Ani3DHuman: Photorealistic 3D Human Animation with Self-guided Stochastic Sampling

City University of Hong Kong 香港城市大學

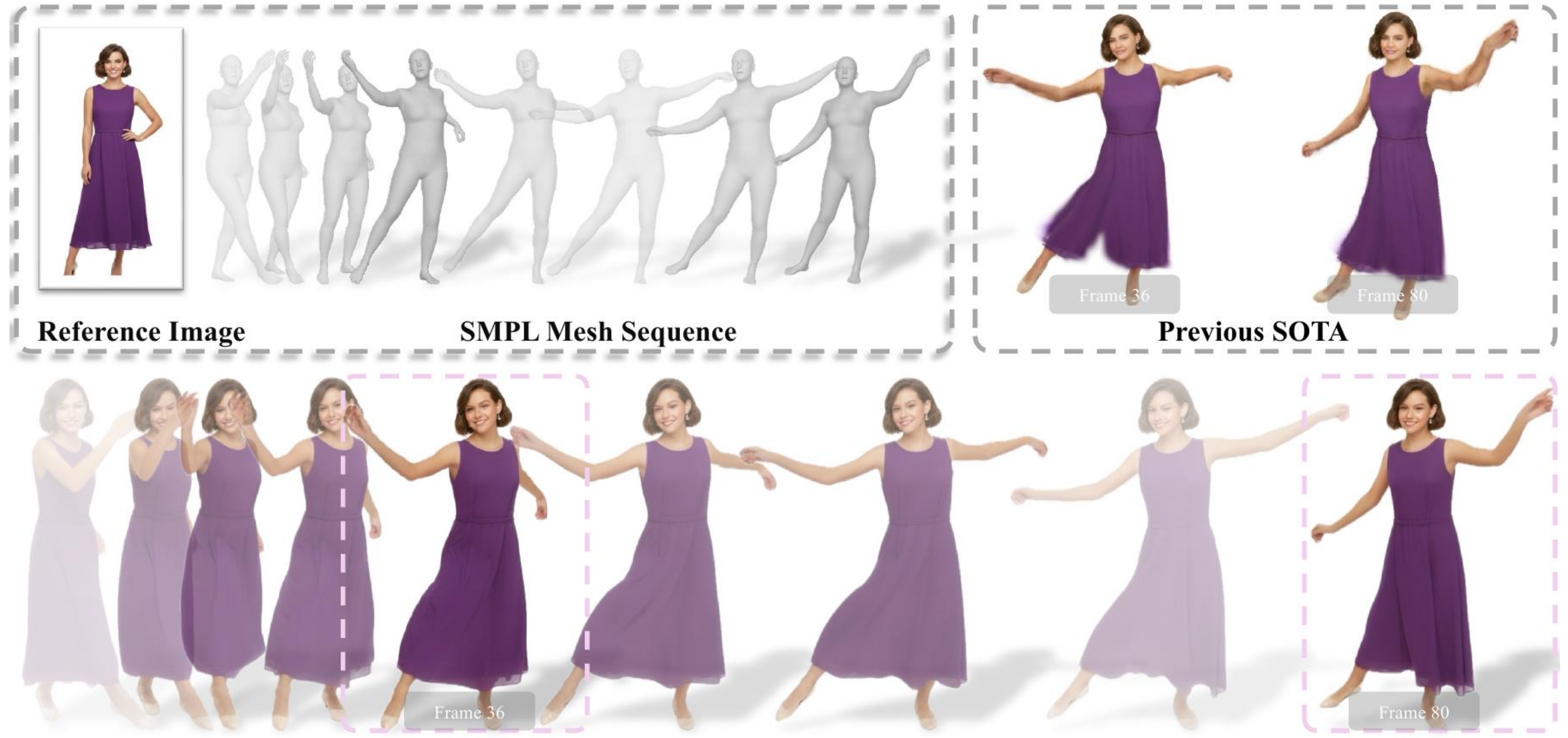
Qi Sun Can Wang Jiaxiang Shang Yingchun Liu Jing Liao

CVPR JUNE 3-7, 2026 DENVER COLORADO

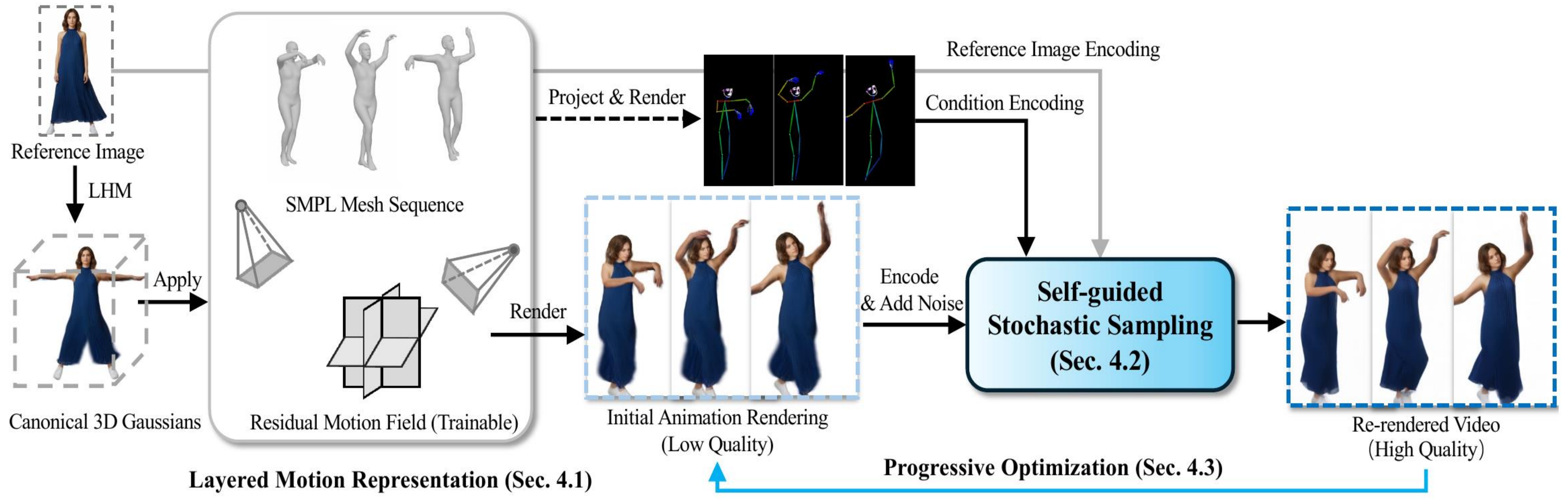
Code Available

Abstract

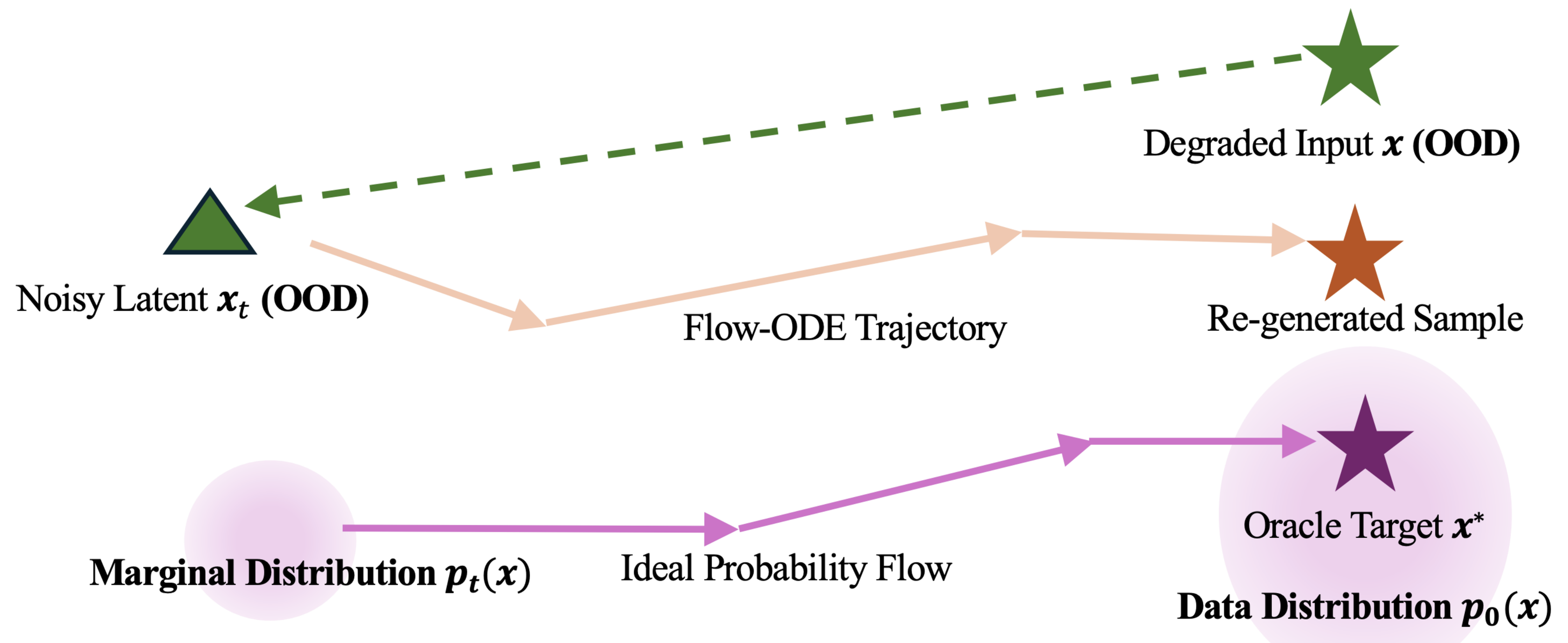
- **Input**: One image + SMPL motion
- **Output**: 4D human Gaussian sequence
- **Motion representation**: rigid body control + non-rigid dynamics.
- **Core Algorithm**: Self-guided stochastic sampling preserves identity and restores realism with *video diffusion prior*.



Framework



How it Works?



Distribution mismatch in deterministic flow matching

Algorithm 1 Self-guided Stochastic Sampling (Practical)

Require: Low-quality video $\boldsymbol{y}$; Pre-trained flow-based model $\boldsymbol{v}_\theta$; Preserved region $\mathcal{M}$; Initial noise step t_0, constant step size λ;

Ensure: Desirable high-quality video $\boldsymbol{x}^*$

Sample $\epsilon \sim \mathcal{N}(0, \mathbf{I})$

$\boldsymbol{x}_t = \sigma_{t_0}\epsilon + (1 - \sigma_{t_0})\boldsymbol{y}$ ▷ Eq. (6)

for $t : t_0 \to 0$ **do** ▷ Sampling loop

  $\hat{\boldsymbol{x}}_{0|t} \leftarrow \boldsymbol{x}_t - \sigma_t \boldsymbol{v}_\theta(\boldsymbol{x}_t, t)$ ▷ Eq. (3)

  $\hat{\boldsymbol{x}}_{1|t} \leftarrow \boldsymbol{x}_t + (1 - \sigma_t)\boldsymbol{v}_\theta(\boldsymbol{x}_t, t)$ ▷ Eq. (4)

  $\boxed{\hat{\boldsymbol{x}}_{0|t} \leftarrow \hat{\boldsymbol{x}}_{0|t} - \lambda \nabla_{\boldsymbol{x}_t} ||\mathcal{M} \odot (\boldsymbol{y} - \hat{\boldsymbol{x}}_{0|t})||^2}$ ▷ Eq. (10)

  Sample $\epsilon \sim \mathcal{N}(0, \mathbf{I})$

  $\boxed{\hat{\boldsymbol{x}}_{1|t} \leftarrow \sqrt{1 - \sigma_t}\hat{\boldsymbol{x}}_{1|t} + \sqrt{\sigma_t}\epsilon}$ ▷ Eq. (8)

  $\boldsymbol{x}_{t_{\text{next}}} \leftarrow (1 - \sigma_{t_{\text{next}}})\hat{\boldsymbol{x}}_{0|t} + \sigma_{t_{\text{next}}}\hat{\boldsymbol{x}}_{1|t}$ ▷ Eq. (5)

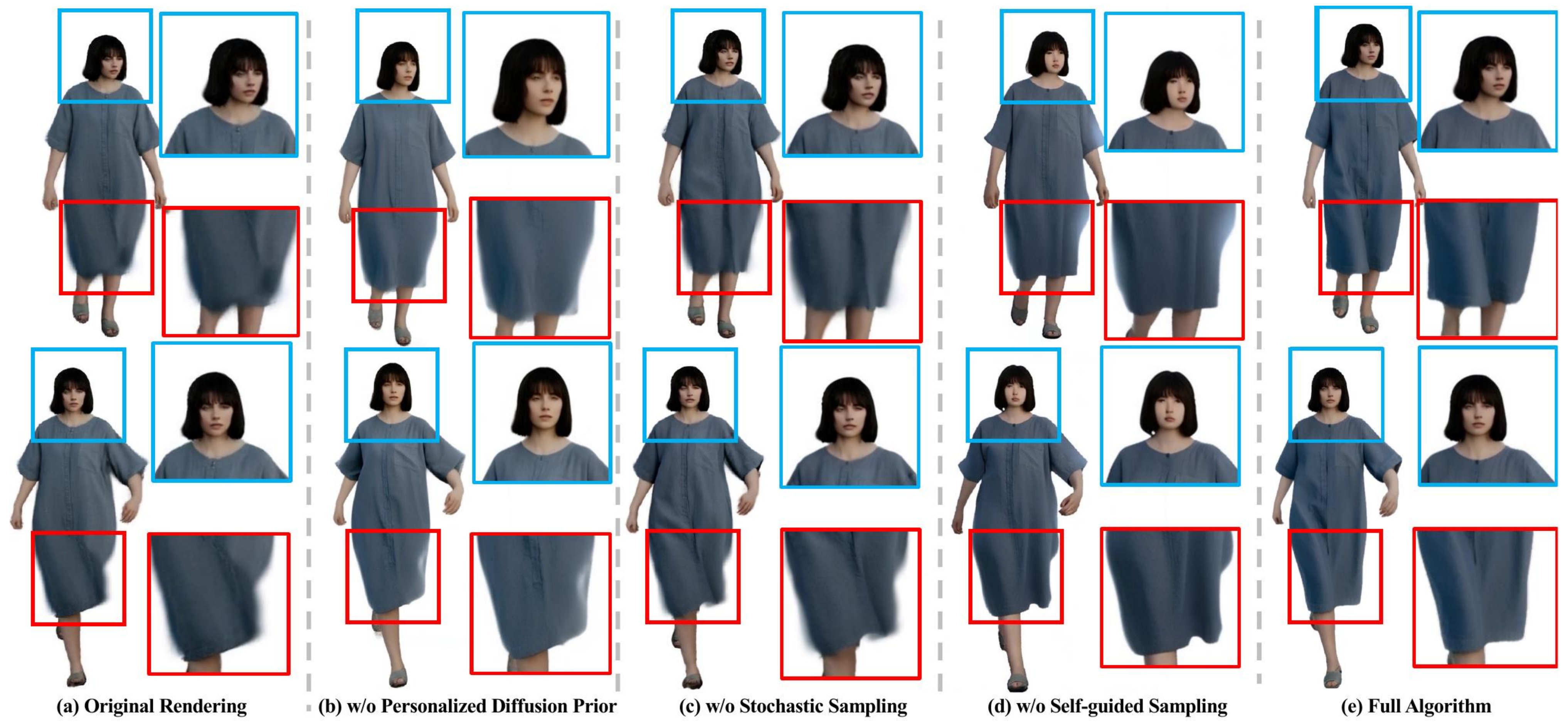
end for

Return $\boldsymbol{x}^* = \boldsymbol{x}_t|_{t=0}$

Eq(10) – DPS-style gradient guidance

Eq(8) – introduce stochasticity in sampling

Ablation – Self-guided Stochastic Sampling



Comparison with SOTA