

doi: 10.7690/bgzdh.2020.03.016

基于立体视觉的 3 维模型重建

杨 靖, 王茂森, 戴劲松

(南京理工大学机械工程学院, 南京 210094)

摘要: 为实现智能机器人清晰把握周围环境的 3 维信息, 提出一种基于坐标变换的 3 维模型重建方式。采用双目立体视觉结构, 经过图像获取、摄相机标定-校正、立体匹配和 3 维特征提取等操作, 获取单一视图下物体的 3 维几何信息。将多个不同坐标系下的 3 维几何信息归一化为同一坐标系下的信息, 并选取具有代表性的长方体纸箱进行实验验证。结果表明: 该方法能实现 3 维模型的重建工作, 对环境几何信息的理解和避障导航有一定的参考价值。

关键词: 数字图像处理; 3 维重建; 相机标定; 坐标变换

中图分类号: TP391.41 **文献标志码:** A

3D Model Reconstruction Based on Stereo Vision

Yang Jing, Wang Maosen, Dai Jinsong

(School of Mechanical Engineering, Nanjing University of Science & Technology, Nanjing 210094, China)

Abstract: In order to realize the intelligent robot to clearly grasp the 3D information of the surrounding environment, a 3D model reconstruction method based on coordinate transformation is proposed. The binocular stereo vision structure is used to obtain the 3 dimensional geometric information of the object in a single view through image acquisition, camera calibration-correction, stereo matching and 3 dimensional feature extraction. Normalize 3D geometric information in multiple coordinate systems to information in the same coordinate system, and a representative rectangular carton is selected for experimental verification. The results show that the method can realize the reconstruction of 3D model, and it also has certain reference value for the understanding of environmental geometric information and obstacle avoidance navigation.

Keywords: digital image processing; 3 dimensional reconstruction; camera calibration; coordinate transformation

0 引言

近年来, 随着计算机的发展、离散数学理论的创立和完善、社会各行业需求的增长, 数字图像处理技术获得了长足的进步, 在航天航空、医学、机器人视觉等领域都发挥着越来越强大的作用。随着计算机视觉技术的不断发展, 3 维重建图像处理技术也成了研究的一大热点^[1]。当前获取物体 3 维信息的方法分为接触式和非接触式。接触式方法主要利用一定的仪器在真实的场景下采集 3 维数据, 但是受到时间、空间、人为因素的影响, 并不适合广泛使用。非接触式测量分为主动和被动 2 类: 主动视觉利用光学原理对物体进行扫描得到数据点云, 实现 3 维重建, 能够得到比较精确的模型, 但是操作不够方便, 费用成本也比较高; 被动视觉通过分析照片或者视频图像中的信息, 从而对实际物体建立模型, 具有操作简单、成本低、使用范围广的特点^[2]。笔者介绍的方法就属于非接触式测量中的被

动视觉。

该方法利用双目相机装置, 获取视场内物体在 2 个相机内的成像, 并分析同一个特征点在不同相机内的成像坐标差异, 利用相似三角形的原理, 得出视场内物体的深度信息。接着对同一物体进行多次测量, 利用坐标转换的方法, 得出物体的 3 维空间模型。该方法具有适应性强、处理迅速、系统资源占用小等优点, 能很好地满足智能机器人在复杂环境下的工作要求。

1 相机标定模块设计

1.1 单目相机标定

图 1 为摄像机成像模型中最简单的针孔模型, f 为相机焦距, 点 $Q(X, Y, Z)$ 由通过投影中心的光线投射到投影平面上, 相应的图像点为 $q(x, y, f)$ 。根据三角形相似的关系有:

$$\frac{x}{f} = \frac{X}{Z} \quad (1)$$

收稿日期: 2019-12-08; 修回日期: 2019-12-29

作者简介: 杨 靖(1995—), 男, 江苏人, 硕士, 从事机械电子工程专业研究。E-mail: q8725@qq.com。

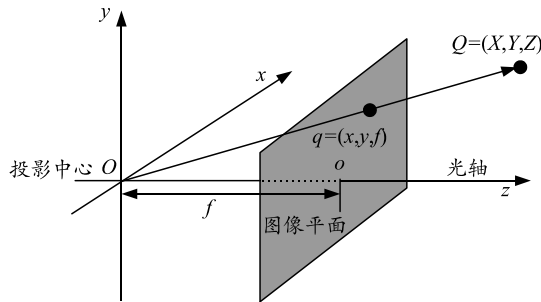


图 1 针孔摄像机模型

相机的标定有 2 方面的作用：1) 为了矫正透镜给针孔模型带来的偏差；2) 将相机和真实世界相互联系起来，使场景不仅是 3 维的，也是物理单位度量的空间。单目相机的标定可以给出相机的几何模型和透镜的畸变模型，利用这 2 个模型可以矫正畸变，建立相机和真实世界的联系^[3]。此外标定结果的好坏直接影响后续图像处理的精度，从而影响最终结果。

假定空间一点 $Q(X,Y,Z)$ 映射到投影平面上点 $q(x,y)$ ，将该变换用齐次坐标的形式表述如下：

$$q = MQ,$$

其中

$$q = \begin{bmatrix} x \\ y \\ w \end{bmatrix}, M = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, Q = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (2)$$

换算矩阵 M 称为摄像机的内参数矩阵。此外还需要确定摄像机的畸变参数，畸变分为径向畸变和切向畸变如图 2 所示。

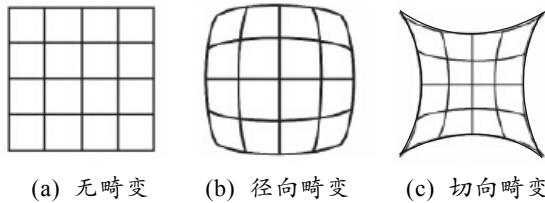


图 2 摄像机畸变

径向畸变是由于透镜的形状，切向畸变是由于摄像机安装过程中透镜不完全平行于图像平面，摄像机的这 2 种畸变可以采用下式进行矫正：

$$\begin{bmatrix} x_{corrected} \\ y_{corrected} \end{bmatrix} = \left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6\right) \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 2p_1 y + p_2 (r^2 + 2x^2) \\ p_1 (r^2 + 2y^2) + 2p_2 x \end{bmatrix}. \quad (3)$$

其中： $\begin{bmatrix} x \\ y \end{bmatrix}$ 为原始点； $\begin{bmatrix} x_{corrected} \\ y_{corrected} \end{bmatrix}$ 为矫正后的坐标点^[4]。

在进行正式标定前，先要采集多组标定棋盘板

的图片。由于环境的光照条件及图片拍摄过程中一些偶然因素及人为因素等，需要预先剔除一些图片，否则会导致标定误差过大，标定结果无效。Matlab 具有强大的图像处理功能，利用 Matlab 中 Stereo Camera Calibrator 工具箱可以分别读入拍摄好的图片，并设置棋盘格的大小。如图 3 所示，标定工具箱会自动剔除一些质量较差的图片，一共 37 组图片，接受 22 组，剔除了 15 组。

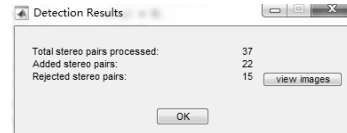


图 3 标定工具箱预选择

如图 4 所示，经过标定后，工具箱会给出每组图片的误差值，可以据此手动剔除误差较大的图片，如最后一幅。

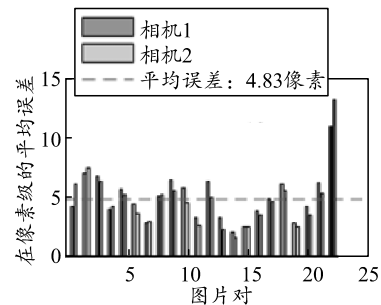


图 4 标定误差

相机采集到的图像是彩色三通道图像，即 B、G、R 三通道，每个通道都对应一组数据，存放在 3 维数组 Mat 中。标定中目标棋盘是黑白色，只需要 1 维通道，存放在 1 维数组中，数据量为 1/3，可以节约系统资源，提高运算速度。利用 OpenCV 自带 cvCvtColor 函数，并设置参数为 CV_BGR2GRAY 即可实现，该函数基于以下公式：

$$Gray = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B. \quad (4)$$

但是该公式使用浮点数进行计算会降低转换速度，可以改为整数运算形式：

$$Gray = (2989 \times R + 5870 \times G + 1140 \times B) / 10000. \quad (5)$$

或者使用运行速度更快的移位操作：

$$Gray = (4898 \times R + 9618 \times G + 1868 \times B) \gg 14. \quad (6)$$

采用改进后的灰度化处理方法可以加快标定阶段的运行速度。主程序中在处理多帧图像信息时也可以使用这种方法，可提高系统运行的速度^[5]。

1.2 双目相机标定

如图 5，立体成像的基本几何学是对极几何，由 2 个针孔摄像机模型和极点组成。 O_l 和 O_r 为 2

个摄像机模型的投影中心。3 维空间中的点 P 在每个投影平面都有一个投影点，记为 P_l 和 P_r 。左(右)相机的成像点 $O_l(O_r)$ 在右(左)透视平面上的成像点 $e_l(e_r)$ 被称为极点。由 P 点和 2 个投影中心 O_l 和 O_r 确定的平面被称为极面，投影点和极点之间的连线称为极线。空间中一点 P 投影到左相机成像平面时， P 点会落在 O_l 指向 P_l 这条射线上，由于单个相机无法知道 P 点到成像平面的距离，故 P 点在直线上的具体位置未知，但能确定该直线被投影到右相机投影平面上为 P_r 和 e_r 确定的极线。

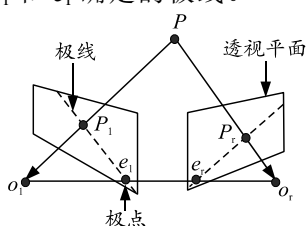


图 5 对极几何模型

综上所述，已知某一图像中的一点，它在另一幅图像上相对应的点一定在对应的极线上。当已知一组相机的对极几何后，寻找 2 幅图像的匹配特征时，可以由 2 维搜索简化到 1 维搜索，这样不仅可以大量减少匹配时间，而且能提高匹配的准确性。

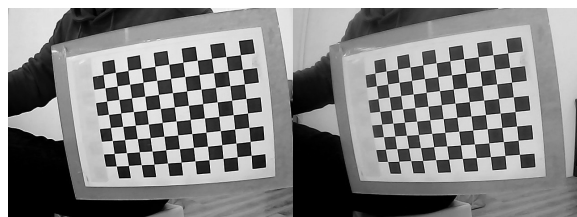
图 6 为双目相机的实物图。双目相机的标定主要作用是确定 2 台相机在空间的位置几何关系。以左相机的投影中心作为摄像机坐标系的原点，通过双目相机标定可以求解出右相机相对于左相机的平移和旋转矩阵 R , T 。此外还要确定本征矩阵 E 和基础矩阵 F ，矩阵 E 包含 2 个相机在空间上的旋转和平移信息，不包含相机的内参数，因此，矩阵 E 利用对极几何的原理将左右物理坐标联系起来。矩阵 F 包含了 2 个相机的内参数还有矩阵 E 的信息，处理的是像素坐标，将左右像素坐标联系起来^[6]。



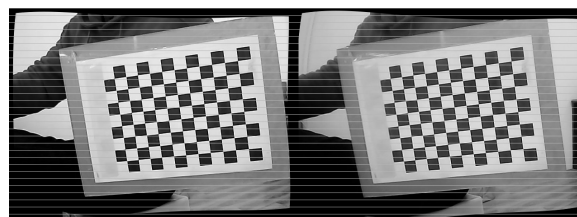
图 6 双目相机实物

图 7(a)为经筛选的双目摄像机标定图片，经过灰度化等预处理和标定，将标定结果作用在原始图像上，进行畸变矫正以及立体匹配，如图 7(b)。从图中可以看到：经过标定，一方面原始图片上弯曲的棋盘边界线被校正成直线，消除了相机畸变的影响；另一方面从图中所做水平辅助线可以看出，左

右相机内成像的对应点都被匹配到同一水平线上，减少了搜索时的运算量。



(a) 原始图



(b) 匹配图

图 7 双目摄像机标定匹配

2 视差及深度信息的计算

经过标定、校正及匹配后的双目摄像机就可以利用三角测量的原理对空间物体进行测距。图 8 为双目摄像机三角测量的示意图，世界中的 P 点在左右图像上成像点为 p_l 和 p_r ，相应的横坐标为 x^l 和 x^r 。定义视差为 $d=x^l-x^r$ ，根据相似三角形的原理可以推导出 Z 值^[7]，其中 f 为相机的焦距，由标定结果中的内参矩阵 M 给出， T 为左右相机成像点的间距，由标定结果中的 T 给出：

$$Z = fT / (x^l - x^r) \tag{7}$$

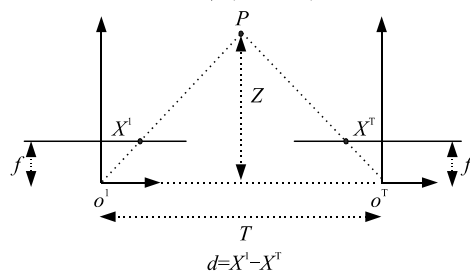
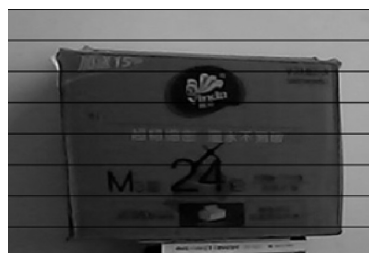


图 8 双目摄像机三角测量

从图 9 可以看到： Z 轴深度值在 1 000 mm，与实验中的实际值基本吻合，在俯视图发现，基本提取出物体轮廓。



(a) 场景图

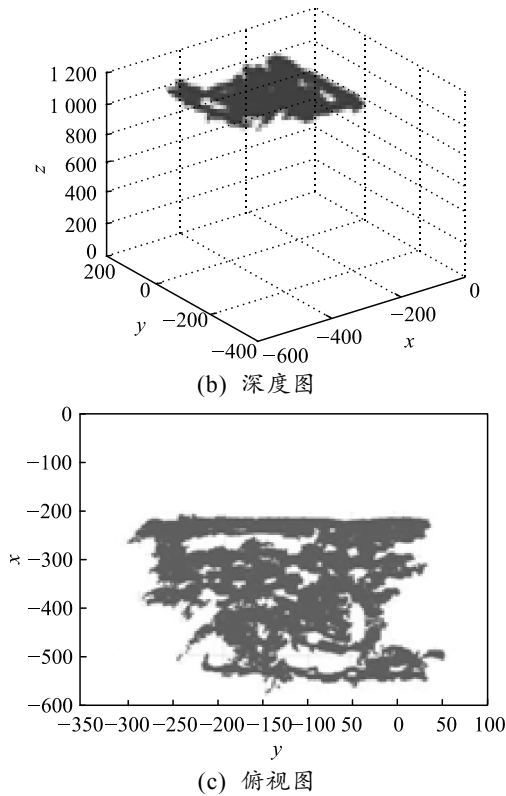


图 9 箱体某面提取效果

3 3 维重建

3 维重建技术通过深度数据获取、预处理、点云配准与融合、生成表面等过程,把真实场景刻画成符合计算机逻辑表达的数学模型,常见的 3 维重建方法有如下几种:

1) 阴影恢复形状法^[8](shape from shading, SFS),该方法利用不同光照条件下图像的明暗程度与阴影来计算物体表面深度信息,并依此实现 3 维重建工作,该方法应用范围广泛,但是数学计算、重建结果不够精细,另外该方法需要准确的光源参数,不适合在室外使用。

2) 机构光法^[9](structured light)通过向表面无特征的物体发射具有特征点的光线,接着利用三角测量的原理计算摄像机与特征光点之间的距离,形成特征点的深度信息,实现 3 维重建。该方法实现简单、精度高,因而被广泛使用。

3) 立体视觉法^[10](multi-view stereo, MVS),包括直接测量获取距离信息,通过利用不同视点上的多幅图像恢复 3 维信息。该方法使用简单方便,不需要有特征光线发生装置,在处理有特征物体时得到了广泛使用。

笔者介绍的基于坐标变换的 3 维重建方法,是基于立体视觉法,由于工作环境中物体一般有丰富

的特征信息,故不需要使用机构光法,而且克服了阴影恢复形状法在实际应用中需要确切了解环境光源参数的弊端。该方法在立体视觉方法的基础上实现了多视角的 3 维信息融合,增强了重建的效果。在已知机器人的运动情况下,可以解算出不同位置的坐标系转化关系,对于场景内物体的坐标点数据进行矩阵的乘法,即可解出主坐标系下的物体空间位置,会大大提高运算速度,很好地满足系统实时性的要求。

坐标变换指一个物体在不同坐标系之间坐标的转换。如图 10 所示,已知空间一点 P 在 xyz 坐标系下的坐标为 (x,y,z) ,求 P 点在新坐标系 $x'y'z'o'$ 下的新坐标,此时需要构造变换矩阵,使得 2 个坐标系重叠,具体分为 2 步:

- 1) 平移坐标轴 $oxyz$,使得坐标原点与新坐标系原点重合;
- 2) 进行旋转变换,使得两坐标系的坐标轴重合。

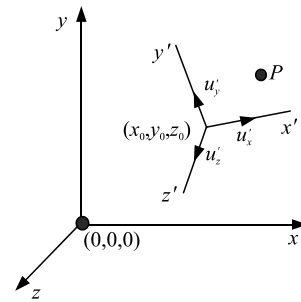


图 10 坐标转换

对于新坐标系 $o'x'y'z'$,其原点坐标为 (x_0,y_0,z_0) ,相对于原坐标系其单位坐标矢量为:

$$\begin{aligned} u'_x &= (u'_{x1}, u'_{x2}, u'_{x3}), \\ u'_y &= (u'_{y1}, u'_{y2}, u'_{y3}), \\ u'_z &= (u'_{z1}, u'_{z2}, u'_{z3}). \end{aligned}$$

如前所述,第 1 步平移坐标系 xyz ,使得 2 个坐标系的原点重合,平移矩阵如下:

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -x_0 & -y_0 & -z_0 & 1 \end{bmatrix}. \quad (8)$$

第 2 步构造旋转矩阵如下:

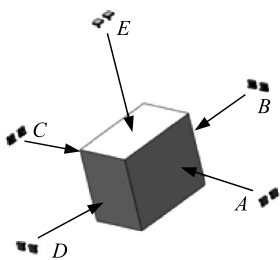
$$R = \begin{bmatrix} u'_{x1} & u'_{y1} & u'_{z1} & 0 \\ u'_{x2} & u'_{y2} & u'_{z2} & 0 \\ u'_{x3} & u'_{y3} & u'_{z3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (9)$$

综上所述,原坐标系下一点 (x,y,z) 转换到新坐标

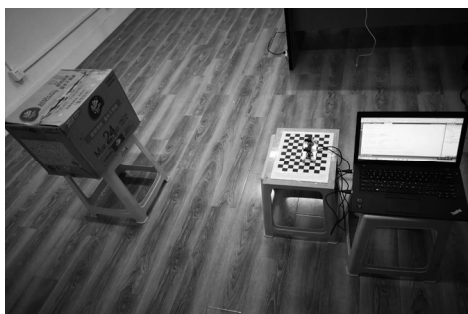
系下点 (x', y', z') 用齐次坐标来描述为：

$$\begin{bmatrix} x' & y' & z' & 1 \end{bmatrix} = \begin{bmatrix} x & y & z & 1 \end{bmatrix} T \cdot R^{[11]}. \quad (10)$$

本次 3 维重建指在不同相机坐标系下的转换。图 11(a)为本次实验的示意图，图 11(b)为本次实验的场景图。在一次实验中，将实验结果即坐标点已经对应的深度信息以坐标点的形式存储在 txt 文档中，并用 Matlab 读取该文档，将对应各面的提取结果展示出来，其中 A 面的提取结果见图 9，其余各面提取结果俯视图见图 12。

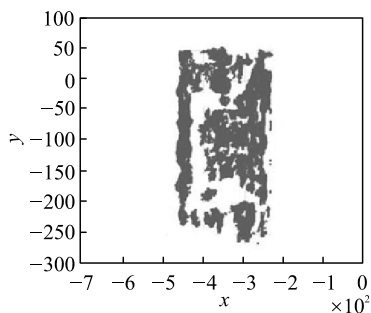


(a) 实验示意图

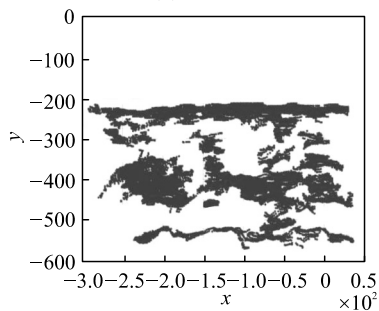


(b) 实验场景

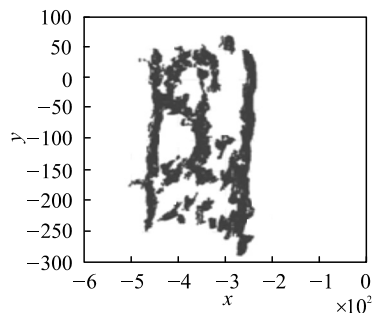
图 11 实验图



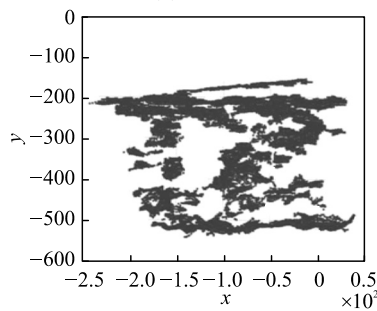
(a) B 面



(b) C 面



(c) D 面



(d) E 面

图 12 B~E 面深度俯视图

利用坐标转换的原理，以 A 面的摄像机坐标系作为主坐标系，将 B~E 面的深度结果进行转换。以 B 面为例：B 面坐标系的 x 轴与 A 面坐标系的 x 轴重合，故 $u'_x = (1, 0, 0)$ ；B 面坐标系的 y 轴与 A 面坐标系的 z 轴重合，故 $u'_y = (0, 0, 1)$ ；B 面坐标系的 z 轴与 A 面坐标系的 y 轴负方向重合，故 $u'_z = (0, -1, 0)$ ，B 面到 A 面的旋转矩阵为：

$$R_{B-A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (11)$$

测量 A 面相机和 B 面相机间距离，x 轴方向为 0，y 轴正方向为 130 cm，z 轴正方向为 138 cm。由式(1)可知：成像面内图像大小 x 与物体大小 X 成正比，与物体距相机距离 Z 成反比。当相机处于与 A 面距离时即 (Z 值一定)，通过实验测得在像素值与距离的比值约为 6.78，因此，可得平移矩阵：

$$T_{B-A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -881 & -936 & 1 \end{bmatrix}. \quad (12)$$

如式(10)所述，将实验测的 B 面所有点坐标，先化为齐次式，再乘以 $T_{B-A} \cdot R_{B-A}$ ，这样就可以得到 B 面反转后的效果。如图 13 所示，完成了 B 面的转

换,水平方向为原始状态,右上角为转化后的效果。其余结果面的做法一样。

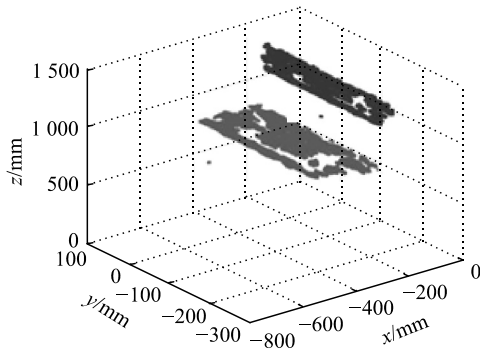


图 13 B面转化前后对比

最终结果如图 14。由图可看出,重建模型把长方体纸箱的轮廓较好地复现出来。

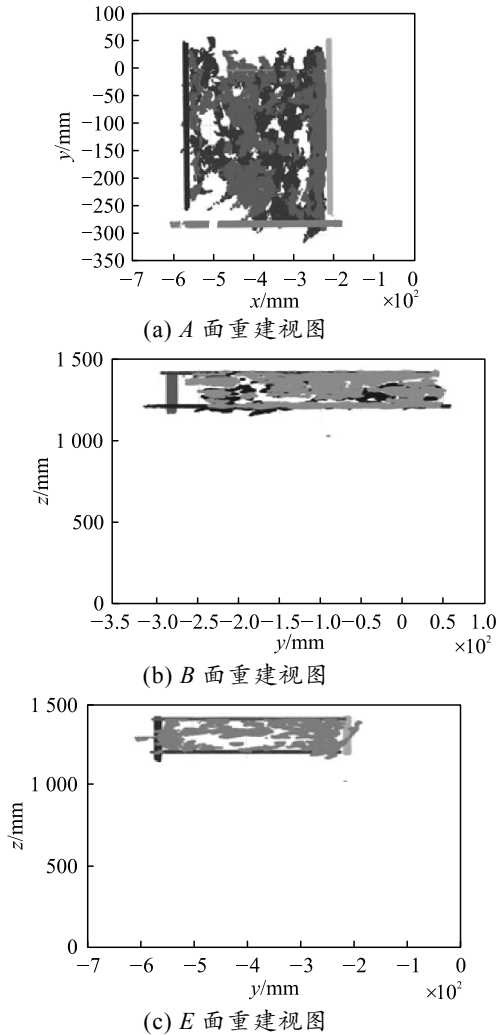


图 14 3 维重建效果图

4 结束语

基于坐标变换的 3 维模型重建系统,利用双目立体视觉系统,通过改进后的灰度算法进行预处理操作,提高运算效率,再经过标定-校正,立体匹配等操作,获取场景中物体的深度信息。设计实验,获取同一个物体 5 个表面(除底面)的深度信息,进而通过坐标变换的方法,将 5 个独立的面转换到同一个坐标系下,得出了最终的重建效果,对实验纸箱做到了较好的 3 维重建效果。笔者对传统的立体视觉 3 维重建方法进行了拓展,与坐标转换法结合,将多视场内的深度信息融合,还原出空间物体的 3 维轮廓,获取更加丰富的 3 维信息,提高了机器人理解环境信息的能力,从而保障了后续工作的正常开展。该方法对于智能机器人环境系统重建以及机器人避障导航系统都有一定的参考价值。

参考文献:

- [1] 李宸鑫. 三维图像重建中的深度信息处理方法[J]. 电子技术与软件工程, 2018(16): 62.
- [2] 王琳, 贾金子, 韩亦扬. 基于计算机视觉的三维重建技术综述[J]. 农村经济与科技, 2017, 28(10): 260.
- [3] 于仕琪, 刘瑞祯. 学习 OpenCV[M]. 北京: 清华大学出版社, 2009: 406-416.
- [4] 王志远, 王茂森. 基于双目视觉的移动机器人测距与定位系统[J]. 兵器装备工程学报, 2017, 38(11): 173-177.
- [5] 周翠香, 李丹, 郭施, 等. 形零部件 3 维工艺设计方法[J]. 兵工自动化, 2018, 37(6): 14-17.
- [6] 冯焕飞. 三维重建中的相机标定方法研究[D]. 重庆: 重庆交通大学, 2013: 8-9.
- [7] 文成维, 陈文建, 詹泳. 嵌入式双目立体视觉测距系统[J]. 计算机系统应用, 2012, 21(10): 143-147.
- [8] 尹贞燧. 浅谈三维重建技术之由阴影信息恢复形状[J]. 中国科技信息, 2013(23): 86-87.
- [9] 段军. 结构光三维扫描技术的应用研究[J]. 现代制造技术与装备, 2019(1): 49-51.
- [10] 曹焱翔. 基于双目视觉的三维重建[J]. 通讯世界, 2018, 25(12): 236-238.
- [11] 陈义, 沈云中, 刘大杰. 适用于大旋转角的三维基准转换的一种简便模型[J]. 武汉大学学报(信息科学版), 2004, 29(12): 1101-1105.