基于深度学习的序列图像深度估计技术

梁欣凯,宋 闯,赵佳佳

(北京机电工程研究所 复杂系统控制与智能协同技术重点实验室,北京 100074)

摘 要:针对单帧图像深度估计的弱泛化性,提出了基于深度学习的序列图像深度估计技术,利用深度卷积神经网络作为基础框架,结合对极几何约束,构建从序列图像到图像对应深度信息的端对端映射,实现仅依赖序列图像信息的无监督深度估计。同时,构建了一类基于场景三维几何信息的损失函数,舍弃原始基于图像间重投影误差的损失函数,提高算法鲁棒性。最后,通过开源数据库验证了算法的准确性和精度,同时,通过红外图像数据集验证了算法的泛化性,为军事领域应用奠定了基础。

关键词:深度学习; 对极几何; 深度估计

中图分类号:V557+.4 **文献标志码**:A **DOI**:10.3788/IRLA201948.S226002

Depth estimation technique of sequence image based on deep learning

Liang Xinkai, Song Chuang, Zhao Jiajia

(Science and Technology on Complex System Control and Intelligent Agent Cooperation Laboratory, Beijing Electro-Mechanical Engineering Institute, Beijing 100074, China)

Abstract: Aiming at the weak generalization of single-frame image depth estimation, a depth estimation technique based on deep learning was proposed, which used deep convolutional neural network as the basic framework and combined the epipolar geometry constraints to construct the end-to-end mapping from sequence images to the depth information, enabling unsupervised depth estimation that only relied on sequence image information. At the same time, a kind of loss function based on three-dimensional geometric information of the scene was constructed, and the original loss function based on the re-projection error between images was discarded to improve the robustness of the algorithm. Finally, the accuracy and precision of the algorithm were verified by the open source database. At the same time, the generalization of the algorithm was verified by the infrared image dataset, which laid a foundation for the military application.

Key words: deep learning; epipolar geometry; depth estimation

收稿日期:2019-05-12; **修订日期**:2019-06-20

基金项目:国防基础科研计划(JCKY2017204B064);国家自然科学基金(61803356)

作者简介:梁欣凯(1992-),男,工程师,硕士,主要从事视觉定位与深度估计方面的研究。Email:lxk820@126.com

0 引 言

对空间的精细化立体感知能力是无人机、机器人等民用应用的基础,也是军事领域“侦查–定位–决策–打击”OODA 环的关键,因此,如何高效方便获得立体空间信息就显得至关重要。

一方面受限于无人飞行器尺寸、载荷与成本,另一方面受限于应用场景,雷达激光等主动传感器设备并不满足实际要求。为了克服上述困难,研究人员借助图像这种信息量丰富的载体,提出了基于图像的深度估计技术。基于图像的深度估计技术通常分为两个领域:基于多视图立体视觉估计和基于图像的深度回归预测。

基于多视图立体视觉估计,通常通过对连续帧图像进行特征提取、特征匹配、匹配代价计算、匹配代价聚合、深度估计、深度优化等一系列步骤进行深度估计 [1]。其进一步延伸出视觉同时定位与地图构建技术(vSLAM)和运动恢复结构技术(SFM),这些技术广泛应用于室内机器人等。但是,当光照、天气环境等发生重大变化,或出现大范围遮挡、小范围高动态变化、重叠或运动模糊等,提取的特征往往随之发生严重畸变,进而造成图像间特征匹配错误,最终导致深度估计失败。

借助卷积神经网络对数据高级抽象特征提取的优势,深度学习技术[1-4]在基于图像的深度回归预测方面取得突破性进展。因其不再利用图像间特征匹配来估计深度信息,从本质上克服了光照、天气环境、遮挡、运动模糊等影响,改善了深度估计的性能。输入带有标注深度信息的图像 [5],首次通过深度神经网络,建立深度信息与图像像素信息间端对端的映射函数,并分别从整合语义信息[6]和利用残差网络[7]进行优化。伦敦大学学院 Godard[8]利用左右图像一致性假设,在训练样本仅为双目立体图像数据的前提下,通过自监督方式完成图像深度估计;利用自身运动信息与重投影误差,伯克利[9]通过无监督策略学习图像间姿态变化关系信息以及图像对应的深度信息,降低了对真实深度信息的保障要求;DEMON[10]与 DEMVS[11]利用堆叠的卷积神经网络,综合利用光流信息、运动信息、场景深度信息,建立图像间像素变化关系信息到场景深度信息和相机姿态变化信息间的双重映射关系,首次将运动视差与深度学习结合起来。但是以上的所有方案,其深度预测的先验知识往往来自于训练的数据样本,当应用场景与训练场景发生明显差异时,过分拘泥于结构化信息的网络往往错误预测图像深度值,引发算法泛化性不足的难题,难以满足军事应用需要。

为了解决以上所提到图像深度估计面临的问题,文中研究了基于深度学习序列图像深度估计技术,一方面,利用深度学习的抽象信息表征优势代替基于多视图立体视觉估计中手工特征提取与匹配过程,一方面,结合对极几何约束弱化网络模型对场景结构化信息的敏感性,另一方面,不再使用图像重投影误差而是直接利用预测深度值与真实深度值在公共三维参考系中的点云误差作为损失函数。通过开源无人驾驶数据库 KITTI 和红外图像数据样本验证了所提出算法的准确性、有效性与泛化性。

1 基于深度学习的序列图像深度估计技术难点与创新点

利用两张图像间的关系信息计算视差,已经在双目领域取得突破性进展[12-14]。然而基于相邻帧图像的运动视差计算与双目立体匹配相比,依然存在以下难点:(1) 不同视角下用于图像匹配的搜索线由相机内参和相机位姿决定,而双目立体视觉则是沿固定极线搜索相匹配的像素进行深度估计,极线信息固化在网络结构中;(2) 匹配点的三角测量关系不仅仅取决于匹配点图像位置信息,同样受相机内参和相机位姿影响;(3) 由于每个像素点位置受到多视图几何约束,所以图像缩放和翻转等手段不能用于训练样本增广。

为了克服以上难点,文中构建输入为连续帧图像,输出为匹配代价聚合的网络结构,之后利用匹配代价聚合预测图像对应深度信息。由于通过网络结构实现了连续帧图像到匹配代价聚合的转变,因此文中提出的运动视差估计算法不再需要进行像素点匹配,降低寻找搜索线的难度,提高了深度估计效率,保证了算法的实时性。此外,通过引入结合图像梯度信息与深度图梯度信息的损失函数项,不仅平滑了预测深度值,也通过其场景分割内涵,降低了“中心死区”风险。最后,将深度值与自身姿态运动结

合转为公共参考系下的场景三维几何信息，并将其作为损失函数项，与图像重投影误差作为损失函数项相比，提高了算法运动视差估计的鲁棒性。

2 基于深度学习的序列图像深度估计技术网络结构

为了利用神经网络结构构建序列图像的匹配代价聚合关系，首先需要明确序列图像间的姿态变化关系，因此，文中所提出的网络结构必然包含两部分：(1) 预测序列图像间位姿变化关系的位姿估计网络；(2) 预测运动视差信息的深度估计网络，其基本架构如图 1 所示。

值得注意的是，文中所提算法的训练数据仅仅是视频流数据，并不包含任何深度值信息、姿态变化信息等，是端对端的无监督学习方式。整个算法流程如下：

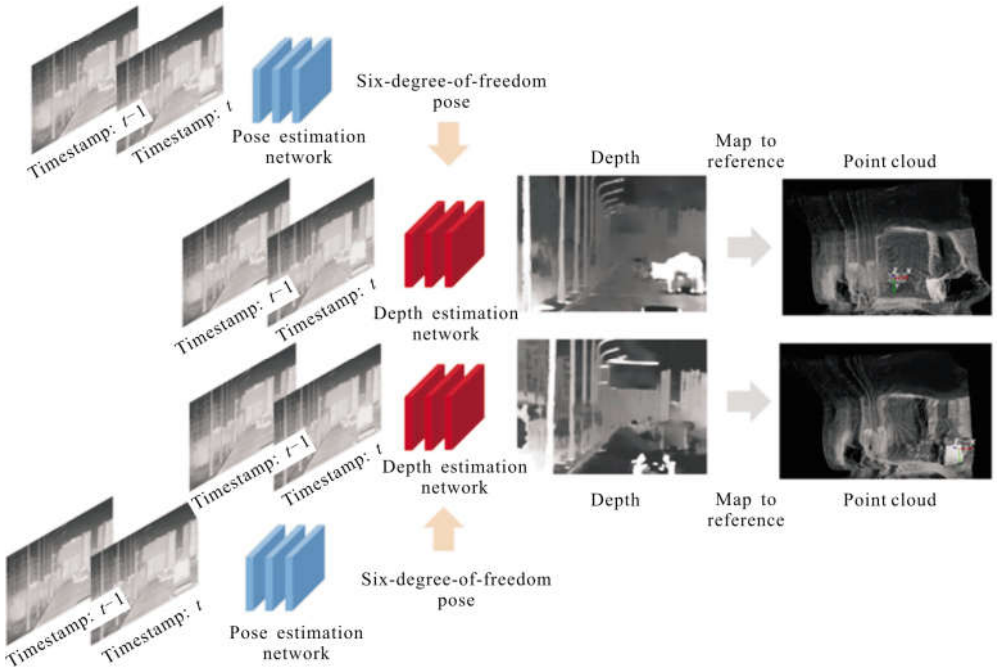


图 1 基于深度学习的序列图像深度估计网络结构

Fig.1 Network structure of depth estimation technique of sequence image based on deep learning

首先，利用位姿估计网络回归连续帧图像之间的位姿变化信息；之后，利用预测信息与连续帧图像像素信息进行融合(详见损失函数部分)，得到初始匹配代价聚合，并输入深度估计网络得到预测深度值，同时进一步将深度值映射为公共参考系下的点云信息；最后，利用预测的点云信息、深度值、位姿变化信息、像素信息等综合考虑用于无监督学习的损失函数(详见损失函数部分)，在提高算法性能的同时降低了数据保证要求。

2.1 位姿估计网络

位姿估计网络输入量为连续帧图像信息，输出为六自由度的位姿变化信息，如图 2 所示，为一系列残差网络结构构成的孪生网络。首先通过权值共享的残差网络分别对连续帧图像进行特征粗提取，在不影响算法性能的同时一方面降低相关训练难度，另一方面缩小训练模型；然后将所提取的特征在通道处连接，最后利用残差网络处理连接后的特征。为了避免全卷积直接回归位姿所引发的精度与算法复

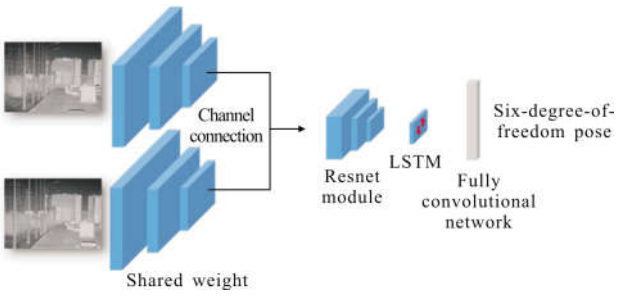


图 2 位姿估计网络

Fig.2 Pose estimation network

杂度问题，文中使用从上到下与从下到上的两类长短期记忆网络(LSTM)首先处理最后一次卷积之后的特征图,最后再利用全卷积进行位姿回归。

2.2 深度估计网络

深度估计网络的输入为连续帧图像及其预测的六自由度位姿变化信息，输出为每张图像在三种分辨率下对应的深度值 (0.062 5 倍,0.25 倍,1 倍),其结构如图 3 所示。为了降低其中降采样过程对特征图分辨率的影响，更好地提取不同尺度空间下有效特征信息,为提高算法精度,在网络结构末端引入了多孔空间金字塔池化提取的特征。其中上层匹配代价聚合的输入为 0.25 倍原始分辨率的图像,多孔空间金字塔池化包含两种不同的卷积，其中扩张率分别为 6、12;下层匹配代价聚合的输入为原始分辨率的图像,多孔空间金字塔池化包含三种不同的卷积,其中扩张率分别为 6、12、18。同时将上卷积的结果与多孔空间金字塔池化的结果进行耦合，若特征维数出现差异,则进行插值保持维数一致。

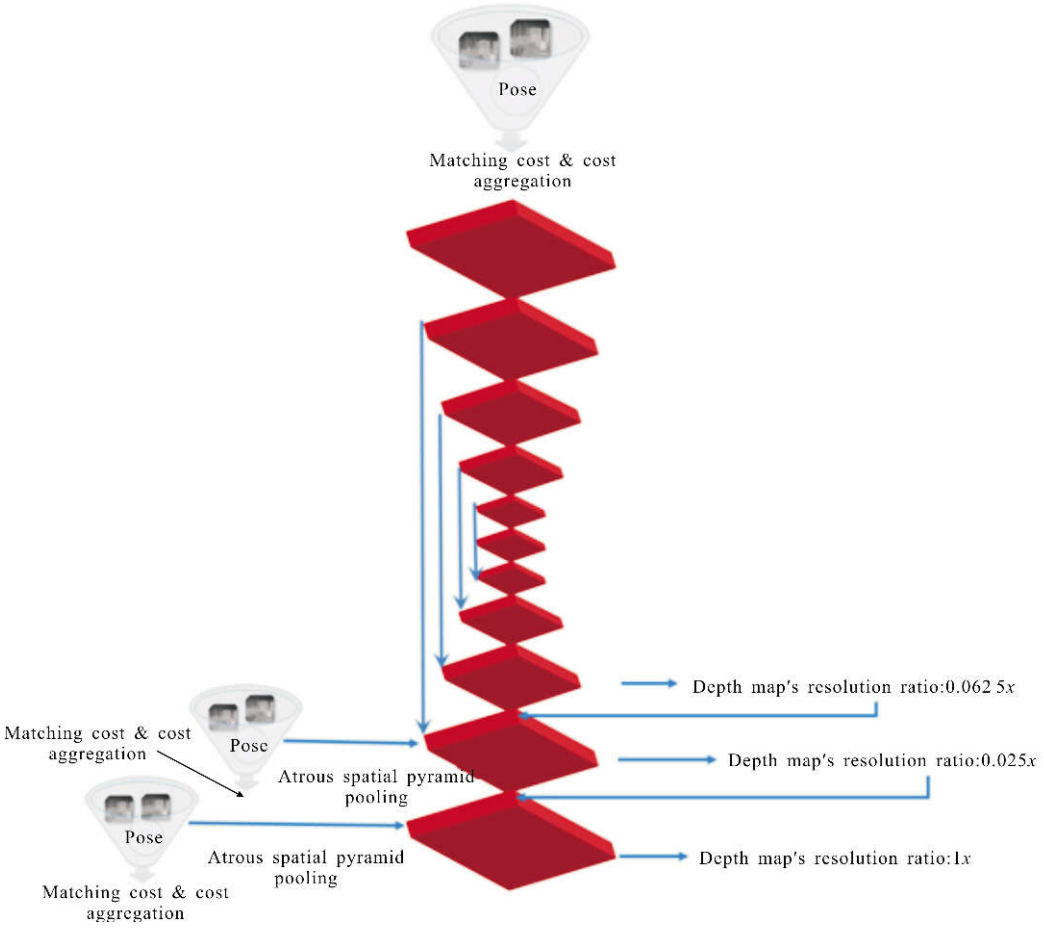


图 3 深度估计网络

Fig.3 Depth estimation network

3 基于深度学习的序列图像深度估计技术损失函数

由于深度估计网络与位姿估计网络深度耦合，因此，必须设计一类能够有效同时训练两类网络的损失函数。目前而言,有关无监督深度学习的深度估计算法往往考虑利用传统二维图像重投影误差作为损失函数，但是该类损失函数更关注图像一致性信息,极易受光照等外界因素影响,尤其不适合应用于复杂的红外图像数据,因此,文中构建有关场景三维几何信息的损失函数，其更重视整个场景几何信息

的一致性，对非结构化信息引起的图像变化具有较强的鲁棒性。

损失函数构建过程如下：

假设 I_t 与 I_{t-1} 分别表示 t 时刻与 $t-1$ 时刻的图像信息，RT_t 表示 $t-1$ 时刻到 t 时刻的相机位姿变化矩阵，D_t 与 D_{t-1} 分别表示 t 时刻与 $t-1$ 时刻图像对应的深度值信息。如果预测得到 D_t，那么其对应的场景空间几何信息 S_t 为：

$$S_t^{ij}=D_t^{ij}*K^{-1}*[i\quad j\quad 1]^{\mathrm{T}} \tag{1}$$

式中：K 为相机内参。

那么当已知 RT_t，那么就可以通过 t 时刻对应的场景空间几何信息 S_t 预测得到 t 时刻图像在 $t-1$ 时刻对应的场景空间几何信息 $\hat{S}_{t-1}=RT_tS_t$，又因为从空间几何信息映射回相机坐标系为 $K\hat{S}_{t-1}$，因此，结合公式(1)可得：

$$\hat{I}_t^{ij}=I_{t-1}^{\hat{i}\hat{j}} \tag{2}$$

式中：$[\hat{i}\quad \hat{j}\quad 1]^{\mathrm{T}}=K*RT_t*(D_t^{ij}*K^{-1}*[i\quad j\quad 1]^{\mathrm{T}})$，$\hat{I}_t^{ij}$ 可以由 I_{t-1}^{ij} 在 $(\hat{i},\hat{j})$ 双线性插值等手段得到。

虽然可以通过公式(2)实现 I_t 与 I_{t-1} 之间的转换，但是由于相机位姿变化的影响，I_{t-1} 的像素点可能在 I_t 边界以外区域，因此，需要设计一个 mask 区域($\mathrm{Mask}_{t-1\to t}$ 意味着影响 $\hat{I}_t$ 的 I_{t-1} 中有效区域)，避免出现非 mask 区域的像素影响整个预测精度。

因此，可以建立如下关于图像重投影的损失函数

$$\mathrm{loss\_IMG}=\|(I_t-\hat{I}_t)\mathrm{Mask}_{t-1\to t}\|_2 \tag{3}$$

然而其必然存在 $\hat{I}_t$ 插值造成的误差，也易受光照、阴影等影响，因此，文中采用基于场景三维几何信息的损失函数代替公式(3)。

基于场景三维几何信息的损失函数，顾名思义，即利用 t 时刻场景空间几何信息真实值 S_t 与预测值 $\hat{S}_t$ 间误差以及 $t-1$ 时刻场景空间几何信息真实值 S_{t-1} 与预测值 $\hat{S}_{t-1}$ 间误差。其本质利用了迭代最近点算法，通过寻找两组点云数据间的刚性变换，使得两组点云数据满足某种度量准则下的最优匹配，即：

$$\underset{RT}{\mathrm{argmin}}\sum_{ij}\|R*S_t^{ij}-\hat{S}_t^{c(ij)}\|_2 \tag{4}$$

式中：$c()$ 意味着迭代最近点算法找到的两个点云之间的匹配关系；R 代表两组点云之间旋转矩阵。综上所述，基于场景三维几何信息的损失函数如公式(5)所示：

$$\mathrm{loss\_3D}=\|R-I\|_2+\sum_{ij}\|R*S_t^{ij}-\hat{S}_t^{c(ij)}\|_2 \tag{5}$$

式中：I 为单位阵。

为了平滑预测的深度值与加强两幅图像的结构相似性，分别引入了 $\mathrm{loss\_SMOOTH}=\sum_{ij}(\|\partial_y D^{ij}\|\mathrm{e}^{-\|\partial_y I^{ij}\|}+\|\partial_x D^{ij}\|\mathrm{e}^{-\|\partial_x I^{ij}\|})$ 和 $\mathrm{loss\_SSIM}=\|[1-\mathrm{SSIM}(I_t,\hat{I}_t)]\mathrm{Mask}_{t-1\to t}\|_2$，其中，$\mathrm{SSIM}(I_t,\hat{I}_t)=\dfrac{(2\mu_{I_t}\mu_{\hat{I}_t}+c_1)(2\delta_{I_t\hat{I}_t}+c_2)}{(\mu_{I_t}^2+\mu_{\hat{I}_t}^2+c_1)(\delta_{I_t}+\delta_{\hat{I}_t}+c_2)}$，代表均值，$\delta$ 代表方差。

故面向超低空避障的运动视差估计技术总损失函数为：

$$\begin{aligned}\mathrm{loss}=&\mathrm{loss\_3D}*\hat{\sigma}_{\mathrm{3D}}^{-2}+\log\hat{\sigma}_{\mathrm{3D}}^{2}+\\&\mathrm{loss\_SMOOTH}*\hat{\sigma}_{\mathrm{SMOOTH}}^{-2}+\log\hat{\sigma}_{\mathrm{SMOOTH}}^{2}+\\&\mathrm{loss\_SSIM}*\hat{\sigma}_{\mathrm{SSIM}}^{-2}+\log\hat{\sigma}_{\mathrm{SSIM}}^{2}+\\&\mathrm{loss\_IMG}*\hat{\sigma}_{\mathrm{IMG}}^{-2}+\log\hat{\sigma}_{\mathrm{IMG}}^{2}\end{aligned} \tag{6}$$

式中：$\hat{\sigma}_{\mathrm{3D}}$，$\hat{\sigma}_{\mathrm{SMOOTH}}$，$\hat{\sigma}_{\mathrm{SSIM}}$ 与 $\hat{\sigma}_{\mathrm{IMG}}$ 是可学习的权值，用来自动平衡各个损失项之间数值关系，避免手工选择的权值导致训练发散。

4 仿真实验

文中通过开源数据库 KITTI 验证了所提出的运动视差估计方案的有效性和准确性，并与其他主流的深度学习深度估计算法进行了比较，最终通过车载红外导引头图像数据验证了算法的泛化性能。

4.1 数据库样本

KITTI 数据集由德国卡尔斯鲁厄理工学院和丰田美国技术研究院联合创办，是目前国际上最大的自动驾驶场景下的计算机视觉算法评测数据集，其中包含市区、乡村和高速公路等场景采集的真实图像数据和图像对应的激光雷达的 3D 点云和车辆轨迹信息。

车载红外导引头图像数据：所提供的外场采集数据为车载拍摄，共三个序列，第一个序列目标障碍物为建筑物楼群、第二个序列目标障碍物为高压线塔、第三个序列目标障碍物为烟囱。图 4 分别为

三个序列图片的示例。

(a) 序列 1 示例 (a) Example of sequence 1　(b) 序列 2 示例 (b) Example of sequence 2　(c) 序列 3 示例 (c) Example of sequence 3

图 4 车载红外导引头图像数据

Fig.4 Vehicle infrared seeker image data

4.2 网络训练

为了加强算法的鲁棒性，必须对原始图像进行数据增广处理，然而传统的数据增广手段如增加高斯与椒盐噪声,左右上下翻转,亮度对比度变换等严重影响了图像间对应的几何信息不利于运动视差的估计,因此,文中只利用缩放手段对图像进行处理。

文中利用 ADAM 求解器进行优化处理，其中 ADAM 的权值 β_1=0.9, β_2=0.999, $\varepsilon=10^{-10}$。将初始学习率设置为 0.000 2，并根据训练迭代次数分阶段将学习率进行指数下降，在 0–1 之间随机采样作为$\hat{\sigma}_{3D}$、$\hat{\sigma}_{SMOOTH}$、$\hat{\sigma}_{SSIM}$ 与 $\hat{\sigma}_{IMG}$ 的初值。基于 tensorflow 框架实现所提出的算法，在 NVIDIA Titan X GPU 训练 500 000 次,受平台内存限制,批处理量为 8,输入图像像素为 640×512。

4.3 性能分析

文中所提出算法在以上提到的数据库进行深度估计验证,其实验结果如表 1 所示,预测深度图如图5 所示。

通过表 1 可以看出：文中提出的运动视差估计技术不仅在深度预测精度方面超过了当前主流的深度估计算法,而且极大程度降低了数据的保障要求。除此之外，由于该算法不同于之前算法过分拘泥于场景的结构化信息与布局，而是利用两张图像间匹配对应关系，所以该算法不仅具有良好的场景泛化能力,而且更擅长捕获场景细节信息。

表 1 文中算法与其他基于深度学习深度估计算法在 KITTI 库下的估计性能对比

Tab.1 Comparison of estimated performance between the proposed algorithm and other deep learning depth estimation algorithms under KITTI dataset

Algorithm	Training data	Distance/m	Absolute error	Relative error	
Ref.[6]	Image & depth	80	0.203	1.548	
Ref.[8]	Binocular images	80	0.148	1.344	
Ref.[9]	Sequence image	80	0.198	1.836	
Proposed algorithm	Sequence image	80	0.141	1.184	
Algorithm	Training data	Distance/m	δ<1.25	$\delta<1.25^2$	$\delta<1.25^3$
Ref.[6]	Image & depth	80	0.702	0.890	0.958
Ref.[8]	Binocular images	80	0.803	0.922	0.964
Ref.[9]	Sequence image	80	0.718	0.901	0.960
Proposed algorithm	Sequence image	80	0.812	0.934	0.976

通过图 5 可以看出：虽然 KITTI 数据库自身具有运动模糊,光照剧烈变化等对于深度估计精度产生影响的特性，但是该方案体现的深度估计效果已经可以满足室外应用并且具有良好的细节感知性

图 5 文中算法在 KITTI 的深度估计效果

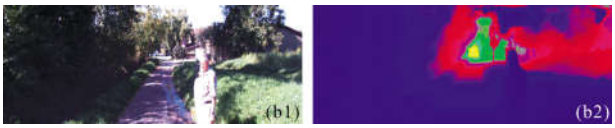
Fig.5 Depth estimation result of the proposed algorithm on KITTI dataset

能,也就证明了该算法对运动模糊、光照剧烈变化具有一定的鲁棒性。

通过图 6 可以看出,该算法对于影响传统深度估计的复杂光照、运动模糊、光照剧烈变化、遮挡、天气环境等因素具有良好的鲁棒性,其中左侧为原始测试图像,右侧为深度估计效果图,以伪彩色形式表现。

(a) 亮度

(a) Illumination



(b) 阴影

(b) Shade

(c) 运动模糊

(c) Motion blur

(d) 遮挡

(d) Occlusion

(e) 天气

(e) Weather

图 6 文中算法在不同环境影响下的深度估计效果

Fig.6 Depth estimation results of the proposed algorithm under different environmental influences

通过图 7 可以看出：即使没有红外数据图像对应的深度信息,通过在其上网络微调,能够在克服红外图像各类复杂特性的基础上，有效地建立图像间的匹配关系信息并预测人为感性认知的深度值。尤其是图 7(c)预测结果,成功判断了近处电线与远处烟囱间的距离关系，为后续军事应用工作奠定了理论基础。

图 7 文中算法对红外图像数据的深度估计效果

Fig.7 Depth estimation result of the proposed algorithm on infrared image data

5 结 论

针对迫切自主空间感知需求，文中提出了基于深度学习的序列图像深度估计技术，一方面继承了深度学习提取高级特征的优势，进而利用对极几何在更高维度更抽象领域进行匹配,预测深度值;另一方面,发扬了深度学习对场景结构化特征的理解,具有对场景三维空间感的感性认知。文中所提出的算法不仅有效克服了纹理不清晰、重叠等传统深度估计技术的桎梏，也突破了现有深度学习深度估计技术泛化性差的瓶颈,达到了良好的深度估计精度,满足实际应用基本要求。此外,该算法对红外序列数据

的试验验证了其强大的泛化性，为后续工程研制工作奠定了基础。

参考文献：

[1] Lin Jinhua, Wang Yanjie. Three-dimentional reconstruction of semantic scene based on RGB-D map [J]. ***Optics and Precision Engineering***, 2018, 26(5): 1231-1241. (in Chinese)
林金花, 王延杰. 三维语义场景复原网络[J]. 光学 精密工程, 2018, 26(5): 1231-1241.

[2] Liu Bo, Xu Tingfa, Li Xiangmin, et al. Adaptive context-aware correlation filter tracking[J]. ***Chinese Optics***, 2019, 12(2): 265-273. (in Chinese)
刘波, 许廷发, 李相民, 等. 自适应上下文感知相关滤波跟踪[J]. 中国光学, 2019, 12(2): 265-273.

[3] Li Yu, Liu Xueying, Zhang Hongqun, et al. Optical remote sensing image retrieval based on convolutional neural networks[J]. ***Optics and Precision Engineering***, 2018, 26(1): 200-207. (in Chinese)
李宇, 刘雪莹, 张洪群, 等. 基于卷积神经网络的光学遥感图像检索[J]. 光学 精密工程, 2018, 26(1): 200-207.

[4] Fan Xiaoting, Li Yi, Luo Xiaowei, et al. Depth estimation based on light field structure characteristic and multiview matching[J]. ***Infrared and Laser Engineering***, 2019, 48(5): 0524001. (in Chinese)
范晓婷, 李奕, 罗晓维, 等. 基于光场结构特性与多视点匹配的深度估计[J]. 红外与激光工程, 2019, 48(5): 0524001.

[5] Eigen D, Puhrsch C, Fergu R. Depth map prediction from a single image using a multi-scale deep network [C]//Proc Conf Neural Information Processing Systems, 2014.

[6] Eigen David, Rob Fergus. Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture [C]//The IEEE International Conference on Computer Vision, 2015.

[7] Laina I, Rupprecht C, Belagiannis V. Deeper depth prediction with fully convolutional residual networks [C]//International Conference on 3D Vision, 2016.

[8] Godard C, Aodha O M, Brostow G J. Unsupervised monocular depth estimation with left-right consistency [C]// The IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[9] Zhou T, Brown M, Snavely N, et al. Unsupervised learning of depth and ego-motion from video [C]//The IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[10] Ummenhofer B, Zhou H, Uhrig J, et al. DeMoN: Depth and motion network for learning monocular stereo [C]//The IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[11] Huang P, Matzen K, Kopf J, et al. DeepMVS: Learning multi-view stereopsis[C]//The IEEE Conference on Computer Vision and Pattern Recognition, 2018.

[12] Kendall A, Martirosyan H, Dasgupta S, et al. End-to-end learning of geometry and context for deep stereo regression [C]//The IEEE International Conference on Computer Vision, 2017.

[13] Wang Xin, Wang Xiangjun. Multiple targets sparse matching for binocular vision positioning system with large field of view [J]. ***Infrared and Laser Engineering***, 2018, 47(7): 0726001. (in Chinese)
王鑫, 王向军. 大视场双目视觉定位系统中多目标稀疏匹配[J]. 红外与激光工程, 2018, 47(7): 0726001.

[14] Qiao Tieying, Cai Lihua, Li Ning, et al. Opposite target measurement based on infrared radiation characteristic system [J]. ***Chinese Optics***, 2018, 11(5): 804-811. (in Chinese)
乔铁英, 蔡立华, 李宁, 等. 基于红外辐射特性系统实现对面目标测量[J]. 中国光学, 2018, 11(5): 804-811.