**Building 3D image from 2D scenes**

1.Principle of the project

The method I use to build 3D image from 2D scenes comes from binocular stereo imaging, which uses two cameras to simulate the process of human eyes recognizing images and constructing 3D scenes in their minds. In experiment environment, it takes four steps to build the 3D image.

To start with, it is necessary to calibrate the two cameras. By doing so, we can obtain the parameters of the two cameras, which plays an important role in the next steps.

After that, we take a picture of our target from each camera and these two pictures are our stereo images. Stereo images are of important since we use them to calculate the depth of our target, however, the images we get from our cameras are often distorted. It is because we can not ensure that the imaging planes of the two cameras are in the same plane. In this case, we should rectify our stereo images before we make use of them.

Then, we extra the characteristic points from both images and match them with one another. For example, if point A is from stereo image 1 and point A’ is from stereo image 2. If and only if point A and point A’ represent the same position on target object, we say A and A’ is a pair of matching points. By doing so, we describe the characteristic of our target.

Finally, we use characteristics of our target to build 3D image.

2.The process of experiment

2.1 Camera calibration

相机标定的目的是为了得到相机的内外参数，此处我们进行的是双目标定。在进行标定之前首先引入坐标系的概念。

The purpose of camera calibration is to get the internal and external parameters of the camera. Here we do binocular calibration. The concept of coordinate system is introduced before calibration.

2.1.1 Image coordinate system（u，v）

这种坐标系以像素为单位，表示的是图像中像素点在图像中所处的位置。

This coordinate system is based on pixels, which represents the position of pixels in the image.

2.1.2 Physical coordinate system of image（x，y）

这种坐标系在像素坐标系下建立，以物理单位表示像素在图像中的位置。其与像素坐标系（u，v）有如下的转换关系，

This coordinate system is established in the pixel coordinate system, and the position of the pixel in the image is represented by physical units. Point *C1* is the intersection of the camera optical axis and the image plane, the coordinate of point C in the (u,v) coordinate system is (u0,v0) and the unit length of the two axes of the coordinate system is *dx*, *dy*. In this case, the Physical coordinate system of image (x,y) has the following transformation relationship with the Image coordinate system (u, v) :

, (2-1)

 (2-2)

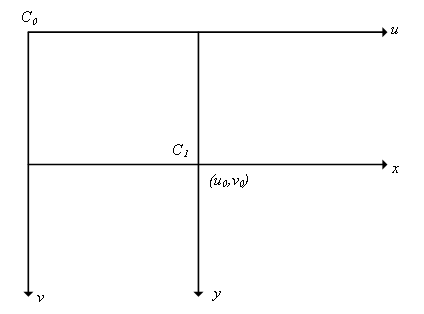


Fig.2-2 The result of calibration

2.1.3 Camera coordinate system

这种坐标系用来表示三维空间中的物体到像平面的投影关系。

This coordinate system is used to represent the projection relationship between the object and the image plane in three-dimensional space.

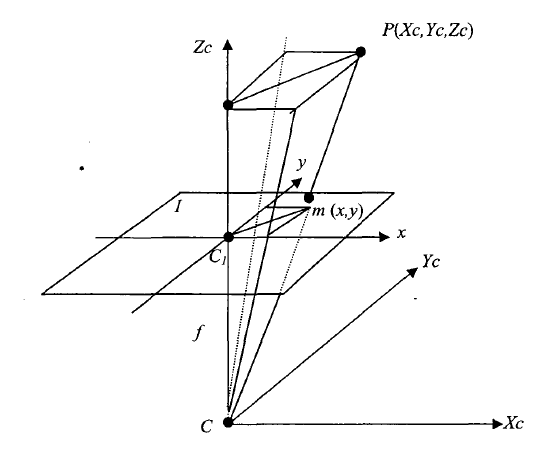


Fig.2-2 The result of calibration

As shown in the figure, *f* is the focal length of the camera and (XC,YC,ZC) is the coordinate of point P in the Camera coordinate system. Therefore, we can obtain the relationship between (XC,YC,ZC) and (x,y,z) :

 (2-3)

2.1.4 World coordinate system

这种坐标系在将相机置于某一场景中时使用。

This coordinate system is used when placing the camera in a scene. It has following relationship with Camera coordinate system:

 (2-4)

Where *R* is the rotation matrix and *t* is the 3D translation.

Simultaneous formula(2-1),(2-2),(2-3),(2-4), We can deduce the following relations:

 (2-5)

2.2 Binocular calibration

由前文转换关系可得，其中M1矩阵代表摄像机的内参数，M2矩阵代表摄像机外参数。通过确定M1M2，即可得到摄像机的参数。由于我们操作时通过用两个摄像机在不同角度拍摄同一物体进行标定，因此还需要确定两个相机之间的相对位置。

From the above transformation relations, *M1* matrix represents the internal parameters of the camera, and *M2* matrix represents the external parameters of the camera. The parameters of the camera can be obtained by determining*M1* and *M2*. Because we use two cameras to shoot the same object at different angles for calibration, we also need to determine the relative position between the two cameras.

To do so, I place two smart phones in the same plane as my cameras and I use Stereo image calibrator in matlab to calibrate the two cameras.

This app can easily calibrate the two cameras by scan pairs of pictures taken by the cameras. Normally, the app expects at least 10 pairs of pictures to calibrate. However, since I did this experiment with a lack of professional environment, I only use two pairs of pictures to calibrate the cameras.



Fig.2-1 Two pairs of pictures from two camera

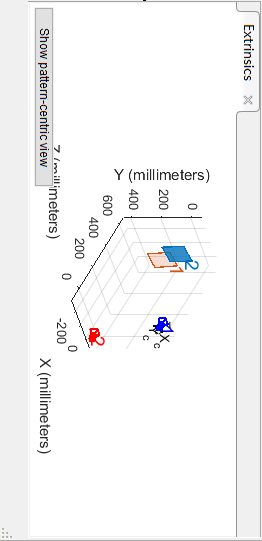


Fig.2-2 The result of calibration

As shown in Fig. ,Cameraparameters1 is a separate calibration parameter for camera1 and cameraparameters2 is a separate calibration parameter for camera2. StereoParams .Rotation of Camera 2 shows the rotation of camera 2 relative to camera 1, Translation of Camera 2 shows the translation of camera 2 relative to camera 1.

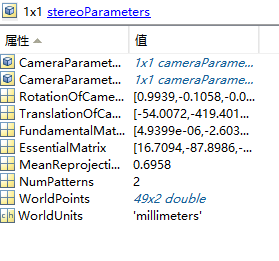
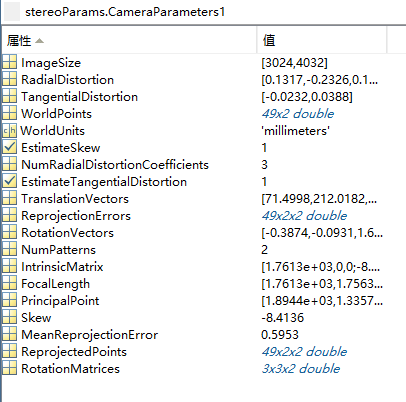


Fig.2-2 The result of calibration

As shown in Fig. ,we can find Radial Distortion,Tangential Distortion and Inreinsic Matrix of camera1.



2.2 Image rectification

此时我们虽然已经得到了相机的参数，但仍需要进行畸变矫正之后才能进行匹配，这是因为匹配时我们希望两台相机的成像平面保证共面，进而减少匹配的计算量。

At this time, although we have obtained the parameters of the camera, we still need to carry out distortion correction before matching. This is because we want the imaging planes of the two cameras to be coplanar so as to reduce the amount of calculation.

I called rectify StereoImages functions in matlab to rectify the two images, this function takes the two images as independent variables and use the camera parameters we got from 2.1 as constraint condition.



Fig.2-3 Stereo image from camera1



Fig.2-4 Stereo image from camera2



Fig.2-5 before rectification

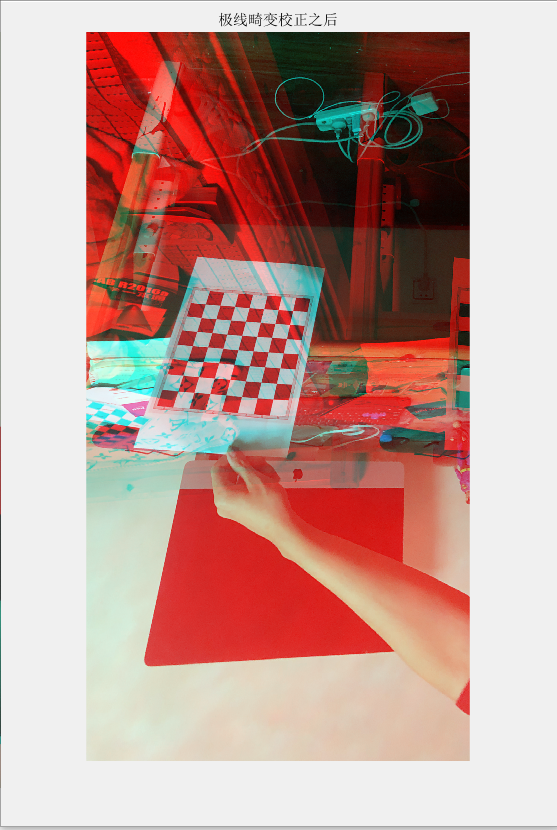


Fig.2-6 after rectification

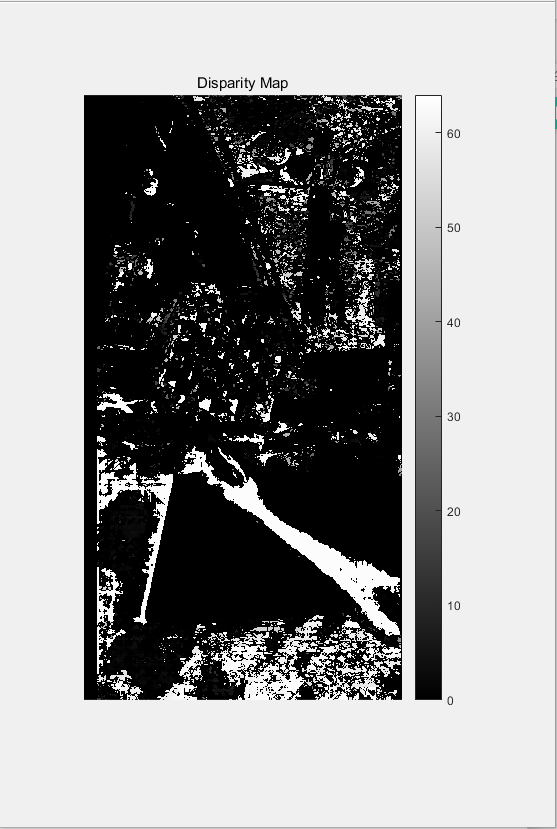
2.3 Block matching

立体匹配是立体视觉研究中的关键部分。其目标是在两个或多个视点中匹配相应像素点，计算视差。通过建立一个能量代价函数，对其最小化来估计像素点的视差，求得深度。

2.3.1极线约束

为了达到这个目的，理论上需要将左右相机拍摄同一物体时各自照片中的对应点匹配起来，这里采用极线约束进行辅助。如下图所示：

By calculating the disparity between characteristic points, I matched them with on another. Calling function in matlab is a feasible way to do so, it allows user to generate disparity map. However, I found this function fails in accuracy.

 Fig.2-7 disparity map

2.4 3D construction

After acquiring the disparity map, we can tell the occlusion relations in our target, which allows us to build 3D image. In this step, the reconstructionScene function in matlab helps building the image.

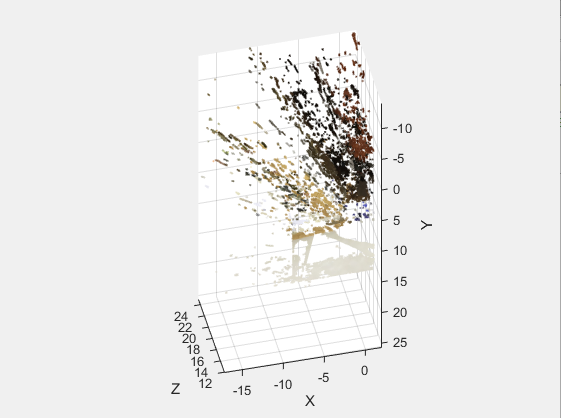


Fig.2-7 3D construction

3.The result analysis

While doing this project, I divided the process into several steps which simplifies my operation. However, as showed in Fig.2-7,the result of my experiment is not unsatisfactory.

Two factors have contributed to this result.

To start with, two pairs of images are not enough for camera rectification, more images may add the accuracy of camera parameters.

Secondly, the disparity function in matlab is not enough to support the goal. Maybe I should try to write my own code in order to solve the problem. Also, some literature shows that build 3D image with a combine of matlab and Open CV solves the problem.