

Chupa: Carving 3D Clothed Humans from Skinned Shape Priors using 2D Diffusion Probabilistic Models



WEB TOON

Byungjun Kim^{1*}

Patrick Kwon^{2*}

Kwangho Lee²

Myunggi Lee²

Sookwan Han¹

Daesik Kim²

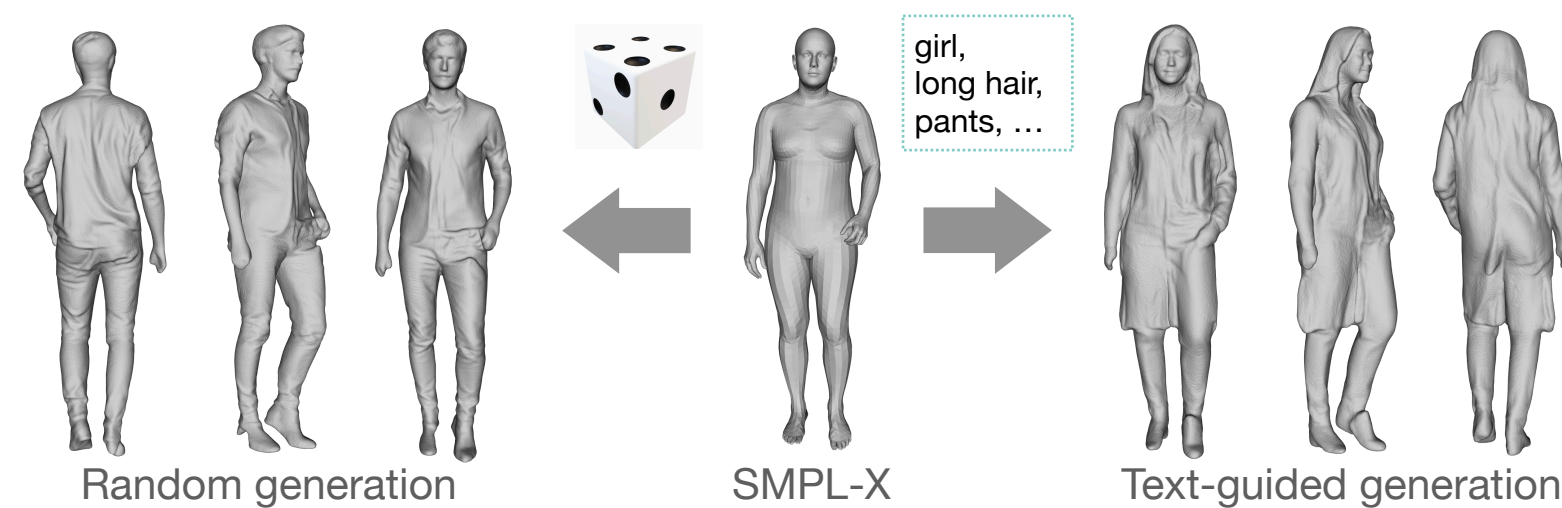
Hanbyul Joo¹

¹Seoul National University

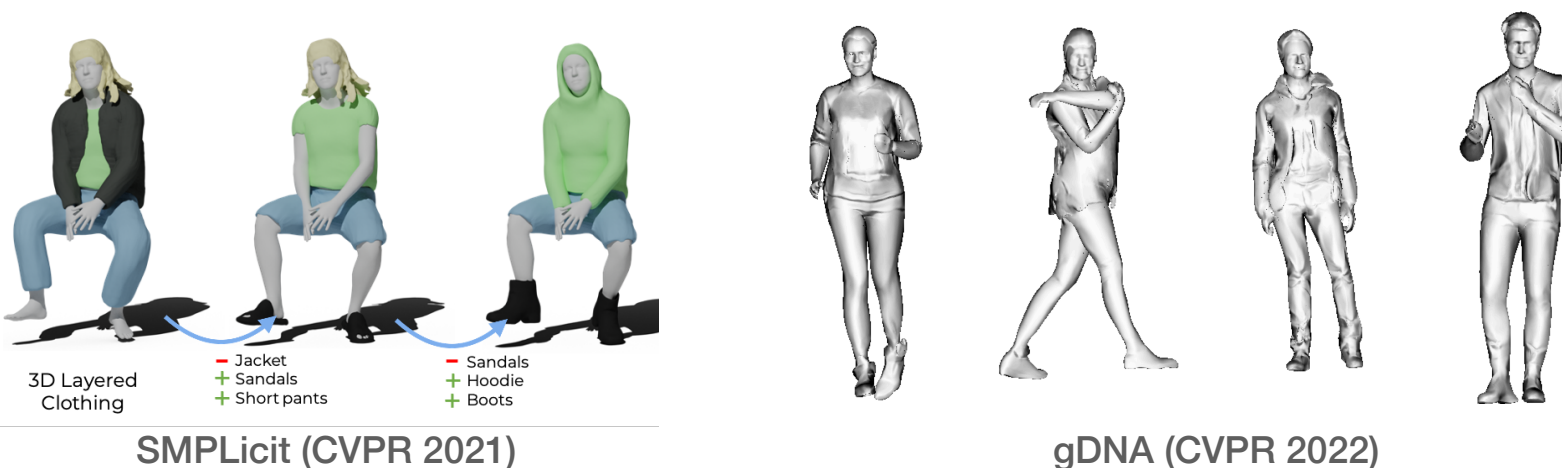
²Naver Webtoon AI

1. Goal

Generate a high-fidelity 3D human mesh from SMPL-X and text prompts



2. Previous Work



- SMPL++ approach
- Rely on templates for body & clothing

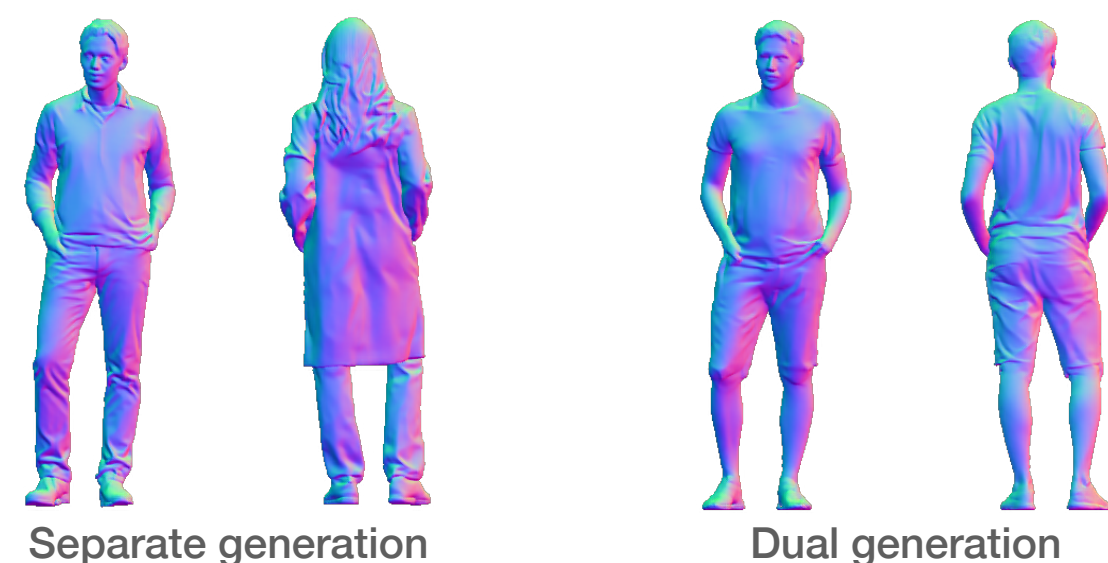
Limited diversity

- 3D implicit field-based approach
- GAN loss for geometric details

Mode collapse, limited quality

3. Key Insights

- Decompose 3D generation into **2D normal map generation + 3D reconstruction**
- **Generating front and back at once** gives holistic geometric details of human

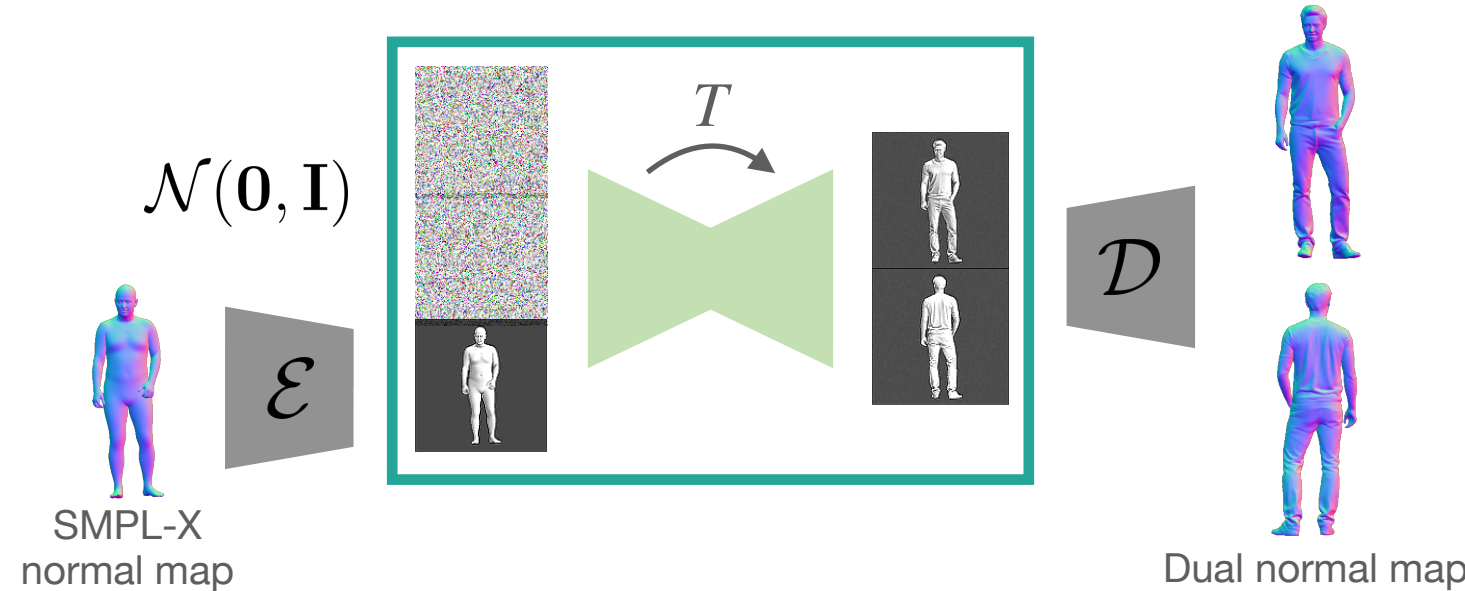


Separate generation

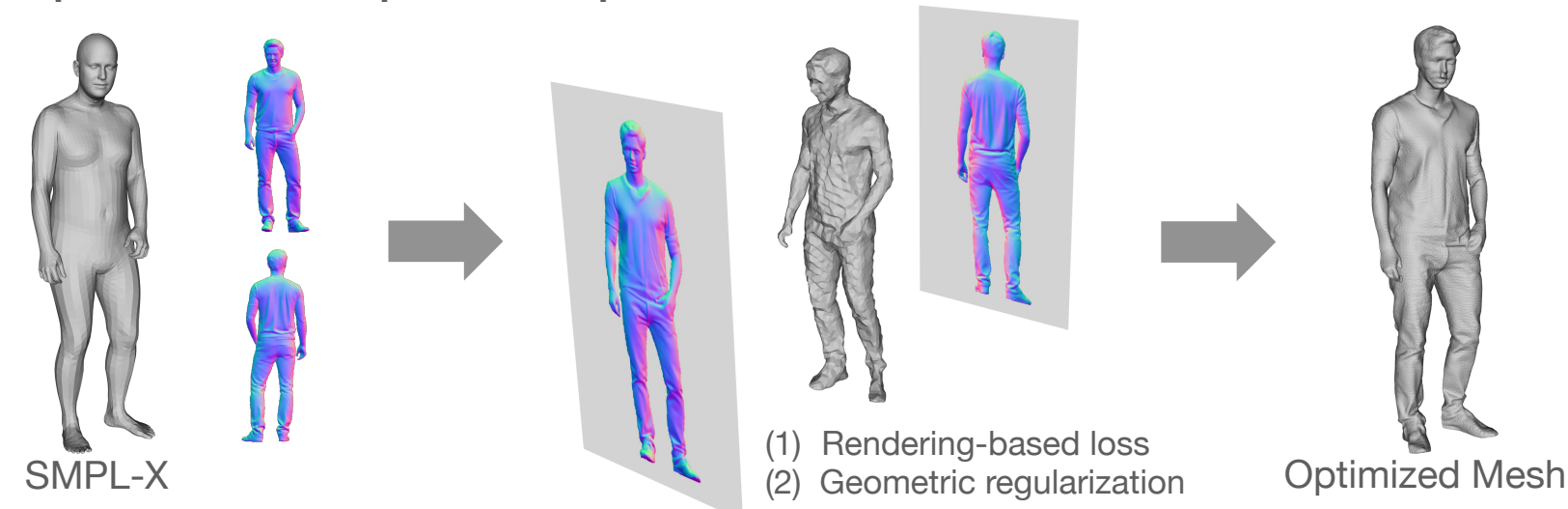
Dual generation

4. Method

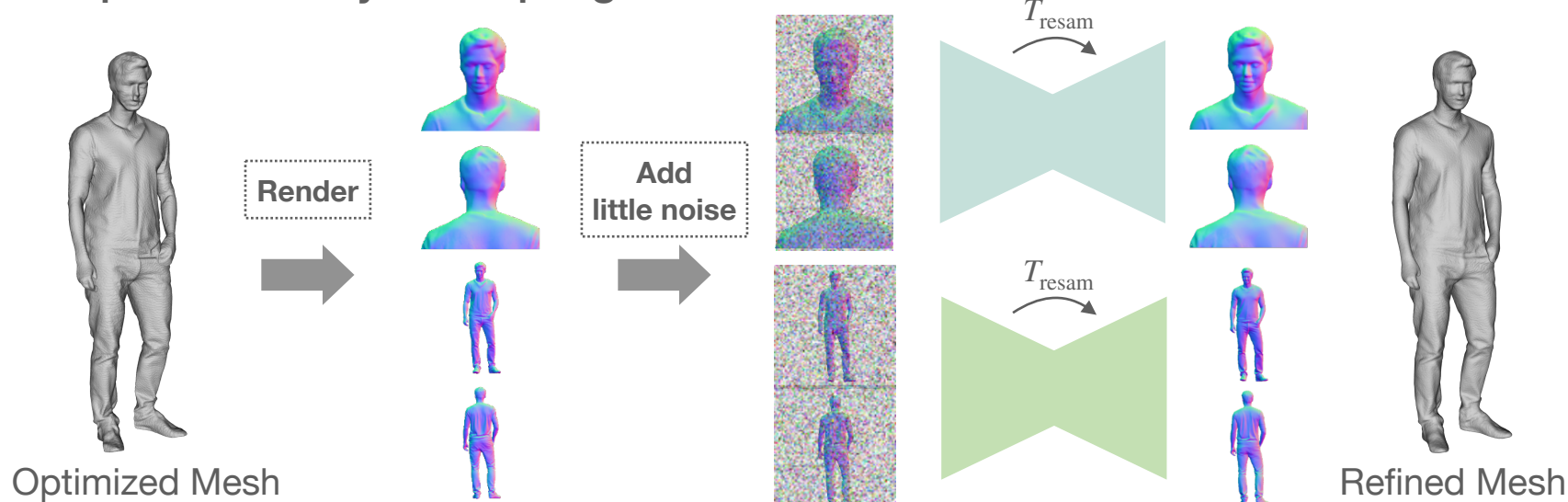
Step 1. Dual normal map generation via diffusion



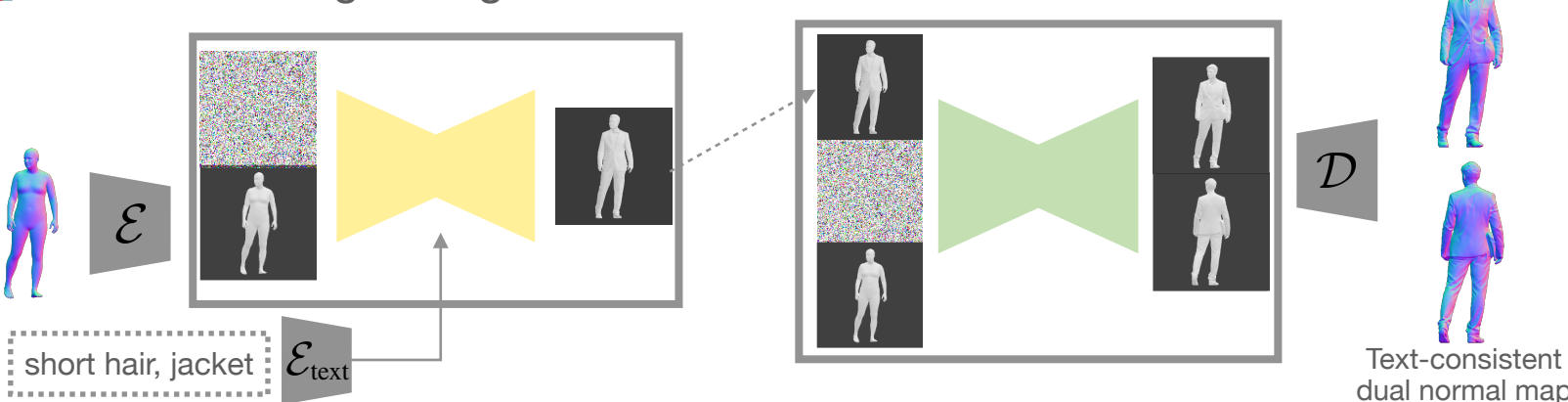
Step 2. Normal map-based optimization



Step 3. Refine by resampling



Extension: Text-guided generation



5. Results

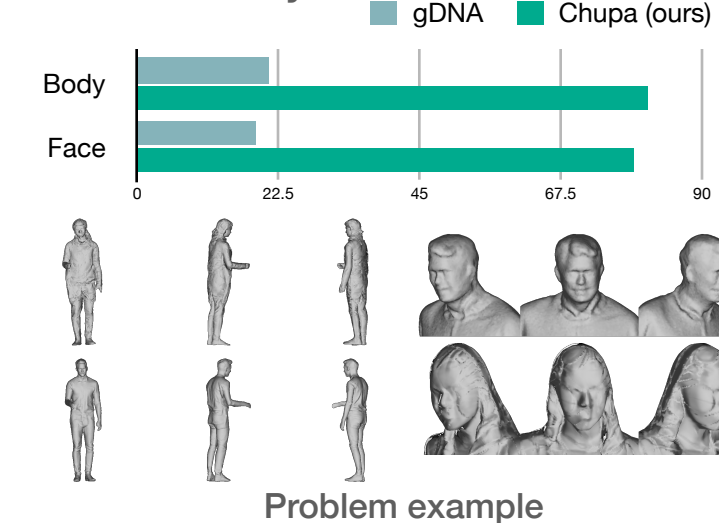
Quantitative results

Method	FID _{normal}	FID _{shade}
gDNA _{coarse} [1]	53.74	68.14
gDNA _{fine} [1]	36.43	45.57
Ours	21.90	36.58

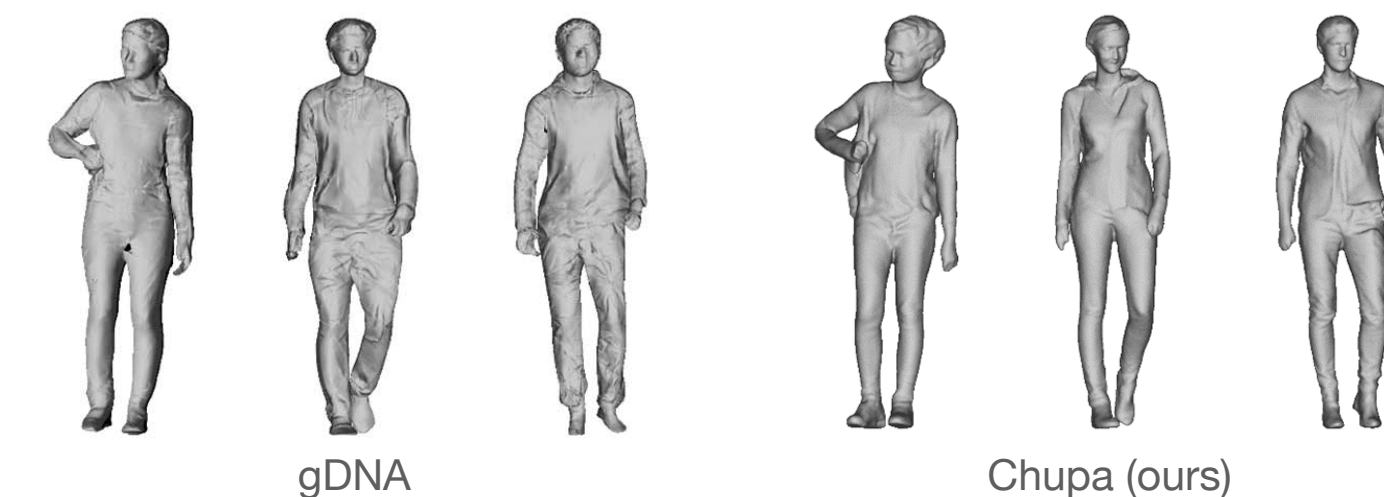
Ablation study

dual.	L _{sides}	refine _{body}	refine _{face}	FID _{normal}
				30.55
✓				26.31
✓	✓			25.50
✓	✓	✓		22.61
✓	✓	✓	✓	21.90

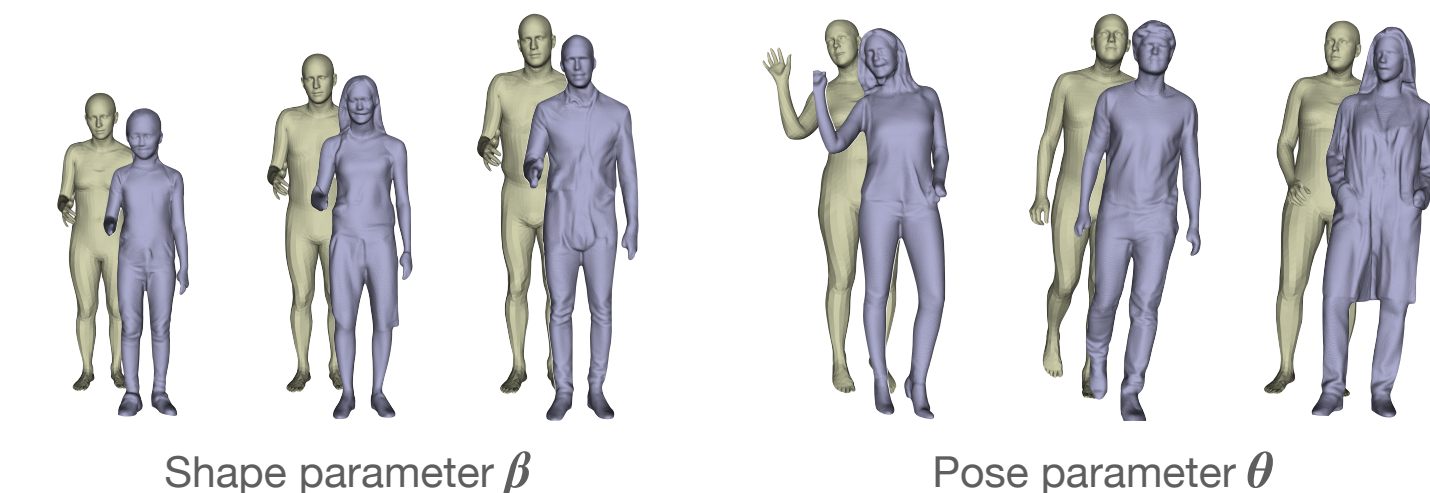
User study



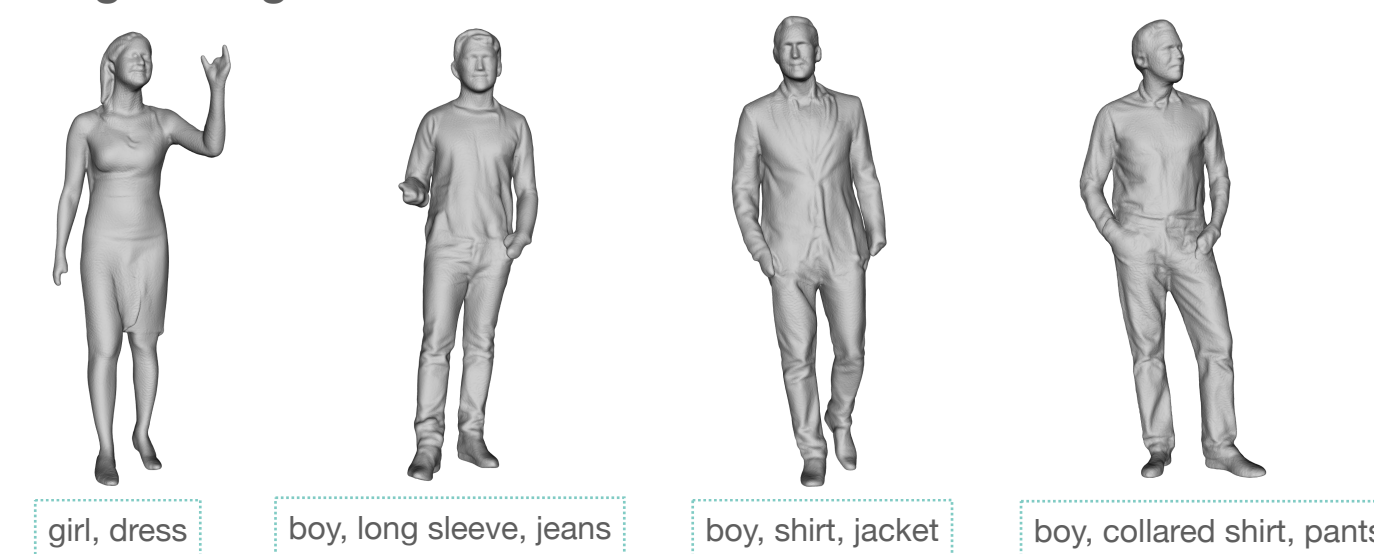
Qualitative comparison with gDNA



Control SMPL-X parameters



Text-guided generation



[1] Chen, Xu, et al. "gdna: Towards generative detailed neural avatars.", CVPR 2022
 [2] Corona, Enric, et al. "Smplicit: Topology-aware generative model for clothed people.", CVPR 2021