

基于多特征 AD-Census 变换的立体匹配算法

陆雨薇^{1,2} 常修源¹ 龙 辉³ 覃高杰¹
LU Yuwei CHANG Xiuyuan LONG Hui QIN Gaojie

摘 要

立体匹配算法的误匹配率一直是影响双目立体视觉精度的关键因素之一。为提高匹配的准确性和鲁棒性,文章提出了一种自适应多特征融合的立体匹配算法。首先,在代价计算阶段,结合图像区域的纹理复杂度,自适应调整 Census 变换窗口的大小,从而增强算法在不同复杂区域中的适应性。此外,在传统 AD 和 Census 变换特征的基础上,引入基于梯度信息的 Census 变换,并结合区域纹理信息计算特征融合权重,从而构建更加精确和可靠的匹配代价。其次,在代价聚合阶段,采用基于距离权重的十字交叉聚合策略,对视差值进行有效的重分配,减少误匹配的发生。并利用赢家通吃(WTA)算法获取初始视差后,通过左右一致性检测和亚像素精细化技术对视差结果进行后处理优化,以获得更加精确的视差图。实验结果表明,在 Middlebury 标准测试数据集中,所提算法在非遮挡区域和全区域的误匹配率分别为 8.02% 和 11.15%,并且在实际三维测量中,最大误差为 0.242%。实验结果验证了所提算法的有效性和可靠性。

关键词

机器视觉; 立体匹配; 自适应权重; AD-Census; 十字交叉法

doi: 10.3969/j.issn.1672-9528.2025.06.002

0 引言

双目立体视觉是计算机视觉领域的一个重要分支,通过模拟人眼的视觉系统来获取场景中的三维信息,被广泛应用于自动驾驶、姿态估计和移动机器人等多个领域^[1]。其核心原理是通过对左右相机获取同一场景下重叠区域同名点进行匹配,进而利用同名点之间的视差值,根据三角形原理计算三维深度信息。其立体匹配精度直接影响最终获取目标三维信息的准确性,因此,如何提高匹配精度仍是立体匹配技术的难点。

目前,传统的立体匹配算法主要分为局部立体匹配算法和全局立体匹配算法^[2]。全局立体匹配算法利用全局能量函数最小化来求解最优视差,尽管其在匹配精度上表现出色,但计算复杂度较高、计算时间较长以及对硬件性能的要求高,限制了其在实际中的广泛应用。典型的全局立体匹配算法有:

动态规划(DP)、图割法(GC)和置信传播(BP)等^[3]。相较之下,局部立体匹配算法则通过在固定大小的窗口内计算代价,并选择代价最小的视差值作为最终匹配结果。由于其具有较高的计算效率,在实际中得到了广泛应用。常见的局部立体匹配算法有:灰度绝对值法(AD)、绝对误差和(SAD)、归一化互相关(NCC)、互信息(MI)、误差平方和(SSD)、Census 变换等。

然而,传统的局部立体匹配算法在无纹理或重复纹理区域的匹配效果较差,并且对光照变化和噪声较为敏感。为了解决这些问题, Lee 等人^[4]提出基于星形 Census 变换的方法,有效降低了无纹理区域的误匹配率,然而,当中心点像素值受噪声影响时,其匹配精度依然会显著下降。王云峰等人^[5]提出了一种自适应权重的 AD-Census 算法,该方法能够根据不同纹理区域的特征自适应地调整权重,从而显著减少纹理复杂区域的误匹配率。Lü 等人^[6]提出了一种基于 HSV 空间的 Census 变换方法,有效减小了光照变化和噪声的影响。然而,该方法丢失了大量颜色信息,导致在视差不连续区域的效果较差。Pang 等人^[7]使用级联残差学习网络,分别计算每段视差图和多尺度的残差,尽管深度学习算法在匹配效果上表现优异,但由于其需要大量的数据标注,因此限制了其广泛应用。通过上述分析可知,视差计算精度不高依然是双目立体匹配面临的重要挑战之一。

为此,本文提出了一种基于动态权重融合的立体匹配方法,该方法结合 AD 代价、Census 代价以及基于梯度信息的

1. 广西科技大学广西汽车零部件与整车技术重点实验室

广西柳州 545616

2. 上汽通用五菱汽车股份有限公司广西新能源汽车实验室

广西柳州 545616

3. 柳州沪信汽车科技有限公司 广西柳州 545616

[基金项目] 广西科技计划“中国—东盟绿色车辆研究院搭建与能力建设”(AA24206060); 柳州市科技计划“基于机器视觉的汽车检具智能化低成本方案开发及应用”(2024AB0401A006)

Census 变换, 构建了一个更加精确的匹配代价计算函数。在代价聚合阶段, 采用基于距离权重的十字交叉法进行代价聚合。最后, 通过进一步优化视差值, 获得更为精确的匹配结果。

1 算法描述

立体匹配通过评估左右图像中视差范围内每对像素点的匹配代价来计算视差值, 其基本流程包括代价计算、代价聚合、视差计算和视差优化。首先, 在代价计算阶段, 结合 AD、Census 变换以及基于梯度信息的 Census 变换, 并根据图像区域特征动态调整各代价权重。其次, 在代价聚合阶段, 采用基于距离权重的十字交叉法进行代价聚合。接着, 在视差计算阶段, 利用赢家通吃 (WTA) 算法, 选取聚合代价最小值对应的视差作为每个像素的初始视差值。最后, 在视差优化阶段, 采用左右一致性检查和亚像素法等技术进一步提升匹配精度, 最终, 通过中值滤波对结果进行平滑处理, 生成最终的视差图, 算法的流程图如图 1 所示。

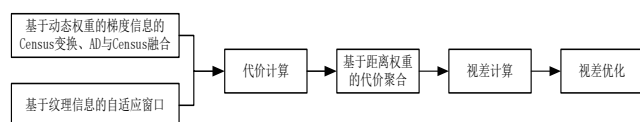


图 1 本文算法流程图

1.1 改进的代价计算

在传统的 AD 代价计算中, 通过评估图像中 RGB 三个颜色通道的差异来计算匹配代价。其计算公式为:

$$C_{AD}(p, d) = \frac{1}{3} \sum_{i=R, G, B} |I_{i,L}(p) - I_{i,R}(p-d)| \quad (1)$$

式中: $C_{AD}(p, d)$ 表示视差为 d 时点 p 处的 AD 代价值; i 表示 RGB 的 3 个通道; $I_{i,L}(p)$ 表示左边图像中点 p 在 i 通道下的像素值; $I_{i,R}(p-d)$ 表示右边图像中点 $(p-d)$ 在 i 通道的像素值。

传统的 Census 变换通过比较中心像素与其邻域像素的大小关系生成二进制编码。具体而言, 当邻域像素值小于中心像素值时编码为 0, 大于或等于时编码为 1。并将窗口内的比较结果按序连接, 生成最终的二进制编码, 其计算公式为:

$$\begin{cases} C_{Census}(p) = \bigotimes_{q \in N_p} \xi[I(p), I(q)] \\ \xi[I(p), I(q)] = \begin{cases} 0 & I(p) < I(q) \\ 1 & I(p) \geq I(q) \end{cases} \end{cases} \quad (2)$$

式中: $C_{Census}(p)$ 表示点 p 处的 Census 变换码; p 表示矩形窗口的中心点; q 表示矩形窗口内的邻域点; $\bigotimes$ 表示按位连接, 即将各个比较结果拼接成最终的二进制编码; $I(p)$ 、 $I(q)$ 分别表示点 p 和邻域点 q 的像素值; N_p 以点 p 为中心的矩形窗口内的所有邻域像素集合; $\xi[I(p), I(q)]$ 计算矩形窗口邻域元素的编码。

完成两张图像的 Census 变换后, 计算左图点 p 与右图在

视差范围内点 $p-d$ 的汉明距离作为匹配代价, 表示左右图像中对应点的匹配程度, 其计算公式为:

$$C_{Census}(p, d) = \text{Hamming}[C_{Census,L}(p), C_{Census,R}(p-d)] \quad (3)$$

式中: $C_{Census}(p, d)$ 表示当视差为 d 时, 左右图像中对应位置的两个二进制编码的相异个数; $C_{Census,L}(p)$ 表示左图像中点 p 处的 Census 变换码; $C_{Census,R}(p-d)$ 表示右图中对应 $p-d$ 处的 Census 变化码。

传统的 Census 变换通过邻域元素与中心像素值比较, 当中心像素受到噪声干扰时, 可能会导致编码发生显著变化, 从而影响匹配精度。为减少噪声对匹配精度的影响, 本文采用邻域均值替代中心值的方式, 以降低噪声干扰, 提升匹配准确性。

此外, 传统的 Census 变换在强纹理和边缘区域的匹配效果较差, 而 AD 代价在纹理丰富区域表现较好, 但对噪声和光照变化较为敏感。梯度可以较好地识别图像边界, 但在无纹理区域和大面积平坦区域中有较大的误差。为了进一步提升算法的匹配精度和鲁棒性, 本文采用 Sobel 算子计算图像的梯度值, 并通过梯度幅值衡量区域灰度值的变化, 从而反映该区域的纹理复杂度。其梯度幅值的计算公式为:

$$\begin{aligned} G_x(x, y) &= \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * I(x, y) \\ G_y(x, y) &= \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * I(x, y) \\ G(x, y) &= \sqrt{G_x^2(x, y) + G_y^2(x, y)} \end{aligned} \quad (4)$$

式中: $G_x(x, y)$ 、 $G_y(x, y)$ 分别为点 (x, y) 处在 x 、 y 方向上的梯度值; $G(x, y)$ 为点 (x, y) 处的梯度幅值。

为进一步提升算法在图像边缘和深度不连续区域的匹配精度, 本文通过将基于梯度信息的 Census 变换与传统 AD、Census 变换结合来提升匹配的准确性。同时, 为充分发挥 AD 代价、Census 代价以及基于梯度信息的 Census 变换在不同区域的优势, 本文提出的算法根据图像的区域特征自适应地调整不同代价的权重。其融合代价计算公式为:

$$\begin{cases} C(p, d) = W_{AD} \cdot \rho(C_{AD}(p, d), \lambda_{AD}) + \\ \quad W_{Census} \cdot \rho(C_{Census}(p, d), \lambda_{Census}) + \\ \quad W_{Grad} \cdot \rho(C_{Grad}(p, d), \lambda_{Grad}) \\ \rho(C, \lambda) = 1 - \exp\left(-\frac{C}{\lambda}\right) \end{cases} \quad (5)$$

式中: $C(p, d)$ 表示在视差为 d 时, 点 p 处的 3 种代价的融合值; ρ 表示用来控制代价归一化, 范围为 $[0, 1]$; λ_{AD} 、 λ_{Census} 和 λ_{Grad} 分别表示 AD 代价、Census 代价和梯度 Census 变换代价的归一化参数; W_{AD} 表示 AD 权重参数; W_{Census} 表示 Census 变换的权重; W_{Grad} 表示梯度的 Census 变换权重。

传统的代价计算方法使用固定大小的窗口，但在处理图像中的深度不连续区域或纹理复杂度变化较大的区域时，固定的窗口在面对不同区域的图像特征时，存在一定的局限性^[8]。例如，在深度不连续区域和高纹理区域，较小的窗口可以有效捕捉细节信息。而在平坦区域，较大的窗口有助于保证匹配的精度。为克服这一局限性，本文通过 Sobel 算子计算的梯度幅值动态调整窗口大小的方法。将梯度幅值作为区域纹理复杂度和深度不连续的评价指标。当梯度幅值较小时，表示该区域灰度变换较为平缓，即区域纹理信息稀疏，因此在该区域需要选择较大的窗口以提升匹配精度；而当梯度幅值较大时，表示该区域灰度变化剧烈，通常出现在边缘或高纹理区域，此时应该选用较小的窗口。动态变换窗口的大小用公式表示为：

$$\text{Window}(x, y) = \begin{cases} 11 \times 11 & G(x, y) < T_1 \\ 9 \times 9 & T_1 \leq G(x, y) < T_2 \\ 7 \times 7 & T_2 \leq G(x, y) < T_3 \\ 5 \times 5 & G(x, y) \geq T_3 \end{cases} \quad (6)$$

式中：Window(x, y) 表示在点 (x, y) 处对应的窗口大小；G(x, y) 表示在点 (x, y) 处的梯度幅值；T₁、T₂、T₃ 分别表示梯度幅值的阈值。

1.2 代价聚合模型

仅通过像素中心及其邻域窗口的代价计算，容易出现较多的误匹配点，因其未考虑与周边像素之间的相关性，导致最终视差图的质量较差。为此，有必要对初步得到的视差值进行进一步优化和调整，以提高匹配结果的精度和稳定性。对此，本文采用十字交叉法（CBCA）^[9]的代价聚合算法。如图 2 所示，十字交叉法通过将当前位置像素作为中心，沿水平和垂直方向进行扩展。在扩展过程中，当相邻像素与当前像素的颜色差异超过设定阈值时，则停止扩展。

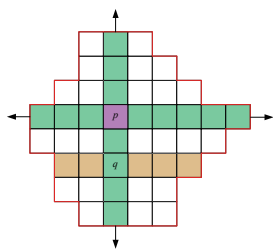


图 2 支持域窗口

其判别用公式表示为：

$$\begin{cases} D_c(p_1, p) < \tau_1 \text{ and } D_c(p_1, p_1 + (1, 0)) < \tau_1 \\ D_s(p_1, p) < L_1 \\ D_c(p_1, p) < \tau_2 \text{ if } L_2 < D_s(p_1, p) < L_1 \end{cases} \quad (7)$$

式中：p 表示当前像素位置，即像素支持域的中心；p₁ 表示当前像素在水平方向或垂直方向上的相邻的任意像素点；

D_c(p₁, p) = |I(p₁) - I(p)| 表示当前像素与相邻像素之间的颜色差异；D_s(p₁, p) = |p₁ - p| 表示两个像素点之间的空间距离；τ₁、τ₂ 表示颜色差异的阈值；L₁、L₂ 表示距离差异的阈值。

传统的十字交叉法代价聚合算法如图 3 所示，通过计算聚合区域内所有像素的代价总和，并将其除以支持域内像素的数量来获取聚合后的代价。这种方式虽然能够有效地整合邻域信息，但未能充分考虑支持域窗口内各个元素对聚合中心的贡献差异，特别是当支持域的臂长较大时，远离聚合中心的像素与聚合中心的相关性较低，其在代价聚合中的影响应予以削弱。

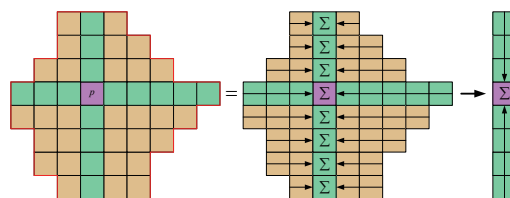


图 3 十字交叉代价聚合

为此，本文采用基于距离权重的十字交叉代价聚合方法。该方法通过为靠近聚合中心的像素分配更高的权重，而对远离中心的像素赋予较低的权重，从而有效提升匹配精度和稳定性。具体而言，首先对十字交叉臂上每个像素根据其距离聚合中心的远近赋予不同的权重，然后基于传统的十字交叉法计算加权后的代价总和。最终，通过将代价总和除以支持域窗口内像素的数量，得到每个像素的聚合代价。基于距离权重的十字交叉聚合公式为：

$$C_{\text{agg}}(p, d) = \frac{1}{N} \sum_{q \in N_p} \rho C(q, d) \quad (8)$$

式中：C_{agg}(p, d) 表示当视差为 d 时，像素点 p 处的聚合代价；q 表示当前像素位置 p 的十字交叉支持域上的像素；N 表示支持域元素总数；ρ 表示距离权重函数，其距离权重的公式为：

$$\rho = 1 - \theta \frac{d - d_{\min}}{d_{\max} - d_{\min}} \quad (9)$$

1.3 视差优化模型

在使用赢家通吃（WTA）算法获取初步的视差结果后，为进一步提高视差图的精度，对视差图进行亚像素插值、左右一致性检查和视差填充等步骤。其中视差图的计算公式为：

$$d(p) = \arg \min_{d \in [d_{\min}, d_{\max}]} C_{\text{agg}}(p, d) \quad (10)$$

式中：d(p) 表示点 p 处的视差值；C_{agg}(p, d) 表示视差为 d 时，点 p 处的聚合代价。

在初步计算得到的视差图中，仍可能存在误匹配点，通过左右一致性检查可以有效剔除这些异常点，用公式表示为：

$$|d_L(p) - d_R\{p - [d_L(p), 0]\}| > \delta_0 \quad (11)$$

式中： $d_L(p)$ 表示左图像中点 p 处的视差值； $d_R\{p-[d_L(p), 0]\}$ 表示右图像中对应点的视差值； δ_0 表示预设的视差阈值。

当左右图像中对应点的视差差值超过该阈值时，则认为该点为异常点并予以剔除。同时，为了进一步提升视差精度，本文采用亚像素插值方法来细化视差值，其计算公式为：

$$d^*(p)=d(p)-\frac{C_{\text{agg}}(p,d+1)-C_{\text{agg}}(p,d-1)}{2(C_{\text{agg}}(p,d+1)-2C_{\text{agg}}(p,d)+C_{\text{agg}}(p,d-1))}$$

(12)

SGM

式中： $d(p)$ 表示像素点 p 处的视差值； $C_{\text{agg}}(p,d+1)$ 、 $C_{\text{agg}}(p,d-1)$ 分别表示视差为 $d+1$ 、 $d-1$ 时像素点 p 处的聚合代价。

最后，通过中值滤波剔除视差图中孤立的离散群点。

2 实验结果与分析

本文采用 C++ 和 CUDA 编程语言实现，开发环境为 Visual Studio 2019，硬件平台为 Intel Core i9-13900H 处理器和 NVIDIA GeForce RTX 4070 GPU。为了验证所提算法的有效性和优越性，采用 Middlebury 标准数据集进行评估，并通过计算误匹配率来衡量算法的匹配精度。其误匹配率的计算公式为：

$$\text{error}=\frac{1}{N}\sum_{(x,y)}(|d(x,y)-d_{\text{GT}}(x,y)|>\delta)$$

(13)

式中： N 表示计算使用的总像素数； $d(x,y)$ 表示计算得到的视差图中点 (x,y) 处的视差值； $d_{\text{GT}}(x,y)$ 表示数据集中标准视差图中点 (x,y) 处的真实视差值； δ 表示预设的视差误差阈值。

2.1 算法对比分析

为验证本文提出的改进立体匹配方法的效果，实验采用 Middlebury^[2] 标准数据集进行测试，并与传统的 SGM 算法^[10]、Census 算法及 AD-Census 算法进行了对比分析。从图 4 中可以看出，在视差不连续区域，尤其是边界部分，SGM 算法、传统 Census 算法和 AD-Census 算法均出现了明显的误匹配现象。而相比之下，本文提出的改进算法在各类测试图像的边界及视差不连续区域相比其他算法显著减少了误匹配现象。尽管仍存在少量误匹配，但总体来看，本文算法生成的视差图质量优于传统的 SGM 算法、Census 算法和 AD-Census 算法，且匹配精度更高。

为客观评价本文所提算法的准确性，本文使用 Middlebury 标准数据集进行测试，通过对比几种算法在非遮挡区域（Nocc）和全部区域（All）的误匹配率，实验结果如表 1 和表 2 所示。

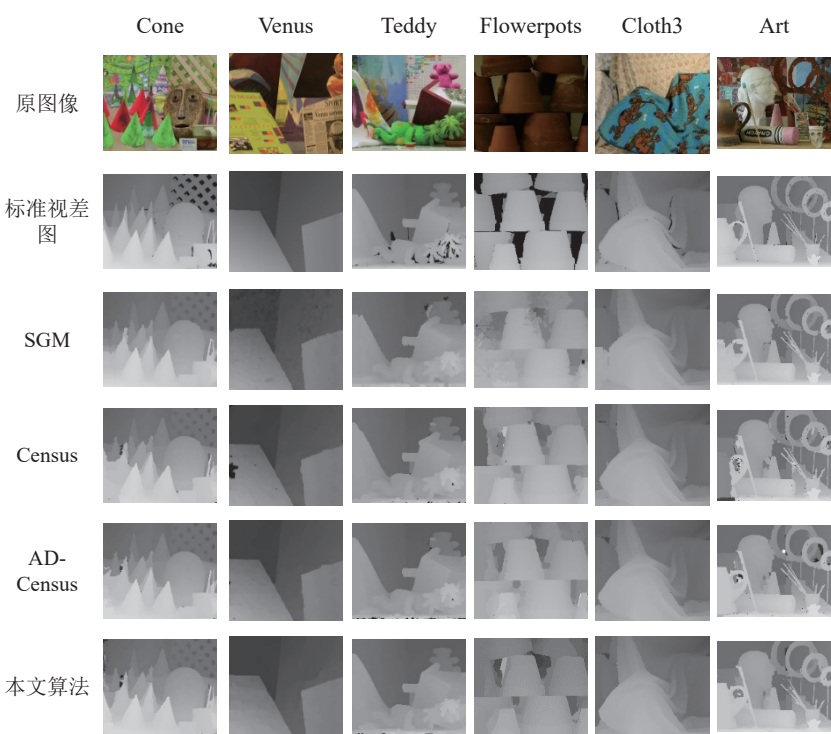


图 4 各算法的实验结果图

表 1 非遮挡区域（Nocc）误匹配率

图片	误匹配率（Nocc）/%			
	SGM	传统 Census	AD-Census	本文算法
Cone	4.33	4.51	4.26	3.41
Venus	2.47	0.99	0.76	0.45
Teddy	6.54	6.33	5.25	4.32
Flowerpots	30.47	23.25	20.59	16.44
Cloth3	6.13	6.64	6.05	5.37
Art	14.72	19.09	18.58	12.51
Baby2	21.94	20.72	19.55	16.33
Rocks2	6.50	10.80	9.79	5.36
平均误差	11.64	11.54	10.60	8.02

表 2 全区域（All）的误匹配率

图片	误匹配率（All）/%			
	SGM	传统 Census	AD-Census	本文算法
Cone	8.60	7.68	6.49	5.36
Venus	3.07	1.69	1.05	0.83
Teddy	8.65	8.72	7.52	5.86
Flowerpots	41.02	34.69	35.73	26.33
Cloth3	7.90	7.65	7.11	6.32
Art	19.07	24.76	24.30	16.53
Baby2	22.61	22.56	21.45	20.35
Rocks2	8.65	11.50	10.57	7.65
平均误差	14.95	14.91	14.28	11.15

通过表 1 和表 2 所示的误匹配率计算结果可知，相比其他算法，本文所提出的改进算法在非遮挡区域和整体区域的

误匹配率均低于SGM算法、Census算法以及AD-Census算法。

2.2 算法的实际三维测量实验

为了验证本文所提算法在实际应用中的测量精度，本文通过对标准件进行三维测量，如图5所示，并将测量结果与实际数据进行比较。实际三维测量结果如表3所示。通过实验分析，其中最小的测量误差仅为0.145 mm，实验结果充分证明了本文提出的算法在实际三维测量中具有较高的测量精度。



图5 实际三维测量效果图

表3 实际三维测量结果

	AB	BC	圆直径
实际距离/mm	100.00	100.00	60.00
测量距离/mm	100.176	99.842	59.855
测量误差/mm	0.176	-0.158	-0.145
误差率/%	0.176	-0.158	-0.242

3 结论

针对双目立体匹配中存在的匹配精度低的问题，本文提出了一种基于动态权重的多特征融合的立体匹配算法。首先，在代价计算阶段，本文通过Sobel算子计算梯度幅值，以区分不同区域的纹理复杂度，并自适应地调整窗口大小，有效地降低了在复杂区域的误匹配率。此外，结合AD代价、Census代价和基于梯度信息的Census变换，并根据区域特征动态地分配权重，进一步降低了高纹理区域及视差不连续区域的误匹配率。在代价聚合阶段，为充分考虑支持域内不同位置对聚合中心的贡献，本文采用了基于距离权重的交叉法进行代价聚合。进一步提升匹配的准确性。同时，为提高立体匹配的运行效率，本文采用了CUDA并行计算技术，显著加速了匹配过程。实验结果证明，在Middlebury标准数据集的遮挡区域和整体区域，本文提出的算法的误匹配率均低于传统方法。同时，在实际的三维测量中，最小的测量误差仅为0.145 mm，充分验证了本文算法在实际应用中的有效性。

参考文献:

[1] 翟国栋, 张文涛, 岳中文, 等. 基于双目视觉技术的煤矿救援机器人研究进展[J]. 煤炭科学技术, 2020, 48(7): 206-217.
[2] 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 (SMBV 2001). Piscataway: IEEE, 2001: 9-10.
[3] 梁培琛. 基于抖动优化的双目三维重建方法[D]. 广州: 华南理工大学, 2021.
[4] LEE J, JUN D, EEM C, et al. Improved census transform for noise robust stereo matching[J]. Optical engineering, 2016, 55(6): 063107.
[5] 王云峰, 吴炜, 余小亮, 等. 基于自适应权重 AD-Census 变换的双目立体匹配[J]. 工程科学与技术, 2018, 50(4): 153-160.
[6] LÜ C, LI J H, KOU Q Q, et al. Stereo matching algorithm based on HSV color space and improved census transform[J/OL]. Mathematical problems in engineering, 2021[2024-05-25]. <https://onlinelibrary.wiley.com/doi/10.1155/2021/1857327>.
[7] PANG J H, SUN W X, REN J S J, et al. Cascade residual learning: a two-stage convolutional neural network for stereo matching[EB/OL]. (2018-07-30)[2024-12-11]. <https://doi.org/10.48550/arXiv.1708.09204>.
[8] 周喆, 沈建新, 韩鹏, 等. 基于 Census 变换和引导滤波的立体匹配算法[J]. 应用光学, 2020, 41(1): 79-85.
[9] ZHANG K, LU J B, LAFRUIT G. Cross-based local stereo matching using orthogonal integral images[J]. IEEE transactions on circuits and systems for video technology, 2009, 19(7): 1073-1079.
[10] HIRSCHMÜLLER H. Stereo processing by semiglobal matching and mutual information[J]. IEEE transactions on pattern analysis and machine intelligence, 2008, 30(2): 32-41.

【作者简介】

陆雨薇(1987—), 女, 山西介休人, 博士, 硕士生导师、副研究员, 研究方向: 机器视觉、算法设计, email: 178108966@qq.com。

常修源(2000—), 男, 河南林州人, 硕士研究生, 研究方向: 机器视觉, email: 750294720@qq.com。

龙辉(1988—), 男, 广西柳州人, 本科, 中级工程师, 研究方向: 工业智能制造、数字孪生, email: longhui@zstarx.com。

覃高杰(2003—), 男, 广西柳州人, 本科在读, 研究方向: 机器视觉, email: 3094281442@qq.com。

(收稿日期: 2025-02-05 修回日期: 2025-06-10)