

基于双目技术的室内三维测量算法研究

胡锦涛, 张飞雄, 史泽坤, 黄浩

(湖北大学 物理与电子科学学院 铁电压电材料与器件湖北省重点实验室, 武汉 430062)

摘要: 研究了基于双目技术的室内三维测量方法, 获取了室内目标的三维信息, 用于室内三维测量、三维定位以及提取目标深度信息; 该方法基于双目视差三角测量原理, 利用张正友棋盘格标定、Bonguet 立体矫正算法、基于区域的立体匹配算法 BM/SGBM 等技术, 计算了室内目标的深度信息; 经过测量, 实验结果满足预期误差, 并从标定误差、实验环境误差和算法误差 3 个方面分析误差原因, 为以后提高测量精度、降低误差提供方向。

关键词: 双目测距; 摄像头标定; 立体矫正; 立体匹配; 深度信息; 误差分析

Research on Indoor Three Dimensional Measurement Algorithm Based on Binocular Technology

Hu Jinbo, Zhang Feixiong, Li Zekun, Huang Hao

(Hubei Key Laboratory of Ferro & Piezoelectric Materials and Devices, Faculty of Physics & Electronic Science, Hubei University, Wuhan, P. R. China, 430062)

Abstract: This paper studies the indoor three dimensional (3D) measurement method based on binocular technology, and acquires the 3d information of the indoor target, which is used for indoor 3D measurement, 3D positioning and extraction of target depth information. Based on the principle of binocular parallax triangulation, this method calculates the depth information of indoor targets by using Zhang Zhengyou's checkerboard calibration, Bonguet stereo rectification algorithm, region-based stereo matching algorithm BM (Block Matching) / SGBM (Semi-Global Block Matching) and other technologies. After measurement, the experimental results meet the expected error. The error causes are analyzed from calibration error, experimental environment error and algorithm error, which provide a direction for improving the measurement accuracy and reducing the error in the future.

Keywords: binocular distance measurement; camera calibration; stereo rectification; stereo matching; depth information; error analysis

0 引言

双目技术作为计算机视觉的一个重要研究分支, 广受人们关注。它在机器人技术、无人驾驶汽车技术等领域有着广泛的前景。双目技术利用双目摄像头感知三维场景的深度信息, 为机器人视觉、三维建模等提供有效依据^[1]。测距方式有很多, 包括量尺测距、电磁波测距、视差法测距等。量尺测距是指用量尺直接测量两点距离, 精度高, 成本低, 但仅适用于短距离测量, 无法做到实时性; 电磁波测距是基于发送的电磁波与接收的电磁波之间的时间差或相位差而计算两点距离, 如红外测距、激光测距, 测距精度高, 实时性强, 但设备复杂、成本也高^[2]。

室内空间有隔离性、封闭性、贴近性、有限性等特点。室内环境被限定性较高的围护实体包围起来, 对光线、声音等有很强的阻隔作用。因此, 传统的测距方法, 在精度、成本、适用环境等方面考虑, 不适用于室内三维测量。利

用视差法测距的双目技术测量方法, 具有高效率、高精度、低成本、简单的系统结构、易于集成在机器人、汽车上等优点, 可以实现室内三维测量。

双目技术是基于类似于人眼的视差视觉原理。它是由双目摄像头在不同位置同时获取目标点的左右视图, 类似于人类或者动物的双眼功能, 并利用双目视差相似三角形测量原理获取目标的三维坐标信息, 由此得到目标点的深度信息, 即测量目标点与双目摄像头之间的距离。

1 系统设计

1.1 双目摄像头的测距原理及系统整体结构

双目摄像头的测距原理是基于人类双目视差的原理, 利用双目视差三角测量原理获取物体的深度信息。双目摄像头测距原理如图 1 所示, 图中点 P 为目标点, O_{left} 为左摄像头的光心, O_{right} 为右摄像头的光心, f 为摄像头焦距 (Focal Length), b 为基线 (Baseline) 长度, d 为视差 (Disparity):

$$d = x_{left} + x_{right} \quad (1)$$

根据相似三角形原理, 即可求得深度信息 z :

$$\frac{z-f}{z} = \frac{b-d}{b} \quad (2)$$

化简得:

$$z = \frac{fb}{d} \quad (3)$$

收稿日期: 2019-06-18; 修回日期: 2019-07-19。

基金项目: 湖北省自然科学基金指导性计划项目(2018CFC797)。

作者简介: 胡锦涛(1998-), 男, 湖北大冶人, 主要从事计算机视觉方向的研究。

通讯作者: 黄浩(1988-), 男, 湖北天门人, 博士, 硕士生导师, 副教授, 主要从事人工智能图像处理方向的研究。

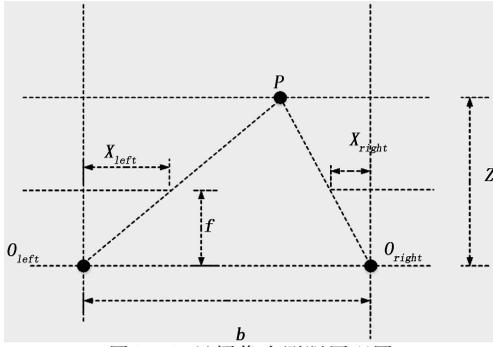


图 1 双目摄像头测距原理图

在双目摄像头测距系统中, 摄像头焦距 f 和基线长度 b 由设备固有参数、自己的需求确定, 为已知量。故在该系统中, 其核心是求解视差 d , 由此可以得到深度信息 z 。

双目摄像头测距系统整体结构如图 2 所示。利用张氏棋盘格相机标定法, 完成双目摄像头标定, 得到摄像头内参数 (焦距、畸变系数、成像原点) 与外参数 (平移向量、旋转矩阵)。待测目标的双目图片可以利用这些参数进行双目立体校准, 使得左右两张视图消除畸变以及行对准。进行立体双目匹配, 将同一物体的左右视图的对应像素点匹配起来, 可以得到视差图, 再利用上述原理公式可计算出深度信息。

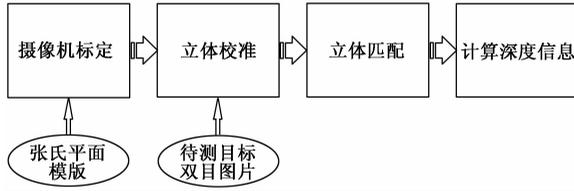


图 2 双目测距系统整体结构

1.2 摄像机标定

摄像机标定是为了获取摄像机的固有参数: 内参数与外参数, 这些参数只与设备本身信息有关。一旦双目摄像头确定, 内参数与外参数就为定值, 这些参数在接下来的实验测试中可以一直使用。内、外参数决定了摄像机光屏的二维图像坐标系与世界三维坐标系之间的映射关系。内参数是一种从平面到像素的变换, 仅决定于摄像机本身的物理特性。外参数反映的是摄像机坐标系与世界坐标系的变换, 决定于内参数与基线长度。

理论上, 任意具有鲜明特征的目标都能成为待标定的物体, 但通常都是以棋盘为标定物体^[3]。棋盘标定法是基于张氏棋盘格相机标定法。本实验采用的张氏棋盘格相机标定法中, 用于标定的模板是角点数为 5×4 、棋盘格边长为 $37\text{mm} \times 37\text{mm}$ 矩形的黑白棋盘。

本实验所选择的图像二维坐标系为世界三维坐标系 $Z=0$ 时的平面平行。设 $P = (x, y, z)^T$ 为三维世界坐标系中标定的棋盘角点坐标, 在相机模型中, 变换为二维摄像机的图像坐标系上的像点 $p = (u, v)^T$ 。它们的增广向量 $\tilde{P} = (x, y, z, 1)^T$, $\tilde{p} = (u, v, 1)$ 它们之间的关系为^[4]:

$$\tilde{s}p = K[R \ T]\tilde{P} \quad (4)$$

其中:

$$K = \begin{pmatrix} \alpha & 0 & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (5)$$

s 是比例系数, 矩阵 K 是摄像机的内参数矩阵, 其中 α , β 为焦距信息, (u_0, v_0) 为坐标主点 (光心), $[R, T]$ 为外参数矩阵, R, T 分别为摄像机平面与物体平面的旋转矩阵与平移矩阵。

摄像头由于透镜的光学特性使得成像存在径向畸变; 两个摄像头由于装配方面存在误差, 不可能绝对平行放置, 故成像存在切向畸变。因此, 需要求解畸变参数对图片进行矫正。在此假设 $q_d(x_d, y_d)$ 为摄像头获取的点, 经过去畸变、矫正后得点 $q_r(x_r, y_r)$, 其关系为:

$$\begin{pmatrix} x_r \\ y_r \end{pmatrix} = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \begin{pmatrix} x_d \\ y_d \end{pmatrix} \begin{pmatrix} 2p_1 x_d y_d + p_2 (r^2 + 2x_d^2) \\ 2p_2 x_d y_d + p_1 (r^2 + 2x_d^2) \end{pmatrix} \quad (6)$$

其中: $D = (k_1, k_2, p_1, p_2, k_3)$ 构成 1×5 的矩阵, 为摄像头的畸变参数矩阵^[5]。其中 k_1, k_2, k_3 确定径向畸变, p_1, p_2 确定切向畸变。

张氏棋盘格相机标定法常用到的工具有 OpenCV 视觉类库、halcon 机器视觉软件以及 Matlab 的 TOOLBOX_calib 或者 Stereo Camera Calibrator 等^[6]。Matlab 的 TOOLBOX_calib 有精度较高、操作方法简单、三维可视化等特点, 因此本文采用 Matlab 的 TOOLBOX_calib 对左右相机进行标定, 标定图片如图 3、图 4 所示, 标定结果如下:

左摄像头:

$$\text{焦距: } [\alpha_l, \beta_l] = [832.19568 \quad 838.64596]$$

$$\text{坐标主点: } [u_{ol}, v_{ol}] = [313.50676 \quad 190.52907]$$

右摄像头:

$$\text{焦距: } [\alpha_r, \beta_r] = [814.96533 \quad 820.46468]$$

$$\text{坐标主点: } [u_{or}, v_{or}] = [348.55656 \quad 190.62929]$$

$$\text{Distortion}_l = (0.00069 \quad 0.20770 \quad -0.00251 \quad 0.01163 \quad 0.00000)$$

$$\text{Distortion}_r = (0.04731 \quad -0.67737 \quad -0.00447 \quad 0.01159 \quad 0.00000)$$

$$OM = (0.01420 \quad 0.01002 \quad -0.00357)$$

$$T = (-65.73752 \quad -2.87428 \quad -12.44283)$$

OM 为旋转关系向量 (对其作 Rodrigues 变换, 变换为 R 旋转关系矩阵), T 为平移向量。

1.3 立体矫正

计算出双目摄像头的内、外参数及畸变参数后, 便可以实现世界三维坐标系与图像二维坐标系之间的转换。但为了求解视差 d , 还需要进行立体矫正。求解视差的过程是将左右视图对应点相互匹配的过程。但是在二维空间中搜索相匹配的点, 将目标点进行严格对应是非常困难的, 内存资源占用、时间耗费、计算复杂度都非常大, 并且由于噪声、曝光、聚焦等因素影响, 匹配结果会出现误匹配现象。因此, 我们可以将二维空间中的搜索匹配降至一维空间, 由此减少搜索匹配范围, 增加匹配准确性。这也是立体矫正的目的所在。

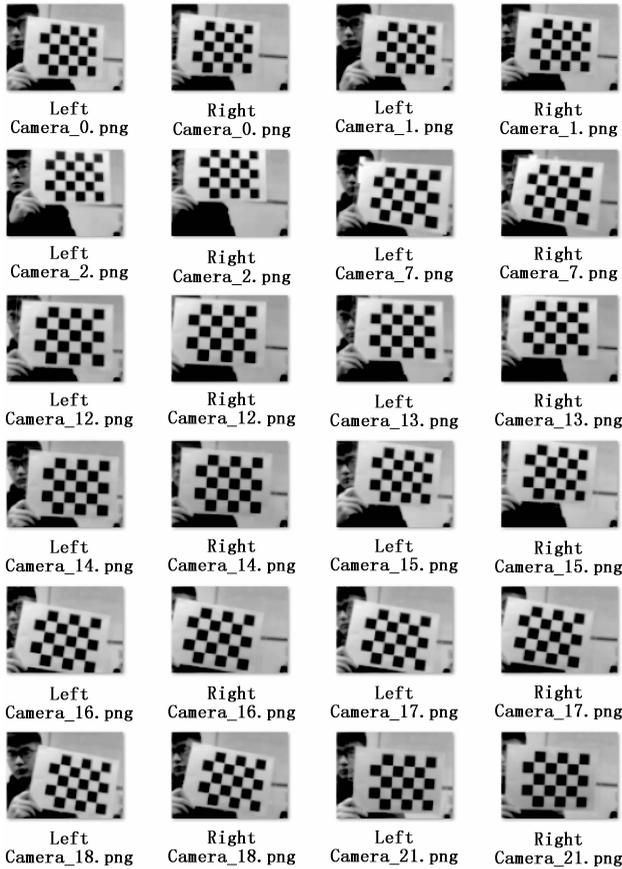


图 3 双目摄像头标定图片 (棋盘格平移)

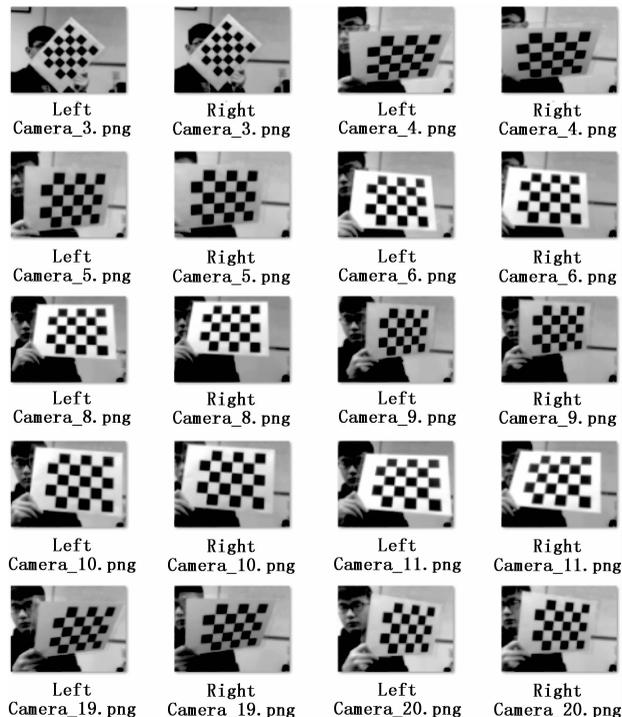


图 4 双目摄像头标定图片 (棋盘格偏转)

立体矫正是利用标定的参数将用双目摄像头获取的图像去畸变处理, 再将消除畸变后的图像进行严格地行对应, 即左右视图对极线处在同一水平线上, 左视图中的任意一

个目标点与右视图中的对应目标点必有相同的行号。在立体匹配的过程中, 搜索匹配同一行号与之相同特征的目标点即可, 这将有效降低计算复杂度, 提高匹配准确性^[7]。

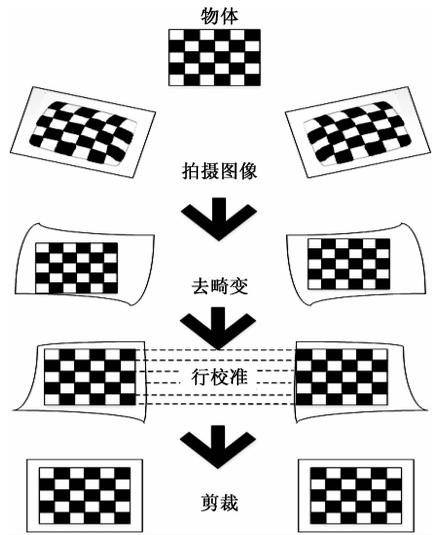


图 5 立体矫正流程图

Bonguet 算法利用双目摄像头的旋转矩阵 R 和平移矩阵 T , 使左右视图中的每幅重投影次数最小、重投影畸变最小、重叠观测面积最大, 立体矫正流程图如上图 5 所示。

Matlab 中的 TOOLBOX_calib 工具箱和 OpenCV 视觉类库都提供了双目矫正的功能。在 OpenCV 中可通过 stereoRectify 函数完成左右视图矫正。该函数通过双目摄像头的内参数矩阵、外参数矩阵、畸变向量, 旋转矩阵和平移矩阵得到左右相机平面间的行对校正旋转矩阵, 左右相机的投影矩阵, 和重投影矩阵。再调用函数 initUndistortRectifyMap 生成图像矫正所需要的映射矩阵。最后调用 remap 函数得到矫正后的图像。矫正前后左右视图如图 6~7 所示。

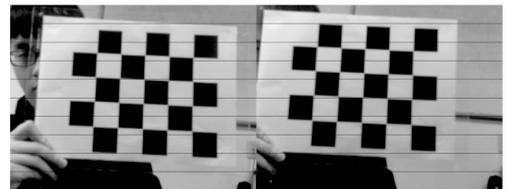


图 6 矫正前左右视图图片 (未行对准)

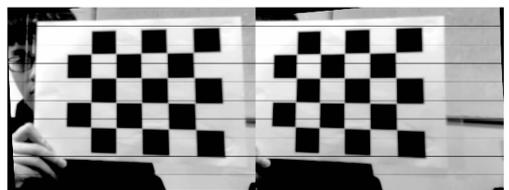


图 7 矫正后左右视图图片 (已行对准)

1.4 立体匹配与计算深度

双目匹配的作用是把同一场景左右视图对应的像点匹配起来, 从而得到视差信息。立体匹配算法依据匹配单元的不同, 可分为基于区域、特征和相位的匹配算法^[8]。基于区域的立体匹配算法通过构建支持窗口的方式估计视差值, 适用于高纹理图像, 采集图像的视点差距不应过大; 基于特征的立体匹配算法顾名思义通过特征匹配估计视差

值, 适用于特征丰富的图像; 基于相位的立体匹配算法依据傅里叶变换的平移不变性, 使得不同频率范围的相位相应做匹配单元, 适用在相位奇点较少、局部结构相同的图像^[9]。根据视差获取方式不同, 又可分为全局立体匹配算法、半全局立体匹配算法和局部立体匹配算法。

基于区域的匹配算法相对于其他两种算法而言能够保证视差图的密集性, 对图像无需预处理, 有实现难度容易、实时性强、易于集成在嵌入式系统等优点, 本实验采用基于区域的匹配算法。

BM (Block Matching) 算法实现快速有效的局部块匹配算法。BM 算法使用小窗口 SAD (绝对误差累计) 查找左右校准图像之间的匹配点。SAD 公式如下:

$$C_{SAD}(x, y, d) = \sum_{x, y \in B} |p_{left}(x, y) - p_{right}(x + r, y)| \quad (7)$$

其中: p_{left} 、 p_{right} 表示匹配窗口灰度均值, B 为匹配区域, r 为窗口偏移量。该算法在强纹理环境下, 能够得到很好的效果, 在弱纹理环境下匹配效果一般, 并且对光照强度、对比度等因素造成的灰度值变化敏感, 但并不影响总体匹配效果。OpenCV 提供了 StereoBM_create 函数实现有效快速的块匹配算法。该算法的两个参数为视差搜索范围 numDisparities 与匹配窗口大小 blockSize。通过规定视差的搜索范围构成了一个双目世界, 寻找合适的匹配窗口大小, 其具体的数值可以根据实验环境和工具参数进行设置。

局部算法是基于块区域代价的累计与优化, 但在深度变化剧烈的区域得到的结果不理想。在噪声等因素的影响下存在匹配模糊现象, 即匹配错误的点比匹配正确的点有更低的代价。因此, 需要增加约束项来约束相邻像素点灰度值的改变^[10]。半全局匹配基于此提出了如下能量函数:

$$E(D) = \sum_p (C(p, D_p) + \sum_{q \in N_p} P_1 T[|D_p - D_q| = 1]) + \sum_{q \in N_p} P_2 T[|D_p - D_q| > 1] \quad (8)$$

半全局块匹配算法 SGBM (Semi-Global Block Matching) 是对 BM 算法的改进。相比较而言, SGBM 提高了匹配准确性, 但同时增加了耗时。在 OpenCV 中设置了 BM/SGBM 的参数后, 使用视觉库中的 stereo_compute 和 normalize 函数可以得到视差图。知道图像点 (x, y) 和视差值 d , 可以使用如下的公式计算物体深度:

$$Q(x, y, d, 1)^T = (X, Y, Z, W)^T \quad (9)$$

由上式可以得到校准映射后的三维坐标 $(\frac{X}{W}, \frac{Y}{W}, \frac{Z}{W})$ 。物体深度可以由函数 reprojectImageTo3D 实现, 函数所得结果的 z 方向的值即为深度信息。如图 8 为左摄像头拍摄图片, 如图 9 为深度图, 像素灰度值体现了深度信息, 但由于算法的局限性, 弱纹理区域无法计算深度。

2 实验结果与误差分析

2.1 双目测距实验结果

双目测距的测试环境是在实验室内, 用鼠标点击如上图 9 所示图中目标点即可获得深度信息, 即目标点与双目摄像机的直线距离。



图 8 左摄像头图片

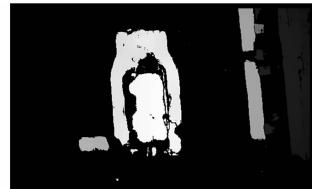


图 9 深度图

使用软尺进行实际测量, 将实验目标摆放在待测实验点。为了计算简便以及结果的准确性, 实际测量距离为整数。用双目摄像机进行拍照得到左右视图, 导入计算机内, 使用之前已经标定好的数据, 对目标点进行双目测距。使用 BM 算法测量结果如表 1 所示。使用 SGBM 算法测量结果如表 2 所示。

表 1 BM 算法实验结果

编号	软尺测距/mm	双目测距/mm	绝对误差/mm	相对误差/%
1	300	315.67	15.67	5.22
2	600	620.94	20.94	3.49
3	900	930.74	30.74	3.41
4	1 200	1 218.34	18.34	1.53
5	1 500	1 521.85	21.85	1.46
6	1 800	1 830.83	30.83	1.71
7	2 300	2 331.87	31.87	1.39

如表 2 所示 SGBM 算法的误差比 BM 算法的误差小, 准确性更高。并且软尺测量的真实距离越大, 相对误差越小。考虑到标定误差、算法误差与实际操作误差, 误差结果在可接受范围内。

2.2 实验误差分析

综合设备环境、实验环境、算法鲁棒性等客观因素考

表 2 SGBM 算法实验结果

编号	软尺测距/mm	双目测距/mm	绝对误差/mm	相对误差/%
1	300	315.21	15.21	5.07
2	600	617.25	17.25	2.88
3	900	910.69	10.69	1.78
4	1 200	1 212.60	12.60	1.05
5	1 500	1 529.84	29.84	0.97
6	1 800	1 828.89	28.89	0.95
7	2 300	2 320.73	20.73	0.87

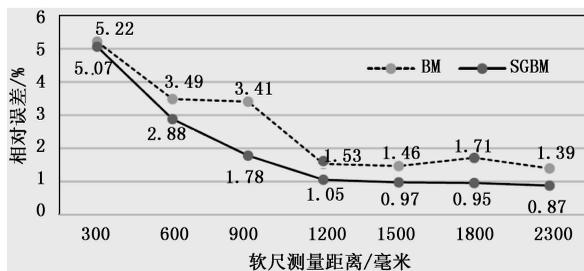


图 10 双目测距相对误差

虑, 实验存在以下误差因素: 双目标定误差, 实验环境误差以及算法误差等。

2.2.1 双目标定误差

双目标定在整个系统里有着至关重要的作用, 后续所

有步骤都是基于双目标定数据的前提下完成的。双目标定误差直接和间接影响着双目测距的结果。

图像采集。首先要保证标定棋盘的左右视图图像清晰,其次采集图像时保证棋盘标定板与摄像机平面有角度变化,最重要的是棋盘标定板需要遍历整个图像区域,在视场的不同范围都要有标定图像。标定图像的多少对精度也会造成影响,尽量多的图像能避免标定数据对噪声的敏感性,本文采集了 21 对图像。每次由于图像的不同,标定的数据也不同。

角点提取。角点提取直接影响标定结果。使用 Matlab 中的 TOOLBOX_calib 标定时,必须手动提取角点,由于人的主观因素影响,角点提取无法恰到好处。例如使用 Matlab 中的 Stereo Camera Calibrator 自动提取角点,标定结果如下:

$$\mathbf{K}_{left} = \begin{bmatrix} 812.3120 & 0 & 288.7609 \\ 0 & 818.8419 & 204.9309 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{K}_{right} = \begin{bmatrix} 814.0528 & 0 & 322.6428 \\ 0 & 820.7897 & 1883.4494 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{D}_{left} = (0.0643 \quad -0.1610 \quad 0.0000 \quad 0.0000 \quad 0.0000)$$

$$\mathbf{D}_{right} = (0.0471 \quad -0.2346 \quad 0.0000 \quad 0.0000 \quad 0.0000)$$

$$\mathbf{T} = (-65.73752 \quad -2.87428 \quad -12.44283)$$

$$\mathbf{R} = \begin{bmatrix} 0.9999 & 0.0033 & 0.0124 \\ -0.0032 & 1.0000 & 0.0079 \\ -0.0125 & -0.0079 & 0.9999 \end{bmatrix}$$

如上式,由于标定棋盘采集角度变化和数量、覆盖相片比例,Matlab 手动角点提取的不确定性,与第 2 节标定结果对比有一定偏差。

2.2.2 实验环境误差

由表 1、表 2 观察可知,实际距离越短,相对误差反而挺大的,绝对误差在 15~32 mm 之中,推测有摄像头摆放角度、摄像头物理结构、噪声等因素影响。

摄像头所在平面与物体所在的平面存在一定角度,距离较短时误差不大,但当距离较长时例如 2 000 mm,与物体夹角为 5°,便会有 7.6 mm 的误差,对实验结果影响较大。

市面上普通的摄像头镜头长度大概约为 2~3 cm,镜头里包含凸透镜、图像传感器等器件,考虑到凸透镜厚度、凸透镜焦距(本实验选定的摄像头焦距为 4 mm)等因素,预估成像点与真实成像点存在毫米级别的误差,这会影响到测量结果。短距离测距时对该误差尤为敏感。

外部电磁波干扰、摄像头内部电路光和电的基本性质等引起的噪声同样会产生误差。

2.2.3 算法误差

本文采用的是基于区域的立体匹配算法,但单纯的区域匹配都有一定限制^[11]。

如图 8、图 9 所示桌上的水杯、旁边的空调有深度信息,而蓝色的桌子、白色的背景没有深度信息,这是由于

BM 算法只查找两幅图像之间的强纹理信息,在背景单一的室内只能得到为数不多的点的深度信息。

在选择好的区域块内,所有的像素都被认为具有相同的视差值,所以该算法不适用于当区域块过大时或者深度变化剧烈的场景。

对光照、对比度和噪声敏感。同一场景,在不同的光照、对比度等因素情况下,得出的结果可能会有很大的偏差。

视差搜索范围(numDisparities)与区域块(blockSize)的大小难以选择。由于 BM/SGBM 算法的特殊性,在不同场景下,所设定的参数不同。匹配窗口大小 blockSize 越大,计算复杂度越小,但是匹配的准确性越低;匹配窗口大小 blockSize 越大,计算复杂度越高,但是同样极易受到噪声影响使之误匹配。

3 结束语

本文研究了双目三维测量技术在室内环境的应用,解决了在室内环境下的目标的三维测量、深度信息提取等问题。从双目视差三角测量原理出发,研究了摄像头立体标定、立体校准、立体匹配等关键技术,计算室内目标深度信息。通过实验对比基于区域的匹配算法 BM 和半全局块匹配算法 SGBM 可知,两种算法的相对误差随着实际距离的增加而减少,且 SGBM 的精度更高。在此基础上又分析了实验误差,从双目标定误差、实验环境误差和算法误差 3 个方面进行分析,研究了引起误差造成的原因,在以后的研究中可以在此方面进行改进。

参考文献:

- [1] 徐杰,陈一民,史志龙. 双目视觉变焦测距技术[J]. 上海大学学报(自然科学版),2009,15(2):169-174.
- [2] 沈彤,刘文波,王京. 基于双目立体视觉的目标测距系统[J]. 电子测量技术,2015,38(4):52-54.
- [3] 王怡萱. 基于双目立体视觉的立木测距系统的研究[D]. 北京:北京林业大学,2013.
- [4] Zhang Z. A Flexible New Technique for Camera Calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2000,22(11):1330-1334.
- [5] 王浩,许志闻,谢坤,等. 基于 OpenCV 的双目测距系统[J]. 吉林大学学报(信息科学版),2014,32(2):188-194.
- [6] 靳盼盼. 双目立体视觉测距技术研究[D]. 西安:长安大学,2014.
- [7] 蒋国良. 结合待测物体强化的双目测距系统设计与实现[D]. 西安:西安电子科技大学,2015.
- [8] 刘欢. 基于双目视觉立体匹配算法的研究与应用[D]. 哈尔滨:哈尔滨工业大学,2018.
- [9] 荆鑫. 计算机视觉中双目立体匹配技术的研究[D]. 南京:南京理工大学,2013.
- [10] 石立. 双目视觉中的半全局立体匹配算法研究[D]. 南京:南京大学,2018.
- [11] 王逸林. 基于区域匹配算法的双目立体匹配技术的研究[D]. 大连:大连海事大学,2012.