

# 基于改进匹配代价和均值分割的 最小生成树立体匹配算法

王道累, 韩洋\*

上海电力大学能源与机械工程学院, 上海 200090

**摘要** 尽管传统的立体匹配模型在精度和鲁棒性方面都表现出了良好的性能,但在弱纹理和深度不连续区域的视差精度问题依然存在。针对上述问题,提出了一种基于改进匹配代价和均值分割的最小生成树立体匹配算法。首先,在匹配代价计算阶段,通过 Census 变换进行初始匹配代价计算,利用 Sobel 算子对输入图像进行边缘信息提取,将提取后的图像边缘信息与 Census 变换后的匹配代价值进行融合,并将其与基于图像亮度信息的代价值进行非线性融合,以提高匹配代价的精度;然后,使用最小生成树代价聚合模型进行聚合操作并利用赢者通吃策略估计图像的初始视差;最后,在视差优化阶段,采用 MeanShift 算法对图像进行分割,结合图像的轮廓信息对误匹配点进行修正,进一步提高在弱纹理及边缘区域的视差精度。实验结果表明,与一些传统算法相比,所提算法具有更高的视差精度,且视差图的边缘、纹理较其他算法更为平滑,具有更强的鲁棒性。

**关键词** 机器视觉; 立体匹配; Sobel 算子; Census 变换; 最小生成树

中图分类号 TP391.41

文献标志码 A

DOI: 10.3788/LOP212760

## Minimum Generation Cube Matching Algorithm Based on Improved Matching Cost and Mean Segmentation

Wang Daolei, Han Yang\*

College of Energy and Mechanical Engineering, Shanghai University of Electric Power, Shanghai 200090, China

**Abstract** The disparity accuracy issue still exists in weak texture and depth discontinuity areas, although the conventional stereo matching model has demonstrated good performance in accuracy and robustness. To address the above issues, a minimum generating cube matching algorithm based on enhanced matching cost and mean segmentation is proposed. First, in the matching cost computation stage, the initial matching cost is computed by the Census to transform, and the input image's edge information is extracted by the Sobel operator. The extracted image edge information is merged with the matching cost value after Census transform, and the nonlinear fusion is conducted with the cost value based on image brightness information to enhance the matching cost's accuracy. Then, the minimum spanning tree cost aggregation model is employed for aggregation operation and the winner-take-all technique is employed to estimate the image's initial parallax. Finally, in the disparity optimization stage, the MeanShift algorithm is employed to segment the image, and the mismatching points are corrected along with the image's contour information to further enhance the disparity accuracy in weak texture and edge areas. Experimental findings demonstrate that compared with some conventional algorithms, the proposed approach has higher disparity accuracy, and the disparity map's edges and textures are smoother and more robust than other algorithms.

**Key words** machine vision; stereo matching; Sobel operator; Census transform; minimum spanning tree

## 1 引言

双目立体匹配技术是图像处理和计算机视觉领域

的研究热点之一,它致力于从标定的立体图像中获取空间点的视差,进而获取图像的深度信息<sup>[1]</sup>。近年来,随着立体匹配技术在计算精度和效率方面的显著提

收稿日期: 2021-10-18; 修回日期: 2021-11-16; 录用日期: 2021-12-21; 网络首发日期: 2021-12-30

基金项目: 国家自然科学基金(61502297)

通信作者: h523393879@163.com

高,立体匹配技术已经被广泛应用于众多视觉任务中,如机器人视觉系统、视频显著性检测与分析、姿态估计与识别、三维重建及许多其他工业领域<sup>[2]</sup>。根据计算策略的不同,传统的立体匹配方法可以分为两类:基于全局的立体匹配及基于局部聚合策略的立体匹配<sup>[3]</sup>。

基于局部聚合策略的立体匹配方法应用最为广泛。局部代价聚合的原理是依次将匹配基元中的像素点作为兴趣点,并以其为中心构造约束区域进行代价聚合,求取参考图像视差。传统的局部立体匹配方法常使用窗口的方式进行代价聚合操作。例如,Xu等<sup>[4]</sup>提出了一种基于固定窗口的局部立体匹配模型,该模型利用梯度差分计算的聚合策略有效提高了局部方法在弱纹理区域的视差精度。Yoon等<sup>[5]</sup>提出了自适应窗口的聚合策略,该策略可以根据像素扩展的约束条件自适应调整窗口尺寸,提高了在视差不连续区域的精度。Einecke等<sup>[6]</sup>提出了多窗口聚合的立体匹配方法,该方法通过多个窗口进行聚合操作,显著提升了局部算法的鲁棒性。尽管局部算法具有运算复杂度低、计算效率高、实时性好的优点,但是局部代价聚合的过程仅考虑了局部邻域的像素关系,因此传统的基于窗口的局部立体匹配模型在几何约束区域容易受到噪声、光照等因素的干扰,无法适用于高精度的实际应用场景。

基于全局的立体匹配方法主要根据 Potts 能量函数理论建模<sup>[7]</sup>。全局的立体匹配模型主要采用能量泛函最小化的策略来获取视差精度。由于在能量函数的构造过程中引入了全局信息,全局方法相较传统的局部方法具有更高的视差精度。常见的全局立体匹配主要包含图割<sup>[3]</sup>、置信传播<sup>[3]</sup>、动态规划算法等<sup>[3]</sup>。这些方法在能量函数中建立数据项、平滑项及遮挡项进行高精度建模,均实现了较好的视差估计效果。尽管全局方法包含了图像的更多信息,能够提供更精准的视差。但全局方法依然存在不足,全局立体匹配算法针对能量泛函的最小化问题,是二维空间的 NP 难问题。因此在能量函数求解的过程中需要不断迭代,这导致全局方法的计算复杂度较高,在真实场景立体图像序列的实时应用中存在不足。

随着神经网络在众多视觉相关领域中的成功应用,基于深度学习的立体匹配方法受到了越来越广泛的关注<sup>[8]</sup>。基于深度神经网络的立体匹配模型通常采用可学习的网络架构,通过训练样本的方式学习图像特征,将视差估计转化为回归问题,直接从立体图像中预测稠密的视差图。受经典光流估计网络<sup>[9]</sup>的启发,Mayer等<sup>[10]</sup>提出了一种用于视差回归的端到端卷积网络,该网络利用由压缩和扩展两部分组成的编解码器结构来提高视差预测的准确性。Pang等<sup>[11]</sup>提出了一个由两个阶段组成的级联剩余学习网络(CRL),该网络通过嵌入的残差学习机制进行视差细化。此

外,Knöbelreiter等<sup>[12]</sup>利用 CNN 和条件随机场(CRF)的优点,提出了一种用于立体匹配的混合网络,该网络使用 CNN 从立体图像中提取特征以产生多个一元和成对代价,然后使用 4 连通 CRF 搜索联合解来优化 CNN 输出。尽管基于神经网络的立体匹配方法在多个标准数据库上表现出了优异的性能,但这些方法大多需要监督,很难直接应用于不易获取真实视差的实际场景。

近来,基于最小生成树(MST)的立体匹配方法因其对高分辨率图像中复杂场景的高精度视差估计而受到广泛关注<sup>[13]</sup>。尽管基于 MST 的立体匹配模型没有采用构造能量函数的方式进行建模,但其在代价聚合过程中通过将图像构造成树的策略引入了全局特征信息,故在实际应用中能兼顾视差精度及实时性。

然而,现有的基于 MST 的立体匹配方法都存在一个典型的局限性,即 MST 的构造可能在弱纹理及视差不连续区域附近视差估计不精确,这主要是因为 MST 的方法在匹配代价计算过程中只考虑颜色信息,没有考虑到图像的边缘特征,故在上述区域匹配效果不理想。

针对基于 MST 的立体匹配模型及局部立体匹配方法中图像弱纹理及深度不连续区域误匹配率较高的问题,本文提出了一种新的立体匹配算法,该算法总体思路如下:在代价计算阶段,首先使用 Sobel 算子对左右图像进行检测获取图像的边缘纹理信息;然后将提取后的图像边缘特征与 Census 变换后的匹配代价值进行线性融合,再和基于图像亮度信息的代价值进行非线性融合,以提高算法在图像弱纹理及深度不连续区域匹配的鲁棒性;随后通过 MST 的聚合策略计算图像初始视差;最后,在视差优化过程中,使用均值漂移算法对图像进行分割,结合图像的轮廓信息与左右一致性原理对视差计算错误区域进行修正,进一步提高在弱纹理及边缘区域的视差精度。

## 2 改进的匹配代价

图 1 是所提算法的整体流程图。

从贝叶斯理论考虑,匹配代价计算可以看成是一种度量方式<sup>[14]</sup>。这种度量代表两个像素具有相同视差的概率。因此,匹配代价的精度对最终的视差结果具有重要影响。Census 变换是一种常见的匹配代价计算方法,其保留了目标像素领域范围内的特征信息,在亮度差异较大的环境下表现出极强的鲁棒性,能够有效应对光照变化等外部噪声的干扰,在现实场景应用中具有较强的可靠性。尽管 Census 方法能够较好地应用于各种现实场景,但其在匹配代价计算的过程中仅考虑像素间的颜色信息,这会导致在弱纹理、边缘及重复纹理区域匹配代价计算不准确的问题。针对该问题,本研究提出了一种新颖的匹配代价计算方法。首先利用 Sobel 算子对输入图像进行边缘信息提取,然

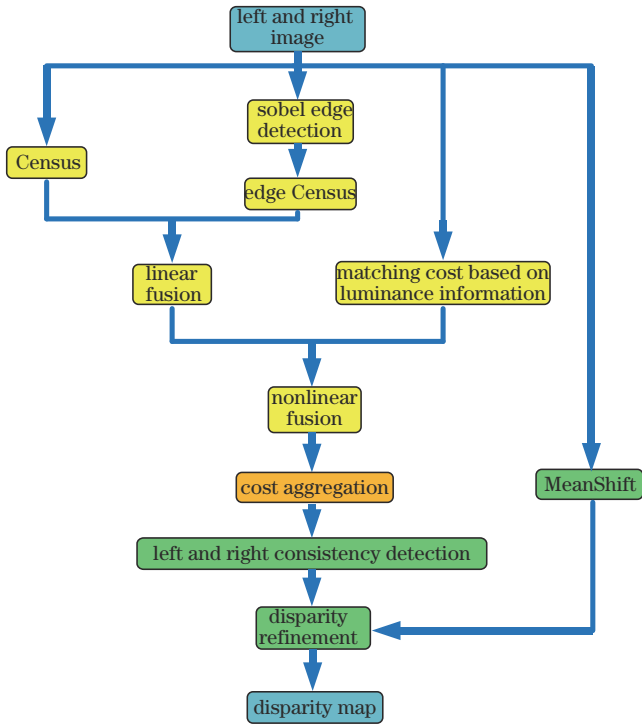


图 1 所提算法流程图

Fig. 1 Diagram of proposed method

后将提取后的图像边缘信息引入匹配代价计算的过程,并结合 Census 与图像亮度信息的代价值进行非线性融合,进而提高匹配代价在复杂纹理区域的计算准确性。

传统的 Census 变换方法通常将目标像素置于窗口中心处,然后通过与其周围领域内的颜色强度进行对比,将颜色的强弱关系转换为比特串,该比特串代表当前目标像素 Census 变换后的变换值。假设目标像素在固定窗口内的领域灰度信息如图 2 所示,在计算 Census 的变换值时,需要将窗口领域内其他像素的灰度值与目标像素依次进行对比。若领域像素的灰度值相较目标像素灰度值更大,则该位处的编码标记为 0,反之设为 1。整个转换过程通过二进制的编码方式进行保存,字符串的长度等于窗口内的像素个数。通过上述过程可以计算出图 2 目标像素处的 Census 变换值为 11010101。然后将对应图像中同一扫描线上的像素点的汉明距离作为目标像素的匹配代价,如图 3 所示。

传统的 Census 方法在匹配代价计算阶段仅考虑图像灰度信息进行变换,这会导致在处理图像边缘区

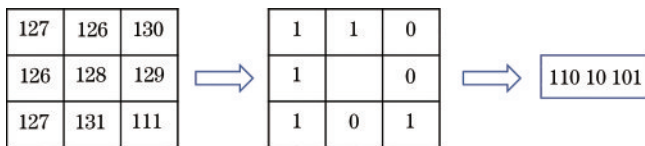


图 2 Census 变换示意图

Fig. 2 Schematic diagram of Census transformation

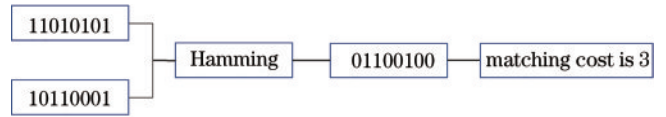


图 3 匹配代价计算过程

Fig. 3 Matching cost calculation process

域时性能表现不佳,易产生匹配代价计算不准确的问题。为了进一步增强匹配代价的鲁棒性,本研究将边缘信息引入匹配代价计算的过程中,从而增强匹配代价计算的鲁棒性。

Sobel 算子是一种常用的边缘检测技术。利用 Sobel 算子对参考图像进行检测,能够通过计算获得的梯度值反映边缘信息的强弱,以有效提取图像的轮廓。图 4 是 Middlebury 标准数据库中 Teddy 图像使用 Sobel 算子检测后的结果。



图 4 Teddy 边缘检测结果

Fig. 4 Teddy edge detection results

为了在匹配代价中加入边缘纹理信息,对边缘提取后的轮廓图进行 Census 变换,然后计算边缘 Census 变换下的匹配代价,并将其与灰度计算得到的初始匹配代价进行线性融合。

传统的 Census 变换计算得到的匹配代价为

$$C_{\text{census}}(p, q, d) = H(p, q), \quad (1)$$

边缘 Census 变换下计算得到的匹配代价为

$$C_{\text{census-edge}}(p, q, d) = H[\text{edge}(p), \text{edge}(q)], \quad (2)$$

$$C_1(p, q, d) = \alpha \times C_{\text{census}}(p, q, d) + (1 - \alpha) \times C_{\text{census-edge}}(p, q, d), \quad (3)$$

式中:  $C_1(p, q, d)$  表示融合后在视差  $d$  下的匹配代价值;  $H(\cdot)$  表明汉明距离计算;  $\text{edge}$  表示 Sobel 计算得到的边缘值;  $\alpha = 0.5$  代表线性融合权值。融合后的匹配代价,在传统 Census 的基础上引入边缘信息,这使得计算得到的匹配代价能够有效提高在边缘纹理处的匹配精度。

Census 变换是一种非参数的局部变换,主要依赖邻域内像素与目标像素间的局部强度关系。因此,在具有复杂纹理的场景中, Census 变换能够有效反映周围纹理的变化特征,进而获得高精度的匹配代价。然而,在重复纹理区域周围邻域像素通常具有相似的结构,目标像素与周围邻域的灰度、边缘等特征变化不明



显。因而在这种情况下,仅采用 Census 变换计算得到的匹配代价具有不可靠性。不同于 Census 对纹理结构较为敏感,亮度信息能够较好地反映重复纹理区域中局部信息的波动,即使像素间仅具有轻微的亮度变化。亮度信息通常采用对左右图像对应像素点灰度值的绝对差和垂直方向上梯度值的绝对差进行线性融合的方法来计算匹配代价<sup>[15]</sup>:

$$C_2(p, q, d) = \theta \times \min \left[ \left| I_l(p) - I_r(p+d) \right|, \tau_1 \right] + (1-\theta) \times \left[ \min \left( \left| \nabla_x I_l(p) - \nabla_x I_r(p+d) \right|, \tau_2 \right) \right], \quad (4)$$

式中: $\theta$ 是比例调节参数; $I(\cdot)$ 代表灰度值; $\tau_1, \tau_2, \nabla_x$ 表示垂直方向的梯度。为了进一步增强弱纹理及重复纹理区域匹配代价的鲁棒性和准确性,将含有边缘信息的 Census 与亮度信息再进行一次非线性融合,构建最终的匹配代价计算函数:

$$C(p, q, d) = \rho \left[ C_1(p, q, d), \lambda_1 \right] + \rho \left[ C_2(p, q, d), \lambda_2 \right], \quad (5)$$

$$\rho(C, \lambda) = 1 - \exp \left[ -C(p, q, d) / \lambda \right], \quad (6)$$

式中: $\lambda_1, \lambda_2$ 是控制参数。改进的匹配代价在传统的 Census 变换的基础上引入边缘信息与亮度信息。边缘及亮度匹配代价的单独计算及融合的改进策略在一定程度上降低了 Census 对纹理结构的依赖,削弱了周围纹理对匹配代价的影响,显著增强了在弱纹理、边缘及重复纹理区域的准确性。

### 3 代价聚合

近来,基于 MST 的代价聚合模型受到越来越广泛的关注,该模型将图像构造为树来定义像素之间的关系,然后通过双边滤波模型驱动相邻像素接近目标像素的方式完成代价聚合步骤。鉴于 MST 的聚合模型具有高精度和高效率的优点,本研究选用 MST 聚合模型进行代价聚合。不同于传统的局部窗口聚合策略,基于 MST 的聚合模型首先将参考图像定义为一个连通的加权无向图:

$$G = (V, E), \quad (7)$$

式中: $V$ 是所有像素点的集合; $E$ 是连接 2 个像素的边集。相邻顶点  $s$  和  $r$  之间边的权重  $w(r, s)$  的表达式为

$$w[s, r] = w(r, s) = |I(s) - I(r)|. \quad (8)$$

然后采用 Kruskal<sup>[10]</sup>算法计算参考图像的 MST。随后,通过两个步骤完成代价聚合:

1) 首先从叶节点到根节点进行初次代价聚合(如图 5 所示):

$$C_d^A(v) = C_d(v) + \sum_{\rho(v_c=v)} S(v, v_c) C_d^A(v_c), \quad (9)$$

$$S(v, v_c) = \exp \left[ -w(v, v_c) / \delta \right], \quad (10)$$

式中: $C_d(v)$ 为第 2 小节中计算得到的对应视差  $d$  下的

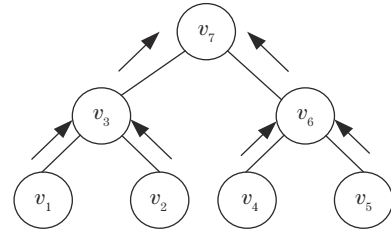


图 5 叶节点到根节点

Fig. 5 Leaf node to root node.

匹配代价值; $C_d^A(v)$ 为节点  $v$  在视差  $d$  下的代价聚合值; $P(v_c)$ 为节点  $v_c$  的父节点; $S(v, v_c)$ 表示  $v_c$  与父节点  $v$  之间的相似性; $\sigma$ 是用来调节 2 个节点的权重参数。

2) 最终的代价聚合值  $C_d^A(v)$  通过将  $C_d^A(v)$  从根节点到叶节点进行聚合获得:

$$C_d^A(v) = S \left[ P(v), v \right] C_d^A \left[ p(v) \right] + \left\{ 1 - S^2 \left[ v, p(v) \right] \right\} C_d^A(v). \quad (11)$$

图 6 阐述了最终的代价聚合值计算过程。

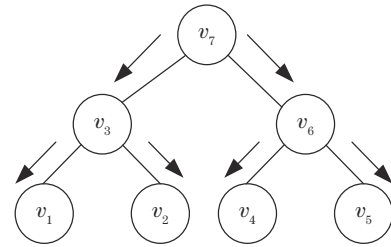


图 6 根节点到叶节点

Fig. 6 Root node to leaf node

### 4 视差计算和细化

在完成代价聚合步骤之后,胜者为王(WTA)<sup>[16]</sup>策略被用来确定每个像素点视差:

$$d_{\text{end}} = \arg \min_{d \in \text{Disp}} C_d^A(v), \quad (12)$$

式中: $d_{\text{end}}$ 表示计算得到的初始视差值;Disp 为候选视差空间值集合,范围在  $0 \sim d_{\text{max}}$  之间; $v$  代表当前像素点。通过聚合策略直接获得的初始视差通常存在遮挡和误匹配的问题,因此需要通过视差细化过程来进一步提高视差图的精度。

本研究提出了一种基于图像分割的视差细化方法。在计算得到左右视差图的情况下,首先利用左右一致性检测的方式从两幅视差图中提取不稳定像素:

$$|d_L - d_R| \leq T, \quad (13)$$

式中: $d_L$ 代表左侧图像计算得到的视差; $d_R$ 表示将右侧图像作为参考图像处理时获得的视差; $T$ 为阈值。满足以上条件的像素是稳定点,否则就判定为不稳定点,即错误匹配或者遮挡点。对于不稳定点像素的视差,采用均值漂移算法进行修正。基于均值漂移的图

像分割<sup>[17]</sup>是一种统计迭代分割算法,具有聚类相似的过程。在分割的过程中选择一个中心点,然后计算该中心点邻域范围内的点到中心点的距离向量的平均值。并通过该平均值得到一个偏移均值,随后将中心点移动到偏移均值位置,重复上述过程,即可实现图像分割。以 Middlebury 标准数据库中的 Teddy 为例,图 7 展示了分割后的结果,其中颜色相同代表像素位于同一分割块。从图 7 可以看出,理论上根据平滑假设,相同分割块内的像素应该具有相似的视差。基于此,本研究从不稳定点的 4 个方向出发去寻找同一分割块内的稳定像素,并将第 1 个遇到的块内稳定像素的视差用来填充当前不稳定点。通过该策略,可以将视差值从稳定像素传播到不稳定像素,进而进一步提高视差估计的准确性。搜索优化过程如图 8 所示。



图 7 图像分割前后结果

Fig. 7 Results before and after image segmentation

		stable point		
		stable point		
	stable point	unstable point	stable point	
		unstable point		
		stable point x		

图 8 搜索优化过程

Fig. 8 Search optimization process

图 8 中,相同颜色为同一分割块,目标像素为窗口中心像素,该点为不稳定点。此时,通过上下左右 4 个方向进行搜索,由于上左右 3 个方向都不在同一像素块内,因此上述领域点的视差不予考虑。尽管下方领域像素包含在同一分割块内,但其不稳定。通过继续向下搜索发现  $x$  点不仅与目标像素在同一分割块内,且该点为稳定点,在这种情况下将选用  $x$  点的视差值对目标像素进行填充。

## 5 实验

为了客观公正评价所提算法性能,采用 Middlebury 标准数据集进行评估。实验在 i5-8264U 1.8 GHz 的 CPU 及 Visual Studio 2019 上进行。此外,还配合了 OpenCV 计算机视觉开源库。在实验对比的过程中,为了获得理想的立体匹配结果,需要为所提算

法的一些自由参数设定合适的值:式(4)中匹配代价的截断阈值  $\tau_1=7$ 、 $\tau_2=2$ 、权重  $\theta=0.11$ ,式(5)中的控制参数  $\lambda_1=15$ 、 $\lambda_2=5$ 。上述实验参数均通过大量实验分析后设定。

Middlebury 是一个被广泛用来评估立体匹配算法性能的标准数据库,其中的图像主要在室内采集,通过结构光的方法获取真实视差值。由于数据库中的图片具有较为复杂的场景及纹理,对现有的立体匹配算法的性能提出了挑战。为了验证所提算法性能,将其与其他先进的算法在 Middlebury V2 及 Middlebury 扩展库中的图像进行了实验对比。

### 5.1 改进匹配代价的实验结果对比分析

为了说明线性权重值设定的有效性,利用 MiddleburyV2 数据集来验证权重参数对匹配代价精度的影响。图 9 是匹配代价精度随线性权重  $\alpha$  的变化示意图。

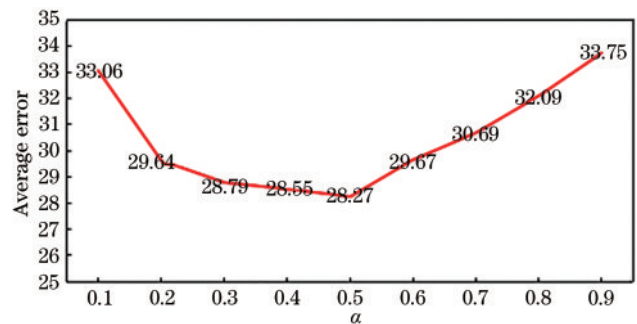


图 9 匹配代价精度随  $\alpha$  变化示意图

Fig. 9 Schematic diagram of matching cost accuracy changing with  $\alpha$

从图 9 可以看出,随着  $\alpha$  的增加,计算得到的匹配代价精度显著上升,直到  $\alpha=0.5$  时平均误差最低。因此,将融合权重参数设定为 0.5。

为了对所提匹配代价计算方法的有效性进行验证。以 Middlebury 中常用的 4 组标准图像 Tsukuba、Venus、Teddy 和 Cones 为例,将所提匹配代价计算方法和 Census、AD、AD-Census、GRD 等常规的匹配代价方法进行了实验对比。

图 10 从上到下依次是各方法仅计算匹配代价得到的视差结果,从左往右的立体图像依次为 Tsukuba、Venus、Teddy 及 Cones。表 1 则是各方法的误差对比结果,从深度不连续区域(disc)、所有区域(all)及非遮挡区域(nonocc)分别来验证各立体匹配代价计算方法的性能。

从图 10 可以看出,GRD 和 AD 的视差图效果最差,生成的视差图中噪声较多,且纹理模糊。表 1 中的数据也表明,GRD 和 AD 的误差相较于其他对比方法较大。在具有简单场景的 Tsukuba 和 Venus 中,AD-Census 明显优于 Census。从 Tsukuba 的台灯头部对比可知,前者轮廓更加清晰。在 Venus 图中也具有相似





图 10 不同算法匹配代价视差图结果对比

Fig. 10 Comparison of different algorithm matching cost disparity map results

表 1 算法匹配代价评测误差对比

Table 1 Comparison of evaluation errors of algorithm matching cost

unit: %

Algorithm	Tsukuba			Venus			Teddy			Cones			Average error
	disc	all	nonocc	disc	all	nonocc	disc	all	nonocc	disc	all	nonocc	
GRD	48.31	49.28	48.74	60.18	62.72	62.1	77.65	79.06	76.75	79.7	84.04	82.21	67.56
AD	46.98	48.76	48.21	58.76	62.35	61.72	77.22	78.82	76.48	79.3	83.76	81.94	67.03
AD-Census	26.96	23.43	22.13	38.14	33.56	32.41	50.44	47.34	41.31	57.87	59.38	54.4	40.61
Census	25.12	30.79	23.76	40.46	42.29	41.4	45.38	43.22	25.44	33.89	32.88	24.67	34.12
Text matching cost	17.17	22.67	15.56	35.02	32.99	31.9	34.74	33.11	36.96	26.03	24.76	15.33	27.19

的情况,AD-Census的噪声明显相较于Census更少。但是在具有复杂场景的Teddy和Cones中,Census性能却优于AD-Census。在Cones中,Census相较于AD-Census在物体上的轮廓更加清晰。这主要是因为复杂的场景通常包含较为复杂的纹理结构,Census在这种情况下匹配代价的准确性较高。尽管AD-Census对灰度及Census进行了非线性融合,但其依然具有明显的局限性,在面对复杂场景时稳定性较差。与上述方法不同,所提方法考虑传统Census的同时,又引入边缘及亮度信息,因此具有更强的鲁棒性。相较于其他对比方法,所提匹配代价方法生成的视差具有更高的视差精度,更接近真实视差图。

此外,由表1的误差对比结果可知,所提方法生成的视差误差最小、准确度更高,分别相较于Census和GRD算法提高了7个百分点和40个百分点。同时精度也优于AD和AD-Census。上述实验结果验证了所提改进Census方法的有效性,其能够适用于不同的应用场景。

## 5.2 视差细化前后效果对比

图11(a)是Cones图片视差细化前后的对比结果,细化后的图像对部分误匹配像素进行了修正。通过进一步观察发现,Cones细化前左端圆锥区域存在明显的空洞区域,但是细化后圆锥左端的空洞得到了有效的填充,边缘更加平滑。

为了进一步验证所提细化方法的有效性,图11(b)展示了Art图片细化前后的视差图对比结果。图中细化前的视差存在较多的空洞区域,但是细化后空洞区域得到很好的填充。此外,在具有复杂纹理的人像周围,视差精度得到了显著提升。

由图11可知,基于图像分割的视差细化可以起到

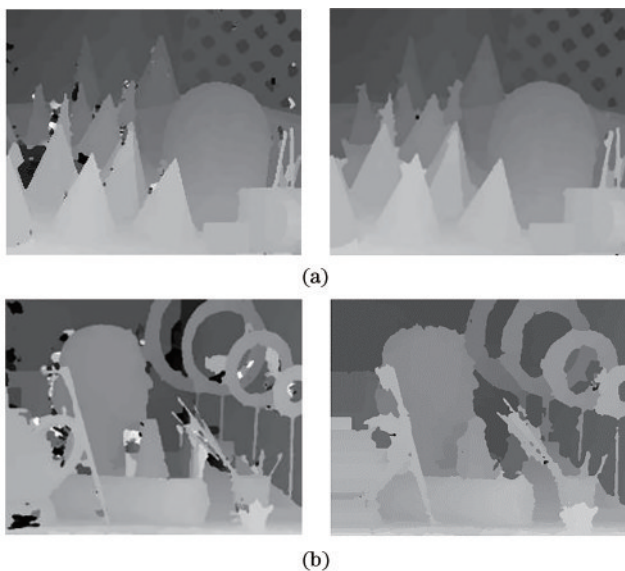


图 11 视差细化前后结果对比。(a) Cones;(b) Art

Fig. 11 Comparison of results before and after parallax thinning.

(a) Cones; (b) Art

去除噪声、填充空洞及使视差更加平滑的作用,能够有效提高视差估计的准确性。

## 5.3 整体算法的实验对比与分析

为了评估所提算法的整体性能,选取了包括NLCA<sup>[13]</sup>、ST<sup>[18]</sup>、ASW<sup>[19]</sup>、SGM<sup>[20]</sup>、ELAS<sup>[21]</sup>及文献[22]、[23]、[24]在内的8个具有代表性的立体匹配算法进行比较。图12从上到下依次是ASW、SGM、ELAS、ST、NLCA、所提算法及真实视差图,从左往右的立体图像依次是Teddy、Cones、Laundry、Baby3,图中白色边框表示边缘及复杂纹理所在区域。从上述立体图像可以看出,ASW、SGM、ELAS、ST这4种算法计算得到的视差结果在左边区域都存在较大的空洞,且复杂纹理区域计算得到的视差比较粗糙,整体匹配效果较差。尽管NLCA也同样采用最小生成树聚合的策略,具有较好的视差结果,但其在复杂边缘依然存在错误匹配。这主要是因为其在视差的计算过程中没有考虑边缘信息,从而导致边缘纹理的缺失。不同于NLCA,所提算法在匹配代价中引入了边缘信息,使得获得的视差更加平滑,纹理更加清晰。此外,相较于其他算法,所提算法计算得到的视差具有明显的提升,特别是在复杂纹理及边缘区域,这再次证明了所提匹配代价方法的有效性。值得注意的是,所提算法获得的视差图中明显空洞及错误匹配区域较少,这得益于基于图像分割的视差细化策略。例如在Teddy左端的弱纹理区域,所提算法完整地计算出了包括遮挡区域在内的视差。在Laundry中,背景的窗户包含大量重复纹理结构,现有的方法很难估计这类区域的视差。与其他经典算法相比,所提算法提供了具有竞争力的表现,有效改善了重复纹理区域的视差精度。以上的实验结果对比说明,所提算法相较于其他算法的视差计算结果更加接近真实视差。

为了定量对比所提算法与其他算法,表2展示了各算法在Middlebury数据库上disc、all、nonocc区域的平均误差对比结果,其中Proposed+Census代表采用Census作为匹配代价的所提算法,Proposed+without代表未加视差优化的所提算法。由表2可以看出,所提算法的平均误差最小,特别是在无遮挡区域,所提算法计算得到的误差远低于其他对比算法。

为了验证所提算法对视差估计精度的提升作用,测试不同模块的有效性。将各算法运行在Middlebury数据集上进行视差结果分析。从表2可以看出,所提算法在视差精度方面明显优于Proposed+Census,特别是在不连续区域,这表明了所提匹配代价方法的有效性。此外,与Proposed+without对比可知,经过视差优化的精度得到了显著提升,即所提视差细化策略可以显著提高视差估计的精度,能对初始视差结果进行修正。



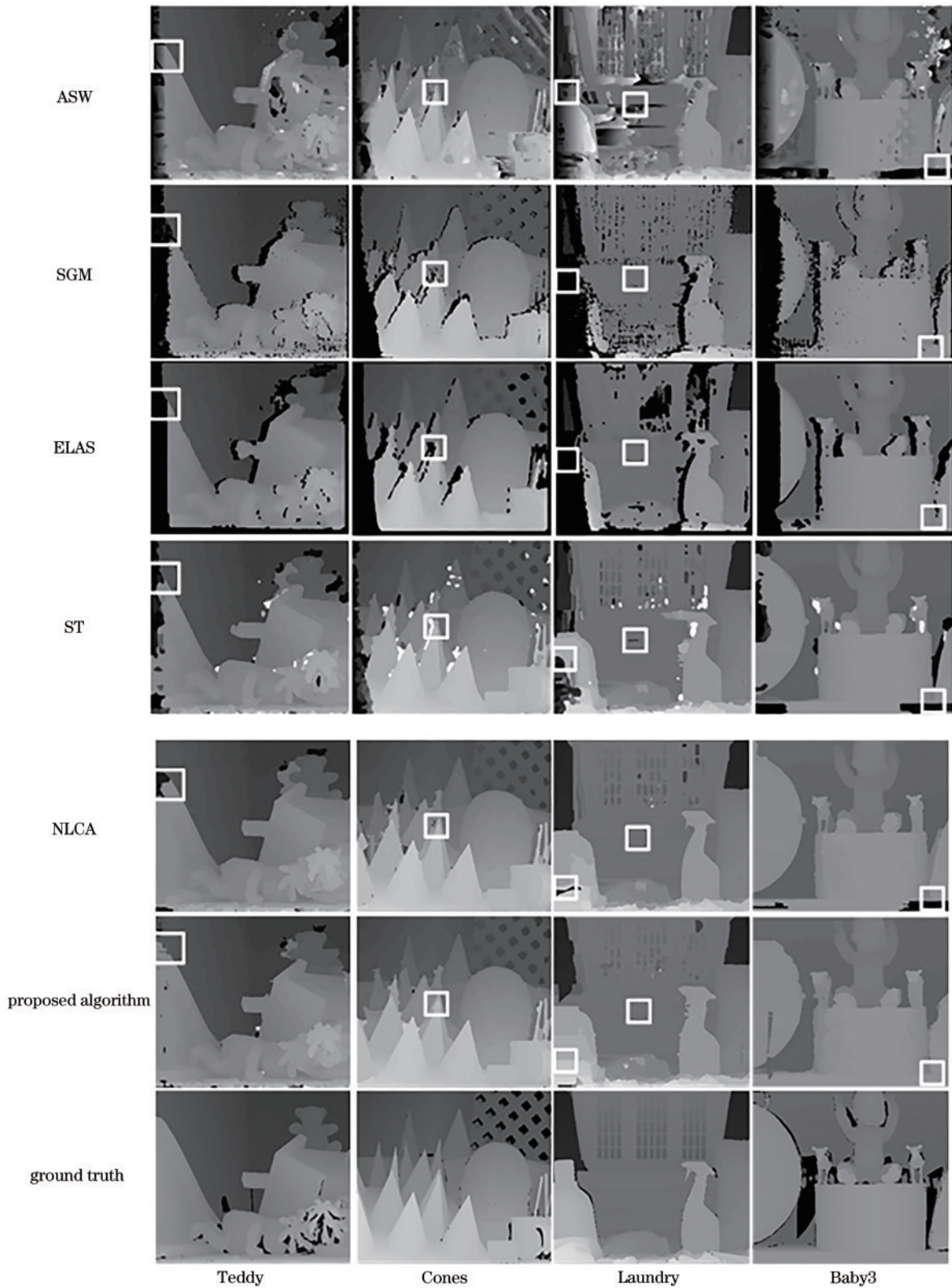


图 12 视差图结果对比

Fig. 12 Comparison of disparity map results

在 Middlebury 扩展集上,所提算法同样展现出了良好的性能,如表 3 所示,在总计 15 张被测试立体图像对中,所提算法计算得到的视差在 13 个测试图中误差最小。在平均误差方面,所提算法的平均误差仅为

8.82%,远远优于其他 6 个对比算法。实验结果再次表明,所提立体匹配方法适用于不同的应用场景,具有极强的鲁棒性和应用前景。



表 2 MID 2 标准集评测结果示意图

Table 2 Schematic diagram of evaluation results of MID 2 standard set

unit: %

Algorithm	Tsukuba			Venus			Teddy			Cones			Average error
	disc	all	nonocc	disc	all	nonocc	disc	all	nonocc	disc	all	nonocc	
ELAS	35.25	19.2	17.46	36.83	11.33	9.81	43.6	30.36	22.45	36.31	27.39	18.25	25.69
SGM	32.03	10.43	9.02	27.77	5.48	3.92	23.78	19.28	9.5	31.72	22.64	13.41	17.42
ASW	12.64	12.68	11.05	13.25	7.55	5.97	23.79	26.09	17.04	24.83	24.23	15.68	16.23
ST	10.68	4.61	3.66	6.37	1.22	0.56	15	7.83	18.64	12.45	12.85	4.57	8.20
NLCA	9.25	3.44	2.91	3.37	<b>0.53</b>	0.34	15.51	12.05	6.44	9.81	9	3.46	6.34
Proposed+Census	16.35	8.63	7.35	8.24	2.37	1.45	20.39	15.51	9.09	9.22	9.77	3.25	9.30
Proposed+without	11.72	8.1	7.14	4.96	1.01	0.56	17.31	13.15	6.75	10.07	10.24	3.48	7.87
Proposed algorithm	8.94	4.35	3.23	<b>2.81</b>	0.89	<b>0.28</b>	15.6	11.8	<b>6.12</b>	7.57	8.81	<b>2.63</b>	<b>6.09</b>
Reference [22]	<b>8.7</b>	3.1	2.8	3.7	2.1	1.7	14.0	11.8	7.4	<b>7.0</b>	<b>6.4</b>	5.1	8.20
Reference [23]	8.77	2.86	<b>2.27</b>	3.16	1.15	0.88	<b>9.35</b>	<b>7.64</b>	7.52	13.08	11.74	8.25	8.37
Reference [24]	10.11	<b>2.72</b>	4.54	3.36	0.95	0.84	17.16	11.72	6.35	8.32	10.33	5.87	6.43

表 3 MID 2 扩展集 15 张评测 nonocc 区域对比结果

Table 3 Comparison results of 15 evaluation nonocc regions in MID 2 extended set

Region	NLCA	ST	ASW	SGM	ELAS	Proposed algorithm
Aloe	5.08	5.62	6.29	9.81	16.54	4.45
Art	10.19	11.70	22.63	20.64	26.02	9.92
Baby1	10.55	5.04	22.36	8.36	13.42	6.84
Baby2	18.01	15.78	25.72	12.79	13.58	12.49
Baby3	7.28	5.58	26.40	10.06	16.36	3.95
Books	10.73	9.87	28.11	16.90	17.08	9.06
Cloth2	4.20	4.43	12.46	6.31	11.15	2.69
Cloth3	2.20	2.36	1.84	5.57	10.37	1.49
Dolls	6.24	6.42	12.61	10.76	15.16	5.10
Flowerpots	22.26	15.46	40.11	20.46	19.56	14.74
Lampshade1	12.81	11.07	44.21	25.15	29.28	9.59
Laundry	11.76	12.81	36.63	21.44	26.16	12.50
Midd1	23.24	34.72	49.24	46.51	53.09	24.30
Moebius	7.89	8.89	23.28	16.61	20.39	7.66
Reinder	9.56	8.73	17.67	11.93	17.51	7.53
Average error	10.80	10.57	24.64	16.22	20.38	<b>8.82</b>

## 6 结 论

为了解决立体匹配在图像弱纹理及重复纹理区域视差估计精度较低的问题,提出了一种基于改进匹配代价和均值分割的最小生成树立体匹配算法。首先使用 Sobel 算子对左右图像进行边缘检测,然后将边缘信息引入匹配代价的计算过程,并与亮度信息进行非线性融合,以提高匹配代价在弱纹理及重复纹理区域的准确性。随后,利用最小生成树进行代价聚合,获取初始视差。最后,在视差计算阶段,使用均值漂移算法完成视差优化,以进一步提高整体视差精度。实验结果表明,所提算法具有较好的视差估计精度,能够适应

于各种复杂场景,具有广泛的应用前景。

## 参 考 文 献

- [1] Guo J D, Diao Z H, Yan S F, et al. Immersive autostereoscopic display based on curved screen and parallax barrier[J]. Chinese Optics Letters, 2021, 19(1): 013301.
- [2] Muminov B, Perry A, Hyder R, et al. Toward simple, generalizable neural networks with universal training for low-SWaP hybrid vision[J]. Photonics Research, 2021, 9(7): B253-B261.
- [3] Scharstein D, Szeliski R, Zabih R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[C]//Proceedings IEEE Workshop on Stereo and Multi-Baseline Vision, December 9-10, 2001, Kauai, HI, USA. New York: IEEE Press, 2001: 131-140.
- [4] Xu J X, Li Q W, Zhou Y, et al. Texture discrimination-enforced matching cost computation and smoothness-weighted cost regularization for stereo matching[J]. Journal of Electronic Imaging, 2019, 28(5): 053025.
- [5] Yoon K J, Kweon I S. Adaptive support-weight approach for correspondence search[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(4): 650-656.
- [6] Einecke N, Eggert J. A multi-block-matching approach for stereo[C]//2015 IEEE Intelligent Vehicles Symposium, June 28-July 1, 2015, Seoul, Korea (South). New York: IEEE Press, 2015: 585-592.
- [7] Kolmogorov V, Zabih R. Computing visual correspondence with occlusions using graph cuts[C]//Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, July 7-14, 2001, Vancouver, BC, Canada. New York: IEEE Press, 2001: 508-515.
- [8] Xu S F, Zou W W. Optical tensor core architecture for neural network training based on dual-layer waveguide topology and homodyne detection[J]. Chinese Optics

- Letters, 2021, 19(8): 082501.
- [9] Dosovitskiy A, Springenberg J T, Brox T. Learning to generate chairs with convolutional neural networks[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 1538-1546.
- [10] Mayer N, Ilg E, Häusser P, et al. A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 4040-4048.
- [11] Pang J H, Sun W X, Ren J S, et al. Cascade residual learning: a two-stage convolutional neural network for stereo matching[C]//2017 IEEE International Conference on Computer Vision Workshops (ICCVW), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 878-886.
- [12] Knöbelreiter P, Reinbacher C, Shekhovtsov A, et al. End-to-end training of hybrid CNN-CRF models for stereo[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1456-1465.
- [13] 杨科, 刘凯. 多测度融合的树形滤波立体匹配算法[J]. 计算机工程与设计, 2021, 42(8): 2277-2283.  
Yang K, Liu K. Tree filtering stereo matching algorithm combining multi measures[J]. Computer Engineering and Design, 2021, 42(8): 2277-2283.
- [14] 李亚旋, 谢红薇, 王春丹, 等. 多分量权值的树结构立体匹配算法[J]. 计算机工程与设计, 2020, 41(9): 2501-2506.  
Li Y X, Xie H W, Wang C D, et al. Tree structure stereo matching algorithm for multi-component weights [J]. Computer Engineering and Design, 2020, 41(9): 2501-2506.
- [15] 刘建明, 何晴, 陈辉. 基于改进 Census 变换与颜色梯度融合的立体匹配[J]. 微电子学与计算机, 2021, 38(9): 38-44.  
Liu J M, He Q, Chen H. Stereo matching based on improved Census transform and color gradient fusion[J]. Microelectronics & Computer, 2021, 38(9): 38-44.
- [16] 陈映光, 周佩, 朱江平, 等. 基于改进 Census 变换和自适应支持域的立体匹配[J]. 激光与光电子学进展, 2021, 58(14): 1433002.  
Chen Y G, Zhou P, Zhu J P, et al. Stereo matching based on improved Census transformation and adaptive support region[J]. Laser & Optoelectronics Progress, 2021, 58(14): 1433002.
- [17] 来毅, 刘颖, 孟义祥, 等. 一种均值漂移与图割相结合的视差估计[J]. 西安邮电大学学报, 2017, 22(6): 56-60.  
Lai Y, Liu Y, Meng Y X, et al. Disparity estimation based on mean shift and graph cut[J]. Journal of Xi'an University of Posts and Telecommunications, 2017, 22(6): 56-60.
- [18] 陈卉, 胡立坤, 黄钰雯. 采用高斯混合模型及树结构的立体匹配算法[J]. 计算机工程与应用, 2017, 53(20): 195-200.  
Chen H, Hu L K, Huang Y W. Stereo matching algorithm based on Gaussian mixture model and tree structure[J]. Computer Engineering and Applications, 2017, 53(20): 195-200.
- [19] 邹进贵, 万荧, 孟丽媛. 一种基于自适应权重 SAD 与 Census 融合的匹配算法[J]. 测绘通报, 2018(11): 11-15.  
Zou J G, Wan Y, Meng L Y. A new stereo matching algorithm based on adaptive weight SAD algorithm and census algorithm[J]. Bulletin of Surveying and Mapping, 2018(11): 11-15.
- [20] 柳庆林, 倪章松, 张平涛, 等. 基于反向合成高斯牛顿的半全局立体匹配算法[J]. 光学技术, 2021, 47(4): 390-397.  
Liu Q L, Ni Z S, Zhang P T, et al. 3D measurement of single-frame surface structured light based on digital speckle projection[J]. Optical Technique, 2021, 47(4): 390-397.
- [21] 张道德, 伍渊, 吴良溢, 等. 基于 ELAS 立体匹配算法的研究与改进[J]. 湖北工业大学学报, 2017, 32(2): 1-6.  
Zhang D D, Wu Y, Wu L Y, et al. Research on improvement of stereo matching algorithm based on ELAS[J]. Journal of Hubei University of Technology, 2017, 32(2): 1-6.
- [22] 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.  
Zhang J, Wang J X. A semi-global stereo matching algorithm combining color information and Census transformation[J]. Science of Surveying and Mapping, 2019, 44(12): 147-152.
- [23] 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.  
Lu G M, Wang J X. Semi global stereo matching algorithm combined with image segmentation[J]. Remote Sensing Information, 2020, 35(6): 85-91.
- [24] 梁海波, 邹佳玲. 改进的非参数 Census 变换立体匹配算法[J]. 吉林大学学报(理学版), 2019, 57(5): 1163-1168.  
Liang H B, Zou J L. Improved stereo matching algorithm based on nonparametric Census transform[J]. Journal of Jilin University (Science Edition), 2019, 57(5): 1163-1168.