How Robust is 3D Human Pose Estimation to Occlusion?

István Sárándi¹, Timm Linder², Kai O. Arras², Bastian Leibe¹
¹Visual Computing Institute, RWTH Aachen University – Aachen, Germany
²Robert Bosch GmbH, Corporate Research – Stuttgart, Germany

Overview

3D Human Pose Estimation Task

- Localize body joints in 3D camera space from an RGB image
- Useful for collaborative robotics
- Much progress in the last few years, as measured on current benchmarks, such as Human3.6M

Real Environments are More Challenging Than Benchmarks

- Occlusion is common in shared human-robot environments
- Current benchmarks don’t systematically model this
- How well do current methods work under occlusion?

Measure Robustness to Synthetic Occlusions

- Add synthetic occlusions to the training data as well
- In addition to usual augmentations
- Turns out to be a good regularizer even in non-occluded cases

Apply Data Augmentation for Improved Robustness and Regularization

Investigated Pose Estimation Model

- Based on two of the best recent methods [2,3]
- Performance on Human3.6M is at state-of-the-art level
- High frame rate inference (204 fps) on Titan X GPU
- Fully-convolutional backbone (ResNet-50) directly predicts a 16x16x16 volumetric heatmap per body joint
- Heatmaps are converted to coordinates with soft-argmax and are back-projected into 3D metric space, where the L1 loss is minimized

Evaluation

- Does the occluder shape matter? Which type of training augmentation improves robustness to which type of test-time occlusion?
- Evaluation measure: mean per joint position error after skeleton alignment at root joint (MPJPE)

Qualitative Example

Baseline model

- View from left (translucent: gr. truth)
- View from left (translucent: gr. truth)

Trained with VOC object aug.

- View from left (translucent: gr. truth)

Key Findings

- Baseline 3D pose estimator is sensitive even to low degrees of occlusion
- Circular occluders are the most difficult
- Training with circles improves robustness to all simple shapes
- Robustness to Pascal VOC occluders not improved by augmenting with simple shapes
- Occlusion augmentation helps even for unoccluded test cases

Won the PoseTrack 3D Challenge at ECCV 2018, ahead of methods using external 2D datasets in training (details in [4])

Acknowledgment

This project has been funded by a grant from the Bosch Research Foundation, by IGA (91628-16CT-2015–727273) and by ERC Consolidator Grant DeeViSe (ERC 2017-CoG-773211).