Stereo Reconstruction

Introduction

Binocular Stereo Vision obtains the 3-dimensional geometric information of an object from multiple images based on the research of the human visual system. The results are presented in the form of depth maps. Images of an object acquired by two cameras simultaneously in different viewing angles, or by one single camera at different times in different viewing angles, are used to restore its 3D geometric information and reconstruct its 3D profile and location.

Binocular stereo vision method requires two identical cameras with parallel optical axes to observe one same object, acquiring two images from different points of view. In terms of trigonometry relations, depth information can be calculated from disparity. Binocular stereo vision method is well developed and stably contributes to favorable 3D reconstruction, leading to a better performance when compared to other 3D construction. Unfortunately, it is computationally intensive, besides it performs rather poorly when baseline distance is large.

Stereoscopic System

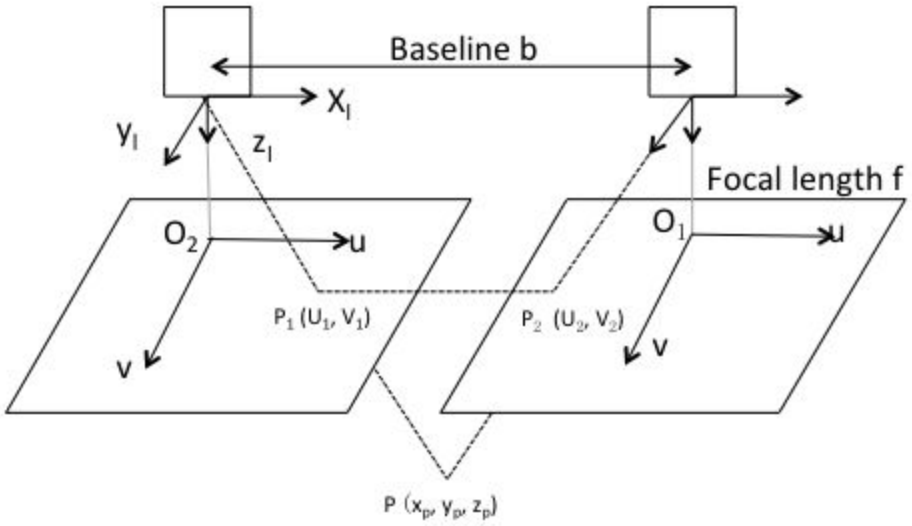


Fig 1 : Horizontally sighted Binocular Stereo Vision

Image acquisition

Two cameras are mounted on a stereo head, which take images constantly.
All the information regarding the cameras is known beforehand.

Given below are the camera parameters :

1. **Intrinsic parameters**
   - Principal point coordinates
   - Focal length
   - Pixel magnification factors
   - Skew (non-rectangular pixels)
   - Radial distortion
2. **Extrinsic parameters**
   - Rotation and translation relative to world coordinate system

Camera calibration

From the above information, camera calibration is done.

Camera calibration in Binocular Stereo Vision refers to the determination of the mapping relationship between the image points P1(u1,v1) and P2(u2,v2), and space coordinate P(xp, yp, zp) in the 3D scenario. Camera calibration is a basic and essential part in 3D reconstruction via Binocular Stereo Vision.

In my implementation, I have used Horizontally sighted Binocular Stereo Vision.
Hence for each pixel P(x,y) in left image, the corresponding pixel will be at the same height but at some different column, i.e. P(x,y) is mapped to P(x',y) in the right image.

Feature extraction

The aim of feature extraction is to gain the characteristics of the images, through which the stereo correspondence processes. As a result, the characteristics of the images closely link to the choice of matching methods.

In my implementation, I have not used feature extraction.
The disparity estimation/ stereo correspondence algorithm processes the whole image.

Stereo correspondence / Disparity Estimation

Stereo correspondence is to establish the correspondence between primitive factors in images, i.e. to match P1(u1,v1) and P2(u2,v2) from two images.

In my submission, I have implemented two methods.

1. **Window Based Search Method** : Algorithm runs once for each pixel in the left image and finds the corresponding pixel in right image by finding the minimum intensity difference between 9*9 windows in both images along the same horizontal line.

2. **Maximum Likelihood Stereo Algorithm** : Algorithm runs for each row in the and finds the set of correspondences that maximize the cost function subjecting to ordering and uniqueness constraints.

Restoration

According to precise correspondence, combined with camera location parameters, 3D geometric information can be recovered without difficulties. Due to the fact that accuracy of 3D reconstruction depends on the precision of correspondence, error of camera location parameters and so on, the previous procedures must be done carefully to achieve relatively accurate 3D reconstruction.

In my implementation, I have not triangulated the mesh. The application displays the point cloud data directly by mapping the left image on the generated data.
The depth estimation is done using the given formula :

$$Z = \text{baseline} * f / (d + \text{doffs})$$

Fig 2 : Depth calculation

The parameters are:

- f : Focal length of the cameras
- d : Estimated disparity by stereo correspondence algorithm.
- doffs : Absolute distance between the cameras in horizontal direction.
- baseline : Absolute distance between the camera lenses in horizontal direction.

Before triangulation using various methods, noise has to be removed from the point cloud data to get a smooth mesh.
Then triangulation can be done using various methods like Marching cubes or Delaunay triangulation.

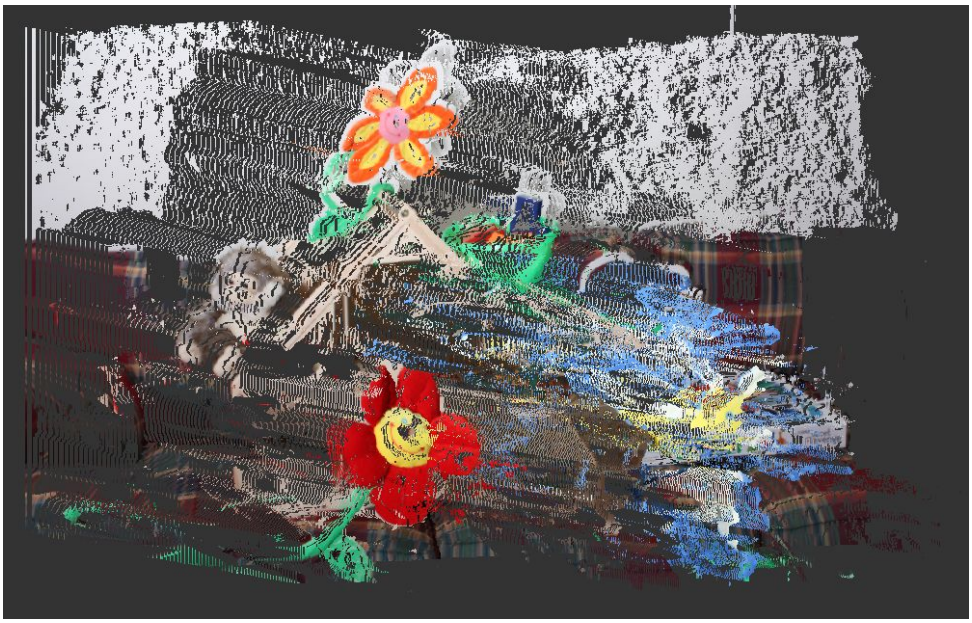
Results

As I was unable to find stereo image data for face with ground truth, I have used the dataset from https://vision.middlebury.edu/stereo/data/scenes2014/

Image 1 :

Point Cloud Outputs :

1. Using 9*9 Window based disparity estimation :



2. Using Maximum Likelihood Stereo Algorithm :

Disparity Map Outputs :

1. Using 9*9 Window based disparity estimation :

2. Using Maximum Likelihood Stereo Algorithm :

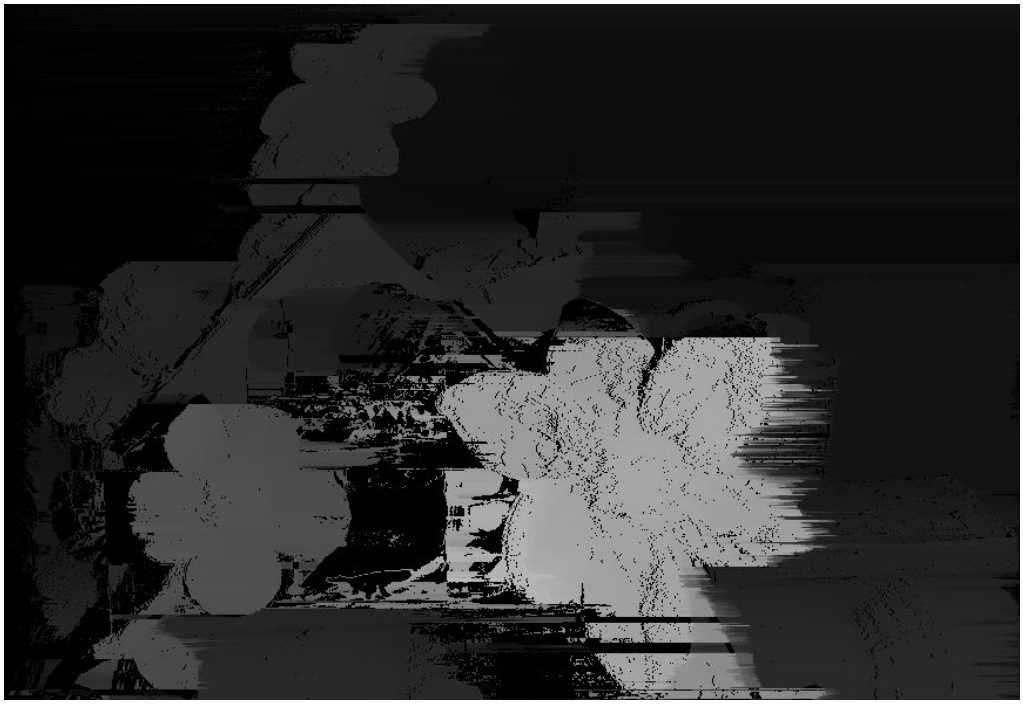
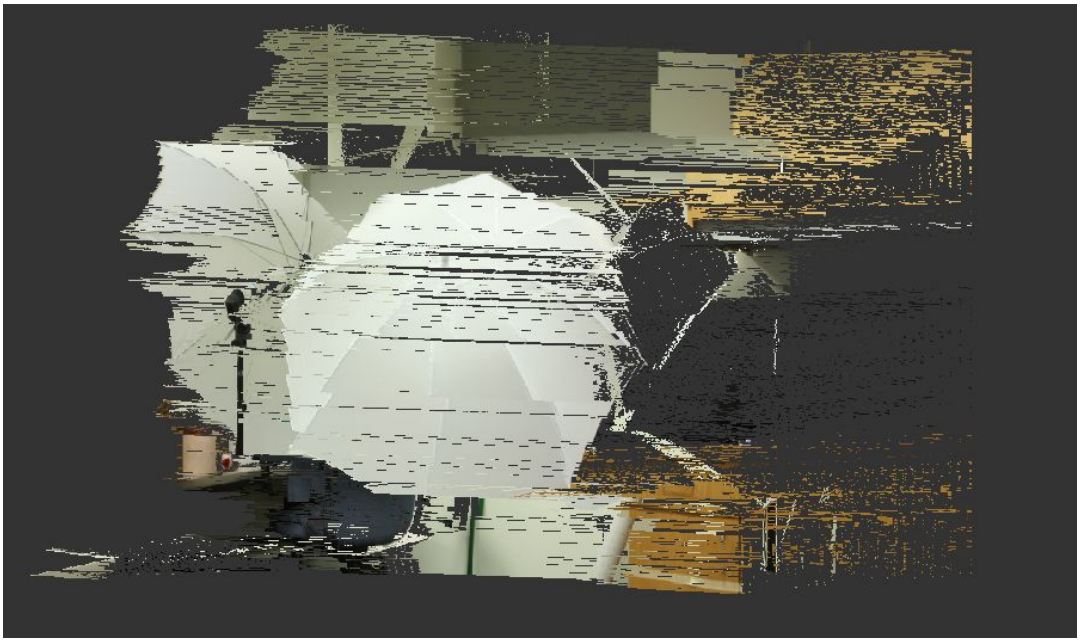


Image 2 :

Point Cloud Outputs :

3. Using 9*9 Window based disparity estimation :

4. Using Maximum Likelihood Stereo Algorithm :



Disparity Map Outputs :

3. Using 9*9 Window based disparity estimation :

4. Using Maximum Likelihood Stereo Algorithm :

Clearly in both cases, the Maximum Likelihood Stereo Algorithm outperforms the Window Based Disparity Estimation Algorithm.

Issues

1. Did not implement point cloud triangulation and feature extraction.
2. Images must already be rectified i.e. they should be made fit to the Binocular Stereo Vision model.
3. Radial Distortion Correction not implemented.

References

Stereo Reconstruction :
9 – Stereo Reconstruction
https://gvv.mpi-inf.mpg.de/files/SIGGRAPH_ASIA_2012/facecap.pdf

Window Based Disparity Estimation :
(PDF) Fast implementation of window-based methods for stereo correspondence

Maximum Likelihood Stereo Algorithm :
A Maximum Likelihood Stereo Algorithm*
https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.707.6410&rep=rep1&type=pdf#:~:text=The%20dynamic%20programming%20approach%20is,estimation%20and%20occlusion%20detection%20simultaneously.

Binocular Stereo Vision :
https://en.wikipedia.org/wiki/3D_reconstruction#Binocular_stereo_vision

Stereo Image Data :
2014 Stereo Datasets