

DOI:10.19651/j.cnki.emt.2108392

顾及图像分割信息的半全局立体匹配算法研究*

李聪聪^{1,3} 方勇^{2,3} 王芮^{1,3} 王振磊^{1,3}

(1.长安大学 西安 710054; 2.西安测绘研究所 西安 710054; 3.地理信息工程国家重点实验室 西安 710054)

摘要: 针对传统半全局立体匹配算法(SGM)在高分辨率图像弱纹理以及视差不连续区域的误匹配问题,提出了一种顾及图像分割信息的SGM算法。该算法在代价计算阶段,首先根据图像分割信息自适应调整匹配窗口大小,采用不同状态信息的改进Census变换计算初始代价,解决传统算法对Census变换窗口中心像素依赖的同时减少了匹配时间;在代价聚合阶段,将图像分割信息与传统SGM算法的全局能量函数进行有机结合,提高算法在弱纹理以及深度不连续区域的匹配效果;最后通过左右一致性检测和子像素细化得到优化后的视差图。所提算法利用Middlebury平台标准数据进行验证,实验结果表明,平均误匹配率为4.54%,与传统SGM算法和一些改进算法相比,该算法能够在影像弱纹理和视差不连续区域获得更高匹配正确率。

关键词: 半全局立体匹配;图像分割;Census变换;全局能量函数;左右一致性检测

中图分类号: TP751 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Research on semi-global stereo matching algorithm considering
image segmentation informationLi Congcong^{1,3} Fang Yong^{2,3} Wang Rui^{1,3} Wang Zhenlei^{1,3}(1. Chang'an University, Xi'an 710054, China; 2. Xi'an Research Institute of Surveying and Mapping, Xi'an 710054, China;
3. State Key Laboratory of Geo-information Engineering, Xi'an 710054, China)

Abstract: Aiming at the mismatching problem of traditional semi-global matching (SGM) in high-resolution images with weak textures and disparity discontinuities, an SGM algorithm that takes into account the image segmentation information is proposed. In the cost calculation stage of this algorithm, the size of the matching window is first adaptively adjusted according to the image segmentation information, and the improved Census transform with different state information is used to calculate the initial cost, which solves the traditional algorithm's dependence on the center pixel of the Census transform window and reduces the matching time. In the cost aggregation stage, the image segmentation information is organically combined with the global energy function of the traditional SGM algorithm to improve the matching effect of the algorithm in weak texture and depth discontinuous regions. Finally, the optimized disparity map is obtained through left-right consistency detection and sub-pixel refinement. The proposed algorithm is verified by using the standard data of the Middlebury platform. The experimental results show that the average mismatch rate is 4.54%. Compared with the traditional SGM algorithm and some improved algorithms, the proposed algorithm can obtain higher matching accuracy in the weak texture and discontinuous disparity areas of the image.

Keywords: semi-global stereo matching; image segmentation; Census transform; global energy function; left and right consistency detection

0 引 言

立体匹配技术作为摄影测量与计算机视觉领域中的一个热点研究方向,其用处不只是用于摄影测量中所涉及的数字表面模型、数字正射模型,还有三维重建、物体跟踪与

检测、机器人导航等。它的本质是通过获取影像对中的匹配点来从二维图像中恢复三维信息。通常,立体匹配算法主要分为两类:局部立体算法和全局立体算法^[1]。局部立体算法主要是在图像上建立一个匹配窗口来进行视差计算,这类算法虽然计算速度较快,但精度较低。全局立体算

收稿日期:2021-11-18

* 基金项目:国家科技重大专项(GFZX04040202)资助

法主要利用能量函数最小化来计算视差,虽然最终的匹配精度较高,但是效率较低,不利于在实际生产中使用^[2,4]。Hirschmuller^[5]提出了一种半全局立体匹配算法,由于该算法同时满足精度和效率这两方面的要求,因此受到许多研究人员的关注。

针对传统 SGM 算法在立体像对弱纹理以及视差不连续区域的匹配精度不高的问题,许多研究人员对此进行了改进。一方面,一些学者将特征匹配得到的匹配点引入立体匹配过程,吴军等^[6]将尺度不变特征转换 SIFT 融入半全局立体匹配,以获取的特征匹配点为约束来对 SGM 算法聚合路径进行分段纠正。王阳萍等^[7]将加速鲁棒特征与 SGM 算法进行结合,用得到的特征点主方向来对聚合路径进行权重赋值。这些算法虽然在一定程度上解决了弱纹理和视差不连续区域的匹配精度问题,但极大增加了计算复杂度,且可能因为错误的特征匹配点产生不好的效果。另一方面,有研究者用图像边缘特征来优化立体匹配过程,Rothersmel 等^[8]采用 Canny 算子检测边缘,并在边缘处减小惩罚参数,从而达到保持边缘的效果。孙丽等^[9]在计算边缘视差图的基础上,利用分段代价聚合来优化 SGM 算法。但是仅利用梯度信息获取物体边缘,仍对噪声敏感,且检测到的边缘由单像素连接而成,实际上很难保证准确定位。朱庆等^[10]采用局部窗口的梯度和标准差提取纹理信息,并根据纹理信息来调整惩罚参数,在一定程度上减小了灰度噪声的影响,但是仅利用局部信息定量表达纹理特征仍存在局限性。

为了提高高分辨率立体图像在弱纹理以及视差不连续区域的匹配精度,本文提出了一种顾及图像分割信息的半全局立体匹配算法,与传统 SGM 算法相比,该算法首先利用自适应窗口的改进 Census 变换计算匹配代价,然后通过全局能量函数与图像分割信息有机结合来优化代价聚合,最终有效地提高了算法在弱纹理和视差不连续区域的匹配正确率。

1 算法原理

本文算法流程如图 1 所示,以生成的核线影像对作为输入,经过图像分割、代价计算、代价聚合、视差计算及视差优化等步骤得到最终的视差图。首先,本文在代价计算过程中对传统的 Census 变换算法进行改进,并且引入自适应窗口,相较于传统 SGM 算法,这在降低运行时间的同时增强了算法的匹配效果。然后在代价聚合阶段将图像分割信息与传统 SGM 算法的全局能量函数进行有机结合,通过改变在视差突变区域的惩罚项来使其更符合实际情况。视差计算采用 WTA 策略,得到初始视差图。视差优化通过左右一致性检查以及子像素细化等方法来进一步剔除错误视差以及提高视差准确度,最终获得优化后的视差图。

1.1 图像分割

图像分割通常是采用颜色、纹理、轮廓等作为分割依据

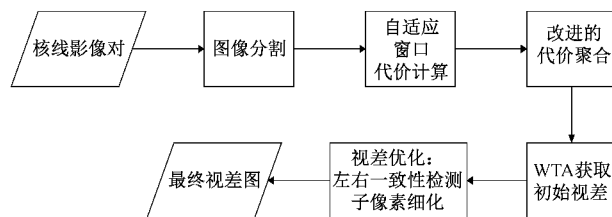


图1 本文算法流程

来将整幅图像分割成不同特性的块状区域。因为图像分割会保留图像中的边缘信息,所以在立体匹配中引入图像分割信息,可提高立体像对在弱纹理以及视差不连续区域的匹配精度^[11]。

MeanShift^[12]是一种利用概率分布的梯度寻找分布峰值的非参数估计方法,由于该算法稳定性、鲁棒性较好,并且不需要预先给定分割块的个数,因此本文采用该算法来获取图像分割信息。实验数据为 Middlebury 公开测试平台提供的 Teddy 图像,分割结果如图 2 所示。



图2 MeanShift 分割结果

图2 MeanShift 分割结果

在图像分割过程中,用核函数 K_{h_s, h_r} 来估计像素点 x 的分布, K_{h_s, h_r} 具体形式如式(1)所示。

$$K_{h_s, h_r}(x) = \frac{c}{h_s^2 h_r^2} k\left(\left\|\frac{x^s}{h_s}\right\|^2\right) k\left(\left\|\frac{x^r}{h_r}\right\|^2\right) \quad (1)$$

式中: x^s 表示像素点的坐标, x^r 表示在点 x^s 的颜色值, h_s 、 h_r 表示带宽, c 为常数。

1.2 代价计算

匹配代价计算的目的是为了衡量匹配点之间的相似性。Census 变换^[13]因鲁棒性较好且计算效率较高而被广泛用于立体匹配中。其原理是通过固定矩形大小的窗口遍历图像,使用像素邻域内的局部灰度差异将像素灰度转换为比特串,最后通过计算比特串间的 Hamming 距离得到像素点间的匹配代价。

由于传统的 Census 变换依赖于中心点像素的灰度值,十分容易受到噪声的干扰,因此仅仅依靠中心点的灰度值进行 Census 变换会较大影响到匹配精度。为了提高匹配效率和质量,本文提出了一种基于中心对称的改进 Census 变换来计算匹配代价,即只比较窗口内中心对称的像素对。与传统 Census 变换相比,改进的 Census 算法只需要 12 位比特串来存储 5×5 窗口需要的数据,大大降低了匹配所需的内存和匹配时间。具体代价计算原理如图 3 所示。

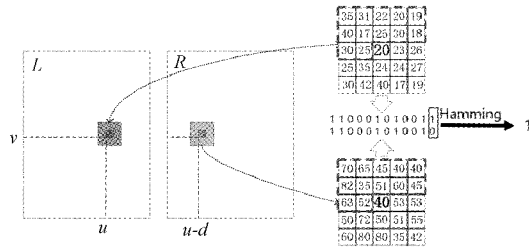


图 3 代价计算原理

在进行代价计算时,选择合适的窗口大小是一个关键问题,窗口尺度必须足够大,以包含足够的强度变化来获得可靠匹配,但又必须足够小,以避免干扰信息的影响。如果窗口太小,没有覆盖足够的强度变化,将会得到一个较差的视差估计。另一方面,如果窗口太大,又会导致计算速度明显降低^[14]。本文提出了一种基于图像分割信息来自适应选择合适窗口的方法,通过该窗口内部其他像素与中心像素位于同一个分割块的个数来确定 Census 变换窗口大小。自适应窗口算法的具体流程如图 4 所示。

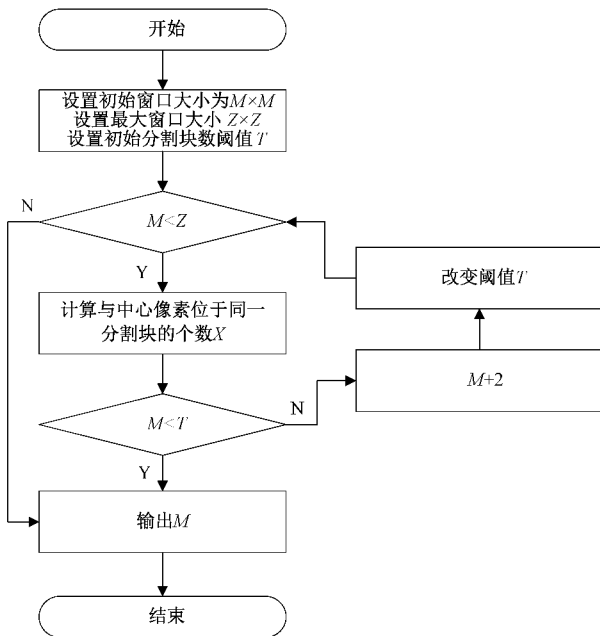


图 4 自适应窗口算法流程

1.3 代价聚合

本文算法是在传统 SGM 算法的基础上结合了图像分割信息来进行代价聚合步骤。其中传统 SGM 算法是通过建立邻接像素之间的关系来构造全局能量函数,然后使得能量函数最小化来完成代价聚合步骤。定义的能量函数公式如下:

$$E(D) = \sum_p C(p, D_p) + \sum_{q \in N_p} P_1 T[|D_p - D_q| = 1] + \sum_{q \in N_p} P_2 T[|D_p - D_q| > 1] \quad (2)$$

$$T[\text{boolean}] = \begin{cases} 1, & \text{true} \\ 0, & \text{false} \end{cases} \quad (3)$$

式中: $C(p, D_p)$ 表示点 p 在视差值为 D_p 时的匹配代价值, D_p, D_q 分别表示点 p 和点 q 的视差值, P_1, P_2 代表着不同的惩罚项, $T[\]$ 为判断函数。对于点 p 与点 q 的视差变化等于 1 时采用 P_1 惩罚项,视差变化大于 1 时采用 P_2 惩罚项。

然而在图像中的视差非连续区域通常存在较大的视差变化,在这些区域给定一个较大的惩罚项是不合理的,因为这会导致在视差结果过度平滑,从而降低算法的匹配精度。针对这个问题,本文在 SGM 算法的基础上加入了图像分割信息,通过改善视差非连续区域的惩罚项来优化全局能量函数。具体就是通过图像分割信息找到视差边缘区域,然后给定这些区域相对较小的惩罚项系数。优化后的全局能量函数如式(4)所示。

$$E(D) = \sum_p C(p, D_p) + \sum_{q \in N_p} P_1 T[|D_p - D_q| = 1] + \sum_{q \in N_p} \sigma P_2 T[|D_p - D_q| > 1] \quad (4)$$

$$\sigma = \begin{cases} \sigma_1, & \text{相同区域} \\ \sigma_2, & \text{不同区域} \end{cases} \quad (5)$$

式中: σ 表示惩罚值, σ_1, σ_2 是给定的阈值。

上式的能量函数最优化问题是一个 NP 完全问题,SGM 算法将其分解为不同路径上的代价聚合来高效解决此问题。具体计算方法如式(6)所示。

$$L_r(p, d) = C(p, d) + \min \begin{cases} L_r(p-r, d) \\ L_r(p-r, d-1) + P_1 \\ L_r(p-r, d+1) + P_1 \\ \min_i L_r(p-r, i) + P_2 \end{cases} \quad (6)$$

式中:下标 r 表示聚合路径, $L_r(p-r, d)$ 表示在聚合路径 r 方向上前一个点 $p-r$ 的视差值为 d 时的聚合代价值, $\min_i L_r(p-r, i)$ 表示聚合路径 r 方向上前一个点 $p-r$ 的所有代价值的最小值。

像素的总路径代价值 S 则通过式(7)计算:

$$S(p, d) = \sum_r L_r(p, d) \quad (7)$$

1.4 视差图生成

在视差计算步骤中,采用 WTA 策略生成初始视差图,即通过选取匹配代价值最小的点对应的视差为初始视差。具体公式表示为:

$$D(p) = \underset{d}{\operatorname{argmin}} S(p, d) \quad (8)$$

式中: $D(p)$ 表示像素 p 的初始视差。

1.5 视差图优化

为了得到更可靠的视差值,本文采用左右一致性检测以及子像素细化来对视差图做进一步的优化。左右一致性检查主要是根据匹配点唯一性的约束来剔除错误匹配点,其公式如下所示:

$$D_p = \begin{cases} D_{bp}, & |D_{bp} - D_{mq}| \leq 1 \\ D_{invalid}, & \text{其他} \end{cases} \quad (9)$$

式中: D_p 、 D_{bp} 表示左视图的点 p 对应的视差, D_{mq} 表示该点在右视差图中对应点的视差值。

由于前面计算出来的视差值是整数级的,无法满足实际应用中一些精度要求,因此采用子像素细化技术,通过二次曲线拟合来获取子像素精度,具体计算方法如式(10)所示。

$$d_{sub} = d_p + \frac{c_1 - c_2}{2(c_1 + c_2 - 2c_0)} \quad (10)$$

其中, d_{sub} 表示子像素拟合后的视差值, d_p 表示点 p 对应的视差值, c_0 表示点 p 在视差为 d 时对应的代价值, c_1 和 c_2 分别是像素 p 在视差 $d \pm 1$ 时对应的代价值。

2 实验验证及分析

为了验证算法的有效性,本文利用 Middlebury 公开测试平台提供的四对标准测试图像 Venus、Teddy、Cone 和 Tsukuba 来进行实验,实验环境是 Visual Studio 2015 下 C++ 编程环境,Windows10, Intel Core i7-6700HQ 处理器,12 GB 内存。实验参数设置如表 1 所示。

表 1 实验参数设置

参数	Z	M	P_1	P_2	σ_1	σ_2
数值	3	11	10	150	1.25	0.75

2.1 算法性能验证

为了验证算法的性能,本文选取了传统的 SGM 算法进行实验对比,实验结果如图 5 所示,其中图 5(c)、(d)两列中的白色方框标明了本文算法的明显改进区域。首先进行定性分析,针对具有较简单场景特征的 Venus 和 Tsukuba 而言,本文算法在白色方框内实现了边界更清晰、轮廓更平滑的匹配效果,表明了本文算法能够有效增强视差不连续区域的匹配效果;同样针对复杂场景特征,本文算法也成功匹配出了 Cone 中具有重复纹理变化的圆锥区,以及 Teddy 中的视差突变区。然后再进行了定量分析,非遮挡区域和全部区域的误匹配率以及 Debug 模式下的运行时间如表 2 所示,可以很明显地看出,本文算法在 4 幅图像对上的平均误匹配率仅为 4.54%,相比于传统算法降低了 7%,这也证明出本文算法在保持时间基本不变的前提下,显著降低了视差图的误匹配率。

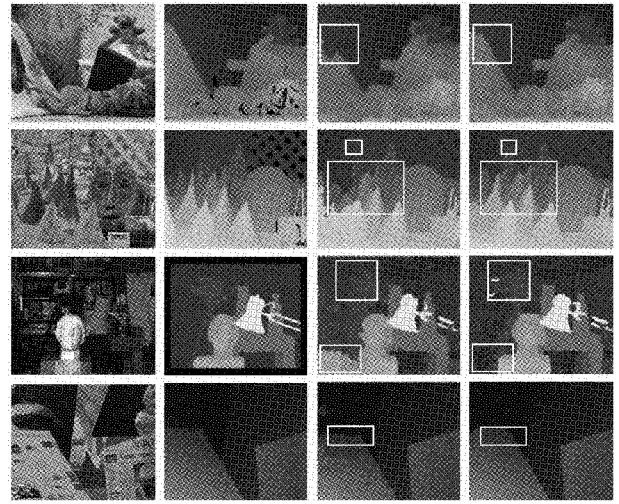


图 5 实验结果对比

表 2 本文算法与传统算法对比

测试图像	传统 SGM 算法			本文算法		
	Nocc/ %	All/ %	时间/ s	Nocc/ %	All/ %	时间/ s
Teddy	13.2	20.0	70.19	5.22	9.70	72.98
Venus	7.16	8.00	24.59	0.58	1.35	25.76
Cone	10.8	17.0	65.89	2.84	8.64	66.86
Tsukuba	7.48	8.62	12.89	3.31	4.71	14.76

为了进一步客观分析本文算法的性能,本文还选取了一些改进 SGM 算法^[15-19]加以对比,不同算法的误匹配如表 3 所示。对于 Venus 和 Tsukuba 这两幅图而言,本文算法的平均误匹配率与表中最优的算法接近,侧面反映了本文算法能够在简单场景特征图像中获得较好的匹配效果。对于 Teddy 和 Cone 这种具有复杂场景的图像来说,本文算法在这两幅图像上的平均误匹配率最低,明显优于其他算法,这也表明了本文算法能够较好地适用于复杂场景。总体来看,本文算法在这 4 幅图像对中的平均误匹配率最低,算法性能明显优于表中其他 SGM 改进算法。

表 3 本文算法与其他改进算法对比

算法	Teddy		Venus		Cone		Tsukuba		平均误差/ %
	Nocc	All	Nocc	All	Nocc	All	Nocc	All	
Ada_SGM	15.68	19.88	4.09	4.23	14.13	18.31	10.30	10.54	12.14
Seg_SGM	7.52	7.64	0.88	1.15	8.25	11.74	2.27	2.86	5.29
Cen_SGM	7.40	11.80	1.70	2.10	5.10	6.40	2.80	3.10	5.05
SGMDDW	6.50	14.50	1.20	2.70	5.60	14.20	2.30	4.40	6.42
planeFitSGM	5.70	11.60	1.10	1.90	3.80	9.30	3.10	4.20	5.09
本文算法	5.22	9.70	0.58	1.35	2.84	8.64	3.31	4.71	4.54

2.2 算法实用性验证

为了验证本文算法在实际遥感影像中的有效性,采用航空大面阵影像来进行测试,其中航空大面阵影像数据由焦距为 182 mm 的 DMZ II 相机获取,影像的画幅大小为 25 080 pixel×23 466 pixel,航向重叠率和旁向重叠率分别为 87%和 16%,地面分辨率为 0.25 m,航高为 8 000 m。

为了体现对于不同场景的匹配准确性,实验选取纹理丰富但是重复性较高的树林区、纹理单一但是视差连续性好的农田区、弱纹理区及视差非连续区域较多的住房区等 3 类场景。实验结果如图 6 所示,由第 1 行视差图可以看出,与传统 SGM 算法相比,本文算法有效减少了树林区内遮挡区域附近的错误视差。由第 2 行可得,本文算法在农田区域的匹配效果与传统算法相当,但由于本文算法降低了视差非连续区域的惩罚项,所以在路两旁的非连续视差区域明显优于传统算法。对比第 3 行视差图可知,自适应窗口的应用使得本文算法在马路等弱纹理区域的匹配效果较传统算法有着明显地提升。

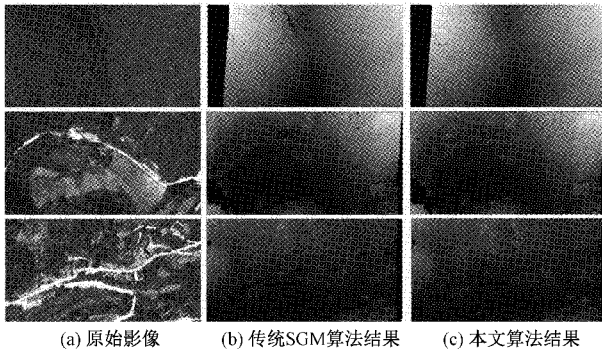


图 6 大面阵影像视差图

3 结 论

本文提出了一种顾及图像分割信息的半全局立体匹配算法。该方法采用 MeanShift 算法得到图像分割信息,然后充分利用图像分割信息优化代价计算和代价聚合步骤,通过自适应窗口的改进 Census 变换以及融合图像分割信息的全局能量函数使算法具有较强鲁棒性,同时提高了立体像对在弱纹理以及视差不连续区域的匹配效果。通过对比实验可知,本文算法在基本不增加时间成本的基础上,明显降低了影像的误匹配率。最后通过实际对航空大面阵影像的实验,证明本文算法在实际应用中的有效性。

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):7-42.
- [2] VEKSLER O. Stereo correspondence by dynamic programming on a tree [C]. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005:384-390.
- [3] FELZENSZWALB P F, HUTTENLOCHER D P. Efficient belief propagation for early vision [J]. *International Journal of Computer Vision*, 2006, 70(1):41-54.
- [4] KOLMOGOROV V, ZABIH R. Computing visual correspondence with occlusions using graph cuts [C]. *Proceedings Eighth IEEE International Conference on Computer Vision ICCV 2001*, 2001:508-515.
- [5] HIRSCHMULLER H. Accurate and efficient stereo processing by semi-global matching and mutual information [C]. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005:807-814.
- [6] 吴军,姚泽鑫,程门门.融合 SIFT 与 SGM 的倾斜航空影像密集匹配[J]. *遥感学报*, 2015, 19(3):431-442.
- [7] 王阳萍,秦安娜,郝旗,等.结合加速鲁棒特征的遥感影像半全局立体匹配[J]. *光学学报*, 2020, 40(16):163-171.
- [8] ROTHERMEL M, WENZEL K, FRITSCH D, et al. SURE: Photogrammetric surface reconstruction from imagery [C]. *LC3D Workshop*, 2012:1-9.
- [9] 孙丽,孙彦玮,李浩.视差边缘优化的 SGM 密集深度估计算法 [J]. *计算机与数字工程*, 2020, 48(11):2729-2733.
- [10] 朱庆,陈崇泰,胡翰,等.顾及纹理特征的航空影像自适应密集匹配方法 [J]. *测绘学报*, 2017(1):62-72.
- [11] 王书侠,王国新,康辉.基于图像分割的改进 Census 变换立体匹配算法 [J]. *高师理科学刊*, 2021, 41(4):35-39.
- [12] COMANICIU D, MEER P. Mean shift: A robust approach toward feature space analysis [J]. *IEEE Trans Pattern Analysis & Machine Intelligence*, 2002, 24(5):603-619.
- [13] ZABIH R. Non-parametric local transforms for computing visual correspondence [C]. *European Conference on Computer Vision*, 1994:151-158.
- [14] 路乾坤,李彦.基于自适应窗口与权重的立体匹配算法 [J]. *电子测量技术*, 2019, 42(14):117-122.
- [15] 黄超,赵华治.根据灰度值信息自适应窗口的半全局匹配 [J]. *中国图象图形学报*, 2019, 24(8):1381-1390.
- [16] 鲁光明,王竞雪.结合图像分割的半全局立体匹配算法 [J]. *遥感信息*, 2020, 35(6):85-91.
- [17] 张晶,王竞雪.结合颜色信息的 Census 变换半全局立体匹配算法 [J]. *测绘科学*, 2019, 44(12):147-152.
- [18] MICHAEL M, SALMEN J, STALLKAMP J, et al. Real-time stereo vision: Optimizing semi-global

matching [C]. 2013 IEEE Intelligent Vehicles Symposium(IV), 2013:1197-1202.

- [19] HUMENBERGER M, ENGELKE T, KUBINGER W. A census-based stereo vision algorithm using modified semi-global matching and plane fitting to improve matching quality[C]. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, 2010:77-84.

作者简介

李聪聪, 硕士研究生, 主要研究方向为立体匹配、三维

重建。

E-mail:1183071645@qq.com

方勇(通信作者), 研究员, 主要从事航空航天遥感精确测绘方面的研究。

E-mail:yong.fang@vip.sina.com

王芮, 硕士研究生, 主要研究方向为多源点云数据融合。

E-mail:1326075550@qq.com

王振磊, 硕士研究生, 主要研究方向为激光卫星测高系统数据处理与精度分析。

E-mail:2605559175@qq.com