E^{i,j}$ 可以表示成如下的形式:

$E^{i,j}(0,0) = A$	$E^{i,j}(0,1) = B$
$E^{i,j}(1,0) = C$	$E^{i,j}(1,1) = D$

为了方便图网络的构造, 将 $E^{i,j}$ 改写成如下的形式:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = A + \begin{bmatrix} 0 & 0 \\ C-A & C-A \end{bmatrix} + \begin{bmatrix} 0 & D-C \\ 0 & D-C \end{bmatrix} +$$

0	$B+C-A-D$
0	0

式中, 第一项 A 为一个常数, 不需要增加额外的边, 第二项和第三项分别依赖于一个变量 x_i 和 x_j , 为了表示最后一项, 增减了边 (v_i, v_j) , 边的权重为 $B+A-C-D$, 由式(12), 可知权值非负, 如图 2(c) 所示, 若 $C-A > 0$ 且 $D-C < 0$, 则构成如图 2(d) 所示的完整的图。

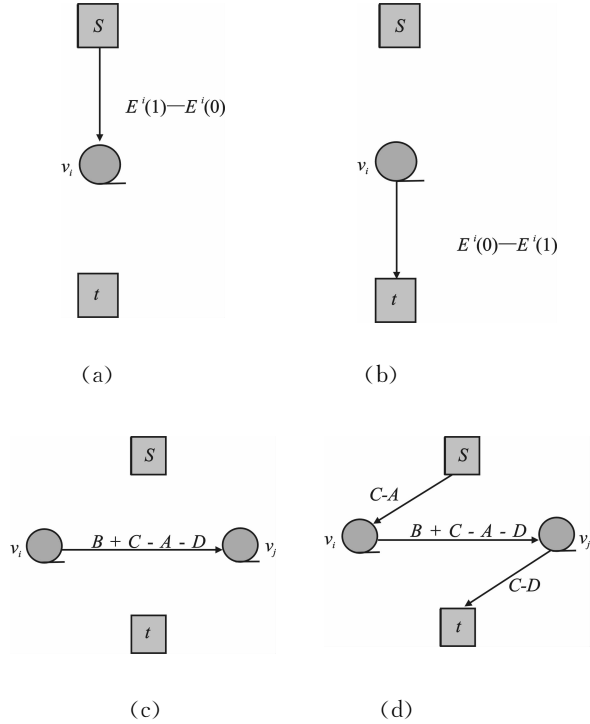


图 2 $E^{i,j}$ 构建的能量图

3.2 最大流算法

本文采用文献[2]介绍的最大流算法, 该算法的运行速度远高于传统的 Fort-Fulkerson 算法和 Dinic 算法, 它是一种改进的增广路算法, 该算法建立了两个互不重合的搜索树, 一个是以源点为根的 S 树, 另一条是以汇点为根的 T 树, S 树中从父节点到子节点是不饱和的, T 树中从子节点到父节点是不饱和的. 不属于 S 树和 T 树中的节点称为自由节点, 搜索树 S 和 T 中的节点或是活跃节点或是非活跃节点, 活跃节点代表了搜索树的边缘, 非活跃节点则在搜索树的内部, 活跃节点允许搜索树向自由节点集中生长, 非活跃节点由于同意树中其他节点的阻止而不能生长, 活跃节点可能会和其他树的节点相接触, 当一个树的活跃节点检测到与它相邻的节点属于另一个树时, 便是找到了一个增广路径。

假设图 $G = (V, \xi)$, 对于任意一个增广路径, 它

的流为 f , 残余图为 G_f , 可以得到活跃节点集 A 和所有的孤点集 O , 该算法可以描述如下:

- (1) 初始化 $S = \{s\}, T = \{t\}, A = \{s, t\}, O = \phi$;
- (2) 当为真, 循环
 - “生长” S 和 T 找到一条从 s 到 t 的增广路径 P ;
 - 如果 $P = \phi$, 则终止;
 - “增广” P ;
 - “吸收”孤点;
 - 直到结束.

其中, 在“生长”阶段, 活跃节点检测与之相邻的非饱和边, 从自由节点集中接受新的子节点, 于是新获得的子节点成为与之相关的搜索树的自由节点, 当给定的活跃节点的所有邻边被检测完后, 该活跃节点变为非活跃节点, 当一个活跃节点检测到其邻近节点是属于另一个搜索树时, 生长阶段终止; “增广”阶段对在生长阶段找到的路径进行增广. 将最大可能的流从源点沿路径推向汇点, 这样路径中的一些边会变成饱和的. 从而, 在 S 和 T 中的一些节点会变成孤点, 实际上, 增广阶段能把树 S 和 T 变成森林, 源点 s 和汇点 t 仍然是两个树的根, 而孤点成为了其它搜索树的根; “吸收”阶段的目的是重构以源点和汇点为根的单树结构 S 和 T . 在这个阶段我们试图为每个孤儿找到一个新的有效的父节点, 新的父节点和孤点属于同一集合. 如果找不到合适的父节点, 我们将孤儿从 S 或 T 中移除并使其成为自由节点, 孤点的所有前子节点都将成为孤点. 当没有孤点留下时, 该阶段结束^[6].

4 实验及结果分析

为了验证本文算法的有效性, 采用 Middlebury 大学计算机视觉研究中心提供的立体图像 baby 图 (370×413) 进行实验, 为分析匹配算法的性能, 本文采用文献[7]中的均方根误差 (记作 R) 和误匹配百分比 (记作 B) 两个指标, 仿真平台选用 matlabR2011a 和 vs2010.

$$R = \left(\frac{1}{N} \sum_{(x,y)} |d_c(x,y) - d_{GT}(x,y)|^2 \right)^{1/2} \quad (13)$$

$$B = \frac{1}{N} \sum_{(x,y)} (|d_c(x,y) - d_{GT}(x,y)| > \delta_d) \quad (14)$$

式中, $d_c(x,y)$ 是算法计算所得的视差值, $d_{GT}(x,y)$ 是真实的视差值, δ_d 是视差的容许错误阈值, 本文中取 $\delta_d = 1$, 参数 $\alpha = 0.1$. 实验图像如图 3 所示.

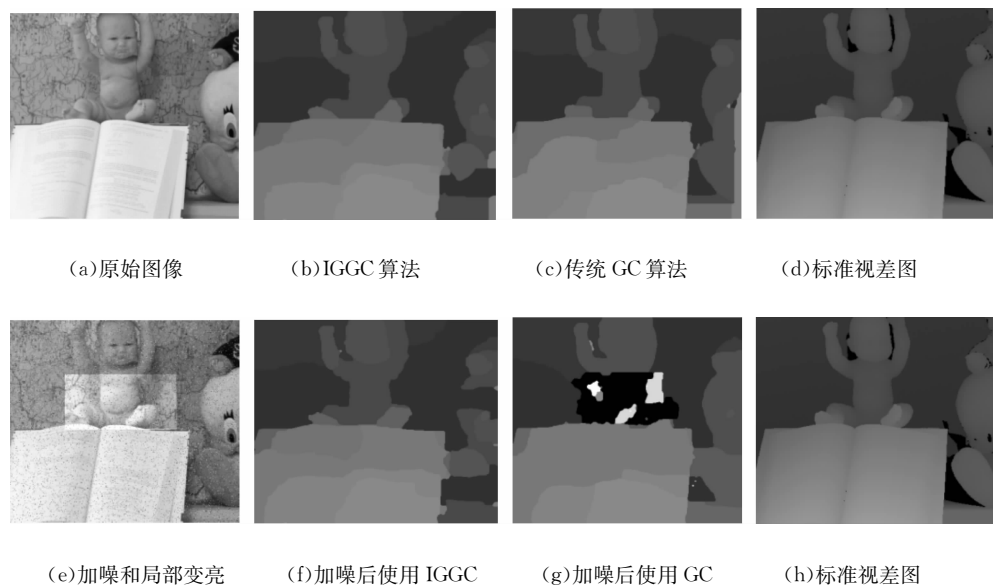


图 3 仿真结果图

在图 3 中,图(a)~图(d)为一组,匹配所用左右两幅图片都是来自数据库中的标准图片,可以看出,GC 算法的匹配效果很理想,当 IGGC 算法与传统的 GC 用于标准图片匹配时,效果几乎相同;图(e)~图(h)为另一组,在匹配所用的右图中加入了 1% 的盐噪声,并改变了图像的局部亮度,如图(e),比较图(f)和图(g)可以看出,局部亮度的变化对传统的 GC 算法影响很大,而 IGGC 算法对噪声和局部亮度变化具有一定的鲁棒性,表 1 和表 2 是利用上文中提到的两个指标定量地研究算法的有效性.

表 1 未处理前视差图的 R 值和 B 值表

	未处理前使用 IGGC(图 3b)	未处理前使用传统 IGGC(图 3c)
均方误差(R)	1.14	1.16
误匹配率(B)/%	3.4	3.8

表 2 右图加噪和局部变亮后视差图的 R 值和 B 值表

	右图处理后使用 IGGC(图 3f)	右图处理后使用传统 GC(图 3g)
均方误差(R)	1.18	3.16
误匹配率(B)/%	3.6	6.1

比较表 1 和表 2 可知,当出现噪声和局部亮度的变化时,利用本文所描述的 IGGC 算法匹配后得到的视差图的均方误差和误匹配率没有发生明显的变化,而利用传统的 GC 算法得到的视差图的均方误差增大,误匹配率增大,证明了本文描述的 IGGC 算法是一种具有鲁棒性的立体匹配算法.

5 结束语

立体匹配技术是机器视觉的关键技术,把立体匹配问题转化求解能量函数最值的问题,然后根据能量函数构造合适的图,最后通过求解图的最小割(最大流)来得到能量函数的最小值,本文提出了一种基于图像的 IG 值的 IGGC 立体匹配算法,对噪声和局部光照的变化具有一定的鲁棒性,实验证明,IGGC 算法在一定程度上能提高 GC 算法的稳健性.

参考文献:

[1] Vladimir Kolmogorov, Ramin Zabih. Mutli-camera scene reconstruction via graph cuts[C]//7th European Conference on Computer Vision. Copenhagen: Springer, 2002: 82-96.

[2] Kolmogorov V. Graph based algorithms for scene reconstruction from two or more views[D]. Ithaca: Cornell University, 2004: 1-139.

[3] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graphcuts[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23 (11): 1222-1239.

[4] Kolmogorov V, Zabin R. What energy functions can be minimized via graphcuts[J]. IEEE Transon Pattern Analysis and Machine Intelligence, 2004, 26 (2): 147-159.

[5] 吴福朝. 计算机视觉中的数学方法[M]. 北京: 科学出版社, 2008: 301-305.

提高.

可以看到,区域页级映射算法相对于传统页级映射算法明显降低了映射表所占用的内存空间,这是区域页级映射算法的最大优势.另外,由于区域映射也是基于页级映射的,其最小的映射粒度是页级的,因此它的性能要高于以块为粒度的块级映射.由于在区域页级映射算法中直接利用了通道间的并行,其读操作性能会优于传统页级映射.

6 结束语

本文介绍了一种基于区域页级映射的闪存转换层映射机制.针对多通道并行固态硬盘的结构,提出适用并行结构的物理页分配算法,通过区域、块、页这三级映射关系完成地址映射,大幅减少了页级映射表占用内存空间太大的问题,由于考虑到固态硬盘多通道的并行结构,对固态硬盘的性能、寿命、功耗等都有很大提升.对固态硬盘技术中的损耗均衡和垃圾回收方法进行了研究,提出了两级损耗均衡的方法,通过分析表明,区域映射机制在内存空间占用、性能、寿命等方面占有优势.

参考文献:

- [1] Intel Corporation. Understanding the flash translation layer (FTL) specification[R]. Santa Clara: Intel Corporation, 1998.
- [2] Gupta A, Kim Y, Ugaonkar B. DFTL: A flash translation layer employing demand-based selective of page-level address mapping [J]. SIGPLAN Not., 2009, 44(3):229-240.
- [3] Zhou You, Wu Fei, Huang Ping, et al. An efficient page-level FTL to optimize address translation in flash memory[C]//Proceedings of the Tenth European Conference on Computer Systems. New York: ACM, 2015: 12-27.
- [4] Ban A. Flash file system optimized for page-mode flash technologies; USA, 5937425 [P]. 1999-08-10.

- [5] Kim J, Kim J M, Noh S H, et al. A space-efficient flash translation layer for compact-flash systems [J]. IEEE Transactions on Consumer Electronics, 2002, 48(2): 366-375.
- [6] Lee S W, Park D J, Chung T S, et al. A log buffer-based flash translation layer using fully-associative sector translation [J]. ACM Transactions on Embedded Computing Systems, 2007, 6(3):1-27.
- [7] Kang J U, Jo H, Kim J S, et al. A superblock-based flash translation layer for nand flash memory [C]//Proceedings of the 6th ACM & IEEE International Conference on Embedded Software. New York: ACM, 2006: 161-170.
- [8] Hu Yang, Jiang Hong, Feng Dan, et al. Exploring and exploiting the multi-level parallelism inside SSDs for improved performance and endurance [J]. IEEE Transactions on Computers, 2013, 62(6):1141-1155.
- [9] Micro. Micro MT29F64G08AKCBBH2 NAND flash memory datasheet [EB/OL]. [2015-03-18]. http://www.micron.com/-/media/Documents/Products/DataSheet/NANDFlash/60Series/M62B_16Gb_32Gb_64Gb_128Gb_AsyncSync_NAND.pdf, 2009.
- [10] Kang J, Kim J, Park C, et al. A mufti-channel architecture for high-performance and flash-based storage system[J]. Journal of Systems Architecture, 2007, 53: 644-658.

作者简介:

蔡晓乐 男,(1991-),硕士研究生,助理工程师.研究方向为计算机应用. E-mail:cxle_avic@163.com.

张亚棣 男,(1968-),博士,研究员.研究方向为计算机体系结构、嵌入式软件.

李亚晖 男,(1976-),博士,高级工程师.研究方向为嵌入式操作系统、计算机网络、信息安全.

海钰琳 女,(1991-),硕士研究生,助理工程师.研究方向为计算机应用.

(上接第 31 页)

- [6] 左森,郭晓松,万敬,等. Graph cuts 及其在立体匹配中的应用[C]//立体图象技术及其应用研讨会论文集.天津:中国图像图形学会,2005.
- [7] 李惠光,翁良宝,姜洪磊.改进的基于扩展移动的立体匹配算法[J]. 计算机工程与设计, 2013, 34(9): 3180-3183.

作者简介:

马东东 男,(1990-),硕士研究生.研究方向为图像处理算法. E-mail:m104104@163.com.

杨 靓 男,(1975-),研究员,博士生导师.研究方向为空间计算机体系结构设计.

刘向增 男,(1981-),博士,工程师.研究方向计算机视觉.