Hierarchical Kinematic Probability Distributions for 3D Human Shape and Pose Estimation from Images in the Wild

Akash Sengupta  Ignas Budvytis  Roberto Cipolla

Problem

Monocular 3D human shape and pose estimation is ill-posed.
• Multiple 3D bodies may explain a given 2D image.

Method

(i) Proxy representation computation
• Input image is converted into an edge + 2D joint heatmap representation.
• Bridges the gap between synthetic training data (with diverse shapes and poses) and real test data.

(ii) 3D Body Sampling and Projection
• SMPL meshes are sampled from the predicted distribution using rejection sampling.
• Rejection sampling is made differentiable using the reparameterisation trick which enables sample re-projection loss.

(iii) 3D Shape and Pose Distribution Prediction

Deep neural network outputs:
(1) Gaussian distribution over shape parameters
(2) Matrix-Fisher distribution over 3D joint rotations
Each joint's distribution parameter is conditioned on its parent's modes, principal axes and dispersions

Results

Comparison with recent 3D human shape and pose distribution estimation methods
• 3D shape (PVE-TSC) and pose (MPJPE/MPJPE-PA) metrics are computed using the minimum error sample for each test image in the 3DPW and SSP-3D datasets.

References
1. N. Koliotouros et al. Learning to reconstruct 3D human pose and shape via model-fitting in the loop. ICCV 2019.
2. M. Loper et al. SMPL: A Skinned Multi-Person Linear model. ACM SIGGRAPH Asia 2015.
3. T. D. Downs. Orientation statistics. Biometrika 59(3):665–676, December 1972.
4. D. Mohlin et al. Probabilistic orientation estimation with matrix fisher distributions. NeurIPS 2020.
5. B. Biggs et al. 3D Multibodies: Fitting sets of plausible 3D models to ambiguous image data. NeurIPS 2020.
6. A. Sengupta et al. Probabilistic 3D human shape and pose estimation from multiple unrelated images in the wild. CVPR 2021.