Report Project 3 & 4: Stereo Vision-Based 3-D Reconstruction Using Feature Detection and Disparity Mapping

Matteo De Angelis   
*College of Engineering*  
*University of Miami*Miami, USA  
mld163@miami.edu

Carter Falkenberg  
*College of Arts and Sciences  
University of Miami*  
Miami, USA  
ckf25@miami.edu

Kelly Rojas  
*College of Engineering*  
*University of Miami*Miami, USA  
knr20@miami.edu

*Abstract* — Abstract.

Keywords — Keywords

# Introduction

Introduction.

# Theory

## Stereo Vision and Depth Perception

Stereo vision is a method that imitates the human visual system to perceive depth. It analyzes disparities between two slightly different images taken from different angles or points of view. The objects at different depths will have different disparities when projected onto the left and right camera sensors. The depth of the objects can be assessed when computing these disparities [1].

## Camera Calibration and Rectification

Camera calibration is crucial to correct lens distortions and establish the intrinsic and extrinsic parameters of a stereo camera system. The intrinsic parameters include principal point and focal length, whereas extrinsic parameters define the orientation and relative position between the two cameras. Camera calibration provides accurate disparity, which leads to precise depth estimations [2].

Rectification turns stereo images into an aligned form where the corresponding points lie on the same scanline. This method helps to simplify the disparity computation. The transformation is accomplished by rectification matrices that adjust the images, where epipolar lines become parallel [2].

## Feature Detection and Matching

Feature detection is significant in stereo matching because it identifies key points in an image which are distinctive and also invariant to rotation, scale, and illumination changes. Some of the feature detection methods include the following:

### Harris corner detector

Identifies areas with significant intensity variations

### Scale Invariant Feature Transform

Extracts key points based on the differences of Gaussian at multiple scales [3].

### Speeded Up Robust Features

Faster alternative to SIFT, and which is based on integral images and the Haar wavelet responses [4].

## Disparity Computation and Depth Estimation

Disparity refers to the pixel shift between the corresponding points in the left and right images, which is computed through the following equation:

where and are the coordinates of the matched feature points in the left and right images. Furthermore, the depth is determined using the disparity equation shown below:

whereis the camera’s focal length, is the baseline distance between the two cameras, and is the computed disparity [5]. The larger disparities correspond to closer objects, whereas smaller disparities correspond to farther objects.

## Dense Disparity Mapping and 3D Reconstruction

The dense disparity mapping extends a disparity estimation to all the pixels, not the detected features only. This method uses block matching or semi-global matching (SGM) algorithms to produce a disparity map, which is then visualized as a depth image [6].

After the disparity values are obtained for the pixels, a 3D point cloud representation of the scene is built. Each point is mapped into a 3-D coordinate system. As a result, visualization and analyses of the reconstructed environment can be completed [6].

# Methodology

## Task 1 – Multimethod Feature Detection for Stereo Image Analysis

In this task, key point detection and feature extraction were performed on rectified stereo images using multiple feature detection methods. The workflow involved feature extraction from given images, merging features from different detectors, applying non-maximum suppression, and visualizing the selected key points.

The input images were read from the specified path using MATLAB’s imread function. These images served as the basis for feature extraction and analysis. The pixel area parameter was set to control the suppression of weak features in the later steps.

To extract robust features from the image, three feature detection algorithms were implemented. These included Scale-Invariant Feature Transform (SIFT), which detects key points based on the difference of Gaussians and extracts the scale invariant features. The second feature detection algorithm was Speeded-Up Robust Features (SURF), which is more efficiently computationally compared to SIFT, based on the Hessian matrix approximation. Finally, the Harris Corner Detector helped with detecting corner-like regions based on autocorrelation matrix analysis. The three algorithms were applied to each image independently with a defined number of strongest features, which were stored for additional processing.

After the features from SIFT, SURF, and Harris detectors were obtained, they were merged using a mergeFeatures function on MATLAB. This step ensured of creating a comprehensive feature set, which included corner-like and blob-like structures in the images. Moreover, nonMaxSuppression function reduced redundancy, keeping the most significant features only. As a results, weaker key points were eliminated within a pre-defined pixel area, focusing on the most relevant features.

The function selectStrongest selected the strongest features from the combined feature set. The images were then displayed using imshow, and detected features overlaid using the plot function, which provided a visualization of extracted key points.

On the feature extraction pipeline, the display\_features flag was set to false for storage of the detected features without immediate visualization. A pre-defined number of strongest features were extracted from each method to ensure consistency in feature selection. Therefore, the methodology for task 1 ensures a robust feature extraction approach by leveraging multiple detection algorithms, and refining detected key points suppression process and selection mechanism.

## Task 2 – Feature Matching Across Stereo Images for Correspondence Estimation

In this task, feature matching was performed between the rectified stereo image pairs to establish correspondences for further disparity estimation. The workflow included loading stereo images, detecting features, extracting descriptors, matching corresponding features, and visualizing results.

The rectified stereo images were loaded from their respective directories (L\_rectified and R\_rectified). Afterward, each image was converted to grayscale using rgb2gray to ensure consistency in feature extraction and matching. A loop was then employed to process the available stereo image pairs.

The feature points were detected in the left and right images using the retrievingFeatures function, which implemented pre-defined feature detection methods. Moreover, the detected features were used to extract feature descriptors through the extractFeatures function. These extracted descriptors represent the detected key points and are crucial for accurate matching.

The extracted descriptors from the left and right images were compared using the matchFeatures function, which identified correspondences between the features through their similarity. It uses parameters such as match threshold, which was set to 10, controlling the maximum allowable matching distance. Furthermore, the max ratio was set to 0.6 to ensure the best match was better than the second-best match and minimize false correspondences. Afterward, the index pairs corresponding to the matched features were extracted. Finally, the valid matched feature points were stored independently for left and right images.

The showMatchedFeatures function helped to visualize the matched feature pairs along with the montage display mode, representing left and right images side by side with lines connecting the corresponding features. As a result, the visualization shows the effectiveness of feature matching.

The program processes multiple stereo pairs in a loop to ensure the available rectified images are analyzed. Such a methodology balances accuracy and robustness, tuning feature matching parameters and reducing mismatches while keeping a high number of valid correspondences.

## Task 3 – Disparity Estimation for Stereo Correspondences

## Task 4 – Depth Estimation and 3-D Reconstruction from Stereo Disparity

# Results

Results.

Text Box to add picture and description and keep it in the same column together

# Conclusion

Conclusion.

##### References

1. R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, 2nd ed. Cambridge, UK: Cambridge University Press, 2004. [Online]. Available: https://www.robots.ox.ac.uk/~vgg/hzbook/
2. Z. Zhang, "A Flexible New Technique for Camera Calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330-1334, Nov. 2000. [Online]. Available: <https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/tr98-71.pdf>
3. D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004. [Online]. Available: https://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf
4. H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features," in *European Conference on Computer Vision (ECCV)*, 2006, pp. 404-417. [Online]. Available: https://people.ee.ethz.ch/~surf/eccv06.pdf
5. G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library*, 1st ed. Sebastopol, CA: O'Reilly Media, 2008, pp. 377-380. [Online]. Available: https://docs.opencv.org/3.4/dd/d53/tutorial\_py\_depthmap.html
6. H. Hirschmüller, "Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005, pp. 807-814. [Online]. Available: https://www.cvlibs.net/publications/Hirschmueller2005PAMI.pdf