LOCAL SCALE ADAPTATION FOR AUGMENTING HAND SHAPE MODELS

PRATIK KALSHETTI*, PARAG CHAUDHURI
INDIAN INSTITUTE OF TECHNOLOGY BOMBAY

ViGIL VISION, GRAPHICS, IMAGING & LEARNING
SIGGRAPH 2022 VANCOUVER+ 8-11 AUG

PROBLEM

Input Sequence of depth frames

Output Registered hand mesh corresponding to each depth frame

Challenge For accurate tracking, the hand mesh needs to be *calibrated* to the user's hand shape [1]

RELATED WORK

The most widely used solution is to use a data-driven hand model, MANO [2], which uses PCA shape blends.

Issue It is learned from 2018 scans of 31 subjects, and thus cannot generalize to unseen hand shapes with substantially large deviations from training data.

OUR SOLUTION

We propose **adaptive MANO (aMANO)**, which augments MANO's shape space with local scale adaptation that enables calibrating to users with substantially different hand sizes than those covered by the original MANO shape space.

METHOD

Local scale parameter $\phi_j = \frac{\|a'_j - b'_j\|}{\|a_j - b_j\|}$ for each bone j

Use modified linear blend skinning (LBS) [3] to avoid stretching artifacts

$$v'_i = \sum_{j=1}^{n_b} W_{b_{ij}} \left(a'_j + R_j \left(W_{e_{ij}} s_j + \left(-a_j + v_i \right) \right) \right)$$

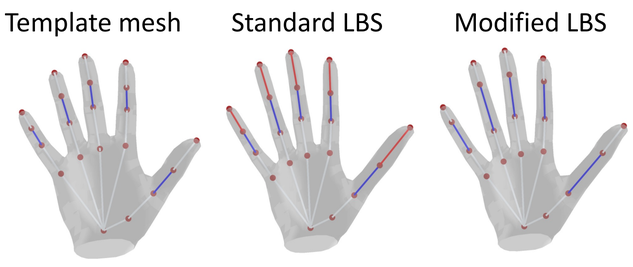
where,

R_j is the rotation that takes bone j's rest vector $(b_j - a_j)$ to its pose vector $(b'_j - a'_j)$,

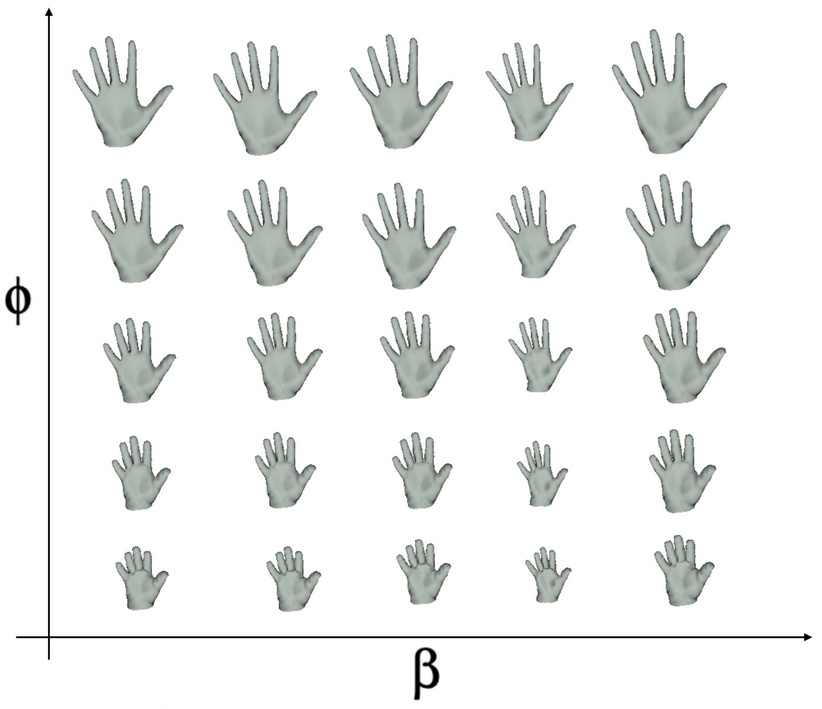
$s_j = (\phi_j - 1)(b_j - a_j)$,

W_b and W_e denote the bone weights and endpoint weights respectively

Endpoint weights W_e for the bones in the index finger



Scenario: Scale PIP-DIP bone (blue). Undesired over scaling(red) beyond endpoint of the bone in standard LBS. This is handled by using endpoint weights in the modified LBS.

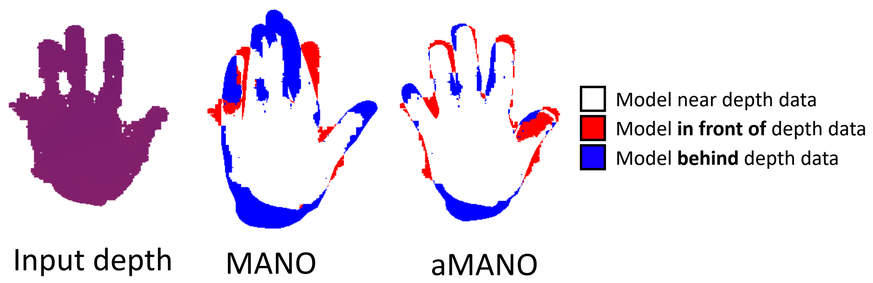


Our local scale parameters ϕ in aMANO compliments MANO's shape space β

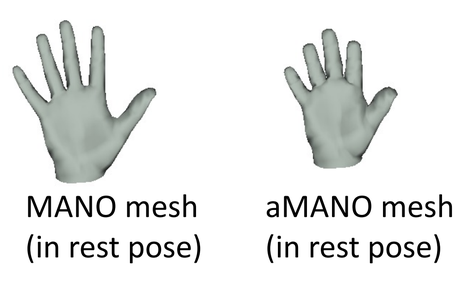
RESULTS

Calibration

Child's hand



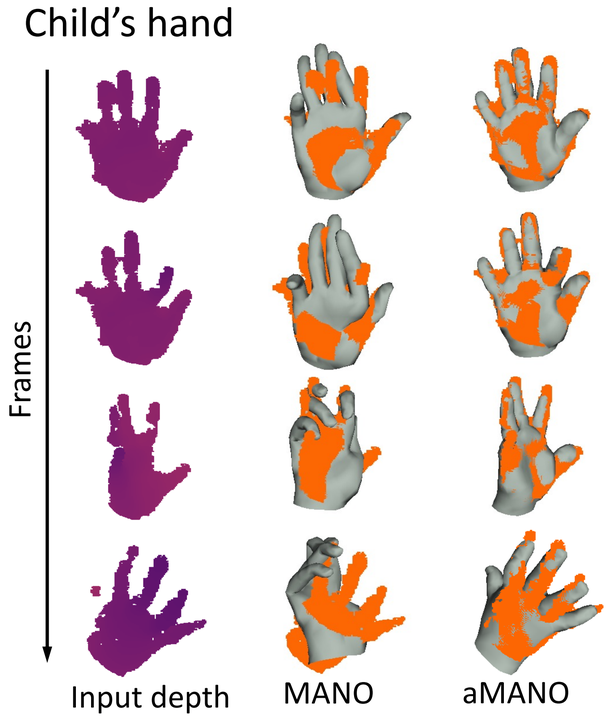
Input depth | MANO | aMANO



MANO mesh (in rest pose) | aMANO mesh (in rest pose)

Tracking

Child's hand



Input depth | MANO | aMANO

data-to-model error metric

$$d2m_{max} = \max_{p \in D^{(data)}} \left\| p - \Pi_{D^{(model)}}(p) \right\|$$

Dataset, User	$d2m_{max}$ (in mm)	
	MANO	aMANO
NYU, User 1	23.2	**18.4**
NYU, User 2	23.6	**19.1**
BigHand, User 1	27.7	**21.6**
BigHand, User 2	23.8	**16.0**
GuessWho, User 1	19.5	**14.2**
Captured data, Adult 1	17.8	**13.6**
Captured data, Child 1	9.2	**7.2**

REFERENCES

1. David Joseph Tan, Tom Cashman, Jonathan Taylor, Andrew Fitzgibbon, Daniel Tarlow, Sameh Khamis, Shahram Izadi, and Jamie Shotton. 2016. Fits Like a Glove: Rapid and Reliable Hand Shape Personalization. In CVPR.
2. Javier Romero, Dimitrios Tzionas, and Michael J. Black. 2017. Embodied Hands: Modeling and Capturing Hands and Bodies Together. ACM TOG 36, 6 (2017), 245:1–245:17.
3. Alec Jacobson and Olga Sorkine. 2011. Stretchable and Twistable Bones for Skeletal Shape Deformation. ACM TOG 30, 6 (2011), 165:1–165:8.

*Corresponding author: pmkalshetti@iitb.ac.in