

DOI:10.19651/j.cnki.emt.2518742

结合纹理区域特性的跨尺度立体匹配算法^{*}卜鹏辉^{1,2} 田隆涛^{1,2} 王航^{1,2} 严亚滔^{1,2}

(1.西安石油大学机械工程学院 西安 710065;2.国家油气钻井装备工程技术研究中心 宝鸡 721002)

摘要: 立体匹配是双目立体视觉感知场景深度信息的关键步骤,针对传统双目立体匹配算法难以有效解决弱纹理区域和复杂光照场景下匹配歧义性的难题,提出了一种结合场景纹理特性的跨尺度立体匹配算法。首先将左右图像进行高斯下采样,得到多种尺度的图像对作为算法的输入图像,然后对不同尺度的图像对进行代价计算得到初始代价体;基于纹理特性将输入图像划分为纹理丰富区域和弱纹理区域,根据纹理区域对初始代价体在各个尺度下进行代价扩散,将纹理丰富区域的匹配代价扩散到弱纹理区域;用优化引导滤波算法对各个尺度视差图进行代价聚合;考虑代价体的多尺度间交互进行代价融合,得到最终代价体;随后经过视差计算和视差后处理得到最终视差图。Middlebury网站的数据集的测试结果表明,引入结合纹理区域特性的跨尺度立体匹配算法后,全部区域的误匹配率相较于引导滤波算法平均降低了2.35%;相较于CSCA算法平均降低了0.77%。非遮挡区域的误匹配率相较于引导滤波算法平均降低了2.29%;相较于CSCA算法平均降低了0.65%。说明所提算法能有效解决弱纹理区域的误匹配问题,并且满足立体匹配过程中高效率、高精度的要求。

关键词: 立体匹配;区域划分;代价扩散;多尺度融合

中图分类号: TN911.73 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Cross-scale stereo matching algorithm combining the characteristics of texture regions

Bu Penghui^{1,2} Tian Longtao^{1,2} Wang Hang^{1,2} Yan Yatao^{1,2}

(1. School of Mechanical Engineering, Xi'an Shiyou University, Xi'an 710065, China;

2. Shanxi Provincial Petroleum Drilling Equipment Engineering Technology Research Center, Baoji 721002, China)

Abstract: Stereo matching is a key step in binocular stereo vision to perceive the depth information of the scene, and in view of the difficulty of the traditional binocular stereo matching algorithm to effectively solve the problem of matching ambiguity in weak texture areas and complex lighting scenes, a cross-scale stereo matching algorithm combining the texture characteristics of the scene was proposed. Firstly, the left and right images are downsampled by Gaussian to obtain image pairs of multiple scales as the input images of the algorithm, and then the cost calculation of image pairs of different scales is carried out to obtain the initial cost body. Based on the texture characteristics, the input image is divided into texture-rich region and weak-texture region, and the initial cost body is difed at each scale according to the texture region, and the matching cost of the texture-rich region is diffused to the weak-texture region. The optimization guidance filtering algorithm is used to aggregate the cost of the parallax map of each scale. Considering the multi-scale interaction between the cost bodies, the cost fusion is carried out to obtain the final cost body. Subsequently, the final disparity map is obtained by parallax calculation and parallax post-processing. The test results of the dataset of Middlebury website show that after the introduction of the cross-scale stereo matching algorithm combined with the characteristics of texture regions, the mismatching rate of all regions is reduced by 2.35% on average compared with the guided filtering algorithm. Compared with the CSCA algorithm, it is reduced by 0.77% on average. Compared with the guided filtering algorithm, the mismatching rate of the unoccluded region is reduced by 2.29% on average. Compared with the CSCA algorithm, it is reduced by 0.65% on average. It shows that the proposed algorithm can effectively solve the problem of mismatching in weak texture regions, and meet the requirements of high efficiency and high precision in the process of stereo matching.

Keywords: stereo matching; region division; match cost spread; multi-scale fusion

0 引言

立体视觉通过模拟人类的双眼视觉机制来感知场景的

深度信息,是三维场景理解和重建的核心技术之一,在智能驾驶、机器人视觉、三维重建、虚拟现实等领域的应用愈发广泛^[1-3]。立体视觉的关键在于立体匹配,它旨在通过比对

收稿日期:2025-05-06

* 基金项目:国家自然科学基金(52074346)、陕西省重点研发计划(2023-YBGY-103)项目资助

立体图像对中的像素点,计算视差图,从而推断场景中的深度信息^[4-5]。

文献[6]将立体匹配划分为代价计算、代价聚合、视差计算和视差优化 4 个步骤,至今仍为多数学者沿用;根据代价计算方式的不同,也可分为全局算法、半全局算法、局部算法和基于深度学习的算法,基于深度学习神经网络的匹配算法通过非线性变换提取图像多层特征来计算匹配代价,通过训练得到模型的参数,提高了算法鲁棒性。但由于需要大量数据进行计算,在实际运用中受到很大限制^[7-9],相比之下,传统算法无需大量数据训练,准确率也满足实际需求,得到更加广泛的应用^[10]。全局算法为尽可能多地获取全局信息,通过全局能量函数来遍历所有像素点,目前常见的全局算法有图割(graph cuts, GC)^[11]、动态规划(dynamic programming, DP)^[12]以及置信传播(belief propagation, BP)^[13],其虽能获得较高质量的视差图,但算法效率低,难以得到实际应用。半全局算法^[14](semi-global matching, SGM)从像素或窗口级别出发计算匹配代价,通过路径代价累积模拟能量最小化中的平滑性约束,部分实现全局一致性,其介于局部算法和全局算法之间,以较低的计算复杂度实现了接近全局优化的效果,是一种效率和精度折中的方法。局部算法通过像素或局部区域计算匹配代价,然后直接选取最佳视差;在计算时主要关注局部信息,其算法简单、计算效率高、适合实时应用,但在遮挡和弱纹理区域表现相对较差^[15]。双边滤波器(bilateral filter, BF)引入匹配代价聚合环节后,表现出较高的边缘保持性^[16];基于十字交叉臂分割方法(cross-based cost aggregation, CBCA)构造简单且精度较高^[17];引导滤波(guided filter, GF)相比双边滤波有较好的边缘保持性和较高的计算效率^[18];跨尺度立体匹配算法(cross-scale cost aggregation for stereo matching, CSCA)有效弥补了传统单尺度代价聚合的不足,在遮挡处理、纹理稀疏区域和复杂光照场景中表现尤为出色^[19]。近年来,局部算法得到极大发展, Xu 等^[20]提出了一种八边形滤波器算法,其性能稳定而且速度较快,畅雅雯等^[21]提出的多特征融合的代价计算和自适应十字窗口聚合的立体匹配算法,有效改善了视差边界和弱纹理区域错误匹配的问题,王笛等^[22]提出了一种改进的特征立体匹配算法,通过双向匹配策略实现区域生长,提升了匹配准确率,陈艺等^[23]提出了一种自适应纹理区域的多尺度融合立体匹配算法,该算法对不同纹理区域具有很好的鲁棒性,吴福培等^[24]将基于十字交叉法的 SGM 聚合与改进引导滤波相融合,具有较强的边缘保持和噪声平滑的能力。各种高效算法的提出,使得立体匹配算法的匹配精度和匹配效率得到了很大发展,但目前各类立体匹配算法仍然受困于遮挡、光照失真和弱纹理等现象,所以针对性地改善各类匹配困难区域的匹配精度是当前的研究热点。

引导滤波是经典的滤波处理器,具有很高的边缘保持性和计算效率,而跨尺度代价聚合算法(CSCA)的多尺度

信息融合在处理噪声、遮挡和纹理弱区域时表现优越,同时能够增强全局一致性;结合两种算法的优点,本文提出了一种结合纹理区域特性的跨尺度立体匹配算法,针对弱纹理区域匹配难度大、匹配精度低的问题,结合引导滤波的边缘保持性和计算效率,以及 CSCA 的多尺度信息融合功能,改善弱纹理区域的匹配效果的同时提升了遮挡、复杂光照场景的匹配效果,增强了视差图的全局一致性。

1 算法描述

本文针对立体匹配弱纹理区域和复杂光照场景下的匹配歧义性问题提出了结合纹理区域特性的跨尺度立体匹配算法,主要采用了两方面策略以提升弱纹理区域的匹配精度,第一,将初始代价体进行了代价扩散操作,将每个视差级下的初始视差图中纹理丰富区域的代价扩散到弱纹理区域,提升了初始视差图的视差精度。第二,利用粗细尺度交互处理的思想,使用跨尺度代价聚合模型,优化不同尺度下的代价聚合精度。算法的整体流程如图 1 所示,其主要流程如下:

- 1)对输入的左右图像高斯金字塔下采样操作,得到各层级待处理图像。
 - 2)使用强度+梯度截断阈值组合算法分别对金字塔各层级进行代价计算操作,得到初级代价体合集。
 - 3)对每个初级代价体的弱纹理区域进行基于区域纹理特性的代价扩散操作,得到二级代价体合集。
 - 4)使用引导滤波核对每个二级代价体进行代价聚合操作,得到三级代价体合集。
 - 5)对各个层级的三级代价体进行跨尺度代价融合,得到最终代价体。
 - 6)由最终代价体计算得到视差图并进行后处理操作。
- 经过视差计算和后处理最终得到结合区域纹理特性的跨尺度立体匹配视差图。

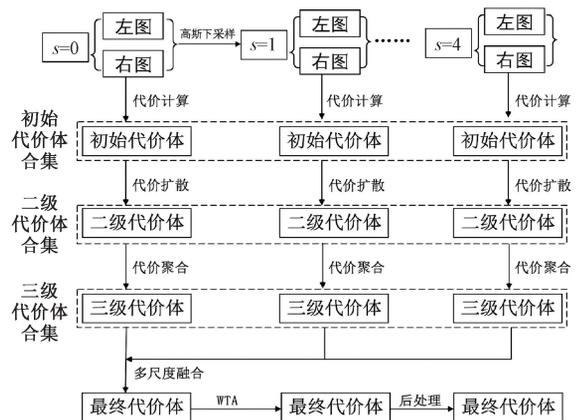


图 1 算法整体流程图

Fig. 1 The overall flow chart of the algorithm

其中,代价扩散是针对弱纹理区域的所有弱纹理点进行代价优化,当所有弱纹理点都被更新后,则终止迭代,

WTA表示赢家通吃策略,具体将在1.4节中介绍。对初始图像进行高斯下采样时, s 表示高斯金字塔的层级,对于高斯金字塔,其高斯下采样示意图如图2所示,第 s 层级与 $s-1$ 层级图像之间满足:

$$I_s(x, y) = \sum_{i=-r}^r \sum_{j=-r}^r I_{s-1}(2x+i, 2y+j) \cdot G(x, y, \sigma) \quad (1)$$

其中, $I_s(x, y)$ 表示第 s 层图像在 (x, y) 处的像素值, $I_{s-1}(2x+i, 2y+j)$ 表示第 $s-1$ 层图像在 $(2x+i, 2y+j)$ 处的像素值。即 I_s 的值通过下采样 I_{s-1} 的偶数位置 $(2x+i, 2y+j)$ 及其周围区域计算得到, $G(x, y, \sigma)$ 为高斯核, $G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{i^2+j^2}{2\sigma^2}\right)$ 是高斯分布函数,决定像素点的权重。 σ 为高斯核标准差,决定模糊程度,值越大就越模糊。 r 为高斯核半径,决定卷积范围。



图2 高斯下采样示意图
Fig. 2 Gaussian downsampling

1.1 匹配代价计算

匹配代价计算环节采用强度和梯度加权计算的方法得到初始匹配代价,同时设置截断阈值防止结果出现突兀,初始代价图如图3所示。对于一组图像对而言,左图中第 i 像素点,即坐标为 (x_i, y_i) 的像素在视差等级为 l 时的匹配代价 $C(i, l)$ 可以表示为:

$$C(i, l) = (1-\alpha) \cdot \min[\|I(i) - I'(i_l)\|, \tau_1] + \alpha \cdot \min[\|\nabla_x I(i) - \nabla_x I'(i_l)\|, \tau_2] \quad (2)$$

其中, i_l 是右图中第 i 像素点,其像素坐标为 $(x_i - l, y_i)$, $I(i)$ 和 $I'(i_l)$ 分别表示像素 i 和像素 i_l 的强度值, $\nabla_x I(i)$ 和 $\nabla_x I'(i_l)$ 分别表示像素 i 和像素 i_l 的梯度值, τ_1 和 τ_2 表示截断阈值, α 平衡强度和梯度的比例。代价计算后, $s=0$ 层级初始代价体在视差为 l 是的初始代价图如图3所示。

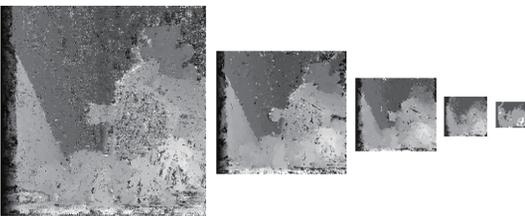


图3 初始代价图
Fig. 3 Initial cost chart

1.2 匹配代价扩散

初始视差图中弱纹理区域存在大量误匹配点,以往的操作是直接对初始代价图进行代价聚合,这样很难对弱纹理区域的匹配代价起到较大的改善,所以本文设计在代价聚合前先利用原图像的灰度图像梯度的局部方差设定阈值将图像划分为纹理丰富区域和弱纹理区域,纹理丰富区域一般具有相对可靠的匹配精度,而弱纹理区域的匹配结果不可靠,于是对弱纹理区域的像素点进行优化处理,将纹理丰富区域像素点的匹配代价扩散到弱纹理区域,从而达到改善弱纹理区域的匹配精度的目的,这就使得后续的代价聚合阶段能取得更加精确的聚合结果。纹理区域划分原理可以表示为:

$$\sigma^2 = \begin{cases} \frac{\sum (G - \mu)^2}{N \times N} & \text{if } \sigma^2 \geq \tau, \text{richregions} \\ & \text{if } \sigma^2 < \tau, \text{weakregions} \end{cases} \quad (3)$$

其中, σ^2 表示窗口内梯度的局部方差, G 表示像素点的梯度, μ 表示窗口内梯度的标准差, τ 表示方差阈值。纹理区域划分的结果如图4所示。



图4 纹理区域划分效果图
Fig. 4 Effect of texture area division

得到区域划分结果后,为了排除窗口内可能不存在纹理丰富点或纹理丰富点太少的情况,本文设计了一种筛选机制,即筛选含有足够数量纹理丰富点的窗口进行扩散操作,若窗口内不存在纹理丰富点或者数量太少,则在下一次迭代时重新筛选,被优化后的弱纹理点将自动变为纹理丰富点,直到所有的弱纹理点都被优化更新则停止迭代。基于上述迭代过程,弱纹理区域的所有像素点都将通过代价扩散得到鲁棒的匹配代价。匹配代价扩散的整体流程如图5所示。代价扩散所使用的核函数如下:

$$\omega_s = \exp\left(-\frac{(i-k)^2}{2s_s^2}\right) \quad (4)$$

$$\omega_r = \exp\left(-\frac{\|f(i, j) - f(k, l)\|^2}{2s_r^2}\right) \quad (5)$$

$$\begin{cases} \omega = \omega_s \times \omega_r \\ \omega_{(i, j) \in \text{weakregions}} = \omega_{(i, j) \in \text{weakregions}} \times 0.0001 \end{cases} \quad (6)$$

其中, ω_s 为空域权重, ω_r 为值域权重, (i, j) 为像素点的坐标, (k, l) 为中心像素点的坐标。 s_s 为空域标准差, $f(i, j)$ 和 $f(k, l)$ 为像素点视差值, s_r 为值域标准差。 ω

表示将空域和值域权重组合后的权重。

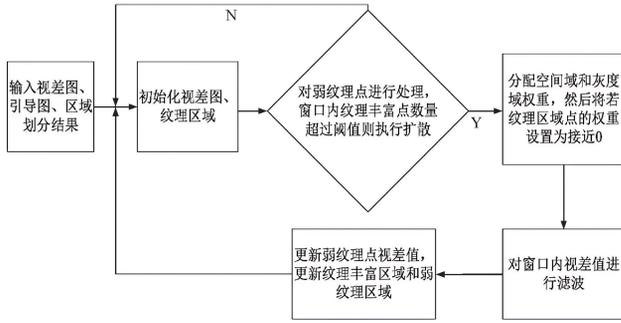


图 5 代价扩散流程图

Fig. 5 Cost diffusion flow chart

在对初始代价体合进行代价扩散操作后,得到二级代价体合集,不同层级经过代价扩散得到的代价图如图 6 所示。

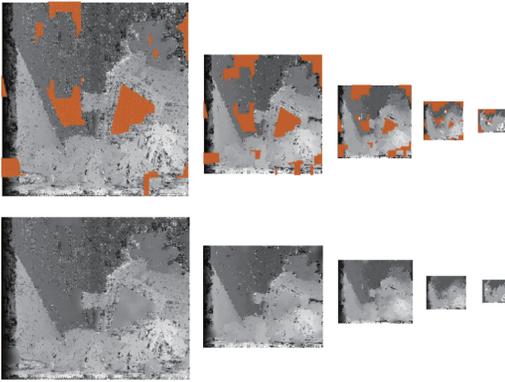


图 6 代价扩散结果

Fig. 6 Cost diffusion results

1.3 匹配代价聚合

1) 代价聚合模型

二级代价体合集集中的每个视差等级上的代价图在弱纹理区域仍然存在很多噪声和误匹配的情况,需要对其进行优化处理,本文采用对每个代价体的每个视差等级进行跨尺度交互的匹配代价聚合模型,主要涉及尺度内和尺度间的处理,采用加权最小二乘法 (weighted least squares, WLS),配合引导滤波核进行代价聚合,加权最小二乘框架可以表示为:

$$\tilde{C}(i, l) = \underset{z}{\operatorname{argmin}} \frac{1}{Z_i} \sum_{j \in N_i} k(i, j) \| z - c(j, l) \|^2 \quad (7)$$

其中, $Z_i = \sum_{j \in N_i} k(i, j)$ 为正则化常量, N_i 为像素 i 的邻域, $k(i, j)$ 为相似核,表示像素 i 和像素 j 的相似度,本文引入引导函数核进行代价聚合, $\tilde{C}(j, l)$ 为邻域 N_i 中第 j 像素在视差为 l 时的代价值, $\tilde{C}(i, l)$ 是第 i 像素在视差为 l 时经过聚合后的代价值。该 WLS 问题的最优解为:

$$\tilde{C}(i, l) = \frac{1}{Z_i} \sum_{j \in N_i} k(i, j) c(j, l) \quad (8)$$

以上为尺度内的聚合模型,为进一步改善弱纹理区域匹配精度,引入代价体尺度间的影响,即考虑尺度间联系的正则化项。在 $0 \sim s$ 尺度间,最优化问题表示为:

$$\hat{V} = \underset{\{Z^s\}_{s=0}^S}{\operatorname{argmin}} \sum_{s=0}^S \frac{1}{Z^s} \sum_{i^s, j^s \in N_{i^s}} k(i^s, j^s) \| Z^s - C^s(j^s, l^s) \|^2 + \lambda \sum_{s=1}^S \| Z^s - Z^{s-1} \| \quad (9)$$

其中, $\{Z^s\}_{s=0}^S$ 表示各代价体合集, $\{i^s\}_{s=0}^S$ 和 $\{l^s\}_{s=0}^S$ 表示各尺度下的强度值和视差值, i^0 表示在最精细层的第 i 像素强度值, i^s 表示 i^0 在第 s 下采样层对应的像素, j^s 表示 s 尺度下邻域 N_{i^s} 中的第 j 像素, N_{i^s} 表示在 s 层级下 i^s 的邻域像素集合, $k(i^s, j^s)$ 为衡量 i^s 和 j^s 的相似度的相似核, $C^s(j^s, l^s)$ 表示 j^s 在视差等级为 l^s 下的代价值, $\sum_{s=1}^S \| Z^s - Z^{s-1} \|$ 为正则化项,考虑了各尺度下的同一像素位置的一致性, Z^s 为 s 尺度下结合多尺度信息的目标代价体, Z^{s-1} 为 $s-1$ 尺度下结合多尺度信息的目标代价体, λ 为控制正则化强度的常量参数;另外:

$$\begin{cases} Z_{i^s}^s = \sum_{j^s \in N_{i^s}} k(i^s, j^s) \\ \hat{V} = [z^0, z^1, \dots, z^s]^T \end{cases} \quad (10)$$

其中, $Z_{i^s}^s$ 表示 s 层级下的正则化常量, \hat{V} 表示各尺度下结合多尺度信息的代价体合集,即为所求目标。将式(9)表示为关于 Z 的函数:

$$F(\{Z^s\}_{s=0}^S) = \underset{\{Z^s\}_{s=0}^S}{\operatorname{argmin}} \sum_{s=0}^S \frac{1}{Z^s} \sum_{i^s, j^s \in N_{i^s}} k(i^s, j^s) \| Z^s - C^s(j^s, l^s) \|^2 + \lambda \sum_{s=1}^S \| Z^s - Z^{s-1} \| \quad (11)$$

对于 $S \in \{1, 2, \dots, S-1\}$, 将 $F(\{Z^s\}_{s=0}^S)$ 对 Z^s 求偏导得:

$$\frac{\partial F}{\partial Z^s} = \frac{2}{Z^s} \sum_{i^s, j^s \in N_{i^s}} k(i^s, j^s) (Z^s - C^s(j^s, l^s)) + 2\lambda(Z^s - Z^{s-1}) - 2\lambda(Z^{s+1} - Z^s) = 2(-\lambda Z^{s-1} + (1+2\lambda)Z^s - \lambda Z^{s+1} - \tilde{C}(j^s, l^s)) \quad (12)$$

令 $\frac{\partial F}{\partial Z^s} = 0$, 可得:

$$\tilde{C}(j^s, l^s) = -\lambda Z^{s-1} + (1+2\lambda)Z^s - \lambda Z^{s+1} \quad (13)$$

同理可得 $s=0$ 和 $s=S$ 的表达式,在考虑多个尺度时,用矩阵运算式表达整个运算则为:

$$\mathbf{A} \hat{V} = \tilde{\mathbf{V}} \quad (14)$$

其中, $\tilde{\mathbf{V}}$ 是由各个尺度下的经过代价聚合的代价体组成的向量, $\tilde{\mathbf{V}} = [\tilde{C}^0, \tilde{C}^1, \dots, \tilde{C}^S]^T$; \hat{V} 为考虑多尺度信息后的目标代价体向量, $\hat{V} = [Z^0, Z^1, \dots, Z^S]^T$, \mathbf{A} 是 $(S+1) \times (S+1)$ 的对角阵,对其进行逆变换得到:

$$\hat{\mathbf{V}} = \mathbf{A}^{-1} \hat{\mathbf{V}} \quad (15)$$

从 $\hat{\mathbf{V}}$ 可以得到最终代价,即 $\hat{\mathbf{V}}$ 在 0 层的代价。

2) 引导滤波

前文提到,代价聚合模型中用到的算法为引导滤波器,引导滤波是一种具有良好的保持边缘特性,能够取得较好的代价聚合效果的滤波算法,同时引导滤波的处理效率相较于全局算法和半全局算法以及双边滤波等算法都更有优势。引导滤波假设图像的局部特性可以用一个线性模型来描述:

$$q_i = a_k \cdot I_i + b_k \quad \forall_i \in \omega_k \quad (16)$$

其中, I 为引导图像,通常为输入图像本身,和待处理图像 p 一起输入, q 为输出图像, I_i 和 q_i 分别为引导图像和输出图像中第 i 像素的强度值; ω_k 为第 k 个半径为 r 的窗口, a_k 和 b_k 为 ω_k 的线性系数。将 p 和 I 下采样后得到 p^s 和 I^s 。其中 p^s 是 s 层的输入图像, I^s 是 s 层的引导图像,为确定线性系数, s 尺度下的输入图像 p^s 和输出图像 q^s 差异最小的最优化问题为:

$$E(a_k, b_k) = \sum_{i \in \omega_k^s} ((a_k I_i^s + b_k - p_i^s)^2 + \epsilon a_k^2) \quad (17)$$

其中, a_k^s 和 b_k^s 为 s 尺度下第 k 窗口的线性系数, ϵ 为正则化因子,防止 a_k^s 过大, ω_k^s 为 s 尺度下的窗口, I_i^s 和 p_i^s 分别为 s 尺度下引导图像和输入图像在窗口 ω_k^s 中第 i 像素的像素强度值,解最小二乘问题可以求得 a_k^s 和 b_k^s :

$$\begin{cases} a_k^s = \frac{\frac{1}{|\omega_k^s|} \sum_{i \in \omega_k^s} I_i p_i^s - \mu_k \bar{p}_k^s}{\sigma_k^2 + \epsilon} \\ b_k^s = \bar{p}_k^s - a_k^s \mu_k \end{cases} \quad (18)$$

其中, \bar{p}_k^s 为 s 尺度下引导图像在 k 窗口的均值, μ_k 和 σ_k 为输入图像在 s 尺度下的 k 窗口下的均值和方差; ϵ 为正则化因子。然后求得 a_k^s 和 b_k^s 在包含 i 的窗口 ω_k^s 中的平均值:

$$\begin{cases} \bar{a}_i^s = \frac{1}{|\omega_k^s|} \sum_{k \in \omega_i^s} a_k^s \\ \bar{b}_i^s = \frac{1}{|\omega_k^s|} \sum_{k \in \omega_i^s} b_k^s \end{cases} \quad (19)$$

将 \bar{a}_i^s 和 \bar{b}_i^s 上采样后得到对应原图第 i 像素处的系数值 \bar{a}_i 和 \bar{b}_i , 可求得滤波后第 i 像素强度 q_i :

$$q_i = \bar{a}_i I_i + \bar{b}_i \quad (20)$$

在本文算法框架下,输入图像为个尺度下经过代价扩散后的代价体,引导图像为各尺度下的左图,得到的匹配代价聚合后的代价图如图 7 所示,经过多尺度代价融合后得到的视差图如图 8 所示。

1.4 视差计算及视差后处理

1) 视差计算

视差计算最常用是赢家通吃法 WTA (Winner-Takes-All), 如图 9 所示, 通过对比该位置在视差范围内的所有代

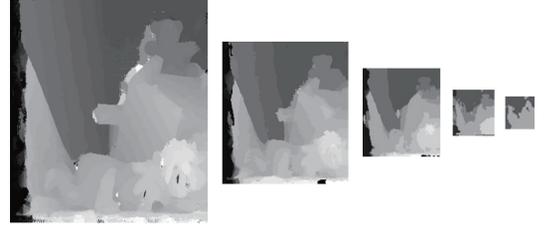


图 7 代价聚合结果

Fig. 7 Cost aggregation results

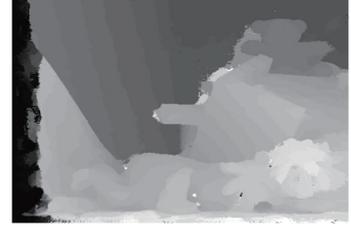


图 8 多尺度代价聚合结果

Fig. 8 Results of multi-scale cost aggregation

价值,选择最小代价值对应的视差值作为最佳匹配视差值,图中 cost 表示匹配代价值,parallax value 为视差值。其公式为:

$$d_{best} = \underset{0 < l < l_{max}}{\operatorname{argmin}} C(j, l) \quad (21)$$

其中, d_{best} 为最佳视差, $C(j, l)$ 是第 j 像素在视差等级 l 下的代价值, l_{max} 为最大视差。

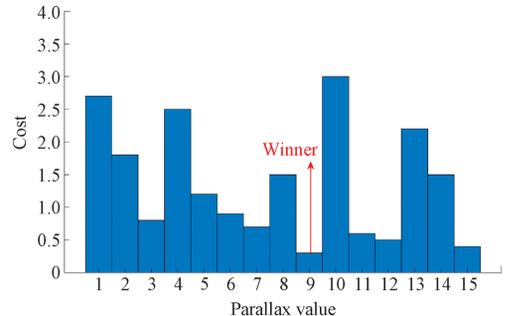


图 9 赢家通吃 ("WTA")

Fig. 9 Winner-takes-all ("WTA")

2) 视差后处理

视差后处理即对得到的视差值进行优化处理,有左右一致性检查(left-right consistency, LRC)、视差填充和中值滤波等方法,能够使得视差值更加接近真实值,本文采用中值滤波算法和左右一致性检查相结合对视差图进行优化处理,得到最终的视差图如图 10 所示。

2 参数分析与实验结果

本文实验数据采用图像全部区域和非遮挡区域的误匹配率来衡量算法的有效性,较低的误匹配率表示算法具有较高的精确性和鲁棒性,能更好地处理各类复杂场景和异常情况。为检验算法效果,本文选用 Middlebury 网站中历



图 10 视差后处理结果

Fig. 10 Parallax post-processing results

年测试图像对中的 10 个图像对来验证算法的有效性,然后将得到的误匹配效果与引导滤波和跨尺度代价聚合算法(CSCA)进行对比分析,从客观角度评判本文算法对误匹配率的改善效果,表征其对代价聚合效果的提升作用。首先,通过误匹配率和视差图误匹配区域大小客观对比本文算法相较于引导滤波算法以及 CSCA 算法的优势,其次,借助视差图效果直观反应不同类型算法的匹配效果,以此验证本文算法相较于其他算法的先进性,最后,通过与不同类型算法视差图进行对比,验证本文算法的高效性。

2.1 参数分析

由于不同图像对的纹理特性和光照条件的不同,在同样窗口大小时,各图像对分别在不同的阈值下得到最佳效果。图 11 为 3 组图像对在同一窗口尺寸下不同的方差阈值对图像非遮挡区域误匹配率的影响,Teddy、Cones 和 Art 分别在方差阈值为 0.023、0.048 和 0.028 时取得最佳扩散效果。可以看出,每组图像对在阈值过小时误匹配率都没有明显改变,在阈值小于最佳数值时,误匹配率随阈值的增加逐渐降低,当阈值超过最佳数值时,误匹配率开始逐渐升高,而当阈值过大使得弱纹理区域包含边界区域时,可能导致在纹理代价扩散时出现错误的扩散或破坏边界。本文借助图像梯度的局部方差来设定阈值进行区域划分,在实际应用中,只要计算出图像梯度的局部方差,就可判断出合适的阈值范围,并验证选择合适的方差阈值进行区域划分。

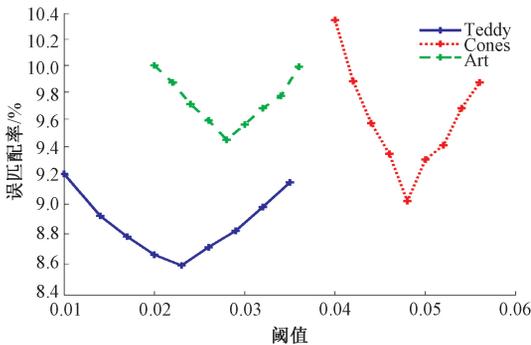


图 11 阈值对结果的影响

Fig. 11 Effect of threshold on results

2.2 实验结果

引导滤波算法是一种具有高效边缘保持性的滤波方法,跨尺度立体匹配算法(cross-scale cost aggregation for stereo matching,CSCA)则是一种集高精度、高边缘保持性

和高效性的一种立体匹配算法;所以在实验过程中,为了客观对比数据的改变效果,将本文算法利用 visual studio 软件配合 opencv 库进行程序编程,然后将本文算法运行结果与引导滤波算法和 CSCA 算法的结果进行数据比较,表 1 给出了在误差阈值为 2.0 情况下,引用 Middlebury Stereo 数据集中覆盖弱纹理、复杂光照条件的 10 个图像对,分别采用引导滤波算法、CSCA 算法和本文算法的全部区域和弱纹理区域的误匹配率的数据对比,Non-occluded 表示非遮挡区域,All 表示全部区域。图 12 中 GF-OURS 表示本文算法误匹配率相对于引导滤波算法在全部区域的降低情况,CSCA-OURS 并表示本文算法误匹配率相对于 CSCA 算法在全部区域的降低情况;图 13 中,GF-OURS 表示本文算法误匹配率相对于引导滤波算法在非遮挡区域的降低情况,CSCA-OURS 并表示本文算法误匹配率相对于 CSCA 算法在非遮挡区域的降低情况。图 14 则给出了 3 种算法下部分图像对的误匹配区域情况。同时,为了进一步验证算法的性能,设计将本文算法分别与全局算法和深度学习算法的效果图进行比较,其中全局算法为 BP 算法,基于深度学习的算法为 GC-Net 算法。从而验证本文算法的匹配精度,对比效果如图 15 所示。

表 1 算法误匹配率对比

Table 1 Comparison of mismatch rates of algorithms

图像对	GF		CSCA		Ours	
	Noc	All	Noc	All	Noc	All
Teddy	7.93	10.19	6.56	9.46	5.40	9.27
Cones	7.48	12.44	6.34	11.58	5.71	9.55
Art	8.32	17.49	7.04	15.47	6.52	14.32
Aloe	8.37	10.16	7.52	9.31	6.77	8.24
Baby1	3.62	4.88	2.67	3.96	2.35	3.48
Adirondack	6.59	8.95	3.30	5.42	2.47	4.59
Jadeplant	10.57	23.23	10.31	22.85	9.34	22.51
Motorcycle	4.76	10.51	4.21	9.61	3.95	9.25
Piano	14.42	22.37	9.41	18.01	9.03	17.45
Recycle	6.81	9.01	5.10	7.79	4.44	7.06

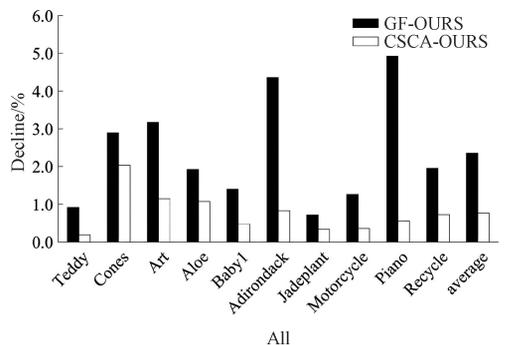


图 12 全部区域误匹配率降低情况

Fig. 12 Decrease in the mismatch rate in all regions

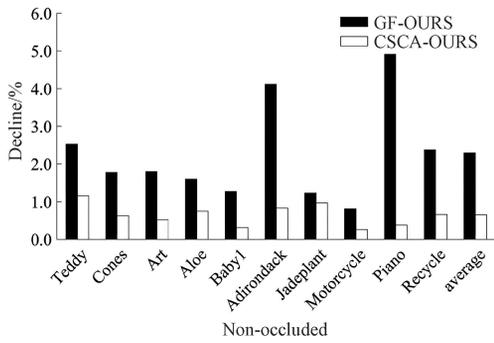


图 13 非遮挡区域误匹配率降低情况

Fig. 13 Decrease in the false matching rate in the unoccluded area

基于以上数据结果,在使用所选的 10 组图像对进行立体匹配得到的视差图后,本文算法与引导滤波和跨尺度立体匹配算法相比;全部区域的误匹配率:本文算法相对于引导滤波算法平均降低了 2.35%,相对于 CSCA 算法平均降低了 0.77%。非遮挡区域的误匹配率:本文算法相对于引导滤波算法平均降低了 2.29%,相对于 CSCA 算法平均降低了 0.65%。所以,本文算法相对于引导滤波算法和 CSCA 算法具备更高的匹配精度,且能有效提高弱纹理区域匹配精度,同时,本文算法的误匹配区域明显少于引导滤波算法和 CSCA 算法,进一步验证了本文算法的匹配效果优于这两种经典算法。

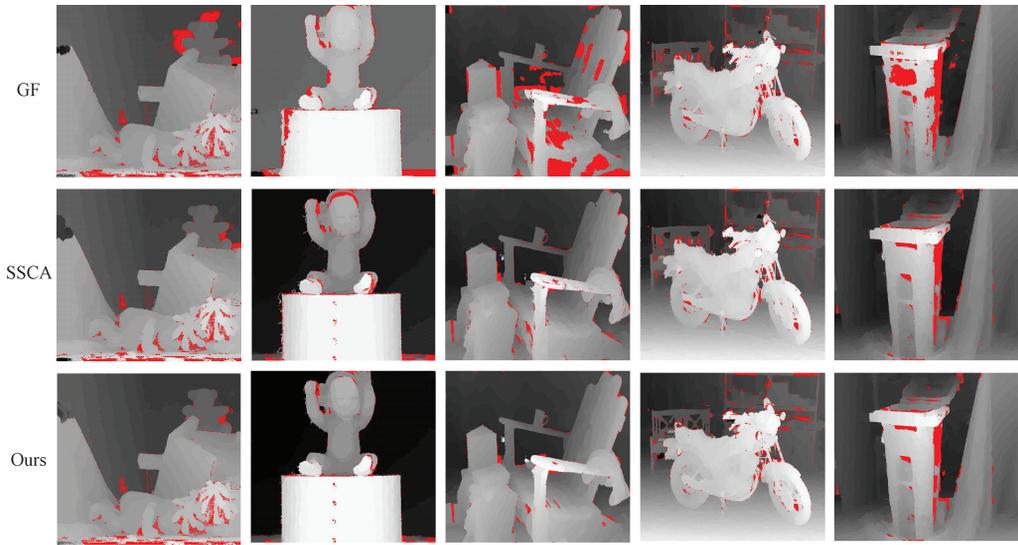


图 14 最终视差效果对比

Fig. 14 Comparison of the final parallax effect

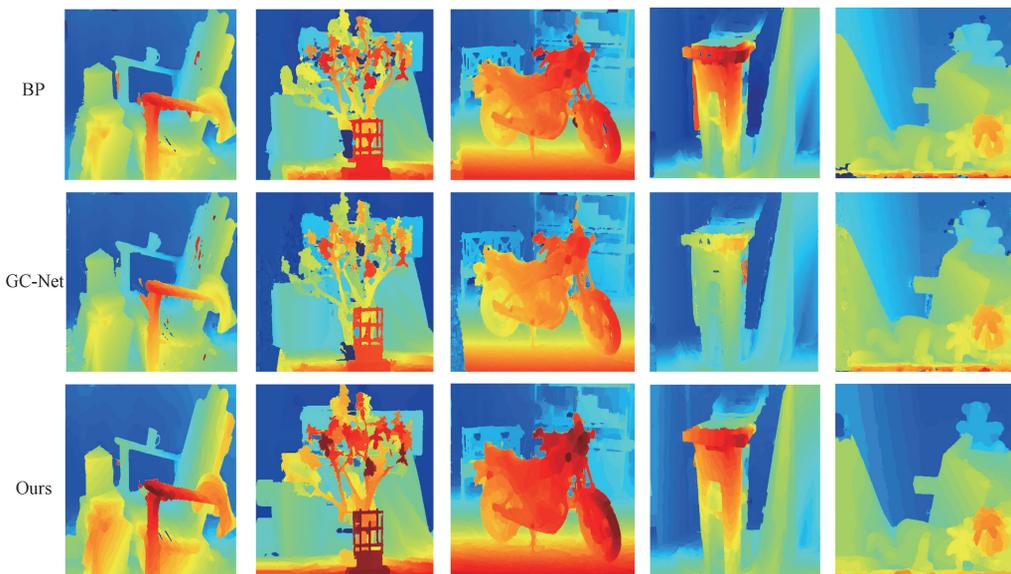


图 15 不同类型算法视差图效果对比

Fig. 15 Comparison of parallax plots of different types of algorithms

另外,全局算法虽匹配效果好,但受制于低下的计算效率和计算复杂度,基于深度学习的算法虽具备较高的全局一致性,但需要大量的训练数据,而本文算法不但有较高的全局保持性,同时具备很好的边缘保持性和计算效率。就图15的最终视差图效果而言,本文算法匹配效果接近于全局算法(BP)且优于深度学习算法(GC-Net)。由此,结合纹理区域特性的立体匹配算法,有效改善了弱纹理区域的匹配效果,同时提升了遮挡、复杂光照场景的匹配效果,增强了视差图的全局一致性。

3 结 论

针对立体匹配中图像弱纹理区域和复杂光照场景下匹配难度大、匹配精确度低的问题,本文提出一种结合纹理区域特性的跨尺度立体匹配算法,通过将图像划分为纹理丰富区域和弱纹理区域,针对性的改善弱纹理区域的匹配代价,同时通过建立跨尺度代价聚合模型,融合多尺度信息,进一步改善匹配代价的可靠性,获得了更加精确的视差图,并将本文算法在全部区域和非遮挡区域的误匹配率与目前广泛引用的高效边缘保持滤波算法进行了分析比较,同时将本文算法最终效果与不同类型算法进行的比较,验证了本文算法对弱纹理区域和光照条件复杂场景视差计算的高精度效果。

参考文献

- [1] 章毓晋. 图象工程: 图象理解与计算机视觉. 下册[M]. 北京:清华大学出版社有限公司, 2000.
ZHANG Y J. Image engineering: Image understanding and computer vision. Volume II[M]. Beijing: Tsinghua University Press Co., Ltd., 2000.
- [2] 张正友, 马颂德. 计算机视觉: 计算理论与算法基础[M]. 北京:北京科学出版社, 1998.
ZHANG ZH Y, MA S D. Computer vision: The basis of computational theory and algorithms[M]. Beijing: Beijing Science Press, 1998.
- [3] 李云廷. 基于立体视觉的三维精确测量方法研究[D]. 武汉:华中科技大学, 2016.
LI Y T. Research on three-dimensional accurate measurement method based on stereo vision [D]. Wuhan: Huazhong University of Science and Technology, 2016.
- [4] 陈炎, 杨丽丽, 王振鹏. 双目视觉的匹配算法综述[J]. 图学学报, 2020, 41(5): 702-708.
CHEN Y, YANG L L, WANG ZH P. Review of matching algorithms for binocular vision[J]. Journal of Graphics, 2020, 41(05): 702-708.
- [5] 陈华, 王立军, 刘刚. 立体匹配算法研究综述[J]. 高技术通讯, 2020, 30(2): 157-165.
CHEN H, WANG L J, LIU G. A survey of binocular vision matching algorithms [J]. High Technology Communications, 2020, 30(2): 157-165.
- [6] 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.
- [7] 李岩, 吴孟男, 刘克平, 等. 融合多尺度信息的各向异性立体匹配[J]. 计算机集成制造系统, 2023, 29(9): 2920-2928.
LI Y, WU M N, LIU K P, et al. Anisotropic stereo matching of multi-scale information [J]. Computer Integrated Manufacturing Systems, 2023, 29(9): 2920-2928.
- [8] 王晓峰, 孙志恒, 喻骏, 等. 基于细节信息增强的无监督双目立体匹配算法[J]. 电子测量技术, 2024, 47(5): 94-101.
WANG X F, SUN ZH H, YU J, et al. Unsupervised binocular stereo matching algorithm based on detail information augmentation [J]. Electronic Measurement Technology, 2024, 47(5): 94-101.
- [9] 覃业宝, 孙炜, 范诗萌, 等. 全距离深度平衡立体匹配网络[J]. 电子测量与仪器学报, 2023, 37(8): 30.
QIN Y B, SUN W, FAN SH M, et al. Full-distance depth balanced three-dimensional matching network [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(8): 30.
- [10] 白明, 庄严, 王伟. 双目立体匹配算法的研究与进展[J]. 控制与决策, 2008(7): 721-729.
BAI M, ZHUANG Y, WANG W. Research and progress of binocular stereo matching algorithm [J]. Control and Decision, 2008(7): 721-729.
- [11] HAMZAH A R, IBRAHIM H, HASSAN A H A. Stereo matching algorithm based on per pixel difference adjustment, iterative guided filter and graph segmentation [J]. Journal of Visual Communication and Image Representation, 2017, 42:145-160.
- [12] FU Y, LAI K, CHEN W, et al. A pixel pair-based encoding pattern for stereo matching via an adaptively weighted cost [J]. IET Image Processing, 2021, 15(4): 908-917.
- [13] 李镛, 段子阳, 张一帆, 等. 基于跨尺度随机游走的立体匹配算法[J]. 华南理工大学学报(自然科学版), 2020, 48(1): 84-92.
LI Q, DUAN Z Y, ZHANG Y F, et al. Stereo matching algorithm based on cross-scale random walk [J]. Journal of South China University of Technology (Natural Science Edition), 2020, 48(1): 84-92.
- [14] HIRSCHMULLER H. Stereo processing by

- semiglobal matching and mutual information [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 30(2): 328-341.
- [15] 向玺蒙, 王竞雪. 多尺度引导滤波加权聚合的立体匹配算法[J]. 测绘科学, 2023, 48(2): 131-139.
XIANG X M, WANG J X. Stereo matching algorithm based on multi-scale guided filtering weighted aggregation[J]. Science of Surveying and Mapping, 2023, 48(2): 131-139.
- [16] 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.
- [17] ZHANG K, LU J, 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.
- [18] HOSNI A, RHEMANN C, BLEYER M, et al. Fast cost-volume filtering for visual correspondence and beyond[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 35(2): 504-511.
- [19] ZHANG K, FANG Y Q, MIN D B, et al. Cross-scale cost aggregation for stereo matching [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27(5): 965-976.
- [20] XU Z, LIU Y, DU S, et al. DFOB: Detecting and describing features by octagon filter bank for fast image matching [J]. Signal Processing: Image Communication, 2016, 41: 61-71.
- [21] 畅雅雯, 赵冬青, 单彦虎. 多特征融合和自适应聚合的立体匹配算法研究[J]. 计算机工程与应用, 2021, 57(23): 219-225.
CHANG Y W, ZHAO D Q, SHAN Y H. Research on stereo matching algorithm based on multi-feature fusion and adaptive polymerization [J]. Computer Engineering and Applications, 2021, 57(23): 219-225.
- [22] 王笛, 胡辽林. 基于双目视觉的改进特征立体匹配方法[J]. 电子学报, 2022, 50(1): 157-166.
WANG D, HU L L. Improved feature stereo matching method based on binocular vision[J]. Acta Electronica Sinica, 2022, 50(1): 157-166.
- [23] 陈艺, 于纪言, 于洪森. 自适应纹理区域的多尺度融合立体匹配算法[J]. 计算机工程与应用, 2023, 59(18): 198-206.
CHEN Y, YU J Y, YU H S. Multi-scale fusion stereo matching algorithm for adaptive texture region [J]. Computer Engineering and Applications, 2023, 59(18): 198-206.
- [24] 吴福培, 黄耿楠, 刘宇豪, 等. 基于多特征 SAD-Census 变换的立体匹配算法[J]. 中国光学(中英文), 2024, 17(2): 278-290.
WU F P, HUANG G N, LIU Y H, et al. Stereo matching algorithm based on multi-feature SAD-Census transform [J]. Chinese Optics (Chinese and English), 2024, 17(2): 278-290.

作者简介

卜鹏辉, 博士, 讲师, 主要研究方向为光学三维测量方法、计算机视觉三维测量/重建、图像处理和基于深度学习的三维场景理解。

E-mail: 230103@xsyu.edu.cn

田隆涛, 硕士研究生, 主要研究方向为计算机视觉三维测量、图像处理。

E-mail: 1622062714@qq.com

王航(通信作者), 博士, 教授, 博士生导师, 硕士生导师, 主要研究方向为油套管柱失效控制、石油装备智能化、机器视觉和三维高精度测量。

E-mail: wanghang220710@163.com

严亚滔, 硕士研究生, 主要研究方向为基于深度学习的三维场景理解。

E-mail: 1289620630@qq.com