Depth calculation using stereovision

Chirag Lamba (15116017)
Electronics and communication engineering dept.
Indian Institute of Technology, Roorkee
Roorkee,India(247667)

***Abstract*—With a single camera we can not calculate the distance of any point in an image, hence we use two cameras separated by some distance known as stereocameras to calculate the distance of any point in the two images captured by the two cameras using the method of triangulation.There are many pre-steps(calibration , rectification) that needs to be done before taking images from two cameras and calculating depth.These methods are used in most of the 3D vision devices today.**

Keywords—stereovision; triangulation; camera calibration; rectification; Depth map;

I. Introduction

This paper will briefly define how to calculate the distance the distance of a point from the center of two cameras. Stereovision in uses in almost all 3D application devices like(3D movies glasses , kinect , 3D maps).human vision can calculate depth because of stereovision(two eyes).If we have only one eye we will not be able to calculate the distance of object in front of us.

If we have only a single pinhole camera all points on line POQ(passing through the optical center of camera) will converge at a single point on the image plane.Hence, we can not calculate the distance of actual point lying on the line PQ.

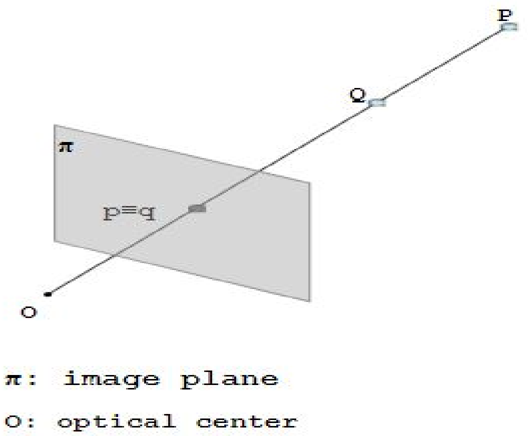

Figure 1. All points on line PQ converge at a single point

If we have two or more cameras separated by a particular distance and aligned almost parallel to each other known as stereocameras (as shown in figure 2) , we can actually calculate the distance of the point using triangulation method(described in section 2). There are some basic pre-processing which includes calibration(removing of lens distortions , calculation of focal length and distance between the cameras) , rectification(the two cameras planes are not aligned perfectly parallel to each other) that needs to be done before taking images from cameras.

After these steps depth map of the scene is created which contains the depth information of the points in front of the cameras.

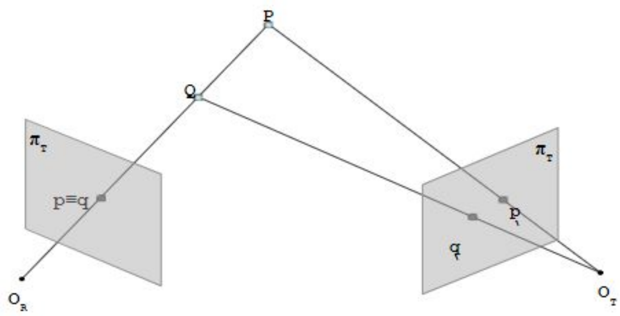

Figure 2. Two cameras to calculate the distance of the point

II. Triangulation method for depth calculation

The triangulation method uses the simple similarity of the two triangles to calculate the depth, it is a method similar to parallax method.

As can be seen if Figure 3.

OL anr OR are left and right cameras.
PL and PR are left and right image planes.

From similar triangles() we can calculate-

d/f = a/xL = c/(-xR) (1)

a + c = d*(xL - xR)/f (2)

b= d*(xL - xR)/f (3)

d = b*f/(xl - xR) (4)

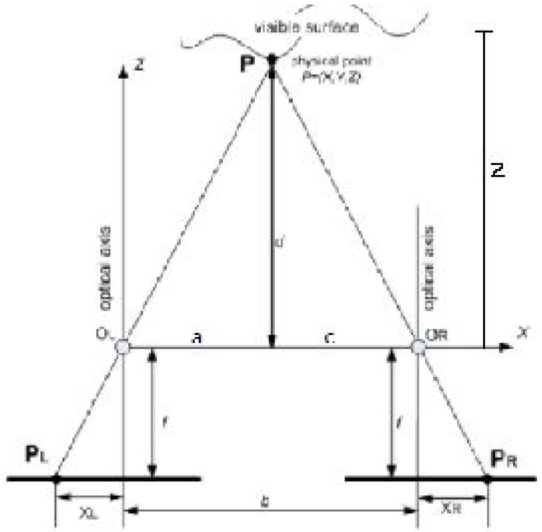

Figure 3. Triangulation method for depth calculation

III. CAMERA CALIBRATION

The first step of the preprocessing is camera calibration.calibration calculates the focal length and distance between the cameras and also removes the distortion caused by camera lenses.

A. removing lens distortions

The lenses of most of the common cameras used today are not perfectly parabolic in nature Hence,the magnification of the image(or the bending of light) is more while going away from the optical axis of the lens(as shown in figure 4) this type of distortions are known as radial distortion.The actual coordinates of the image points appears greater than they actually are.

The equation for correcting radial distortions are-

$$x(corrected) = x(1 + k_1r^2 + k_2r^2 + k_1r^2) \quad (5)$$

$$y(corrected) = y(1 + k_1r^2 + k_2r^2 + k_1r^2) \quad (6)$$

where k1 , k2 , k3 are constants which needs to be calculated.

these equations are in the form of taylor series up to 3 terms.

Another type of distortion known as tangential distortion occurs because the image plane is not perfectly parallel to the camera plane.

the equations for removing tangential distortion-

$$x(corrected) = x + (p2(r^2 + 2y^2) + 2p1xy) \quad (7)$$

$$y(corrected) = y + (p1(r^2 + 2y^2) + 2p2xy) \quad (8)$$

where p1 and p2 are constants which needs to be calculated

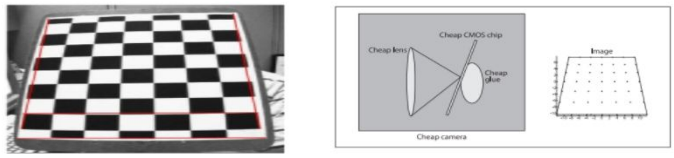

Figure 4. radial distortion(left) tangential distortion(right)

B. calculation of focal length and distance between cameras.

we can calculate the focal length and distance between the cameras using 10-15 images of a standard object.computer vision software opencv uses chessboard as the standard object.The images are taken in different orientation and different distance from cameras.Some mathematical calculation is done by opencv and the final result is returned in the form of a matrix.

.

IV. RECTIFICATION

After the calibration step the distortions in the image , focal length and distance of cameras is also known.To apply the triangulation method the images must be parallel to each other and the image planes must be levelled.we need the information of the rotation and translation between cameras the rotational and translational matrices are also calculated by opencv using inbuilt functions.

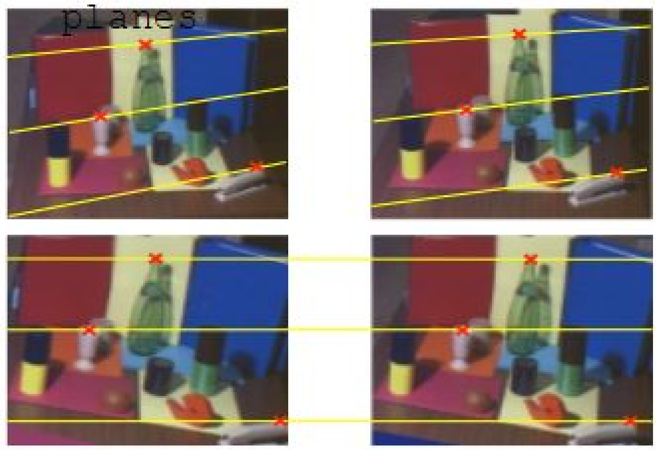

Figure 5.Image rectification

Opencv uses SIFT(scale invariant feature transform algorithm) to calculate the matching features between the two images and calculate the epipolar line and then finding the rotation and translation vectors between cameras.

As seen in figure 5, after the rectification step the image planes become totally parallel to each other.the yellow lines shown in the figure are the epipolar lines which represents the angle of rotation between the cameras.

V. Depth map

After the rectification step depth map of the scene is created using the triangulation method on the two stereo images.

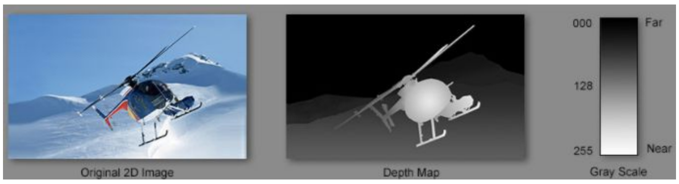

Figure 6. Depth map from stereo images

.

Depth map is a 2D image which contains 3D information in grayscale values. more grayscale value represents farther points and vice versa.
Depth map is used in a variety of 3D reconstruction problems(like 3D maps).

VII. Devices and Applications of Stereovision

Human depth perception is based on stereovision, if we have only a single eye we would not be able to calculate the depth of objects around us.

Microsoft Kinect which is used in video games for depth perception is also based on calculating the depth map. However , the approach is somewhat different.

3D maps - we have seen 2D maps (like google maps) which only contains distance information between two places. recently google has added 3D features in in google maps which also allows us to extract the height information about features in a map.

3D pictures and VR's - 3D pictures and VR's also use stereovision in its 3D cameras to create a 3D view for the user.

Figure 7(a). 3D maps

Figure7.(b) A VR headset

VI. Conclusions

Stereovision is used in a variety of devices. However the current algorithms are not as fast and efficient as expected. we can not use this distance calculation method when there is a delay between the frames taken by our two cameras hence this limits the use of opencv in moving cameras approach. Moreover using more than two cameras is an efficient solution but that will reduce the field of view between which is the intersection of the field of views of all the cameras.Still stereovision is a reliable method in a variety of 3D vision applications.

References

[1] S. Solak and E. D. Bolat, "Distance estimation using stereo vision for indoor mobile robot applications," *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)*, Bursa, 2015, pp. 685-688.

[2] https://www.google.co.in/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwiloMyI-M_XAhUKvY8KHYstA5kQFggnMAA&url=http%3A%2F%2Fwww.bogotobogo.com%2Fcplusplus%2Ffiles%2FOReilly%2520Learning%2520OpenCV.pdf&usg=AOvVaw2EchJ-LMDHNC3Zf6GKwLZ1