Virtual Rigid Body: Design Optimization and Performance Analysis
600.446 CISII Project 14 Final Report

David Lee (dslee@cis.jhu.edu)
Mentors: Alexis Cheng (acheng22@jhu.edu), Dr. Emad M. Boctor (eboctor1@jhmi.edu)

Introduction

In image-guided surgery, optical tracking system fulfills an important task of locating surgical tool in preoperative or intraoperative images for surgical operations. Optical tracking system largely consists of fiducial markers, which are rigidly attached to a tool to be tracked, and camera, which detects the fiducials. Detected positions of the rigid body fiducials serve to define a transformation from coordinate system of the camera to coordinate system of the tool. Using the transformation and some known, or calibrated, positions of the tooltip in tool coordinates, the tooltip can be identified in the camera space.

Several types of rigid body markers have been developed to provide accurate and convenient means of optical tracking. Active markers use sources such as LED to generate visible or infrared (IR) light; pseudo-passive markers have retro-reflective coating to reflect lights emitted from an external source; fully passive markers do not produce or reflect light and rely on visual patterns for detection. Fully passive markers, represented by a commercial system MicronTracker, have addressed problems in active and pseudo-passive markers associated with natural property of light, including quadratic decrease of intensity with increased marker-camera distance, and saturation when the markers are placed close to the camera.

A novel type of rigid body, called virtual rigid body (VRB), has been introduced to further promote accurate and robust optical tracking. The discussed conventional rigid bodies compute the pose of an object via position measurement of physically attached fiducials. On the contrary, VRB system recovers the pose of a projector attached to the object from position measurement of "virtual" light projected onto some surface, for instance, the patient's abdomen, as illustrated below.
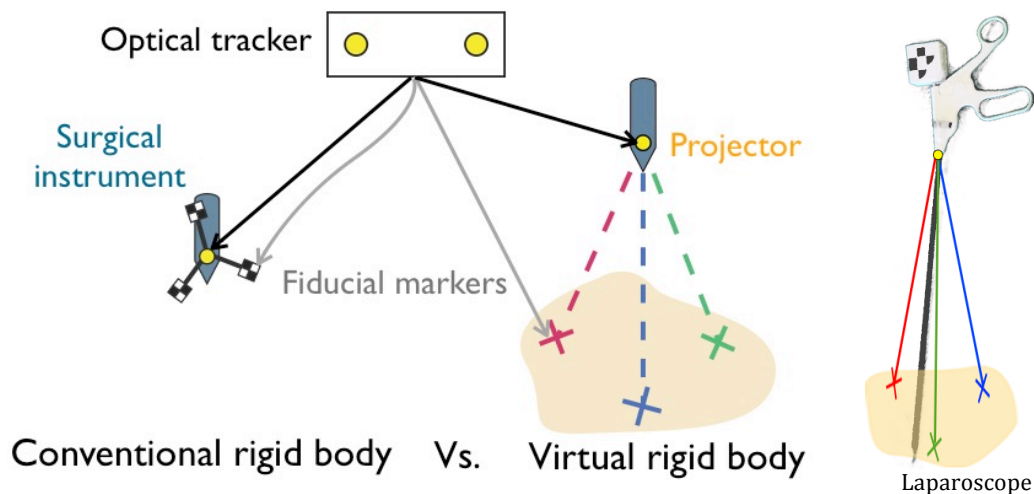
VRB provides a few advantages over the conventional rigid bodies, especially in spatially constrained situations such as laparoscopy. Wider distribution of fiducials, which encourages higher tracking accuracy, can be difficult for conventional rigid body, because physical attachments interfere with surrounding surgical environment. However, the only physical component of VRB is a projector, whereas the projected patterns that serve as fiducial markers do not interfere in any way with physical space. By projecting a large pattern within the camera's field of view, therefore, higher accuracy can be promoted. Furthermore, rigid bodies require the line of sight between the camera and fiducial markers; occlusion of a physical marker by, for example, surgeon's hand therefore prevents detection of the marker. In the case of VRB, the pose can still be recovered from projection on the surgeon's hand, instead of on the original surface of interest. Facile implementation of redundancy adds to the robustness of VRB. Finally, because pattern is projected in the direction to where surgical procedure is to be performed, the camera can always view both the target location and the virtual fiducial markers, presenting a more natural and useful way of tracking.

Problem

While Cheng et al. reported comparable accuracy with conventional rigid body upon introducing the concept of VRB, VRB's performance has not been extensively assessed. In this light, the current project aimed to perform three main tasks. First, an experiment and analysis pipelines were developed to conduct repeatable and consistent assessment of VRB as well as conventional optical tracking system. Second, geometric configurations of VRB have been designed to identify optimal conditions for VRB application. Third, using the developed pipelines and VRB designs, tracking error of VRB was evaluated and compared with that of conventional physical rigid body for simple translational and rotational motions.

Approach

Virtual rigid body grid projection/detection

VRB configuration selection
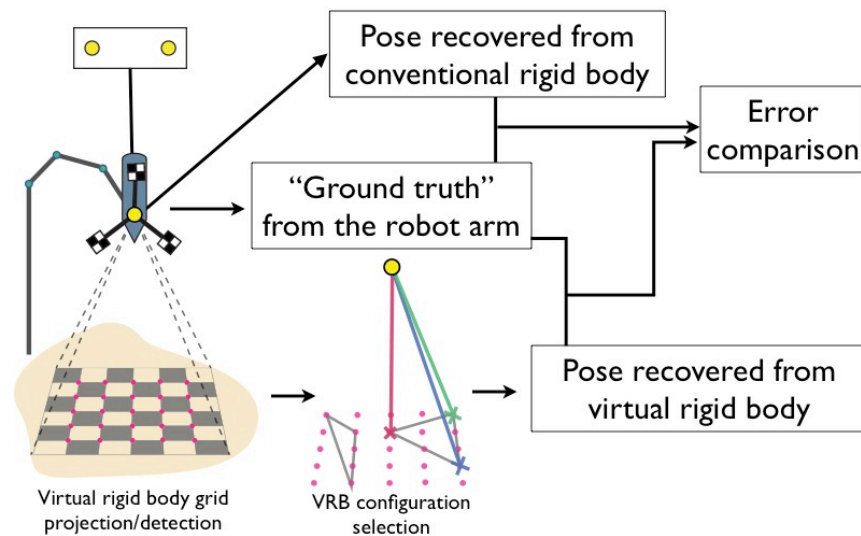
Overview

The overall approach is described in the diagram above. On the functional level, three streams of poses are recorded: "robot", "marker" and "VRB". "Robot" (Universal Robot

5) defines series of poses that serve as the ground truth for optically tracked poses. A series of poses is measured from physical rigid body fiducial "markers", and another series of poses of the projector is recovered from "VRB". Both the projected patterns for VRB, and physical markers for conventional rigid body are detected by a commercial optical tracking system, MicronTracker. Deviations of the optically tracked ("marker" and "VRB") poses from the robot poses are defined as the tracking error. Then, "marker" and "VRB" errors are compared for different set of VRB configurations under translational and rotational trajectories.

To achieve this functional goal, the project pipeline consists of two large parts: experiment and analysis. Experiment pipeline serves to acquire data, and is composed of three subparts: 1) Projected VRB grid, 2) MicronTracker interface for measurement of VRB grid position and physical marker pose, and 3) robot programming for reference pose measurement and precise motion trajectories. Analysis pipeline serves to process acquired data for tracking error evaluation, and is composed of three subparts: 1) data processing, 2) VRB design selection and 3) tracking error calculation.
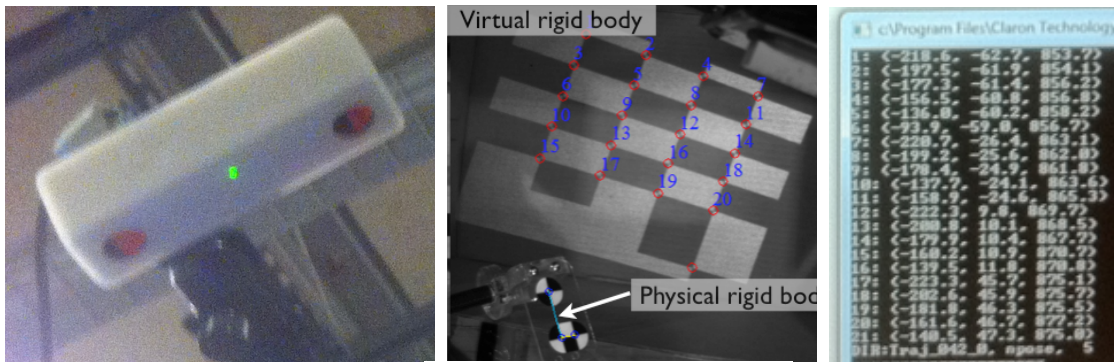
Experiment Pipeline
  1) Projected VRB grid

Optical tracking using VRB begins with projecting some pattern. A pico-projector (Showwx, Microvision) connected to a laptop displays the grid pattern on the right. Projecting a grid pattern enables determining and evaluating different VRB designs after performing a single experiment. Therefore, noise factors that may randomly arise when examining data across multiple trials are controlled.
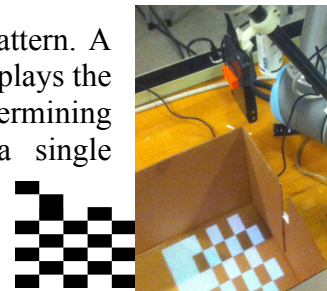
VRB grid, projector

2) MicronTracker Interface

Commercial optical tracking system MicronTracker detects cross-points of "checkerboard" pattern, therefore both the physical marker and projected pattern, controlling the effect of segmentation algorithm on tracking performance of conventional and virtual rigid body. Using MicronTracker's algorithm and its demo software, C++ based optical tracking data acquisition interface was developed.
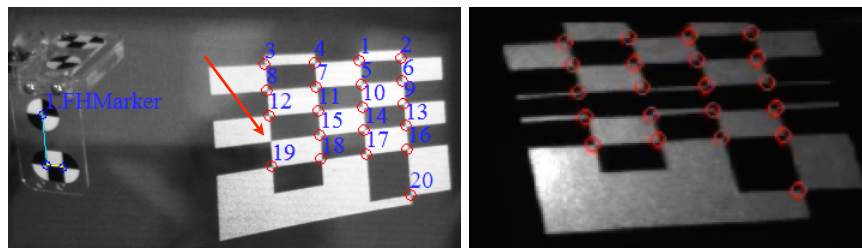

MicronTracker Interface

From the MicronTracker interface, two sets of data are acquired. First, the *pose*, the position *and* orientation, of physical marker is directly obtained by using a predefined marker, a unique group of checkerboard patterns. The output has the format of [x, y, z, R_x, R_y, R_z], where the first three entries denote position, and last three rotations around the axis denoted by the subscript. Example physical marker is provided on the right. Second, the *positions* of the cross-points of the projected checkerboard are recorded, as shown above. The output has the format of [x, y, z]. A subset of these points is selected according to desired VRB

Marker

design in the analysis pipeline, and is used to estimate the pose of the projector. Different designs of VRB will be related to accuracy of the estimated projector pose.

MicronTracker has some limitations. First, MicronTracker does not guarantee detection of all cross-points on the VRB checkerboard grid in a single frame of image. Also, due to the shutter speed of the MicronTracker camera and refresh rate of projector, flickering artifacts are often detected. Most importantly, MicronTracker depth perception is considerably noisy. Therefore, at a single pose of the projector, 100~1000 frames were collected for single pose.


Undetected point   Flickering

MicronTracker yet holds significant merit by establishing a unified platform for consistently examining the performance of two types of rigid bodies.

3) Robot (UR5) component

The 6-joint UR5 robot arm provides simple means of programming precise translational and rotational motion trajectories via a GUI controller. Four types of motion trajectories have been programmed: in-plane translation, normal translation, screw rotation and swinging rotation, as illustrated in the diagram below.

in-plane translation   normal translation   screw rotation   swinging rotation

UR5 and projector
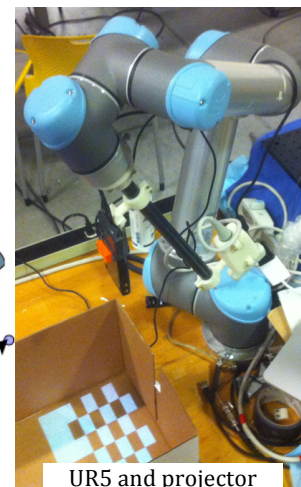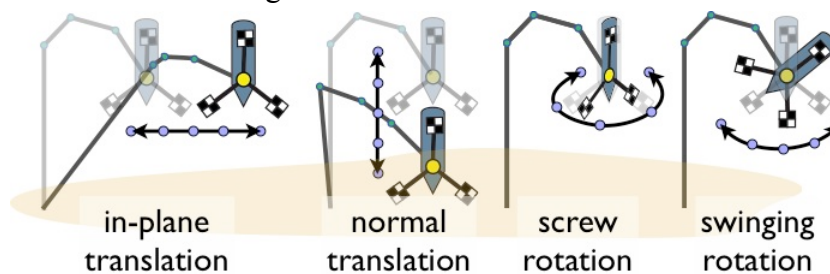
"Palette" function assigns some number of "waypoints" equally spaced across the path defined by user input of starting and end points. The entire translational or rotational

path is defined as a *trajectory*. The sub-movement between two waypoints within the trajectory is defined as a *motion*. At each waypoint, the robot waits for collecting multiple frames of optical tracking data as discussed above. In the meantime, the reference robot pose is recorded. The output has the format of [x, y, z, R_x, R_y, R_z].
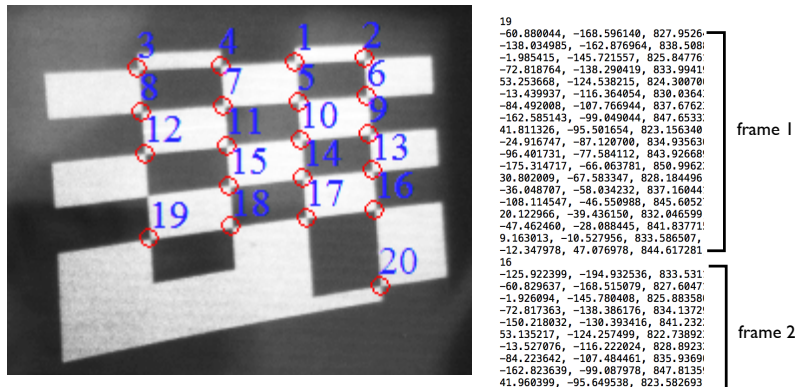
To summarize, the robot arm defines translational and rotational trajectories, which consists of smaller motions between a waypoint to the next. Robot *pose* is recorded at each waypoint. MicronTracker interface measures the *pose* of physical marker, and *position* of all cross-points of the projected checkerboard.

<u>Analysis pipeline</u>
1) Data processing
   From experiments, robot *pose*, marker *pose* and VRB grid cross-point *positions* are acquired. For the robot and marker, therefore, data processing is merely parsing and averaging multiple frames of data. For the VRP gird cross-point positions, there are a number of steps: 0. Raw data of multiple frames. 1. Frame collation. 2. Outlier truncation and averaging. 3. Plane fitting. 4. Correspondence assignment.

   0. Raw data


```
19
-60.880044, -168.596140, 827.9526
-138.034985, -162.876964, 838.508
-1.985415, -145.721557, 825.84776
-72.818764, -138.290419, 833.9941
53.253668, -124.538215, 824.30070
-13.439937, -116.364054, 830.0364
-84.492008, -107.766944, 837.6762
-162.585143, -99.049044, 847.6533
41.811326, -95.501654, 823.156340
-24.916747, -87.120700, 834.93563
-96.401731, -77.584112, 843.92668
-175.314717, -66.063781, 850.9962
30.002009, -67.583347, 828.184496
-36.048707, -58.034232, 837.16044
-108.114547, -46.550988, 845.6052
20.122966, -39.436150, 832.046599
-47.462460, -28.088445, 841.83771
9.163013, -10.527956, 833.586507,
-12.347978, 47.076978, 844.617281
16
-125.922399, -194.932536, 833.531
-60.829637, -168.515079, 827.6047
-1.926094, -145.780408, 825.88358
-72.817363, -138.386176, 834.1372
-150.218032, -130.393416, 841.232
53.135217, -124.257499, 822.73892
-13.527076, -116.222024, 828.8923
-84.223642, -107.484461, 835.9369
-162.823639, -99.087978, 847.8135
41.960399, -95.649538, 823.582693
```
frame 1

frame 2

   1. Frame collation
       Multiple frames of grid position data are stacked on each other. For each frame, distance map is created such that if the distance is less than 0.5 mm, the two points from different frames are regarded as the same points. x and y coordinates are quite consistent. However, as can be seen from the figure on the right, z coordinate data is very noisy. Outlier rejection and averaging is necessary.
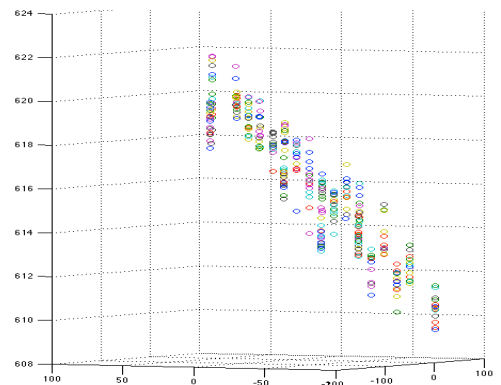

   2. Outlier rejection and averaging
       Simply averaging the collated data does not produce planar results, as can be seen on the left. Therefore, z-score thresholding of 0.5~1 is performed and then averaged, whose results are more consistent.
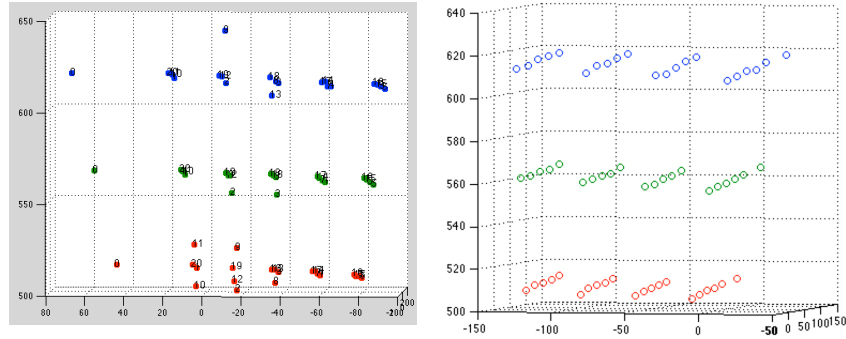
3. Plane fitting

The z-coordinate disturbances still remain, and are further dealt by fitting a plane using RANSAC within the measured VRG grid points and projecting the measured grid points onto the estimated plane. RANSAC and plane fitting implementation was adapted from the codes available online, credited to Peter Kovesi, Centre for Exploration Targeting, The University of Western Australia.
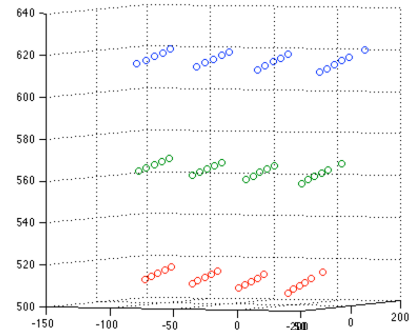

4. Correspondence assignment

Finally, each point in refined VRB grid is assigned correspondences, which are necessary in VRB pose estimation algorithm. Going back to the discussion of VRB grid, the extra pattern that gives asymmetry to the grid serves as a reference point, or point 0. The following points are numbered in the closest order, and therefore a zig-zag correspondence is assigned.
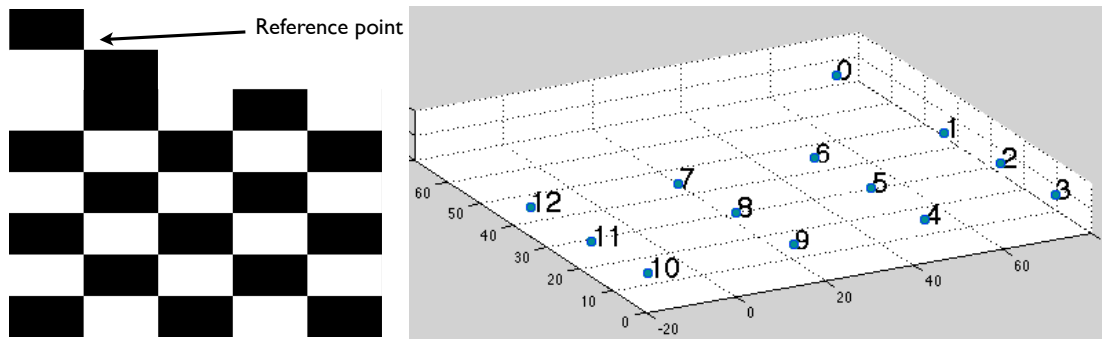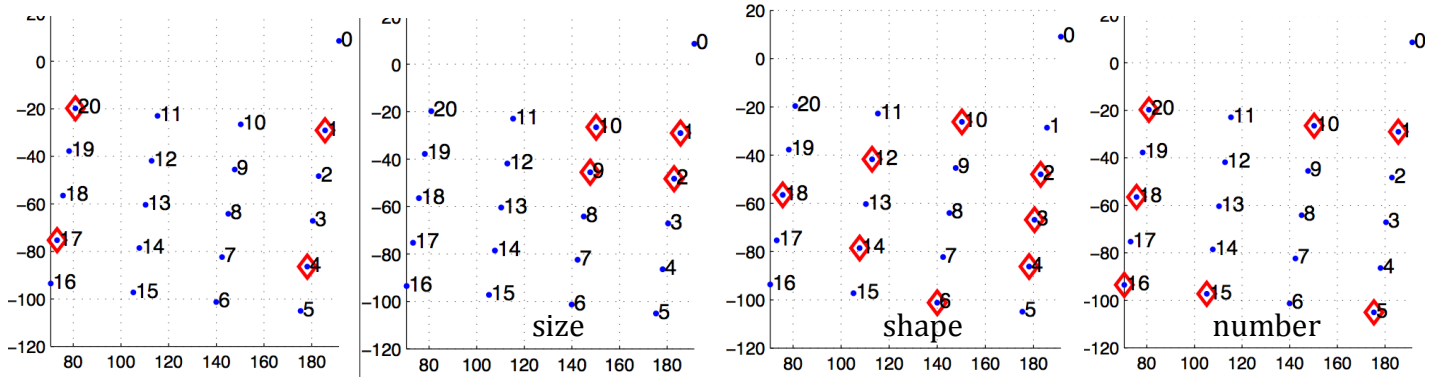

2) VRB configuration selection and projector pose recovery

Some design ideas of VRB configuration include size, shape and number of projected light rays.
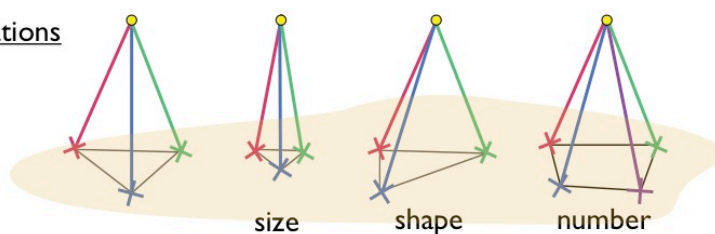
From these different VRB configurations, the pose of the projector, "VRB pose" is recovered. While fully explained in Cheng et al., 2014, the algorithm is breifly illustrated below.

- Pyramid model:
  - $T_a = [R, \vec{t}]$: Unknown pose of apex in tracker coordinate
  - $\vec{u}_i$: Known unit projection vectors from apex
  - Distance between projected bodies
    - $d_{ij} = ||\vec{p}_i - \vec{p}_j||$: from tracker coordinate
    - $||s_i\vec{u}_i - s_j\vec{u}_j||$: from apex coordinate
  - Find sets of s_i such that optimization problem $d_{ij}^2 = ||s_i\vec{u}_i - s_j\vec{u}_j||^2$ is satisfied.
  - T_a is recovered by registration between $\vec{p}_i$ and $s_i\vec{u}_i$

$T_a = [R, \vec{t}]$?
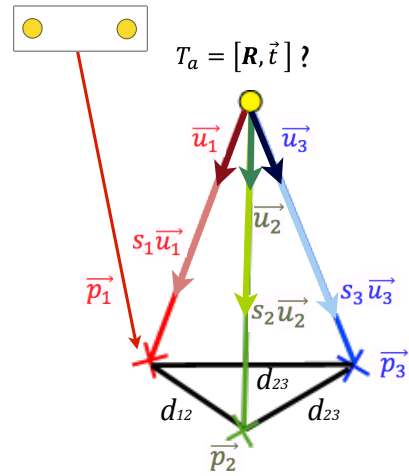
3) Error calculation

The current study focuses on the relative motions between VRB and robot pose, and between marker and robot pose. The deviation of VRB and marker motions from the robot motion is defined as the error. Translational error Δt is defined for translational motion, and rotational error $\Delta\theta$ is defined for rotational motion. The idea and basic implementation, which are undergoing current research, were provided for by the mentor.

A motion can be described by a transformation $F_{ij} = [R, \vec{t}]$ between a pose i to another pose j. The distance $t = |\vec{t}|$, and angle of rotation, $\theta = \arccos\left(\frac{trace(R)-1}{2}\right)$ define how much the pose moved from i to j. Ideally for translational motion, t should be equal for robot, marker and VRB from pose i to j. The differences in t, from pose i to j, $\Delta t_{marker} = |t_{marker} - t_{robot}|$ and $\Delta t_{VRB} = |t_{VRB} - t_{robot}|$ define the translational error. Ideally for rotational motion, θ should be equal for robot, marker and VRB from pose i to j. The differences in θ, from pose i to j, $\Delta\theta_{marker} = |\theta_{marker} - \theta_{robot}|$ and $\Delta\theta_{VRB} = |\theta_{VRB} - \theta_{robot}|$ define the rotational error. This is summarized in the diagram below.

For $F_{ij} = [R, \vec{t}]$ between poses i and j,

$$\theta = \arccos((trace(R) - 1)/2)$$
$$t = |\vec{t}|$$

For translation,
$$\Delta t_{marker} = |t_{marker} - t_{robot}|$$
$$\Delta t_{VRB} = |t_{VRB} - t_{robot}|$$

For rotation,
$$\Delta\theta_{marker} = |\theta_{marker} - \theta_{robot}|$$
$$\Delta\theta_{VRB} = |\theta_{VRB} - \theta_{robot}|$$
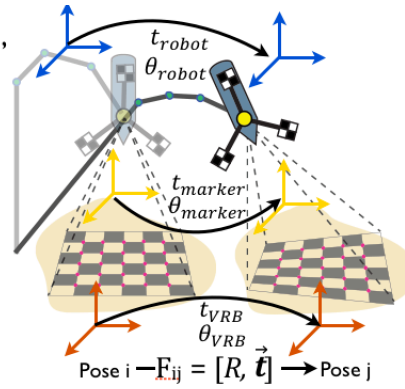
Results

Some error analysis results from combinations of motion trajectories (translational and rotational), and VRB pattern configurations (size, shape, number of projection) are presented.

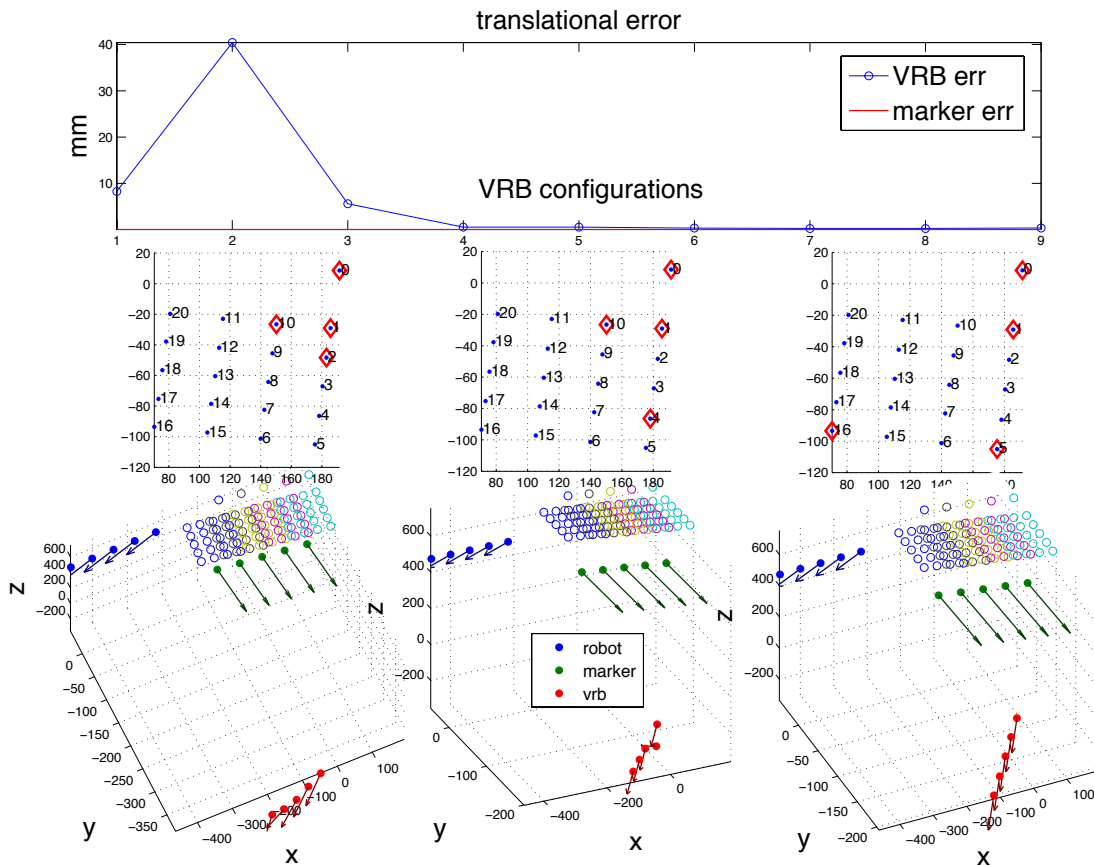1) In-plane translational motion, increasing pattern size

Figure above illustrates the performance of VRB of increasing size in recovering in-plane translational motion. The middle row shows the increasing pattern size (red diamonds) from left to right. The bottom row shows the recovered VRB poses (red) from corresponding VRB configuration, along with pose measured from physical marker (green) and robot (blue). The top row shows the error between robot and marker poses (red), and between robot and VRB poses (blue). Physical marker reports translational error of 0.04mm. While VRB error did not get better than the marker error, VRB error decreased from about 10 mm to 0.2 mm from configuration 1 to 9, which increased in

size. The improvement in VRB error is reflected in the visualization of the VRB poses from different configurations (bottom row). From left to right, the recovered VRB pose (red) forms more consistent straight line. Outliers exist. For VRB configuration 2, whose pattern size is larger than configuration 1, the error reaches 40mm, which is much bigger than expected. Although not shown, it is hypothesized that the position of acquired "point 3" included in configuration 2 is noisy, since other configurations including point 3 seems to consistently reflect high error regardless of the configuration.

2) Screw rotational motion, increasing pattern size
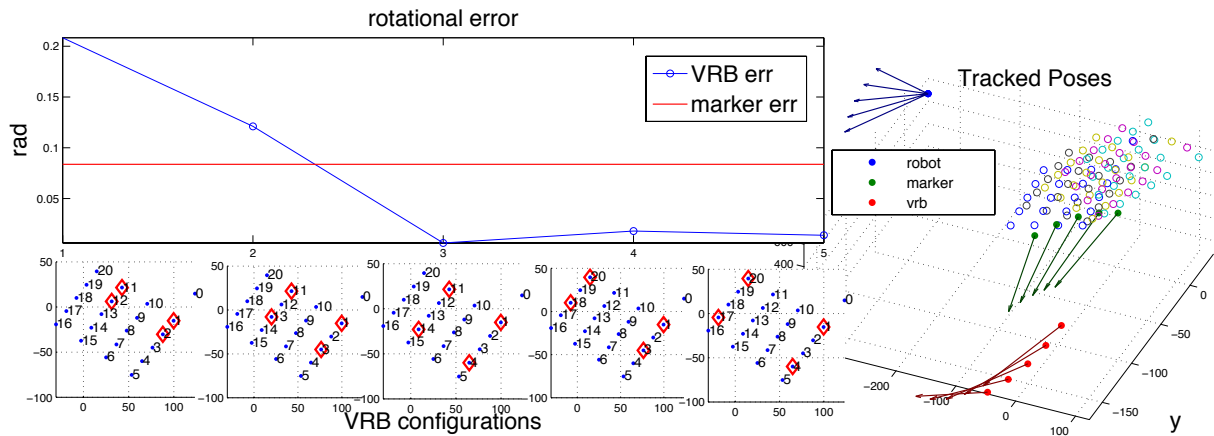
Similarly, VRB error in screw rotational motion is analyzed against increasing VRB pattern size. Marker reported error of 0.08 rad. As the pattern became wider, the rotational error decreased from 0.2 to 0.01 radians. From the third VRB configuration, VRB performed better than the physical marker.

3) Normal translation motion, polygons with increasing number of sides
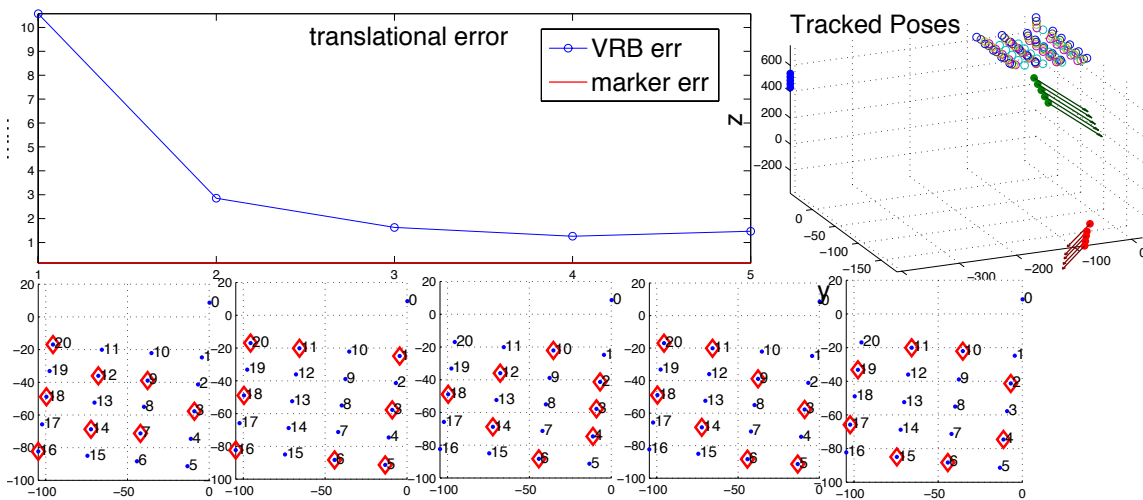
Marker reported error of 0.14mm. Different VRB shape was analyzed in normal translational motion. From left to right, polygon patterns of VRB with increasing number

of sides were analyzed, from triangle to octagon. Interestingly, the error decreased quite consistently from 11 to 1.8 mm.

4) Normal translation motion, number of projection
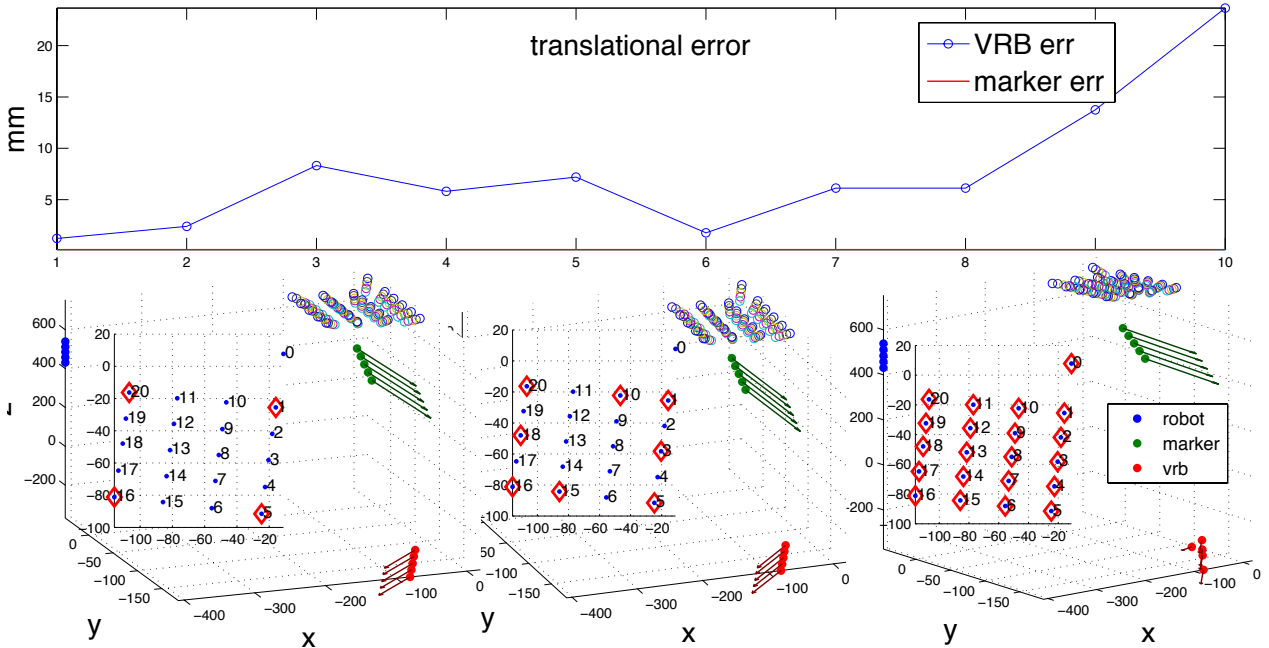
Marker reported error of 0.14mm. Number of projection was increased from 4 to 21 (entire grid). In contrast to our expectations, VRB error increased with increased number of grid point selection, from 1.2 to 24 mm.

Discussion

Preliminary understanding of pattern between VRB configuration and tracking accuracy can be established. In general, wider – larger distance of each virtual fiducials from the centroid of the selected virtual fiducials, and larger distance between each virtual fiducial – projected pattern led to lower error. Therefore, in the results describing different polygon pattern versus tracking error, it is believed that how extensively the fiducials are distributed, rather than the polygonal shape itself, influences tracking error. This suggests that the novel system of VRB can be understood in the classic framework regarding optical tracking error and fiducial distribution, that higher tracking accuracy is achieved from wider distribution of fiducial markers.

More often than not, however, results deviate from general understanding. For instance, larger number of fiducials is in general expected to provide higher tracking accuracy. As described above, however, using larger number of projection led to higher error. This suggests that there are yet much to study regarding virtual rigid body.

Much emphasis is placed on the fact that this study is only preliminary. The choice of patterns was not random, and the influence of experimenter's bias cannot be neglected. Another shortcoming is that MicronTracker, which was used with the intention of controlling the influence of detection algorithm, in fact may affect performance of VRB in tracking. As discussed, big portion of analysis pipeline consists of identifying planar

grid points from very noisy depth information. The nature of projected pattern can also affect accuracy of MicronTracker's detection algorithm. Projection may entail blurring; when the projection is slanted, detection performance may be influenced because in general physical markers assume orthogonal cross-points; surface conditions such as reflectivity can introduce noise as well. Therefore, the foremost necessity is development of external segmentation method specifically tailored to detecting virtual projected markers.

Still, the results reconfirm the prospects of VRB in optical tracking. With large enough virtual fiducial distribution, lower error was achieved in estimating poses from VRB for rotational motion. Why rotational motion observes such significant improvement as opposed to translational motion is placed for further investigation.

The current project takes a first step of finding optimal design for the newly developed virtual rigid body system, which attempts to track an object from projected light pattern. An experimental and analysis setup has been configured to acquire and analyze optical tracking data under controlled motion. With the pipeline, basic trend between virtual rigid body designs and tracking accuracy of simple translational and rotational motion has been identified in comparison to the conventional physical marker system. Much more extensive and quantitative analysis will follow to provide more concrete idea on optimal design of virtual rigid body.

Literature
•Cheng et al.,*Virtual Rigid Body:A New OpticalTracking Paradigm in Image Guided Interventions,* CARS, 2014
• West et al., *Designing Optically Tracked Instruments for Image-Guided Surgery,* IEEE Transactions On Medical Imaging, 2004
• J.A.Sánchez-Margallo, *Technical Evaluation of a Third Generation Optical Pose Tracker for Motion Analysis and Image-Guided Surger*y, CLIP, 2012

Management

Task distribution

    As a single-person project, I was in primary charge of the project. However, Alexis always gave me advice and guidance with the problems that I had to overcome, and the next steps I needed to take. Dr. Boctor also provided me valuable insights on the large picture of the project, and advice on skills and attitude when presenting my work.

Deliverables

    I had underestimated the depth of the project because the logic was straightforward. I had planned too much, yet too little, by over-planning a publication and even mechanical design, yet underestimating the size of experiment and analysis pipelines. I was not able to reach original maximum deliverables. I achieved concrete experimental and analysis pipelines, solid data for simple trajectories, and basic ideas on optimal design of VRB.

Planned	Achieved
Minimum • Marker grid • Experimental routines in form of python or C++ codes • Experimental data	Minimum - Experimental Pipeline Setup • Virtual rigid body (VRB) grid • Detection component • Processing component • Robot component
Expected • Analysis and evaluation of the virtual markers • Optimal design of virtual markers	Expected - Experiment/Analysis • Run pipeline for data collection • Analysis pipeline • Comparison between virtual and physical rigid body • Basic design of virtual rigid body
Maximum • Publication • Experimental data on non-level surfaces. • Introductory ideas on projector design.	Maximum • Documentation

Further works

    The foremost task is to more quantitatively evaluate the virtual fiducial distribution, for instance, distance from the centroid, and from each other. Another important project is developing an external detection and segmentation methods. As discussed, depth information from the MicronTracker is too noisy for projected light patterns, which seems to be primary reason that VRB accuracy is often not as high as expected. Then would follow experiments on non-level surfaces, and mechanical design for smaller projectors.

What I learned

    I became much more familiar with general optical tracking problem, and with using the robot arms. I learned how research is a continuous rollercoaster of hope and misery, and how it all sums into a valuable experience. I learned that I have to learn.