

# 一种基于图割理论的快速立体匹配算法

朱程辉, 任冉冉

(合肥工业大学 电气与自动化工程学院, 安徽 合肥 230009)

**摘要:** 针对传统的立体匹配算法中存在的低纹理区域和遮挡区域匹配精度低、实时性不好等问题, 提出了一种基于图割理论的立体匹配算法。把图像分割成色彩单一的不同区域; 计算初始视差图, 利用可靠点求取各分割区域的平面模板参数, 对模板参数相同的相邻区域进行融合; 构造全局能量函数, 采用图割算法求取全局能量最小的视差最优分配。实验结果表明, 该算法对低纹理区域和遮挡区域均有较好的匹配结果, 能够满足高精度、高实时性的要求。

**关键词:** 立体匹配; 图割; 图像分割; 能量函数; 视差

中图分类号: TP391.4

文献标识码: A

文章编号: 1674-7720(2010)10-0035-04

## A fast stereo matching algorithm based on graph cuts theory

ZHU Cheng Hui, REN Ran Ran

(School of Electric Engineering and Automation, Hefei University of Technology, Hefei 230009, China)

**Abstract:** Aiming at the problems that traditional stereo matching algorithm has its drawbacks of low resolution and poor real time, this paper presents a fast stereo matching algorithm based on graph cuts theory. First, the image is divided into different segments of a single color. Then calculating the initial disparity, the plane parameters of single segment are computed only using reliable pixels, and group neighboring segments with the same plane parameters. Finally construct a global energy function, minimization of the cost function is achieved by graph cuts. The experimental results indicate that the algorithm has a better performance in low texture regions and occluded portions, and can meet both demands for high resolution and real time.

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

所谓立体匹配,是指针对同一场景从不同视点拍摄图像,找出图像对中像素点之间的对应关系,求出该场景的视差图。目前,立体匹配已被广泛地应用在机器人视觉导航、医学成像、工业检测<sup>[1]</sup>和三维测量<sup>[2]</sup>等领域。然而,在空间三维景物投影到二维图像过程中,由于受光照条件、噪声干扰、景物几何形状畸变等因素的影响,立体匹配问题始终没有统一的解决方式,特别是在低纹理区域及遮挡区域,如何获得场景的稠密视差图仍然是一个极富挑战性的研究课题<sup>[3]</sup>。

近年来,国内外许多学者对立体匹配问题做了大量的研究。参考文献[4]提出了一种基于图像分割的立体匹配算法框架,用一系列的平面模板来描述经过图像分割产生的色彩单一的区域,这样就可以把传统的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,从而大大提高了算法的实时性,但在求解模板参数时没有考虑遮挡区域和低纹理区域对计算带来的误差,匹配精度不高。参考文献[5]针对立

体匹配问题提出了全局匹配方法,通过构造一个全局能量函数,同时计算所有像素的视差,将对应点的匹配问题转化为寻找能量函数的全局最优问题。目前常用的优化能量函数的算法有模拟退火法、贪婪算法、遗传算法、动态规划法<sup>[6]</sup>、图割算法<sup>[7-8]</sup>等,其中图割算法的匹配精度较高,误差较小,但耗时太长。

本文在参考文献[4]和参考文献[5]的基础上,提出了一种新的快速立体匹配算法。首先求取彩色图像分割后的图像的初始匹配图,只利用可靠点去计算每个分割区域的平面模板参数,以减少遮挡区域和低纹理区域对计算带来的误差。然后构造一个全局能量函数,把初始视差范围作为全局能量函数的一个参考项,限制有效视差范围小于事先定义的阈值,缩小视差搜索范围,从而提高了运算效率。最后利用图割算法优化全局能量函数,用分割区域代替像素进行构图,大大改善了算法的实时性。实验结果表明,本文算法在得到较精确的稠密视差图的同时,显著缩短了匹配时间,提高了算法的鲁

棒性。

### 1 彩色图像分割

因为彩色图像分割的质量好坏会影响到后期的匹配结果,所以选择合适的图像分割算法非常关键。目前常用的算法有均值平移算法、分水岭算法等,这些算法计算效率高、实时性好,但精度较低。Pedro 和 Daniel 提出的基于图论的高效图像分割算法<sup>[9]</sup>,适用范围广,精度高且实时性好,因此本文采用该算法进行彩色图像分割。

由于本文对图像分割的目的是将分割后的区域作为匹配基元,而不是去识别区域,因此只要不把前景和背景分割到同一区域就满足要求。为了使各分割区域色彩尽量单一,在计算过程中可将区域间色彩差异的阈值参数适当调低,这样分割的区域偏小些。

### 2 平面模板计算

平面模板计算一般可分为以下几个步骤:首先选择适当的匹配窗口对图像进行初始匹配,得到较精确的初始视差图;然后根据可靠点计算各分割区域所对应的平面模板参数;最后构造匹配代价公式进行模板优化,把模板参数相同的相邻区域进行融合。

#### 2.1 初始匹配

假设两幅图像已经过外极线校准,如果参考图 I 中的点 $(x,y)$ 和匹配图 J 中的点 $(x',y')$ 相匹配,则应满足如下关系:

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

其中, $d(x,y)$ 是有效视差范围内的任意视差值。

本文采用 SAD(Sum of Absolute Differences)算法计算各像素在不同视差时的匹配代价。大匹配窗口虽然可以在低纹理区域获得较好的初始匹配结果,但同时增强了前景膨胀效应<sup>[10]</sup>,因此在算法中选择  $3 \times 3$  的小窗口进行计算。设  $c(x,y,d)$  是参考图 I 中点 $(x,y)$ 其视差为  $d(x,y)$  的匹配代价,计算公式如下:

$$c(x,y,d) = \frac{1}{9} \sum_{i=-1}^1 \sum_{j=-1}^1 |I(x+i,y+j) - J(x+i+d,y+j)| \quad (2)$$

点 $(x,y)$ 的可能视差值中,使匹配代价最小的视差值即为要求的初始视差值。

#### 2.2 初始模板计算

TAO 提出一种立体匹配算法框架<sup>[4]</sup>:假设在单一的色彩区域中视差变化是平滑的,该区域可以通过一个平面模板公式(3)来描述:

$$d(x,y) = c_1x + c_2y + c_3 \quad (3)$$

其中, $c_1, c_2, c_3$  为平面模板参数, $(x,y)$  是图像像素的坐标, $d(x,y)$  为像素 $(x,y)$  的视差值。

$$\text{将式(3)改写成 } A [c_1 \ c_2 \ c_3]^T = B \quad (4)$$

式中矩阵  $A$  的第  $i$  行为  $[x_i \ y_i \ 1]$ , 矩阵  $B$  的第  $i$  行为  $d(x_i, y_i)$ , 模板参数  $c_1, c_2, c_3$  用最小二乘法计算。

由于初始视差值的精度对计算模板参数的影响很

大,因此在求解模板参数时只利用可靠点去求解。本文采用交叉检验法去检测可靠点:以左图像的某像素区域窗口作为模板,在右图像中进行匹配,计算的视差值记为  $d_{left}$ ; 然后以右图像该像素匹配点区域窗口作为模板,在左图像中进行匹配,计算的视差值记为  $d_{right}$ 。如果  $d_{left} = -d_{right}$ , 则该像素为可靠点,否则为不可靠点。

将计算出来的初始模板依次放入模板集合,每次都与集合内已存在的模板进行比对,如果集合中存在与当前模板的相同模板,则当前模板不放入模板集合,否则加入模板集合。一般来说,模板集合里的模板数量远远小于图像分割区域的数量。

#### 2.3 区域和模板优化

区域分割后,很可能将原属于一个模板的区域分割成若干个独立的区域,使得整个算法效率降低,因此需要把模板参数相同的相邻区域进行融合。具体步骤如下:

- (1) 计算出模板集合里每个模板对应区域的匹配代价;
- (2) 给每个分割区域分配一个匹配代价最小的模板 ID;
- (3) 对模板 ID 相同的相邻区域进行融合;
- (4) 计算融合后区域的匹配代价。

步骤(2)~(4)较容易实现,步骤(1)是算法实现的关键所在。一般求解最小匹配代价的公式为:

$$C(S,P) = \sum_{(x,y) \in S} c(x,y,d) \quad (5)$$

然而,这种算法并没有考虑到区域中不可靠点对匹配代价产生的偏差,因此需要对公式(5)进行修正。由于遮挡点一般出现在物体边界处,而这些地方都是高纹理区域,因此,有理由认为交叉检验法检测出的不可靠点只有出现在高纹理区域内才可能是遮挡点;而出现在低纹理区域内,则认为不是遮挡点,称为不支持点。对于如何划分高低纹理区域,本文采用了参考文献[11]中的判断方法,令:

$$g(x,y) = \frac{1}{9} \sum_{i=y-1}^{y+1} \sum_{j=x-1}^{x+1} \left[ \left( \frac{\partial I}{\partial x} \right)_{(i,j)} \right]^2 \quad (6)$$

预先设定阈值  $T_g$ , 当  $g(x,y) < T_g$  时,点 $(x,y)$ 即属于低纹理区域,否则属于高纹理区域。

假设  $n$  是区域  $S$  内非遮挡点的数量, $s$  是区域  $S$  内可靠点的数量, $O$  是  $S$  内的遮挡部分,则公式(5)变成:

$$C(S,P) = \sum_{(x,y) \in S-O} c(x,y,d) e^{1-\frac{s}{n}} \quad (7)$$

其中, $S$  为分割区域, $P$  为平面模板, $C(S,P)$  为平面模板  $P$  相对  $S$  的相似匹配代价, $d$  为模板  $P$  在点 $(x,y)$  处计算所得的模板视差, $c(x,y,d)$  为当点 $(x,y)$  视差为  $d$  时的匹配代价。

新的匹配代价公式既包含了非遮挡点的匹配代价之和,又包含了一个新的指数函数,使匹配代价可随着不支持点比例的改变而变化。在低纹理区域里,由于包含了不支持点的信息,从而扩大了计算面积,降低了模

## 图形、图像与多媒体

Image Processing and Multimedia Technology

板匹配代价的误差;在高纹理区域里,由于去除了遮挡部分,使得正确的模板与其他模板的匹配代价差异很大,提高了算法的精度。因此,改进的匹配代价公式特别适用于低纹理区域及遮挡区域。

## 3 模板视差最优分配

本文采用基于全局的算法进行模板视差最优分配,构造一个全局能量函数,将模板视差的最优分配问题转化为求解全局能量函数最小化问题<sup>[5]</sup>。常用的全局能量函数为:

$$E(f) = E_{\text{data}}(f) + E_{\text{smooth}}(f) \quad (8)$$

定义公式(8)中的数据项  $E_{\text{data}}(f) = \sum C(S, P)$

平滑项  $E_{\text{smooth}} = \sum_{s_i, s_j} u_{(s_i, s_j)} \delta(P_i \neq P_j)$

其中  $C(S, P)$  为公式(7)定义的函数。假设第  $i$  块区域与第  $j$  块区域为相邻区域,  $u_{(s_i, s_j)}$  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。

远远偏离实际的视差无助于降低优化函数的能量,只会增加算法的计算量。为了提高运算效率,本文根据初始匹配的计算结果,把初始视差范围作为能量函数的一个参考项,重新构造全局能量函数:

$$E(f) = E_{\text{data}}(f) + E_{\text{smooth}}(f) + E_{\text{range}}(f) \quad (9)$$

其中,视差范围约束项为:

$$E_{\text{range}}(f) = \begin{cases} 0 & \text{if } f(p) \in (f_i - \beta, f_i + \beta) \\ \infty & \text{else} \end{cases}$$

$\beta$  是允许的视差变化范围。

对上面建立的能量函数,采用 Boykov 的构图方法进行构图,并运用  $\alpha$  扩张局部优化逐步逼近全局最优的方法进行能量函数最小化<sup>[12]</sup>。最小能量所对应的标号即为最佳标号,也就是要求的视差。

每一次  $\alpha$  扩张的最小能量采用图割理论中的最大流/最小割算法来求解。首先将分割区域作为构造图中的节点,将模板视差作为构造图中边的容量,从而构建

三维网络,添加源  $s$  和汇  $t$ ,将从源到汇所有边的代价和的最小值转化为构造图的最小割。图像中每个区域都有一条属于最小割的视差边,该边表明分配的视差值。采用最大流/最小割算法求解构造图的最大流<sup>[6]</sup>,根据最大流/最小割定理(在任一网络中,最大流的值等于最小割的容量),也就获得了图的最小割,最终求解此次  $\alpha$  扩张的视差分布。

## 4 实验及其结果

为了检验本文算法的有效性,使用 Matlab 在普通 PC 机上对本文算法进行验证,并与采用 Layered 算法和 Belief 算法进行了比较,结果如图 1 所示。实验所采用的立体图像对来自于美国 Middlebury 大学计算机视觉研究中心提供的 Tsukuba 图像对,大小为  $288 \times 384$  (pixel)。Tsukuba 图像对的特点是包含比较复杂的场景,既有高纹理区域,也有低纹理的平滑区域,物体之间的边界相对较复杂,还有一些细长条的物体,如台灯的灯管,对于立体匹配算法来说都是个挑战。

从图 1 中可以看出,本文所得视差图在满足视差平滑约束的同时保留了细节信息,能够处理低纹理区域。为了更好地评价视差图,本文计算出参考文献[3]中根据误差大于 1 个像素计算的 3 个量化指标:  $B_{\text{unex}}$  (非纹理区域匹配错误百分比)、 $B_{\text{all}}$  (所有像素匹配错误百分比)和  $B_{\text{dis}}$  (视差不连续区域匹配错误百分比),并将本文算法与其他算法进行了比较,结果如表 1 所示。从表 1 中可以看出,本文算法的运行速度较快,实时性好,匹配准确率较高。除本文算法的数据外,文中其他数据均来自 <http://vision.middlebury.edu/stereo>。

本文提出了一种新的快速立体匹配算法,把 TAO 提出的基于图像分割的立体匹配算法和基于全局的算法相结合,利用可靠点计算分割区域的平面模板参数,并在计算最小匹配代价时充分考虑了不可靠点带来的误

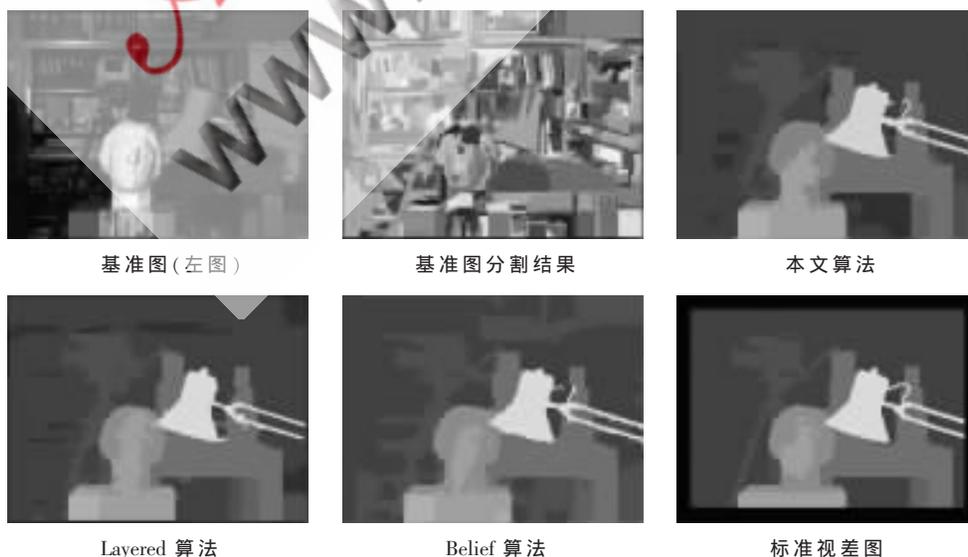


图 1 Tsukuba 图像对各种算法计算的视差图比较

表 1 不同立体匹配算法的评价结果

算法	$B_{\text{unrec}}/\%$	$B_{\text{all}}/\%$	$B_{\text{disc}}/\%$	运行时间/s
本文算法	1.21	0.26	6.71	1.016
Layered 算法	1.58	1.06	8.82	1.813
Belief 算法	1.15	0.42	6.31	1.348

差,提高了算法在低纹理区域和遮挡区域的匹配精度。在构造全局能量函数时,把初始视差范围作为一个参考项,降低了网络规模。利用图割理论中的最大流/最小割算法优化能量函数,用区域代替像素作为优化单元,大大提高了运算效率。实验结果表明,本算法对低纹理区域和遮挡区域有很好的匹配效果,且显著缩短了匹配时间,能够满足计算机视觉应用中高精度、高实时性的要求。

## 参考文献

- [1] 马颂德,张正友.计算机视觉—计算理论与算法基础[M].北京:科学出版社,1998.
- [2] 张爱武,李明哲,胡少兴.一种用于三维曲面视觉测量的立体精匹配方法[J].光学技术,2001,27(2):115-117.
- [3] HONG L, CHEN G. Segment-based stereo matching using graph cuts[C]. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004,(1):74-81.
- [4] TAO H, SAWHNEY H S, KUMAR R. A global matching framework for stereo computation[C].Proceedings of the 8th International Conference on Computer Vision, Vancouver, Canada,2001:532-539.
- [5] BLEYER M, GELANTZ M. Graph-cut-based stereo matching using image segmentation with symmetrical treatment of occlusions[J].Signal Processing:Image Communication,Special Issue on Three-Dimensional Video and Television,2007,22(2):127-143.

- [6] FORSTMANN S, OHYA J, KANOU Y. Real-time stereo by using dynamic programming[J].CVPR,2004,3(7):29-37.
- [7] KOLMOGOROV V, ZABIH R. Computing visual correspondence with occlusions using graph cuts[C]. Proceedings of the 8th International Conference on Computer Vision,Vancouver, Canada,2001:508-515.
- [8] BOYKOV Y, KOLMOGOROV V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in computer vision[J].IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004,26(9):1124-1137.
- [9] FELZENSZWALB P F, HUTTENLOCHER D P. Efficient graph-based image segmentation[J]. International Journal of Computer Vision,2004, 59(2):167-181.
- [10] ALI M, SHAH J R, AHMED M. 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.
- [11] SCHARSTEIN D, SZELISKI R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J].International Journal of Computer Vision, 2002,47(1): 7-42.
- [12] BOYKOV Y, VEKSLER O. Fast approximate energy minimization via graph cuts [J].IEEE Trans on Pattern Analysis and Machine Intelligence,2001,23(11):1222-1239.

(收稿日期:2010-01-22)

## 作者简介:

朱程辉,男,1959年生,副教授,硕士,主要研究方向:控制技术应用,图像处理与模式识别。

任冉冉,男,1984年生,硕士研究生,主要研究方向:检测技术与图像处理。