

利用图像分割的基于图割的立体匹配算法

胡春海 熊 英*

(燕山大学测试计量技术及仪器河北省重点实验室, 河北 秦皇岛 066004)

摘要 立体匹配通过寻找同一空间景物在不同视点下投影图像的像素间的一一对应关系, 最终得到该景物的视差图。在对匹配算法作了深入研究的基础上, 提出了一种利用图像分割的基于图割的立体匹配算法。算法把参考图分割成多个区域, 然后用平面公式在一个分割中建立视差。视差模板是从初始视差分割中提取的。每一个分割被分配到精确的视差模板。构建全局能量函数, 能量函数的鲁棒最小化是由基于图割的最优化获得的。算法对低纹理区域和接近视差边界区域有很好的匹配效果, 同时, 又解决了传统的基于全局算法中计算量过大, 实时性不好的问题。实验表明, 本算法能满足高精度、高实时性要求。

关键词 立体匹配; 视差; 图割; 图像分割; 能量函数

中图分类号 TP391 文献标识码 A doi: 10.3788/AOS200828s2.0043

Graph-Cut-Based Stereo-Matching Algorithm Using Image Segmentation

Hu Chunhai Xiong Ying

(Measurement Technology and Instrumentation Key Lab of Hebei Province, Yanshan University, Qinhuangdao, Hebei 066004, China)

Abstract Stereo-matching is an algorithm that through searching the corresponding relations between the projection images' pixels of the same scene on different visual point, gains a disparity map of the scene finally. This paper presents a graph-cut-based stereo-matching algorithm using image segmentation on the basis of in-depth study on the images matching algorithms. In proposed algorithm, the reference image is divided into segments. Then modeling disparity is built inside a segment by a planar equation. A set of disparity layers is extracted from initial disparity segments in a clustering process. A global energy function is constructed. Robust minimization of the cost function is achieved by graph-cut-based optimization. The proposed algorithm produces good-quality results, especially in regions of low texture and close to disparity boundaries. The expensive computing cost for traditional global algorithms can also be reduced. Experiments demonstrate that the algorithm can meet both demands for high resolution and real time.

Key words stereo-matching; disparity; graph cut; image segmentation; energy function

1 引 言

立体匹配通过寻找同一空间景物在不同视点下投影图像的像素间的一一对应关系, 最终得到该景物的视差图。本文讨论的是立体匹配中双目视觉的情况, 即只对两张图片进行处理。立体匹配是目前计算机视觉研究中的一个难点和热点, 在许多计算

机视觉的应用中都是很关键的一个步骤, 如三维重构(3D reconstruction)^[1]、基于图像的绘制(Image based rendering)^[2]等。

根据视差数据的分布情况, 匹配算法分为稀疏视差匹配和密集视差匹配两大类。稀疏视差匹配^[3]一般以过边缘轮廓、线段等图像特征为匹配基元, 并仅匹配两幅图像的特征区域。这些方法通常都比较

基金项目: 河北省教育厅科学研究计划(2006445)资助课题。

作者简介: 胡春海(1966—), 男, 教授, 博士生导师, 主要从事多传感器数据融合、视觉检测、精密测控技术等方面的研究。

E-mail: fred-hu@ysu.edu.cn

* 通信联系人: E-mail: 187323901@163.com

鲁棒,但一般情形下获得比较稀疏的结果。密集视差算法^[4]通常以图像灰度或其它特性作为匹配基元,由候选点邻域间的相关程度作为判别依据,常用的方法有:区域相关法^[5],动态规划法^[6],以及基于光流场^[7]和基于图割^[8]的方法等。其优点在于给出所有(或几乎所有)像素的视差信息,但会产生一些误匹配。

通过构建能量函数试图获得图像的某些全局性质,在早期的视觉中得到广泛的应用。但不可避免的是通常很难获得能量函数的全局最小化。鉴于此,很多作者更倾向于寻找局部小的求解。然而传统的匹配算法^[9~12]存在的最大问题是计算量过大,在实时性要求比较高的情况下不适合采用。与传统的8种算法(例如模拟退火法^[13]、M-估计法^[14]等)相比,图割算法不仅总体精度高,且在不连续区域和低纹理区域的精度也比其它算法都高。不仅如此,即便有些算法(例如模拟退火法)的精度与图割算法的精度接近,但图割算法在优化过程中收敛更快。

本文提出一种利用图像分割的基于图割的立体匹配算法。该算法框架基于平滑表面假设,即在单一的色彩区域中视差变化是平滑的,这样就可以通过一个平面模板(以下简称平面模板为模板)公式 $d(x, y) = a + bx + cy$ 来描述该区域($[a \ b \ c]^T$ 为模

板参数),由该公式计算出来的视差称之为模板视差。基于这个假设,就可以把传统的基于全局算法中对每个点分配最优视差的问题转化为对每个区域分配最优模板的问题。构造一个全局能量函数,测量了在像素和分割水准上的分配的品质。能量函数的鲁棒最小化是由基于图割的最优化获得的。本算法大大提高了算法的实时性。

2 图像分割算法

图像分割是整个算法的基础,本算法采用 Pedro 和 Daniel 提出的基于图论的高效图像分割算法,该算法具有精度高和实时性好的特点。一般来说,为了使平滑表面假设更加可靠,应使各区域色彩尽量单一,可将区域间色彩差异的阈值参数适当调低,使划分的区域偏小一点。

表1给出了各张检验图进行图像分割的运算时间。图1为 Tsukuba 图用该算法进行图像分割后的结果。本文中算法结果均在 PC 机上运行所得。用于检验本算法的图片全部来自 Middlebury 图片库,这些图片可以检验立体匹配算法在各种不同环境下的表现情况。

表1 各张检验图进行图像分割的运算时间

Table 1 Cost time of image segmentation for four testing images

Picture	Venus (434×383)	Sawtooth (434×380)	Tsukuba (384×288)	Map (284×216)
Testing time /s	0.703	0.703	0.438	0.219

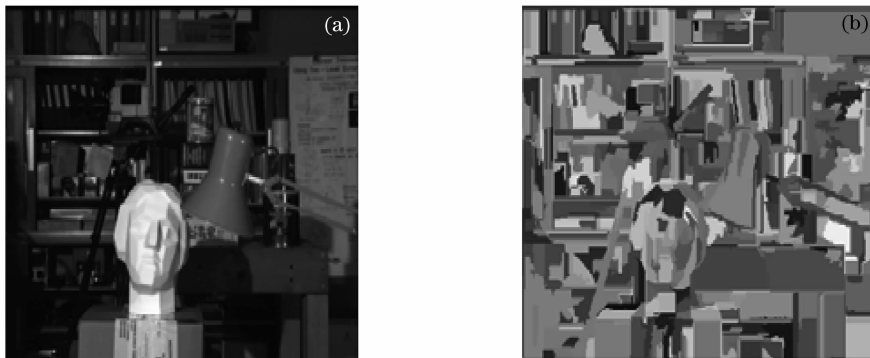


图1 Tsukuba 图的 (a) 参考图和 (b) 分割效果

Fig. 1 Original image and computed segmentation. (a) Reference image; (b) segment result

3 模板计算

3.1 视差初始值获取

为了讨论方便,假设两幅图像的外极线已经被校准过了,如果参考图中的一点 (x, y) 与匹配图中

的一点 (x', y') 相匹配,则应满足以下的对应关系:

$$x' = x + sd(x, y), y' = y, \quad (1)$$

式中 $s = \pm 1$, 用于保证视差 $d(x, y)$ 始终为非负值。本文中均选取左视图为参考图,右视图为匹配图,故

取 $s = -1$ 。

本算法先采用了绝对残差(SAD)计算各像素在不同视差时的匹配代价。如果采用大窗口进行计算,在低纹理区将获得更加可靠的初始匹配结果,但同时也会不可避免地增强前景膨胀效应,所以在算法中所有点的匹配代价都用 3×3 小窗口进行计算。

视差的初始值的精确性对模板的估计影响很大,除了通常采用的交叉校验外,又加入了相似点滤除的处理,从而增强了初始值的可靠性。若参考图中的某点在匹配图中存在可靠匹配点,该点被称为可靠点,否则为不可靠点。

3.2 初始模板计算

利用已获取的可靠点,便可进行模板参数的计算。算法采用的是加权最小二乘法求模板参数,并进行多次迭代直至参数收敛。显然,一个区域可靠点数越多,该区域计算所得的模板参数越可靠。设某个区域可靠点数为 N_{visible} ,若满足式(2),该区域被定为可靠区域,否则为不可靠区域。本算法只对可靠区域进行模板计算。

$$N_{\text{visible}} \geq N_v. \quad (2)$$

模板的误差 Δ 定义为

$$\Delta = f(a, b, c) = \sum_{i=1}^{N_{\text{visible}}} w_i [d_i - (a + bx_i + cy_i)]^2, \quad (3)$$

当 $\frac{\partial f}{\partial a} = 0, \frac{\partial f}{\partial b} = 0, \frac{\partial f}{\partial c} = 0$, Δ 取得最小值, w_i 为第 i 个点的加权系数。

第 1 次计算模板参数时,令 $w_i = 1, d_i$ 取各点的视差初始值。以后每迭代一次, d_i, w_i 都要更新一次。利用前一次的模板参数可计算出各点模板视差,在模板视差附近一定范围内重新选取匹配代价最小的视差作为新的 d_i 。加权二乘法的基本思想是若某点的 d_i 与模板视差的偏差越大,该点的权系数应该越小,这里取

$$w_i = e^{-Kt}, t = |(a + bx_i + cy_i) - d_i|. \quad (4)$$

4 用图割法进行模板分配

本算法采用基于全局的算法中常用的能量函数:

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

$$E_{\text{data}}(f) = \sum C(S_i, P_i), \quad (6)$$

$$E_{\text{smooth}}(f) = \sum_{(S_i, S_j)} L_{(i,j)} \delta(P_i \neq P_j), \quad (7)$$

式中 $C(S_i, P_i)$ 为第 i 块区域分配的模板 P_i 后的相似度匹配代价,第 i 块区域与第 j 块区域为相邻区域, $L_{(i,j)}$ 为它们的边界长度。当 $P_i \neq P_j$ 时, $\delta(P_i \neq P_j) = 1$, 否则为 0。

为了最小化上述能量函数见式(5),我们构建一个无向网络 N ,使得标号函数集和网络 N 的图割构成双射,并且对每一个标号 f ,其能量 $E(f)$ 即为其所对应的割的容量。根据上述映射关系,能量函数式(5)的最小化问题等价于 N 的最小割求解问题,而最小割问题可以通过计算最大流来解决。

网络流可以用来轻易解决图像匹配问题。给定图像 G ,分为 A 和 B 两部分。通过附加两个新结点 s 和 t ,由 G 建立网络 N ,从 s 到 A 的每个结点增添一条弧,将所有的边变为从 A 到 B 的弧,从 B 的每个结点到 t 增添一条弧。然后将所有弧的容量设为 1。在 N 上应用最大流-最小割定理(即在任一网络中,最大流的值等于最小割的容量),当求出最大流,也就获得了 N 的最小割,从而获得了问题的求解。

5 实验及其结果

为了检验算法的匹配效果,利用 Matlab 在普通 PC 机上进行了实验。表 2 反映了各算法对四组序列图像进行匹配所得的参数比较。从表中可以看出,本算法不仅运算速度相对较快,而且匹配准确率也较高。图 2 展示了算法对 Tsukuba, Sawtooth, Venus 和 Map 四组序列图像的运算效果。从图中可以看出视差图边界清晰,定位准确,尤其在低纹理区域和视差不连续区域得到了很好的效果。

表 2 各算法对四组序列图像进行匹配所得的参数比较

Table 2 Comparison of each algorithm's image matching parameters

		Our method	Layered	Belief prop.
Testing time /s	Tsukuba	1.035	1.813	1.348
	Sawtooth	1.070	2.391	1.626
	Venus	1.072	2.360	1.661
	Map	1.028	1.547	1.278
Error matching rate /%	Tsukuba	0.97	1.58	0.15
	Sawtooth	0.19	0.34	0.98
	Venus	0.16	1.52	1.00
	Map	0.16	0.37	0.84

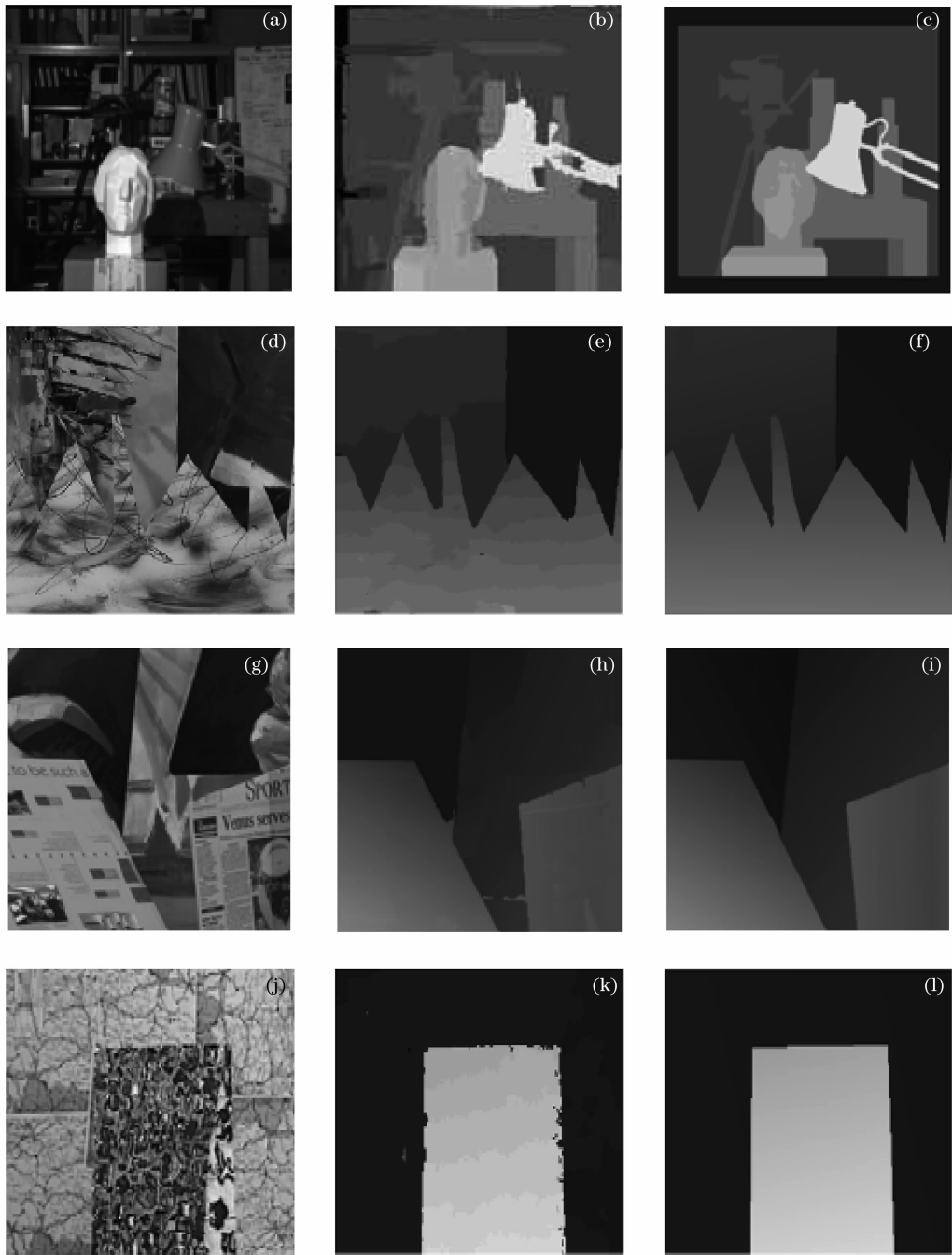


图2 本算法对四组序列图像的运算效果。从上到下的顺序为：Tsukuba图，Sawtooth图，Venus图和Map图。
从左到右的顺序为：参考图，本算法所得的视差图和真实视差图

Fig. 2 Results for four testing images. From top to down order: Tsukuba, Sawtooth, Venus, and Map. From left to right order: reference images, extracted disparity maps, and the ground truth disparity maps

6 结 论

提出了一种利用图像分割的基于图割的立体匹配算法。假定在单一的色彩区域中视差变化是平滑的，而且假定视差边界与分割边界相符。视差模板是从初始视差分割中提取的。每一个分割被分配到精确的视差模板。一个全局能量函数测量了在像素

和分割水准上的分配的品质。能量函数的鲁棒最小化是由基于图割的最优化获得的。实验表明，本算法对低纹理区域和接近视差边界区域有很好的匹配效果，同时，又解决了传统的基于全局算法中计算量过大，实时性不好的问题。因而，本算法能很好地满足计算机视觉应用场合的高精度、高实时性要求。

参 考 文 献

- 1 M. Ali, J. R. Shah, M. Ahmed. 3D Reconstruction and Model Acquisition of Objects in Real World Scenes Using Stereo Imagery[C]. Proceedings of the 7th International Multi Topic Conference, Islamabad, Pakistan, 2003, 32~37
- 2 H. T. Nguyen, N. Minh. Do Image-based rendering with depth information using the propagation algorithm[C]. *Proceedings of Acoustics, Speech, and Signal Processing*, Philadelphia, USA, 2005;589~592
- 3 D. Scharstein, R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J]. *International Journal of Computer Vision*, 2002, **47**(1): 7~42
- 4 V. Kolmogorov, R. Zabih. Computing visual correspondence with occlusions using graph cuts[C]. *Proc. the 8th International Conference on Computer Vision*, Vancouver, Canada, 2001, 508~515
- 5 Xu Yi, Zhou Jun, Zhou YuanHua. Phase matching and occlusion detection based on dynamic programming [J]. *Acta Electronica Sinica*, 2004, **32**(4): 591~595
徐奕, 周军, 周源华. 基于动态规划的相位匹配和遮挡检测[J]. *电子学报*, 2004, **32**(4): 591~595
- 6 H. Tao, H. S. Sawhney, R. A. Kumar. Global matching framework for stereo computation [C]. *Proc. the 8th International Conference on Computer Vision*, Vancouver, Canada, 2001, 532~539
- 7 H. Hirschmuller. Improvements in real-time correlation-based stereo vision[C]. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 2001, 141~148.
- 8 P. F. Felzenszwalb, D. P. Huttenlocher. Efficient graph-based image segmentation [J]. *International Journal of Computer Vision*, 2004, **59**(2):167~181
- 9 He Bin. *Digital Image Processing* [M]. Bei Jing: Demotic Mail Publishing Company, 2003
何斌. 数字图像处理[M]. 北京: 人民邮电出版社, 2003
- 10 Zhang yujin. *Image Processing and Analysis*. Beijing: Qinghua University Publishing Company, 2002
章毓晋. 图像处理和分析[M]. 北京: 清华大学出版社, 2002
- 11 Yang J., Lc Z. H. A domain of interest based novel search scheme in P2P networks[J]. *Journal of Computer Research and Development*, 2005, **42**(5): 804~809
杨舰, 吕智慧等. 一种基于兴趣域的高效对等网络搜索方案[J]. *计算机研究与发展*, 2005, **42**(5): 8804~809
- 12 Wang Hongmei, Zhang Ke, Li Yanjun. Research progress on image matching [J]. *Computer Technology and Application*, 2004, (19): 41~45
王红梅, 张科, 李言俊. 图像匹配研究进展[J]. *计算机工程与应用*, 2004 (19): 41~45
- 13 Y. Boykov, O. Veksler, R. Zabih. Fast approximat energy minimization via graph cuts [J]. *IEEE Trans.*, 2006, **23**(11): 1222~1239
- 14 V. Kolmogorov, R. Zabih. Computing visual correspondence with occlusions using graph cuts [A]. *Proc. the Eighth International Conference on Computer Vision (ICCV-01)* [C]. Vancouver, British Columbia, Canada: IEEE Computer Society, 2001. 2. 508~515