Fast and Robust Hand Tracking Using Detection-Guided Optimization

Srinath Sridhar¹ Franziska Mueller¹,² Anti Oulasvirta³ Christian Theobalt³

1 Max Planck Institute for Informatics, Saarland University
2 Computer Science Department, Boston University
3 Department of Computer Science, Aalto University

**Results & Evaluation**

AVERAGE FINGERTIP ERROR: DEXTER 1 DATASET

19.6 mm average error

**Input and Model Representation**

- **Input:**
  - Depth features
  - Part labels

- **Model:**
  - 3D hand model

**Part Labelling Random Forest**

- Random forest trained with 50,000 real examples
- Multi-class classification into 11 hand parts
- 3 trees each with 22 levels
- 4000 feature tests per node

**Pose Optimization**

- **Objective:**
  - Maximize similarity between C₁ and C₂ while avoiding collisions and choosing smooth, biomechanically plausible poses

- **Formulation:**
  - Mixture models

**Model Fitting**

- **Objective:**
  - Detect hand model
  - Account for pose variation in hand width, height and thickness
  - Fitting takes < 1 second

Depth-only similarity
Detection-guided similarity

- Pose particles are optimized using depth-only and detection-guided energies

- Particle with best final pose fitting energy is chosen

**Contributions**

- **DETECTION-GUIDED POSE OPTIMIZATION**
  - Particle-based pose optimization
  - Depth and part labels

**Hand Tracking**

- **Challenges:**
  - Complex hand motions
  - Uniform skin color
  - Self occlusions
  - Varying camera-scene configurations
  - Real-time pose optimization

- **PRACTICAL:**
  - 25 FPS without GPU
  - Supports arbitrary moving camera-scene arrangements

**Interaction with Future Devices**

- Impractical to use keyboard and mouse. Hands are mobile, accessible and always with us.