GANerated Hands for Real-Time 3D Hand Tracking from Monocular RGB

Franziska Mueller[1,2] Florian Bernard[1,2] Oleksandr Sotnychenko[1,2]
Dushyant Mehta[1,2] Srinath Sridhar[3] Dan Casas[4] Christian Theobalt[1,2]

[1] max planck institut informatik
[2] SIC Saarland Informatics Campus
[3] Stanford University
[4] Universidad Rey Juan Carlos

Motivation & Challenges

Natural Interaction

Robotics

Activity Recognition

- **Complex** and **fast** motions
- **High-dimensional** problem
- **Runtime** constraint

- (**Self-**)**occlusions** and **self-similarities**
- Depth and scale **ambiguity**
- **Hard to annotate** data in **3D**

Contributions

Real-time full **3D** hand tracking from **monocular RGB** video

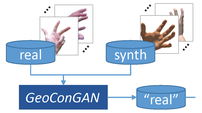
Single RGB camera/ community video

@ 50 FPS

occlusion-robust

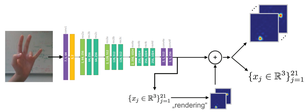
various viewpoints

Technical Novelties



New GAN for **geometrically consistent unpaired image-to-image translation**

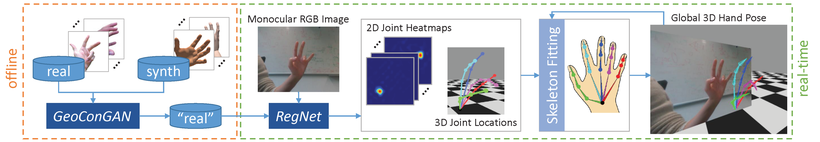
Novel **enhanced RGB dataset** with **3D** hand joint **annotations** (>260k frames)



CNN with **projection layer** for tightly coupled regression of **2D and 3D joint locations**

Smooth and plausible global 3D hand pose by combination with kinematic model

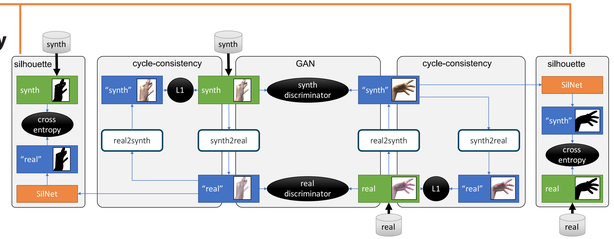
Method



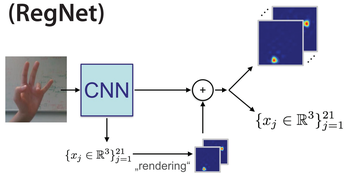
GANerating Training Data (GeoConGAN)

Geometric consistency loss:

- forces hand pose to not change during translation, hence
- perfect synthetic annotation can be transferred



Joint Position Regression (RegNet)



- **2D joint location heatmaps:** determine global position
- **Root-relative 3D joint positions:** distinguish hand poses with the same 2D projection

Kinematic Skeleton Fitting

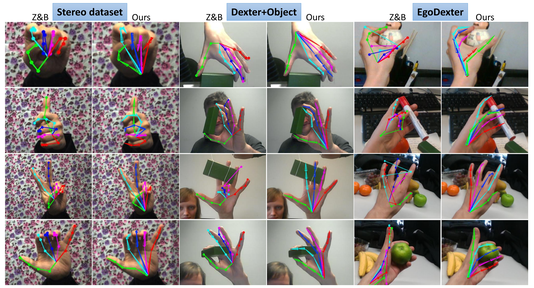
Minimize energy:

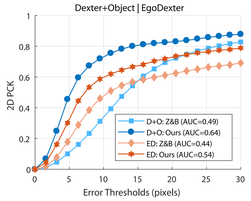
$$E(\Theta) = E_{2D}(\Theta) \quad \text{2D Inverse Kinematics}$$
$$+ E_{3D}(\Theta) \quad \text{3D Inverse Kinematics}$$
$$+ E_{\text{limits}}(\Theta) \quad \text{Joint Angle Limits}$$
$$+ E_{\text{temp}}(\Theta) \quad \text{Temporal Smoothness}$$

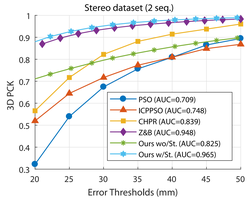
Final output: parameters of kinematic model (global transform, joint angles)

Results

Comparison to Zimmermann and Brox, ICCV 2017

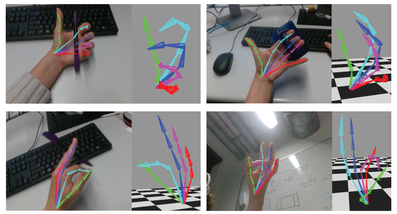




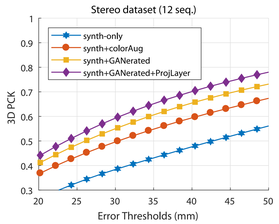


Community Video

Live Capture Setup



Ablative Analysis



GANeratedHands dataset available!
handtracker.mpi-inf.mpg.de

Also check out our other datasets at:
gvverfcapeva.mpi-inf.mpg.de

Graphics Vision & Video

Supported by the ERC Starting Grant project CapReal (335545).
Dan Casas was supported by a Marie Curie Individual Fellow, grant agreement 707326.