HuManiFlow: Ancestor-Conditioned Normalising Flows on SO(3) Manifolds for Human Pose and Shape Distribution Estimation

Akash Sengupta
Ignas Budvytis
Roberto Cipolla

Multiple 3D human reconstructions can correspond to a 2D image due to depth ambiguity, occlusion and truncation. Motivates a probability distribution over 3D pose and shape, which should exhibit 3 properties…

Motivation

We use the SMPL\(^{3}\) 3D body model.

\[
\text{Shape}\ 
\{R_i\}_{i=1}^{23}
\]

Body Mesh

3D rotation of each body-part about its parent joint. Body-part rotations belong to the Lie group SO(3).

We predict a distribution over SMPL pose and shape conditioned on a 2D input \(X\).

\[
P_{\text{shape}}(\{R_i\}_{i=1}^{23}, \beta | X) = P_{\text{shape}}(\beta | X)P_{\text{pose}}(\{R_i\}_{i=1}^{23} | \beta, X)
\]

Full-body pose is factorised into per-body-part rotation distributions conditioned on ancestor body part rotations.

\[
P_{\text{pose}}(\{R_i\}_{i=1}^{23} | \beta, X) = \prod_{i=1}^{23} P_{\text{SO}(3)}(R_i | \beta, \phi)\text{ }\text{Aggregated into context }\beta_i
\]

Ancestors of part \(i\)

Method

3D input consistency and 3D sample diversity of probabilistic pose and shape methods on 3DPW.

Input consistent in 2D because we:

(i) Use kinematic tree to factorise body pose.
(ii) Consider domain of body-part rotations SO(3).

Not diverse in 3D because:

(i) Use expressive distribution models (flows).
(ii) Don’t use \(\ell_{1}\) posed loss functions (e.g. 3D MSE).

3D Multibodies\(^{5}\)

Not Input-Consistent in 2D

Diverse in 3D

HuManiFlow

Input Consistent in 2D

Diverse in 3D

2D Input Consistency

3D Sample Diversity

Distribution Accuracy

Trade-off in current probabilistic methods\(^{1,2}\)

Input Image

Note 2-axial (slightly) uncertainty for arm vertices.

ProHMRF\(^{1}\)

Input-Consistent in 2D

Not Diverse in 3D

Samples (front + side)

Pre-Vertex Variance

Input Proxy Representation

Results


References