**Object Pose Estimation**

**Metric**

Object pose estimation is a kinematic measure of how well a robotic hand can estimate the pose of an object. The pose of an object is described in Cartesian coordinates, and the estimation fidelity will be captured in terms of the error between the hand-estimated Cartesian pose versus the reference-measured Cartesian pose. Object pose estimation is useful feedback for in-hand manipulation control and hand-arm coordination and control, particularly since visual occlusions for an external vision system typically occur when grasping an object.

**Dependencies**

System dependencies for object pose estimation can vary considerably based on strategy. Strategies can fall into one of two main categories – contact or non-contact. If requiring object contact, then pose estimation capabilities will likely involve proprioceptive and cutaneous sensory systems as well as the driving algorithm for making the estimations. In this case, estimation performance will depend on these underlying sensors and the overarching estimation algorithm. For the non-contact strategy, a vision strategy is likely used that will depend on the vision sensor and supporting algorithms. Overall performance is likely to also depend on the object’s properties as well including morphology, orientation, and optical traits.

**Test Method**

Artifact:

Due to object dependency, the test should be conducted across a range of objects. Any object of interest must be retrofitted with reflective markers in order to measure the Cartesian pose of the object with a reference motion capture system (MOCAP).

Description:

Of the previously listed dependencies, only the object will be taken as a controlled test variable. It will be up to the user to place the object appropriately within the hand and establish an object-fixed coordinate system and a ground coordinate system. The object-fixed coordinate system should be known to both the robotic hand and the reference MOCAP system with the relevant transformations. The object should be moved through a variety of poses with respect to the hand, while object poses are estimated by the hand and “ground-truth” poses are measured by the reference MOCAP system.

Performance Measures:

The main performance measure should be the Root Mean Squared Error of $e\_{estimation}$, $RMSE\_{e,estimation}$, where $e\_{estimation}= r\_{c,estimation}- r\_{c}\in R^{6×1}$. Furthermore, $r\_{c}$ is the pose of the object as measured by the reference MOCAP system, and is defined as, $r\_{c}=[x,y,z,γ,β,α]$, where $x,y,z$ are translations and $γ$,$ β$, $α$ are rotations about the $X$, $Y$, and $Z$ axes. Finally, $r\_{c,estimation}$ is the pose of the object as estimated by the robotic hand. $RMSE\_{e,estimation}$ is calculated separately for a variety of parts. For thorough experimentation, several runs should be conducted per object, and the mean and 95% confidence interval of $RMSE\_{e,estimation}$ can be calculated to capture a more accurate representation of performance.

**Example Implementation**

Test Setup:

A three-fingered, 7 degree-of-actuation robotic hand retrofitted with bio-inspired tactile sensors was used as the test platform to manipulate several objects (see Fig. 1). Also shown in this figure are three geometrically primitive artifacts – sphere, cuboid, and cylinder. The sphere has a diameter of 120 mm and mass of 286 g, the cuboid has dimensions of 90 mm by 90 mm by 75 mm and a mass of 178 g, and the cylinder has a diameter of 90 mm, a height of 75 mm, and a mass of 143 g. The artifacts are retrofitted with reflective markers for position tracking using a motion capture system (MOCAP). The time-variant desired translation and rotation trajectories were defined as follows:

1) $r\_{cd,z}=-0.0075sin(t)+0.1425$ (m)

2) $r\_{cd,γ}=-\frac{π}{25}sin(1.25t)$ (rad)

3) $r\_{cd,β}=-\frac{π}{25}sin(0.75t)$ (rad)

This implementation is a contact-based object pose estimation solution, and therefore the hand was tasked with manipulating the object along the above-defined trajectories in order to induce object motion and estimate the object’s pose. See Manipulation Control test method.



Figure 1. Robotic hand holding a a) sphere, b) cuboid, and c) cylinder with attached reflective markers for object motion tracking.

Results:

Given these trajectories, object pose estimation performance was captured in Table 1. Interestingly, the orientation error remained relatively low. The translation performance was most accurate in the Z-axis across all objects. Substantial translation error accrued in the X and Y axes.

1. Total manipulation performance for object translation and object orientation.

|  |  |  |
| --- | --- | --- |
| Object | $RMSE\_{e,estimation}$ (mm) | $RMSE\_{e,estimation}$ (deg) |
| Sphere | [18.00,4.57,2.34] | [2.98,3.37,3.23] |
| Cuboid | [8.45, 6.82,2.76] | [2.05,1.81,2.87] |
| Cylinder | [11.88,5.66,2.50] | [1.97,1.29,3.07] |

Notes:

It is observed that estimation errors are largely due to finger-object slipping that was not detectable or incorporated by the robotic hand.

Data (to do):

|  |  |
| --- | --- |
| *Data File Archive:*   |  |
| *Data Files:*  |  |
|  |  |
| *File Format:*  |  |
| *Data Values:*  |  |
| *Units:* |  |
| *Data Sample Rate:* |  |