H-V Epipolar Geometry Calibration and Slope Extraction for Advance Depth Camera to Estimate Target Distance

Patel Ronak
Masters of Engineering in VLSI and Embedded System
GTU PG School, Ahmedabad, Gujarat, India

***Abstract*— In this Paper, Post Processing method for distance estimation exploiting the special linear structure of epipolar geometry plane image (EGPI) and locally linear vector stitching (LLVS). Point Distance mapping estimated locally by locating the slope of each linear vector point on EGPIs, which are projected from the corresponding 3D world coordinates. Depth camera calibration process is carried for extracting the depth parameters using special target. Depth vs. slope database is generated from the slope extraction algorithm. Then point distance method for generating complete pixel depth map.**

***Index Terms*- Disparity, Depth Map, Stereo Matching, Parallax Shift**

I. INTRODUCTION

The estimation of the virtual depth from an image can be considered as a multi-view stereo problem since each virtual image point occurs in multiple parallax images. To understand the concept consider an example, when looking out of the side window of a moving car, the distant scenery seems to move slowly while the lamp posts closer to viewer flash by at a high speed. This effect is called parallax, and it can be exploited to extract geometrical information from a scene. From multiple captures of the same scene from different viewpoints, it is possible to estimate the distance of the objects, i.e. the depth of the scene. Advanced Light Field cameras are used to capture the multiple parallax shift images [1].

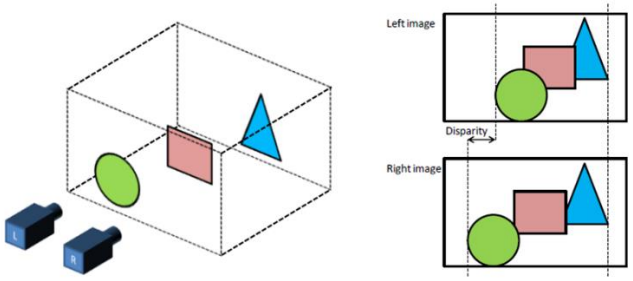


Figure 1. Disparity in a stereo pair. The green circle is closest to the camera, so its location in the image changes the most when the camera is displaced, while the distant blue triangle does not seem to be moving at all.

By tracking the displacement of points between the alternate images, the distance of those points from the camera can be determined. Different disparities between the points in the two images of a stereo pair are the result of parallax. When a point in the scene is projected onto the image planes of two horizontally displaced cameras, the closer the point is to the camera baseline, the more difference is observed in its relative location on the image planes. Stereo matching aims to identify the corresponding points and retrieve their displacement to reconstruct the geometry of the scene as a depth map.

II. RECTIFICATION OF PARALLAX SHIFT FOR DEPTH IMPROVEMENT

For efficient parallax matching, the concept of epipolar lines is essential. This allows for the matching to happen only in one direction, along an epipolar line, greatly reducing the search space. However, this is rarely the case in a realistic capture setting. Camera misalignment makes it impossible to reliably search only on the epipolar lines. This is why a software rectification step is performed after capture. In short, the misalignment of the image planes can be corrected by rotating them in a 3-dimensional space. The rotation matrices doing this can be computed from the images themselves or from specifically captured calibration images.

III. CAMERA CALIBRATION TO REMOVE LENS ABERRATION

A. Parameters from Calibration

Geometric camera calibration, also referred to as camera re-sectioning, estimates the parameters of a lens and image sensor of an image or video camera. These parameters to correct for lens distortion, measure the size of an object in world units, or determine the location of the camera in the scene. These tasks are used in applications such as machine vision to detect and measure objects. They are also used in robotics, for navigation systems, and 3-D scene reconstruction.

B. Ideal Pin Hole Camera Model

The pinhole camera parameters are represented in a 4-by-3 matrix called the camera matrix. This matrix maps the 3-D world scene into the image plane. The calibration algorithm calculates the camera matrix using the extrinsic and intrinsic parameters. The extrinsic parameters represent the location of the camera in the 3-D scene. The intrinsic parameters represent the optical center and focal length of the camera [6].

$$w\,[x\ y\ 1] = [X\ Y\ Z\ 1]\,P$$

Scale factor · Image points · World points

$$P = \begin{bmatrix} R \\ t \end{bmatrix} K$$

Camera matrix · Extrinsics (Rotation and translation) · Intrinsic matrix

Calibration algorithm uses the camera model proposed by assuming pin hole camera .The model includes:

- Lens distortion
- Plot the relative locations of the camera and the calibration pattern
- Calculate the re-projection errors
- Calculate the parameter estimation errors

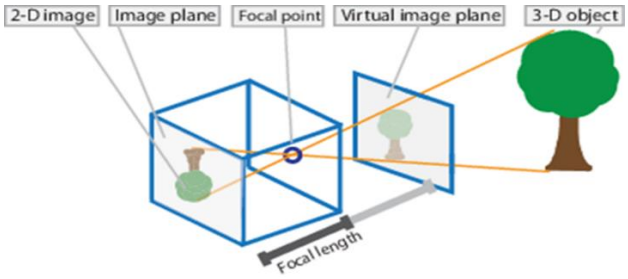


Figure 2. A Pinhole camera model

The world points are transformed to camera coordinates using the extrinsic parameters. The camera coordinates are mapped into the image plane using the intrinsic parameters.

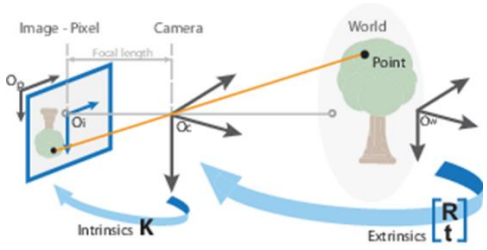


Figure 3. Mapping of world points are transformed to camera coordinates

IV. CALIBRATION EXPERIMENT AND RESULTS

Special target image dataset is used for calibrating camera is shown in below image:

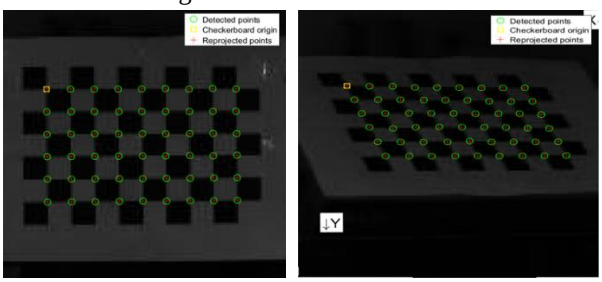


Figure 4. Detection of checkerboard pattern are shown through green circle, the yellow square shows the origin point and the red plus symbols are the re-projection points.

A. *Pattern Centric View of Calibration Result*

Here in the pattern centric view we can see the different position of the checkerboard image that we have used for the calibration.

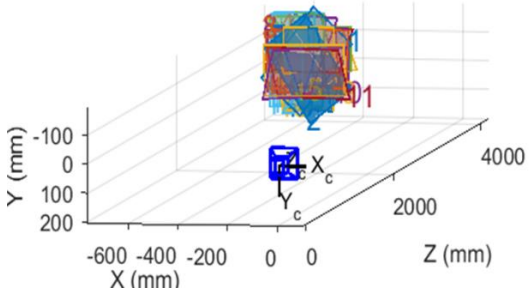


Figure 5. Above figure shows the pattern centric view that is showing the different checkerboard plane

B. *Overall Mean Error results in Pixel*

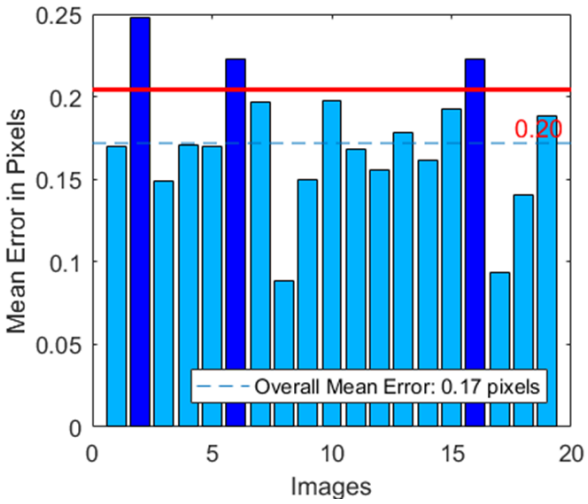


Figure 6. Mean Overall Error Correction in intrinsic Parameters

V. SLOPE EXTRACTION AND DEPTH MAP GENERATION FROM EPIPOLAR

This section shows how EPI blocks are obtained from the raw light field sub images. A small portion of the raw light field image. EPI based scene representation can be obtained when either multiple cameras are in place capturing the same scene from different views or when a single camera is present which captures the scene from different positions. Using the set of multiple view images so captured the EPI image is generated.

The whole concept of stereo matching is based on finding correspondences between the input images. Correspondence between two points is determined by inspecting the n x n pixel neighborhood N around both points. The pairing that has the lowest sum of absolute differences,

Is selected as a corresponding point pair. In practice, a matching block is located for each pixel in an image. The relative difference in the location of the points on the image planes is the disparity of that point. Due to the assumption of being constrained into a 1-dimensional search space, these disparities can be represented as a 2D disparity map which is the same size as the image. Disparity of a point is closely related to the depth of the point. Although in this application the search space can be constrained to be horizontal only (with certain assumptions). The matching block size is one of the most important parameters that affect the outcome of the estimation. Smaller blocks can match finer detail, but are more prone to errors, while large blocks are more robust, but destroy detail [5].

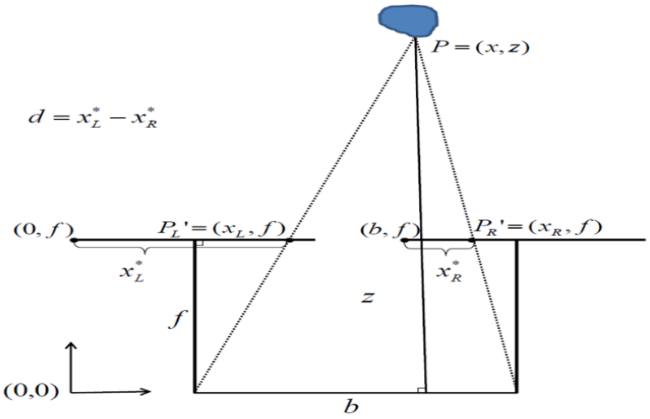


Figure 7. Relation of Extreme pair image with Depth point P(x, z)

To find the pixel in a certain micro image which corresponds to our pixel of interest x_R we search for the minimum intensity error along the epipolar line in the corresponding micro image [8]. If there was no inverse virtual depth observation obtained yet for the pixel of interest x_R, an exhaustive search along the epipolar line has to be performed. For that case the search range is limited on one end by the micro lens border and on the other end by the coordinates of x_R with respect to the micro lens center. A pixel on the micro lens border results in the maximum observable disparity x_R and thus in the minimum observable virtual depth v, while a pixel at the same coordinates as the pixel of interest in the corresponding micro image equals a disparity $P_R = 0$ and thus a virtual depth $v = \infty$.

If there exists already an inverse virtual depth hypothesis $Z(\boldsymbol{x}_R)$, the search range can be limited to $z(\boldsymbol{x}_R) \pm n\sigma z(\boldsymbol{x}_R)$, where n is usually chosen to be $n = 2$.

$$Z(x_R) \sim N(z(x_R), \sigma^2 z\,(xR))$$

In the following we define the search range along the epipolar line as given in following equation.

$$x^s{}_R(px) = x^s{}_{R0} + p_x \cdot e_p$$

Here $\boldsymbol{x}^s{}_{R0}$ is defined as the coordinate of a point on the epipolar line at the disparity $px = 0$, as given below.

$$x^s{}_{R0} = x_R + d \cdot e_p$$

Within the search range we calculate the sum of the squared intensity error e over a 1-dimensional pixel patch $(1 \times N)$ along the epipolar line, as defined in below equation.

$$Distance = \sum_{k=-(N-1)/2}^{(N-1)/2} [I1(xR + ke) - I2((px) + ke)]$$

$$Distance = \sum_{k=-(N-1)/2}^{(N-1)/2} [E * \Delta I(m)]$$

In the above equation to improve accuracy, error parameter E is introduced for different slope conditions. The only parameter required in the above equation is slope (m) or angle which can be extracted from the epipolar image.

From multiple parallax images two types of epipolar geometry image can be generated i.e. horizontal geometry epipolar and vertical geometry epipolar. In Fig. 8 pixel rendering for generating vertical epipolar image is shown. Similarly considering horizontal vectors and rendering it to generate horizontal epipolar [2].

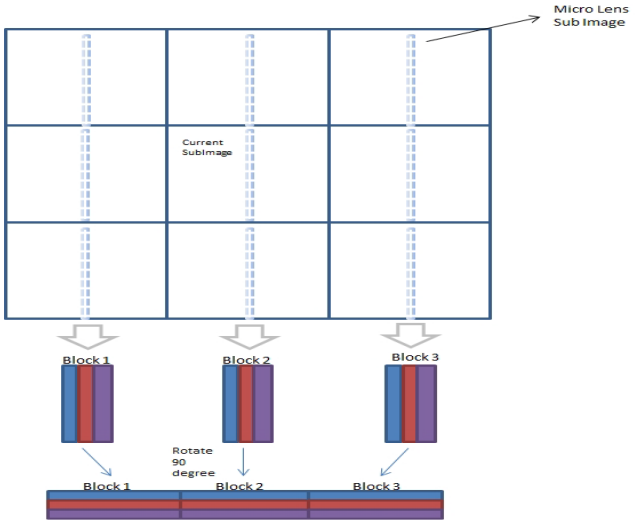


Figure 8. Rendering Raw Image to Generate Epipolar Image

A. *Epipolar Image Slope Extraction*

The below figure is Epipolar image generated as method shown in Fig 9.

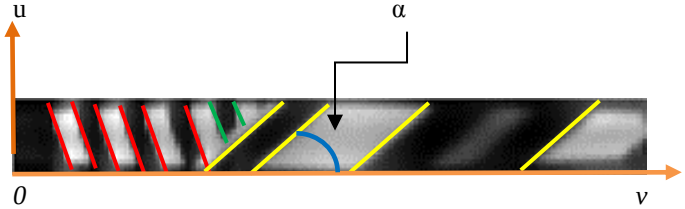


Figure 9. Epipolar image with red line having negative slope, green line having occlusion and yellow line having positive slope

B. Slope – Depth Relation

For establishing depth slope relation according to optical design of the camera by performing an experiment and generated depth-slope database.

In experiment carried for depth calibration, special target image with uniform pattern and opposite color difference is more preferable. In our case we are using checkerboard as target image.

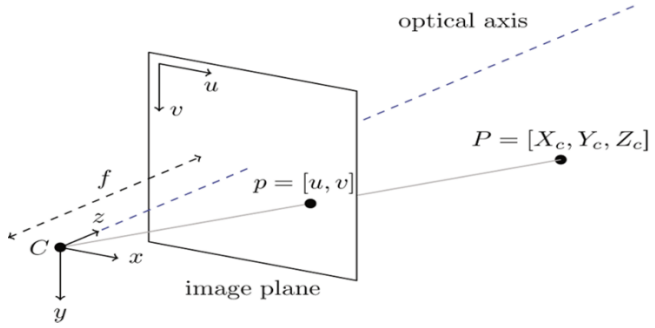


Figure 10. Illustration for camera Depth-Slope Calibration. Here target is move along optical axis with known focus distance. Then recording the depth and slope values for particular distance.

As shown in Fig 10. Target board is moved along the optical axis with finite interval. For particular distance multiple data samples are obtained and taking the average gives the final output slope. The corresponding depth-slope values are recorded in the database. Then Depth Map is generated for given slope of pixel.

Observing Fig .12 show the graph of distance vs. slope. In graph observe the linearity in the distance and the slope for target object before focus distances and target object after focus distances.

VI. CONCLUSION

For depth generation using stereo pair there are many error in the depth map due to non-linearity in the successive approximation distribution shown in Fig. 12. In the proposed method using epipolar geometry error free depth map can be generated. In the proposed method when angle tends closer to right angle i.e. the occlusion point non sharp depth edges are obtained. In the future work to improve the edges of the Depth Map using middle order half profile vector [11].

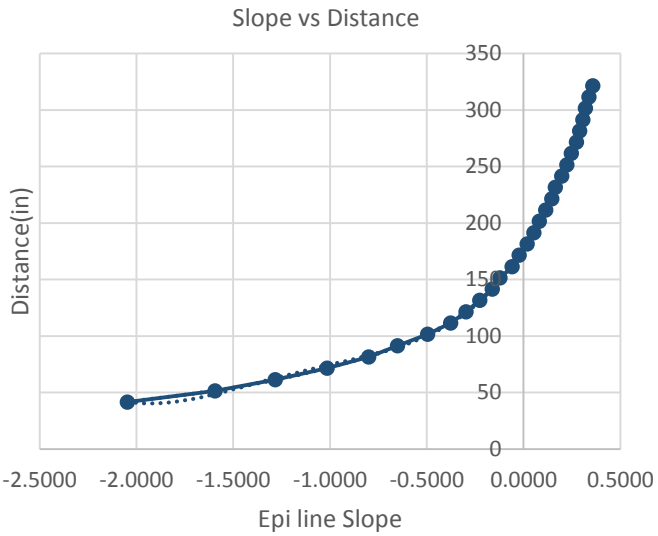
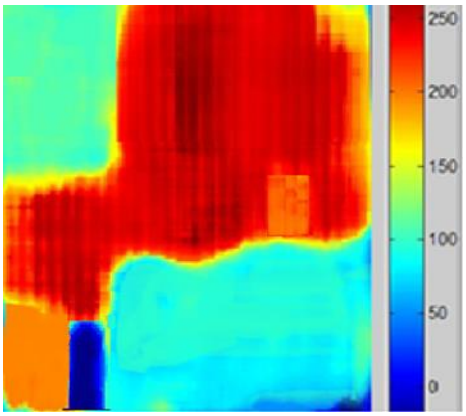


Figure 11. Resultant Graph for Slope vs. Distance from Experiment



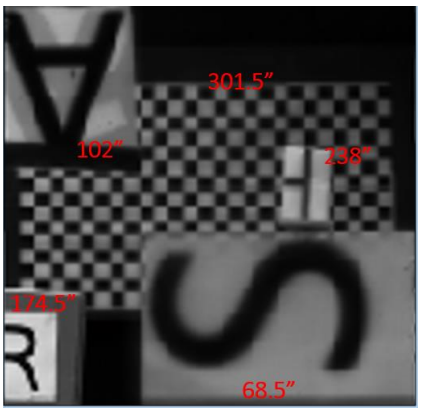


Figure 12. Figure shows Proposed algorithm for depth map generation from the Epipolar image and the estimated target distance from proposed method

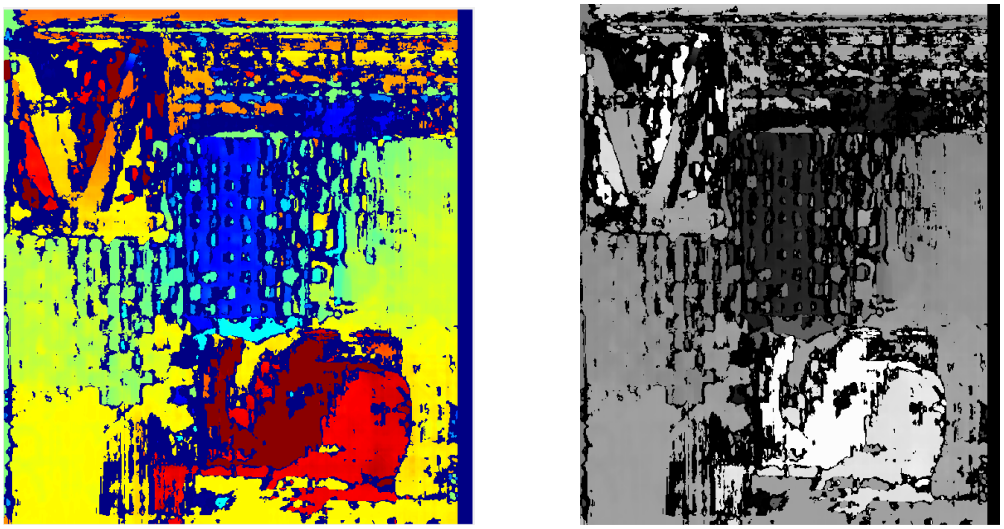


Figure 13. Figure shows the depth map generateed from the stereo pair with successive aproximation distribution error.

REFERENCES

[1] E. H. Adelson and J. Y. A. Wang. Single lens stereo with a plenoptic camera. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(2):99–106, February 1992.

[2] T. E. Bishop and P. Favaro. Full-resolution depth map estimation from an aliased plenoptic light field. In Computer Vision ACCV 2010, volume 6493 of Lecture Notes in Computer Science, pages 186–200. Springer Berlin Heidelberg, 2011.

[3] T. Georgiev and A. Lumsdaine. Depth of field in plenoptic cameras. In Eurographics, 2009.

[4] S. Heber and T. Pock. Shape from light field meets robust pca. In Proc. European Conference on Computer Vision (ECCV), 2014.

[5] A. Lumsdaine and T. Georgiev. Full resolution lightfield rendering. Technical report, Adobe Systems, Inc., 2008.

[6] Bouguet, J. Y. "Camera Calibration Toolbox for Matlab." Computational Vision at the California Institute of Technology.

[7] Bradski, G., and A. Kaehler. Learning OpenCV: Computer Vision with the OpenCV Library. Sebastopol, CA: O'Reilly, 2008.

[8] Levoy M., “Light Fields and Computing Imaging”, IEEE06, August 2006.

[9] Ren Ng , “Fourier Slice Photography” Stanford University.

[10] M. Ito and A. Ishii, “Three view stereo analysis,” IEEE Trans. Patt. Machine Intell., vol. PAMI-8, pp. 524-531, 1986.

[11] S. B. Kang, R. Szeliski, and J. Chai. Handling occlusions in dense multi-view stereo. In CVPR, 2001.

[12] J. Sun, N.-N. Zheng, and H.-Y. Shum. Stereo matching using belief propagation. IEEE TPAMI, 2003.

[13] S. Wanner and B. Goldluecke. Globally consistent depth labeling of 4D light fields. In CVPR, 2012. 1,5.

[14] S. Wanner and B. Goldluecke. Variational light field analysis for disparity estimation and super-resolution. IEEE TPAMI, 2013. 5