Two View Depth Estimation

Aditya Jain , 2020554
Devansh Arora , 2020053
Harsh Goyal , 2020562

Problem Statement

The goal of this project is to estimate the depth an image using Stereo-Camera pair. We will be using classical methods of Computer Vision to generate the depth map of a stereo-pair of images including epipolar geometry , image rectification and disparity computation for generating depth maps.

Dataset Description

- For this project we have used Middlebury Mobile Datasets - 2021 version.
- It consists of 24 datasets of 11 different scenes captured from 1-3 different viewing angles under varies lightning conditions and exposures.
- It utilizes a structured light acquisisition system consisting of a mobile device mounted on a UR5 robot arm.

Depth Map Pipeline

1. Calibration

Detect feature points and draw corresponding lines

2. Rectification

Apply perspective transformation to make images parallel for easy triangulation

3. Correspondence

Appling SSD on each epipolar line to get disparity map.

4. Depth Image

Using disparity info. to calculate the depth map of each image.

Calibration

- First we need to find the keypoints in the two images . For this we have used ORB feature descriptor.
- We have used BruteForce matcher to match the detected feature points and finally stored the X and Y coordinate of these features in two separate list.
- Computed Fundamental matrix using 8 point algorithm by applying SVD on the feature points obtained above.
- Estimating Essential matrix using F-matrix and intrinsic camera matrix. Decomposing it to Rotation and Translation matrices.
- Computing the possible camera alignments of stereo-pair using essential matrix and finding the best alignment based on triangulation check on chirality condition :

Rectification

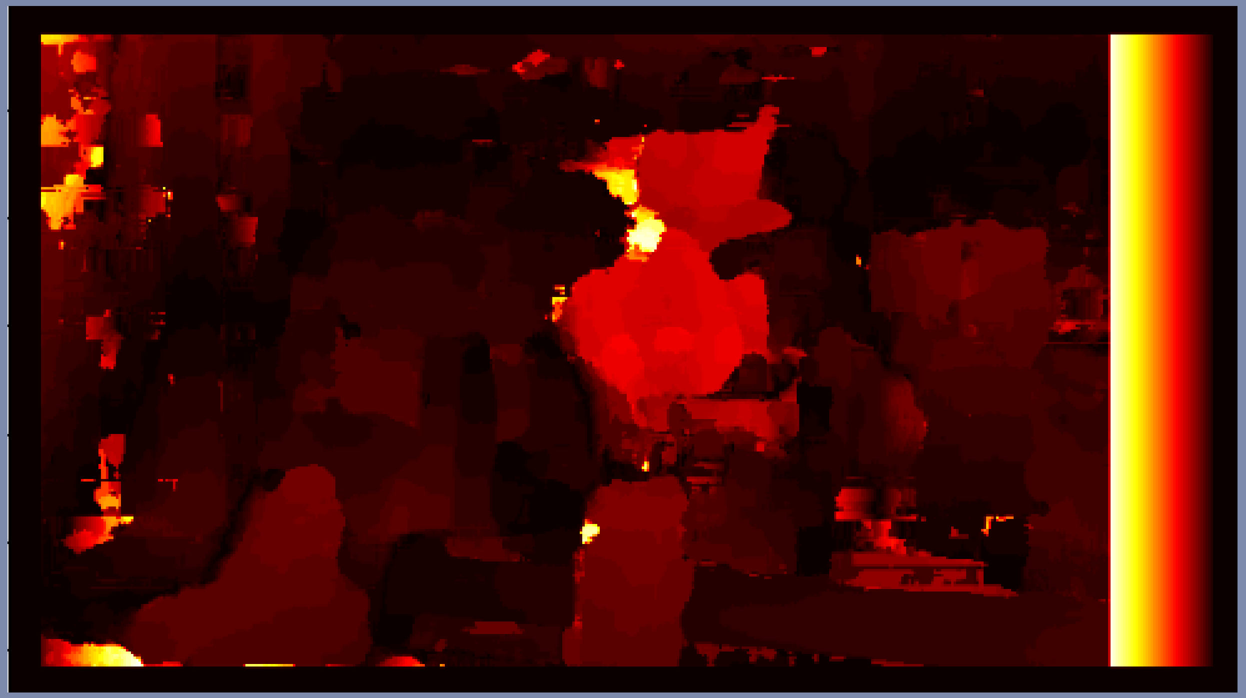
- Applying perspective transformation using StereoRectifyuncalibrated function of opencv library to get the homography matrices H1 and H2.
- Rectifying the feature points using homography matrices H1 and H2 for the two feature point sets obtained in the calibration step.
- Using wrapPerspective function to rectify the image-pair.
- Plot the epipolar lines on both parallel images obtained along with feature points.

Correspondences and Depth Map

- For each epipolar line , applying SSD correspondence and computing disparity map. For this we have used block_comparison method to compare left and right images and computed minimum SSD value for each pixel.
- Computing Depth map from disparity map using depth information for each pixel, camera baseline and focal length.
- Generating both grayscale and color images for both the disparity and depth maps.

Some Results

Artroom



Disparity

Depth

Traproom

Novelty

Problem:

The main issues during disparity matching are

- Finding the right size of matching window
- Dealing with occluded area

Solution:

- Exploiting the epipolar constraints of finding the correspondences at the same horizontal level by enforcing row consistency.
- Smoothening the effect of occlusions

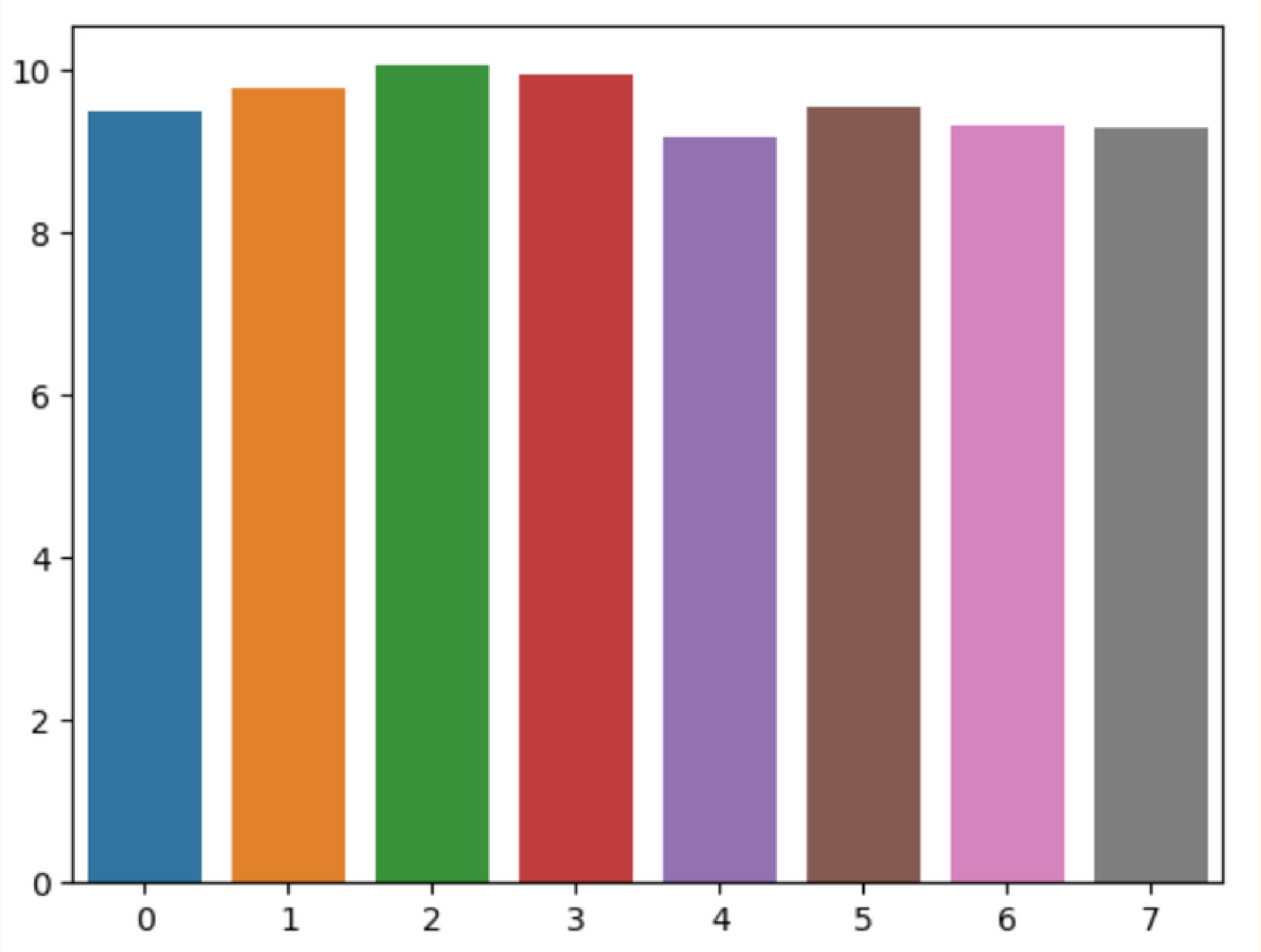
Novelty Procedure

- The procedure for removing inconsistency by matching and warping images remains similar to previous methods. In the next step, we introduce a pairwise matching pixel cost based on the Sum of Squared Differences (SSD) computed along each row of the images.
- We create a cost matrix for every row containing all possible pixel matches between the left and right images. This cost matrix is known as the Disparity Space Image. Using this matrix, we aim to find a path that maps from the end to the start of each row with the minimum sum.
- In this process, we can incorporate row consistency constraints and occlusion information to define the disparity values. We can use these constraints to differentiate between the disparity caused by linearization and the disparity caused by occlusion cost. Finally, the minimum-value path is chosen, and the resulting disparity map is drawn.
- However, this method produces occluded spots in between, which can be smoothed out through additional processing techniques.

Analysis

- For the best camera alignment, the rotation and translation matrix came close to :
  Rotation : [[0.99 -0.02 -0.03], [0.02 0.99 0.006], [0.03 -0.006 0.99]]
  Translation : [-0.2 -0.15 0.96] for all images .
- The RMSE error between the ground truth disparity and obtained disparity image came close to 10.
- With Novelty , the final error shows an average RMSE error of about 9.5

Bar Graph is showing an RMSE error in analyzing 8 images.



Individual Contribution

Aditya Jain	Implementation + Ppt
Devansh Arora	Implementation + Ppt
Harsh Goyal	Implementation + Ppt

Thank you!

Do you have any questions for us?