CS9645B: Introduction to Computer Vision Techniques
Assignment #1

Mena SA Kamel
Student ID: 251064703
MESc Candidate, ECE, Robotics and Control
Due Date: February 27, 2020

I. Objective

The purpose of this assignment is to investigate the problem of cross calibration between stereoscopic vision systems and 3D remote eye and gaze trackers. The problem is solved as a Perspective-n-Point (PnP) problem, where a rotation matrix and translation vector are used to perform a rigid body transformation of the points from the stereoscopic vison system frame of reference to the eye and gaze tracker's frame of reference. In addition, the algorithm's noise robustness is studied by introducing zero-mean Gaussian noise to the gaze vector data and calculating the respective re-projection error. The results are discussed in the "Results and Analysis" section. The source code is provided in "Assignment1_Source-code_Mena-SA-Kamel.py".

II. Methods

Since the main goal of this study is to cross-calibrate the stereoscopic depth camera with the gaze trackers, we treat this problem as a Perspective-n-Point (PnP) problem. The aim of a PnP problem is to determine the pose (position and orientation) of a camera given its intrinsic parameters. This requires n points of correspondence between the world points (3D) and their respective 2D projections [1].

Monero-Noguer et al proposed an efficient *O(n)* solution to the PnP problem that achieved better accuracy and noise robustness compared to the other $O(n^5)$ and $O(n^8)$ solutions at the time. Their approach involves expressing the coordinates of the n 3D points as a weighted sum of four virtual control points [1]. For the purpose of this assignment, Monero-Noguer et al's solution implemented in OpenCV's **solvePnP (CV_EPNP)** [2] is used to obtain a matrix M_t, composed of a rotation matrix and translation vector that brings points from the stereo camera's frame of reference to the gaze tracker's frame of reference.

To study the noise robustness of the PnP solution, zero-mean Gaussian noise is added with variable spread, σ equal to *K* % of the magnitude of each gaze vector, g_i. The behavior of the solution is examined for increasing values of *K* by calculating the re-projection error which inherently describes the average squared norm difference according to the following equation:

$$\frac{1}{n}\sum_{i=1}^{n}||P_i - \hat{P}_i||^2$$

Where n is the number of points, P_i is the i^th^ 3D point in the stereo system's frame of reference and $\hat{P}_i$ is the 3D point after adding noise to the gaze vector, g_i and transforming it from the gaze tracker's frame of reference to the stereo system's frame of reference.

III. Results and Analysis

Using the updated assignment data and assuming that the camera has no radial and tangential distortion, the following matrix of extrinsic parameters is obtained M_t:

$$M_t = \begin{bmatrix} 0 & 0 & 1 & -50.00000012 \\ -0.00000001 & 1 & 0 & 0.00000017 \\ -1 & -0.00000001 & 0 & 30.00000063 \end{bmatrix}$$

A. Examining the Matrix M_t

In order to verify the matrix M_t, we translate the world points from the stereovision system's frame of reference to the eye and gaze tracker's frame of reference using the obtained matrix M_t. The obtained coordinates are then perspective projected unto a virtual plane perpendicular to the eye tracker's line of sight at a distance of 1 from the tracker's projection center. The results are shown in figure 1.

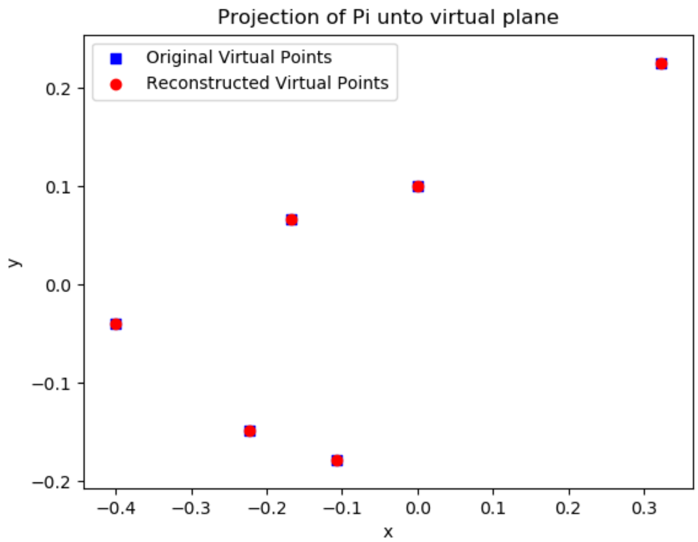

Figure 1 Projecting world 3D points to the eye tracker's 2D virtual plane using the camera extrinsic parameters matrix, M_t and intrinsic parameters M_c.

Initially, the reconstructed points do not exactly match the original points from the gaze tracker. This can be attributed to noisy or inaccurate points in the dataset provided and gives a preliminary indication of the algorithm's sensitivity to noise. This is proven by removing point #4 from the original dataset *((5,-1,40), (-10,1,25))*, and obtaining a new M_t matrix. Figure 2 shows that the new M_t matrix results in no re-projection error. Therefore, since Monero-Noguer et al's solution requires n > 4, removing one point from the data file will still allow us to obtain an accurate M_t.

$$M_{t\_new} = \begin{bmatrix} 0.00000014 & 0.00000006 & 1. & -50.00000039 \\ -0.00000031 & 1. & -0.00000006 & 0.00000308 \\ -1. & -0.00000031 & 0.00000014 & 29.99999385 \end{bmatrix}$$

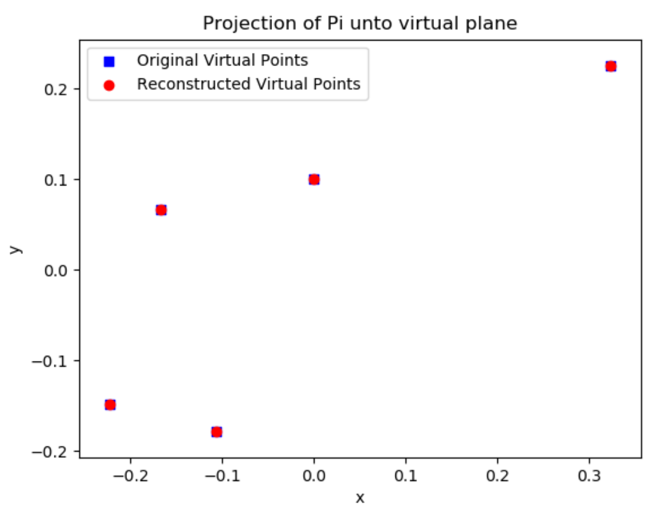

Figure 2 Projecting world 3D points to the eye tracker's 2D virtual plane using the camera extrinsic parameters matrix, M_t and intrinsic parameters M_c after correcting for noisy data points.

B. Examining the Noise Results

It is evident from figure 3 that the average squared norm differences in increasing when *K* varies from 1% to 19% of the magnitude of the gaze vectors, g_i. However, the curve is not smooth and consecutive runs give plots with different slopes, with an overall positive trend.

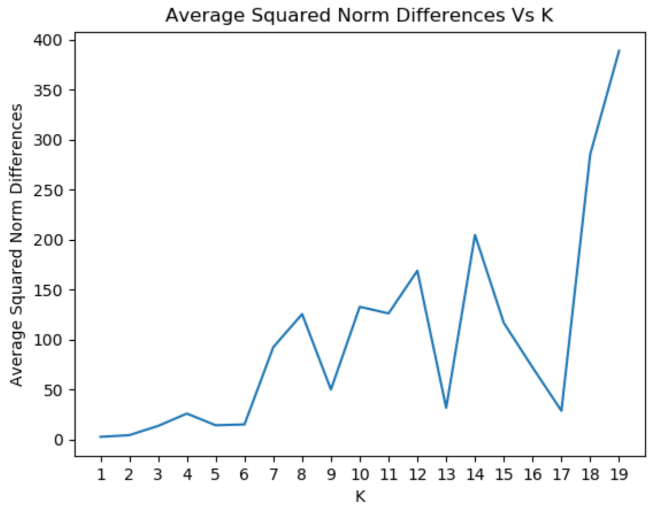

Figure 3 Effect of increasing Gaussian noise spread on re-projection error (1 < K < 20).

In order to determine if the algorithm's error is increasing in a linear or exponential manner, *K* is allowed to vary from 1% to 100%. The resulting average squared norm differences plot is shown in Figure 4. We can therefore conclude that the solution's re-projection error increases exponentially with *K*, with variance significantly increasing with *K*.

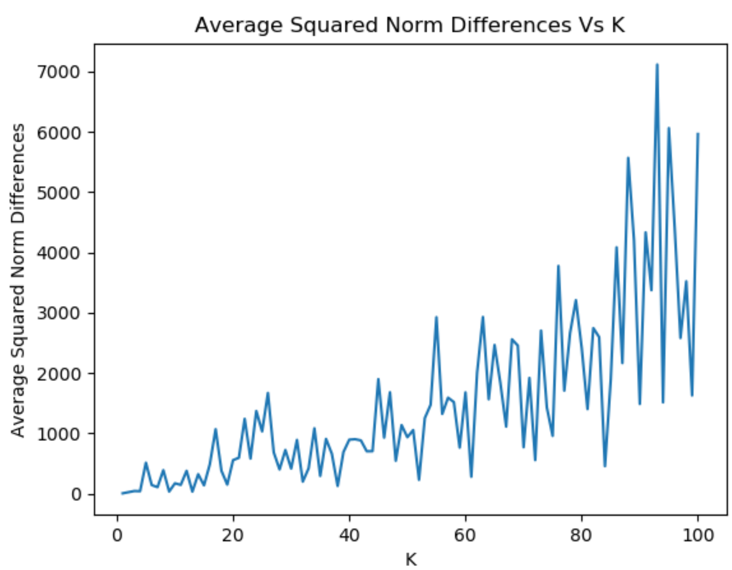

Figure 4 Effect of increasing Gaussian noise spread on re-projection error (1 < K < 100).

References

[1] F. Moreno-Noguer, V. Lepetit, and P. Fua, "Accurate Non-Iterative O(n) Solution to the PnP Problem," *2007 IEEE 11th International Conference on Computer Vision*, Oct. 2007.

[2] "Camera Calibration and 3D Reconstruction," Camera Calibration and 3D Reconstruction – OpenCV 2.4.13.7 documentation. [Online]. Available: https://docs.opencv.org/2.4/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html?highlight=solvepnp. [Accessed: 20-Feb-2020].