MS-MANO: Enabling Hand Pose Tracking with Biomechanical Constraints

Pengfei Xie[1,*], Wenqiang Xu[2,*], Tutian Tang[2], Zhenjun Yu[2], Cewu Lu[2]

[1]Southeast University [2]Shanghai Jiao Tong University [*]Equal Contribution

Limitations of Previous Visual Hand Dynamic Analysis

Focus Mostly on Static Estimations

$$\text{Position:} \quad \mathbf{x}(t) \tag{1}$$

Overlooked Dynamic Information

$$\text{Velocity:} \quad \dot{\mathbf{x}}(t), \quad \text{Acceleration:} \quad \ddot{\mathbf{x}}(t) \tag{2}$$

Limitations of Previous Visual Hand Dynamic Analysis

Mistakenly Treat Human Hand as **Multi-Body Systems**

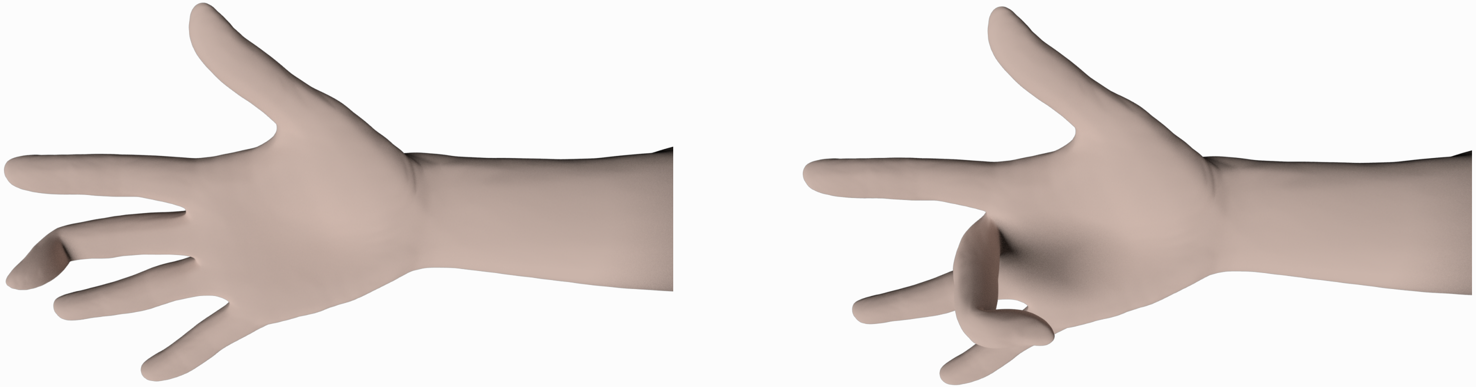

Figure: The illustrations showcase two hand configurations: **Left.** Fingers bent only at the tip joints. **Right.** Fingers bent at both tip and base joints. These poses are achievable for robots, but not human hands due to joint limitations.

Musculoskeletal Model

- Explicitly emulates the dynamics of muscles & tendons to drive the skeleton
- Impose physiologically realistic constraints on the resulting torque trajectories

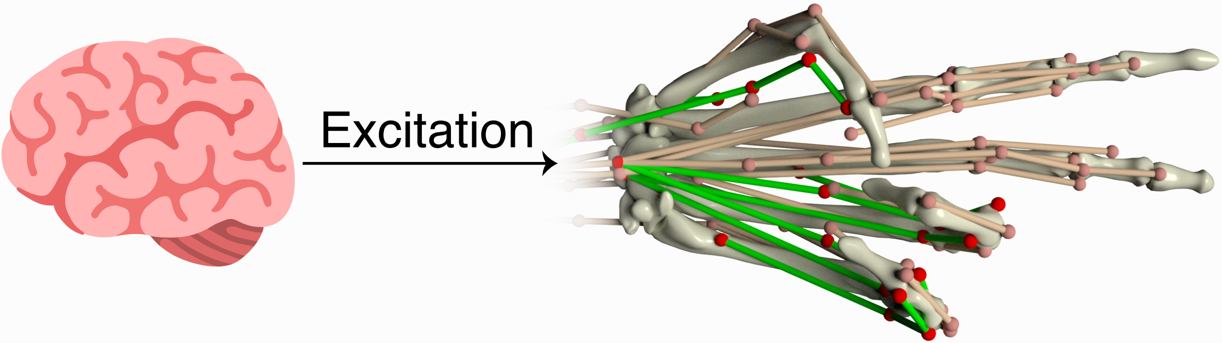

Figure: The excitation signal originating from the brain triggers the contraction or relaxation of muscles. The triggered muscle segments are illustrated in green, while the relaxed ones are in brown.

Our Contributions

To address the limitations of previous visual hand dynamic analysis methods, we

- Introduce **MS-MANO**, a musculoskeletal extension of the MANO hand model, with learning support and shape adaptability.
- Demonstrate MS-MANO's effectiveness in hand pose tracking using the **BioPR** framework, benchmarking on DexYCB and OakInk datasets.

Figure: MS-MANO model.

MusculoSkeletal-MANO

MS-MANO integrates hill-type-based musculoskeletal dynamics with MANO, mapping OpenSim's bone-centric muscles into a joint-centric representation.

(a) (b) (c)

Figure: Joint-centric muscle adaptation. **(a)** A set of smaller bones is mapped into a single joint. **(b)** Bone-centric muscle segments can adapt to different shapes. **(c)** (Left) Raw skeleton after automatic mapping results intersection. (Right) Revised skeleton can perfectly fit with MANO model.

Biomechanical Pose Refiner

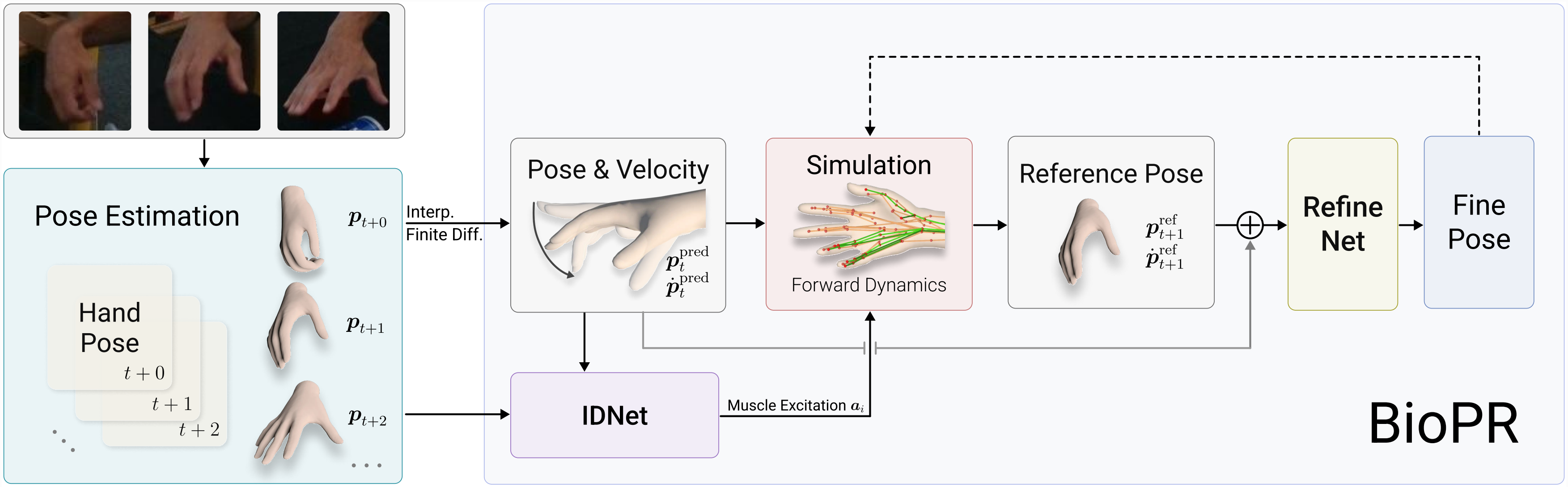

Figure: BioPR's simulation-in-the-loop pipeline. It interpolates and differentiates poses, infers muscle excitation signals using IDNet, and generates the next reference pose via forward dynamics. The Refine Net performs final refinement based on initial pose, velocity, and reference pose.

Results

Methods	MPJPE↓	AUC↑	AE↓
VIBE	16.95	67.5	36.4
TCMR	16.03	70.1	34.3
MeshGraphormer	16.21	69.1	35.9
gSDF	14.4	89.1	30.3
gSDF + BioPR	**12.81**	**89.7**	**29.9**
Deformer	13.64	89.6	31.7
Deformer + BioPR	**12.92**	**90.4**	**30.7**

Table: Quantitative results.

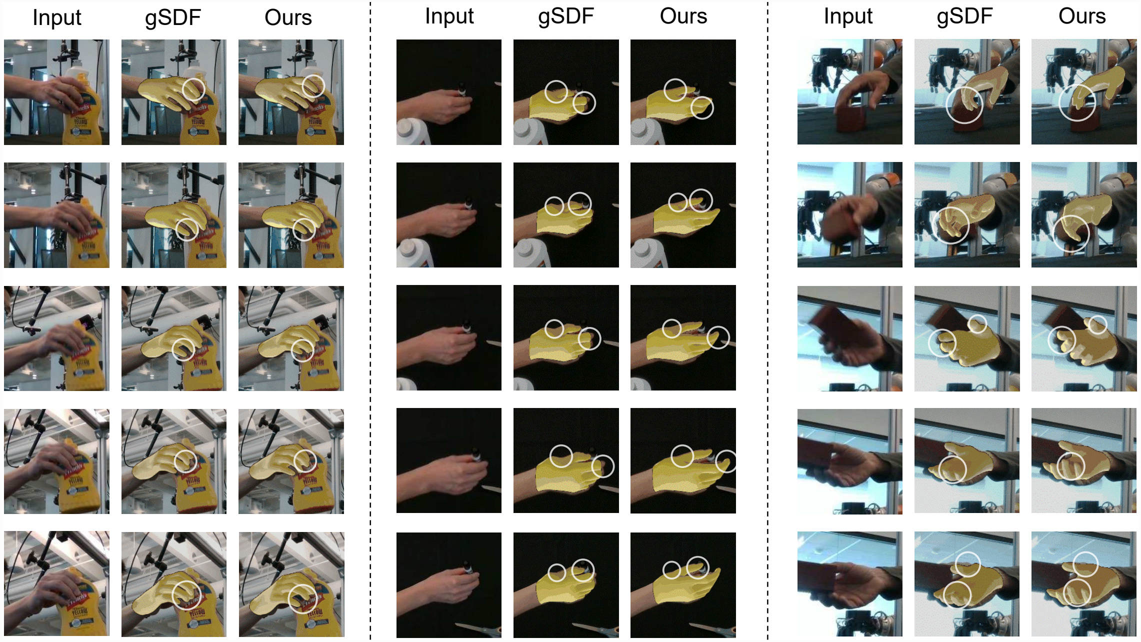

Figure: Qualitative results.

Conclusion

- We propose **MS-MANO**, a musculoskeletal extension of the MANO hand model, to enable hand pose tracking with biomechanical constraints.
- We introduce **BioPR**, a simulation-in-the-loop framework, to refine hand pose estimates by incorporating muscle dynamics.
- We demonstrate the effectiveness of MS-MANO and BioPR on the DexYCB and OakInk datasets.