Our project focuses on scaling monocular ORB\_SLAM during initialization using a metric sensor, downward facing IR.

ORB\_SLAM initializes with either Homography or Fundamental. Since both the methods are correct up-to scale, we are extracting the translation vector from the two matrices and scaling it using the metric information from IR. We have implemented both the Homography and Fundamental matrices and extracted the translation vector from it. However, in our preliminary tests, the triangulated 3d points don’t look right and the translation vector makes sense but isn’t scaling properly. Figure 1 shows the feature matches used to extract the fundamental then the essential matrix ().

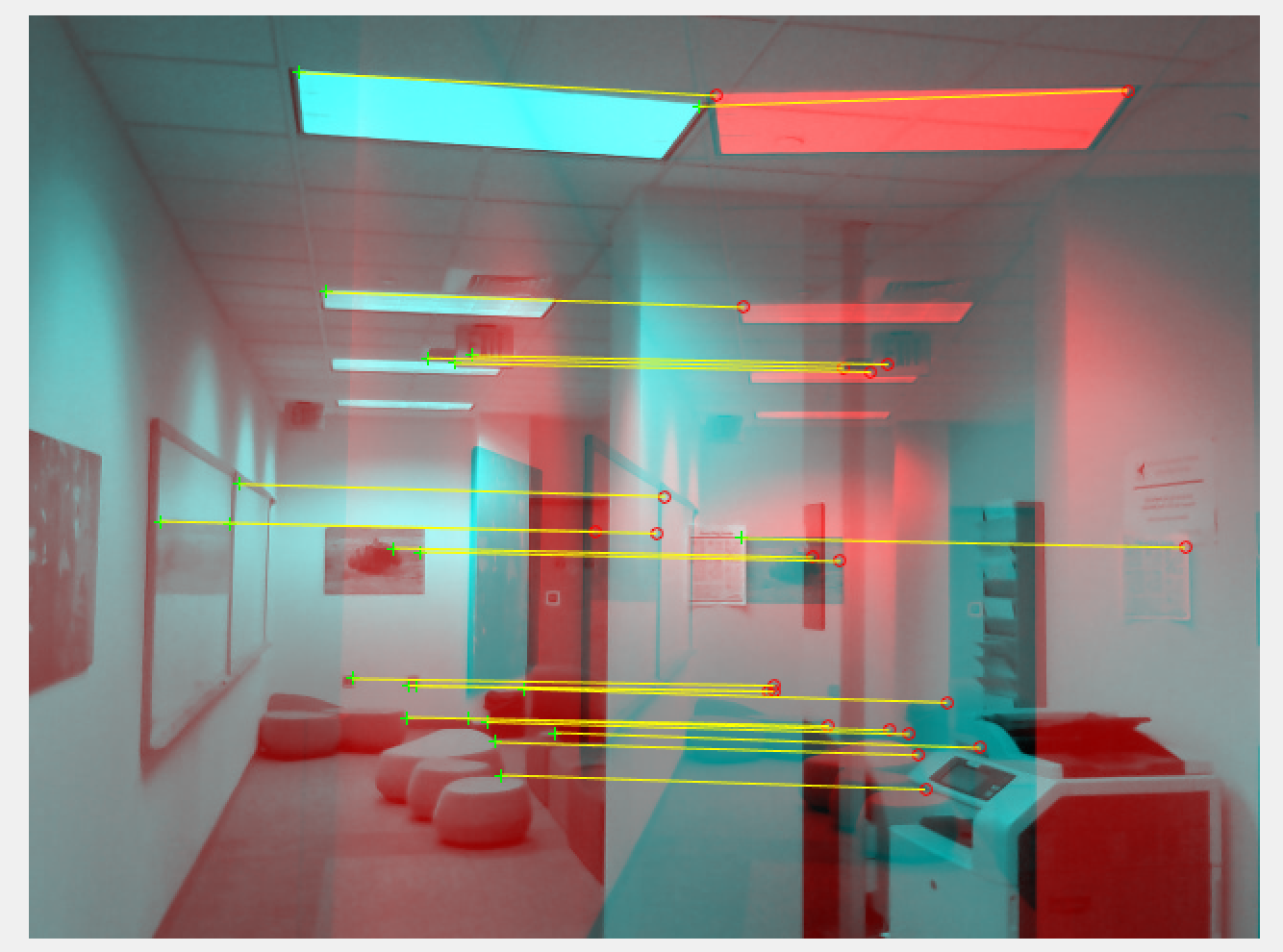


Figure 1: Feature Matches

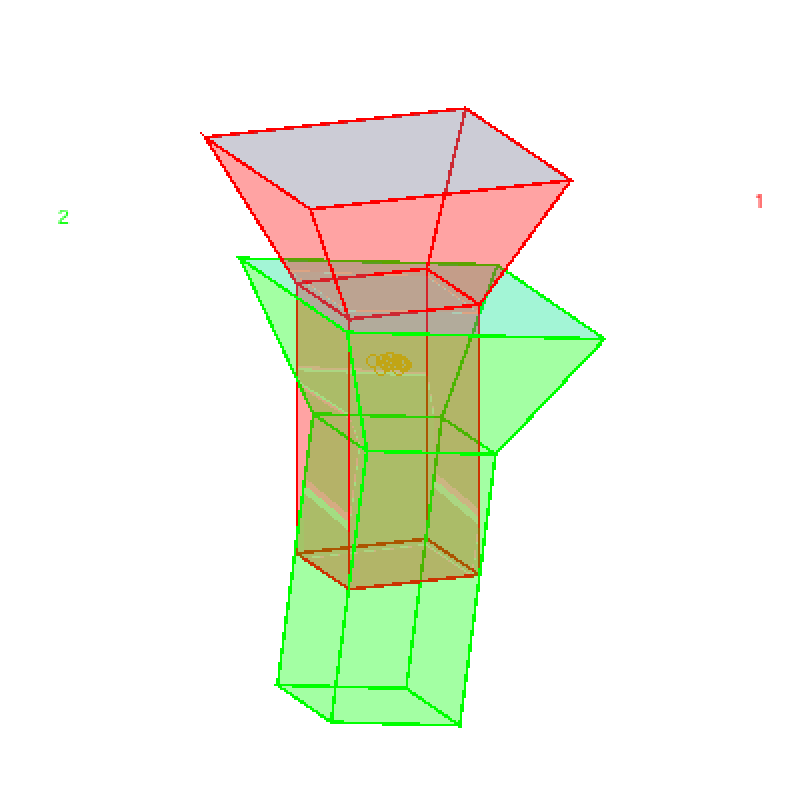
When extracting rotation and translation from the essential matrix, there are four potential transformations, neither of which makes sense. Although the relative motion of the camera makes sense, when triangulated, the points are bizarre. As shown in Figure 2, the relative camera motion makes sensor, but the points aren’t being triangulated correctly. Similar problems exist in the homography case. 

Figure 2: Relative Camera motion with triangulated points

Sensor & Validation Setup

The sensor setup is shown in Figure 4. The sensor setup contains a Kinect RGB-D sensor to enable both RGB-D and monocular SLAM, an UM7 IMU and 2 Sharp distance sensors mounted underneath. The IMU is used to accept or reject distance readings since we need to have near-zero roll and pitch to obtain good distance readings with respect to the ground, and a simple bubble level is used to help the user acquire useful data by keeping the sensor level. We have acquired an initial dataset to perform RGB-D Slam using ORB-SLAM, which will be used as a basis of comparison against our modified monocular ORB-SLAM. RGB-D ORB-SLAM was successful tracking using the depth image from the kinect, shown in Figure 3. We have extracted the key frame map points to create a sparse pointcloud of the scene, and will eventually use that as a basis of comparison.

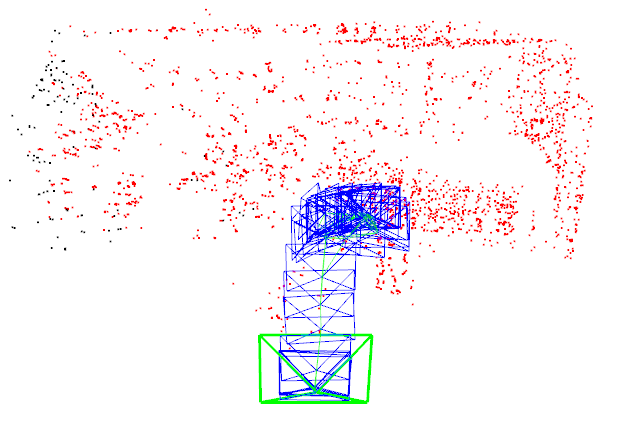
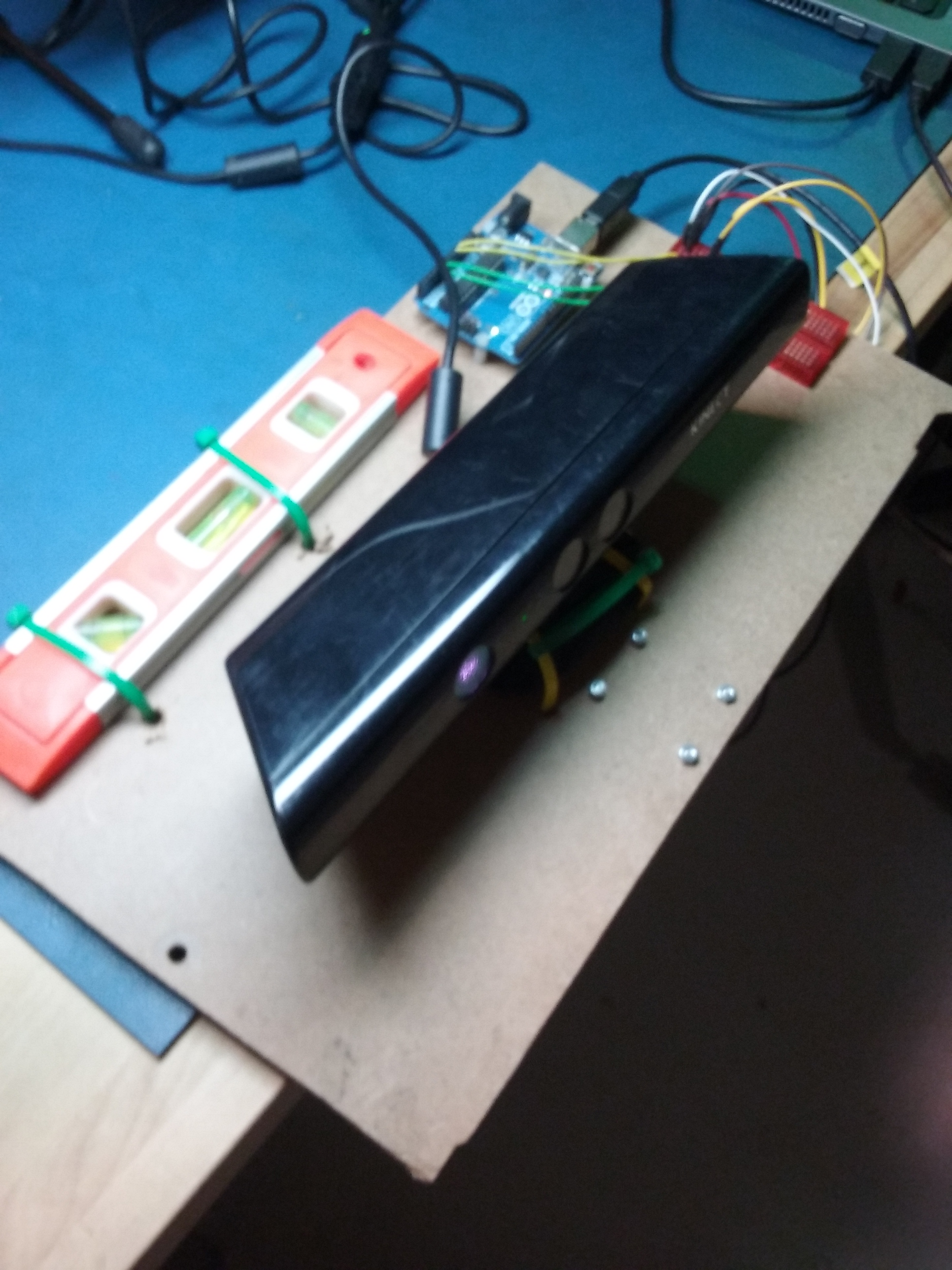
Figure 3: RGB-D ORB-SLAM results using sensor rig

Figure 4: Sensor Rig ( Sharp distance sensors mounted underneath)