PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

An experimental scene dynamic perception algorithm for intelligent beaker

Zhiquan Feng, Hongyue Wang, Changsheng Ai, Haiyan Shao

Zhiquan Feng, Hongyue Wang, Changsheng Ai, Haiyan Shao, "An experimental scene dynamic perception algorithm for intelligent beaker," Proc. SPIE 12079, Second IYSF Academic Symposium on Artificial Intelligence and Computer Engineering, 120790A (1 December 2021); doi: 10.1117/12.2622843

Event: 2nd IYSF Academic Symposium on Artificial Intelligence and Computer Engineering, 2021, Xi'an, China

An experimental scene dynamic perception algorithm for intelligent beaker

HongYue Wang[1,a,b], ZhiQuan Feng* [a,b], ChangSheng Ai[2,c], HaiYan Shao[3,c]
[a]School of Information Science and Engineering, University of Jinan
[b]Shandong Provincial Key Laboratory of Network Based Intelligent Computing, Jinan, China
[c]Department of mechanical and Electronic Engineering, University of Jinan
[1]652458257@qq.com
* Corresponding author: ise_fengzq@ujn.edu.cn
[2]me_aics@ujn.edu.cn
[3]meshaohy@hotmail.com

Abstract

In middle school intelligent chemistry experiments, experimenters often perceive the scene information through Kinect and other fixed cameras. There is a serious occlusion problem, which makes this method almost infeasible. Therefore, this paper puts forward a research idea of setting a micro camera on the beaker for the first time. The main contribution of this work is to propose a 3D reconstruction algorithm of experimental scene based on mobile camera.The main innovation lies in the process of obtaining two images of the same laboratory scene from different angles through a moving binocular camera, then using the point cloud to reconstruct the three-dimensional shape of the experimental instrument in the scene and restore the spatial position information of the object. In order to reconstruct the experimental instrument information better in the chemical experiment scene, this paper integrates the model of transforming the picture from RGB to HSV based on the traditional binocular 3D reconstruction algorithm, which is also another innovation of this paper.

Keywords-3D reconstruction; Binocular calibration; Computer vision; OpenCV

1. INTRODUCTION

In the real middle school intelligent chemistry experiment scene, there are many problems, such as the variety of instruments and the huge volume of traditional recognition equipment. These problems will lead to serious occlusion between instruments and between instruments and interlocutors, so the system is difficult to effectively and accurately perceive the experimenter's operation behavior and experimental scene information. Therefore, this paper proposes an experimental scene dynamic perception algorithm for intelligent beaker. By installing the binocular camera on the intelligent beaker [1], various experimental instruments in the traditional chemical laboratory scene can be accurately reconstructed in three dimensions during the movement of the intelligent beaker.

In the early stage of technology development, we usually use image information for three-dimensional reconstruction. For example, Xiaohui Hao et al. [2] used ultrasonic image for three-dimensional reconstruction of medical information; Huiling Qing[3] used lidar technology for 3D reconstruction of terrain scanning. With the development of technology, Qin Zhang et al. [4] used monocular camera to restore the target three-dimensional structure by using pictures, but monocular three-dimensional reconstruction is prone to mismatching in dynamic environment and is not competent for complex scenes. In 2010, Microsoft Research Institute [5] launched a three-dimensional stereoscopic camera Kinect, which can reconstruct the target in three dimensions by obtaining the depth map. This technology is different from 3D point cloud stitching. It mainly uses a Kinect to continuously scan around the object and reconstruct the 3D model of the object in real time, which effectively improves the reconstruction accuracy. With the development of deep learning in recent two years, Jia Wu et al. [6] reconstructed the input image through convolution neural network.And its application forms are also diverse. For example, vijayanarasimhan et al. [7] reconstruct three-dimensional by robot, and Rui Liu et al. [8] reconstruct three-dimensional by combining images taken by UAV.

The above research progress has good results for their respective application scenarios, but they generally have the problems of high price of application equipment and high time complexity. Aiming at the application scenario of traditional laboratory, this paper proposes an experimental scenario dynamic perception algorithm for intelligent beaker.

Second IYSF Academic Symposium on Artificial Intelligence and Computer Engineering,
edited by Wei Qin, Proc. of SPIE Vol. 12079, 120790A · © 2021 SPIE
0277-786X · doi: 10.1117/12.2622843

2. MATERIALS AND METHODS

The three-dimensional reconstruction algorithm of binocular stereo vision consists of five parts: image acquisition, camera calibration, image correction, stereo matching and depth value calculation. The specific process is to select two images from different perspectives of the same scene, and use the two-dimensional information to restore the three-dimensional information of the visible surface of the scene. Based on the traditional binocular 3D reconstruction algorithm, this paper integrates the HSV color conversion model to extract the instrument information required by the laboratory scene more accurately. The algorithm framework is shown in Figure 1.

Figure 1 Algorithm block diagram of binocular 3D reconstruction system

2.1 Image acquisition

This paper uses the binocular camera to obtain the original image, uses the videocapture() function of OpenCV to obtain the binocular camera scene, and cuts the scene video into picture frames every 2 seconds to obtain the image.In this paper, the binocular camera is installed on the intelligent beaker for middle school chemistry experiment. During the movement of the experimenter taking the beaker, the image information of the experimental desktop is captured in real time. The experimental equipment is shown in Figure 2.

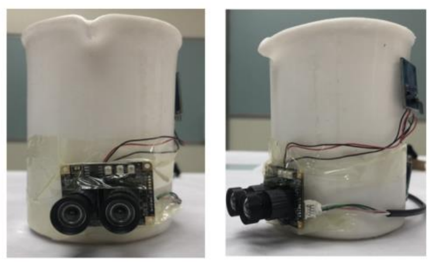

Figure 2 Equipment diagram of intelligent beaker combined with binocular camera

2.2 Binocular calibration

The goal of dual target determination is to obtain the internal parameters, external parameters and distortion coefficients of the left and right cameras. The internal parameters include the focal lengths f_x and f_y of the left and right cameras and the imaging origin c_x and c_y the external parameters include the rotation matrix R and translation vector t of the left camera relative to the right camera, and the distortion coefficient includes the radial distortion coefficient (k_1, k_2, k_3) and the tangential distortion coefficient (p_1, p_2), so as to determine the mathematical relationship between the object point and the image point. The binocular camera calibration system includes four coordinate systems, as shown in Figure 3.

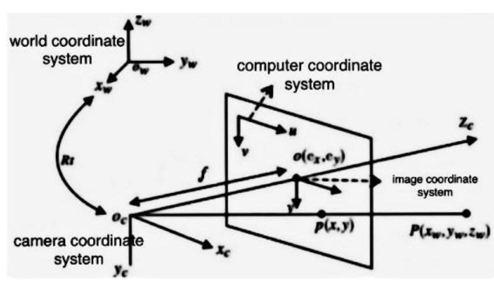

Figure 3 Double target fixed coordinate system

They are: world coordinate system$o_w X_w Y_w Z_w$, camera coordinate system$o_c X_c Y_c Z_c$, image coordinate systemoxyand computer image coordinate system uv. According to the pinhole imaging principle, the transformation from camera coordinate to image coordinate system can be obtained, as shown in formula (1):

$$\begin{cases} x = f\dfrac{x_c}{z_c} \\ y = f\dfrac{y_c}{z_c} \end{cases} \rightarrow z_c\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \quad (1)$$

The transformation from the image coordinate system to the computer coordinate system is shown in formula (2):

$$\begin{cases} u = f\dfrac{x}{d_x} + c_x \\ v = f\dfrac{y}{d_y} + c_y \end{cases} \rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{d_x} & 0 & c_x \\ 0 & \dfrac{1}{d_y} & c_y \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

d_x is the physical size of each pixel in the x direction, d_y is the physical size of each pixel in the y direction, and (c_x, c_y) is the center coordinate of the image.

The transformation from the world coordinate system to the camera coordinate system is shown in formula (3):

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = R\begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = t \rightarrow \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix}\begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (3)$$

From formulas (1), (2) and (3), the computer image coordinate transformation relationship between the point P coordinate of the spatial world coordinate system and its projection point can be obtained, as shown in formula (4):

$$z_c\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix}\begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = M_1 M_2 x_w \quad (4)$$

M_1 is the internal parameter matrix, f_x and $f_y$is the effective focal length of u-axis and v-axis respectively, and u_0、v_0 is the optical center; M_2 is the external parameter matrix, R and t are the rotation matrix and translation vector; The internal parameters of the nonlinear model also include radial distortion parameter k_1、k_2 and tangential distortion parameter p_1、p_2.

2.3 Image correction

2.3.1 DISTORTION CORRECTION

In the process of image data acquisition, the lens of binocular camera often turns a straight line in the real environment into a curve in the picture. The closer to the edge of the image, the more obvious this phenomenon is. This distortion phenomenon is mainly divided into two categories: barrel distortion and pillow distortion, as shown in Figure 4:

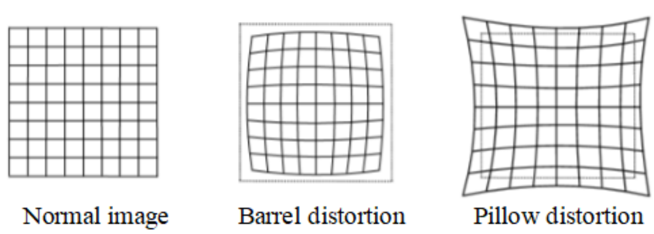

Figure 4 Distortion diagram

2.3.2 STEREO CORRECTION

The purpose of stereo correction is to perform mathematical projection transformation on the left and right views taken in the same scene, so that when the two camera image planes are on the same plane and the same point is projected on the two camera image planes, it should be aligned on the same line of the two pixel coordinate systems, which is simply called coplanar line alignment.

2.4 Stereo matching

The essence of stereo matching is to give a point in one image and find the corresponding point in another image, so that these two points are the projection of the same object point in space. The commonly used stereo matching methods can be basically divided into two categories: (a) local methods: BM, SGM, ELAS, Patch Match, etc. (b) Global methods: Dynamic Programming, Graph Cut, Belief Propagation, etc. In this paper, we use SGBM algorithm to generate parallax map.

2.5 Depth calculation

The depth value calculation is based on the principle of triangulation, and the three-dimensional point cloud information of the scene is reconstructed by using the camera model and parallax map. The reconstruction accuracy is mainly affected by the matching parallax accuracy and baseline width. It is generally directly proportional to the matching parallax accuracy and inversely proportional to the camera baseline length.

2.6 The image is converted from RGB to HSV color space

RGB color space uses the linear combination of three color components to represent color. Any color is related to these three components. Therefore, continuous color transformation is not intuitive. If you want to adjust the color of the image, you need to change these three components.

Based on the above reasons, HSV color space is more used in image processing. In HSV color space, it is easier to track objects of a certain color than BGR. It is often used to segment objects of a specified color. In this paper, we use HSV color space to segment the required chemical instruments, which can reconstruct and identify the scene information more accurately.

2.7 Dynamic perception algorithm of experimental scene for intelligent beaker

The specific process is to select two images from different perspectives of the same scene and restore the scene with the two-dimensional information. In this paper, we use an experimental scene perception algorithm based on binocular motion vision to perceive the beaker, test tube, test tube rack, spoon, brown jar and other objects in the experimental scene.The algorithm process is shown in Table I.

TABLE I. ALGORITHM

Dynamic Perception Algorithm of Experimental Scene for Intelligent Beaker(DPAES)
Input: Real time video sequence **VS**,binocular calibration result **BC**
Output: 3D point cloud of chemical experiment scene **S**
1:**While** RGBImg!=Empty**While** VS != NULL **and** BC **do**
2: The binocular video sequence is cut into the image information of the left and right cameras respectively,
$Img^0_{left}, Img^0_{right} = CutVideo(VS)$
3: The binocular calibration information **BC** is used to correct the distortion of the image information in the second step,
$Img^1_{left}, Img^1_{right} = Undistort(Img^0_*, BC)$
4: The binocular calibration information BC is used to stereo correct the image information in the third step,
$Img^2_{left}, Img^2_{right} = StereoRectify(Img^1_*, BC)$
5: Convert RGB image to HSV image,
$Img^3_{left}, Img^3_{right} = CvtColor(Img^2_*)$
6: Generate point cloud 3D reconstruction scene,
$S = PointCloud(Img^3_*)$
7: **End**

3 RESULTS & DISCUSSION

3.1 Experimental environment

The camera used in this paper is a 4mm focal length 100 degree wide-angle distortion free Weixin binocular camera with a resolution of 1920 * 1080.The computer system is Windows system, the CPU is i7-8750H, and the graphics card isRTX3090 * 2.

3.2 The process of experimental double target determination

This process can be roughly divided into two parts: (a) calibrate the monocular camera of the left and right cameras respectively to obtain the calibration parameters of the left and right cameras. (b) The rotation matrix and translation vector between the left and right cameras are further calculated to complete the stereo calibration. Figure 5 shows the checkerboard calibration board of this paper.

Figure 5 Schematic diagram of chessboard calibration board

In the process of binocular calibration in this paper, the calibration toolbox of Matlab is used to calibrate the checkerboard pictures taken by 50 groups of binocular cameras. Finally, the data parameters we obtained are shown in Table II.

TABLE II. MATLAB CALIBRATION PARAMETER TABLE

	Left camera			Right camera		
k_1	0.0933			0.0927		
k_2	0.1693			0.1531		
k_3	-0.1804			-0.1468		
p_1	0.0175			0.0181		
p_2	0.0118			0.0082		
Internal reference	440.25	0	328.80	440.25	0	342.74
	0	436.30	243.01	0	436.30	237.59
	0	0	1	0	0	1
Rotation matrix	1		0.0019782258		0.0086644	
	-0.00197752		1		-8.632944937e-05	
	-0.0086646		6.919156687e-05		1	
Translation matrix	-16.453938		1.734179209		-0.220235644	

3.3 The process of correction

In order to accurately reconstruct the experimental instrument information of the chemical laboratory, this paper uses the DPAES algorithm for one frame scene. First, distortion correction and stereo correction are performed by using the undisport() and remap() functions of OpenCV, as shown in Figures 6 and 7. After the calibration process, we extracted the chemical experimental instrument red beaker, green spoon, brown jar, test tube rack and yellow test tube according to the corresponding HSV values, as shown in Figure 8. Finally, the PCL (point cloud Library) point cloud library is used to generate the experimental instrument for reconstructing the chemical experiment scene, as shown in Figure 9.

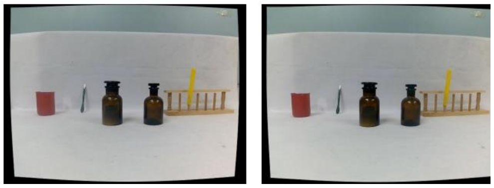

Figure 6 Distortion correction diagram

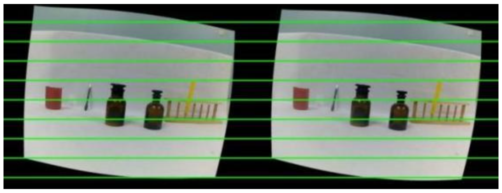

Figure 7 Schematic diagram of stereo correction

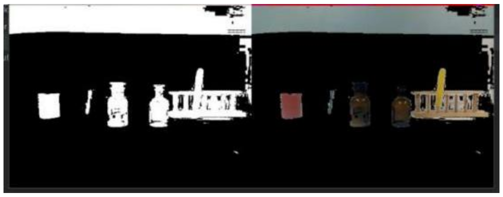

Figure 8 Schematic diagram of HSV color space transformation

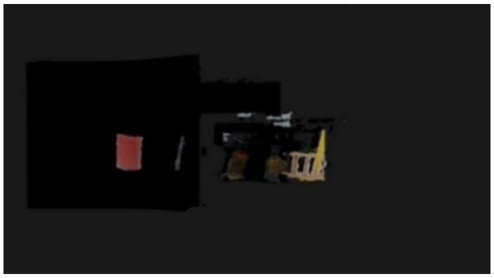

Figure 9 3D point cloud generation diagram of experimental instrument

4 CONCLUSION

In domestic middle school chemistry courses, many chemical experiments have the characteristics of high cost and relatively dangerous operation and so on. So, many experiments can not be operated and displayed. Therefore, the emergence of intelligent beaker can improve the accuracy of user behavior perception and the authenticity and immersion of user experimental operation in the virtual chemistry experiment system. However, in the real experimental scene, due to the complex types of instruments and the huge volume of traditional recognition equipment, there are serious occlusion problems between instruments and instruments, and between instruments and interlocutors, so the system is difficult to effectively and accurately perceive the experimenter's operation behavior and experimental scene information. Therefore, based on the interaction and intelligence of the intelligent beaker, this paper integrates the binocular 3D reconstruction technology, which can reconstruct the experimental scene in real time and perceive the location information in the program.

ACKNOWLEDGMENTS

This paper is supported by the the Independent Innovation Team Project of Jinan City (No. 2019GXRC013)

REFERENCES

[1] ZhiQuan Feng,Di Dong. An intelligent beaker based on experimental scene situation awareness[P]. ShanDong：CN112295617A,2021-02-02.

[2] XiaoHui Hao, ShangKai Gao,XiaoRong Gao. An ultrasonic 3D image reconstruction method based on irregular sampling plane[J]. Journal of Tsinghua University:Natural Science Edition, 1999(09):9-13.

[3] Huiling Qing. Research and design of 3D reconstruction system based on lidar data[D].Central South University,2005.

[4] Qin Zhang,QingXuan Jia. 3D reconstruction of indoor scene based on laser and monocular vision[J]. Journal of system simulation, 2014(02):357-362.

[5] Han, J, Shao, et al. Enhanced Computer Vision With Microsoft Kinect Sensor: A Review[J]. IEEE Transactions on Cybernetics, 2013, 43(5):1318-1334.

[6] Vijayanarasimhan S, Ricco S, Schmid C, et al. Sfm-net: Learning of structure and motion from video[J]. arXiv preprint arXiv:1704.07804, 2017.

[7] Li Huang. Research on 3D environment reconstruction technology for mobile robot visual navigation[D]. Zhejiang University of Technology.

[8] Rui Liu,BinBing Li,Jie Luo.Visual 3D reconstruction based on UAV monocular video[J]. Computer engineering and design, 2017, 38(01): 160-164.