Hybrid Analytical and Neural IK for Human Pose and Shape Estimation

Sihui Wang

April 11th, 2023

Overview

Introduction

Method

Algorithm

Evaluation

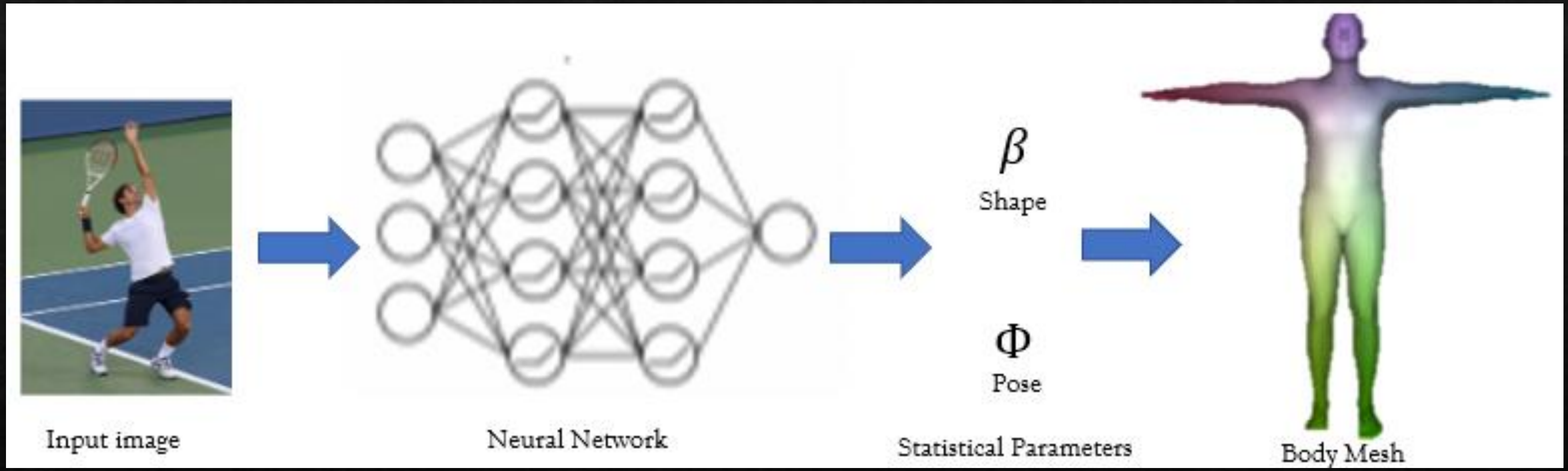
Conclusion & Further work

Introduction

- Task: Recover 3D human poses and shapes from one monocular RGB image
- Previous Approaches:
  - Model – based
  - 3D Keypoint – based

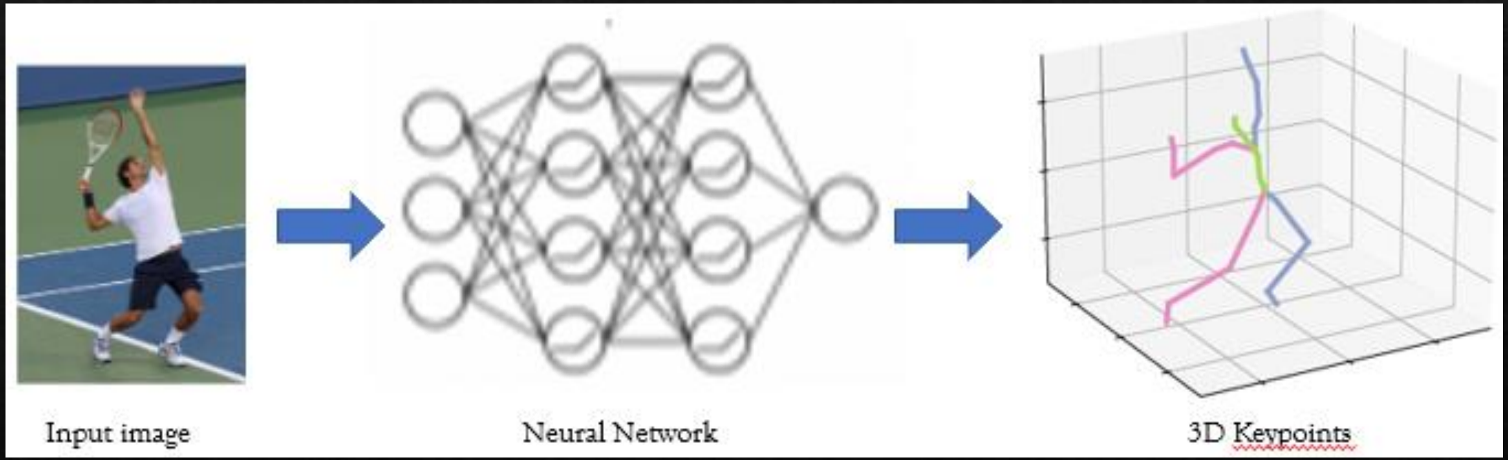
Introduction – Model Based Method

Drawback: Hard to learn the mapping from pixels to abstract statistical parameters.

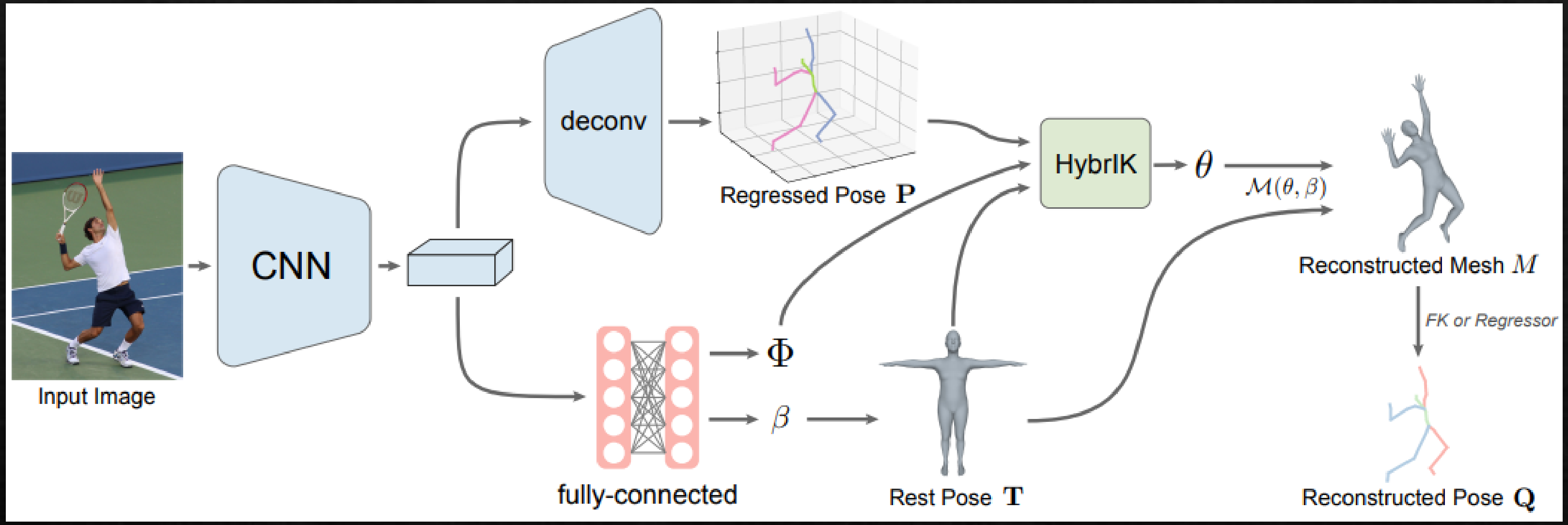


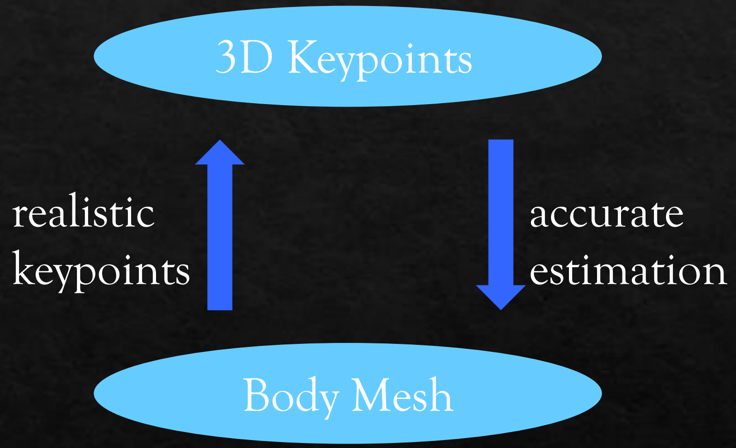
Introduction – 3D Keypoint Based Method

Drawback: The keypoint estimation can be unrealistic.



Introduction – Hybrid Method

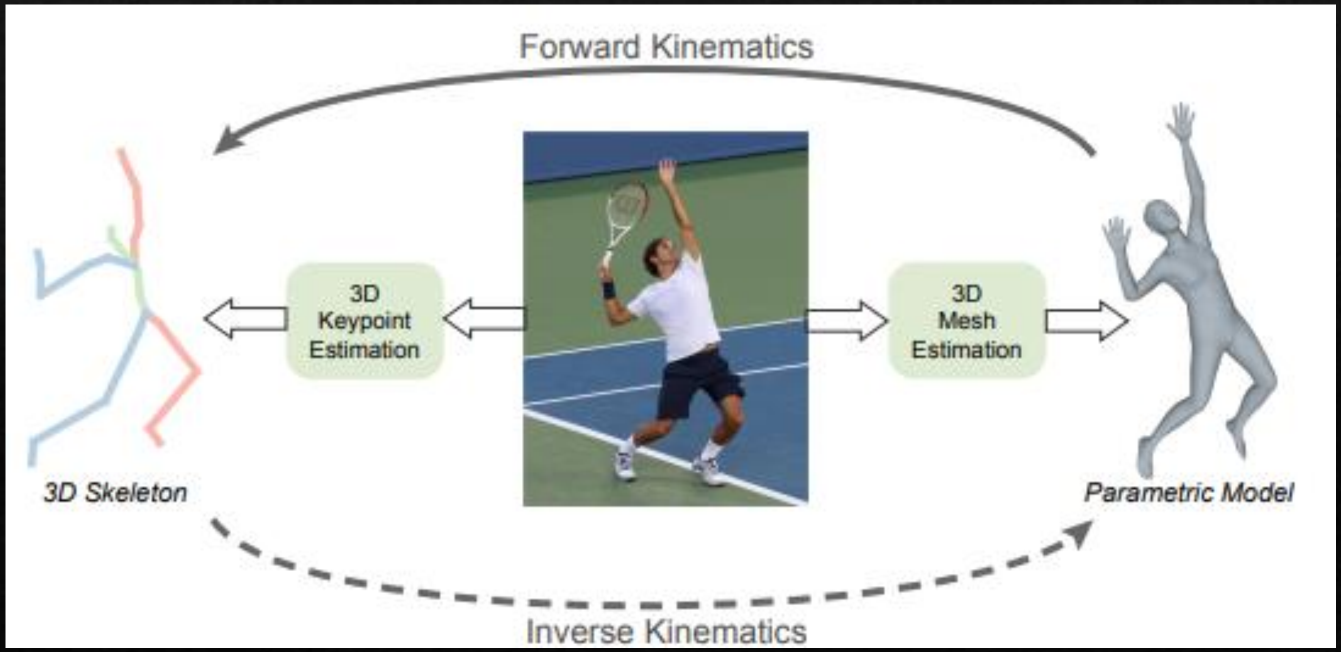




Method – Inverse Kinematics

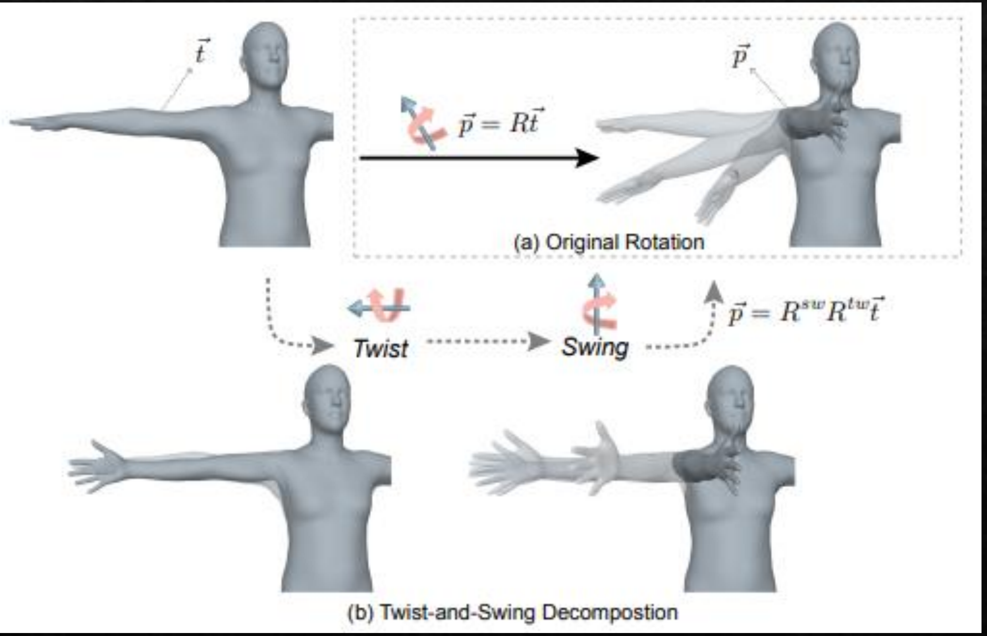
- Find pose parameters to align body mesh templates with keypoint estimation.

$$R = IK(P, T)$$



Method – Decomposition and Hybrid IK

- Twist: longitudinal rotation – estimated by neural networks
- Swing: In plane rotation – analytical solution



Algorithm – Naïve HybrIK

- $p_k - p_{pa(k)} = R_k(t_k - t_{pa(k)}) = R_{pa(k)} \cdot R_{pa(k),k}(t_k - t_{pa(k)})$
- $R_{pa(k)}^{-1}(p_k - p_{pa(k)}) = R_{pa(k),k}(t_k - t_{pa(k)})$
- $R_{pa(k),k} = R_{pa(k),k}^{sw} \cdot R_{pa(k),k}^{tw}$

Algorithm 1: Naive HybrIK

Input: $\mathbf{P}, \mathbf{T}, \Phi$

Output: $\mathbf{R}$

```
Determine R_0;
for k along the kinematic tree do
    p⃗_k ← R^{-1}_{pa(k)} (p_k − p_{pa(k)});
    t⃗_k ← (t_k − t_{pa(k)});
    R^{sw}_{pa(k),k} ← D^{sw}(p⃗_k, t⃗_k);
    R^{tw}_{pa(k),k} ← D^{tw}(t⃗_k, φ_k);
    R_{pa(k),k} ← R^{sw}_{pa(k),k} R^{tw}_{pa(k),k};
```

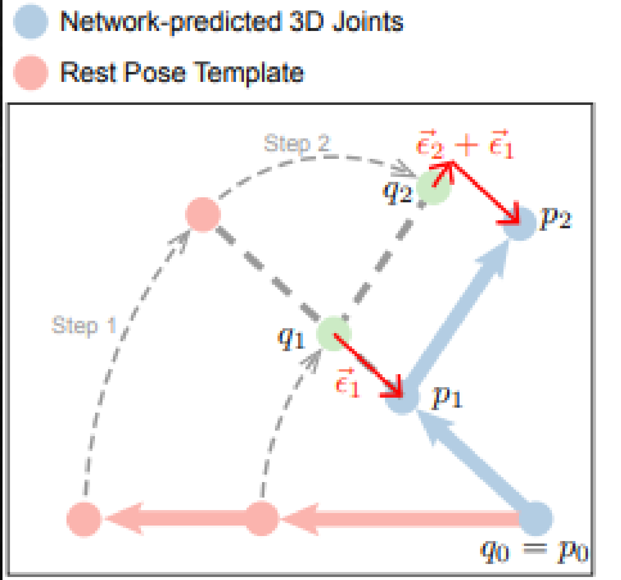
Determine R_0;

for k *along the kinematic tree* **do**

- $\vec{p}_k \leftarrow R_{\mathrm{pa}(k)}^{-1}(p_k - p_{\mathrm{pa}(k)})$;
- $\vec{t}_k \leftarrow (t_k - t_{\mathrm{pa}(k)})$;
- $R_{\mathrm{pa}(k),k}^{sw} \leftarrow \mathcal{D}^{sw}(\vec{p}_k, \vec{t}_k)$;
- $R_{\mathrm{pa}(k),k}^{tw} \leftarrow \mathcal{D}^{tw}(\vec{t}_k, \phi_k)$;
- $R_{\mathrm{pa}(k),k} \leftarrow R_{\mathrm{pa}(k),k}^{sw} R_{\mathrm{pa}(k),k}^{tw}$;

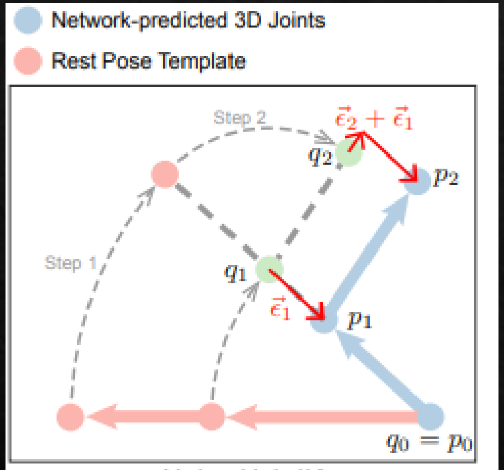
Algorithm – Naïve IK

- Problem:
  - Bone Length Inconsistency
  - Reconstruction Error is Accumulated

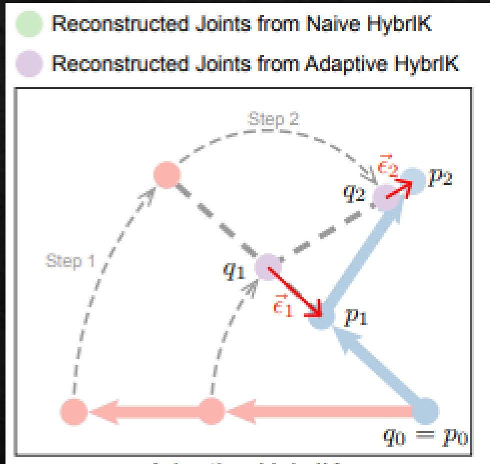


Naive HybrIK

Algorithm – Adaptive IK



Naive HybrIK



Adaptive HybrIK

Algorithm 2: Adaptive HybrIK

Input: P, T, Φ

Output: R

Determine R_0;

for *k along the kinematic tree* **do**

$q_{\text{pa}(k)} \leftarrow R_{\text{pa}(k)}(t_{\text{pa}(k)} - t_{\text{pa}^2(k)}) + q_{\text{pa}^2(k)}$;

$\vec{p}_k \leftarrow R^{-1}_{\text{pa}(k)}(p_k - q_{\text{pa}(k)})$;

$\vec{t}_k \leftarrow (t_k - t_{\text{pa}(k)})$;

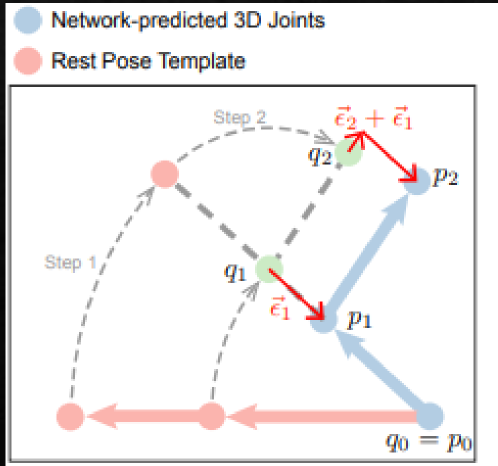
$R^{sw}_{\text{pa}(k),k} \leftarrow \mathcal{D}^{sw}(\vec{p}_k, \vec{t}_k)$;

$R^{tw}_{\text{pa}(k),k} \leftarrow \mathcal{D}^{tw}(\vec{t}_k, \phi_k)$;

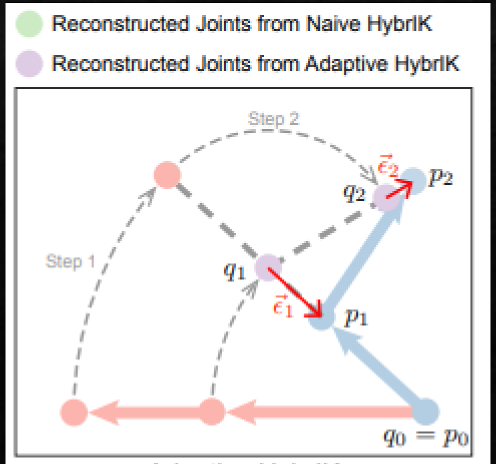
$R_{\text{pa}(k),k} \leftarrow R^{sw}_{\text{pa}(k),k} R^{tw}_{\text{pa}(k),k}$;

Algorithm – Our Algorithm

- The adaptive algorithm is heuristic – not guarantee to minimize reconstruction errors.
- We propose λ – adaptive:
  - Use some point in the middle: $\lambda p_1 + (1-\lambda) q_1$

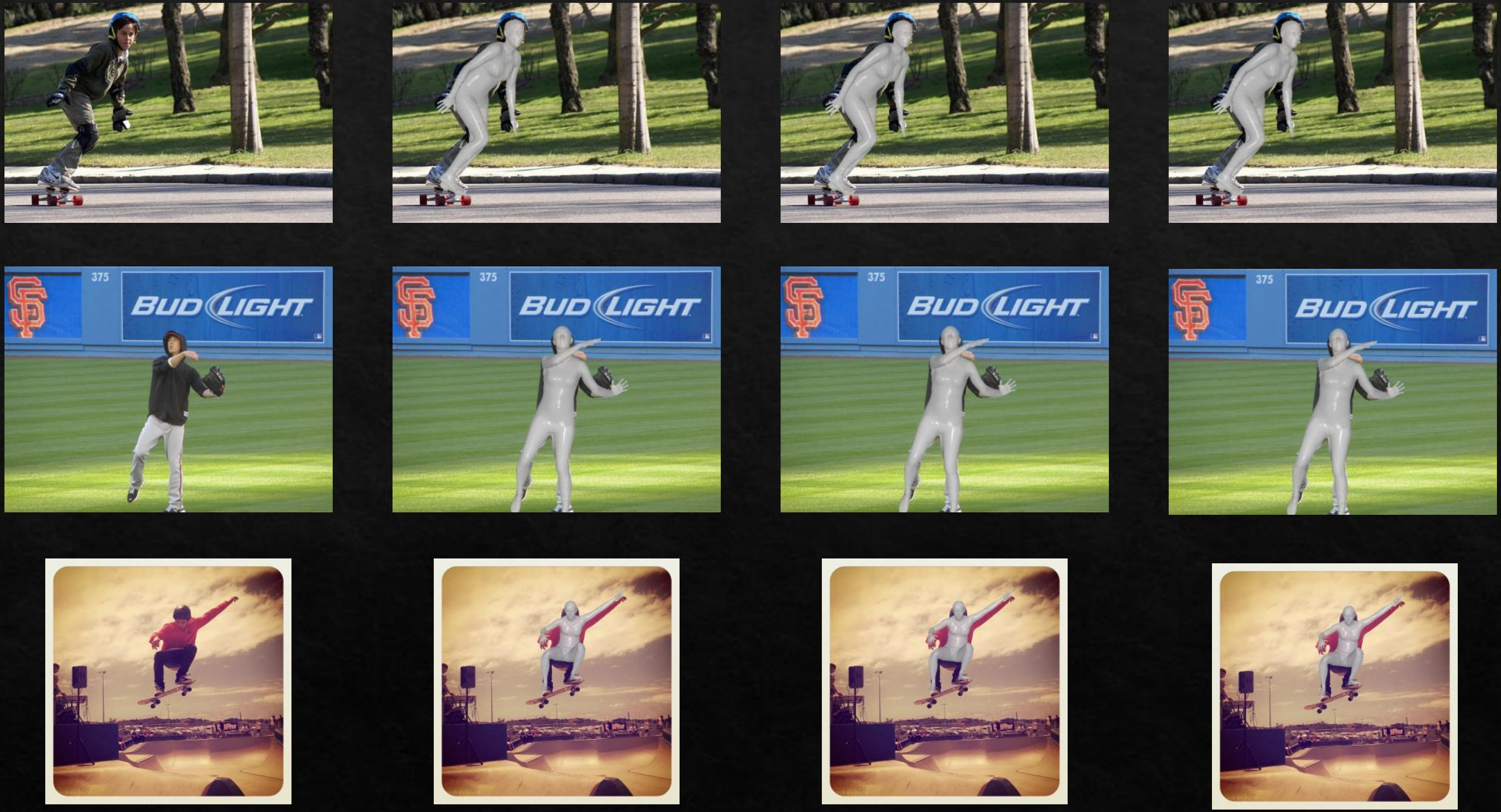


Naive HybrIK



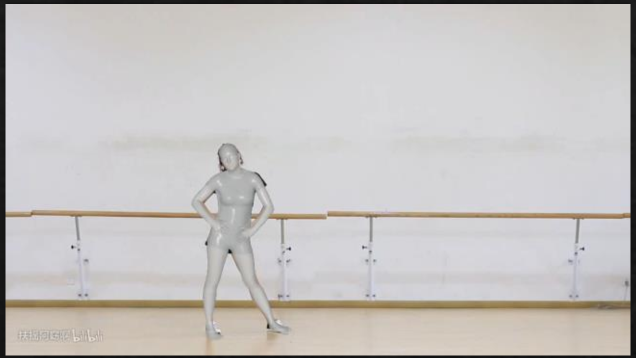
Adaptive HybrIK

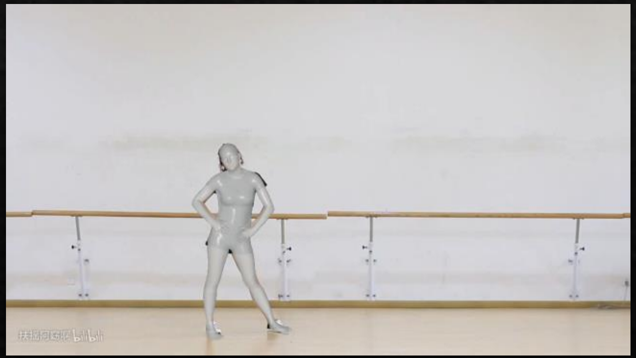
Evaluation – Qualitative Results

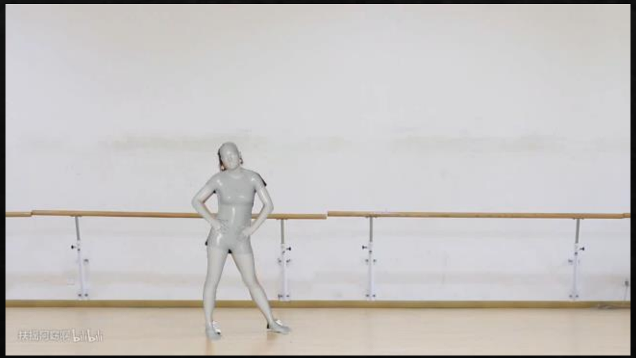


$\lambda = 0.0$ $\lambda = 0.5$ $\lambda = 1.0$

Evaluation – Qualitative Results







Conclusion & Next Steps

- Conclusion So Far:
  - Hypothesis: $\lambda = 0$: Accurate directions; $\lambda = 1$: Accurate position
  - Finding: Visual differences are negligible when tuning λ
- Next Steps:
  - Larger Scale Quantitative Evaluations are Needed.
  - Comparison of Reconstruction Error Under different λ
  - Robustness Under different λ

Q & A