Hands and Interaction Tracking

Nimble Team

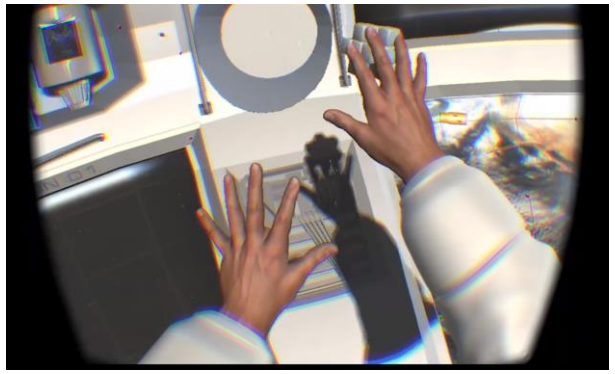
Yuting Ye

Oculus / Facebook Reality Labs

Why do we care about hands for VR?

Self Presence

Seeing your own hands

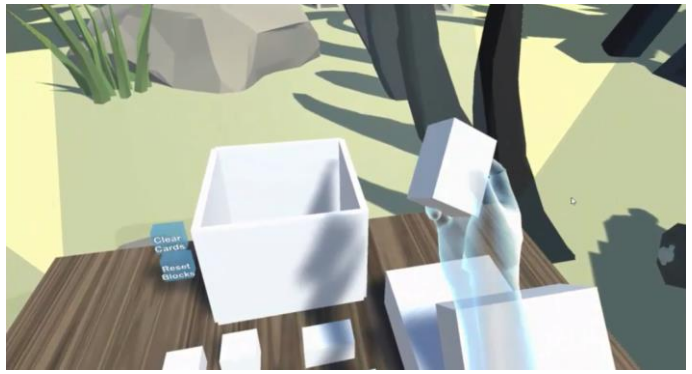


Social (avatar) hands

Incorporating hands into a social avatar

Simple input

Affect the virtual world with pushes, grabs, flicks and gestures

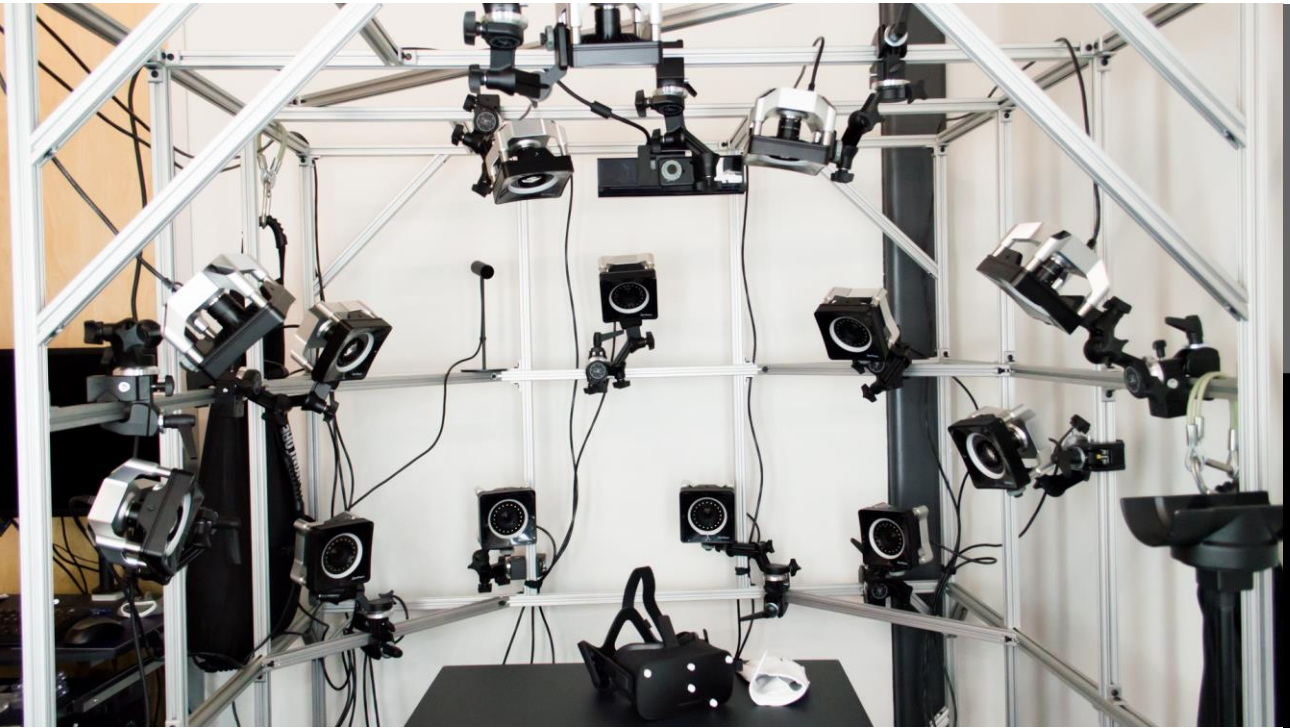


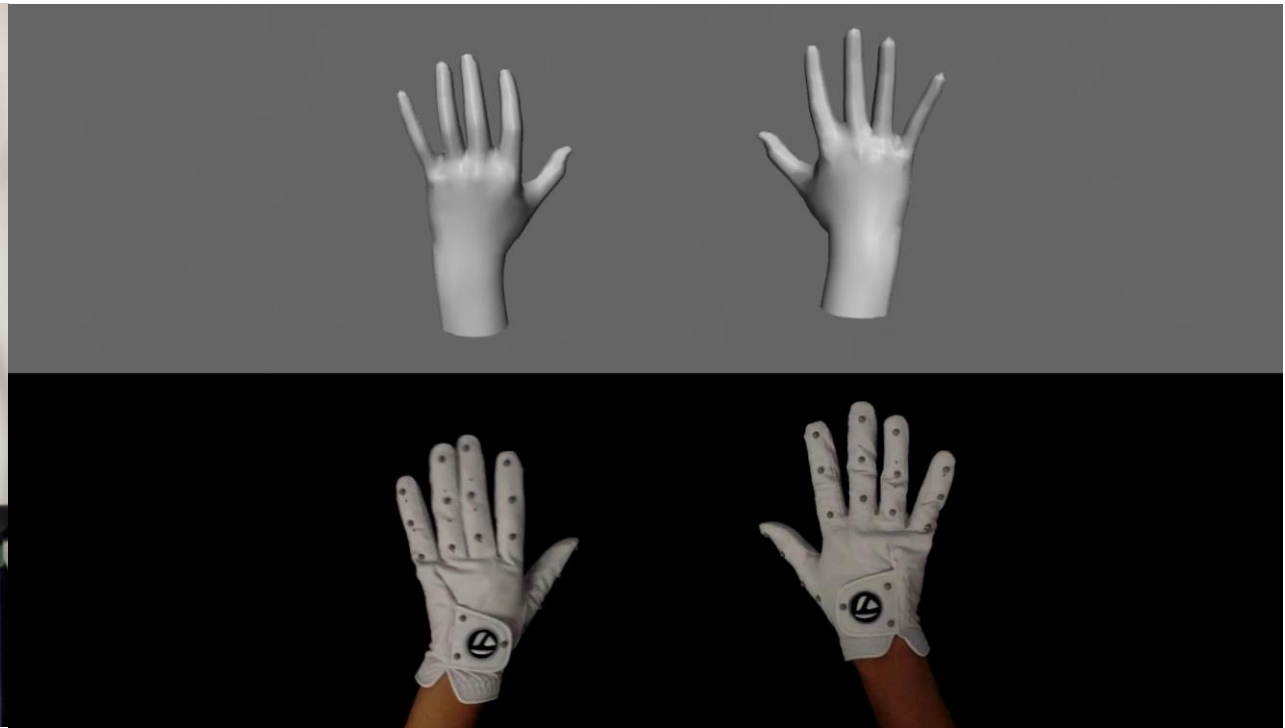
Marker-based hand-tracking

180Hz mocap and neural networks

Highest quality real-time hand tracking

Time machine for interaction prototyping





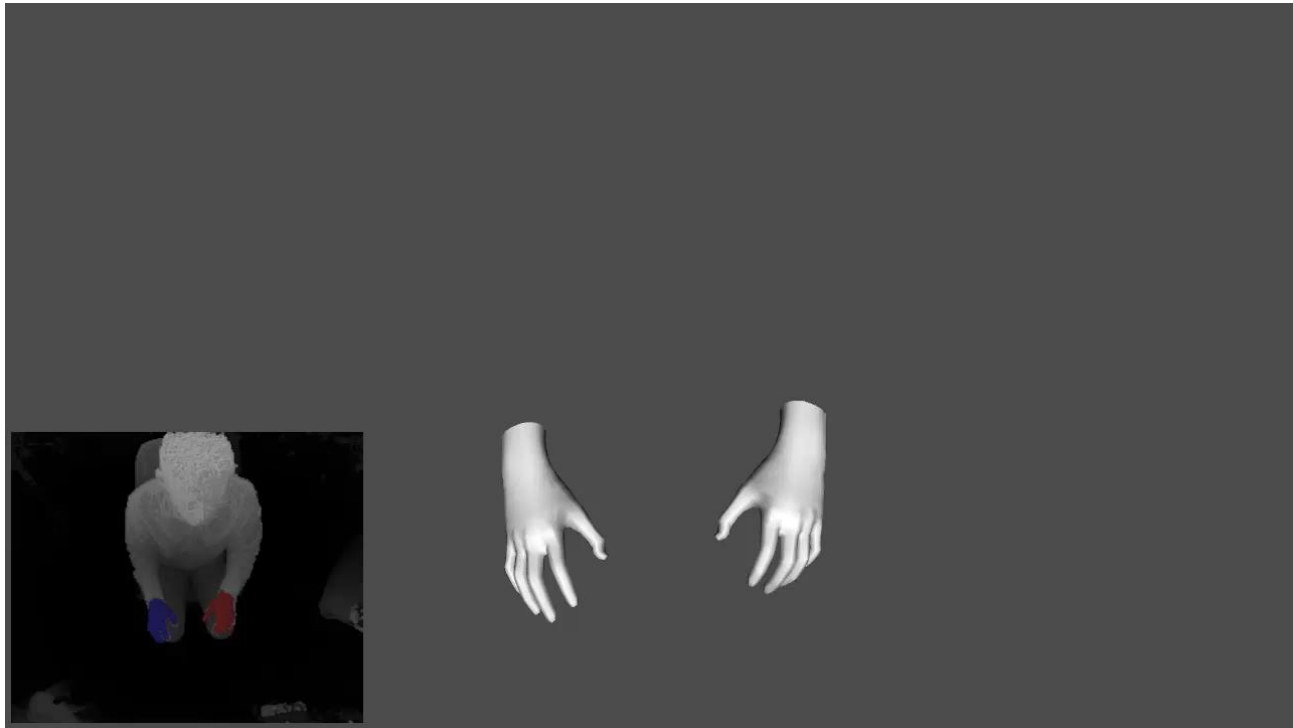

CLOSE
CLEAR
CLOSE
CLEAR
Blue

Depth-based hand-tracking

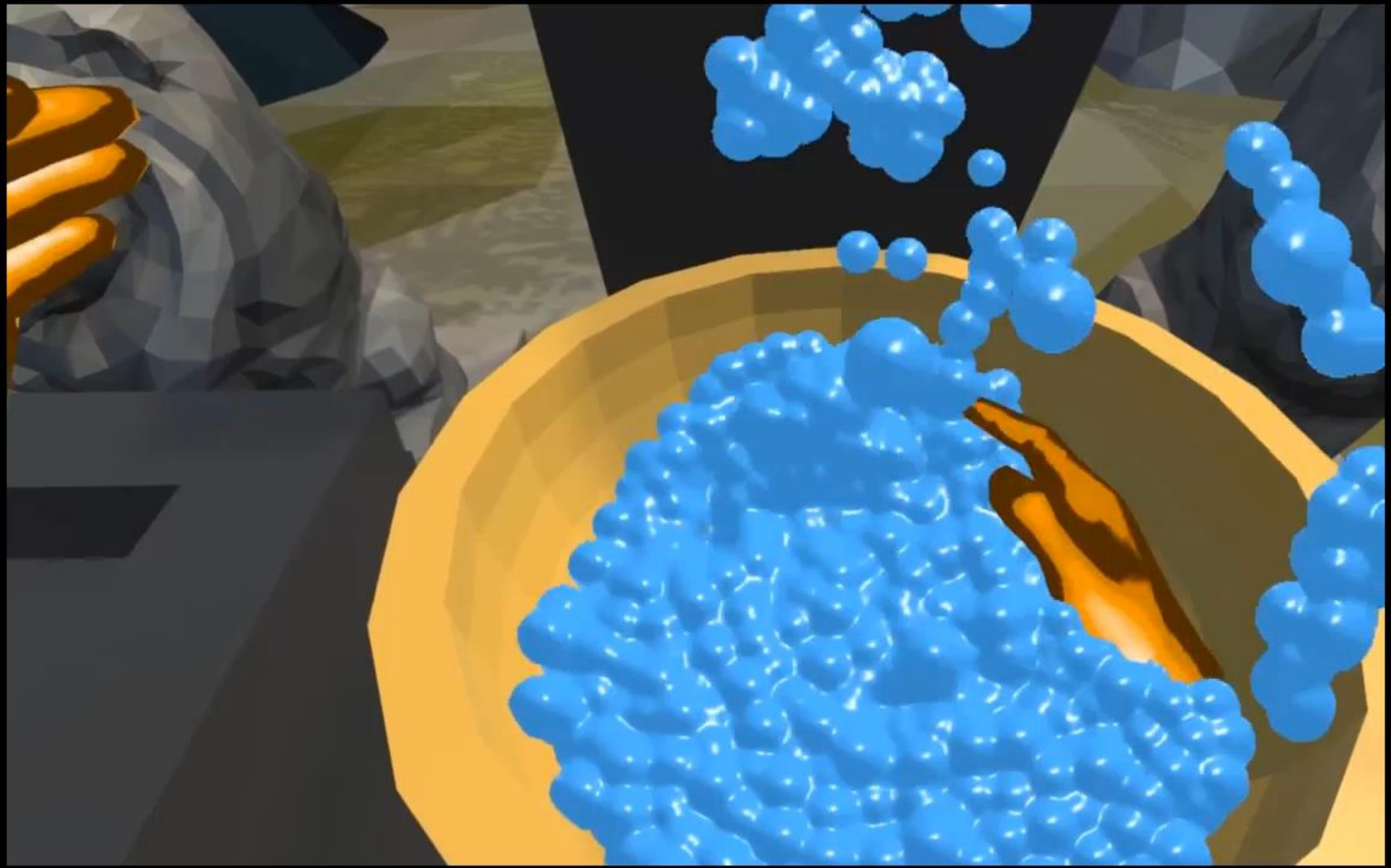
Real-time monocular hand tracking on depth sensing cameras

More practical and easily deployable

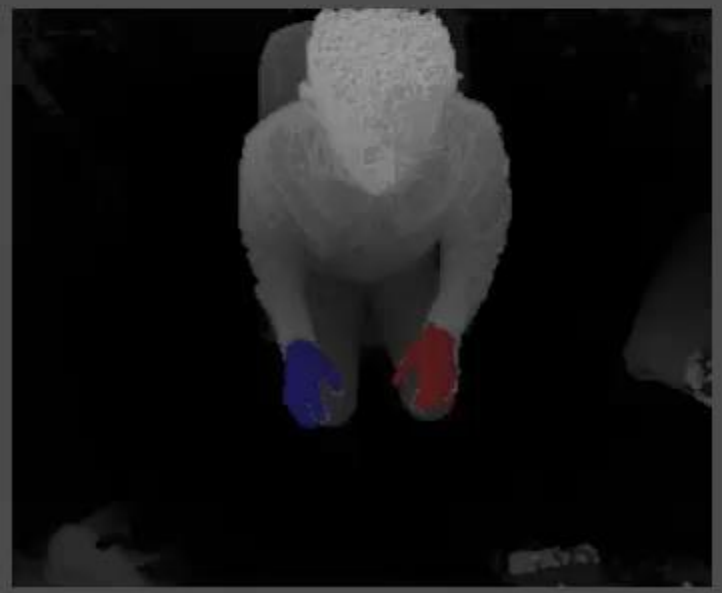
Kinect v2



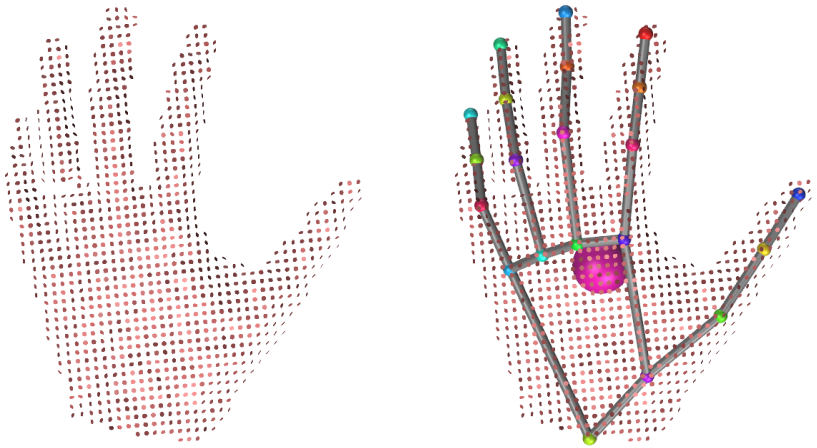

Clear
Cards
Reset
Blocks



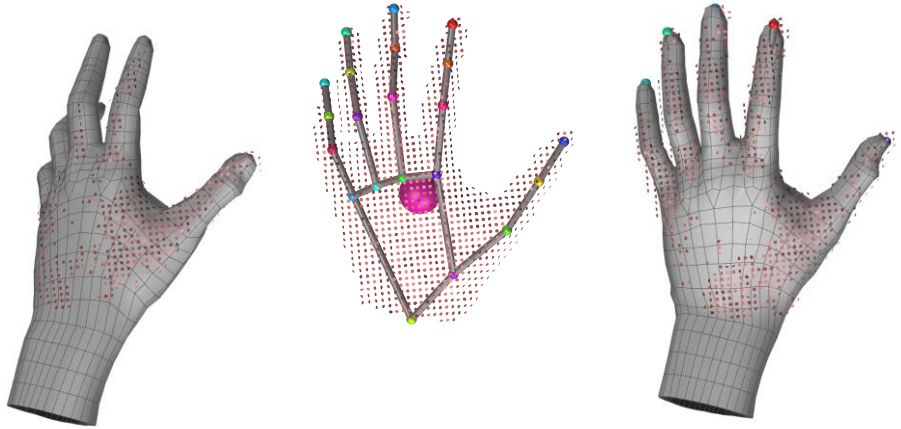
Skeletal hand tracking from a Kinect v2



How do we do it?

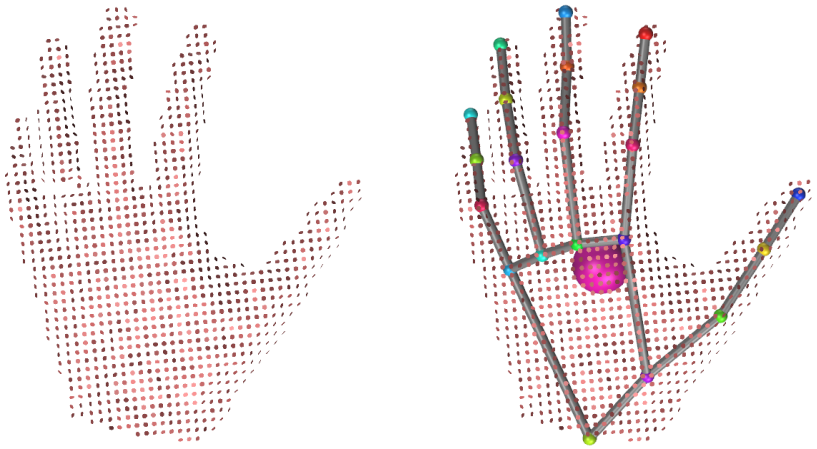


Discriminative models –
trained with machine learning

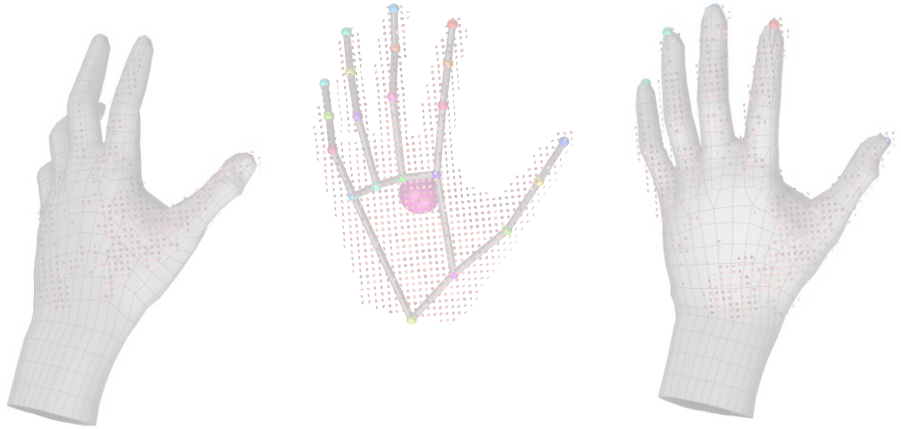


Generative models –
model-based tracking

Making discriminative models work



Discriminative models –
trained with machine learning



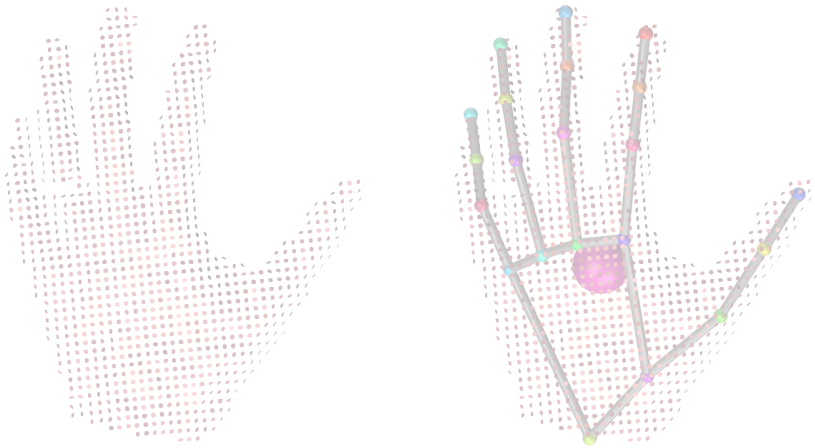
Generative models –
model-based tracking

Collect lots of high quality ground truth

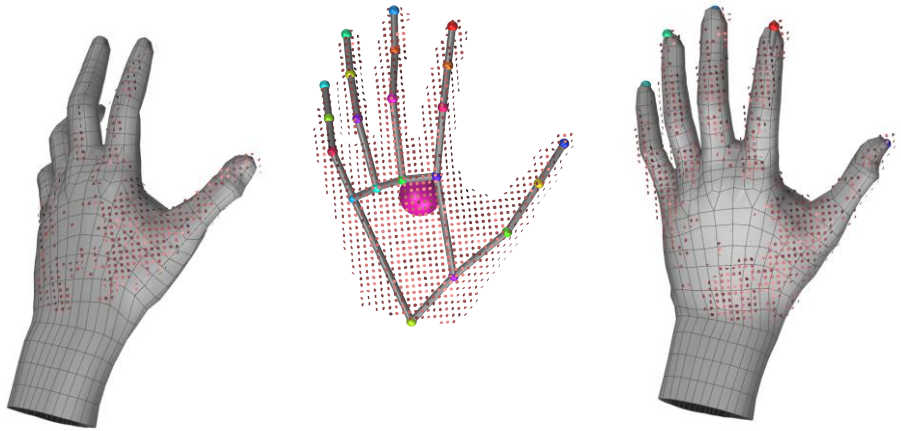
First generation multi-camera ground truth

Second generation multi-camera ground truth

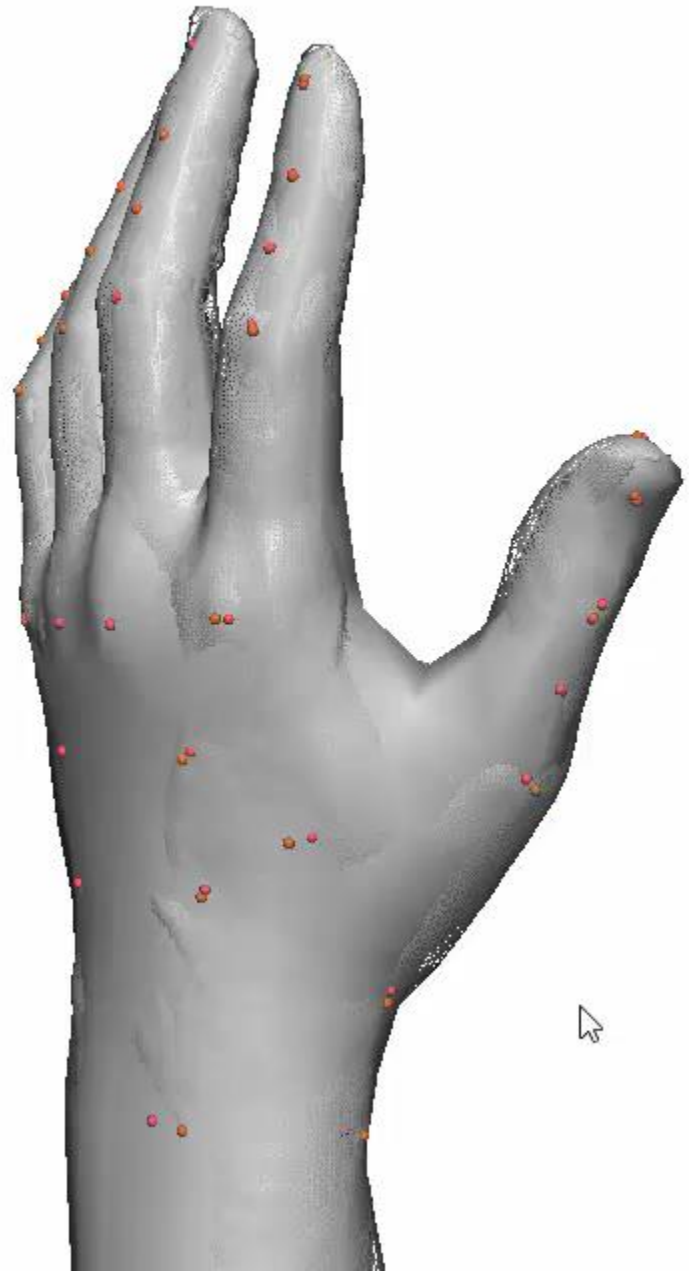
Making model-based tracking work

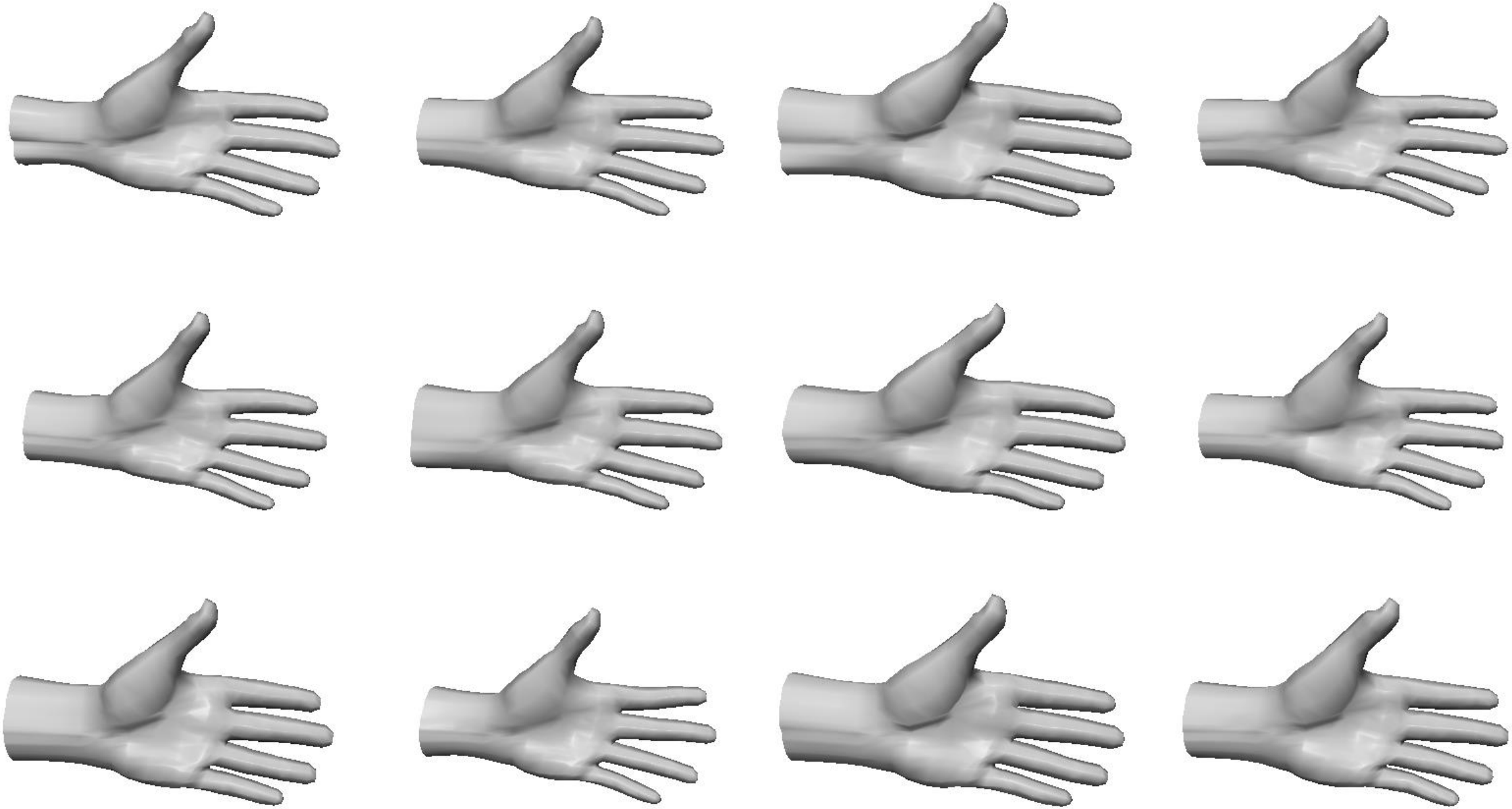


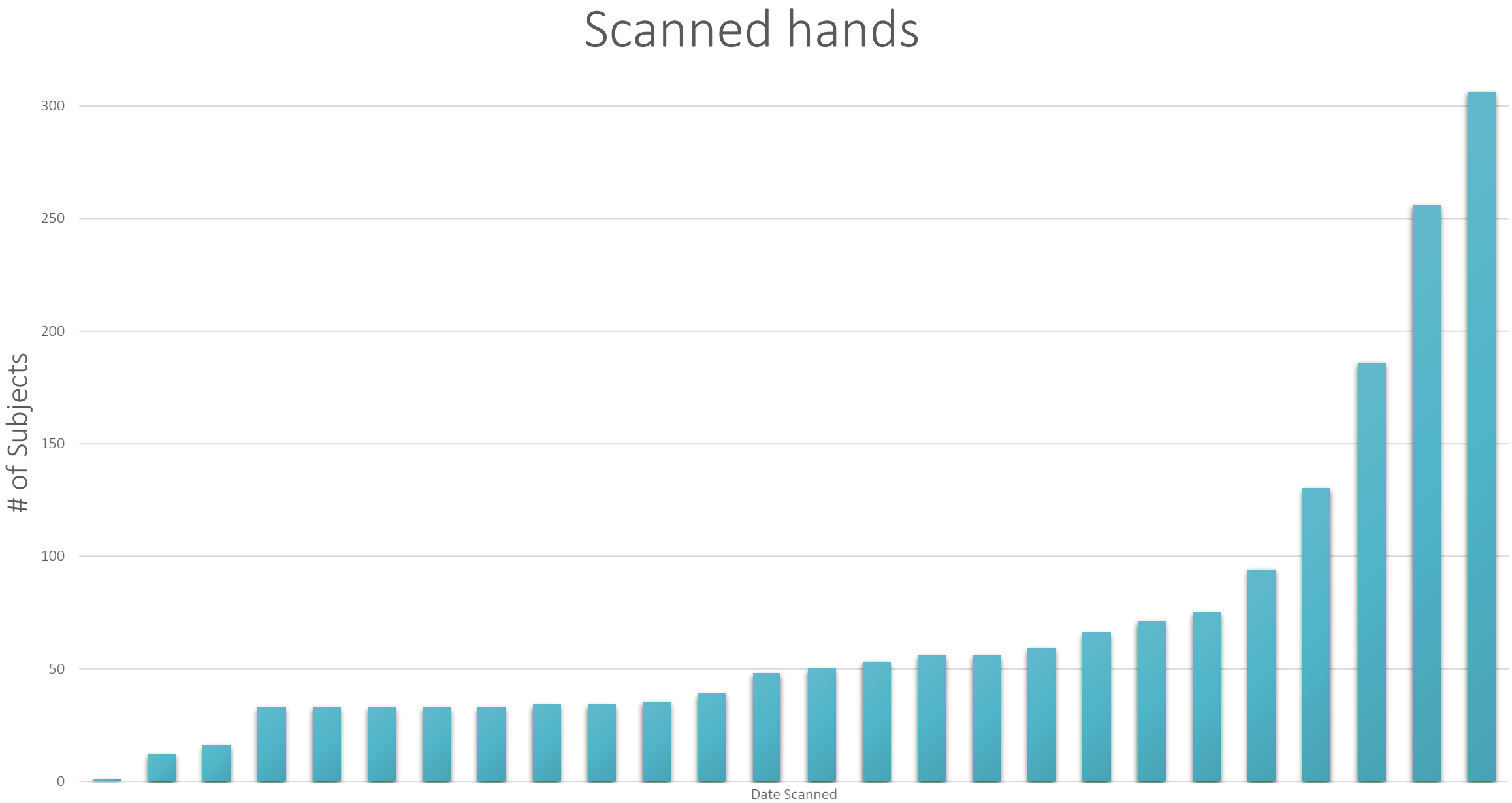
Discriminative models –
trained with machine learning



Generative models –
model-based tracking

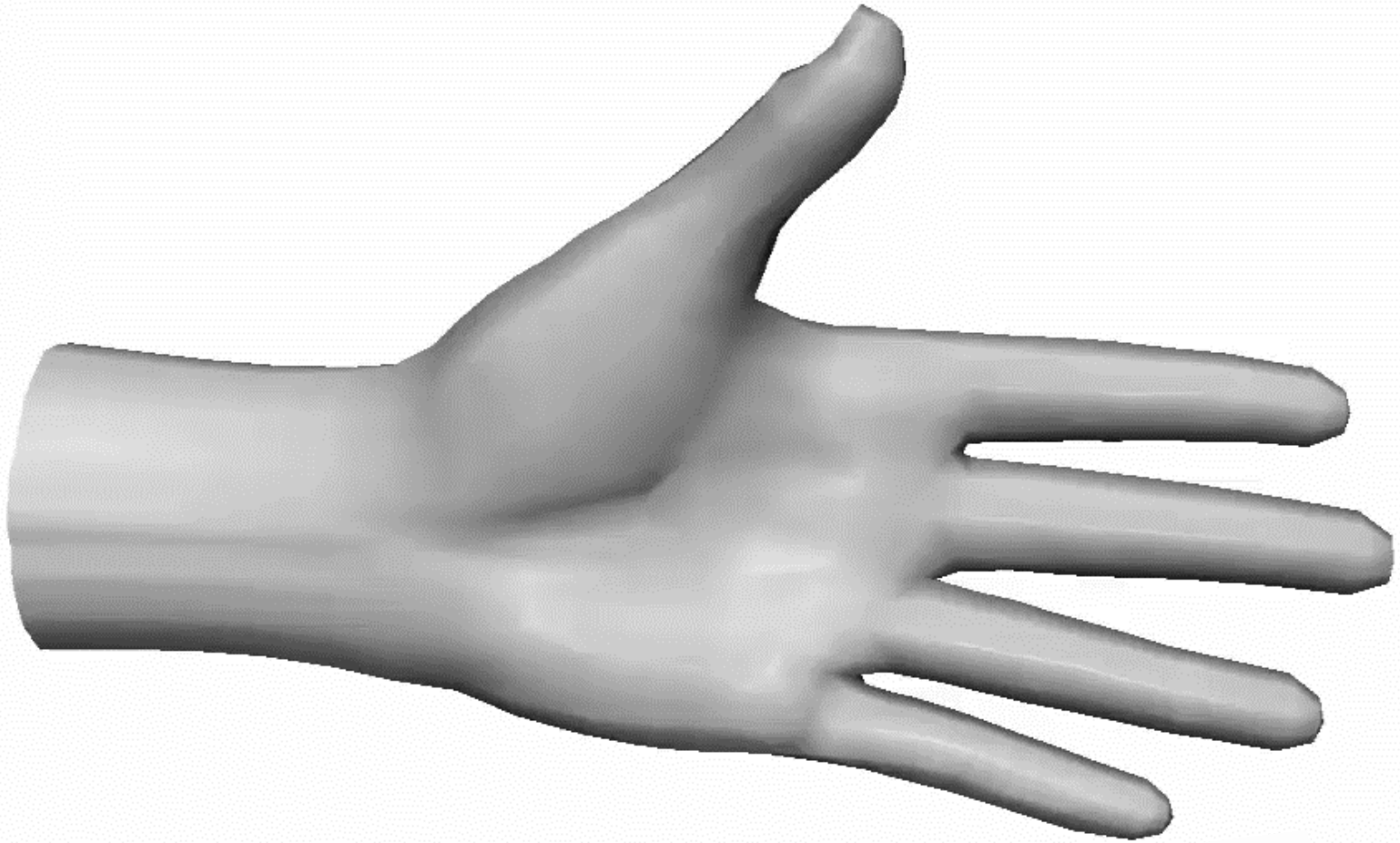


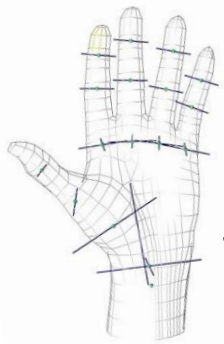




Scanned 300+ hand models – largest ground truth hand model dataset known

Mean position



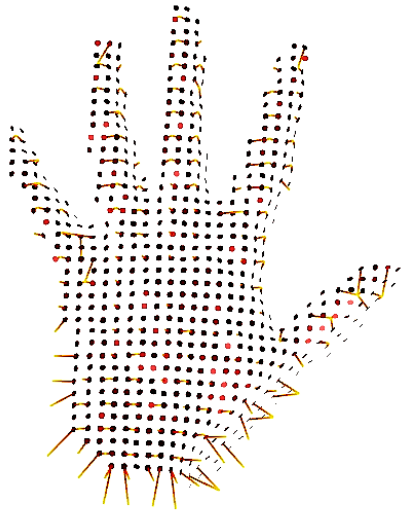


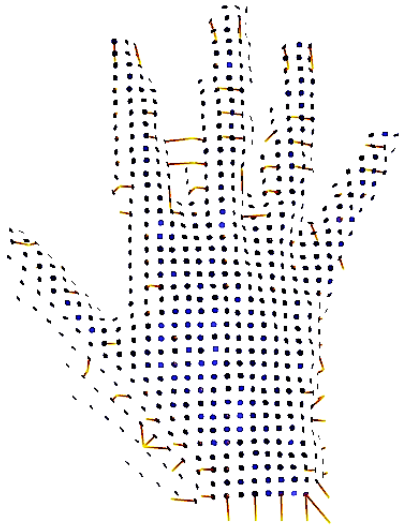
$w_0 +$

$w_1 +$

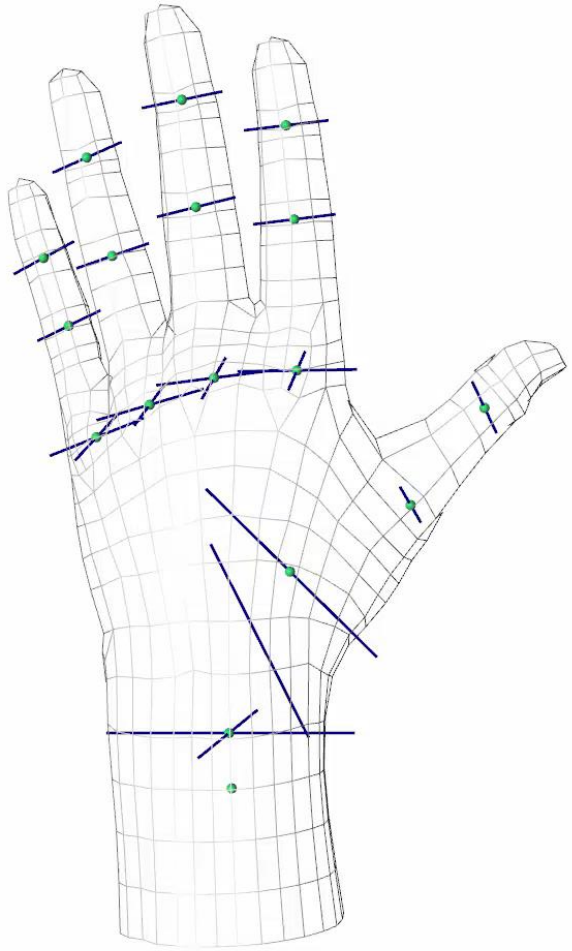
$w_2 + \ldots$

Online hand calibration

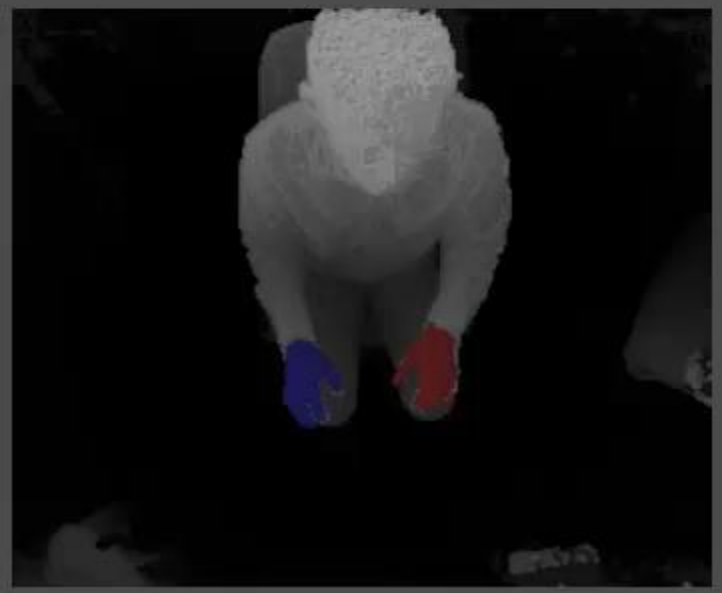




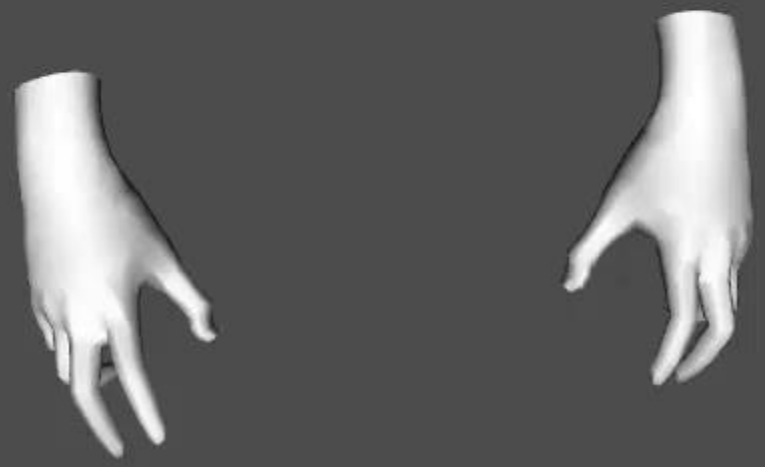
Solve



Isolated hand tracking works well



But, occlusion and body interactions are the common case



Handheld interactions

DodecaPen

Accurate 6DoF Tracking of a Passive Stylus

Po-Chen Wu[*†]
Robert Wang[†]
Kenrick Kin[†]
Christopher Twigg[†]
Shangchen Han[†]
Ming-Hsuan Yang[‡]
Shao-Yi Chien[*]

[*]National Taiwan University
[†]Oculus Research
[‡]UC Merced

Published at UIST 2017!

DodecaPen

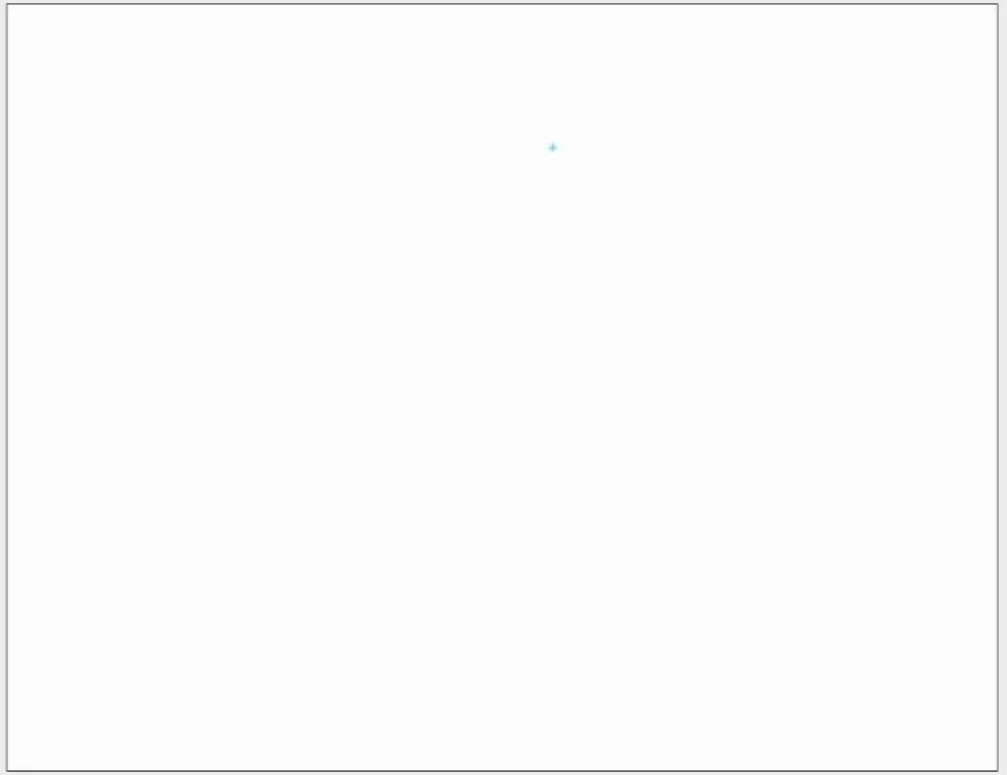
Dodecahedron + Pen

Passive
+
Accurate

&

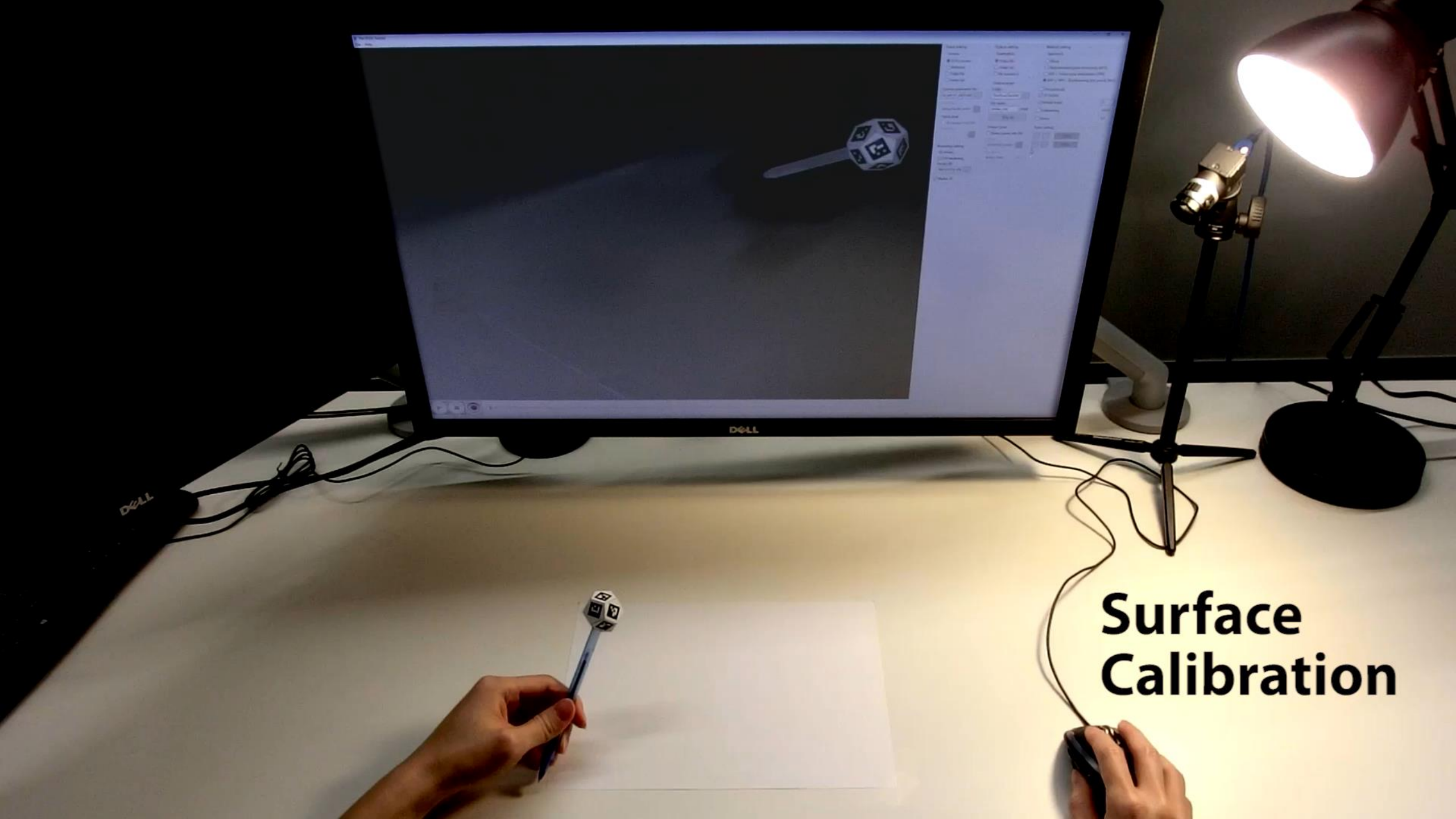

POINT GREY
Manfrotto

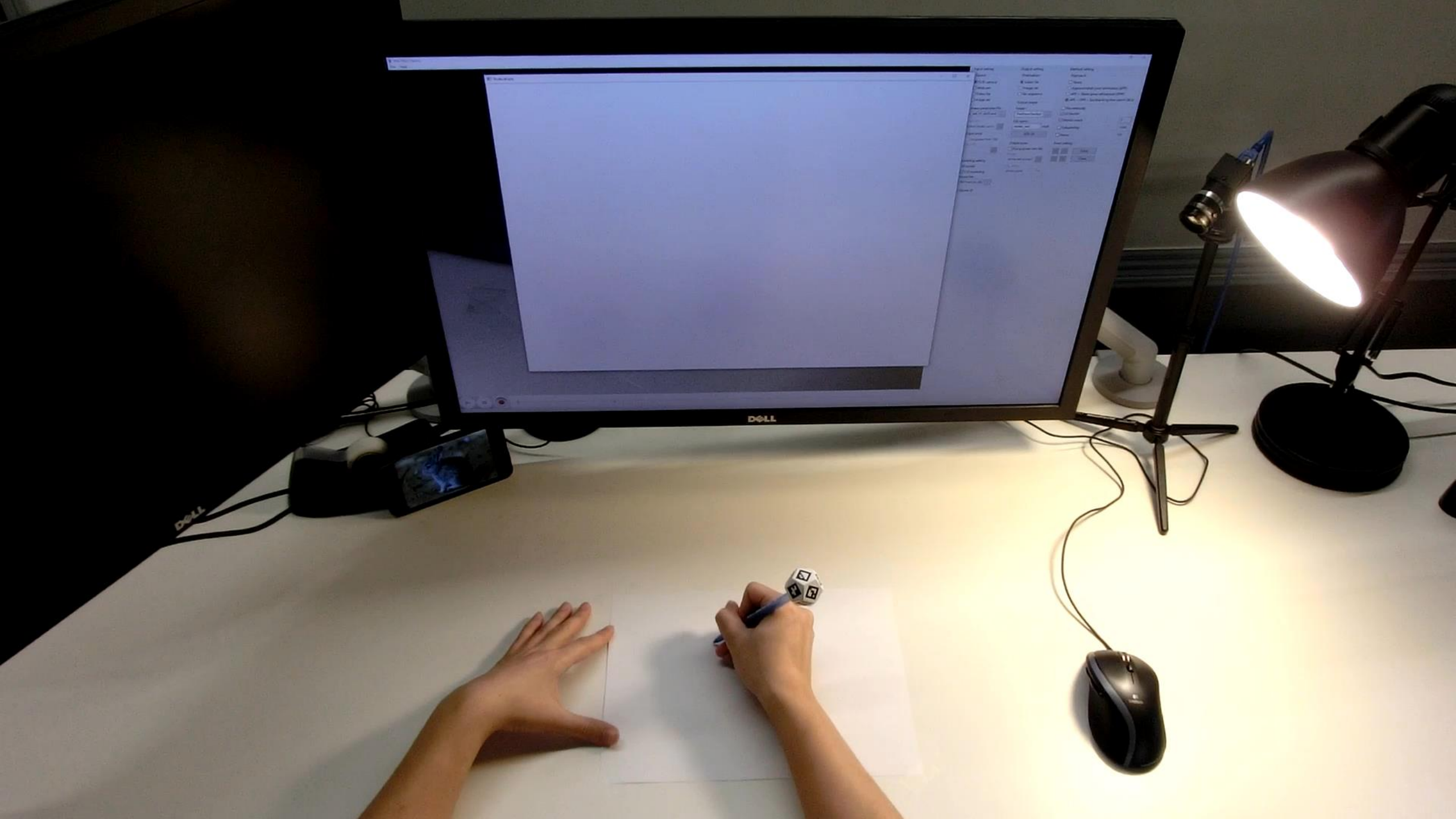
DodecaPen: Puppy



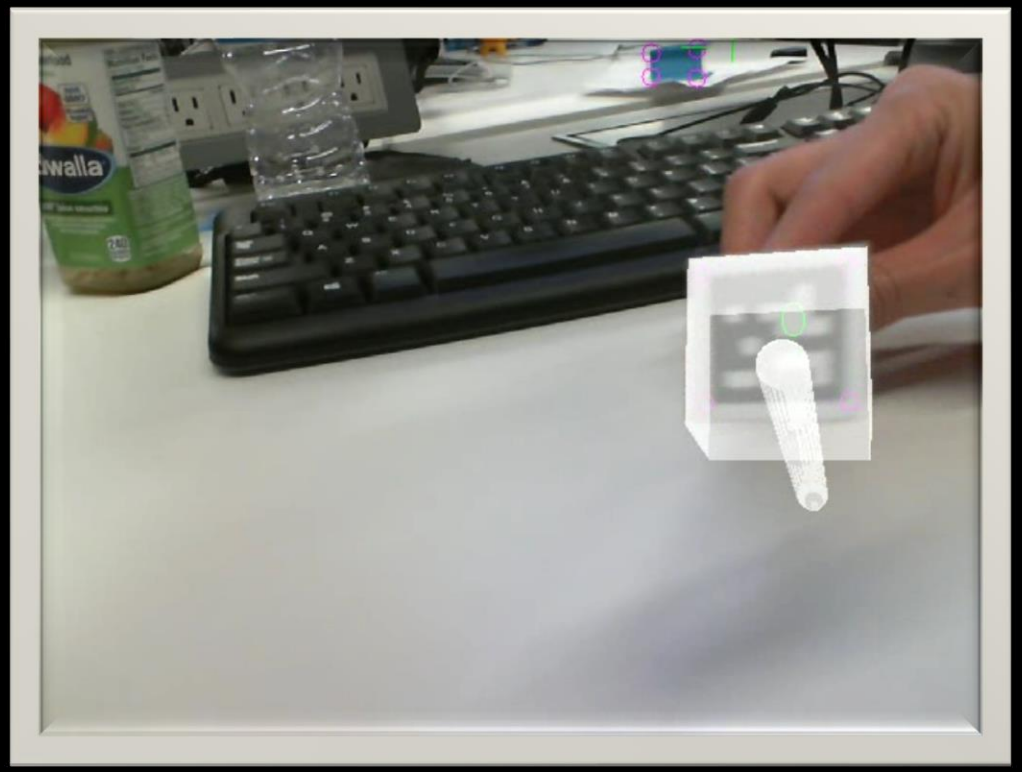
Input Frames

Pen-tip Trajectory

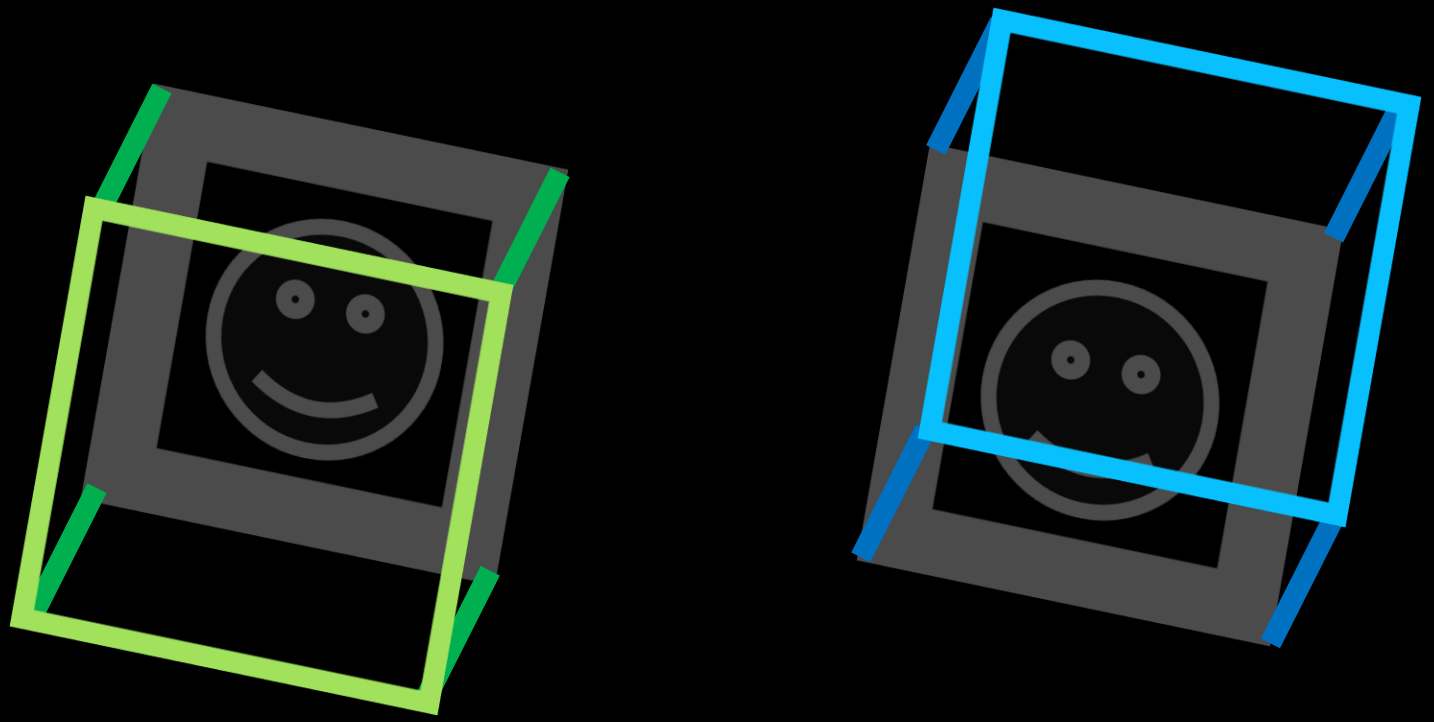

Surface
Calibration



Why Dodecahedron?



Pose Jumping!



Multiple Candidates due to Coplanar Points

Proposed 6DoF Pose Tracking System

Camera + DodecaPen

Input Frames

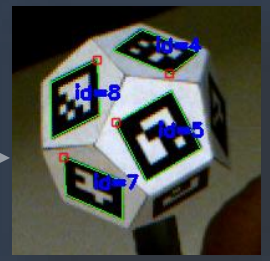
Inter-frame Corner Tracking (**ICT**)
- Pyramidal Lucas-Kanade marker corner tracking
- PnP algorithm to get the initial pose $\mathbf{p}'$

Marker Intensity Normalization

No

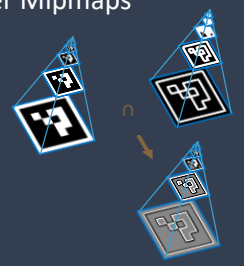
Did APE Succeed?

Yes



Marker Mipmaps

Mipmap Masks

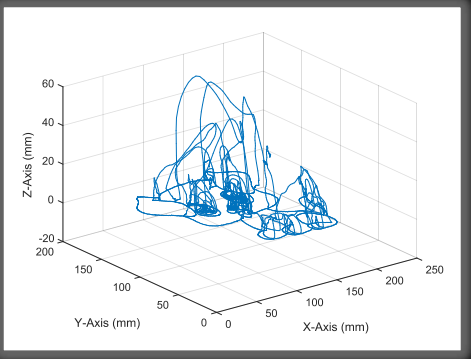


Masked Mipmaps

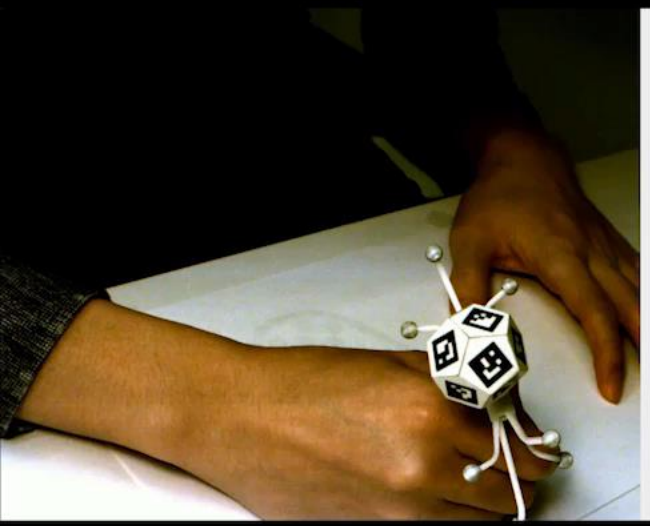
Approximate Pose Estimation (**APE**)
- Marker Detection
- Minimize reprojection error with PnP algorithm to get the initial pose $\mathbf{p}'$

Dense Pose Refinement (**DPR**)
- Minimize appearance distance with Gauss Newton and backtracking line search (**BLS**) to get the final pose $\mathbf{p}^*$
- Marker & mask mipmaps

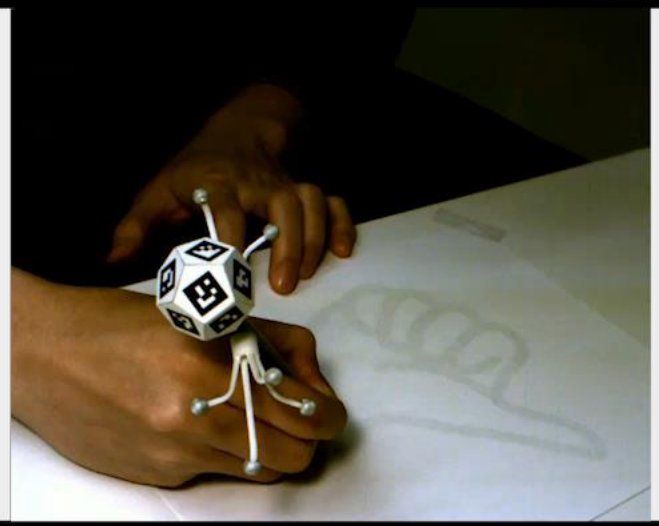
Digital 2D Drawing



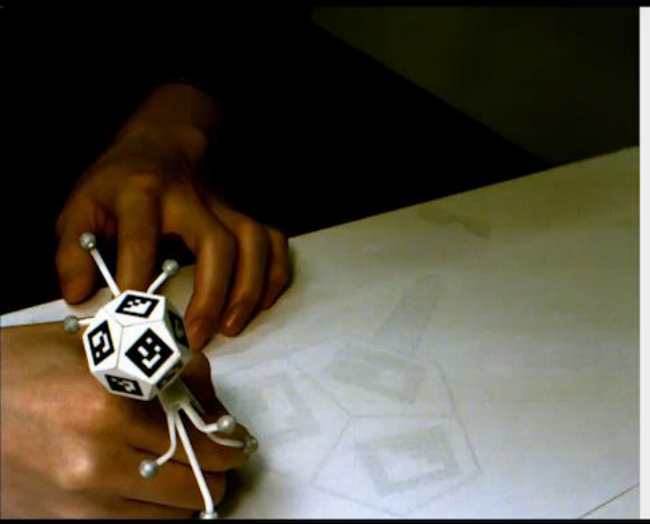
Output Pen-tip Trajectory (Based on DodecaPen Poses)



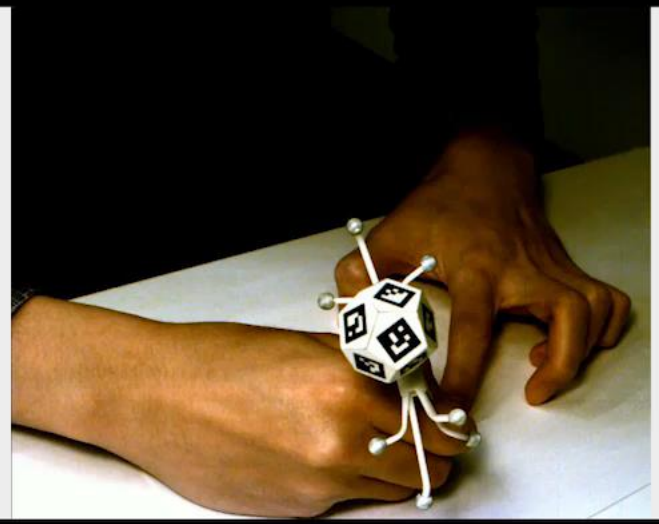
Boba



Thumb

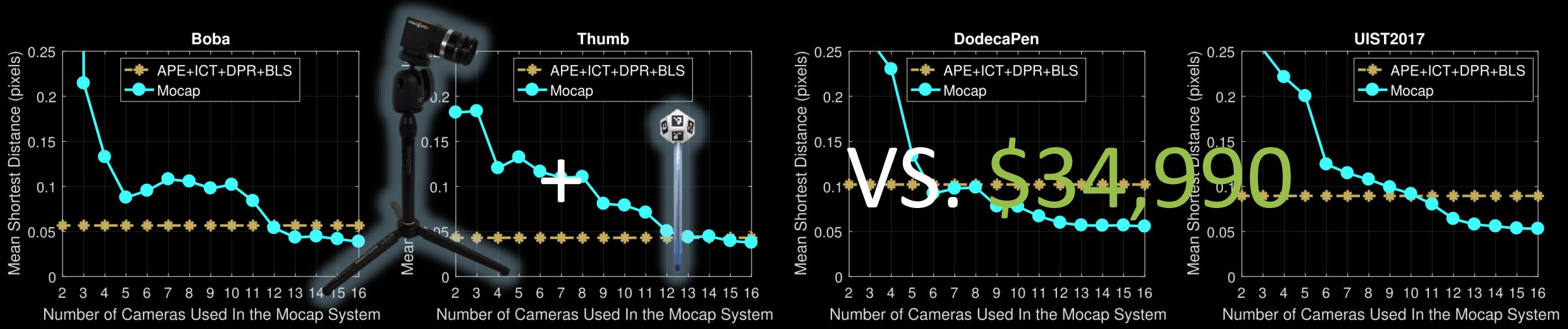


DodecaPen



UIST2017

DodecaPen VS. Mocap

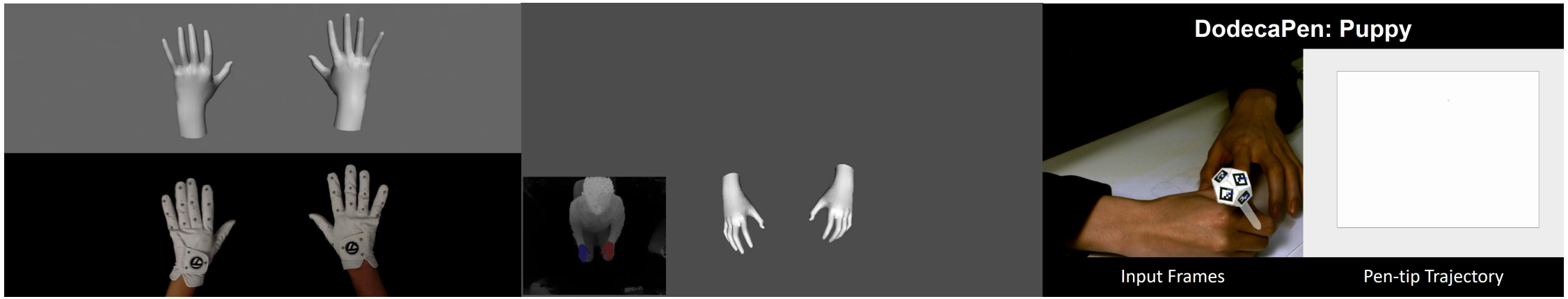


Comparable to a mocap system with 10 active cameras

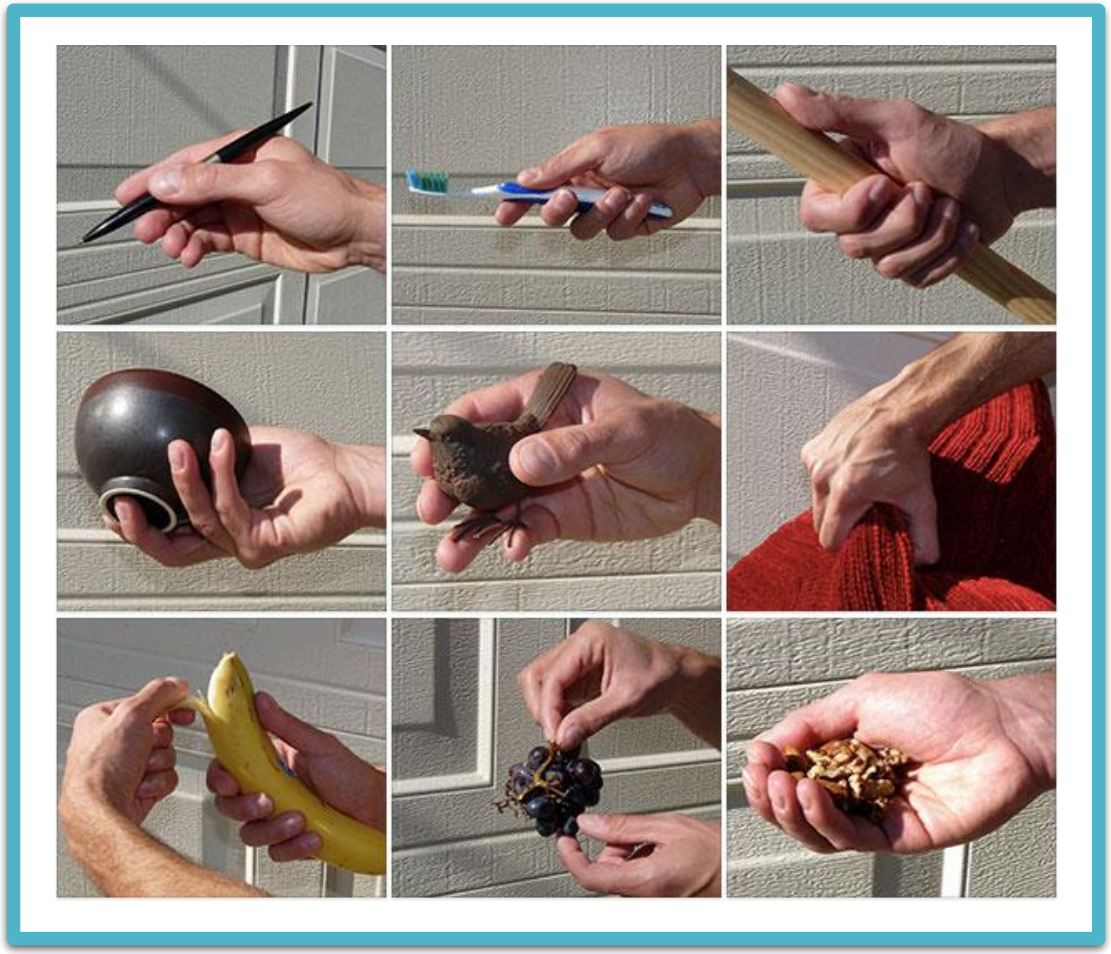
Take Home Message

1. Sub-millimeter accurate 6DoF tracking using a set of readily available and easy-to-assemble components
2. Single camera pose estimation can be fast enough and robust enough for drawing in 2D, 3D and in VR

Hands and interaction tracking



Hands and interaction tracking



From [Victor, “A Brief Rant” 2011]