Team 3 (CMPE-297 Computer Vision) Project Report
Stereo Depth Estimation

Authors: Taylor Maurer, Nelson Paz, Marshall Siss
Professor: Sanja Damjanovic

Our project will focus on a past data competition with the main goal of generating disparity maps using self-driving imagery. ArgoVerse AI is a company that is working to generate self-driving cars and released a competition earlier this year to find the quickest and most effective way of creating 3D depth maps. Using their own stereo cameras they generated a huge dataset of rectified stereo images as well as ground truth depth maps generated from lidar sensors. The imagery is mostly of tracking sequences which basically means imagery from cars driving around. Using computer vision algorithms, either deep learning or more classical approaches, the goal is to generate depth/disparity maps using the stereo imagery. Over a given tracking sequence we would then have a stream of depth information. Theoretically this would then be used within the sensor suite for object detection and obstacle avoidance.

The specific dataset is listed on this page: https://www.argoverse.org/av1.html#stereo-link, and contains a total of 113 3D tracking sequences. The stereo images are 2056x2464 (very large!) and sampled at 5Hz. This totals at around 6000 stereo image paris. Additionally of the 113 tracking sequences, 74 of these also has ground truth disparity maps to compare against. If the dataset here proves to be too difficult to deal with there are 'backup' datasets provided by Middlebury.edu of other stereo images located here: https://vision.middlebury.edu/stereo/data/. In addition to being a good plan B to the argoverse dataset they will also allow us to evaluate our methods on other types of imagery.

Generally the typical approach is to use deep learning approaches. In a previous class we explored training and using MobileStereoNet to generate disparity maps. While this approach didn't quite work, it is an area that we can explore more of for this project. As well as Deep Learning and CNN approaches to bypass the need for an extremely high performing computer typical computer vision approaches will also be used. These include some of the following:

1. OpenCV Approach
2. A Hybrid Algorithm for Disparity Calculation From Sparse Disparity Estimates Based on Stereo Vision
3. ORB Based Stereo Depth Map Generation

And then there is another approach of using another lightweight deep learning model , called AnyNet. While some of these approaches may not be used they will be looked at to see which fit the project scope and will give good results. In the end the methods that are evaluated will be compared against each other. Using the ground truth depth and disparity maps.

Work Division:

- Taylor Maurer: Implementation of "A Hybrid Algorithm for Disparity Calculation From Sparse Disparity Estimates Based on Stereo Vision"
- Nelson Paz: Algorithm implementation
- Marshall Siss: Algorithm implementation