Non-Linear Hand-Pose Undersensing: Preliminary Results

Pawan Subedi, Mathias Antonin Schoen and Edoardo Battaglia

***Abstract*— Hand pose visualization through full Hand Pose Reconstruction (HPR) is an important aspect of haptic experiences, and can enable more natural and intuitive interaction with virtual or remote objects. However, the high number of degrees of freedom that need to be sensed can make the task challenging, especially when using approaches based on wearable sensing. In this paper, we propose preliminary results from applying a non-linear data-driven approach for full HPR using a reduced number of sensors, and compare the outcome with results using a previously developed linear method on the same dataset. Our preliminary model yields a Root Mean Squared Error of 13.2 degrees on the test set, with an average error reduction of approximately 2 degrees across the dataset when compared with the previous linear method. Visual results have been obtained in CHAI3D using the open source cHand library.**

I. Introduction

Measuring hand pose is important in many haptic applications, since hands are our primary means of interaction with the environment. There is a rich state of the art on the topic, with different approaches relying either on vision-based techniques or wearable systems. While vision-based approaches have the advantage of leaving the hand free, their performance is significantly worse when occlusions are present (e.g. grasping inside a cup). Wearable systems are robust to this, but usually require a high number of sensors to measure full hand posture, leading to higher costs and cumbersome calibration procedures. In this paper we present preliminary results from our work on an undersensing approach, using deep learning to reconstruct the full hand pose in a 25 DoF hand model from five joint angles measurements. This led to an improved performance when compared to a previous linear method [1].

II. Methodology

Results in this paper were obtained from an open source dataset [2], which was divided into training, validation, and testing datasets following standard neural network training practices. We designed a deep neural network consisting of four dense layers with non-linear activation functions, which were chosen together with other hyperparameters (e.g., number of neurons in each layer) through an iterative tuning procedure. The input angles for the network were chosen to be the joint angles of each knuckle, i.e. $\theta_4, \theta_7, \theta_{12}, \theta_{17}$ and θ_{22} in the model shown in Figure 1. The target output is a vector containing the remaining 20 joint angles in the chosen hand model, which combine with the previous five inputs to produce the full 25 DoF hand pose.

All authors are with the Department of Mechanical Engineering at the University of Utah. e-mail: edoardo.battaglia@utah.edu

This work was supported in part by the National Science Foundation under award number #1846726.

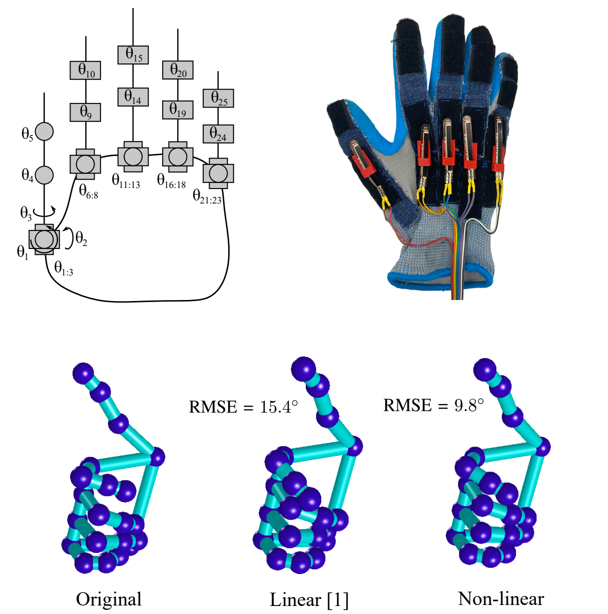

Fig. 1: Hand model, glove and visual representations of the ground truth and reconstructions from the dataset from [2], for a sample hand pose.

III. Preliminary Results and Future Work

Our preliminary neural network yielded a Root Mean Squared Error of 12.9, 13.3, and 13.2 degrees on the training, validation, and testing datasets respectively, while the method used in [1] yielded 16.2, 16.0, and 15.2 degrees. We are currently setting up a glove with five bend sensors attached to it, which will be used to reconstruct hand postures in real time. Work on the implementation of our machine learning model on a glove-embedded board is also currently ongoing. After that, we will validate the system's performance with live data against a motion capture system. In parallel, we are considering different network architectures to reduce the error further, and evaluating results for a lower number of input angles. Finally, we will explore the possibility of implementing a feature selection approach where the network automatically selects the best subset of angles to use as input, for a given number of input angles.

References

[1] S. Ciotti, E. Battaglia, N. Carbonaro, A. Bicchi, A. Tognetti, and M. Bianchi, "A synergy-based optimally designed sensing glove for functional grasp recognition," *Sensors*, vol. 16, no. 6, p. 811, 2016.

[2] O. Glauser, S. Wu, D. Panozzo, O. Hilliges, and O. Sorkine-Hornung, "Interactive hand pose estimation using a stretch-sensing soft glove," *ACM Transactions on Graphics (ToG)*, vol. 38, no. 4, pp. 1–15, 2019.