ENAE788M Assignment 4a - Estimating 3D trajectory of a stereo sensor

Estefany Carrillo, Mohamed Khalid M, and Sharan Nayak

I. INTRODUCTION

This project involves the estimating the 3D camera trajectory of a stereo sensor. The stereo camera data bag files provided by team BRZ is taken as input for estimating the 3D camera trajectory.

II. FEATURE MATCHING

The first step involves acquiring the left and right images of the stereo sensor. These are got by subscribing to topics "/duo3d/left/image_rect" and "/duo3d/right/image_rect" for the left and right images respectively. Once the images are acquired, feature points are extracted from the left and right images using the Features from Accelerated Segment Test (FAST) algorithm. The feature points from the left and right images are brute forced compared using the L2 norm to find matching features. The matching features are sorted by closest match and then only the first few (5 to be exact) are used for depth estimation.

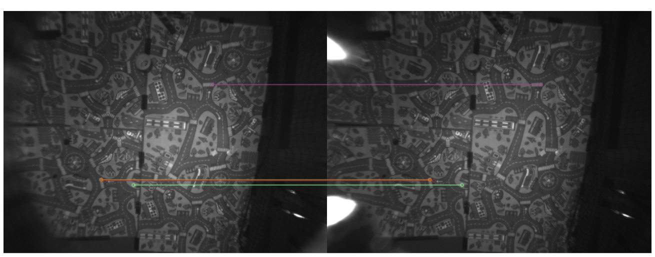

Fig. 1: Features obtained from the FAST algorithm in the two images obtained from the stereo camera.

III. DEPTH ESTIMATION

For depth estimation, the disparity of each of the first five matches is calculated. The disparity is calculated using the equation (1)

$$d = (X_{CL} - X_L) - (X_{CR} - X_R) \quad (1)$$

where X_{CL} is x-pixel coordinate of center of left Image, X_L is x-pixel coordinate of feature point in left image, X_{CR} is x-pixel coordinate of right image and X_R is x-pixel coordinate of feature point in right image. Once the disparity d is calculated, the depth is determined using the equation (2):

$$Z = f_P B/d \quad (2)$$

where Z is the depth, f_P is the focal length in pixel units, B is the baseline distance and d is the disparity. Using the Duo3d MLX datasheet, the focal length f=2mm and the baseline distance d=30 mm. The focal length f_P is calculated as f/w where w is the pixel width (in physical units) obtained as $6\mu m$ from DUO 3D datasheet.

The average of the depth calculated from the best 5 matched points is determined and run through a 3-point moving average low pass filter to filter out high frequency noise. It is observed through testing that were a some bias error of 30 cm associated with the calculated depth. This bias value is subtracted from the filtered Z value to get final Z depth estimate.

IV. OPTICAL FLOW

In order to estimate the linear and angular velocities, the sparse optical flow between two consecutive frames is computed using the Kanade-Lucas-Tomasi Tracker (KLT). First, corner points are detected using the function from OpenCV, *goodFeaturesToTrack*. This function is an implementation of the Shi-Tomasi corner detection, in which we set the maximum number of corners that should be detected to 100, the quality level of detection to 0.3, the minimum distance between detected corners to 10, and the block size to 10. Once we have the corner points, we then pass them to the OpenCV function *calcOpticalFlowPyrLK* along with two consecutive frames to return the positions of the detected corner in the both images. In Figs. 2-4, we can observe the tracks generated from the optical flow computed.

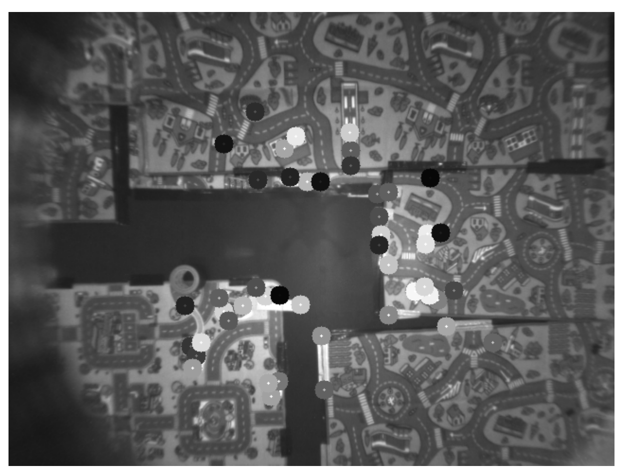

Fig. 2: Corner points detected using the KLT algorithm. Gray and black circles represent these corner points detected in the first left image of the stereo camera.

Once we have the corresponding positions of corner points in the consecutive images, we then compute the delta time Δt by extracting the timestamp from each image. With the