

Supplementary Material

Anonymous Author(s)

Affiliation

Address

email

Below is the supplementary material regarding the work done in “One-Shot Imitation Learning: A Pose Estimation Perspective”. More information can be found on the project’s website <https://sites.google.com/view/one-shot-il-posest>. On the latter you’ll be able to find a concise summary of the proposed one shot imitation learning framework, along with a series of real-world robotics videos showing the potential of such formulation.

1 Relationship Between T_δ Expressed in the Camera Frame and the World Frame

In the main paper we state that the equation below changes the frame of reference of T_δ from the camera’s frame to the robot’s one.

$${}^R T_\delta = T_{RC} {}^C T_\delta T_{CR} \quad (1)$$

Now we will derive the latter expression to give a better understanding of the underlying algebra of the Special Euclidean group in three dimensions, SE(3). Firstly, we refer to the following equation which has been derived in the paper

$${}^C T_\delta = T_{CO}^{Test} T_{OC}^{Demo} \quad (2)$$

It is important to note that in the above equation, the camera could be replaced by any fixed frame. As a result we can derive a similar expression also for T_δ in the robot frame.

$${}^R T_\delta = T_{RO}^{Test} T_{OR}^{Demo} \quad (3)$$

Additionally we know that

$$T_{RO}^{Test} = T_{RC} T_{CO}^{Test} \quad (4)$$

By substituting equation 2 in the above equation we get

$$\begin{aligned} T_{RO}^{Test} &= T_{RC} T_{CO}^{Test} = T_{RC} {}^C T_\delta (T_{OC}^{Demo})^{-1} \\ &= T_{RC} {}^C T_\delta T_{CO}^{Demo} \end{aligned} \quad (5)$$

Finally we can substitute equation 6 into 3 which concludes this derivation.

$${}^R T_\delta = T_{RO}^{Test} T_{OR}^{Demo} = T_{RC} {}^C T_\delta T_{CO}^{Demo} T_{OR}^{Demo} \quad (6)$$

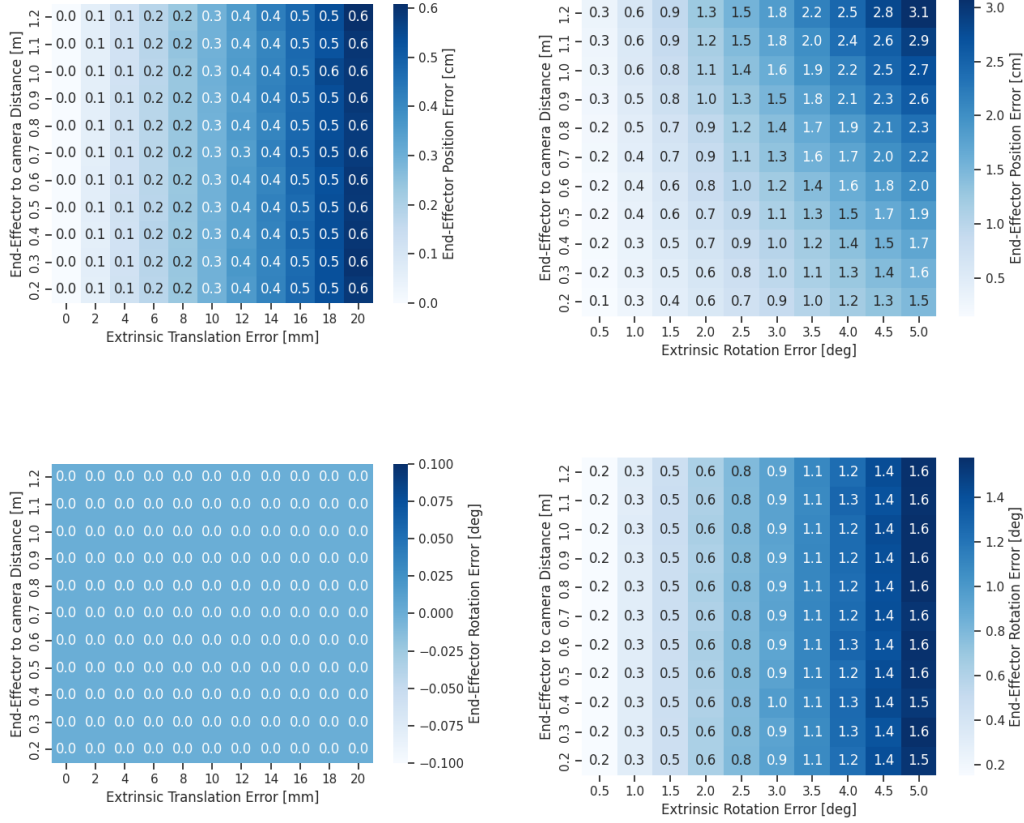
$$= T_{RC} {}^C T_\delta T_{CR} \quad (7)$$

2 Sensitivity Analysis of One-Shot Imitation Learning

Here we present the full results obtained from our controlled noise experiments

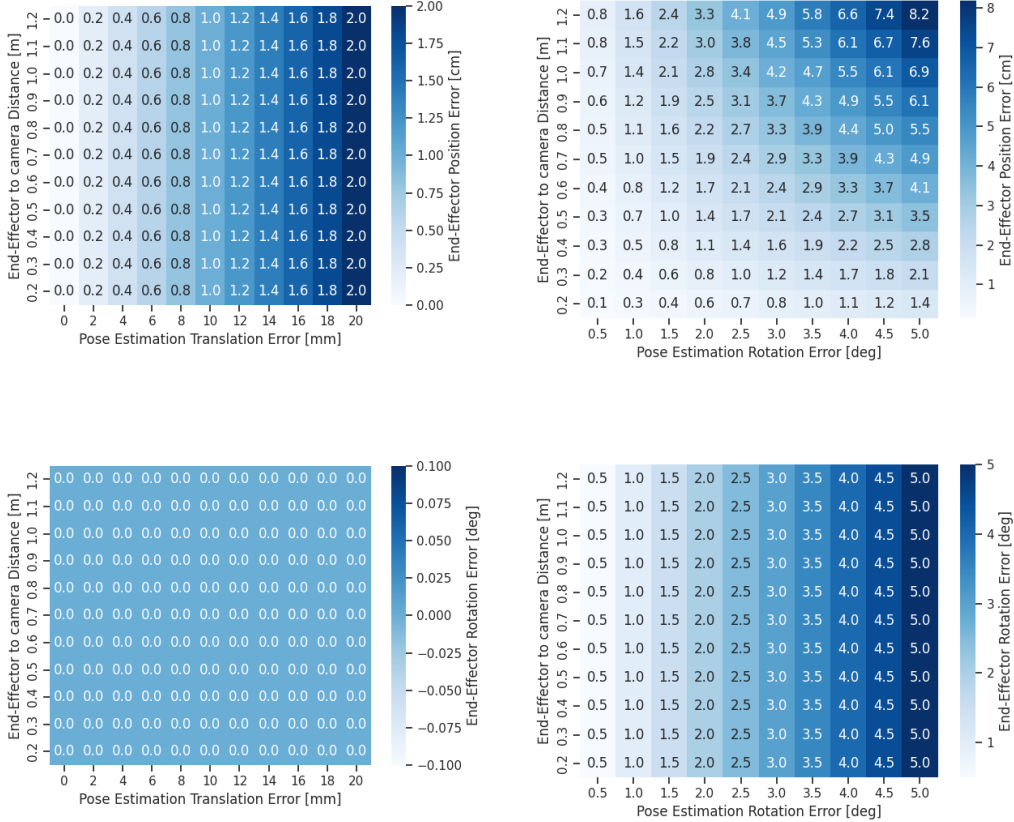
2.1 Sensitivity to Error in Extrinsic Camera Calibration

In the figures below we can show that the relationship between rotation errors in T_{RC} and orientation errors in trajectory alignment is linear and independent of the EEF-to-camera distance. Similarly, translation errors in T_{RC} correlate with translation errors in alignment, while there is no relationship between translation errors in T_{RC} and rotation errors in alignment. The scarce impact of errors in the translation compared to the orientation component could be attributed to the fact that rotations can induce translations, but not vice versa.



2.2 Sensitivity to Error in Relative Pose Estimation

We experience a similar behaviour to the previous section.



3 Considered One-Shot Unseen Object Pose Estimation Baselines

In the main paper we discuss the performance of eight pose estimation baselines within the context of a simulated experiment.

Firstly, we generated a custom dataset using blender. More specifically, we generated scenes with a single object in them. Afterwards, we collected data consisting in a set of RGB-D image pairs all taken from the same camera pose. Each pair representing the same object in different poses, one for each image. Along with these observations we also saved the relative pose between the objects in the two images, which will then be used as ground truth for pose estimation. To test the environment generalisation of the baselines, each scene had a random lighting condition and background.

Moreover we divided the objects in five categories. 1) **Non-symmetric objects** which are objects that are not symmetric around any of their axes. 2) **Fully-symmetric objects** which are objects that are symmetric around their z-axis. This means that when estimating the pose of them, the predicted orientation around their z-axis does not constitute an error. 3) **Potentially-symmetric objects** which are objects that were not in the fully-symmetric category because they are visually non symmetric, although they are fully symmetric geometrically. In the context of robotic skills such objects might still be considered symmetric. For instance a can is symmetric around one of its axes but its label might not be uniform, therefore resulting in a non symmetric visual appearance. 4) **Partially-symmetric objects** which are objects that are symmetric around their z-axis but only by certain rotations. For instance the orientation of a uniform cube is equivalent when rotated by multiples of ninety degrees. 5) Lastly **Potentially-partially-symmetric objects** follow the same reasoning as the potentially fully symmetric objects but with only a finite number of valid orientations around the z-axis.