Estimating Correspondences of Deformable Objects “In-the-wild”

Yuxiang Zhou*, Epameinondas Antonakos*, Joan Alabort-i-Medina*, Anastasios Roussos*, Stefanos Zafeiriou*†

*Department of Computing, Imperial College London, U.K. †Center for Machine Vision and Signal Analysis, University of Oulu, Finland

IEEE 2016 Conference on Computer Vision and Pattern Recognition

CVPR 2016

Introduction

Motivation

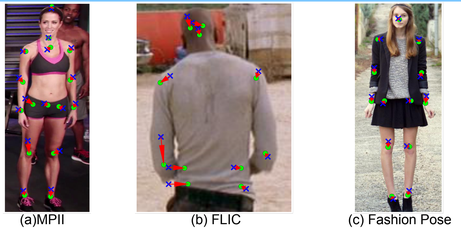
- Very difficult to consistently annotate most deformable objects
- Annotations are labour intensive

Solution

- Estimate dense correspondences from inconsistent annotations
- Construct accurate Statistical Deformable Models

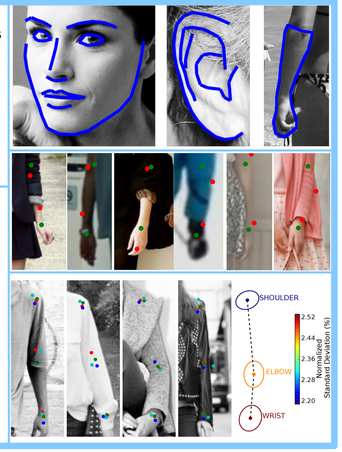
Contribution

- Corrected annotations for FLIC and MPII publicly available

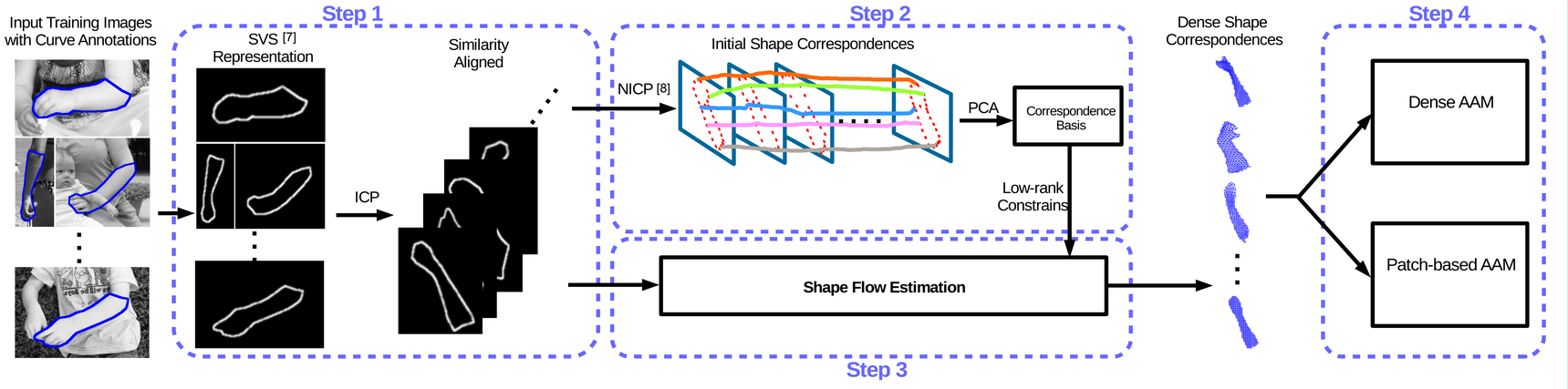


(a)MPII (b) FLIC (c) Fashion Pose

Inconsistent human annotation among different dataset

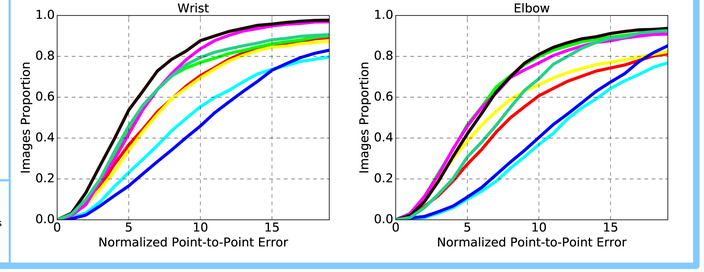


Architecture

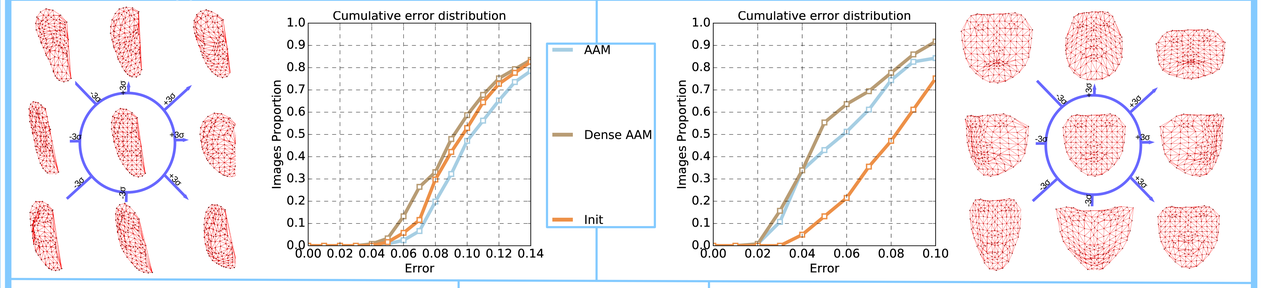


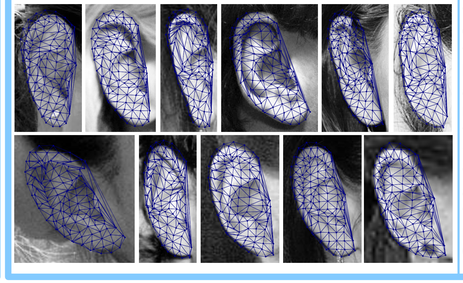
Patch-based Active Appearance Models

- Outline VS Skeleton
- State-of-the-art performance
- Applicable to objects with challenging interior texture
- Experiments trained on 891 line annotated “in-the-wild” data and tested on BBCPose testset

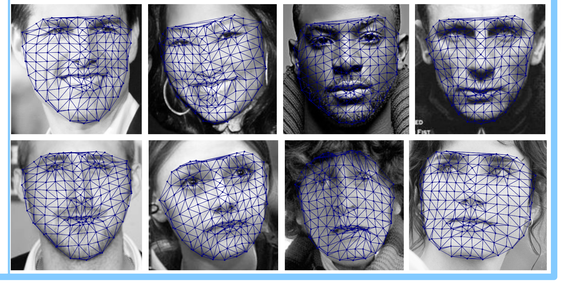


Dense Active Appearance Models





- Explicit use of pixelwise one-to-one correspondence
- Motion model of sparse holistic AAMs (piece-wise affine, thin-plates splines) is replaced by sampling all pixel values onto the reference frame
- Face experiments trained and tested on LFPW dataset
- Ear experiments on our private in-the-wild database of 605 images



Reference

[1] P. Buehler et al. Upper body detection and tracking in extended signing sequences. IJCV2011.
[2] J. Charles et al. Upper body pose estimation with temporal sequential forests. BMVC2014.
[3] J. Charles et al. Domain adaptation for upper body pose tracking in signed TV broadcasts. BMVC2013.
[4] T. Pfister et al. Deep convolutional neural networks for efficient pose estimation in gesture videos. ACCV2015.
[5] Y. Yang et al. Articulated human detection with flexible mixtures of parts. PAMI2013.
[6] T. Pfister et al. Flowing convnets for human pose estimation in videos.ICCV2015
[7] B. Amberg et al. Optimal Step Nonrigid ICP Algorithms for Surface Registration. CVPR2007
[8] V Nguyen et al. Support vector shape: A classifier-based shape representation. PAMI2013

Download the corrected annotations:

http://www.ibug.doc.ic.ac.uk/resources/bodypose-anno-correction