

# 基于边缘的多尺度立体匹配算法

夏林元<sup>1,2</sup> 肖俊<sup>1</sup> 林丽群<sup>3</sup>

(1 武汉大学测绘遥感信息工程国家重点实验室, 武汉市珞喻路 129 号, 430079)

(2 中山大学地理科学与规划学院, 广州新港西路 135 号, 510275)

(3 湖北大学资源与环境学院, 武汉市学院路 11 号, 430062)

**摘要:**提出了一种面向对象的多尺度递进的立体匹配算法, 实现了由粗到细的快速立体匹配。首先对参考图进行粗分割获取目标对象, 针对目标区域采用边缘线动态规划的方法获取视差, 根据精度需求进一步判断该目标是否需要细分以获取更加精细的视差, 得到有效的视差图。实验结果表明此方法快速有效。

**关键词:**立体匹配; 图像分割; 动态规划; 多尺度

**中图法分类号:** P237.3

立体匹配是视觉导航关键的一步, 国内外许多学者对立体匹配问题作了相当多的研究<sup>[1,2]</sup>。目前, 立体匹配算法按照基元的不同, 主要分为基于特征和基于区域两类。基于特征的方法处理对象主要为角点、边缘线等, 但其只能得到稀疏的视差图, 难以保证匹配的准确性和精度; 而基于区域的匹配方法又主要分为基于局部窗口与基于能量最小化法。基于窗口的方法由于窗口的选取存在着很大的局限性, 窗口太小会产生太多的不连续结构; 太大则会导致视差的过平滑, 所以效果相对较差, 但其容易在硬件上进行优化, 速度较快<sup>[3]</sup>。基于能量的方法通过构造能量函数, 搜索能量最小值得到视差图, 代表算法为图切割法<sup>[4-5]</sup>、置信度传播<sup>[6]</sup>和动态规划法<sup>[7-9]</sup>。能量法解得的视差图精度高, 但由于能量法需要进行全局优化, 速度较慢<sup>[10-11]</sup>。而动态规划是具有代表性的一维能量最小化方法, 具有速度快、容易硬件实现等优点, 并且还具有比窗口法更高的准确率<sup>[8]</sup>。Ohta<sup>[9]</sup>等提出了动态规划立体匹配算法, 在用区域匹配法求得初始匹配代价之后, 用动态规划进行全局优化, 使视差图的质量得到明显提高, 但由于传统的动态规划是基于扫描线进行处理, 其视差图结果容易产生条纹瑕疵。

图像分割的目的是为获取目标对象, 也是后

续视差求解的基础。本文算法中需要一种可以进行多尺度分割的算法, 因此, 采用自下而上的区域融合分割算法, 该算法在分割过程中模拟人类的感知过程, 充分考虑到形状、灰度和空间语义等信息, 通过对图像进行分割、区域合并, 对图像快速地进行分割, 获取符合人的视觉对象区域, 根据参数设置的不同可以进行多尺度分割。分割算法和分割尺度对算法的精度都存在影响<sup>[12-13]</sup>, 本文的多尺度视差提取能在一定程度上避免算法和尺度的影响。

## 1 多尺度的立体匹配方法

整个算法流程如图 1 所示。本文提出的多尺度立体匹配并非是全局多尺度<sup>[14]</sup>, 而是在满足条件的情况下, 对部分区域实行细分, 从而符合人类观察远处物体具有模糊性, 而近处物体则具有精确性这样的规律。这一点也满足立体视觉导航中的要求, 而且处理过程中, 以目标对象为视差求解对象可以避免逐像素求解。

整个算法包括以下几个关键过程: ① 分割; ② 目标边缘链码提取; ③ 以边缘链码进行动态规划, 求取边缘线视差; ④ 若该区域满足条件, 则需要进一步分割, 重复步骤①~步骤③直至满足

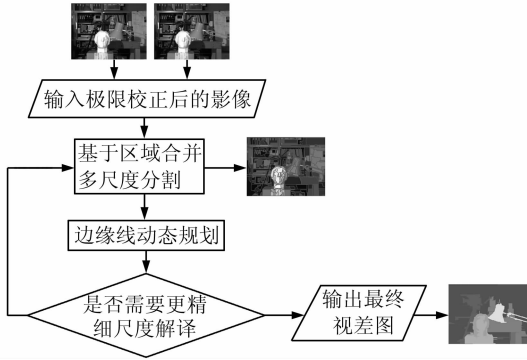


图1 立体匹配流程图

Fig. 1 Flow of Stereo Matching

条件,从而实现该区域的多尺度视差求解;⑤ 重复步骤②~步骤④过程,直至遍历整幅图像。整个算法中及涉图像分割和对象的视差获取两个重要内容。

将动态规划用于针对沿分割区域内边缘线求取视差,相比较求取两幅图像的视差空间图像,运算量大大减小。而且以分割区域的内边缘线为对象,求取视差精度高。在以边缘进行动态规划的过程中,将梯度等特征引入相似性测度计算。分割区域的边缘是梯度较明显的一条链码,在此算法中引入梯度可以增强动态规划相似性测度的有效性。这种方法具体包括以下3个步骤。

1) 边缘线视差空间的建立。为方便起见,本文假设输入图像都经过极线校正。以左图像作为参考图,并以  $d(x, y)$  表示参考图点  $(x, y)$  的视差值。视差值在视差范围内取整数值。

考虑到边缘线上的各点都为边缘点,这样左右图像上,同名点应具有同样的作用,即右图像上对应的同名点也应为边缘点。由此,综合考虑到图像的颜色信息与边缘特征,构成相似性测度函数  $C$ :

$$C_{\text{SAD}}(x, y, d) = \sum_{(i, j) \in N(x, y)} \sum_{k \in \{r, g, b\}} |I_i^k(i, j) - I_r^k(i - d, j)| \quad (1)$$

$$C_{\text{GRAD}}(x, y, d) = \sum_{(i, j) \in N(x, y)} \sum_{k \in \{r, g, b\}} |\nabla_x I_i^k(i, j) - \nabla_x I_r^k(i - d, j)| + \sum_{(i, j) \in N(x, y)} \sum_{k \in \{r, g, b\}} |\nabla_y I_i^k(i, j) - \nabla_y I_r^k(i - d, j)| \quad (2)$$

式中,  $I_i^k(i, j)$ 、 $I_r^k(i, j)$  表示像素点  $(i, j)$  在色彩通道  $k$  的值;  $N(x, y)$  是一个以  $(x, y)$  作为中心点的  $3 \times 3$  窗口;  $\nabla_x$ 、 $\nabla_y$  表示在  $x$ 、 $y$  方向上的梯度。因此,考虑到针对边缘点的动态规划处理,引入梯度算子,对求取相似性更加有效。构建相似

性测度函数  $C$ :

$$C(x, y, d) = (1 - \rho)C_{\text{SAD}}(x, y, d) + \rho C_{\text{GRAD}}(x, y, d) \quad (3)$$

当对应的像素匹配(也就是互为同名点)时,它的相似性度量函数应为最小。

2) 边缘线上的动态规划。沿着分割区域建立了边缘的视差空间图后,沿着边缘线进行动态规划,以搜索出边缘线上各点的最优视差。建立包括相容性约束和平滑性约束的能量评价函数为:

$$E(d) = E_{\text{data}}(d) + \lambda E_{\text{smooth}}(d) \quad (4)$$

式中,  $E_{\text{data}}$  代表匹配单元沿着边缘线视差空间的惩罚值;  $E_{\text{smooth}}$  是相邻单元视差值差距的平滑值;  $\lambda$  为相容性约束和平滑性约束的比例系数。目的就是计算出边缘线上所有点的视差值,使得目标函数取值最小。

$$E_{\text{smooth}}(d) = S(d_p, d_q) = \begin{cases} 0, & d_p = d_q \\ \omega_{pq}, & \text{其他} \end{cases} \quad (5)$$

式中,  $p$ 、 $q$  为参考图像中边缘线上相邻的两个像素点;  $d_p$ 、 $d_q$  是赋予这两个像素点的视差值。这个  $S$  常常用于基于能量的立体匹配<sup>[9]</sup>, 源于马尔科夫随机场的波茨模型。如果参考图像中相邻像素拥有相同视差值,则没有惩罚。如果相邻像素的视差值不同,则惩罚值  $\omega_{pq}$  就由  $|d_p - d_q|$  决定。

3) 视差赋值。对于区域的视差,充分考虑了人类视觉对其的影响,即远处物体和近处物体的重要程度是不同的。因此,本文引入了以下两个条件来实现多尺度的视差求解:① 若区域边缘线视差均值和视差的方差小于给定阈值则不再作进一步处理,考虑到速度,选用最简单快捷的投票法计算边缘线所在分割区域的视差值。统计边缘线上各点视差值,选取出现次数最多的值为分割区域的视差值。② 若区域边缘线视差大于给定阈值,且面积大于一定情况下,则需要将该区域进行更细尺度的区域分割,然后采用上述方法用动态规划的方式搜索边缘线视差求解,从而求解出更加精确的视差。这种条件下,即认为近处物体的视差是更为关注的对象。在粗尺度分割的情况下,很容易将近处物体与其他物体划分为一个区域。如果该区域的整体视差较大,则需要进一步分割,将该区域划分为若干个更小的区域,进一步获取小区域的精确视差。

## 2 实验结果与分析

本文实验均在 CPU 为酷睿双核 2.8 GHz 的 PC 机上运行,操作系统为 Windows7,编程工具

为 VC++ 6.0,为验证本算法的正确性,选用 middlebury 图片库<sup>[15]</sup>的 Tsukuba and Teddy 图片进行实验。这些立体图满足平行立体视觉测距原理和外极线约束条件,且无噪声。

本文实验系用图 Tsukuba 和 Teddy 的参考图如图 2(a)所示。图中仅用了两个尺度来获取视差,图 2(d)为粗分割图,作为基准进行后续的视差求取,图 2(e)为进一步分割后的视差结果图。

从表 1 来看,本文算法的结果明显优于传统基于像元的动态规划结果,对大部分区域能够有效求取其视差;而且从整个运算过程来看,仅以区域的边缘为计算单元,有效地减少了计算数量,减少了算法的复杂度,提高了运算时间。实验中没有直接以细尺度分割来求解视差,这样能够加速算法的运行。

针对 Tsukuba 的实验,图 2(d)左图为第一个分割尺度对应的视差图,其中大部分区域的视差精度都很高,除了台灯所在位置。这主要是因为台灯在这个尺度中被分为一个整体,根据区域视差求取的条件,该区域会被进一步分割,在下一个尺度中进一步获取各个区域的视差,因此,最终的

多尺度分割的视差结果图 2(e)左图所示,台灯区域的视差能够较好地获取;同时图中摄像机的位置在多尺度的结果中比单一尺度要精确。

从 Teddy 实验结果看,图 2(b)右图所示的粗尺度分割。对近处物体区分不够,近处物体在粗尺度下与其他物体分割成一个区域,而在求解视差的过程中,近处物体的视差求解也不够准确。根据限定条件对近处的区域进行了细分割,最终得到图 2(e)右图所示的较为精确的视差结果。该实验中近处和远处物体的视差能够较为准确地体现,对于有一致视差的区域能够很好地区分。该图的真实视差图 2(c)右图中大部分区域的视差均为渐变,而本文算法以整体为对象,认为整个区域的视差是相同的,因此结果上与真实值之间有一定的出入,但与渐变的范围差异不大,并不影响视差图的精度。目标区域内部视差应该是一致或是渐变的,而且渐变的范围不会太大。如果区域渐变太大,在影像分割的过程中,也会被分割成多个区域而非一个区域,同时多尺度的获取视差中也能获取更精确的视差。从实际用途考虑,算法能够大致体现视差的渐变过程,也能够满足精度的需求。

表 1 不同算法测试结果比较

Tab. 1 Comparison of Different Algorithms

算法	Tsukuba			Venus			Teddy			Cones		
	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc
本文算法	2.36	3.33	10.1	0.83	0.96	4.98	7.42	10.7	12.4	7.44	12.7	14.7
DP	3.43	4.23	9.85	6.50	7.43	17.4	7.11	14.9	13.4	6.52	15.1	15.1
SO	4.23	6.21	10.7	5.14	6.58	17.0	11.3	20.2	18.6	8.76	18.8	16.1

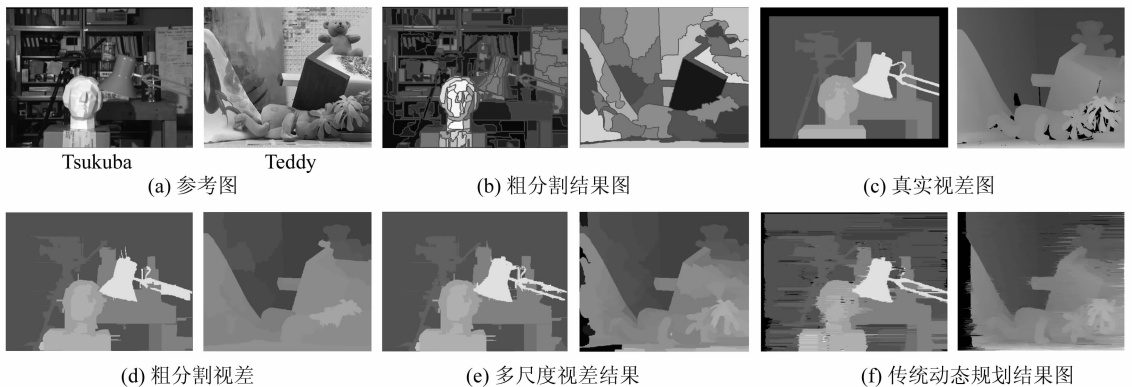


图 2 实验的结果图

Fig. 2 Results of Experimental

## 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, 2001, 47(1/2/3):7-42
- [2] Brown M Z, Burschka D, Hager G D. Advances in Computational Stereo [J]. IEEE Trans Pattern Anal. Mach Intell, 2003, 25(8): 993-1 008
- [3] Kande T, Okutomi M. A Stereo Matching Algorithm with an Adaptive Window, Theory and Experiment[J]. IEEE Transaction on Pattern Analysis

- and Machine Intelligence, 1994, 16(9):920-932
- [4] Hong L, Chen G. Segment-based Stereo Matching Using Graph Cut[C]. Conference on Computer Vision and Pattern Recognition, Washington, D C, 2004
- [5] Boykov Y, Veksler O, Zabih R. Fast Approximate Energy Minimization Via Graph Cuts[R]. IEEE Conference on Computer Vision and Pattern Recognition, Tech Rep, 2003
- [6] Sun J, Shum Y H, Zheng N N. Stereo Matching Using Belief Propagation[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2003, 25(7):787-800
- [7] Bobick A F, Intile S S. Large Occlusion Stereo[J]. International Journal of Computer Vision, 1999, 33(3): 181-200
- [8] Abdollahifard M, Faez K, Pourfard M. Fast Stereo Matching Using Two Stage Color-based Segmentation And Dynamic Programming[C]. The 6th International Symposium on Mechatronics and Its Application, Tehran, Iran, 2009
- [9] Ohta Y, Kanade T. Stereo by Intra-and Inter-scanline Search Using Dynamic Programming[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1985(7): 139-154
- [10] Bleyer M, Gelautz M. A Layered Stereo Matching Algorithm Using Image Segmentation and Global Visibility Constraints[J]. Photogrammetry and Remote Sensing, 2005, 59:128-150
- [11] Tao H, Sawhney H. Global Matching Criterion and Color Segmentation Based Stereo[C]. IEEE Workshop on Applications of Computer Vision, Palm Springs, CA, USA, 2000
- [12] 沈占锋, 骆剑承, 胡晓东, 等. 高分辨率遥感影像多尺度均值漂移分割算法研究[J]. 武汉大学学报·信息科学版, 2010, 35(3):313-316
- [13] Bruzzone L, Carlin L. A Multilevel Context-Based System for Classification of Very High Spatial Resolution Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2006, 44:2 587-2 600
- [14] 林丽群, 舒宁, 龚龔, 等. 基于像斑的多光谱影像跨尺度分类研究[J]. 武汉大学学报·信息科学版, 2009, 34(1):40-43
- [15] Scharstein D, Szeliski R. Middlebury Stereo Vision Research Page [OL]. <http://vision.middlebury.edu/stereo>, 2011

第一作者简介:夏林元,教授,博士生导师,现主要从事 GNSS 的教学与研究。

E-mail: xialiny@mail.sysu.edu.cn

## Segment-based Stereo Matching Using Edge Dynamic Programming

XIA Linyuan<sup>1,2</sup> XIAO Jun<sup>1</sup> LIN Liqun<sup>3</sup>

(1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing,

Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

(2 School of Geographical Science and Planning, Sun Yat-Sen University, 135 Xingangxi Road, Guangzhou 510275, China)

(3 School of Resources and Environmental Science, Hubei University, 11 Xueyuan Road, Wuhan 430062, China)

**Abstract:** Most current stereo matching method is based on pixel processing, but object-oriented approach is more conform to the human visual habit. This paper put forward a scale-span multi-scale stereo matching method, achieved by coarse to fine stereo matching. Firstly the reference image is rough segmented by an object-oriented segmentation method, then by employing a refined DP algorithm which is based on the inner edgeline of the segment, disparity are obtained. According to the accuracy demand, determine whether need to do the refined segmentation. Finally, the final disparity images are obtained. The experimental results indicate that the algorithm has high matching precision and fast calculating speed.

**Key words:** stereo matching; image segmentation; dynamic programming; scale-span