



引文格式:沙洪俊,袁修孝.双目影像密集匹配方法的回顾与展望[J].武汉大学学报(信息科学版),2023,48(11):1813-1833.
DOI:10.13203/j.whugis20230037
Citation:SHA Hongjun, YUAN Xiuxiao.State-of-the-Art Binocular Image Dense Matching Method[J].Geomatics and Information Science of Wuhan University, 2023, 48(11):1813-1833.DOI:10.13203/j.whugis20230037

双目影像密集匹配方法的回顾与展望

沙洪俊¹ 袁修孝¹

¹ 武汉大学遥感信息工程学院,湖北 武汉,430079

摘要:以摄影测量为应用背景,综述了双目影像密集匹配方法。首先借助图表对局部密集匹配与全局密集匹配两类方法进行了简明的比较,指出各类方法的优缺点以及面临的主要挑战;然后分析了密集匹配中难以处理的建筑物遮挡问题,将航摄影像中普遍存在的建筑物遮挡现象分为5种类型,阐述了现有遮挡检测与填充算法的针对性,为解决密集匹配中的空洞问题提供技术思路;最后展望了双目影像密集匹配的发展趋势。以期帮助读者全面了解传统双目影像密集匹配技术,对基于深度学习的密集匹配研究有所裨益。

关键词:双目影像;密集匹配;全局密集匹配;局部密集匹配;遮挡处理

中图分类号:P237

文献标识码:A

收稿日期:2023-03-20

DOI:10.13203/j.whugis20230037

文章编号:1671-8860(2023)11-1813-21

State-of-the-Art Binocular Image Dense Matching Method

SHA Hongjun¹ YUAN Xiuxiao¹

¹ School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

Abstract: Based on the application in the field of photogrammetry, this paper retrospects the binocular image dense matching method. First, the two categories of local dense image matching and global dense image matching methods are concisely compared, and the advantages, disadvantages, and main challenges of each method are pointed out. Then, the building occlusion problem that is difficult to deal with in dense matching is analyzed. The common building occlusion phenomenon in aerial photography is divided into five types, and the pertinence of the existing occlusion detection and filling algorithms is expounded, which provides technical ideas for solving the bottleneck problem in dense image matching. Finally, the development trend of dense image matching of binocular images is prospected. This paper can help readers fully understand the traditional binocular image dense matching technology, and it will be beneficial to the research of dense image matching based on deep learning.

Key words: binocular images; dense image matching; global dense image matching; local dense image matching; occlusion processing

自进入数字地球时代以来,智慧城市的建设工作一直是推动其发展的关键,而三维点云数据的快速获取则是智慧城市建设的 key 一环。目前,三维点云数据主要通过激光雷达(light detection and ranging, LiDAR)技术和影像密集匹配技术来获取。前者虽然可以在短时间内快速获取到数以百万计的地物点三维坐标并得到相应的三维点云模型^[1-2],但相对于获取成本低、效率高、

纹理信息丰富、空间分辨率高、测绘范围大、人工交互少^[3]的影像密集匹配技术,依然存在一些不足^[4-6],因而后者受到了摄影测量与遥感界的广泛关注。

自20世纪50年代起,基于航摄影像的匹配算法被提出后,越来越多的学者开始致力于影像的稀疏匹配方法研究。直到20世纪90年代,随着具有生成数字表面模型和数字地形模型功

基金项目:国家自然科学基金(41771479);国家高分专项(50-H31D01-0508-13/15)。

第一作者:沙洪俊,博士生,研究方向为遥感影像密集匹配理论与方法。hongjun_sha@whu.edu.cn

通讯作者:袁修孝,博士,教授。yuanxx@whu.edu.cn

能的摄影测量商业软件的出现^[1],人们纷纷着手于密集匹配算法的研究。密集匹配技术在计算机视觉领域也被称为立体匹配技术,在其数十年的发展过程中,已有许多文献对其算法进行了归类对比分析^[1,7-11],美国明德学院的Scharstein和Szeliski还创建了广受关注的计算机视觉评测网站^[12],不仅提供了一系列权威的测试数据,还提供了量化误差的算法测试平台,只需要将匹配结果提交到平台就可以获得权威的性能分析并参与算法排名。该网站自2001年开放以来,已公布了430多种影像密集匹配算法,图1为Middlebury网站公布的各年度密集匹配算法数量(统计时间截至2022-11-20)。通过分析不难发现:

1) 双目影像的密集匹配算法数量呈逐年上升趋势。这表明双目影像的密集匹配问题正在被越来越多的学者所关注,同时也说明双目影像的密集匹配技术依然是一个热点研究问题。

2) 多视影像密集匹配算法的年度数量明显偏少,自2014年开始逐年下降。多视密集匹配中的多角度影像可以有效改善遮挡区域的匹配效果,特别是近些年发展起来的倾斜摄影测量技术^[13],融合了传统的航空摄影测量技术和近景摄影测量技术^[14]。现有的多视密集匹配模型主要划分为两类:基于多视的匹配模型及基于多视立体融合的匹配模型^[15]。前者同时考虑多条交会光线以获得更为精确稳健的匹配结果,匹配模型较为复杂,且大部分方法匹配效率较低^[16],并不适合于大范围的影像匹配,更多的此类匹配模型方法介绍可参见文献^[17-18];后者则是将多视密集匹配视为以单个立体像对(双目影像对)为匹配单元的多个立体像对的匹配结果融合问题,该类模型可以通过粗差剔除的方法更有效地改善遮挡区的匹配精度,效率更高、操作简便、易于并行加速,适用于大范围的影像匹配,其核心仍然是双目密集匹配技术。因此,相比于多视影像密集匹配,双目影像的密集匹配更为重要,是更为常用的密集匹配方法,这也是本文着重对双目密集匹配进行综合论述的初衷之一。

近年来,具有相当热度的基于深度学习的密集匹配方法大体上可以分为分布式方法和端到端的方法^[11],后者为当前主流算法。然而,分布式方法对于弱纹理、视差不连续区域以及遮挡区域的匹配效果并不理想,端到端方法则需要较高

的计算机内存消耗。基于深度学习的方法虽然在常规区域可以获得更优的匹配效果,但对于弱纹理、视差不连续以及遮挡区域的匹配效果仍然存在劣势,多数的方法受限于训练标签,文献^[19]详细分析了基于深度学习的影像密集匹配算法,此处不再赘述。

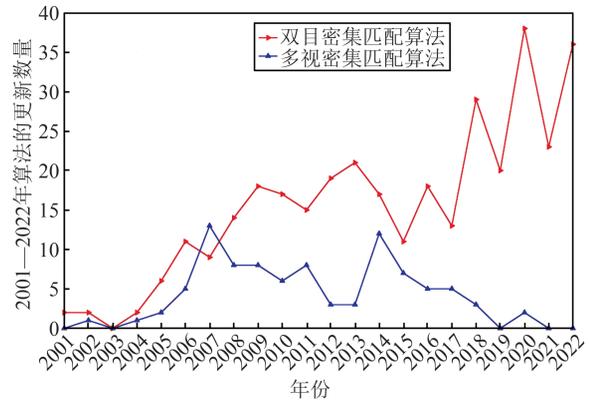


图1 Middlebury网站公布的2001—2022年密集匹配算法数量

Fig. 1 Number of Dense Matching Algorithms Submitted to Middlebury Website from 2001 to 2022

本文通过对现有的双目影像密集匹配算法进行综合分析,着重对传统的密集匹配算法进行总结,以期能系统地阐述影像密集匹配技术的发展现状。本文首先对密集匹配算法的匹配过程进行了梳理,按照局部密集匹配与全局密集匹配对双目密集匹配算法进行了系统的分类概述,并指出各类算法的优缺点;然后对摄影测量中的密集匹配技术进行讨论,对比分析了摄影测量场景与计算机视觉场景之间的差异,同时对影像遮挡机理进行了剖析,详细阐述了遮挡检测与遮挡填充的处理方法,为解决密集匹配的空洞问题奠定了理论基础;最后着重阐述了笔者对双目密集匹配方法发展趋势的展望。

1 双目密集匹配

双目密集匹配即从一个立体像对的二维影像数据信息中提取并恢复对应物体的三维空间信息。现有的双目密集匹配算法依据同名点搜索空间的不同可分为基于像方空间的密集匹配算法和基于物方空间的密集匹配算法^[20]。

基于像方空间的双目密集匹配算法通常采用水平核线影像作为输入影像,又称为单基线密集匹配算法,是一种常用的匹配算法,其核线影像包括框幅式影像的核线影像和线阵推扫式影

像的核线影像。前者是从航空摄影测量和地面摄影测量中经常使用的中心投影影像所获得的核线影像,其核线模型基本分为两种,基于数字影像的几何纠正模型和基于共面条件的模型。其中,基于数字影像几何纠正的核线纠正模型更为简单常用;线阵推扫式影像的核线影像多为从卫星遥感影像所获得的核线影像,其核线模型也分为两种,基于多项式拟合的核线模型和基于投影轨迹法的核线模型^[21]。基于投影轨迹法的核线模型因其具有更加严密的理论而被广泛使用^[22-23]。单基线密集匹配算法通过核线约束对左右影像进行逐像素的匹配,并利用三角测量原理计算立体像对的视差值,最终将视差转换为物方三维坐标,即获得 2D 影像点的 3D 地面坐标(详见图 2)。具体计算过程可以概括为 4 步:匹配代价计算、匹配代价聚合、视差计算与精化以及物方三维坐标计算^[7]。匹配代价计算是以左核线影像为参考影像、右核线影像为目标影像,在指定

的视差搜索范围 $[d_{\min}, d_{\max}]$ 内,计算待匹配像素 p_i 与每个视差下的候选像素之间的视差值 $d^{[24]}$,并将其存储为 $W \times H$ 阶的视差矩阵,其中 W 为影像宽度, H 为影像高度,从而获得视差图;匹配代价聚合是指对初始视差值在基于单个像素 p_i 的支持窗口内进行滤波,剔除不合理的视差值;通过聚合而获得的每个像素会对应多个不同的视差值 d ,视差计算与精化是为每个像素选取最佳的视差值 $d_i^{[25]}$,一般采用胜者为王策略来优选视差值,即在单个像素 p_i 的所有初始视差值中,选择最小的代价值作为最佳视差,剔除错误视差,同时检测并修复无效视差^[26];物方三维坐标计算是在已知影像内外方位元素的前提下,利用视差值求解目标点坐标,即首先根据已知的相机内方位元素与视差值 d_i 求出像素点 p_i 的像方坐标 $p(x, y, z)$,然后结合已知的影像外方位元素将像方坐标 p 转换为大地测量坐标,从而求出目标点 P 的物方三维坐标 $P(X, Y, Z)$ 。

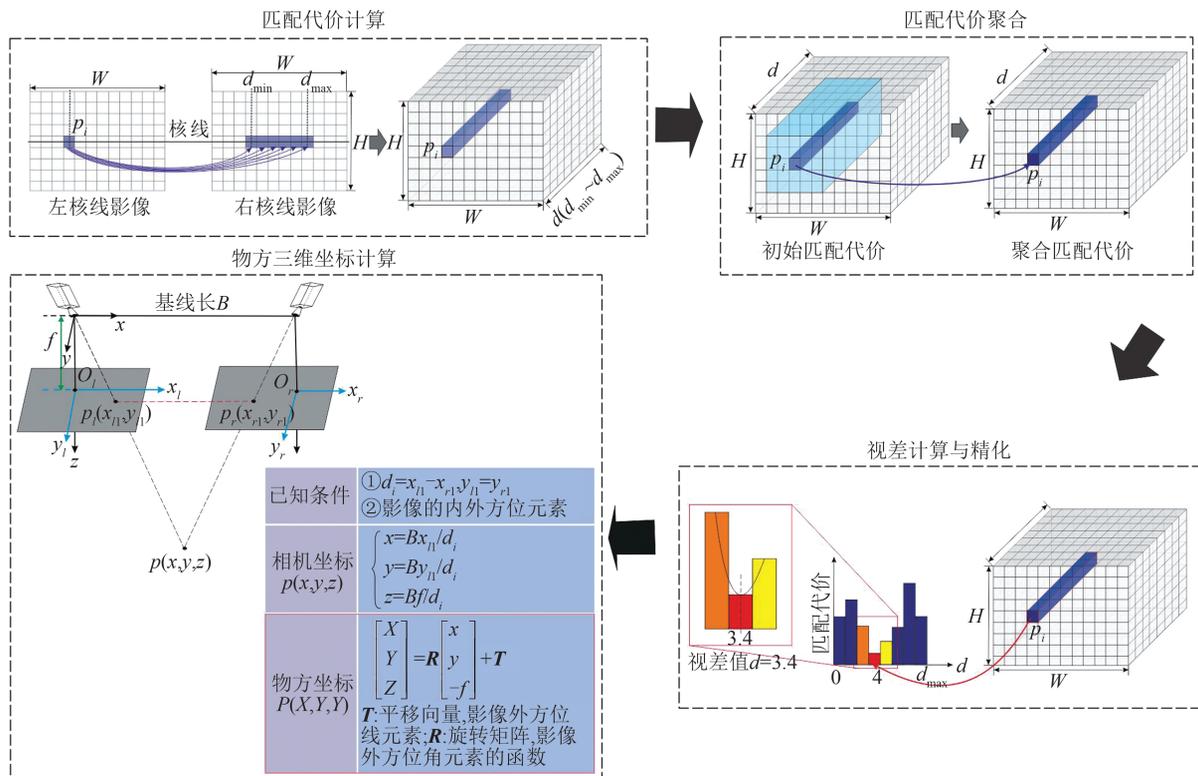


图 2 基于像方空间的双目密集匹配

Fig. 2 Binocular Dense Image-Matching Based on Image Space

基于物方空间的密集匹配是在基于像方空间密集匹配算法之后发展起来的算法,即将密集匹配算法从像方空间转至物方空间。通常直接以地面高程方向或影像视线方向为同名点搜索方向,并将指定的物方面元反投影至影像中,以

影像中投影面元之间的相似性来确定物方点。现有的基于物方空间密集匹配算法多是在铅垂线轨迹法(vertical line locus, VLL)的基础上发展起来的。本文以 VLL 算法为例对基于物方空间的双目密集匹配过程进行介绍。图 3 展示了

VLL原理,假定在物方有一条铅垂线轨迹,其对应影像上的投影也是一条直线(图3中的红色虚线),沿此铅垂线方向搜索地面目标点 P 所对应的左右影像中的像点 p_{li} 与 p_{ri} ($i=0, 1, 2, \dots$)即可确定目标点 P 的高程。具体步骤如下:首先给定目标点 P 的平面坐标以及近似最低高程,并根据精度需求确定高程搜索步距 ΔZ ;然后根据目标点 P 的平面坐标及候选高程值计算目标点在左右影像上的像点坐标 $(x_{p_{li}}, y_{p_{li}})$ 、 $(x_{p_{ri}}, y_{p_{ri}})$,分别以这两点为中心在各自影像上选取匹配窗口,计算匹配测度;令 $i=i+1$,重复前述步骤;最后根据匹配测度最大值确定目标点的高程值 Z ,目标

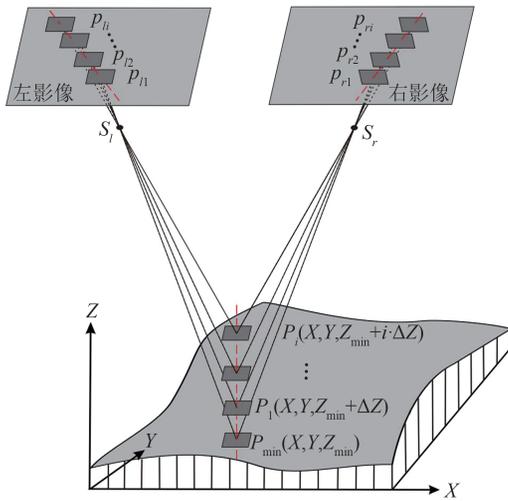


图3 基于物方空间的双目密集匹配

Fig. 3 Binocular Dense Image-Matching Based on Object Space

双目密集匹配当前采用的多是基于像方空间的双目密集匹配算法,依据其是否隐式地使用了光滑假设而被区分为局部密集匹配和全局密集匹配两种^[7]。具体说来,在匹配算法中使用了代价聚合和视差确定组合的为局部密集匹配方法,使用了能量函数优化的为全局密集匹配方法。

1.1 局部密集匹配

利用各个像素的邻域信息求解对应视差的局部密集匹配算法又被称为基于窗口的密集匹配算法^[33],此类算法包含以上双目密集匹配算法完整的步骤且在代价聚合阶段隐式地使用了光滑假设约束。对于局部密集匹配算法而言,在匹配代价聚合时,选择适当的匹配窗口是决定匹配精度的关键。

早期使用的单一固定窗口法因限制了窗口的尺寸大小,使其易在对象边界处发生重叠,从而错误地聚合了不同区域的匹配代价,导致地物

点 P 的物方三维坐标为 $P(X, Y, Z)$,详细过程参见文献^[27]。

基于物方的密集匹配算法实质依然是基于影像的几何定向参数对同名像点进行反复前方交会,从而获取物方点的过程^[16]。由于基于物方空间的密集匹配算法可以同时处理多幅影像,这种方法多被用于多视影像的密集匹配中^[28],如常见的GC³(geometrically constrained cross-correlation)^[29]及其改进模型AMMGC(adaptive multi-image matching with geometric constraints)^[30]、MVLL(modified vertical line locus)^[31]、SGVLL(sem-global vertical line locus)^[32]等。

已知条件	①目标点 P 的平面坐标 ②近似最低高程 Z_{\min} ③影像的内外方位元素
候选高程 Z_i	$Z_i = Z_{\min} + i \Delta Z$ ($i=0, 1, 2, \dots$)
左影像中的像点 P_{li} 坐标	$\begin{cases} x_{p_{li}} = -f \frac{a_{11}(X-X_{s_l}) + b_{11}(Y-Y_{s_l}) + c_{11}(Z-Z_{s_l})}{a_{13}(X-X_{s_l}) + b_{13}(Y-Y_{s_l}) + c_{13}(Z-Z_{s_l})} \\ y_{p_{li}} = -f \frac{a_{21}(X-X_{s_l}) + b_{21}(Y-Y_{s_l}) + c_{21}(Z-Z_{s_l})}{a_{23}(X-X_{s_l}) + b_{23}(Y-Y_{s_l}) + c_{23}(Z-Z_{s_l})} \end{cases}$
右影像中的像点 P_{ri} 坐标	$\begin{cases} x_{p_{ri}} = -f \frac{a_{11}(X-X_{s_r}) + b_{11}(Y-Y_{s_r}) + c_{11}(Z-Z_{s_r})}{a_{13}(X-X_{s_r}) + b_{13}(Y-Y_{s_r}) + c_{13}(Z-Z_{s_r})} \\ y_{p_{ri}} = -f \frac{a_{21}(X-X_{s_r}) + b_{21}(Y-Y_{s_r}) + c_{21}(Z-Z_{s_r})}{a_{23}(X-X_{s_r}) + b_{23}(Y-Y_{s_r}) + c_{23}(Z-Z_{s_r})} \end{cases}$

边缘粗化模糊^[34],使其在深度不连续处、弱纹理区域以及重复纹理区域具有较高的误匹配率。为此,基于窗口的密集匹配方法逐渐衍生为自适应窗口法、自适应权重法以及非前向平行算法等。

1.1.1 自适应窗口法

局部密集匹配所选择的窗口大小既要足够大,以包括足够的强度变化,又要足够小,以避免因覆盖对应场景点深度(即视差)变化的整个区域而引起的失真问题^[35],因而演化出了可以根据局部强度和视差变化而自适应地选择窗口的自适应窗口法。主要是依据局部区域的像素颜色、空间距离、纹理熵等属性特征的相关性来改变聚合窗口的大小、形状、位置和数量,从而提高视差图的精度。将现今较为常用的具有代表性的自适应窗口法概括为单窗口算法与多窗口算法,算法的简要描述和存在的优缺点具体见表1。

表 1 自适应窗口法

Tab. 1 Methods of Adaptive Window

方法	简要描述	存在的优点与缺点
单矩形自适应窗口法 ^[35]	从最小的窗口 3×3 像素开始,在水平与垂直方向选择窗口的大小和形状迭代更新每个点,直到收敛,即通过视差分布的统计模型来评估局部强度和视差方差,进而选择不确定性最小的窗口。	优点:窗口根据图像噪声的不同在垂直或水平方向被拉长,始终可以避免视差跳跃的角落。 缺点:对初始视差估计敏感,高度依赖于初始视差估计;计算昂贵,支持窗口的形状被限制为矩形,不适用于接近任意形状、深度不连续点的像素。
基于最小比循环的自适应窗口法 ^[36]	通过最小比循环(小的周长与面积之比)算法实现图形在紧凑窗口上的快速优化,最大和最小的窗口分别设置为 31×31 像素和 3×3 像素。	优点:包含所有的矩形形状,但是大多数形状不是矩形; 缺点:使用的支持窗口的形状并不通用,需要许多用户指定的参数。
基于多尺度优化的各向异性自适应窗口法 ^[37]	源于星形估计邻域拓扑的几何思想。通过选择逐点的理想星形邻域来实现最优的自适应,将原几何问题转化为实际的多尺度优化问题。	优点:产生各向异性的估计量对于变化点非常敏感,可以提供干净且准确的边缘适应性和出色的恢复性能,即可以揭示出图像的精微元素。 缺点:内存消耗大,计算成本高。
基于鲁棒估计的相关函数自适应窗口法 ^[38]	首先通过相关函数选择匹配候选点的连通性来考虑感兴趣像素附近的视差,然后通过寻找具有正强度梯度相似度的最大连通匹配候选点的视差平面来估计整数视差。	优点:可以提高深度不连续区域和连续光滑区域的匹配精度,且对噪声具有鲁棒性;亚像素测量可减少偶然的局部变形。 缺点:对于较大的噪声或不同于正态分布的灰度像素,在相关测度上很可能会产生误匹配。
自适应窗口增长算法 ^[39]	从信息论的角度,通过调整匹配窗口的大小来实现信息传递量的最大化,因为在匹配窗口中,随着零位值的增加,熵也增加到了了一定的值。	优点:可以避免无法匹配的无特征区域,同时保持匹配窗口尽可能小;能够适应自定义硬件架构的高效实现。 缺点:涉及复杂的统计计算,并依赖于众多的假设。
基于 SA-DCT (shape-adaptive discrete cosine transform) 与各向异性的自适应窗口法 ^[40]	将 SA-DCT 与各向异性局部多项式逼近-置信区间相交技术结合使用,以点为单位自适应地将邻域构造为八边形。	优点:重建的边缘是干净的,没有伪影引入拟合变换; 缺点:当没有足够的相似的面片时,会因为假设的图像先验不再有效而使性能下降。因此,只有在一定条件下,这种图像先验模型才能比其他模型更有效,当条件不再满足时,它就可能变得不合适。
基于 Go-Light 优化的自适应窗口法 ^[41]	采用一个任意形状的 5 个像素窗口和一个规则的固定窗口进行同时匹配,当一个离散的视差值满足特定的约束条件时,该视差值则被替换为其周围视差值的平均值。	优点:在准确地捕捉纹理密集区域的视差的同时,利用固定窗口匹配无纹理区域。 缺点:每对像素需要尝试所有形状和方向,计算时间是 5 个像素的匹配时间乘以 148,相当于匹配一个大小为 27 的方形窗口。
基于自适应多边形逼近的自适应窗口法 ^[42]	基于局部自适应多边形逼近技术与强度和空间信息结合,自适应地确定方向尺度,将邻域结构简化成 8 个方向。	优点:不仅可以对局部自适应多边形逼近技术的执行速度进行优化,而且 8 方向估计在深度不连续区域和均匀区域均获得了较高的重建质量。 缺点:当处理具有重复纹理但深度不同的图像区域时,可能会使基于双边滤波器的方法出现问题。
基于正交积分图像的自适应窗口法 ^[43]	根据颜色相似度和连通性约束,对每个锚点自适应地构造一个直立交叉局部支持骨架和一个形状自适应支持区域,在垂直邻域中合并交叉的水平段。	优点:窗口形状或像素级能够适应于底层的场景结构。 缺点:每一个像素都需要存储 4 个臂长来表示一个自适应的局部支持交叉,因而需要大量计算机内存。
基于多水平分段的自适应窗口法 ^[44]	根据像素的绝对强度差和梯度相似度来确定 4 个不同臂长,以此来构建局部支持骨架。	优点:可以有效减少位于不正确深度区域的不良像素数量,在深度不连续性和分段平滑区域都有较好的性能。 缺点:RSW 在近似各种场景结构时缺乏灵活性,并且沿深度不连续处存在明显的模糊性。
基于颜色相似度和空间距离的自适应窗口法 ^[45]	利用局部颜色相似度和空间距离确定 4 个方向的臂长,以此构造局部窗口。	优点:用局部颜色相似度和空间距离构造变量核窗口,有效减少了在遮挡区域和深度不连续区域的离群值。 缺点:虽然可以减弱固定窗口的限制,但支持窗口之外的信息仍然缺失。

续表 1

方法	简要描述	存在的优点与缺点
对称多窗口算法 ^[46]	通过9个不对称窗口来计算每个像素点的相关性并保留其与最小SSD(sum of squared differences)误差值的差异,因为一个产生较小SSD错误的窗口更有可能覆盖一个恒定深度的区域。	优点:除了改善接近遮挡边界的立体匹配性能外,还可以通过提供更充足的支撑区域来减少纹理少的小图像区域中的错误估计。 缺点:窗口大小仍然是一个自由参数;纹理较差的区域也有问题,但这是典型问题。
关联窗口组合的多窗口算法 ^[47]	基于小的中间窗口来选择周围的最佳相关窗口,实现大的关联窗口组合,分为3种多窗口情况,并配置不同的窗口相关值获取策略。	优点:相关函数误差过滤器,可以使不确定的匹配无效并减少一般误差,减少了对对象边界错误。 缺点:多重关联窗口配置的计算成本偏高。
最小化边界过伸的多窗口算法 ^[48]	顾及水平、垂直和对角线方向的边缘特征,根据唯一性或左右一致性检查,将多个窗口向四个方向进行 9×9 像素与 15×15 像素的窗口大小扩展。	优点:引入了特殊的多窗口来最小化边界过伸长问题,较少有边界越界问题。 缺点:要同时使用8个扩展窗口,需要很高的计算成本。
基于单矩形的多窗口算法 ^[49]	使用4个窗口围绕感兴趣的像素,并基于评价参考图像中的相关统计特征,确定最佳窗口。窗口的大小由平均和水平/垂直平均偏差来统计确定。	优点:可以捕获在低纹理区域更多的信息。 缺点:精度、准确率仍有待提高。
选择性多窗口算法 ^[50]	在多个不同大小的窗口中选择最可靠的窗口,利用自适应方案自动确定适当的窗口大小,其中最小的窗口为 3×3 像素,最大的窗口等于最大的视差。	优点:为低纹理区域提供了一种局部/全局匹配策略,大窗口用于全局匹配,小窗口用于局部匹配,通过使用一系列的窗口来确保至少一个窗口产生一个可靠的视差估计。 缺点:在深度不连续点附近会产生不正确的视差估计,具有高计算复杂度。

总而言之,基于自适应窗口法可以总结如下:

1)最早兴起的自适应矩形单窗口法多是基于积分图像与四象限掩模等技术通过改变矩形窗口的大小或长宽比例来减少位于不正确深度区域的不良像素数量,提高准确率。相比于固定窗口法,其具有更高的匹配精度,但却要高度依赖于初始视差的估计,并且矩形窗口并不适合于接近具有任意形状的深度不连续区。因而该类方法在视差边缘处具有相对较低的准确度。

2)自适应形状单窗口算法多是基于各向异性局部多项式逼近技术或是通过局部属性构建直立交叉局部支持骨架等方式来构建自适应形状窗口。相较于自适应矩形单窗口法,这种方法的窗口形状自适应能力有了较大的提高,以至于在深度不连续区域的匹配准确度得到了较大的提升,但窗口形状或是模型的通用性仍然不高,很多算法需要特定的参数甚至是需要特定的条件下才能得以实现。

3)自适应多窗口算法通常是结合多个单窗口来匹配相对复杂的图像区域,相比于自适应单窗口法,这类方法在低纹理区域以及边缘区域的

匹配准确度有了更大的提升,具有更广泛的适用性。但因其较高的计算复杂度,需要很高的计算成本,匹配效率有所降低。

近年来,为提高自适应窗口算法的运算速率、加强匹配的时效性,很多基于现场可编程门阵列(field programmable gate array, FPGA)或图形处理器(graphics processing unit, GPU)硬件加速的局部方法开始流行^[51]。其中,FPGA开发的灵活性以及对高速信号处理的优越性,使其成为了近些年数字系统开发的核心部件;GPU技术利用硬件的并行计算能力,将算法设计为高度并行,同时使用多个计算核心进行算法运算,具有明显的加速效果^[52]。

1.1.2 自适应权重法

自适应权重法是推动局部密集匹配算法发展的关键。自2006年Yoon等^[53]首次提出可以在匹配质量与全局匹配算法相媲美的局部匹配方法后,局部匹配算法再次被推向密集匹配研究的高潮。这种方法大多是依据局部区域的像素颜色、空间距离、甚至是梯度等属性特征的相似度来分配匹配窗口内原始代价的权值,权值计算方式是实现该类方法的关键。基于不同的权重值主函数,表2对其进行了归类分析。

表 2 自适应权重法

Tab. 2 Methods of Adaptive Weight

方法	简要描述(权重值主函数)	存在的优点与缺点
基于颜色相似度和几何邻近度的自适应权重法 ^[53]	$w(p, q) = \exp\left(-\frac{\Delta c_{pq}}{r_c} + \frac{\Delta g_{pq}}{r_p}\right)$ <p>式中, Δc_{pq}、Δg_{pq} 分别为像素 p 与 q 的色差与欧氏颜色距离; r_c、r_p 为自定义参数。</p>	<p>优点: 不依赖于初始视差并可以保留任意形状的深度不连续面, 且对不同尺寸的窗口具有较好的鲁棒性。</p> <p>缺点: 每个像素都需要进行独立的计算, 导致出现大量的重复计算, 且易受噪声影响。</p>
测地线距离变换的自适应权重法 ^[54]	$\begin{cases} w(p, c) = \exp\left(-\frac{D(p, c)}{r}\right) \\ D(p, c) = \min_{P \in Q_{p,c}} d(P) \end{cases}$ <p>式中, $D(p, c)$ 为像素 p 与窗口中心 c 的测地线距离; r 为用户自定义参数; $Q_{p,c}$ 为从 p 到 c 的所有路径集合; $d(P)$ 为路径 P 的代价。</p>	<p>优点: 加强了前景的连通性, 防止了高权值被错误地分配给背景对象。</p> <p>缺点: 窗口必须足够大, 才能更好地处理低纹理区域, 计算缓慢, 对于结构高度复杂的区域, 由于定权方式所导致的聚合区域过小会使得效果不佳。</p>
Mini-census 自适应权重法 ^[55]	$w(p, c) = \exp\left(-\frac{\Delta C_{p,c}}{r_c}\right)$ <p>式中, $\Delta C_{p,c}$ 为像素 p 和 c 之间的曼哈顿颜色距离; r_c 为自定义参数。</p>	<p>优点: 边缘检测器的集成减少了搜索空间。</p> <p>缺点: 视差估计算法复杂, 对硬件缺少友好, 横向和纵向的匹配代价聚合方法可能会降低匹配质量。</p>
生物和心理物理启发的自适应权重法 ^[56]	$\begin{cases} w_{\text{total}} = w_{\text{dist}} * w_{\text{dissim}} * w_{\text{discon}} \\ w_{\text{dist}}(x', y', d) _{x,y,d} = -k_1 \ln(\text{dist}(x', y', d) _{x,y,d}) \\ w_{\text{dissim}}(x', y', d) _{x,y,d} = -k_2 \ln(\text{dissim}(x', y', d) _{x,y,d}) \\ w_{\text{discon}}(x', y', d) _{x,y,d} = -k_3 \ln(\text{discon}(x', y', d) _{x,y,d}) \end{cases}$ <p>式中, w_{total} 为左、右影像的总权重; w_{dist} 为 $\text{dist}(x', y', d) _{x,y,d}$ 的权重; w_{dissim} 为 $\text{dissim}(x', y', d) _{x,y,d}$ 的权重; w_{discon} 为 $\text{discon}(x', y', d) _{x,y,d}$ 的权重; $\text{dist}(x', y', d) _{x,y,d}$ 为中心像素 (x, y) 到像素 (x', y') 的欧氏距离; d 为当前视差值; $\text{dissim}(x', y', d) _{x,y,d}$、$\text{discon}(x', y', d) _{x,y,d}$ 分别为像素 (x', y') 和 (x, y) 的颜色不相似度和不连续性; $-k_i (i = 1, 2, 3)$ 为自定义参数。</p>	<p>优点: 在聚合过程中, 利用格式塔理论可以很好地确定像素间的相似程度, 且不涉及迭代过程; 3 个基本格式塔定律含义如下: (1) 接近: 两个像素越接近, 它们的相关性越强; (2) 强度相似: 两个像素的颜色越相似, 相关性就越强; (3) 连续性: 两个像素的深度越相似, 相关性越强。</p> <p>缺点: 左右图像中容易出现支撑区域形状不一致的情况, 对匹配易造成干扰。</p>
基于颜色相似度、接近度和梯度相似度的自适应权重法 ^[57]	$w(\vec{p}, \vec{q}) = f_s(\Delta c_{\vec{p}\vec{q}}) \times f_p(\Delta \text{dis}_{\vec{p}\vec{q}}) \times f_g(\Delta \text{gra}_{\vec{p}\vec{q}})$ <p>式中, $\Delta c_{\vec{p}\vec{q}}$、$\Delta \text{dis}_{\vec{p}\vec{q}}$ 和 $\Delta \text{gra}_{\vec{p}\vec{q}}$ 分别代表像素 \vec{p} 与像素 \vec{q} 的颜色相似度、欧氏距离相似度和梯度相似度; $f_s(\Delta c_{\vec{p}\vec{q}})$、$f_p(\Delta \text{dis}_{\vec{p}\vec{q}})$ 和 $f_g(\Delta \text{gra}_{\vec{p}\vec{q}})$ 分别代表上述三者对应的分组强度。</p>	<p>优点: 在不降低并行度和精度的前提下节省了计算量。</p> <p>缺点: 物体的边界和精细纹理区域的匹配效果不佳。</p>
代价滤波 ^[58]	$w_{p,q}(I_n) = \frac{1}{ w ^2 \sum_{q \in w_c} (1 + \frac{(I_{p,n-1} - \mu_{c,n-1})(I_{q,n-1} - \mu_{c,n-1})}{\sigma_{c,n-1}^2 + \epsilon})}$ <p>式中, I_n 为第 n 次迭代时的灰度图像; q、c 分别为像素 p 的相邻像素与中心像素; w_c、w 分别为自定义大小的窗口及像素数量; σ 和 μ 分别为影像强度值对应的方差和均值; ϵ 为平滑项的控制参数。</p>	<p>优点: 不但可以提高低纹理区域影像匹配的鲁棒性, 而且能够改进和保留视差边缘信息。</p> <p>缺点: 由于迭代而使得算法的执行时间较长。</p>
基于 ASW 的三边滤波法 ^[59]	$w_f(p, q) = e^{-\frac{\Delta c_{pq}}{r_c}} e^{-\frac{\Delta g_{pq}}{r_s}} e^{-\frac{\Delta E_{pq}}{r_c}} e^{-\frac{\Delta c_{pq}}{r_c}} e^{-\frac{\Delta g_{pq}}{r_s}}$ <p>式中, $e^{-\frac{\Delta c_{pq}}{r_c}}$、$e^{-\frac{\Delta g_{pq}}{r_s}}$ 分别为双边滤波权函数的颜色相似项和空间接近项; $e^{-\frac{\Delta E_{pq}}{r_c}}$ 与 r_e 分别为边界强度项与自定义参数。</p>	<p>优点: 解决相邻像素在不同视差处及相近颜色处产生的模糊问题, 三边滤波权值函数优于双边滤波权值函数, 特别是对于深度不连续区域。</p> <p>缺点: 很可能出现空洞像素, 匹配精度不高。</p>

续表2

方法	简要描述(权重值主函数)	存在的优点与缺点
基于水平方向和垂直方向的可分离逐次加权求和法 ^[60]	$w_x = \begin{cases} \prod_{k=i+1}^x \exp\left(\frac{\sqrt{\sum_{n=1}^3 [I_R(k,y,n) - I_T(k-1,y,n)]^2}}{-2\alpha}\right), & i < x \\ 1, & i = x \\ \prod_{k=x+1}^i \exp\left(\frac{\sqrt{\sum_{n=1}^3 [I_R(k,y,n) - I_T(k-1,y,n)]^2}}{-2\alpha}\right), & i > x \end{cases}$ $w_y = \begin{cases} \prod_{k=j+1}^y \exp\left(\frac{\sqrt{\sum_{n=1}^3 [I_R(x,k,n) - I_T(x,k-1,n)]^2}}{-2\alpha}\right), & j < y \\ 1, & j = y \\ \prod_{k=y+1}^j \exp\left(\frac{\sqrt{\sum_{n=1}^3 [I_R(x,k,n) - I_T(x,k-1,n)]^2}}{-2\alpha}\right), & j > y \end{cases}$	<p>优点:可以正确地对接不连续深度的区域进行精化,在平滑视差图的同时保持物体边界和深度不连续性。</p> <p>缺点:生成的视差图处于离散空间,不适合基于图像的渲染。</p>
加权指导滤波 ^[61]	$w_G(p') = \frac{1}{N} \sum_{p=1}^N \frac{\sigma_{G,1}^2(p') + \epsilon}{\sigma_{G,1}^2(p) + \epsilon}$ <p>式中,$\sigma_{G,1}^2(p')$为影像G在3×3窗口内的方差;ϵ为取值为$L^2 \times 10^{-6}$的常数;L为影像的动态范围。</p>	<p>优点:可以减少结果图像中出现的光晕伪影,适用于单幅图像细节增强、单幅图像去雾以及不同曝光图像的融合。</p> <p>缺点:准确度不高,仍然有可见的光晕伪影。</p>
基于重力模型的自适应权重法 ^[62]	$\begin{cases} a^d(q) = \frac{M^d(q)}{R^2(p,q)} \\ R(p,q) = \exp(\Delta C(p,q) + \Delta D(p,q)) \end{cases}$ <p>式中,$a^d(q)$为权重系数;$M^d(q)$、$R^2(p,q)$分别为当前视差d的支持区域内中心像素p的邻域像素q的质量和距离;$\Delta C(p,q)$、$\Delta D(p,q)$分别为颜色空间和几何空间中像素p和q的归一化欧氏距离。</p>	<p>优点:可以为很小的支持区域内的像素分配合理的权值,在非连续区域、高度纹理区域以及遮挡区域具有更好的效果。</p> <p>缺点:计算时长会随着窗口的增加而迅速增加。计算速度仍有待提高。</p>
简单的线性迭代聚类分割法 ^[63]	$w(p,q) = \begin{cases} \exp(-N_s(p,q)/r), N_s(p,q) \leq 0.5K \\ \exp(-\Delta c(I_c(p), I_c(q))/r_c), N_s(p,q) > 0.5K \end{cases}$ <p>式中,$N_s(p,q)$为分割的不相似性;$I_c(p)$、$I_c(q)$分别为像素p与q的颜色强度;r与r_c为用户自定义参数;Δc为颜色相似性函数。</p>	<p>优点:可以在提高代价聚合效率的同时改善边缘区域的误匹配视差空洞。</p> <p>缺点:只使用每个像素的颜色和坐标作为特征,当超像不够小时不能很好地贴合物体边界。</p>
自适应指数加权移动平均滤波法 ^[64]	$w_{ij} = \exp\left(-\left(\frac{\ \vec{c}_i - \vec{c}_j\ ^2}{2\sigma_c^2} + \frac{\ \vec{g}_i - \vec{g}_j\ ^2}{2\sigma_d^2}\right)\right)$ <p>式中,\vec{c}_i、\vec{c}_j分别为CIE-Lab颜色空间中超像素i和j的像素平均值;σ_c、σ_d分别为控制色衰减速率与距离权值衰减率的参数;\vec{g}_i、\vec{g}_j分别为i和j在欧氏空间中的重心坐标。</p>	<p>优点:该滤波器将平滑滤波器和自适应权值滤波器相结合,解决了有限聚集窗口内信息不足的问题。</p> <p>缺点:内存需求大,耗时长。</p>

根据自适应权重法的权重确定方式可总结如下:

1)应用最为广泛的基于滤波的自适应权重

法多是根据颜色较为一致的区域视差值较为接近、连续区域具有更加相近的视差值或是梯度等原理来为匹配窗口内的每个像素分配权重。该

类方法中的双边滤波法^[42,53,59]具有相对较好的匹配精度,特别是在深度或是视差断裂区域,但计算效率并不高。为了解决双边滤波的计算效率问题,人们引入了基于图像的引导滤波算法^[34,45,61,65-67],因其具有更高的匹配精度与计算效率而被更广泛使用。总体而言,此类方法可以改善视差边界的匹配效果,最终获得一个较为满意的匹配结果。

2) 基于分割的自适应权重法是基于假定分割块的边缘与深度不连续区域保持一致,且每个分割块具有相同视差或满足某种关系(如仿射变换)的前提下,以待匹配点是否落在分割块内作为定权的参考依据。这类方法在遮挡区域与无纹理区域会具有较好的匹配效果。

3) 自适应权重法可以被划分为基于滤波的自适应权重法与基于分割的自适应权重法。相比于自适应窗口法,这种方法具有更高的匹配精度和更小的计算代价。但这类方法需要对每个像素进行独立的加权计算,计算复杂度相对较高。

1.1.3 非前向平行法

自适应窗口法和自适应权重法是基于匹配窗口内的像素具有相同视差的假设,即假设影像平面与影像的匹配窗口所对应的实际场景表面相平行,称之为前向平行。然而前向平行假设对于深度不连续区域以及表面倾斜区域的影像匹配并不适用。为此,针对具有倾斜结构的场景,衍生出了非前向平行法。

相对于自适应窗口法与自适应权重法,非前向平行法的兴起较晚,至今仅有十余年的发展历史^[68]。这种方法通过倾斜窗口来估计与具有倾斜结构场景相符的支持区域。比较常见的方法当属 PatchMatch 算法,这种算法通过随机采样快速寻找影像面片之间的近似最近邻匹配,再根据影像中的自然一致性将这些匹配信息快速传播至周围区域。但此方法对于病态输入具有较差的收敛性,而且容易产生重影或羽状伪影^[68]。为此,基于 PatchMatch 算法衍生出了一系列常用的方法,包括 PMBP^[69]、PM-Huber^[70]、PM-PM^[71]、PMF^[72]、PMSC^[73] 以及近两年来较受欢迎的 LocalExp 算法^[74]。

基于 PatchMatch 衍生出的非前向平行算法是目前较受欢迎的双目密集匹配方法之一。该系列算法主要是通过改变搜索策略和引入外来项策略实现对原有算法的改进,具体表

现为:

1) 改进搜索策略。将原算法扩展到跨尺度搜索、基于核线的最近邻搜索以及多个最近邻匹配搜索等,可以更有效地顾及全局信息,获得具有亚像素级精度的视差结果。

2) 引入外来项。通过引入外来项,包括图割以及置信度传播等来建立 PatchMatch 与外来项之间的联系,即将原有算法的最佳特征与外来项的最佳特征相结合以提高计算效率和准确性。

使用 3D 标签的 PatchMatch 算法不仅避免了前向平行偏差,而且基于此算法改进的很多方法目前已经达到了亚像素级的匹配精度,但无论是改变搜索策略还是引入外来项的改进算法都高度依赖于影像面片之间的初始匹配结果,均需要更长的计算时间。

总之,非前向平行法解决了非平行表面的匹配歧义问题,填补了局部匹配方法关于非前向平行假设的空缺,但对于低纹理区域的匹配精度以及算法的整体匹配效率仍具有提升空间。

1.2 全局密集匹配

基于马尔可夫随机场(Markov random field, MRF)模型优化的全局密集匹配方法不同于局部密集匹配方法会在代价聚合阶段使用隐式光滑假设,这种方法通常不执行聚合步骤,而是通过能量函数的平滑项来实现显式光滑假设约束,并通过解算全局最优能量函数来求解。全局能量函数包括表示像素自身相似度度量的数据项和表示相邻像素间平滑程度的平滑项,该函数将密集匹配问题转换为求取全局能量函数最小化问题^[75]。因此,全局能量函数的优化方式以及能量函数的最小化求解是全局密集匹配方法的研究重点。

全局密集匹配的能量优化算法中应用较广泛的方法包括动态规划法、置信度传播法和图割法。

1.2.1 动态规划法

动态规划法又称为扫描线优化法^[76],在 20 世纪 90 年代被引入到全局密集匹配算法中,是最早兴起的一类用于解决密集匹配问题的全局优化方法^[77]。早期的动态规划算法仅是通过单独对影像的每一条扫描线进行优化来完成一维动态规划,但仅考虑一维层面的优化并不能够实现影像区域的整体最优。因此,针对不同扫描线之间的断层问题即条纹问题,又提出了诸如文献^[78]等解决方案,但这些方法也同时会产生垂直条纹的影响。

Hirschmuller^[79]为避免在扫描线优化算法中单一路径产生的条纹效应,首次提出了一种基于多路径动态规划策略的半全局匹配(semi-global matching, SGM)算法。这种算法将多个不同的扫描线路径以不同的角度运行来逼近二维优化,

即对影像进行多次一维能量优化。SGM算法因考虑了匹配目标边界的精确性、光照变化的鲁棒性以及计算效率^[80]而具有很好的应用价值。因而后续涌现出了大量的基于SGM的改进算法(见表3)。

表3 基于SGM改进的算法

Tab. 3 Improved Algorithm Based on SGM

方法	改进之处	改进后的效果
CSGM ^[81]	采用强度一致的视差选择非纹理区域和保留不连续插值填充视差图空洞。	能够处理复杂的形状和精细的结构,并产生清晰的物体边界,但增加了算法的复杂度。
FitSGM ^[82]	将SGM与基于census的相关性和基于分割的平面拟合相结合来增强无纹理和遮挡区域的匹配效果。	不仅提高了遮挡和无纹理区域的匹配质量,而且可以显著地减少内存消耗。
由粗到精的快速SGM ^[83]	首次提出了由粗到精的匹配概念来降低计算复杂度。	减少了算法的运行时间,同时保证具有原始算法结果的质量,但只支持互信息代价函数。
tSGM ^[84]	将图像降采样为金字塔,根据多分辨率方法来确定每个像素的搜索空间,其中低分辨率金字塔的匹配结果被用来限制高分辨率金字塔的视差搜索范围。	克服了差异搜索空间的显著增加以及由此产生的更高的处理时间和内存需求,但对于高度不连续区域的匹配结果没有经典方法清晰。
eSGM ^[85]	通过选择性地存储稀疏聚合代价来减少临时内存使用量,并保证没有显著地降低精度。	可以减少内存带宽的要求,并允许大量影像一次性匹配,但计算成本增加了50%。
SORT-SGM ^[86]	将SGM和census变换相结合。	实现了两个亚像素级插值函数,提高了亚像素级的精度。
实时优化SGM ^[87]	对不同的路径方向使用单独的自适应惩罚,每个路径有自己的权重且4个惩罚参数依赖于强度梯度。	算法运行时不受扩展参数化的影响且算法的准确性有所提高。
MGM ^[88]	将多个方向的信息合并到每个扫描线相关的代价中。	克服了SGM一维扫描线优化的局限性,减弱了特征条纹效应,提高了视差结果的质量。
NG-fSGM ^[89]	搜索相邻像素的流相似度并对代价聚合步骤进行优化。	显著地降低了计算的复杂度和内存空间。
SGM-P ^[90]	通过预先计算的表面方向先验获得具有深度变化的先验知识来调整惩罚项。	增强了SGM算法处理非纹理或弱纹理倾斜区域的能力,但算法只适用于离散的视差值。
高效内存参数SGM ^[91]	一种基于高斯混合模型的参数化方法。将数百个代价值表示为一组高斯参数(如权重、均值和方差)。	减少了对存储空间的要求,因为内存需求的增加远远小于视差级别数量的比例。
NG-fSGM ^[92]	通过使用邻近像素的信息,积极地削减流向搜索空间的大小。	在显著地降低了复杂度的同时,也成功地解决了无纹理区域和遮挡区域的局部模糊问题。

基于SGM的改进算法多数是出于降低计算复杂度和内存空间以及提高计算性能的目的来对算法进行改进的,常见的较为有效的改进策略包括由粗到精的分层策略、基于影像边缘的惩罚策略以及与census相关性相结合等策略。这类算法不仅可以避免动态规划法产生的条纹效应,而且在无需迭代的同时对参数的选择也不具有敏感性,是近年来最受欢迎的一种密集匹配算法。但很多算法在提高算法准确度的同时,无形中也增加了算法的复杂度及运行时间。于是算法的硬件可实现性成为提高算法效率的关键,其中比较常用的硬件优化策略为基于GPU的硬件并行优化与基于FPGA的硬件并行优化。

1.2.2 置信度传播法

置信度传播(belief propagation, BP)算法不同于需要单独计算每条扫描线最优解的动态规划法,该算法是通过多次迭代来计算整个影像的最优解。BP算法将影像作为MRF来解释,影像中的每个像素视为观测节点,影像中的像素灰度值作为观测值,视差是观测的隐藏节点标签,相邻节点通过每一次迭代来互相传递消息,最终的视差则是通过具有最大概率的标签来确定^[93]。

自2003年Sun等^[94]将BP算法引进到密集匹配算法中后,越来越多的基于BP的改进算法被提出,包括HBP^[95]、CSB^[96]、BP-M^[97]、SBP^[98]、LBP^[99]以及SDT-BP^[100]等算法。这些算法具有更好的匹配效果,但同时也会带来复杂的计算问

题,导致需要更长的计算时间,实时性较差。这也是为什么近年来所提出的BP算法大都是出于降低算法的复杂度、提高算法速度以及内存优化的目的来对BP算法进行改进的。

1.2.3 图割法

图割法是除了BP算法以外的另一类基于二维层面的较为有效的全局密集匹配方法。这种方法将视差分配问题转化为全局能量函数最小化问题。具体而言,图割法将影像中的像素作为节点,将像素的视差值作为标签构建无向图,采用最大流和最小割算法求解对应能量最小的标签赋值方案,即最优视差。不同的是,BP方法通过在MRF图中迭代地将消息从当前节点传递到邻近节点,从而使能量函数最小化,图割法是将最大流和最小割算法应用于从MRF图中提取的能量流结构中,从而给出最小能量解^[101]。

图割法因构建无向图时会有大量节点且运算过程需要进行多次迭代而使整体算法的空间复杂度和时间复杂度相对较高。因此,基于图割法的全局密集匹配方法多是通过减少搜索空间或是搜索路径以及简化图等方式来对算法进行

加速。虽然图割法在最大流算法的效率方面已经有了很大的提高,但当视差搜索范围较大时,仍无法满足实时性应用的需求。

2 摄影测量密集匹配

综合摄影测量领域的密集匹配算法不难发现,基于像方空间的双目密集匹配算法几乎都是基于计算机视觉场景中的匹配算法,这就使得其无法获得更为突出的匹配效果。因此,对于摄影测量场景而言,密集匹配问题除了要考虑密集匹配算法本身的特点外,还要考虑由于场景差异所带来的影响。

2.1 计算机视觉场景与摄影测量场景的差异

摄影测量场景不同于计算机视觉场景,后者一般包括多数的室内场景以及少数的街景(自动识别)。这里主要讨论计算机视觉的室内场景与航空航天摄影测量的场景差异。根据大量的参考文献以及公开的数据集,可将两种场景的差异大致概括为相机位置、光源位置、光照特点、背景特点以及像幅大小等,详见图4。

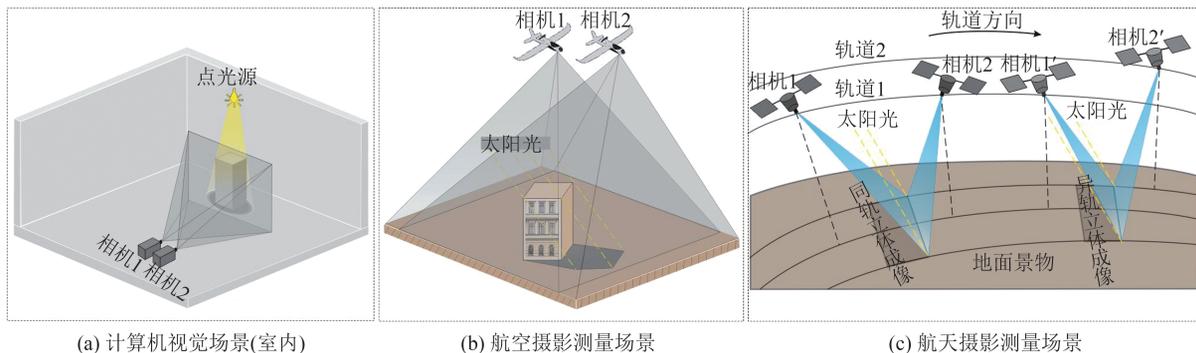


图4 影像数据的摄取场景

Fig. 4 Scene of Image Data Uptake

计算机视觉场景与摄影测量场景的差异包括:

1) 在计算机视觉场景中,相机通常是置于场景的侧前方,近距离地摄取极小图幅影像,因而获取的影像可以明显地区分前景与背景。前景的物体轮廓也格外的清晰完整,地物间所存在的遮挡问题通常也可以通过人工干预尽量规避;而在摄影测量场景中,相机是通过自上而下的方式远距离地摄取大图幅影像,地物结构更为复杂,前景与背景会发生变化,地面上的所有地物即为前景,背景即为地面,地物间存在的遮挡问题是不可避免的。

2) 计算机视觉室内场景中的人造点光源的

位置、光线强度属于人为可控因素,既可以尽量规避阴影的影响,也可以通过增强前景与背景的对比度来将前景与背景清晰区分,几乎不存在阴影的干扰。虽然有少数影像数据中的背景存在阴影,但背景一般为平整的平面,即同一视差背景,因而对匹配结果并无明显的影响;而摄影测量的室外场景的光源位置与强度是人为不可控的太阳光,近乎平行光传播。因而,影像质量更容易受到地物阴影的影响,特别是城区内的高层建筑物。阴影包括自影、本影以及半影,自影与本影则是影响影像质量的主要因素。

3) 计算机视觉室内场景的环境特点较为单一可控且稳定,而摄影测量室外场景具有更少的

人为可控因素,环境更为复杂。摄影测量场景更易受天气变化、辐射畸变等因素的影响,使得密集匹配影像对之间存在更多的差异,严重影响了影像对之间的相关性表达,特别是高分辨率卫星遥感影像。对于高分辨率卫星遥感影像而言,不论是同轨立体影像对间存在的几分钟时间差,还是异轨立体影像对间更长的时间差,都会带来不可忽视的辐射差异,致使影像对之间存在非线性辐射差异,影响影像对之间的相关性表达。

4) 摄影测量场景的立体像对获取要求更为严格,需要地理参考数据的同时又要考虑像对的基线长短、重叠度、交会角大小等问题(包括卫星遥感影像,影像的获取和分析通常被看作是航空摄影测量的直接扩展)^[16,102]。在摄影测量影像密集匹配中,具有一定地面分辨率的近似平行的影像对需要具有一定的航向与旁向重叠度,同时需要精确地获取相机的空间位置与姿态信息以及地面控制点的三维坐标等定位参数信息。而对于立体视觉场景而言,三维信息的恢复通常以左右相机的相对位置为准,并不需要地理参考,特别是计算机视觉领域的热门研究 SfM (structure from motion) 技术以及 SLAM (simultaneous localization and mapping) 技术,它们并不需要预先知道相机的三维坐标和姿态或是地面控

制点的三维坐标,即无需环境的先验信息,只需要已知相机的相对位置和场景几何以及高度重叠的像片集,便可实现影像的获取与计算工作,三维点云数据也可以用任意坐标系来表示^[103-104]。

5) 基于深度学习方法的差异在于应用场景的差异。作为计算机视觉领域中的研究热点,基于深度学习的密集匹配算法可以获得更好的匹配效果。然而,对于大尺度的摄影测量场景影像而言,一方面,大量真实样本的模型参数训练必然需要用极高的成本来获取真值数据,更为复杂且不同的室外环境令预先训练好的模型的迁移学习能力受到影响^[105-106];另一方面,基于深度学习的方法很难直接应用到完整的高分辨率影像数据中,即使是通过分块处理来适应 GPU 的计算能力,也会因为单块像素信息太少而无法保证匹配结果的精度。

2.2 遮挡处理

对于摄影测量密集匹配中的大幅面城市影像而言,影响双目密集匹配质量的一个关键因素当属建筑物的遮挡问题。而不同的建筑物高度或是摄影中心(视点)相对建筑物的位置以及建筑物之间的位置关系都会带来不同的遮挡效果,建筑物间存在的遮挡问题可被简要概括为5种情况,详见图5。

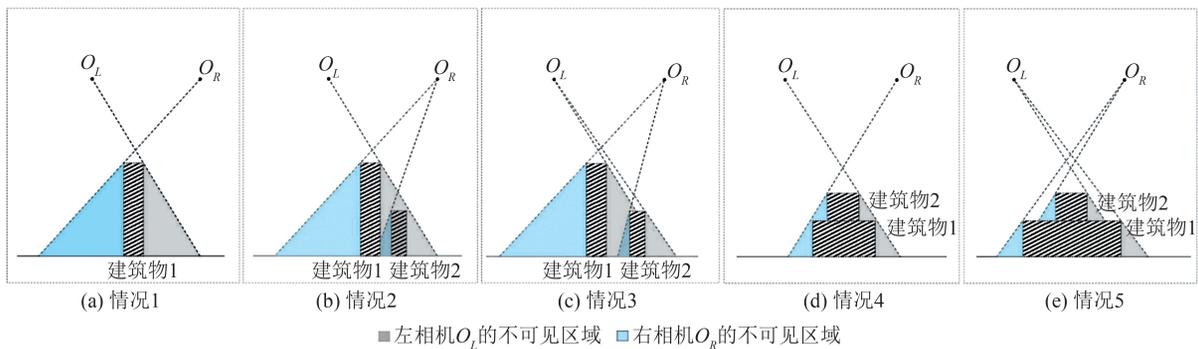


图5 不同建筑物场景的遮挡可见性分析

Fig. 5 Analysis of Occlusion Visibility of Different Building Scenes

图5中,有5种不同的建筑物遮挡情况。情况1为独栋建筑物1被左右相机 O_L 、 O_R 拍摄所呈现的可见性情况;情况2与情况3则表示对于建筑物1,具有不同相对位置关系的相邻建筑物2被左右相机 O_L 、 O_R 拍摄所呈现的两种可见性情况;情况4与情况5则表示建筑物1的屋顶附属建筑物2被左右相机 O_L 、 O_R 拍摄所呈现的两种可见性情况。从以上几种遮挡情况不难发现:对于高层建筑物而言,由于自上而下的拍摄角度以及视点

与建筑物之间的相对位置关系,建筑物的底部结构以及边缘特征会因为遮挡而不可见,进而无法完成正确的匹配,最终会使得匹配结果中被建筑物遮挡的部分因为信息缺失而出现空洞。因此,遮挡处理也成为了提高密集匹配质量的关键。

遮挡处理包括遮挡检测与遮挡填充。对于遮挡检测,多数的方法都是选择利用一定的约束条件去估计被遮挡像素的视差,进而实现遮挡区域的检测。表4列出了几种常见的约束条件。

表 4 用于检测遮挡像素的约束条件
Tab. 4 Constraints for Detecting Occluded Pixels

约束条件	主要描述	主要存在的优点与缺点
顺序约束	通过保持两幅输入图像沿扫描线的匹配顺序提供线索,它要求匹配邻近的像素。	在单薄的物体或狭窄的孔周围,顺序约束无效。
唯一性约束	两个图像中的像素之间的对应点是唯一的,每个像素必须最多有一个视差向量。	不适用于包含水平倾斜表面的场景,会导致像素数量不等的对应。因为水平倾斜的表面在一幅图像中比另一幅图像中表现出更多的水平拉伸。
可见性约束	如果右侧图像中至少有一个像素与左侧图像中的像素匹配,那么左侧图像中的像素将在两幅图像中都可见。可见性约束允许多对一匹配。	可见性约束只加强了一幅图像的遮挡和另一幅图像的视差之间的一致性。当图像背景具有不同深度时,这种假设是无效的。
左右一致性约束	又被称为交叉检查约束。分别基于左图像与右图像之间的对应关系来计算左右视差图,然后将左右视差图中不一致的视差值声明为遮挡。	在噪声和无纹理区域会产生较大的误差,且需要计算双向流,这对于一些高级匹配算法来说是非常耗时的。
连续性约束	即视差连续性约束。假设除遮挡区域和视差不连续区域外,其他区域都是光滑的。	可以有效地减弱噪声,增强视差连续性。
弯曲约束	使用左视差图将左图像中的所有像素投影到右图像上,以生成左遮挡图。在多对一映射的情况下,如果视差图可靠,左边图像中的两个像素对应右边图像中的同一个像素,则视差值较小的像素被认为是被遮挡的像素。	依赖于视差图的可靠性,计算复杂度较高,耗时较长。

除了这些常见的约束条件外,地面控制点约束^[107]、对称约束^[108]、分割约束^[109]、光度约束^[110]、亮度差约束^[111]等约束条件也被结合使用于遮挡检测。

遮挡处理的另一个重要工作是对遮挡检测到的异常值进行遮挡填充处理。遮挡填充是指异常值在其邻域中搜索具有最相似像素的视差值,并对其进行迭代填充的过程。根据搜索策略的不同,将遮挡填充方法分为扫描线填充与最近邻填充。

扫描线填充是指针对每个异常值像素所在的扫描线,向左侧和右侧搜索最近的有效像素的视差值来填充所有被遮挡的像素。同时,由于被遮挡区域在通常情况下具有与背景对象相似的性质,因而会将异常值分配给左右两个有效视差值之间的最小值。然而,扫描线填充法因其是考虑在扫描线范围内进行遮挡填充,所以这类方法会在视差图结果中出现水平条纹,因而也有许多学者选择使用中值滤波来对该种方法进行改进^[34,58,112-113]。扫描线填充法虽然是一种简单有效的遮挡填充方法,但对于具有倾斜表面的场景并不具有可观的填充效果。

不同于扫描线填充法的搜索策略,最近邻插值法是选择在不同的方向上搜索最近的有效像素,然后通过将有效视差传播到相邻的无效视差区域来执行插值^[9,66,110,114-118]。最近邻插值法相

比于扫描线填充法是更为常用的一类遮挡填充方法,但该类方法常常会因为像素的处理顺序是基于几何顺序而具有无效的填充结果。

简言之,借助条件约束检测遮挡像素,并对检测到的遮挡像素进行视差值填充,来完成遮挡处理工作,虽然可以在一定程度上提高密集匹配结果的质量,但也因为此过程需要大量的迭代计算,使得整个密集匹配算法的计算复杂度有所增加。而且就目前的遮挡填充效果而言,遮挡填充仍是一个具有挑战性的任务。

2.3 匹配速度与精度

影像匹配的速度、精度和可靠性是评价密集匹配算法质量的 3 个重要指标^[18],尤其是当匹配算法已经确保具有相当的可靠性时,如何提高算法的匹配速度与匹配精度便成为更进一步提高或衡量算法质量的关键。

匹配速度的提升是满足高效率匹配需求的关键,通常可以通过改变匹配算法本身的搜索策略和基于计算机硬件层面的并行优化策略来实现。常用的基于搜索策略的优化算法包括上文所述的动态规划算法、松弛迭代算法、人工神经网络以及遗传算法^[18],其基本思想是通过改变搜索策略来尽可能多地避免遍历搜索,提高匹配速度。基于计算机硬件层面的并行优化策略包括基于中央处理器(central processing unit,CPU)的并行优化、基于 GPU 的并行优化以及基于 FPGA 的并行优

化,这种优化策略是利用硬件的并行计算能力将算法设计为高度并行,同时使用多个计算核心进行算法运算,提高匹配速度^[51],而对于摄影测量场景中的高分辨率大幅面影像而言,显然具有更为充裕的存储空间的CPU更具优势。

匹配精度的提升是满足高精度三维点云数据获取的关键,除了目视判别外,现有的较为常见的精度评价指标包括物方精度评定与像方精度评定两种,常见的精度评价指标及主要描述见表5。

表5 精度评价指标
Tab. 5 Evaluation Metrics

评价指标	精度	主要描述
点云检测质量	物方精度	$Q(p) = \frac{t(p)}{t(p)+m(p)} \cdot \frac{s(p)}{t(p)+s(p)}$ <p>式中, $\frac{t(p)}{t(p)+m(p)}$、$\frac{s(p)}{t(p)+s(p)}$ 分别表示完整率与正确率; $t(p)$、$s(p)$、$m(p)$ 分别代表正检点数量、虚检点数量、漏检点数量。</p>
平面位置中误差	物方精度	$S_{xy} = \pm \sqrt{\frac{\sum_{i=1}^n (x_i - x_i^0)^2 + (y_i - y_i^0)^2}{n}}$ <p>式中, x_i、y_i 分别为第 i 个检查点的量测坐标值; x_i^0、y_i^0 为第 i 个检查点的标准坐标值; n 为检测点总数。</p>
高程平均中误差	物方精度	$H_{(i)} = \pm \sqrt{\frac{\sum_{i=1}^n (h_i - h_i^0)^2}{n}}$ <p>式中, h_i、h_i^0 分别为第 i 个检查点的量测高程值与标准高程值; n 为检测点总数。</p>
高程准确率与错误率	物方精度	$P_H = \frac{\sum_{i=1}^n h_i - h_i^0 < \mu}{n} \text{ 与 } P_{H1} = \frac{\sum_{i=1}^n h_i - h_i^0 > \mu}{n}$ <p>式中, h_i、h_i^0 分别为第 i 个检查点的量测高程值与标准高程值; n 为检测点总数; μ 为阈值。</p>
高程平均绝对误差	物方精度	$M_H = \frac{\sum_{i=1}^n h_i - h_i^0 }{n}$ <p>式中, h_i、h_i^0 分别为第 i 个检查点的量测高程值与标准高程值; n 为检测点总数。</p>
视差平均绝对误差	像方精度	$\text{Avg} = \frac{\sum_{i=1}^n d_i - d_i^0 }{n}$ <p>式中, d_i、d_i^0 分别为第 i 个检查点的视差值与标准视差值; n 为检测点总数。</p>
视差均方误差	像方精度	$\text{RMS} = \sqrt{\frac{\sum_{i=1}^n d_i - d_i^0 ^2}{n}}$ <p>式中, d_i、d_i^0 分别为第 i 个检查点的视差值与标准视差值; n 为检测点总数。</p>
像素正确率与错误率	像方精度	$P_d = \frac{\sum_{i=1}^n d_i - d_i^0 < \mu_1}{n} \text{ 与 } P_{d1} = \frac{\sum_{i=1}^n d_i - d_i^0 > \mu_1}{n}$ <p>式中, d_i、d_i^0 分别为第 i 个检查点的视差值与标准视差值; n 为检测点总数; μ_1 为阈值。</p>
重投影误差 ^[119]	像方精度	$Q(p) = \sqrt{\frac{\sum_{i=1}^n (\text{dis}(p_L, \tilde{p}_L) + \text{dis}(p_R, \tilde{p}_R))^2}{n}}$ <p>式中, p_L、\tilde{p}_L 分别为左像点坐标与右影像在左影像上同名像点坐标估值; p_R、\tilde{p}_R 分别为右像点坐标与左影像在右影像上的同名像点坐标估值; $\text{dis}(\cdot)$ 表示像点坐标与其估值的距离; n 为检测点总数。</p>

匹配精度的提升通常可从两方面入手:一方面可以通过在基本的匹配算法的基础上引入几何约束的点、线、面以及尺度空间等特征约束条

件来提升视差初值的可靠性以实现匹配精度的提升;另一方面可以通过后处理技术来实现视差的优化以实现匹配精度的提升。通过引入特征

约束条件来提高算法的匹配精度是提高深度不连续区域以及弱纹理区域匹配精度的重要手段,特别是具有丰富平面几何结构特征的城市景观^[120-124],但其局限性在于几何约束条件的适用范围有限,多数方法仅适用于范围较小,几何特征相对丰富的区域,且约束条件也多为定制式,并不适用于其他的匹配算法^[16,106];尺度空间特征约束条件是利用不同尺度的输入图像来检测匹配关键点,也即尺度不变性,这种方法被更多地引入自适应窗口法^[37,125-127]、自适应权重法^[128-132]以及SGM算法^[133-135]中,然而这种方法却会降低原算法的运算速度。后处理技术即是对视差值的优化,视差优化通常包括误匹配剔除、整像素精度精化与视差平滑。误匹配剔除是针对弱纹理、重复纹理以及遮挡等区域出现的孤立异常点等误匹配结果进行左右一致性检测、中值滤波、加权中值滤波^[136]、异常值剔除滤波^[116]的剔除;整像素精度^[20]精化则是对视差结果的整像素精度进行子像素或亚像素的细化,常见的方法有基于抛物线函数插值法^[137]、最小二乘匹配^[20]、拟合抛物线^[138]等;视差平滑则是对视差结果通过双边滤波、引导滤波等手段进行匹配结果的平滑操作。

3 结 语

本文结合图表对传统双目密集匹配算法进行了全面分析。局部密集匹配的复杂度并不高,具有匹配效率高、实时性强的优势,但缺点是匹配精度较低,特别是在模糊的弱纹理区域;全局密集匹配算法具有较高的匹配精度,特别是在深度不连续区域、遮挡区域、弱纹理等区域均具有较好的效果,但由于这类算法大都是通过迭代计算来实现的,算法的实时性与匹配效率都不太理想。此外,针对遥感影像所具有的宽基线、大幅面的特点,本文对高层建筑物的遮挡情况进行了详细的分析。在航摄场景中,由于采用自上而下的拍摄方式,高层建筑物因为遮挡问题而隐藏掉了许多的纹理信息,如何处理遮挡问题是提高密集匹配质量的关键,自然会增加整个算法的计算复杂度。

就摄影测量而言,从上文所述的密集匹配整体算法来看,现有的双目密集匹配主要存在匹配效率和匹配精度两方面的问题。匹配效率问题指的是算法无法完全满足实时获取匹配结果的需求,需要进一步缩短算法的计算时间,提高匹配效率。匹配效率的提高,一方面需要降低算法

的复杂度,分析算法是否必须迭代运算,在满足大幅面的遥感影像匹配需求的同时考虑搜索策略的优化,另一方面也可以通过基于计算机硬件层面的并行优化策略来缩短计算时间。匹配精度问题则是指匹配结果所存在的边缘毛刺与匹配空洞等问题,该部分数据的不精确与缺失大大降低了算法的整体匹配精度。精度的提升方法一方面可以通过直接在算法中引入特征约束条件来提升视差初值的可靠性,另一方面可以通过后处理技术对视差结果更进一步处理以实现匹配精度的提升。

就密集匹配技术的迅猛发展而言,本文难以穷尽所有的密集匹配方法,但愿以上内容能够给读者带来更进一步的认识。我们坚信,在不久的将来,影像密集匹配技术在算法的实时性以及匹配精度上一定会取得更大的进步。

参 考 文 献

- [1] Remondino F, Spera M G, Nocerino E, et al. State of the Art in High Density Image Matching[J]. *The Photogrammetric Record*, 2014, 29 (146) : 144-166.
- [2] Nebiker S, Bleisch S, Christen M. Rich Point Clouds in Virtual Globes — A New Paradigm in City Modeling? [J]. *Computers, Environment and Urban Systems*, 2010, 34(6): 508-517.
- [3] Remondino F, Spera M G, Nocerino E, et al. Dense Image Matching: Comparisons and Analyses [C]//2013 Digital Heritage International Congress, Vancouver, Canada, 2014.
- [4] Niemeyer J, Rottensteiner F, Soergel U. Contextual Classification of LiDAR Data and Building Object Detection in Urban Areas [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 87: 152-165.
- [5] Zhang J, Lin X. Filtering Airborne LiDAR Data by Embedding Smoothness-Constrained Segmentation in Progressive TIN Densification [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2013, 81: 44-59.
- [6] Cheng L, Tong L H, Li M C, et al. Semi-automatic Registration of Airborne and Terrestrial Laser Scanning Data Using Building Corner Matching with Boundaries as Reliability Check [J]. *Remote Sensing*, 2013, 5(12): 6260-6283.
- [7] Scharstein D, Szeliski R. A Taxonomy and Evaluation of Dense Two-frame Stereo Correspondence Algorithms [J]. *International Journal of Computer Vi-*

- tion, 2002, 47(1): 7-42.
- [8] Seitz S M, Curless B, Diebel J, et al. A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms [C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, USA, 2006.
- [9] Lazaros N, Sirakoulis G C, Gasteratos A. Review of Stereo Vision Algorithms: From Software to Hardware [J]. *International Journal of Optomechatronics*, 2008, 2(4): 435-462.
- [10] Sunyoto H, van der Mark W, Gavrilu D M. A Comparative Study of Fast Dense Stereo Vision Algorithms [C]//IEEE Intelligent Vehicles Symposium, Parma, Italy, 2004.
- [11] Yuan Xiuxiao, Yuan Wei, Xu Shu, et al. Research Developments and Prospects on Dense Image Matching in Photogrammetry [J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48 (12): 1542-1550. (袁修孝, 袁巍, 许殊, 等. 航摄影像密集匹配的研究进展与展望[J]. 测绘学报, 2019, 48 (12): 1542-1550.)
- [12] Scharstein D, Szeliski R. Middlebury Stereo Vision Research Page [EB/OL]. [2023-05-06] <https://vision.middlebury.edu/>
- [13] Yang Guodong, Wang Minshui. The Tilt Photographic Measurement Technique and Expectation [J]. *Geomatics & Spatial Information Technology*, 2016, 39(1): 13-15. (杨国东, 王民水. 倾斜摄影测量技术应用及展望[J]. 测绘与空间地理信息, 2016, 39(1): 13-15.)
- [14] Li Deren, Wang Shugen. Introduction to Photogrammetry and Remote Sensing [M]. 3rd ed. Beijing: Surveying and Mapping Publishing House, 2021. (李德仁, 王树根. 摄影测量与遥感概论 [M]. 3版. 北京: 测绘出版社, 2021.)
- [15] Huang Xu. LiDAR Data Constrained Multi-View Dense Matching and Point Clouds Fusion [D]. Wuhan: Wuhan University, 2016. (黄旭. LiDAR点云约束下的多视影像密集匹配与融合方法研究[D]. 武汉: 武汉大学, 2016.)
- [16] Han Yilong. Dense Image Matching and Building Edge Optimization with High-Resolution Multiview Satellite Imagery [D]. Wuhan: Wuhan University, 2020. (韩轶龙. 高分辨率卫星多视影像密集匹配与建筑物边缘优化[D]. 武汉: 武汉大学, 2020.)
- [17] Yasutaka F, Carlos H. Multi-View Stereo: A Tutorial [M]. New York: Now Foundations and Trends, 2015.
- [18] Zhang Yongsheng, Liu Jun. High-Resolution Remote Sensing Satellite Applications [M]. Beijing: Science Press, 2020. (张永生, 刘军. 高分辨率遥感卫星应用[M]. 北京: 科学出版社, 2020.)
- [19] Ji Shunping, Luo Chong, Liu Jin. A Review of Dense Stereo Image Matching Methods Based on Deep Learning [J]. *Geomatics and Information Science of Wuhan University*, 2021, 46(2): 193-202. (季顺平, 罗冲, 刘瑾. 基于深度学习的立体影像密集匹配方法综述[J]. 武汉大学学报(信息科学版), 2021, 46(2): 193-202.)
- [20] Zhang Jianqing. Pan Li, Wang Shugen. Photogrammetry [M]. Wuhan: Wuhan University Press, 2009. (张剑清, 潘励, 王树根. 摄影测量学[M]. 武汉: 武汉大学出版社, 2009.)
- [21] Yuan Xiuxiao, Cao Jinshan. Theory and Method of Precise Ground Target Positioning by High-Resolution Satellite Remote Sensing [M]. Beijing: Science Press, 2012. (袁修孝, 曹金山. 高分辨率卫星遥感精确对地目标定位理论与方法[M]. 北京: 科学出版社, 2012.)
- [22] Hu Fen, Wang Mi, Li Deren, et al. Approximate Epipolar Rearrangement Algorithm of Satellite Stereo-imagery by Projection Trajectory [J]. *Geomatics and Information Science of Wuhan University*, 2009, 34(12): 1431-1435. (胡芬, 王密, 李德仁, 等. 利用投影轨迹的卫星影像近似核线重排快速算法[J]. 武汉大学学报(信息科学版), 2009, 34 (12): 1431-1435.)
- [23] Wang Fenfei, Ren Xin, Liu Jianjun, et al. The Methods of CE-2 Image Dense Matching and Lunar DEM Extraction [J]. *Geomatics and Information Science of Wuhan University*, 2017, 42(7): 918-923. (王奋飞, 任鑫, 刘建军, 等. 嫦娥二号影像密集匹配及DEM制作方法[J]. 武汉大学学报(信息科学版), 2017, 42(7): 918-923.)
- [24] Wu W H, Zhu H, Yu S Y, et al. Stereo Matching with Fusing Adaptive Support Weights [J]. *IEEE Access*, 2019, 7: 61960-61974.
- [25] Salehian B, Fotouhi A M, Raie A A. Dynamic Programming-based Dense Stereo Matching Improvement Using an Efficient Search Space Reduction Technique [J]. *Optik*, 2018, 160: 1-12.
- [26] Lin Y M, Lu N G, Lou X P, et al. Matching Cost Filtering for Dense Stereo Correspondence [J]. *Mathematical Problems in Engineering*, 2013, 2013: 1-11.
- [27] Kou Rong. Image Matching Experiments Using the VLL Algorithm [J]. *Value Engineering*, 2010, 29 (22): 144. (寇蓉. 关于VLL影像匹配的实验[J]. 价值工程. 2010, 29(22): 144.)
- [28] Jiang Wanshou. Multiple Aerial Image Matching

- and Automatic Building Detection[D]. Wuhan: Wuhan University, 2004. (江万寿. 航空影像多视匹配与规则建筑物自动提取方法研究[D]. 武汉: 武汉大学, 2004.)
- [29] Li Z. Automatic Digital Surface Model (DSM) Generation from Linear Array Images[J]. *Mitteilungen – Institut fuer Geodasie und Photogrammetrie an der Eidgenoessischen Technischen Hochschule*, 2005, 88:32-48
- [30] Fan Dazhao, Ji Song, Lei Rong. Theory, Method and Application of Multi-Baseline Stereo Matching Technique [M]. Beijing: Surveying and Mapping Publishing House, 2017. (范大昭, 纪松, 雷蓉. 多基线立体匹配理论、方法与应用[M]. 北京: 测绘出版社, 2017.)
- [31] Ji Song. Study on the Strategy and Improvement Method of Multi-View Matching Technology [D]. Zhengzhou: Information Engineering University, 2012. (纪松. 多视匹配策略与优化方法研究[D]. 郑州: 信息工程大学, 2012.)
- [32] Zhang Yanfeng. Rough Landform Guided Semi-Global Optimization for Dense Image Matching and Terrain Surface Reconstruction [D]. Wuhan: Wuhan University, 2017. (张彦峰. 粗略地形引导下影像匹配与地表重建的半全局优化方法[D]. 武汉: 武汉大学, 2017.)
- [33] Zhou K, Meng X X, Cheng B. Review of Stereo Matching Algorithms Based on Deep Learning [J]. *Computational Intelligence and Neuroscience*, 2020, 2020: 1-12.
- [34] Yang Q, Ji P, Li D, et al. Fast Stereo Matching Using Adaptive Guided Filtering[J]. *Image and Vision Computing*, 2014, 32(3): 202-211.
- [35] Kanade T, Okutomi M. A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994, 16(9): 920-932.
- [36] Veksler O. Stereo Correspondence with Compact Windows via Minimum Ratio Cycle [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(12): 1654-1660.
- [37] Foi A, Katkovnik V, Egiazarian K, et al. A Novel Anisotropic Local Polynomial Estimator Based on Directional Multiscale Optimizations[C]//International Conference Math in Signal Processing, Cirencester, UK, 2004.
- [38] Kim G, Chung S. An Accurate and Robust Stereo Matching Algorithm with Variable Windows for 3D Measurements [J]. *Mechatronics*, 2004, 14 (6) : 715-735.
- [39] Cyganek B. Adaptive Window Growing Technique for Efficient Image Matching[C]// Pattern Recognition and Image Analysis, Estoril, Portugal, 2005.
- [40] Foi A, Katkovnik V, Egiazarian K. Pointwise Shape-Adaptive DCT for High-quality Denoising and Deblocking of Grayscale and Color Images[J]. *IEEE Transactions on Image Processing*, 2007, 16 (5): 1395-1411.
- [41] Su X Y, Khoshgoftaar T M. Arbitrarily-Shaped Window Based Stereo Matching Using the Go-light Optimization Algorithm [C]//IEEE International Conference on Image Processing, San Antonio, USA, 2007.
- [42] Lu J B, Lafruit G, Catthoor F. Anisotropic Local High-Confidence Voting for Accurate Stereo Correspondence[C]// Image Processing: Algorithms and Systems VI, San Jose, USA, 2008.
- [43] Lu J B, Ke Z, Lafruit G, et al. Real-time Stereo Matching: A Cross-based Local Approach [C]// IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, China, 2009.
- [44] Xu Y F, Zhao Y, Ji M Q. Local Stereo Matching with Adaptive Shape Support Window Based Cost Aggregation [J]. *Applied Optics*, 2014, 53 (29) : 6885-6892.
- [45] Zhu S P, Yan L N. Local Stereo Matching Algorithm with Efficient Matching Cost and Adaptive Guided Image Filter [J]. *The Visual Computer*, 2017, 33(9): 1087-1102.
- [46] Fusiello A, Roberto V, Trucco E. Sharp and Dense Disparity Maps Using Multiple Windows [C]//The 3rd IEEE Pacific Rim Conference on Multimedia, Hsinchu, Taiwan, China, 1997.
- [47] Hirschmüller H, Innocent P R, Garibaldi J. Real-time Correlation-based Stereo Vision with Reduced Border Errors [J]. *International Journal of Computer Vision*, 2002, 47(1): 229-246.
- [48] Jeon J, Kim C, Ho Y S. Sharp and Dense Disparity Maps Using Multiple Windows[C]// The 3rd IEEE Pacific Rim Conference on Multimedia, Hsinchu, Taiwan, China, 2002.
- [49] Demoulin C, Droogenbroeck M V. Disparity Map Determination by Multiple Adaptive Windows [C]// IEEE 17th Signal Processing and Communications Applications Conference, Antalya, Turkey, 2005.
- [50] Adhyapak S, Kehtarnavaz N, Nadin M. Stereo Matching via Selective Multiple Windows[J]. *Journal of Electronic Imaging*, 2007, 16(1): 013012.

- [51] Li Yingsong. Research on Key Techniques of Fast Stereo Matching for Photogrammetric Images [D]. Wuhan: Wuhan University, 2018. (李迎松. 摄影测量影像快速立体匹配关键技术研究[D]. 武汉: 武汉大学, 2018.)
- [52] Li Y, Zheng S, Wang X, et al. An Efficient Photogrammetric Stereo Matching Method for High-Resolution Images [J]. *Computers & Geosciences*, 2016, 97: 58-66.
- [53] 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.
- [54] Hosni A, Bleyer M, Gelautz M, et al. Local Stereo Matching Using Geodesic Support Weights [C]// The 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, 2009.
- [55] Chang N Y C, Tsai T H, Hsu B H, et al. Algorithm and Architecture of Disparity Estimation with Mini-census Adaptive Support Weight [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2010, 20(6): 792-805.
- [56] Nalpantidis L, Gasteratos A. Biologically and Psychophysically Inspired Adaptive Support Weights Algorithm for Stereo Correspondence [J]. *Robotics and Autonomous Systems*, 2010, 58(5): 457-464.
- [57] Geng Y N, Zhao Y, Chen H X. Stereo Matching Based on Adaptive Support-Weight Approach in RGB Vector Space [J]. *Applied Optics*, 2012, 51(16): 3538-3545.
- [58] 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*, 2013, 35(2): 504-511.
- [59] Chen D M, Ardabilian M, Chen L M. A Novel Trilateral Filter Based Adaptive Support Weight Method for Stereo Matching [C]// British Machine Vision Conference 2013, Bristol, UK, 2013.
- [60] Zhu C T, Chang Y Z, Wang H M, et al. Efficient Stereo Matching with Decoupled Dissimilarity Measure Using Successive Weighted Summation [J]. *Mathematical Problems in Engineering*, 2014, 2014: 1-9.
- [61] Li Z G, Zheng J H, Zhu Z J, et al. Weighted Guided Image Filtering [J]. *IEEE Transactions on Image Processing*, 2015, 24(1): 120-129.
- [62] Zhang Z H, Wang Y Q, Huang T, et al. A Weighting Algorithm Based on the Gravitational Model for Local Stereo Matching [J]. *Signal, Image and Video Processing*, 2020, 14(2): 315-323.
- [63] Huang C S, Huang Y H, Chan D Y, et al. Shape-reserved Stereo Matching with Segment-based Cost Aggregation and Dual-path Refinement [J]. *EURASIP Journal on Image and Video Processing*, 2020, 2020(1): 1-19.
- [64] Yang S, Lei X Y, Liu Z F, et al. An Efficient Local Stereo Matching Method Based on an Adaptive Exponentially Weighted Moving Average Filter in SLIC Space [J]. *IET Image Processing*, 2021, 15(8): 1722-1732.
- [65] Hosni A, Bleyer M, Rhemann C, et al. Real-time Local Stereo Matching Using Guided Image Filtering [C]//IEEE International Conference on Multimedia and Expo, Barcelona, Spain, 2011.
- [66] Hong G, Kim B. A Local Stereo Matching Algorithm Based on Weighted Guided Image Filtering for Improving the Generation of Depth Range Images [J]. *Displays*, 2017, 49: 80-87.
- [67] He K M, Sun J, Tang X O. Guided Image Filtering [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(6): 1397-1409.
- [68] Barnes C, Shechtman E, Finkelstein A, et al. PatchMatch [J]. *ACM Transactions on Graphics*, 2009, 28(3): 1-11.
- [69] Besse F, Rother C, Fitzgibbon A, et al. PMBP: PatchMatch Belief Propagation for Correspondence Field Estimation [J]. *International Journal of Computer Vision*, 2014, 110(1): 2-13.
- [70] Heise P, Klose S, Jensen B, et al. PM-Huber: PatchMatch with Huber Regularization for Stereo Matching [C]//IEEE International Conference on Computer Vision, Sydney, Australia, 2013.
- [71] Xu S, Zhang F, He X, et al. PM-PM: PatchMatch with Potts Model for Object Segmentation and Stereo Matching [J]. *IEEE Transactions on Image Processing*, 2015, 24(7): 2182-2196.
- [72] Lu J B, Li Y, Yang H S, et al. PatchMatch Filter: Edge-Aware Filtering Meets Randomized Search for Visual Correspondence [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(9): 1866-1879.
- [73] Li L C, Zhang S L, Yu X, et al. PMSC: PatchMatch-based Superpixel Cut for Accurate Stereo Matching [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, 28(3): 679-692.
- [74] Taniat T, Matsushita Y, Sato Y, et al. Continuous 3D Label Stereo Matching Using Local Expansion Moves [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(11): 2725-

- 2739.
- [75] Haq Q M U, Lin C H, Ruan S J, et al. An Edge-aware Based Adaptive Multi-feature Set Extraction for Stereo Matching of Binocular Images[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2022, 13(4): 1953-1967.
- [76] Gong M L, Zhang Y, Yang Y H. Near-Real-time Stereo Matching with Slanted Surface Modeling and Sub-pixel Accuracy [J]. *Pattern Recognition*, 2011, 44(10/11): 2701-2710.
- [77] Ney H. Dynamic Programming as a Technique for Pattern Recognition[J]. *Journal of Applied Probability*. 1982, 3(1):268-271.
- [78] Gong M L, Yang Y H. Fast Unambiguous Stereo Matching Using Reliability-based Dynamic Programming [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(6): 998-1003.
- [79] 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, San Diego, USA, 2005.
- [80] Hirschmüller H. Semi-global Matching — Motivation, Developments and Applications [C]// Photogrammetric Week, Stuttgart, Germany, 2011.
- [81] Hirschmuller H. Stereo Vision in Structured Environments by Consistent Semi-global Matching [C]// IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, USA, 2006.
- [82] 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]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 2010.
- [83] Hermann S, Klette R. Evaluation of a New Coarse-to-fine Strategy for Fast Semi-global Stereo Matching [M]//Advances in Image and Video Technology. Berlin, Heidelberg: Springer, 2011: 395-406.
- [84] Rothmel M, Wenzel K, Fritsch D, et al. SURE: Photogrammetric Surface Reconstruction from Imagery [C]//LC3D Workshop, Berlin, Germany, 2012.
- [85] Hirschmüller H, Buder M, Ernst I. Memory Efficient Semi-global Matching [J]. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2012, 3(1): 371-376.
- [86] Pantilie C D, Nedeveschi S. SORT-SGM: Subpixel Optimized Real-time Semiglobal Matching for Intelligent Vehicles [J]. *IEEE Transactions on Vehicular Technology*, 2012, 61(3): 1032-1042.
- [87] Michael M, Salmen J, Stallkamp J, et al. Real-time Stereo Vision: Optimizing Semi-global Matching [C]//IEEE Intelligent Vehicles Symposium (IV), Gold Coast, Australia, 2013.
- [88] Facciolo G, Franchis C D, Meinhardt E. MGM: A Significantly More Global Matching for Stereovision [C]//British Machine Vision Conference, Swansea, UK, 2015.
- [89] Xiang J, Li Z Y, Blaauw D, et al. Low Complexity Optical Flow Using Neighbor-Guided Semi-global Matching [C]//IEEE International Conference on Image Processing (ICIP), Phoenix, USA, 2016.
- [90] Scharstein D, Tanai T, Sinha S N. Semi-global Stereo Matching with Surface Orientation Priors [C]//2017 International Conference on 3D Vision (3DV), Qingdao, China, 2018.
- [91] Lee Y, Park M G, Hwang Y, et al. Memory-Efficient Parametric Semiglobal Matching [J]. *IEEE Signal Processing Letters*, 2018, 25(2): 194-198.
- [92] Li Z, Xiang J, Gong L, et al. Low Complexity, Hardware-Efficient Neighbor-Guided SGM Optical Flow for Low-Power Mobile Vision Applications [J]. *IEEE Transactions on Circuits and Systems for Video Technology*. 2019, 29(7): 2191-2204.
- [93] Ji Yanhua. Research on Semi-Global Dense Matching Algorithm for Aerial Images Using Edge Information [D]. Wuhan: Wuhan University, 2020. (纪艳华. 顾及边缘信息的航摄影像半全局密集匹配算法研究 [D]. 武汉: 武汉大学, 2020.)
- [94] Sun J, Zheng N N, Shum H Y. Stereo Matching Using Belief Propagation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(7): 787-800.
- [95] Felzenszwalb P F, Huttenlocher D P. Efficient Belief Propagation for Early Vision [J]. *International Journal of Computer Vision*, 2006, 70(1): 41-54.
- [96] Yang Q X, Wang L, Ahuja N. A Constant-Space Belief Propagation Algorithm for Stereo Matching [C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 2010.
- [97] Xiang X Q, Zhang M M, Li G X, et al. Real-time Stereo Matching Based on Fast Belief Propagation [J]. *Machine Vision and Applications*, 2012, 23(6): 1219-1227.
- [98] Noorshams N, Wainwright M J. Stochastic Belief Propagation: A Low-Complexity Alternative to the

- Sum-product Algorithm [J]. *IEEE Transactions on Information Theory*, 2013, 59(4): 1981-2000.
- [99] Komodakis N, Tziritas G. Image Completion Using Efficient Belief Propagation via Priority Scheduling and Dynamic Pruning [J]. *IEEE Transactions on Image Processing*, 2007, 16(11): 2649-2661.
- [100] Ayoobi H, Rezaeian M. Swift Distance Transformed Belief Propagation Using a Novel Dynamic Label Pruning Method [J]. *IET Image Processing*, 2020, 14(9): 1822-1831.
- [101] Hamzah R A, Ibrahim H, Abu Hassan A H. 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.
- [102] Hu Xin, Wang Renli, Wang Jianrong. Principle and Methods of Photogrammetric Positioning for Space Linear Imagery [M]. Beijing: Surveying and Mapping Publishing House, 2018. (胡莘, 王仁礼, 王建荣. 航天线阵影像摄影测量定位理论与方法 [M]. 北京: 测绘出版社, 2018.)
- [103] Zhang Jinming. Structure from Motion in the Geosciences [M]. Beijing: Publishing House of Electronics Industry, 2017. (张锦明. 运动恢复结构(SfM): 理论方法及地球科学应用 [M]. 北京: 电子工业出版社, 2017.)
- [104] Gao Xiang, Zhang Tao, Liu Yi. Fourteen Lectures on Visual SLAM: From Theory to Practice [M]. 2nd ed. Beijing: Publishing House of Electronics Industry, 2019. (高翔, 张涛, 刘毅. 视觉SLAM十四讲: 从理论到实践 [M]. 2版. 北京: 电子工业出版社, 2019.)
- [105] Tian Mao. Image Hierarchical Dense Matching with Slanted Plane Structures Constrains [D]. Wuhan: Wuhan University, 2019. (田茂. 倾斜平面约束的影像层次密集匹配 [D]. 武汉: 武汉大学, 2019.)
- [106] Yang Wenhuan. Digital Surface Model Generation of Weak Texture Regions in Optical Satellite Remote Sensing Images [D]. Wuhan: Wuhan University, 2020. (杨文环. 光学卫星遥感影像弱纹理区域数字表面模型提取方法研究 [D]. 武汉: 武汉大学, 2020.)
- [107] Bobick A F, Intille S S. Large Occlusion Stereo [J]. *International Journal of Computer Vision*, 1999, 33(3): 181-200.
- [108] Chambon S, Crouzil A. Towards Correlation-based Matching Algorithms That Are Robust near Occlusions [C]//Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, UK, 2004.
- [109] Bleyer M, Gelautz M. Graph-Cut-based Stereo Matching Using Image Segmentation with Symmetrical Treatment of Occlusions [J]. *Signal Processing: Image Communication*, 2007, 22(2): 127-143.
- [110] Wang Y C, Tung C P, Chung P C. Efficient Disparity Estimation Using Hierarchical Bilateral Disparity Structure Based Graph Cut Algorithm with a Foreground Boundary Refinement Mechanism [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2013, 23(5): 784-801.
- [111] Jang W, Ho Y. Discontinuity Preserving Disparity Estimation with Occlusion Handling [J]. *Journal of Visual Communication and Image Representation*, 2014, 25(7): 1595-1603.
- [112] Pang X, Zhang Y, Han J, et al. A Stereo Matching in Complex Scene and Occlusion Handling Model [J]. *Optical Review*, 2014, 21(6): 826-832.
- [113] De-Maeztu L, Villanueva A, Cabeza R. Near Real-time Stereo Matching Using Geodesic Diffusion [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(2): 410-416.
- [114] Lin C H, Liu C W. Accurate Stereo Matching Algorithm Based on Cost Aggregation with Adaptive Support Weight [J]. *The Imaging Science Journal*, 2015, 63(8): 423-432.
- [115] Yang Q X, Wang L, Yang R G, et al. Stereo Matching with Color-Weighted Correlation, Hierarchical Belief Propagation, and Occlusion Handling [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(3): 492-504.
- [116] Hirschmuller H. Stereo Processing by Semiglobal Matching and Mutual Information [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(2): 328-341.
- [117] Malathi T, Bhuyan M K. Asymmetric Occlusion Detection Using Linear Regression and Weight-based Filling for Stereo Disparity Map Estimation [J]. *IET Computer Vision*, 2016, 10(7): 679-688.
- [118] Sun J, Li Y, Kang S B, et al. Symmetric Stereo Matching for Occlusion Handling [C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, USA, 2005.
- [119] Yuan Wei. Dense Image Matching and Its Application Based on Optical Flow Field [D]. Wuhan: Wuhan University, 2020. (袁巍. 基于光流场的密集影像匹配理论与实践 [D]. 武汉: 武汉大学, 2020.)
- [120] Wu B, Zhang Y S, Zhu Q. A Triangulation-based Hierarchical Image Matching Method for Wide-Baseline Images [J]. *Photogrammetric Engineering &*

- Remote Sensing*, 2011, 77(7): 695-708.
- [121] Li H, Zhang X G, Sun Z. A Line-based Adaptive-weight Matching Algorithm Using Loopy Belief Propagation [J]. *Mathematical Problems in Engineering*, 2015, 2015: 1-13.
- [122] Xie Y, Feng D, Chen H, et al. An Omni-scale Global-local Aware Network for Shadow Extraction in Remote Sensing Imagery [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 193(7): 29-44.
- [123] Qin R J, Chen M, Huang X, et al. Disparity Refinement in Depth Discontinuity Using Robustly Matched Straight Lines for Digital Surface Model Generation [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, 12(1): 174-185.
- [124] Lu X, Qin R, Huang X. Using Orthophoto for Building Boundary Sharpening in the Digital Surface Model [C]//ASPRS 2019 Annual Conference Denver, Colorado, USA, 2019.
- [125] Zhang K, Lu J B, Lafruit G. Scalable Stereo Matching with Locally Adaptive Polygon Approximation [C]//2008 15th IEEE International Conference on Image Processing, San Diego, USA, 2008.
- [126] Jen Y H, Dunn E, Fite-Georgel P, et al. Adaptive Scale Selection for Hierarchical Stereo [C]// 22nd British Machine Vision Conference, Dundee, Scotland, 2011.
- [127] Zeglazi O, Rziza M, Amine A, et al. A Hierarchical Stereo Matching Algorithm Based on Adaptive Support Region Aggregation Method [J]. *Pattern Recognition Letters*, 2018, 112: 205-211.
- [128] Ma N, Men Y B, Men C G, et al. Accurate Dense Stereo Matching Based on Image Segmentation Using an Adaptive Multi-cost Approach [J]. *Symmetry*, 2016, 8(12): 159.
- [129] Kordelas G A, Alexiadis D S, Daras P, et al. Enhanced Disparity Estimation in Stereo Images [J]. *Image and Vision Computing*, 2015, 35: 31-49.
- [130] Furuta R, Ikehata S, Yamaskai T, et al. Efficiency-Enhanced Cost-Volume Filtering Featuring Coarse-to-Fine Strategy [J]. *Multimedia Tools and Applications*, 2018, 77(10): 12469-12491.
- [131] Ham B, Cho M, Ponce J. Robust Guided Image Filtering Using Nonconvex Potentials [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(1): 192-207.
- [132] Zhu C T, Chang Y Z. Hierarchical Guided-Image-Filtering for Efficient Stereo Matching [J]. *Applied Sciences*, 2019, 9(15): 3122.
- [133] Chuang T Y, Ting H W, Jaw J J. Dense Stereo Matching with Edge-Constrained Penalty Tuning [J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 15(5): 664-668.
- [134] Patil S, Prakash T, Comandur B, et al. A Comparative Evaluation of SGM Variants (Including a New Variant, tMGM) for Dense Stereo Matching [J]. arXiv: 1911.09800.
- [135] Yao P, Feng J Q. Ensemble Learning with Advanced Fast Image Filtering Features for Semi-global Matching [J]. *Machine Vision and Applications*, 2021, 32(4): 1-15.
- [136] Mozerov M G, van de Weijer J. Accurate Stereo Matching by Two-step Energy Minimization [J]. *IEEE Transactions on Image Processing*, 2015, 24(3): 1153-1163.
- [137] Haller I, Nedeveschi S. Design of Interpolation Functions for Subpixel-Accuracy Stereo-Vision Systems [J]. *IEEE Transactions on Image Processing*, 2012, 21(2): 889-898.
- [138] Yang Q X, Yang R G, Davis J, et al. Spatial-depth Super Resolution for Range Images [C]//IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, USA, 2007.