

基于像素点分类和颜色分割的树型滤波立体匹配

高申勇^{1,3}, 戈豪豪², 张桦²

(1. 浙江大学电气工程学院 杭州 310027; 2. 杭州电子科技大学计算机学院 杭州 310018;

3. 浙江水利水电学院信息工程与艺术设计学院 杭州 310018)

【摘要】针对树型滤波匹配算法只需颜色一个要素计算权重而引起的匹配错误,提出一种基于像素点分类和颜色分割的树型滤波局部立体匹配算法。首先,在计算初始匹配代价时,按照稳定度将像素点分类;其次根据参考图像的颜色信息将其建立为代价树,并在建树的过程中根据颜色分割约束获得颜色分割图像;利用颜色分割图像和像素点分类信息,改进代价树中各边的权值;最后执行树型滤波,并获得稠密的视差图,从而完成立体匹配。采用Middlebury数据集进行的实验结果表明,该算法相比传统的树型滤波算法,在各区域的精度上都有一定的提升。

关键词 颜色分割; 像素点分类; 立体匹配; 树型滤波

中图分类号 TP391.41

文献标志码 A

doi:10.3969/j.issn.1001-0548.2017.03.021

Tree Filter Matching Method Based on Pixel Classification and Color Segmentation

GAO Shen-yong^{1,3}, GE Hao-hao², and ZHANG Hua²

(1. College of Electrical Engineering, Zhejiang University Hangzhou 310027;

2. School of Computer Science and Technology, Hangzhou Dianzi University Hangzhou 310018;

3. School of Information Engineering, Zhejiang University of Water Resources and Electric Power Hangzhou 310018)

Abstract Tree filter matching method will cause wrong matching results since it considers only one single component with color to obtain the weight of support region. This paper presents a local tree filter stereo matching method based on pixel classification and color segmentation. The pixels of an image are classified according to their stability during their initial matching cost computation phase. A tree structure based on color information of reference image is constructed, meanwhile, a segmented image with color segmentation constraint is generated. The weight value of each edge is improved by utilizing the information of segmented color image and classified pixels. Finally, the tree filter is executed and the dense disparity is achieved. The experimental results on Middlebury datasets show that our proposed method has higher accuracy than other original tree filter methods in each special region.

Key words image segmentation; pixel classification; stereo matching; tree filter

立体匹配算法在计算机视觉的发展上起着举足轻重的作用,是近年来计算机视觉研究领域的一个大方向。立体匹配算法旨在找出左右图像对中的对应点,获得其视差值,进而利用视差值和深度值的关系,得到深度信息,以完成立体匹配的过程。文献[1]根据其研究与总结,将立体匹配算法划分为初始匹配代价计算、代价聚合、视差初始计算、视差求精4个步骤。立体匹配算法分为全局算法和局部算法两种。全局算法主要预先定义一个能量函数,通过迭代使其能量函数最小化来获得匹配结果。该方法精度非常高,但不断迭代耗费了大量的时间。常

用的全局算法有图割^[2-3]、置信传播^[4]、动态规划等。局部算法则主要通过窗口和各相邻点权值的选取实现像素的单匹配,局部匹配算法的复杂度很低,精度不如全局算法。相比全局算法,局部算法一步到位,没有迭代过程。因此局部匹配算法的核心问题,就是在代价聚合的过程中选取合适的匹配窗口和权值以获得高精度的视差图,然后通过视差和深度的关系,由视差图所提供的信息推导出深度信息。

最初的匹配算法通常采用固定的匹配窗口和权值,算法实现较简单。但是在低纹理区域、遮挡区域、视差不连续区域等特殊区域,匹配误差极大,

收稿日期:2016-03-08;修回日期:2016-05-05

基金项目:国家自然科学基金面上项目(61471150);科技部国际科技合作专项项目(2014DFA12040);浙江省重点科技创新团队项目(2011R50009);浙江省自然科学基金面上项目(LY13F020033)

作者简介:高申勇(1981-),男,博士生,副教授,主要从事立体视觉匹配和智能机器人控制技术方面的研究。

所以局部匹配算法逐渐演变成基于匹配窗口和基于权重两种。文献[5]提出了一种自适应的窗口算法, 该算法以待匹配点为中心, 根据待匹配点与周围像素的颜色关系为每个像素点构建支持窗口。这种方法选取待匹配点周围和颜色信息更接近的像素点作为匹配窗口, 能有效地避免视差不连续区域的误匹配, 聚合代价更加可信。文献[6]提出了一种自适应权重(adaptive weight)算法, 利用空间距离和颜色信息两个成份计算每一个像素点对于待匹配点的权重。这样获得的权值信息极具代表性, 获得的视差图结果在精度上也取得了巨大的进步。文献[7]提出一种基于测地距离作为支持权重的方法, 也取得了很好的效果。但是这些自适应权重的计算方法时间复杂度较大, 在一些实时的应用场景下无法实现。

文献[8]提出了一种树型滤波算法, 该算法根据颜色信息, 通过相邻像素点连接, 将整个参考图像构建成一个最小生成树模型, 树模型上任一像素点到其他像素点上只有一条通路, 利用双向滤波计算代价值。该算法突破了局部支持窗口的概念, 利用整个参考图像上所有像素点聚合单一像素点, 但直接参与计算的却只有待匹配点的父亲节点和孩子节点, 这使得算法在复杂度和精度上的表现十分出色。然而由于建树过程中每个节点只和自己的相邻节点连接, 导致节点的空间距离都为1, 所以计算权值时利用颜色信息作为唯一标准, 使得权重分配存在问题。文献[9]针对该问题, 提出了基于分割树的滤波算法, 并引入了初始深度信息调整权重, 但是初始深度需要一次滤波计算获得, 并且需要重新建树, 时间复杂度消耗较大。

基于此, 本文提出一种基于颜色分割和像素点分类的树型滤波立体匹配算法。该算法利用颜色分割和像素点分类中的一些限定条件, 修改参考图像构建的树结构中部分边的权值, 使得权值更为合理, 以获得更精确的稠密视差图, 从而得到更精确的立体匹配结果。

1 算法概述

本文算法利用颜色分割和像素点分类信息对树结构中的权重进行修改, 构建新的生成树结构。该算法基于以下两条假设: 同一颜色分割上的像素点的视差值平滑变化; 同一区域中稳定点的代价值应该向不稳定点传播。

算法的主要流程如图1所示。获得左右图像输入后, 首先, 计算初始代价值, 根据初始代价中蕴含

的信息对像素点进行分类, 将像素点分为稳定点与不稳定点。同时, 根据输入图像的颜色信息为图像构建对应的生成树, 并在构建生成树同时利用分割约束进行区域分割以获得颜色分割图像。然后, 基于像素点分类信息、颜色分割图像以及初始生成树共同构建改进的生成树, 增加稳定点在代价聚合中的权重, 减少不稳定点在聚合中的权重, 为边缘区域设置惩罚值, 减少不连续区域的像素点对各自区域造成的影响。最后, 一次执行代价聚合和视差的计算及精化, 获得稠密视差图。

本文针对原始树型滤波算法权值计算方式单一的缺点, 引入像素分类集与颜色分割图改进其权值, 使获得的视差图具有更高精度、更加平滑的边缘, 并且在时间复杂度上并没有太大的变化。本文的贡献主要是以下几点:

- 1) 将像素点分类运用到权值改进的领域, 使得稳定点在代价传播中拥有更高的权值。
- 2) 在构建生成树的过程中添加分割约束以获得颜色分割图像, 对时间复杂度的影响很小, 并且将颜色分割图像直接运用以改进权重。

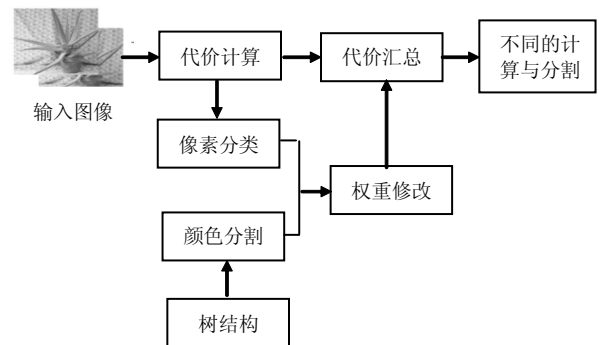


图1 算法流程图

2 本文算法

2.1 生成树的构建

构建生成树的主要思想是先将参考图像视作一幅图 $G=(V,E)$, 其中, V 代表节点集, 即参考图像中的所有像素点, E 代表边集。点集中的节点 p 和 q 所连成的边 $e_{p,q}$ 的值 W_e 即为两相邻像素点的距离, 像素点 p 与 q 的距离为:

$$D(p,q) = |I(p) - I(q)| \quad (1)$$

式中, $I(p)$ 与 $I(q)$ 为像素点 p 和 q 的颜色信息。针对每一个节点 $v_p \in V$ 构建一个子树 T_p , 并通过不断的合并这些子树, 最终生成一个满足以下条件的最小生成树:

- 1) 图 G 中所有的节点都在子树 T 中。

2) 子树 T 上的任一节点到另一节点有且仅有一条通路。

3) 子树 T 中边的和值 $D(T)=\sum_{e_{p,q}\in E_T} D(p,q)$ 最小。

构建生成树的伪代码如下:

算法: 参考图像的最小生成树构建。输入: 参考图像, 将参考图像视作一幅图 $G=(V,E)$ 。输出: 最小生成树 $T=(V,E')$, 树 T 中包含图 G 中的所有节点 V , 但是 E' 是 E 的一个子集, 即 $E'\in E$ 。

- 1) 计算边集 E 的距离信息。
- 2) 边集 E 按距离排序, 以升序重新排列边集。
- 3) 初始化边集 $E'\in\emptyset$ 。
- 4) for 每一个节点 $v_p\in V$ do
初始化子树 $T_p=(V_p,E_p)$, 其中, $V_p=V\{v_p\}$,

$E_p=\emptyset$ 。

5) end for

6) for 每一条边 $e_{p,q}\in E$ do

7) if $T_p\neq T_q$ then

将 T_p 与 T_q 合并成一棵新的子树

$T_{p,q}=(V_{p,q},E_{p,q})$, 其中, $E_p\cup E_q\cup\{e_{p,q}\}\rightarrow E_{p,q}$,

$V_p\cup V_q\rightarrow V_{p,q}$ 。

8) 更新 E' , $E'\cup\{e_{p,q}\}\rightarrow E'$ 。

9) end if

10) 当满足 $|E'|\equiv|V|-1$ 时, 跳出循环。

11) end for

12) return $T=(V,E')$ 。

2.2 基于树结构的颜色分割

2.1节中的生成树构建方法, 在构建生成树的同时, 利用生成树结构以及其距离信息, 对参考图像进行颜色分割, 生成对应的颜色分割图像。

由于物体的边缘通常都是颜色信息变化较大的区域, 所以在构造生成树时, 引入颜色分割约束, 子树 T_p 与 T_q 的分割约束为:

$$CS(T_p,T_q)\leq\min\left(\max(e_{T_q})+\frac{\tau}{|T_q|},(\max(e_{T_p})+\frac{\tau}{|T_p|})\right) \quad (2)$$

式中, e_{T_q} 和 e_{T_p} 分别是子树 T_q 与 T_p 中的边集; τ 是一个预定义的固定值, 本文所有实验中 $\tau=1200$ 。图2为本文方法获得的颜色分割图像与经典的MeanShift算法^[10]获得的分割图像的对比。可以看出, 本文算法所获得的分割图像的分割效果基本准确, 并且相对于MeanShift算法而言, 本文算法融合在建树的过程中, 不需要消耗多少额外的计算量, 非常适合应用到本文的后续算法中。

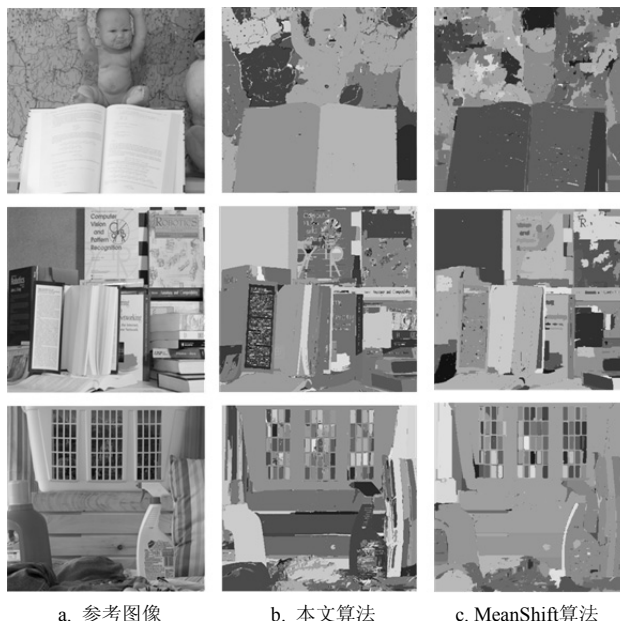


图2 本文与MeanShift分割算法结果对比

2.3 像素点分类方法

根据AD-Gradient公式^[11]获得初始代价结果。将像素点根据其初始代价值中极值的辨识度分为稳定点和不稳定点两种。在计算初始代价的同时, 根据像素点各深度下的代价值判断其稳定性。像素点 p 的稳定性为:

$$P(p)=\begin{cases} \text{Stab} & S(p)>\varphi \\ \text{Unst} & S(p)\leq\varphi \end{cases} \quad (3)$$

式中, φ 是一个预定义的固定值, 本文实验中 $\varphi=0.04$; Stab表示该值稳定; Unst表示该值不稳定; $S(p)$ 是像素点 p 的极值代价值辨识度, 有:

$$S(p)=\left|\frac{C_1(p)-C_2(p)}{C_2(p)}\right| \quad (4)$$

式中, $C_1(p)$ 与 $C_2(p)$ 分别是代价值结果最好和次好的代价值。

2.4 基于颜色分割和像素点分类的聚合权值改进

初始的权重计算为:

$$W(p,q)=\exp\left(-\frac{D(p,q)}{\sigma}\right) \quad (5)$$

式中, σ 是一个预定义的值, 本文中所有实验 $\sigma=0.08$ 。以上权值计算中, 只使用了颜色信息一个要素, 而由于生成树中只能邻接节点相连, 距离信息均为1, 所以不能参与权值的计算。本文引入颜色分割与像素点分类, 改进后的权值为:

$$W_m(p,q)=\begin{cases} \exp\left(-\frac{D(p,q)}{S(p,q)}\right) & \text{seg}(p)=\text{seg}(q) \\ \exp\left(-\frac{D(p,q)+\mu}{\sigma}\right) & \text{其他} \end{cases} \quad (6)$$

式中, $seg(p)$ 与 $seg(q)$ 代表像素点 p 和 q 的颜色区域分割信息; μ 是一个预定义的惩罚值; $S(p,q)$ 是一个自适应的调整参数, 有:

$$S(p,q)=\begin{cases} \sigma & P(q) == \text{Stab} \\ \rho^2\sigma & P(q) == \text{Unst} \\ \rho\sigma & \text{其他} \end{cases} \quad (7)$$

式中, ρ 是一个调节值, 本文实验中 $\rho=0.5$ 。改进后的权值边缘区域因为添加了一个惩罚值, 受另一区域的影响减小, 而同一区域中的稳定点的权重增加, 不稳定点的权重减弱, 实现了视差从稳定点向不稳定点的传播。

2.5 树型滤波

在获得权值之后, 根据构造好的生成树结构与初始代价值进行代价聚合。如图3所示, 树型滤波分为两步, 分别为从叶子节点到根节点的聚合与从根节点到叶子节点的聚合。

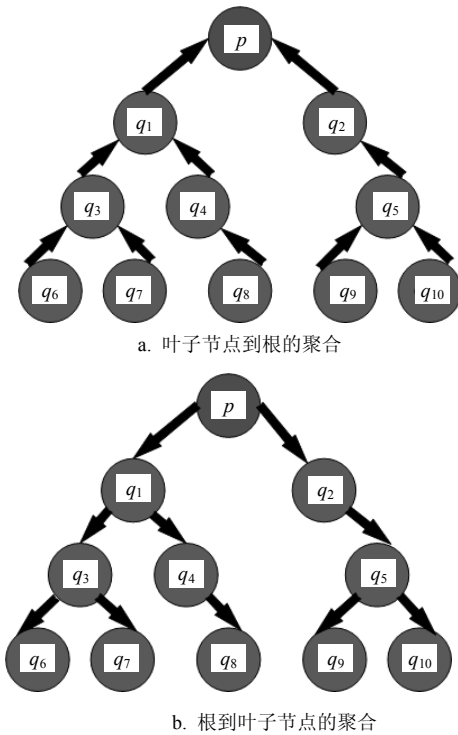


图3 树型滤波算法

通过两步聚合, 每个节点实际上只直接从其双亲节点和孩子节点获得聚合代价, 但由于树型滤波的可传播性, 每一个待匹配点都间接地从所有像素点中获得聚合值。 $C_d^{A\uparrow}(p)$ 为像素点 p 的聚合代价, 从叶子节点到根节点方向的聚合为:

$$C_d^{A\uparrow}(p)=\sum_{q\in\text{Ch}(p)} W_m(p,q)\cdot C_d^{A\uparrow}(q) \quad (8)$$

式中, $\text{Ch}(p)$ 是像素点 p 的孩子节点集合; $W_m(p,q)$ 为

改进之后的权值信息。从根节点到叶子节点方向的聚合为:

$$C_d^A(p)=W_m(\text{Pa}(p),p)C_d^A(\text{Pa}(p))+ (1-W_m^2(\text{Pa}(p),p))C_d^{A\uparrow}(p) \quad (9)$$

式中, $\text{Pa}(p)$ 是像素点 p 的父亲节点。两步聚合后, 算法的代价聚合部分完成, 然后通过简单的WTA公式获得初始视差图, 引入左右一致性检测等方法, 优化初始视差图以获得最后的精确结果。

2.6 时间复杂度

由以上方法可知, 相对于原始的树型滤波算法而言, 本文的时间复杂度增量, 主要集中在颜色分割、像素点分类以及权重改进3个方面。假设参考图像边的数量为 e , 节点数量为 n , 视差范围为 d , 由文献[8]可知, 原始树型滤波算法在建树方面的时间复杂度为 $O(e+n+2e\log_2n)$, 而滤波算法的时间复杂度为 $O(nd)$ 。

在本文算法中, 颜色分割中的判定条件是逐边执行的, 所以颜色分割的时间复杂度为 $O(e)$, 权重修改的复杂度也同为 $O(e)$, 另外像素点分类的时间复杂度为 $O(nd)$ 。一般情况下, 建树的时间复杂度要大于滤波算法, 所以本文算法在时间复杂度上的增量是非常小的。

3 实验结果和分析

将几个传统的树型滤波算法与本文算法的实验结果作比较。代价初始计算的方法均采用AD-Gradient方法, 处理方法均采用树型结构非局部后处理方法。测试数据集是Middlebury网站^[12]上提供的标准图像。首先测试4幅图像(Tsukuba、Venus、Teddy Cones), 分别测试非遮挡区域、全部区域以及边缘区域3个区域的误匹配点(occlusion、all、disc), 本文算法的参数设置如下: $\tau=1\ 200$, $\varphi=0.04$, $\sigma=0.08$, $\mu=5$, $\rho=0.5$ 。

图4为本文算法与其他4个经典的结果图对比。从图4可知, 本文提出的算法相比于原始的树型滤波算法, 有一定的优势。在Tsukuba的桌角位置, ST-1和MST算法的前景物体的视差都会溢出, 导致前景物体的视差影响到背景。而本文算法中权值经过改进之后, 这样的情况可避免。另外, Cones图片中许多圆锥体的边角位置也会出现前景视差值溢出到背景范围的情况, 本文的算法相对于传统的树型滤波算法, 有一定的改进和提升。

表1是更直观的数据结果。从表1中可以发现, Tsukuba图像的整体效果提升十分明显, 无论是在边

缘区域还是在非遮挡区域效果均有较大改善, Venus和Cones的边缘区域的错误率, 有一定程度的下降,

另外, 4幅标准数据集的整体的误匹配率也下降了3.1%左右。

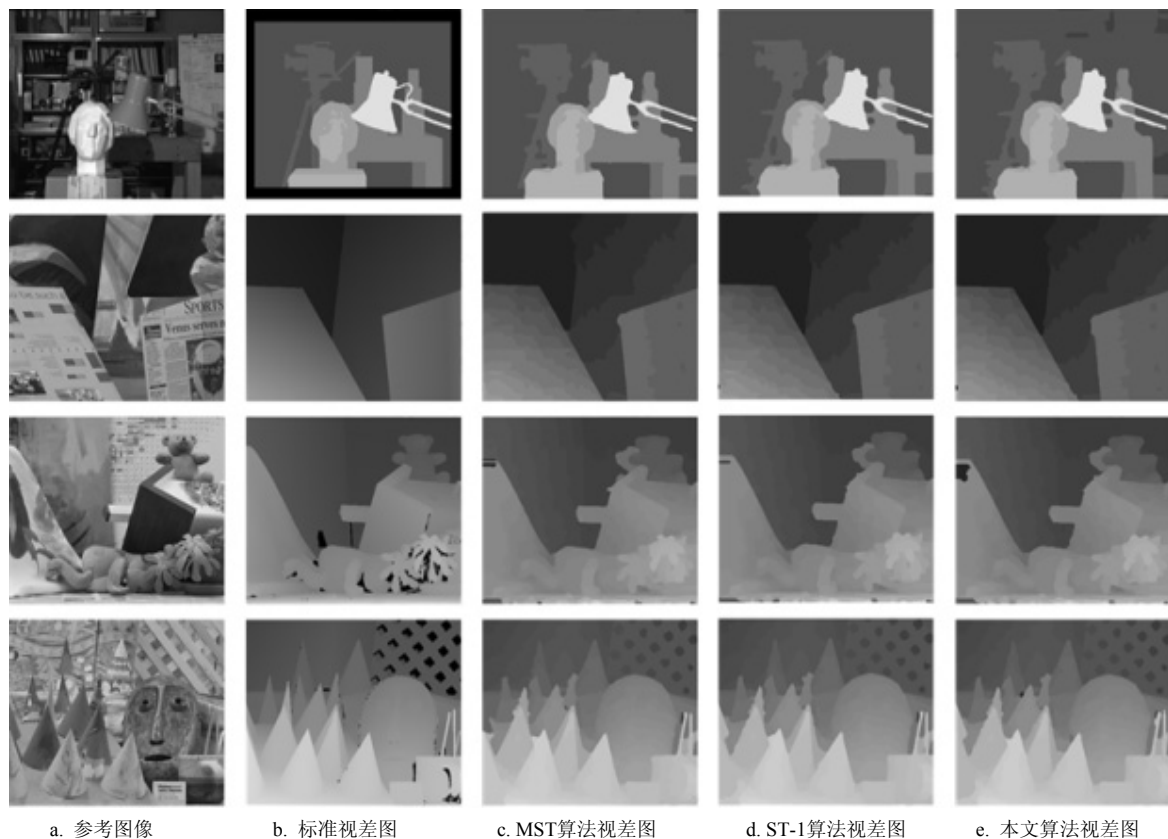


图4 视差图对比

表1 算法误匹配率对比

算法	Tsukuba			Venus			Teddy			Cones			平均误差
	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	
MST	1.86	2.29	8.75	0.67	0.89	5.66	5.20	9.89	12.97	2.51	8.38	7.40	5.54
ST-1	2.01	2.50	9.07	0.64	0.91	5.27	5.18	9.76	12.98	2.48	8.50	7.40	5.56
本文	1.50	1.99	7.99	0.64	0.99	5.29	5.17	10.18	12.99	2.39	8.40	7.07	5.39

4 结束语

本文提出了一种新的树型滤波算法。通过像素点分类与颜色分割改进树型滤波算法中生成树的权值, 使稳定点的代价值向不稳定区域扩散, 同时也保护了视差边缘区域, 使得无论是在弱纹理区域, 还是在视差不连续区域, 代价值都变得更加合理。实验结果表明, 本文算法在原始树型滤波算法的基础上, 有明显的进步, 并且时间复杂度上没有太大的改变。

参考文献

- [1] SCHARSTEIN D, SZELISKI R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[J]. *International Journal of Computer Vision*, 2002, 47(1-3): 7-42.
- [2] HONG L, CHEN G. Segment-based stereo matching using graph cuts[C]//IEEE Computer Society Conference on Computer Vision & Pattern Recognition. [S.l.]: IEEE, 2004: 74-81.
- [3] 祝世平, 杨柳. 基于自适应分水岭的图割的立体匹配算法[J]. *光学学报*, 2013(3): 221-229.
ZHU Shi-ping, YANG Liu. Stereo matching algorithm with graph cuts based on adaptive watershed[J]. *Acta Optica Sinica*, 2013(3): 221-229.
- [4] SUN J, SHUM H Y, ZHENG N N. Stereo matching using belief propagation[M]//Computer Vision — ECCV 2002. Berlin Heidelberg: Springer, 2002: 787-800.
- [5] ZHANG K, LU J, LAFRUIT G. Cross-based local stereo matching using orthogonal integral images[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2009, 19(7): 1073-1079.
- [6] YOON K J, KWEON I S. Locally adaptive support-weight approach for visual correspondence search[C]//Computer Vision and Pattern Recognition. San Diego, CA, USA: IEEE,

- 2005: 924-931.
- [7] HOSNI A, BLEYER M, GELAUTZ M, et al. Local stereo matching using geodesic support weights[C]//IEEE International Conference on Image Processing. [S.l.]: IEEE, 2009: 2069-2072.
- [8] YANG Q. A non-local cost aggregation method for stereo matching[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE Computer Society, 2012: 1402-1409.
- [9] MEI X, SUN X, DONG W, et al. Segment-tree based cost aggregation for stereo matching[C]//Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.]: IEEE, 2013: 313-320.
- [10] COMANICIU D, MEER P. Mean shift: a robust approach toward feature space analysis[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 24(5): 603-619.
- [11] RHEMANN C, HOSNI A, BLEYER M, et al. Fast cost-volume filtering for visual correspondence and beyond[C]//Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE Computer Society, 2011: 3017-3024.
- [12] SCJARSTEIND, SZELISKI R. Middlebury stereo evaluation[EB/OL]. [2015-12-20]. <http://vision.middlebury.edu/stereo/eval/>.

编辑 漆蓉

(上接第587页)

- [18] 吴少智, 吴跃, 徐鹏, 等. 支持向量回归的颅内压时间系列无损估计方法[J]. 电子科技大学学报, 2011, 40(6): 956-960.
WU Shao-zhi, WU Yue, XU Peng, et al. Support vector regression based time series mining approach for non-invasive ICP assessment[J]. Journal of University of Electronic Science and Technology of China, 2011, 40(6): 956-960.
- [19] R-Project. Rcasting functions for time series and linear models[EB/OL]. [2015-11-01]. <http://cran.r-project.org/web/packages/forecast/index.html>.

编辑 漆蓉