Two-View Depth Estimation

Group 11
Kunal | 2019430
Nikhil | 2019259
Rohit Kumar Vishwas| 2019269

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Problem Statement

- Estimating the 3D coordinates of objects from two images captured from different viewpoints

- Challenging task due to the differences in the orientations, positions, and scales of the two images and presence of noise and outliers

- Many practical applications in robotics, augmented reality, 3D modeling

Dataset Description and Demonstration

- We are using Middlebury Dataset for Two-view Depth Estimation
- One of the most popular and widely-used benchmarks
- Dataset consists of pairs of images of the same scene taken from two different viewpoints
- Contains intrinsic parameter of camera, baseline, x-difference of principal points and ground truth
- Includes multiple resolutions and color spaces, such as gray, RGB, and infrared

Dataset Description and Demonstration

- Import data online and convert into matrix
- Stored that matrix on drive using pickle
- There are 23 pairs of images and camera intrinsic parameters

Left Image

Right Image

Left Image

Right Image

Methodology

Calculated the Matching points using SIFT (Scale-Invariant Feature Transform) algorithm

Scale-space extrema detection: The algorithm detects potential interest points at different scales by looking for local extrema in the scale space of the image.

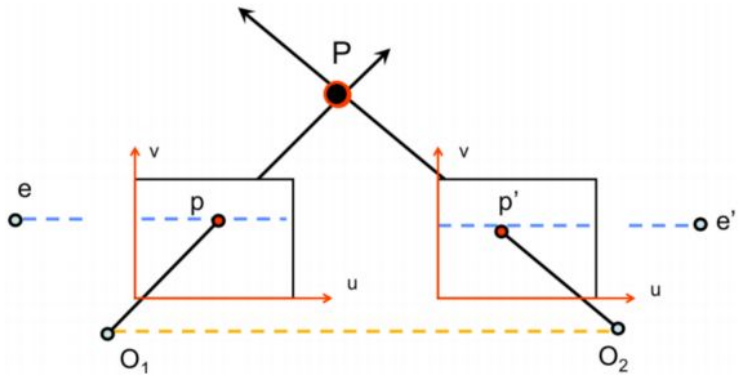
Keypoint localization: It refines the detected keypoints' location based on the scale-space extrema and fits a precise location using a Taylor series expansion.

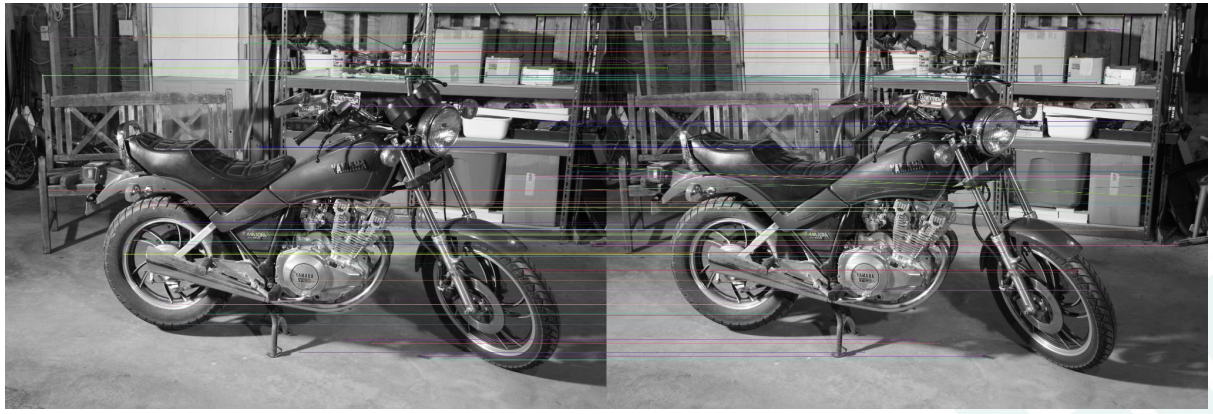
Orientation assignment: It assigns an orientation to each keypoint based on the gradient direction in the image's local neighborhood.

Keypoint descriptor generation: The algorithm creates a unique descriptor for each keypoint by considering the local image gradients' magnitude and direction.

Keypoint matching: It matches keypoints between two images based on the similarity of their descriptors.

Keypoint filtering: It filters out unreliable matches based on various criteria, such as the ratio of the closest and second-closest match distances.





Methodology

Calculated Fundamental matrix:

used SVD to find the F

Used RANSAC (Random Sample Consensus) to find the best Fundamental matrix.

Calculated Essential matrix using Fundamental matrix

find the epipolar constraint first using F

then calculated the E, using epipolar constraint

Computed Camera pose (R and t) using Essential matrix.

using equation [tx]R = E

Methodology

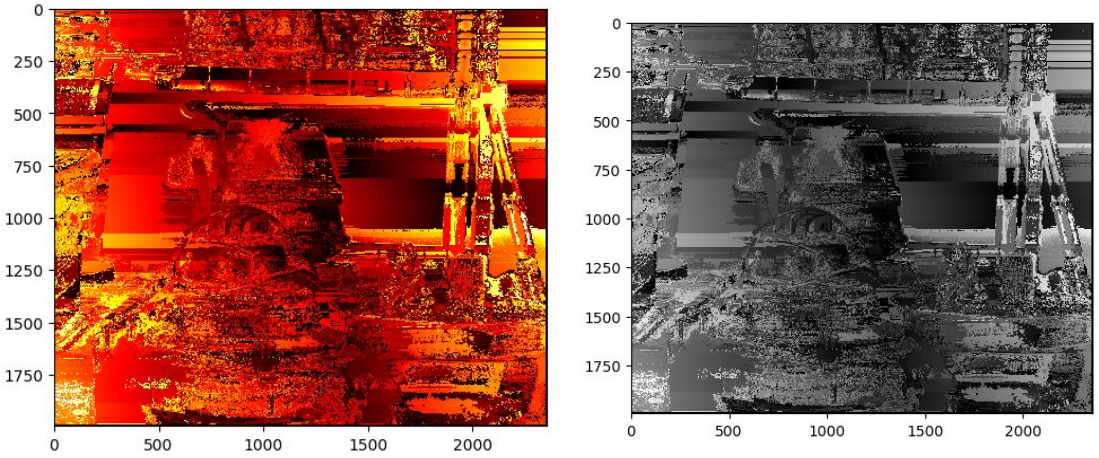
- Mapped 2D points into 3D points with the help of TriangulatePoints
- Made a kernel of window size 5 and iterate over all points of 3D point
- Compute Disparity map
- Compute Depth map

Results and Analysis

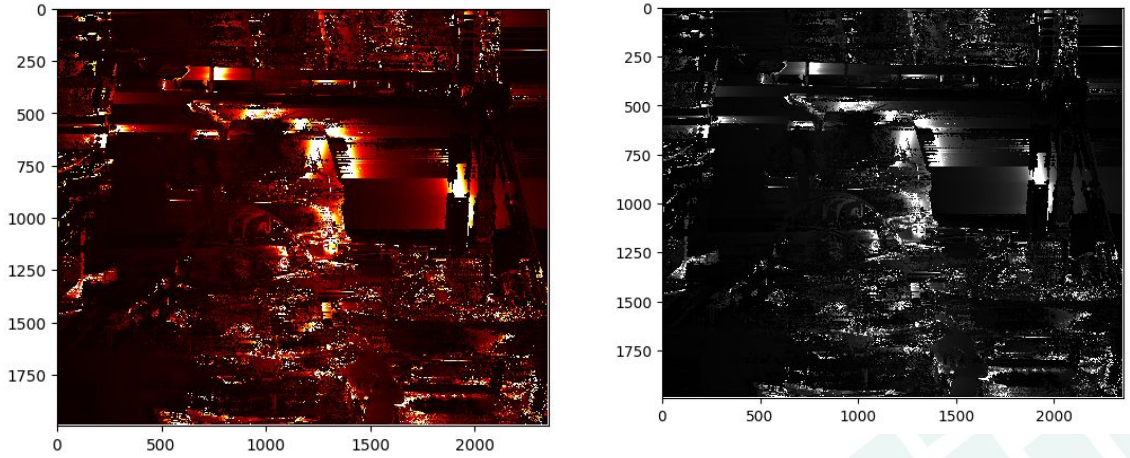
Image 1

Image 2

Image showing those point matches in both images



Images Disparity map for cmap = hot and gray



Images Depth map for cmap = hot and gray

Results and Analysis

Image 1

Image 2

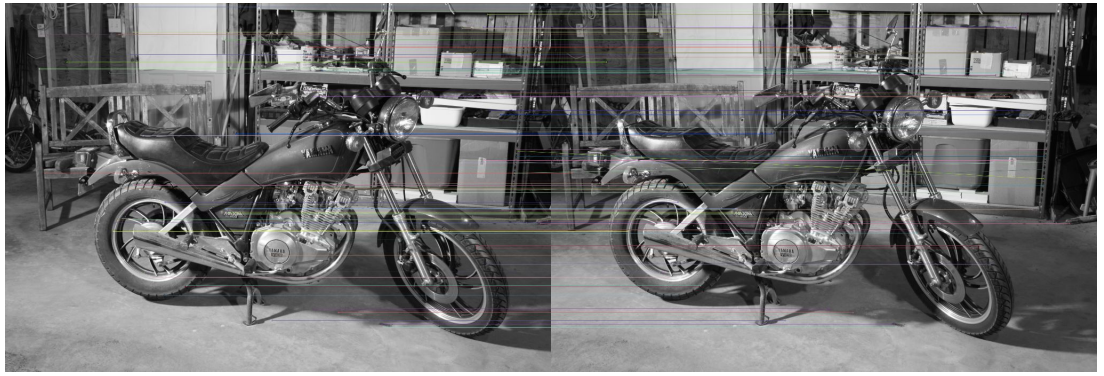
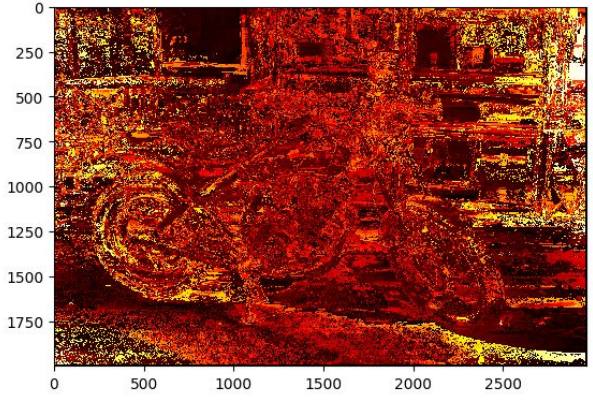
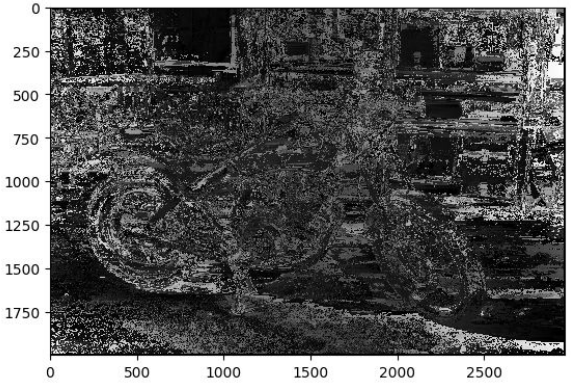
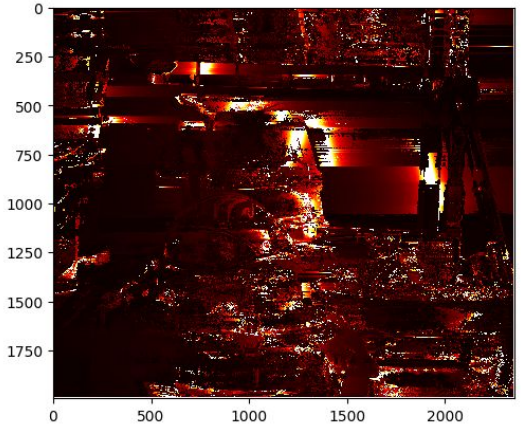


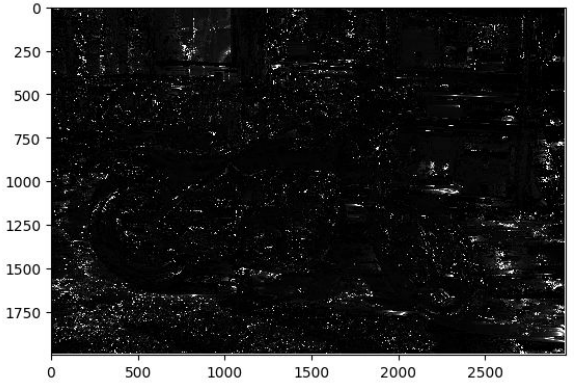
Image showing those point matches in both images





Images Disparity map for cmap = hot and gray





Images Depth map for cmap = hot and gray

Thank You…