

基于红外立体视觉的行人感知方法研究

王向军^{1,2}, 杨寿常^{1,2}, 陈瑞祥^{1,2}

(1. 天津大学 精密测试技术及仪器国家重点实验室, 天津 300072;

2. 天津大学 微光机电系统技术教育部重点实验室, 天津 300072)

摘要: 基于热红外特性, 红外立体视觉路况行人感知方法可以在夜间、雾霾环境下有效检测道路场景中的行人等目标, 提高驾驶安全性。针对红外图像中纹理细节少, 传统稠密双目立体匹配算法效果差的问题, 本文首先根据目标在红外图像下的亮度、边缘特征提取感兴趣区域 (Region of interest, ROI); 然后在 ROI 中提取图像特征点并匹配, 进而计算原始稀疏深度图; 最后根据目标表面深度变化较小的特点, 结合 ROI 和原始深度图估计半稠密深度图。本文搭建了实验系统验证该方法的有效性。实验结果表明, 在系统约 120°观测视场角内, 该方法对行人等目标深度感知相对误差在 15 m 范围内优于 1.5%, 30 m 范围内优于 3%。

关键词: 长波红外; 立体视觉; SURF; 深度图; 辅助驾驶

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1001-8891(2021)07-0702-07

Pedestrian Perception Method Based on Infrared Stereo Vision

WANG Xiangjun^{1,2}, YANG Shouchang^{1,2}, CHEN Ruixiang^{1,2}

(1. State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China;

2. MOEMS Education Ministry Key Laboratory, Tianjin University, Tianjin 300072, China)

Abstract: Based on the thermal infrared characteristics, the infrared stereo vision pedestrian perception method can effectively detect and measure pedestrians in road scenes at night and hazy environments, with the aim of improving driving safety. Owing to less texture details in infrared images, the traditional dense binocular stereo matching algorithm performs poorly. To solve this problem, the region of interest (ROI) is extracted according to the brightness and edge features of the targets in the infrared image. Then, the image feature points are extracted and matched in the ROI to calculate the original sparse depth map. Finally, according to the small depth difference in the surface of the targets, the semi-dense depth map was estimated by combining the ROI and the original depth map. We designed an experimental system to verify the effectiveness of the proposed method. The experimental results showed that the relative error of the depth perception of pedestrians was better than 1.5% at 15 m and 3% at 30 m in the field of view of approximately 120°.

Key words: long-wave infrared, stereo vision, SURF, depth map, assisted driving

0 引言

根据世界卫生组织的统计, 2016 年道路交通伤害导致 140 万人死亡, 已经是全球第八大死亡原因^[1], 道路交通安全已经成为人类不可忽视的安全威胁之一。为减少道路交通对人类造成的伤害, 目前已经有多种传感器系统被应用于汽车辅助驾驶, 以期在危险情况中及时对驾驶者做出提醒, 如雷达、激光测距雷

达、超声波和摄像机等技术。其中雷达、激光测距雷达和超声波等属于主动测量方案, 在复杂环境下易受干扰; 而基于摄像机的测量方案属于被动测量方法, 抗干扰能力更强, 在辅助驾驶和自动驾驶领域具有广阔的应用前景, 已经成为计算机视觉领域的热门研究方向。

在夜间、雾霾等低能见度天气条件下, 由于驾驶员的视线受阻, 更易发生交通事故。红外线由于波长

较长,具有穿透雾霾的能力,且所有温度高于绝对零度的物体都会产生红外辐射^[2],故对红外线成像的红外相机具有全天候工作能力,特别在夜间、雾霾等对可见光相机成像具有严重干扰的天气条件下具有明显优势。因此,利用红外成像设备作为辅助观测手段以提高行车安全性的方法,得到了广大相关研究人员和国内各大汽车制造厂商的关注^[3]。

目前,基于红外相机的车载辅助驾驶系统主要利用红外相机的全天候工作特性,帮助驾驶员在恶劣条件下观察路面情况,技术手段主要集中在行人检测与跟踪领域。如文献[3]-[4]介绍了车载夜视技术以及红外行人检测算法的发展现状,文献[5]介绍了一种基于YOLO模型的远红外车载图像快速行人检测方法,文献[6]提出了一种相关滤波框架下的红外图像行人跟踪方法,行人的位置估计需依靠驾驶员的经验判断。文献[7]使用激光雷达成像与红外成像融合的方式获取可视图像及深度信息,但成本较高。

基于立体视觉原理获取深度信息,仅需两个相机采集双目图像,然后利用双目视差原理计算目标深度,成本较低。但是在红外图像中,由于纹理细节较少,在可见光立体视觉中使用广泛的BM(Block Matching)和SGBM(Semi Global Block Matching)等传统稠密立体匹配算法不能很好地完成红外图像立体匹配。因此,本文结合道路场景的特点,设计了一种基于红外立体视觉的路况行人感知方法,该方法可以生成感兴趣目标处稠密的全局半稠密深度图。经实验验证,该方法在道路环境下可以有效地提供车前行人等目标的深度信息。

1 双目测距模型

1.1 三角测量模型

本文双目测距模型如图1所示,两相机光心 O_1 、 O_2 存在 x 、 z 方向明显位移 dx 、 dz 。考虑到实际系统的安装误差因素,两相机坐标系间同时存在 y 方向位移 dy 和绕 x 、 y 、 z 三轴的相对旋转 α 、 β 、 γ 。 dx 、 dy 、 dz 和 α 、 β 、 γ 共同构成相机的外部参数,可通过双目相机标定获得。

通过标定获得相机内部及外部参数后,本文采用线性三角测量法对空间点进行解算,如图1,空间点 N 分别投影到两相机像平面 π_1 、 π_2 的 n_1 、 n_2 点处,投影关系可由式(1)表示,其中 n_1 、 n_2 用 π_1 、 π_2 像面坐标系下的齐次坐标表示, P_1 、 P_2 分别表示空间点 N 到像面 π_1 、 π_2 的投影关系,为 3×4 矩阵,由相机内、外部参数计算得到:

$$n_1 = P_1 N, n_2 = P_2 N \quad (1)$$

通过 π_1 、 π_2 像面上的一对匹配点 $n_1 \leftrightarrow n_2$,可根据式(2)^[8]计算得到 N 点在世界坐标系下的齐次坐标。其中, x_i 、 y_i 分别表示 π_i 像面上 N 的像点 n_i 的横、纵坐标, p_i^j 表示 P_i 矩阵的第 j 行($i=1, 2, j=1, 2, 3$)。

$$\begin{bmatrix} x_1 p_1^3 - p_1^1 \\ y_1 p_1^3 - p_1^2 \\ x_2 p_2^3 - p_2^1 \\ y_2 p_2^3 - p_2^2 \end{bmatrix} N = 0 \quad (2)$$

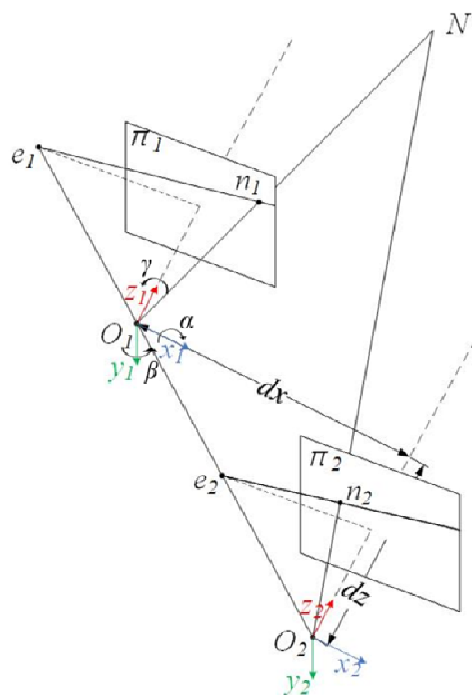


图1 双目测量模型

Fig.1 Binocular measurement model

1.2 极线约束

极线约束是指双目图像中某一图像上的点必然对应于另一图像上的一条线,如图1,直线 $O_1 N$ 上任一点都成像于 π_1 像面上的 n_1 点处,而其在像面 π_2 上的像位于直线 $e_2 n_2$ 上,即 $e_2 n_2$ 为 n_1 的对极线, e_2 为光心 O_1 在像面 π_2 上的投影点,其数学关系可用式(3)表示:

$$I_{e_2 n_2} = F n_1 \quad (3)$$

式中: $I_{e_2 n_2}$ 表示直线 $e_2 n_2$; F 为基本矩阵,可通过相机内、外部参数获得。

由于 n_2 是 N 在像面 π_2 的像点,故 n_2 在直线 $I_{e_2 n_2}$ 上,即 $n_2^T I_{e_2 n_2} = 0$,所以一对匹配点 $n_1 \leftrightarrow n_2$ 满足式(4)关系。

$$n_2^T F n_1 = 0 \quad (4)$$

因此,可利用式(4)对初始匹配结果进行筛选,去除误匹配点。

2 系统设计与搭建

2.1 系统方案设计

驾驶员在驾驶车辆时，一般需要观察车辆正前方和两侧路面情况以确保行车安全。因此，为了满足实际道路应用条件，本文设计了一种大视场红外双目立体视觉环境感知单元，有效视场角约为 120° 。由于目前红外相机成像单元分辨率普遍不高，以及红外相机镜头视场角等因素的制约，本文提出了一种双红外双目测量方案。方案设计如图2所示，即使用两对双目相机分别覆盖左右两侧，组成等效 120° 测量视场。图2中4只相机分别组成左视、右视双目相机组，两相机组的基线长度均设计为75cm，两组双目相机测量视场构成 120° 等效测量视场。该方案将4只红外相机安装在同一直线上，将系统尺寸限制在 $100\text{cm}\times 10\text{cm}\times 10\text{cm}$ 内，有利于在不改变车辆结构的条件下安装使用。

2.2 实验系统平台

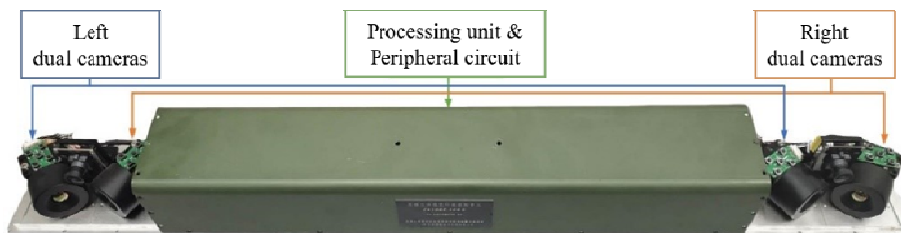


图3 大视场红外立体视觉实验系统

Fig.3 Large field of view infrared stereo vision experimental system

2.3 红外棋盘格标定板

由于三角测量模型需要已知的相机内参和双目相机外参解算匹配点坐标，因此需对相机进行标定。目前应用最广泛的相机标定算法为张正友标定法^[9]，该方法需要单平面棋盘格作为标定板。由于红外辐射量和物体温度有关，物体温度越高，红外辐射越强，红外图像体现的是物体温度差异。根据红外成像原理，本文设计了一种红外棋盘格标定板（下称标定板），如图4所示。

图4(a)为该标定板的可见光图像，图4(b)为红外图像。棋盘格白格为良导热材料制成，背层贴有电热片，黑格为隔热材料制成，背层无电热片，通电后白格区域电热片发热使其温度升高，黑格区域的隔热材质使其保持相对较低的温度，从而使棋盘格的红外图像具有与可见光图像相似的图像特征。本文设计的棋盘格可用于红外相机标定，可见光相机标定和红外、可见光相机联合标定。

依据2.1节所述方案，搭建实验平台，如图3所示，该平台由4只长波红外相机及广角镜头、嵌入式处理单元、外围电路、刚性载体和外壳组成。

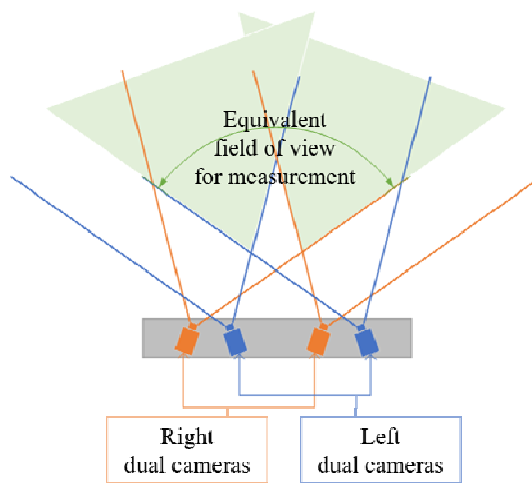
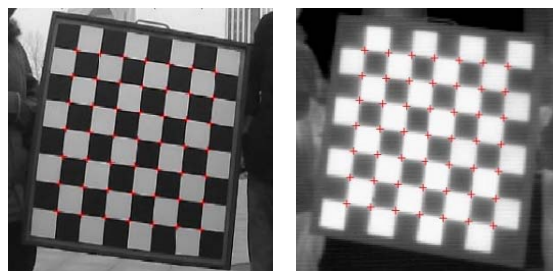


图2 系统设计方案

Fig.2 System design scheme



(a) 可见光图像

(b) 红外图像

(a) Visible image

(b) Infrared image

图4 红外棋盘格标定板

Fig.4 Infrared chessboard

3 深度图生成方法

本文的路况行人感知问题研究定位于道路使用者中的行人和骑行者等目标。由于人体具有较为明显的温度、边缘特征，因此可以此为依据在双目相机左、右图像中寻找感兴趣区域，并在感兴趣区域中进行特征点提取、配准并最终生成可视化半稠密深度图。

3.1 图像幂次变换

图像幂次变换是指通过对图像灰度值的 k 次幂变换,使灰度值较大的位置得到增强,而灰度值较小的位置得到抑制。图像幂次变换可用式(5)表示:

$$i_{\text{new}} = M \left(\frac{i_{\text{old}}}{i_{\text{max}}} \right)^k \quad (5)$$

式中: i_{old} 、 i_{new} 为图像上任一点转换前、后灰度值; i_{max} 为原图像的最大灰度值; M 为处理后的图像最大灰度值,指数 k 越大,图像增强效果越明显,本文应用场景中 k 值的取值范围为 1.6~2.0。

由于红外图像是对红外热辐射的成像,温度高的物体成像位置灰度值大,温度低的物体成像位置灰度值小。一般红外场景中的人体温度较高,背景温度较低,对应红外图像中,行人目标位置灰度值较大,背景位置灰度值较小,通过图像幂次变换,可有效增强

行人目标,抑制背景。如图 5(b),相对原图(图 5(a)),其背景得到明显抑制,行人目标得到增强。

3.2 感兴趣区域提取

由于行人、骑行者的竖向边缘较多^[10],因此可以使用 Sobel 边缘检测算法检测图像中的竖向边缘,如图 5(c) (为了便于观察,图像有亮度增强)所示,图中实线框内为行人目标,虚线框内为无效目标,利用目标灰度约束可以剔除大部分无用边缘信息,如图 5(d)所示,图像中剩余的轮廓信息主要集中在行人位置处。由于轮廓位置所占图像面积较小,仅将该位置作为特征提取区域不利于提取到足够多的特征点,因此使用图像膨胀算法对图 5(d)所示基于灰度约束的边缘提取图像进行处理,得到图 5(e)所示感兴趣区域 (Region of interest, ROI),通过在左右图像 ROI 区域内提取和匹配特征点,可有效减少特征点提取和匹配数量,减少误匹配数量,提高算法运行效率。

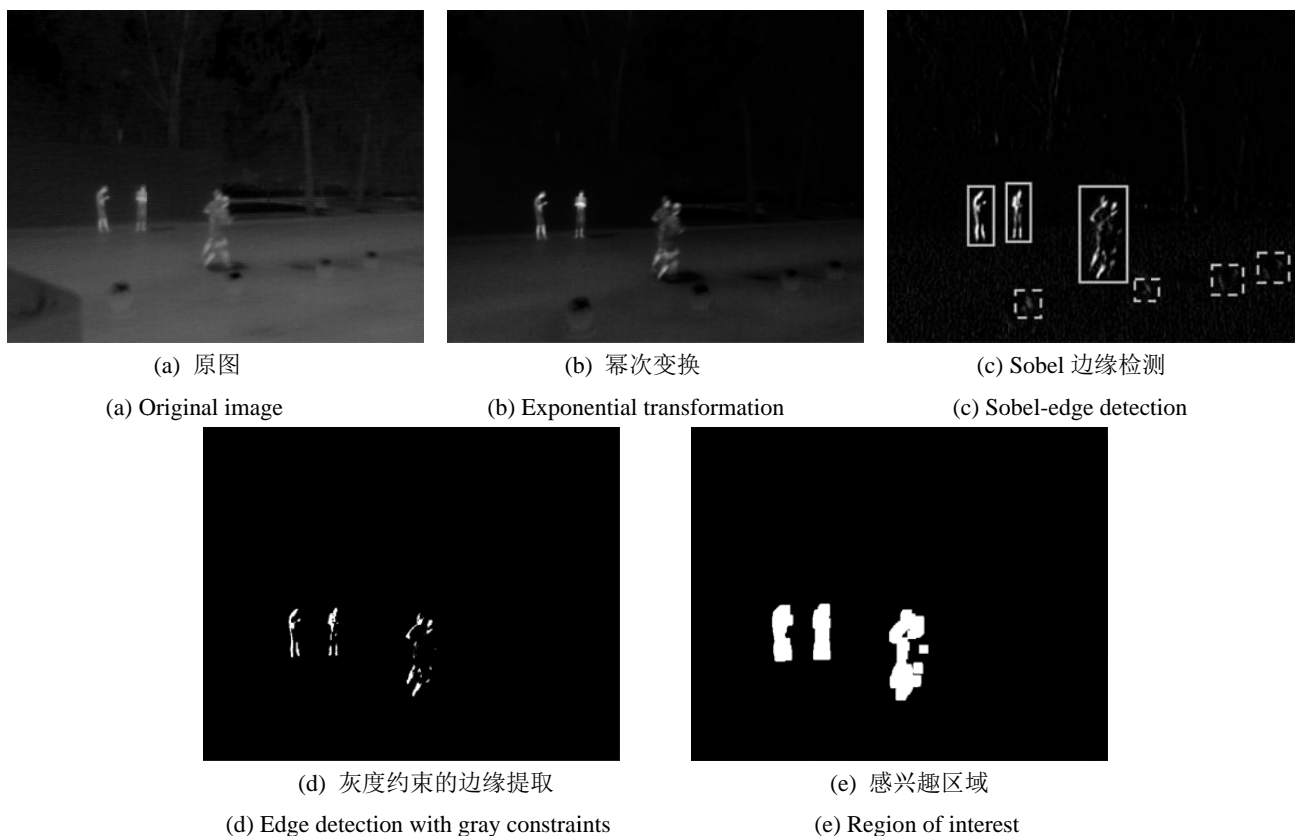


图5 感兴趣区域提取

Fig.5 Region of interest (ROI) detection

3.3 特征点检测与匹配

由于红外图像缺乏一般可见光图像中丰富的纹理细节特征,仅在目标轮廓边缘等梯度变化较大的地方存在明显图像特征,因此本文采用 SURF (Speeded-Up Robust Features) 算法^[11]在 ROI 区域对

图像进行特征提取。SURF 算法基于图像边缘检测特征点,且运行速度快,经实验验证,其在长波红外图像上具有良好的特征点提取效果。利用 SURF 算法在 ROI 区域进行特征点检测和匹配的效果如图 6(a)所示,图中线段的端点为特征点位置,其两端为一对初

始匹配点,可以看出其中存在明显错误匹配点。利用式(4)对匹配点进行筛选,结果如图6(b)所示,图中线段连接的匹配点为筛选后的匹配点,可见图6(a)中的明显错误匹配点被有效剔除。

3.4 半稠密深度图估计

获取筛选后匹配点后,利用三角测量模型可以解算匹配点对应的空间点坐标,获取原始稀疏深度图(如图7(a))。由于原始深度图是稀疏的,因此其直接用于显示效果较差。本文采用种子填充法(Seed-Filling)标记图5(e)中不同连通域,在各连通域中选取原始深度图对应区域中的所有有效深度值的中值作为该区域的代表深度值,得到半稠密深度图,如图7(b)。

由于ROI图像中不同连通域表示场景中的不同目标,同一目标的不同位置的深度值是相近的,因此使用代表深度值估计整个目标的深度是可行的,且代表深度值的中值选取策略可以有效避免少量错误匹配点产生的异常深度值对目标深度估计的影响。

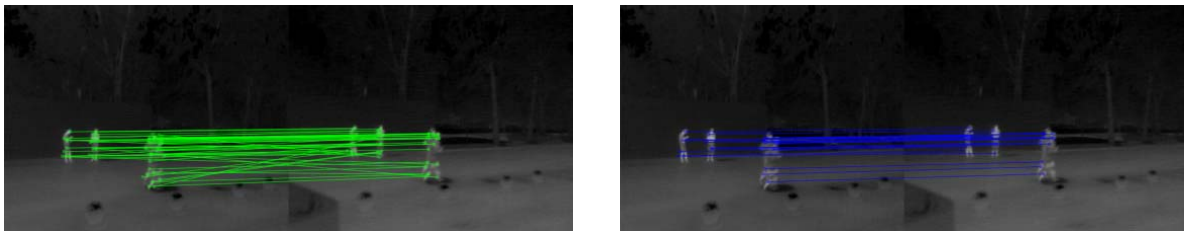
3.5 深度图显示

为便于观察,本文采用从浅到深的渐变颜色表示0~30m内的距离远近。距离越远,颜色越深;距离越近,颜色越浅。图8给出了6组夜间路面测试结果,对照图和显示图依次对应。其中对照图经过幂次变换对背景进行了抑制,显示图在对照图的基础上引入了

渐变颜色表示图像中目标的深度信息,并在图像右侧给出了颜色图例。图8中6组显示图中包含不同距离的行人、跑步者、骑行者等目标。可以看出,行人、跑步者、骑行者等目标均得到较好的颜色标记,参考颜色图例可以判断目标的相对远近,从而为驾驶员在夜间驾驶提供良好的辅助观测信息。

4 测距精度实验

为验证系统测距精度,本文在夜间室外环境下,采用合作目标对左、右视双目相机分别进行精度实验,合作目标为一个2×2红外棋盘格,如图9所示。以左视双目相机为例,合作目标参考位置如图10所示,取3个方向,每个方向5~30m范围内的18个点(图中未全部画出)验证系统测距精度。合作靶标的位置约定真值由精度为2+2ppm的全站仪间接测得,以左相机为基准点,利用全站仪分别测量左相机和合作靶标的中心角点坐标,通过计算得到合作靶标中心距左相机基准点的距离约定真值。将系统测距结果和约定真值作比较,图11给出了系统测距误差,系统测距结果在30m范围内绝对误差均小于1m,相对误差小于3%;15m范围内绝对误差小于0.5m,相对误差小于1.5%。考虑到本文的路况行人感知方法的主要应用场景为向驾驶者提供辅助观测信息,系统的测距精度满足实用需求。



(a) 特征点提取及初始匹配 (b) 极线约束筛选后匹配点
(a) Feature points extraction and initial match (b) Matching points selected by epipolar constraint

图6 特征点提取及匹配

Fig.6 Feature points extraction and match



(a) 原始深度图 (b) 半稠密深度图
(a) Original depth map (b) Semi-dense depth map

图7 深度图(局部)

Fig.7 Depth map (local)



(a) 对照图
(a) Images for Comparison



(b) 显示图
(b) Images for show
图8 路面测试结果
Fig.8 Results of road test



图9 2×2 红外棋盘格
Fig.9 2×2 infrared chessboard

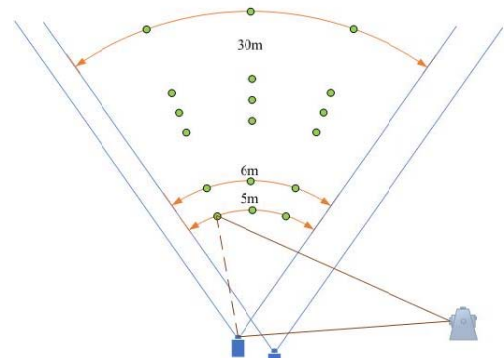


图10 左视双目视场目标参考位置
Fig.10 Reference position of targets in left binocular field of view

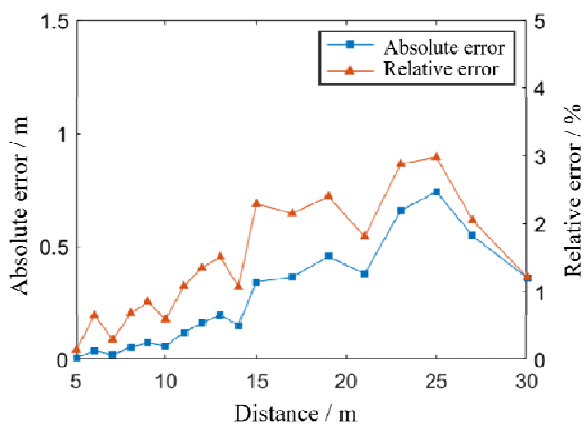


图 11 测距实验结果

Fig.11 Results of distance measurement

5 结论

本文设计了一种基于红外立体视觉的路况行人感知方法。针对红外图像纹理细节少,传统稠密双目立体匹配算法效果差的问题。本文首先提取图像 ROI,然后在其中提取特征点并匹配以生成原始稀疏深度图,最后结合 ROI 和原始稀疏深度图估计 ROI 内稠密的全局半稠密深度图。并设计了实验系统对该方法进行实验验证。实验结果表明,该方法具有良好的行人等目标的检测效果,在系统约 120°观测视场角内,目标深度感知相对误差在 15 m 范围内优于 1.5%, 30 m 范围内优于 3%, 满足实际应用场景。

参考文献:

[1] World Health Organization. The Top 10 Causes of Death [DB/OL]. [2018-05-24]<https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>.

[2] Gade R, Moeslund T B. Thermal Cameras and Applications: a Survey[J]. *Machine Vision and Applications*, 2014, **25**(1): 245-262.

[3] 石永彪, 张湧. 车载红外夜视技术发展研究综述[J]. *红外技术*, 2019, **41**(6): 504-510.

SHI Yongbiao, ZHANG Yong. Survey on development about vehicular infrared night vision technology[J]. *Infrared Technology*, 2019, **41**(6): 504-510.

[4] 许腾, 黄铁军, 田永鸿. 车载视觉系统中的行人检测技术综述[J]. *中国图象图形学报*, 2013, **18**(4): 359-367.

XU Teng, HUANG Tiejun, TIAN Yonghong. Survey on pedestrian detection technology for on-board vision systems[J]. *Journal of Image and Graphics*, 2013, **18**(4): 359-367.

[5] 于博, 马书浩, 李红艳, 等. 远红外车载图像实时行人检测与自适应实例分割[J]. *激光与光电子学进展*, 2020, **57**(2): 293-303.

YU Bo, MA Shuhao, LI Hongyan, et al. Real-time pedestrian detection for far-infrared vehicle images and adaptive instance segmentation[J]. *Laser & Optoelectronics Progress*, 2020, **57**(2): 293-303.

[6] DING M, ZHANG X, CHEN W H, et al. Thermal Infrared Pedestrian Tracking via Fusion of Features in Driving Assistance System of Intelligent Vehicles[J]. *Proceedings of the Institution of Mechanical Engineers Part G-Journal of Aerospace Engineering*, 2019, **233**(16): 6089-6103.

[7] ZHANG W Y, FU X H, LI W. The Intelligent Vehicle Target Recognition Algorithm Based on Target Infrared Features Combined with Lidar[J]. *Computer Communications*, 2020, **155**: 158-165.

[8] Richard Hartley, Andrew Zisserman. 计算机视觉中的多视图几何[M]. 韦穗, 章权兵, 译. 北京: 机械工业出版社, 2019: 240-242.

Richard Hartley, Andrew Zisserman. *Multiple View Geometry in Computer Vision*[M]. WEI Sui, ZHANG Quanbing. Beijing: China Machine Press, 2019: 240-242.

[9] ZHANG Z. A flexible new technique for camera calibration[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(11): 1330-1334.

[10] Bertozzi M, Broggi A, Caraffi C, et al. Pedestrian detection by means of far-infrared stereo vision[J]. *Computer Vision and Image Understanding*, 2006, **106**(2): 194-204.

[11] Bay H, Ess A, Tuytelaars T, et al. SURF: Speeded Up Robust Features[J]. *Computer Vision and Image Understanding*, 2008, **110**(3): 346-359.