

Abstract: The Impact of Synthetic Data On The Reliability of Hand Tracking Systems

Daniel Desmond Dennis

April 2020

This dissertation investigates the merits of using synthetic data for training a hand tracking model to address the problem of the lack of training data for such systems. This is achieved by looking at two different synthetic data generation strategies, and training an existing hand tracking model to see how it performs on those generated synthetic datasets. An existing real dataset, the MSRA dataset is used as a control. The MANO model is used as the basis for generating synthetic data, and V2V-Posenet is used as the existing hand tracking model. The first strategy generates the data based on random parameters for MANO, this is referred to as the *Random MANO Dataset*. The second strategy generates the data based on parameters determined from the groundtruth of the real dataset to recreate that real dataset using inverse kinematics, this is referred to as the *IK MANO Dataset*.

In recent years, the accuracy of hand tracking systems has improved with the use of *convolutional neural networks* (CNNs). However, these systems need comprehensive training data, and such a dataset does not exist currently. This need is underlined by the fact that current hand tracking systems do not generalise well for a given training dataset, and performs poorly for hand poses that are not covered by that training dataset. The reasons for this lack of data are twofold. Firstly, annotating these datasets is difficult. Secondly, the human hand is a uniquely complex object, and a comprehensive dataset needs to cover a diverse range of; camera perspectives, hand poses, hand shapes, and interactions with objects, so the amount of data that needs to be collected to cover these ranges is also unknown. Synthetic data offers a way around this issue, since the annotation is theoretically perfect, and the capture of that data can be automated.

This dissertation shows that generating synthetic based on the *Random MANO Dataset* strategy shows promise, but more testing is required by generating more synthetic images. State-of-the-art results on this test dataset is demonstrated at an 8.90mm average per-joint mean squared error, but poor performance when comparing with the real test dataset. It also shows that trying to recreate a real dataset with the *IK MANO Dataset* is challenging, with efforts to achieve this showing poor performance at 77.44mm. The key contributions of this dissertation are (1) performing a side-by-side comparison between a synthetic and real hand tracking dataset using the state-of-the-art synthetic hand data generation method, MANO, (2) giving the basis for a new synthetic dataset based on the MANO model with the *Random MANO Dataset*, and (3) providing a new framework for generating a synthetic dataset by using a real dataset as a basis with the *IK MANO Dataset*.