

基于双目视觉的深度图拼接

袁瑞峰^{**}, 刘明^{*}, 惠梅, 赵跃进, 董立泉

北京理工大学光电学院精密光电测试仪器及技术北京市重点实验室, 北京 100081

摘要 提出了一种基于双目视觉的深度图像拼接方法。利用预先标定好的双目深度传感器,通过运动采集到具有一定重叠区域的两张深度图,深度图与左目图像的像素点一一对应,对左目图像进行特征提取并匹配,计算出单应性矩阵,进而对两幅深度图进行拼接,并结合单应性矩阵对深度图深度进行矫正,得到最终深度图拼接结果。仿真和实验结果表明,所提方法可以有效扩大双目深度传感器的视场,得到的拼接深度图与单传感器采集到的基本相同。

关键词 图像处理; 双目视觉; 深度图; 单应性矩阵; 拼接

中图分类号 TP391

文献标识码 A

doi: 10.3788/LOP55.121013

Depth Map Stitching Based on Binocular Vision

Yuan Ruifeng^{**}, Liu Ming^{*}, Hui Mei, Zhao Yuejin, Dong Liquan

*Beijing Key Laboratory for Precision Optoelectronic Measurement Instrument and Technology,
School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China*

Abstract A depth image stitching method based on binocular vision is proposed, in which a pre-calibrated binocular depth sensor is used to acquire two depth maps with certain overlapping areas through motion. These depth maps correspond to the pixel points of left-view images one by one. The feature extraction is performed on the left-view images and after matching, the homography matrix is calculated and obtained. Then the two depth maps are further stitched and the depth of these maps is corrected based on the homography matrix, and thus the final depth map stitching results are obtained. The simulation and experimental results show that, the proposed method can effectively expand the field of view of the binocular depth sensor and the obtained stitching depth map is basically same with the depth map acquired by a single sensor.

Key words image processing; binocular vision; depth map; homography matrix; stitching

OCIS codes 100.2960; 330.1400

1 引言

双目立体视觉利用双摄像机模拟人眼成像^[1],利用视差原理得到三维空间中物体的轮廓尺寸及位置。传统的双目深度传感器只能在重叠区域获得深度图,具有视场小的不足,通过深度图像拼接的方法可以得到大视场的深度图,可有效扩大双目深度传感器的视场角^[2],具有很高的应用研究价值,可广泛应用于机器人、虚拟现实、机械制造、三维测量、生物医学、考古勘探等领域。

图像拼接一般通过图像获取、图像配准及图像融合来完成。其中图像配准最为关键,涉及特征提

取与匹配^[3]。常见的特征提取方法有 Harris 角点算法、加速稳健特征 (SURF) 算法、快速特征点提取/描述 (ORB) 算法和尺度不变特征变换 (SIFT) 算法等^[4]。由于 Harris 角点对尺度的敏感性,基于 Harris 角点的特征提取方法稳定性不高。SURF 特征提取方法的计算速度优于 SIFT 算法的,但是在进行丰富纹理图像拼接时效果稍逊于 SIFT 算法的。Rublee 等^[5]提出的 ORB 算法在保证尺度和旋转不变性的前提下,通过角点检测 (FAST) 使特征提取速度有了大幅度的提升,但是在丰富纹理的特征匹配时效果远不如 SIFT 算法的。Lowe^[6]提出的 SIFT 特征提取方法具有尺度不变性和特征描述

收稿日期: 2018-06-05; 修回日期: 2018-06-14; 录用日期: 2018-07-12

基金项目: 国家自然科学基金(61475018,61301190)

* E-mail: bit411liu@bit.edu.cn; ** E-mail: 601663546@qq.com

精准的优势,虽然耗时较长,但是在一些需要高精度而对实时性要求不高的场合仍有很大的应用空间。基于摄像机旋转的深度图像拼接方法^[7]与传统的图像拼接方法原理基本一致,不适用于摄像机一般性运动后的拼接。

为了克服传统双目深度传感器视场小的缺点,提出了一种基于双目视觉的深度图像拼接方法。考虑到左目彩色图与深度图一一对应,在保证一定重叠区域的前提下自由移动双目深度传感器,采集移动前后的彩色图及深度图。以彩色图进行特征提取及匹配,并计算出单应性矩阵,以此单应性矩阵对深度图像进行拼接,并结合空间位置关系对深度图的深度值进行了矫正,最终得到了大视场的深度图拼接结果。

2 双目深度图的拼接算法原理

2.1 双目深度图的获取

两个完全相同的摄像机平行放置,测距原理^[8]

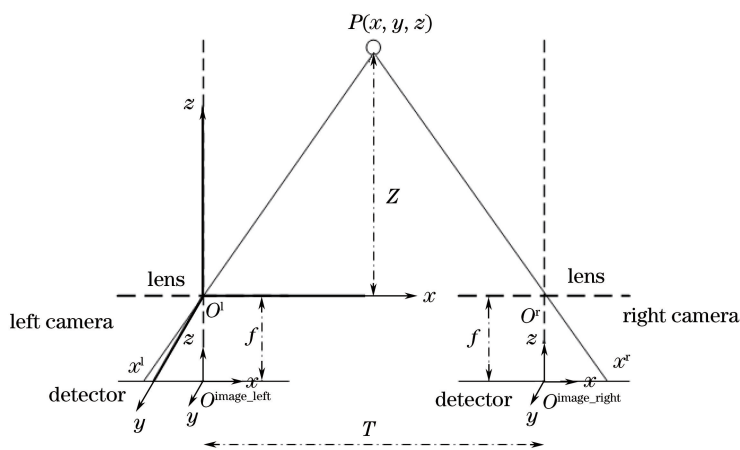


图 1 双目测距原理图

Fig. 1 Schematic of binocular ranging

采用张正友标定法^[9]对双目摄像机进行联合标定,并根据标定结果对左、右摄像机进行极线校正和畸变校正,使左右相机对应的像素点处于同一水平线。采用绝对误差和(SAD)算法进行立体匹配,计算公式为

$$S_{\text{SAD}}(\mu, \nu) = \sum \{ |L_{\text{left}}(\mu, \nu) - R_{\text{right}}(\mu, \nu)| \}, \quad (4)$$

式中 $S_{\text{SAD}}(\mu, \nu)$ 为差的绝对值之和; $L_{\text{left}}(\mu, \nu)$ 为左图一定大小窗口下的像素点灰度值; $R_{\text{right}}(\mu, \nu)$ 为在右图以相同大小窗口进行水平线搜索得到的像素点灰度值。找到两个窗口覆盖像素点差的绝对值之和最小的位置,即 $S_{\text{SAD}}(\mu, \nu)$ 最小的位置,该位置为匹配的像素点。以此方法遍历全图,得到最终的深

度图。如图 1 所示。其中 O^l 、 O^r 分别为左右摄像机的光心, O^l 亦即世界坐标系的坐标原点, $O^{\text{image_left}}$ 、 $O^{\text{image_right}}$ 分别为左、右摄像机图像坐标系的坐标原点, $P(x, y, z)$ 为三维空间中任意一点, Z 为 P 点在世界坐标系中的纵坐标,即深度值。 x^l 为 P 点在左摄像机的像点, x^r 为 P 点在右摄像机的像点, T 为左右摄像机光心间的距离,即基线距离。 f 为左、右摄像机的镜头焦距。 P 点在左、右摄像机像面像点的横向坐标直接存在的差异即视差 $d = x^r - x^l$, 根据相似关系可得

$$\frac{T + x^r - x^l}{Z + f} = \frac{T}{Z}. \quad (1)$$

经过数学计算可得

$$Z = \frac{fT}{x^r - x^l}. \quad (2)$$

将(1)、(2)式代入 $d = x^r - x^l$ 可得

$$Z = \frac{fT}{d}. \quad (3)$$

度图。

2.2 双目深度图的拼接

单应性变换原理^[10]如图 2 所示,主要是通过 $A_1A'_1$ 、 $A_2A'_2$ 、 $A_3A'_3$ 、 $A_4A'_4$ 四对匹配得到的特征点计算出单应性矩阵,以此单应性矩阵直接对平面 π 投影变换到平面 θ , 得到最终的图像拼接结果。考虑到深度图所包含的特征信息较少,双目深度传感器所得的深度图与左目图像的像素点一一对应,移动前后左目彩色图像作为输入,使用 SIFT 算法对两张图像进行特征提取,得到匹配点对后使用随机抽样一致性(RANSAC)算法^[11]剔除误匹配点,最终计算出单应性矩阵。一般情况下考虑单应性矩阵有 8 个自由度,因此只需四对点就能求得单应性矩阵。

通过多对匹配点的迭代优化进一步提高了单应性矩阵计算的准确度。其中 RANSAC 算法能够在离群数据少于内群数据的情况下通过迭代优化的方式估

计出内群数据模型的参数。以此单应性矩阵对移动前后的深度图进行拼接,得到初步的拼接结果。

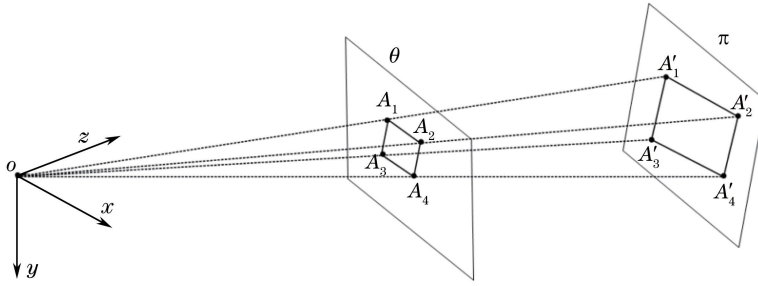


图 2 单应性变换原理

Fig. 2 Schematic of homography transformation

2.3 双目深度图的矫正

传统的图像拼接仅需将移动后的图像投影到移动前的图像像面就能得到最终的结果^[12]。而深度图的像素值具有不同于一般彩色图的意义,深度图的灰度值代表物体的深度值,即空间物体到垂直于

相机光轴的过镜头光心的平面的垂直距离。考虑到移动前后双目传感器的空间位置发生了变化,提出了一种基于单应性矩阵的深度图矫正方法。具体原理图如图 3 所示。

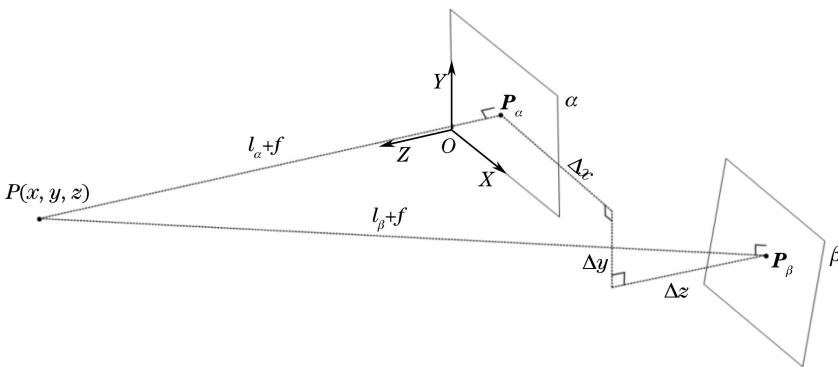


图 3 深度矫正原理图

Fig. 3 Schematic of depth correction

P_α 为点 P 在 α 面上的投影向量, P_β 为点 P 在 β 面的投影向量, $(\Delta x, \Delta y, \Delta z)$ 为两处投影点的坐标差。 α 面为移动前的深度图位置, β 面为移动后的深度图位置。以 α 面为基准, 将 β 面以单应性矩阵进行投影变换。变换关系^[13]为

$$P_\alpha = \mathbf{H}P_\beta, \quad (5)$$

式中 \mathbf{H} 为变换矩阵, \mathbf{H} 的表达式为

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}, \quad (6)$$

式中 h_{ij} ($i, j=1, 2, 3$) 为矩阵元。

将(6)式代入(5)式得到

$$\begin{bmatrix} x_\alpha \\ y_\alpha \\ z_\alpha \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_\beta \\ y_\beta \\ z_\beta \end{bmatrix}. \quad (7)$$

故

$$\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} h_{11} - 1 & h_{12} & h_{13} \\ h_{21} & h_{22} - 1 & h_{23} \\ h_{31} & h_{32} & h_{33} - 1 \end{bmatrix} \begin{bmatrix} x_\beta \\ y_\beta \\ z_\beta \end{bmatrix}, \quad (8)$$

式中 $\Delta x = |x_\alpha - x_\beta|$, $\Delta y = |y_\alpha - y_\beta|$, $\Delta z = |z_\alpha - z_\beta|$; l_α 为 α 面深度图的深度值; l_β 为 β 面深度图的深度值。

根据几何关系可得

$$l_\alpha = \sqrt{(l_\beta + f)^2 - (\Delta x)^2 - (\Delta y)^2} - \Delta z - f. \quad (9)$$

根据(9)式对 β 面深度图的深度值进行矫正, 得到最终的深度图拼接结果, 具体算法流程图如图 4 所示。通过移动前后双目深度传感器采集到具有一定重叠区域的彩色图和深度图, 以彩色图作为输入

计算单应性矩阵,以此单应性矩阵对深度图进行初步拼接,结合(9)式对深度图的初步拼接结果进行矫正,得到最终拼接结果。

正,得到最终拼接结果。

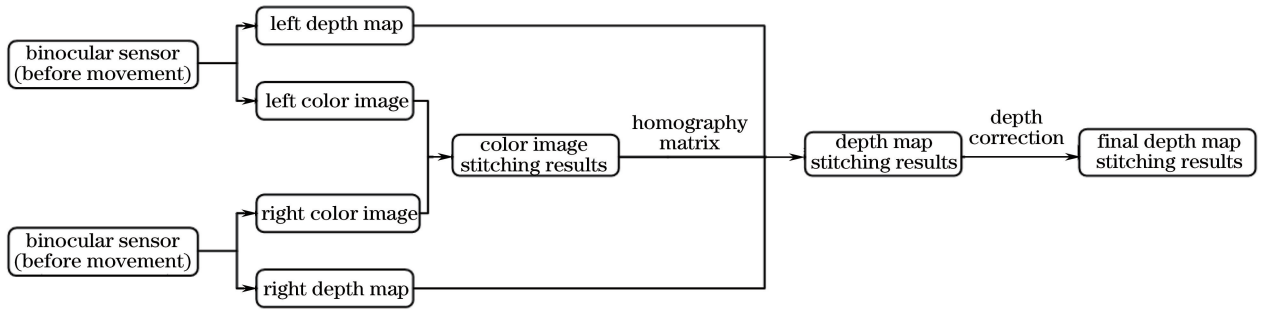


图4 深度图拼接算法的流程图

Fig. 4 Flow chart of depth map stitching algorithm

3 理论仿真与实验

3.1 理论仿真

用 Matlab 进行理论仿真,如图 5 所示,设基础面 A 为绿色部分,其四顶点坐标分别为 $(-30.1, -12, 0.9820)$ 、 $(-10.6, 7.5, 1.0054)$ 、 $(28.4, 9.45, 1.0171)$ 、 $(8.9, -10.05, 0.9937)$ 。对此面以任意方向进行透视变换,为了更直观地显示结果,取单应性矩阵 H_1 ,对 A 平面的三维坐标进行透视变换,即 $B = H_1 A$ 得到红色平面 B ,其对应的四顶点坐标分别为 $(-20, -20, 1)$ 、 $(-20, 19, 1)$ 、 $(19, 19, 1)$ 、 $(19, -20, 1)$ 。同理取单应性矩阵 H_2 ,进行透视变换,得黑色平面 C ,其顶点坐标分别为 $(-30.1, -12, 1)$ 、 $(-10.6, 7.5, 1)$ 、 $(28.4, 9.45, 1)$ 、 $(8.9, -10.05, 1)$ 。

考虑到 B 、 C 两平面共平面,取四对顶点为对应特征点,代入 x 、 y 坐标计算得到矩阵 H_{BC} ,通过 $B = H_1 A$ 、 $C = H_2 A$,计算得到矩阵 $H_{12} = H_2 H_1^{-1}$ 。对比最终结果,在误差允许的范围可得 $H_{12} = H_{BC}$ 。

$$H_1 = \begin{bmatrix} 1.0529 & -1.0517 & -0.9464 \\ -0.1058 & 2.1033 & 2.0927 \\ -0.0003 & -0.0009 & 0.9990 \end{bmatrix}, H_2 =$$

$$\begin{bmatrix} 1.0000 & -0.0000 & -0.0000 \\ 0.0000 & 1.0000 & -0.0000 \\ -0.0003 & -0.0009 & 0.9990 \end{bmatrix}, \text{计算得到}$$

$$H_{12} = \begin{bmatrix} 1.0000 & 0.5000 & -0.1000 \\ 0.0500 & 0.5000 & -1.0000 \\ -0.0000 & -0.0000 & 1.0000 \end{bmatrix},$$

$$H_{BC} = \begin{bmatrix} 1.0000 & 0.5000 & -0.10000038 \\ 0.0500 & 0.5000 & -1.0000001 \\ -0.0000 & -0.0000 & 1.0000 \end{bmatrix}。$$

仿真实验结果表明,通过图像拼接获得的单应性矩阵能够反映对应点空间位置间的关系,为深度

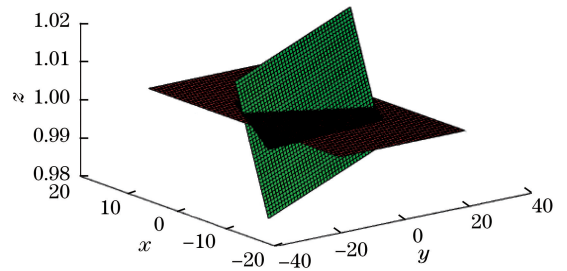


图5 仿真结果图

Fig. 5 Simulation results

矫正提供了理论基础。

3.2 双目摄像机的标定与畸变校正

以可变基线 USB3.0 双目视觉相机作为双目深度传感器,所用镜头为 2.8 mm 广角镜头,所用单目图像大小为 640 pixel \times 480 pixel。采用张正友标定法对双目视觉模块进行标定,所用棋盘格标定板的方格尺寸为 30 mm \times 30 mm。

3.3 双目深度图的获取

对双目传感器采集到的图像进行畸变校正、极限矫正、立体匹配后得到深度图。双目传感器移动

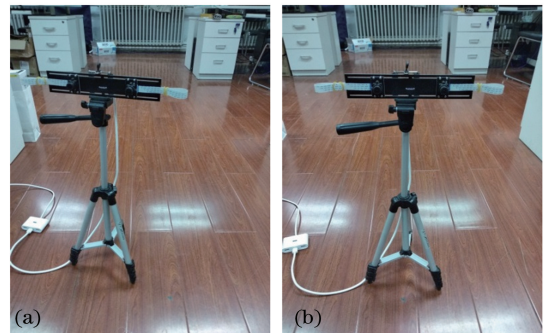


图6 双目测距模块移动前后的实物图。

(a) 移动后; (b) 移动前

Fig. 6 Real maps of binocular ranging module before and after moving. (a) After moving; (b) before moving

后与移动前的位置分别如图 6(a)、(b)所示。移动

前后采集到的左目彩色图和深度图如图 7 所示。

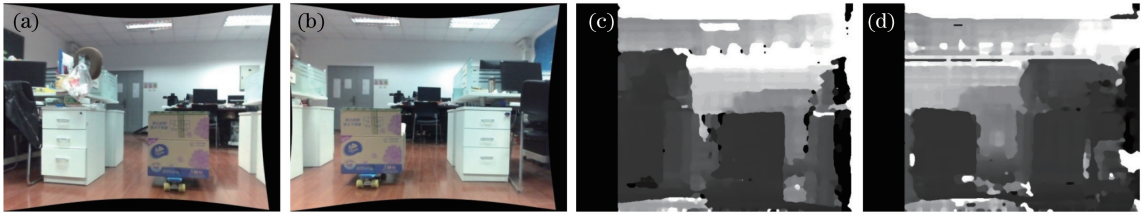


图 7 移动前后的彩色图及深度图。(a)移动前左目彩色图;(b)移动后左目彩色图;(c)移动前深度图;(d)移动后深度图

Fig. 7 Color images and depth maps before and after moving. (a) Left-view color image before moving;

(b) left-view color image after moving; (c) depth map before moving; (d) depth map after moving

3.4 双目深度图的拼接及深度矫正结果的评价

对彩色图像进行特征点提取与匹配,计算出单应性矩阵,根据此单应性矩阵对彩色图进行拼接,得

到彩色图的拼接结果,并以此单应性矩阵对深度图进行拼接与深度矫正,结果如 8 所示。



图 8 拼接结果。(a)彩色图;(b)深度图;(c)矫正后的深度图

Fig. 8 Stitching results. (a) Color image; (b) depth map; (c) depth map after correction

拼缝在 485 pixel 处,可以发现图 8(c)右半部分的深度(灰度)比图 8(b)的右半部分的更深,表明经矫正后位于后方的平面映射到前方标准平面上,结果符合实际情况。为了更直观地显示矫正结果,随

机抽取四个像素高度,取拼缝左右 70 pixel 点位置即像素宽度为[415,555]处的像素值绘制折线图,如图 9 所示。右深度图经矫正后深度值明显减小,符合从远处映射到近处的实际情况。为了进一步显示

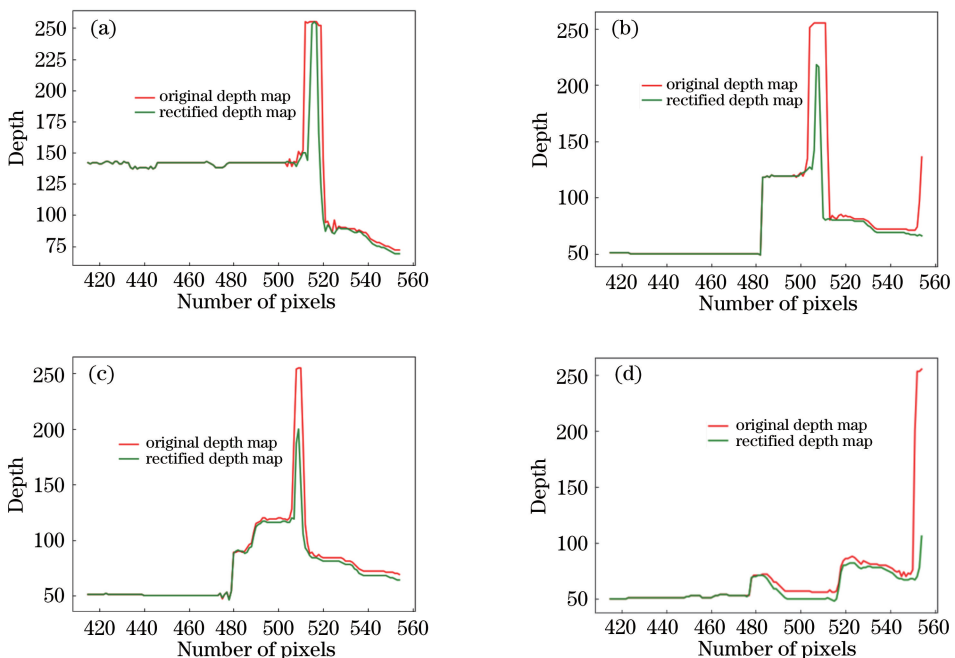


图 9 不同像素高度下的深度矫正折线图。(a) 220;(b) 280;(c) 360;(d) 420

Fig. 9 Depth correction images under different pixel heights. (a) 220; (b) 280; (c) 360; (d) 420

矫正效果,在图 8(b)、(c)中纸箱位置截取 60 pixel×60 pixel 大小的矫正后深度图与标准深度图以计算峰值信噪比 f_{PSNR} 。截取的深度图及对

应的 f_{PSNR} 见表 1。可以看出,经矫正后 f_{PSNR} 值由 50.0165 变为 51.4927,表明矫正方法效果良好,矫正后的结果与标准图基本一致。

表 1 矫正前后的 f_{PSNR} Table 1 f_{PSNR} before and after correction

Depth map	(Standard)	(Before correction)	(After correction)
f_{PSNR}		50.0165	51.4927

4 结 论

提出了一种基于双目视觉的深度图拼接方法,并进行了理论仿真与实验。所提方法具有普适性和实用价值,能有效地扩大双目深度传感器的视场,得到的拼接图像与单传感器图像基本相同。推导了投影变换前后深度值变换与单应性矩阵之间的关系,通过仿真和实验结果验证了所推导公式的正确性。

参 考 文 献

- [1] Wang S, Xu X. 3D reconstruction based on horopter[J]. Acta Optica Sinica, 2017, 37(5): 0515004.
王珊,徐晓.基于双目单视面的三维重建[J].光学学报,2017,37(5):0515004.
- [2] Li X M. Research and implementation of mosaic technology based on depth images[D]. Chengdu: University of Electronic Science and Technology of China, 2013.
李小妹.基于深度图像的拼接技术研究及实现[D].成都:电子科技大学,2013.
- [3] Zhang X, Jin Y X, Xue D. Image matching algorithm based on SICA-SIFT and particle swarm optimization [J]. Laser & Optoelectronics Progress, 2017, 54(9): 091002.
张鑫,靳雁霞,薛丹. SICA-SIFT 和粒子群优化的图像匹配算法[J].激光与光电子学进展,2017,54(9):091002.
- [4] Liu L, Kuang G Y. Overview of image textural feature extraction methods[J]. Journal of Image & Graphics, 2009, 14(4):622-635.
刘丽,匡纲要.图像纹理特征提取方法综述[J].中国图象图形学报,2009,14(4):622-635.
- [5] Rublee E, Rabaud V, Konolige K, et al. ORB: An

efficient alternative to SIFT or SURF[J]. Proceedings of IEEE International Conference on Computer Vision, 2011, 58(11): 2564-2571.

- [6] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [7] Cho S I, Kim J C, Ban K J, et al. 3D panorama generation using depth-map stitching[J]. Journal of Information & Communication Convergence Engineering, 2011, 9(6): 780-784.
- [8] Jiang Y T. Binocular ranging system and calibration method [D]. Changchun: Changchun University of Science and Technology, 2013.
姜雨彤.双目测距系统及标定方法研究[D].长春:长春理工大学,2013.
- [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] Zou P P, Zhang Z L, Wang P, et al. Binocular camera calibration based on collinear vector and plane homography[J]. Acta Optica Sinica, 2017, 37(11): 1115006.
邹朋朋,张滋黎,王平,等.基于共线向量与平面单应性的双目相机标定方法[J].光学学报,2017,37(11):1115006.
- [11] Bahraini M S, Bozorg M, Rad A B. SLAM in dynamic environments via ML-RANSAC [J]. Mechatronics, 2018, 49: 105-118.
- [12] Zhang C P, Wang Q. Survey on imaging model and calibration of light field camera[J]. Chinese Journal of Lasers, 2016, 43(6): 0609004.
张春萍,王庆.光场相机成像模型及参数标定方法综述[J].中国激光,2016,43(6):0609004.
- [13] Stockman G, Shapiro L G. Computer vision [M]. Upper Saddle River: Prentice Hall, 2001.