CaPhy: Capturing Physical Properties for Animatable Human Avatars

Zhaoqi Su[1], Liangxiao Hu[2], Siyou Lin[1], Hongwen Zhang[1], Shengping Zhang[2], Justus Thies[3], Yebin Liu[1]

[1]Tsinghua University, [2]Harbin Institute of Technology, [3]Max Planck Institute for Intelligent Systems
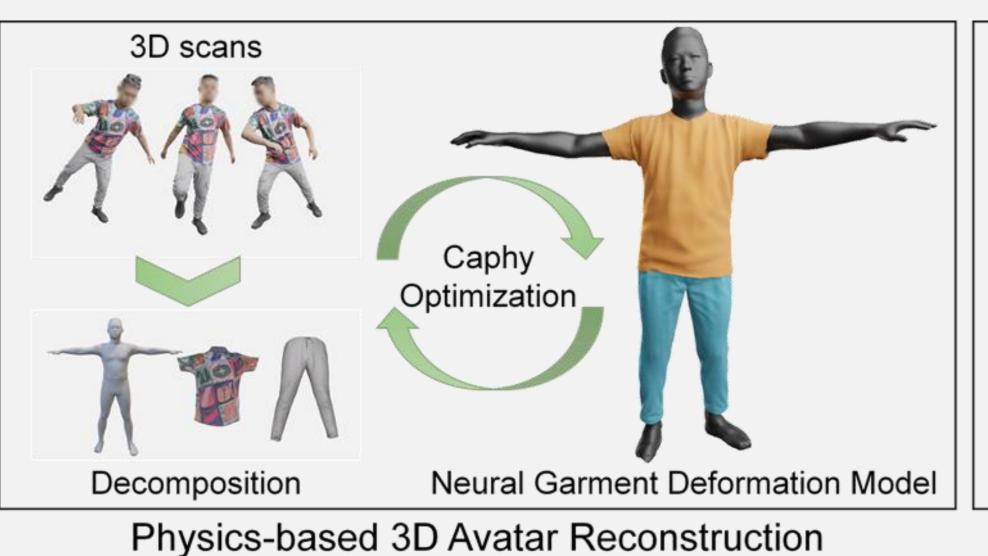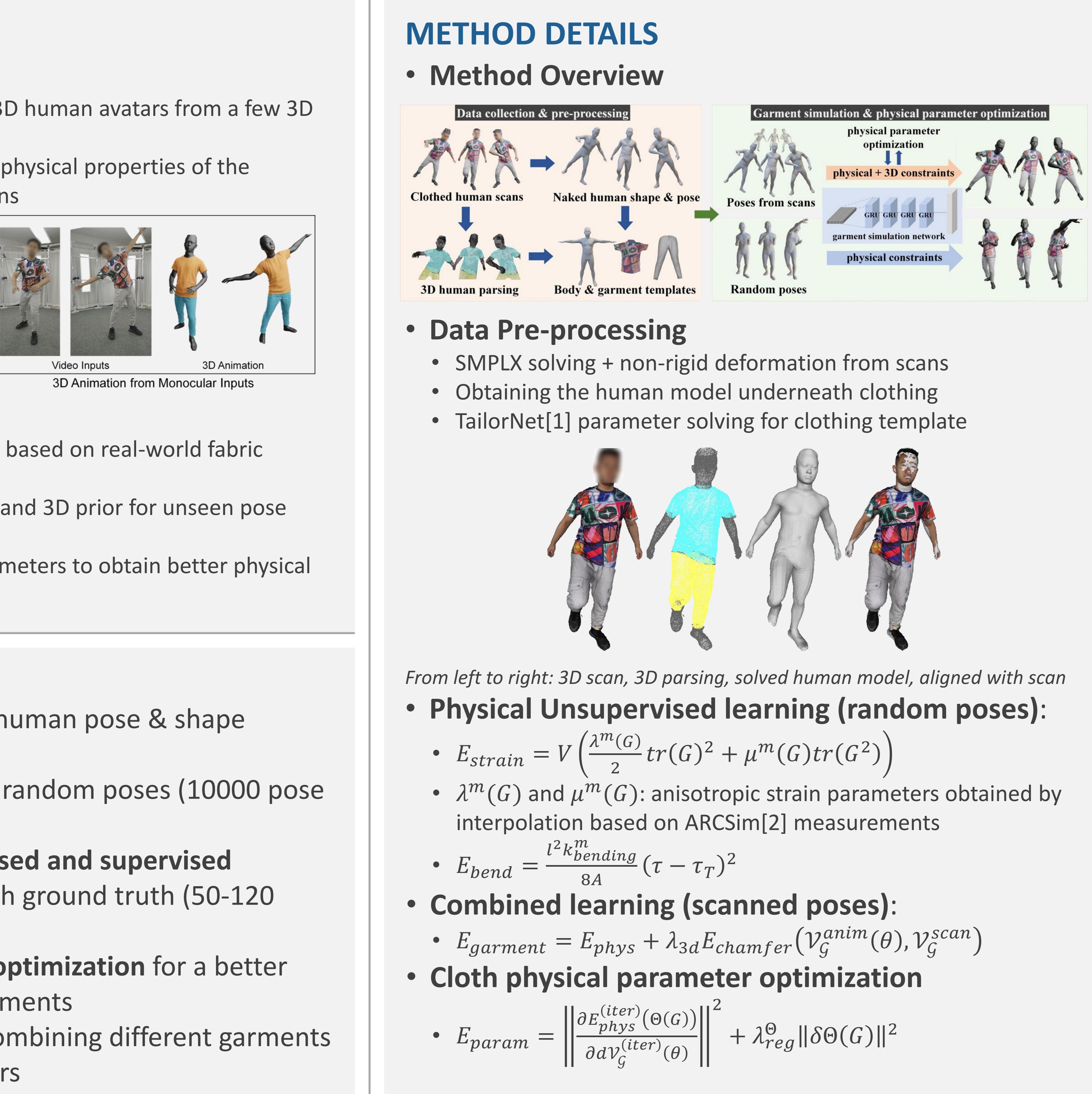
ICCV23 PARIS

INTRODUCTION

- **Goal**:
  - A method for physics-based 3D human avatars from a few 3D scans
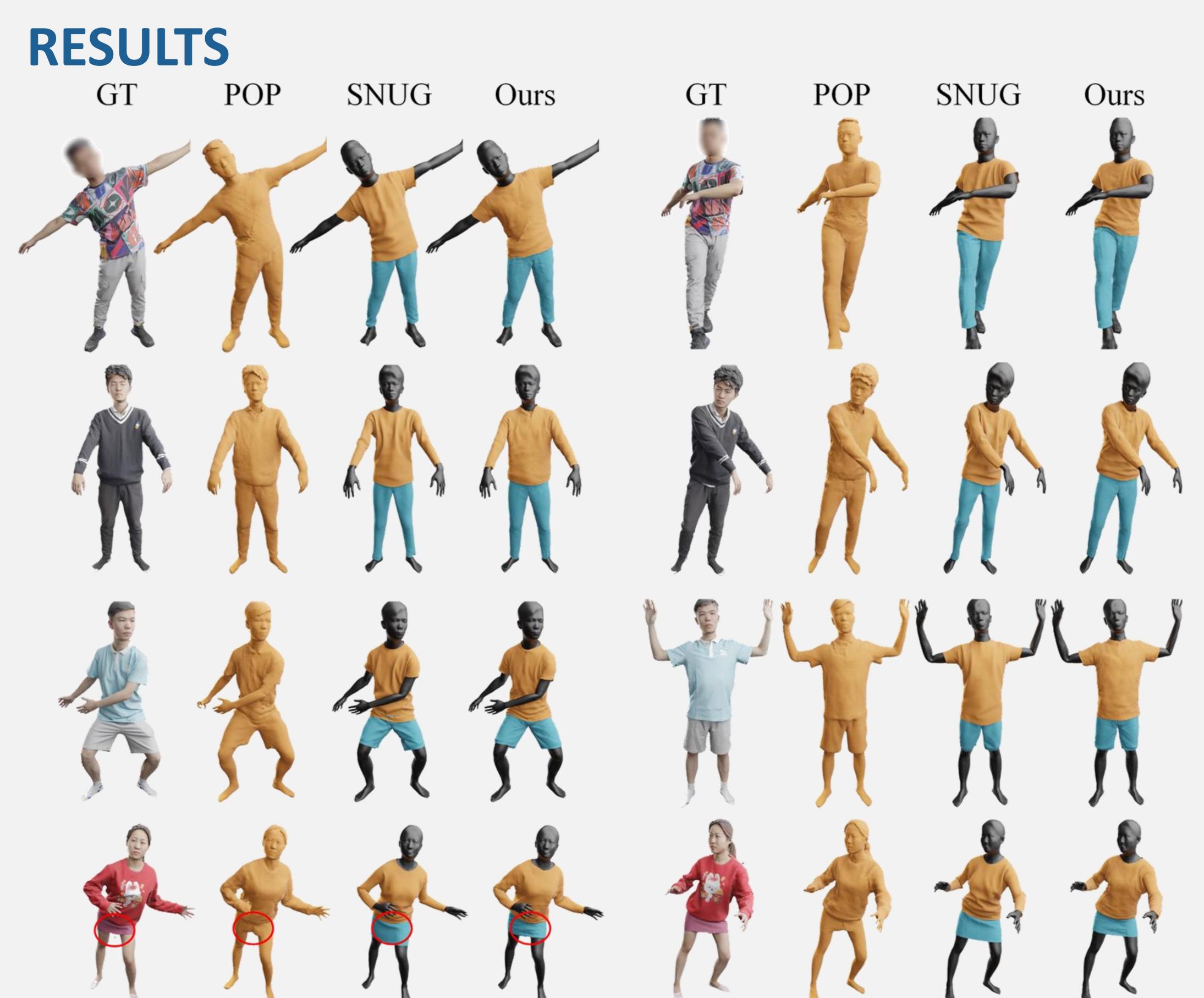 - Capturing the geometric and physical properties of the clothing from real observations

Physics-based 3D Avatar Reconstruction

3D Animation from Monocular Inputs

- **Contributions**:
  - A physical model formulation based on real-world fabric measurement
  - Combination of physics prior and 3D prior for unseen pose generation
  - Optimization of physical parameters to obtain better physical properties

KEY IDEAS

- **Pre-processing**: obtaining human pose & shape underneath clothing
- **Unsupervised learning** for random poses (10000 pose sequences)
- **Combination of unsupervised and supervised learning** for scan poses with ground truth (50-120 static poses)
- **Cloth physical parameter optimization** for a better physical formulation of garments
- **Collision Fine-tuning** for combining different garments with different physical priors

METHOD DETAILS

- **Method Overview**

Data collection & pre-processing

Clothed human scans → Naked human shape & pose

3D human parsing → Body & garment templates

Garment simulation & physical parameter optimization

physical parameter optimization

Poses from scans

physical + 3D constraints

garment simulation network

GRU GRU GRU GRU

physical constraints

Random poses

- **Data Pre-processing**
  - SMPLX solving + non-rigid deformation from scans
  - Obtaining the human model underneath clothing
  - TailorNet[1] parameter solving for clothing template

From left to right: 3D scan, 3D parsing, solved human model, aligned with scan

- **Physical Unsupervised learning (random poses)**:
  - $E_{strain} = V\left(\frac{\lambda^m(G)}{2}tr(G)^2 + \mu^m(G)tr(G^2)\right)$
  - $\lambda^m(G)$ and $\mu^m(G)$: anisotropic strain parameters obtained by interpolation based on ARCSim[2] measurements
  - $E_{bend} = \frac{l^2 k_{bending}^m}{8A}(\tau - \tau_T)^2$

- **Combined learning (scanned poses)**:
  - $E_{garment} = E_{phys} + \lambda_{3d}E_{chamfer}\left(\mathcal{V}_\mathcal{G}^{anim}(\theta), \mathcal{V}_\mathcal{G}^{scan}\right)$

- **Cloth physical parameter optimization**
  - $E_{param} = \left\|\frac{\partial E_{phys}^{(iter)}(\Theta(G))}{\partial d\mathcal{V}_\mathcal{G}^{(iter)}(\theta)}\right\|^2 + \lambda_{reg}^\Theta\|\delta\Theta(G)\|^2$

RESULTS

GT POP SNUG Ours   GT POP SNUG Ours

Test set results compared with POP[3] and SNUG[4]. Our results are visually more similar to the ground-truth observations.

Two examples of texture reconstruction results. Each from left to right: the test set model, results without texture, and results with texture.

REFERENCE

[1] Patel et al. TailorNet: Predicting Clothing in 3D as a Function of Human Pose, Shape and Garment Style. CVPR 2020.

[2] Wang et al. Data-driven elastic models for cloth: modeling and measurement. TOG 2011.

[3] Ma et al. The Power of Points for Modeling Humans in Clothing. ICCV 2021.

[4] Santesteban et al. SNUG: Self-Supervised Neural Dynamic Garments. CVPR 2022.