Mapping rooms with a Stereo camera from a fixed location

A. Turner, R. Green

Dept. Computer Science and Software Engineering,
University of Canterbury,
Christchurch,
New Zealand

Email: atui5@student.canterbury.ac.nz,
richard.green@canterbury.ac.nz

Abstract

This paper described a method for a stereo camera with one degree of freedom to generate a map of a room. The relative location of features is used to translate them, along with the depth information, to the correct position to build a map of the environment.

Keywords: Map generation, Stereo camera

1 Introduction

When building a map of an unknown environment, adding a second camera can provide sufficient information to generate such a proximity map. However, the mapping and localization problems can be considered independent [1]. In this paper we will be presenting an algorithm to map a room using a fixed location where only the yaw of the camera is changed. Work into mapping an environment using sonar has been performed by [2]. Others have worked on using a single camera to build a panoramic image. [3] uses an off set camera to build their panorama. [4] describes a similar method using polycentered panorama’s to look at the 3D scene.

The position of a camera can be estimated using vision when in a known environment [5]. Stereo vision has also been used to navigate unknown areas [6] as we as known [7] and mapping them when they are unknown [8, 9, 10, 11].

2 Method

To map a room using a stereo camera we need to know the angle the camera is relative to previous frames. A similar method to the one the authors chose is described in [12].

When the relative angle is found the system calculates the location of the found objects using the depth information from the disparity between the separate camera’s.

2.1 Finding and matching good features

We first find a number of corners within each image. The corners are recorded as points on the frame where \( p = (x, y) \). The point is then matched both between camera’s and between frames. First the point is matched to a corresponding point on the other camera. Any point on either camera that is unmatched is then removed as the pair of points will be used to calculate the 3 dimensional vector from the camera to the item. When the points have been matched they are matched to a point on the same camera, but between frames. As with the previous case the unmatched point’s are removed as they are used to calculate the rotation of the camera. Finally the points are matched between camera’s and between frames. This is to ensure the algorithm found the correct point between frames.

The matched points are then stored in the set \( P \) for further processing.
2.2 Calculate the end vectors of the rotation

To calculate the rotation of the camera we first need to know the start and end vectors of the features in $P$. To calculate the vector we take the location of the point in each camera and use the equations:

$$Z = \frac{1}{p_{c,x} - p_{2,x}}$$ (1)

$$X = Z \cdot p_{1,x}$$ (2)

$$Y = Z \cdot p_{1,y}$$ (3)

Where $p_{c,x}$ is the $x$ value of the point in cameras $c$.

These vectors are then paired together to explain the rotation of the camera.

2.3 Change in angle of points

To calculate the change in angle of the camera we first have to project the vector pair found in section 2.2 onto the $XZ$ and $YZ$ planes. The $XZ$ projection will be used to calculate the yaw, while the $YZ$ projection will be used to calculate if the camera is parallel to the floor. To project the vector onto the $XZ$ plane the $y$ component is set to 0. With the $YZ$ plane the $x$ component is set to 0.

Next the vectors are normalized so $|p| = 1$ and the dot-product of the projected vector is calculated:

$$r = p_f \cdot p_{f+1}$$ (4)

where $r$ is the dot product, $p_f$ is the projected vector. As unit vectors are used the dot-product will be the cosine of the angle between $p_f$ and $p_{f+1}$.

The calculation of the angle, $\theta$, between the vectors is

$$\theta = \arccos\left(\frac{r}{|p_f||p_{f+1}|}\right)$$ (5)

As $|p_f| = 1$ and $|p_{f+1}| = 1$ equation 5 can be simplified to

$$\theta = \arccos(r)$$ (6)

The calculations are then repeated with the projection on the $YZ$ plane to calculate the change in pitch.

2.4 Generation of the Map

To generate a map of the room the depth map was created using the disparity of the stereo cameras. This will create an grey scale image where the intensity related to the distance the object is from the camera.

The images are then converted to a set of vectors representing the location of the object. These vectors are rotated to the camera’s current offset. This is because the vectors are around $(0, 0, 0)$ where, except in the first frame, the camera is not pointing directly towards this vector.

When enough frames have been captured the vectors may then be sent to an algorithm like the Sparse Bundle Adjustment described in [13]. This can then be used to generate a mapping of the room.

3 Experiment

An experiment was setup to test the software to map a wall and calibrate the calculations. The camera used was a Point Grey Research Bumblebee 2 and was placed on a tripod. One image for each camera was captured per frame for offline analysis. The captured images were 640 pixels wide by 480 pixels high.

A wall in the Computer Vision lab at the University of Canterbury Department of Computer Science and Software Engineering was setup with the camera and a checkerboard. The checkerboard was used to help with the calibration of the camera and to increase the number of points available to the software. Four frames were captured rotating the camera $18^\circ$ anti-clockwise between each frame.

3.1 Calibration

Figure 2 shows the four images captured by the camera in the first two frames used to calibrate the camera. The software then computes the optical flow of each of the points in the camera’s producing Figure 3. The mean angle for the first two frames was 3.11 (sd. 0.98). There was a single data point outside the standard deviation. Then this outlier was removed the mean angle was still 3.11 however the standard deviation was reduced
4 Conclusion

It has been shown that it is possible to generate a map using stereo camera’s from a fixed location. The method is able to generate a map of the wall in a room from a series of frames.

In the future the proposed method will be generalized to allow for a free moving robot to take continuous footage as it travels around an environment. The use of stereo camera’s can also be merged with other range finding equipment, for example laser’s and sonar equipment.

References


