A Stretch-Sensing Soft Glove for Interactive Hand Pose Estimation

Oliver Glauser
ETH Zurich
oliver.glauser@inf.ethz.ch

Shihao Wu
ETH Zurich

Daniele Panozzo
New York University

Otmar Hilliges
ETH Zurich

Olga Sorkine-Hornung
ETH Zurich

Figure 1: Our stretch-sensing soft glove captures hand poses in real time and with high accuracy. It functions in diverse and challenging settings, like heavily occluded environments or changing light conditions, and lends itself to various applications.

ABSTRACT

We present a stretch-sensing soft glove to interactively capture full hand poses with high accuracy and without requiring an external optical setup. Our device can be fabricated with simple tools available in most fabrication labs. The pose is reconstructed from a capacitive sensor array embedded in the glove. We propose a data representation that allows deep neural networks to exploit the spatial layout of the sensor itself. The network is trained only once, using an inexpensive off-the-shelf hand pose reconstruction system to gather the training data. The per-user calibration is then performed on-the-fly using only the glove.

ACM Reference Format:

Oliver Glauser, Shihao Wu, Daniele Panozzo, Otmar Hilliges, and Olga Sorkine-Hornung. 2019. A Stretch-Sensing Soft Glove for Interactive Hand Pose Estimation. In *Proceedings of SIGGRAPH '19 Emerging Technologies*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3305367.3327975>

1 INTRODUCTION

Hands are our primary means for object manipulation and important for communication. Applications such as gaming, robotics, biomechanical analysis, rehabilitation, AR and VR depend on accurate means to recover the full hand pose.

Vision-based methods have made significant progress in hand motion capture, but they require externally mounted cameras with full visibility of the whole hand. This cannot be guaranteed in many practical scenarios, due to object manipulations, body occlusion, self-occlusion or the hand moving out of the field of view. Furthermore, changing light conditions or motion blur due to fast translational movements are challenging for camera-based techniques.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SIGGRAPH '19 Emerging Technologies, July 28 - August 01, 2019, Los Angeles, CA, USA
© 2019 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-6308-2/19/07.
<https://doi.org/10.1145/3305367.3327975>

Thus they are limited to applications with a controlled environment and impose physical constraints on immersive user experiences. Mounting inertial measurement units (IMUs), bend sensors, strain sensors or combinations thereof directly onto the hand removes the need for direct line-of-sight and can improve robustness and reliability. Such sensorized data gloves have been proposed in research (e.g. [Chossat et al. 2015]) and are available commercially (e.g., [Cyb 2019; Man 2019]), they have at most 22 sensors which are fewer than DOFs in the human hand. This inherently restricts the reconstruction accuracy.

Therefore, we introduce a new data glove featuring as many as 44 sensors in combination with a data-driven model for hand pose reconstruction while retaining a thin, unobtrusive form-factor.

2 TECHNICAL CONTRIBUTION

Goals. We focus on the core task of a data glove — capturing accurate hand poses. Furthermore, an optimal data glove should be comfortable to wear, real-time, low cost, easy to use and not require any external setup, e.g., cameras.

Sensor and Fabrication. Our glove senses local stretch magnitude exerted on embedded silicone sensors by measuring their capacity changes. The use of a large number of sensors is made possible by recent work introducing stretch-sensing sensor arrays entirely from silicone [Glauser et al. 2019]. We contribute an adapted fabrication process (shown in Fig. 2) for wearable gloves and propose a custom sensor design for the task of fine-grained hand-pose estimation. Furthermore, we integrate the sensor array with a textile cut pattern, and propose a readout scheme to enable sensing at 60 Hz instead of 8 Hz. The resulting glove has a material cost of around 75 USD. It can be fabricated in a modern fablab, the only special tool required is a laser cutter.

Data-driven approach. We use more sensors than the hand's DOFs and leverage modern data-driven techniques to transfer the resulting sensor readings to hand poses. We discover that a data

Figure 2: The fabrication consists of two main stages: Fabricating the silicone sensor (1-8) and the textile base glove (9-12).

Figure 3: From left to right: the sensor readout is normalized by the hand-specific min-max per sensor, arranged in two 5x5 stretch maps, and fed forward through a fully-convolutional neural network (U-Net) that predicts a 5x5 pose map.

representation based on prior knowledge of the geometric neighborhood and spatial correspondence, both in the input and output domain, allows a neural network to more efficiently discover the inter-dependencies between the joints in the human hand (Fig. 3).

Efficiently capturing a sufficiently large and diverse training data set for hand pose estimation is a notoriously difficult problem. Our glove design is so unobtrusive that it is almost invisible to a depth-camera. This allows us to leverage a state-of-the-art model-fitting based hand tracking approach [Tkach et al. 2017] to capture one million samples of sensor readings and the corresponding joint-angle configurations.

Figure 4: A gallery of real-time session frames showing diverse poses predicted by a personalized model.

3 RESULTS

We performed experiments on a dataset captured from 10 people. The hand length varies from 17 to 20.5 cm, the width from 9 to 11 cm, and the aspect ratio of length to width from 1.6 to 2.1. We report mean pose reconstruction errors (ME) in the angular form.

Our glove allows for different applications scenarios depending on the availability of time, equipment, and the requirement of accuracy. Possible scenarios include a *generic* model tested on an unseen hand with an on-the-fly min-max calibration (average ME of a leave-one-out experiment of 10 people: 7.6), a generic model *fine-tuned* with personal data (average ME: 6.5) and a fully *personalized* model (average ME: 6.2) where both training and testing data are captured from the same hand.

In a comparison sequence, with a ME of 6.8, our glove outperforms the ManusVR glove (ME: 11.9) and the CyberGlove (ME: 10.5). Fig. 4 shows a gallery of frames from a live capture session. It is possible to take a model trained on data from one glove and predict poses with another glove (Glove-II) with a mean error of 8.80 degrees for a test session. If we fine-tune the model with just 5 minutes of captured Glove II training data, we can reduce the mean error to 5.73 degrees. During data acquisition and method development, one of our gloves was in use for over 25 hours (cumulative).

Note that the current version of our device does not come with the tracking of global hand translation and orientation, which, however, is a well-studied problem, and can be easily added in conjunction with virtual reality applications (e.g., using the HTC Vive tracker or an Oculus touch controller).

4 EXPERIENCE

The emerging technology visitors have the chance to try our new stretch-sensing glove in interactive demos. We will encourage the visitors to touch, feel, and wear our comfortable and suitable glove. After a short on-the-fly calibration of about one minute, the visitor can explore how accurate our glove can capture hand poses in real-time: the currently estimated hand pose is continuously displayed (like in a mirror) on a screen. Additionally, various data glove applications in human-computer interaction and virtual reality will be showcased.

REFERENCES

- 2019. CyberGlove. <http://www.cyberglovesystems.com/>.
- 2019. ManusVR glove. <https://manus-vr.odoo.com/hardware>.
- Jean-Baptiste Chossat, Yiwei Tao, Vincent Duchaine, and Yong-Lae Park. 2015. Wearable soft artificial skin for hand motion detection with embedded microfluidic strain sensing. In *Proc. ICRA*. IEEE, 2568–2573.
- Oliver Glauser, Daniele Panizzo, Otmar Hilliges, and Olga Sorkine-Hornung. 2019. Deformation Capture via Self-Sensing Capacitive Arrays. *ACM Trans. Graph.* 38, 2 (2019).
- Anastasia Tkach, Andrea Tagliasacchi, Edoardo Remelli, Mark Pauly, and Andrew Fitzgibbon. 2017. Online Generative Model Personalization for Hand Tracking. *ACM Trans. Graph.* 36, 6 (2017).