Learning Dense Facial Correspondences in Unconstrained Images

Ronald Yu *[1,3], Shunsuke Saito †[1,3], Haoxiang Li ‡[2], Duygu Ceylan §[2], and Hao Li¶[1,3,4]

[1]University of Southern California
[2]Adobe Research
[3]Pinscreen
[4]USC Institute for Creative Technologies

1. Additional Evaluations

1.1. Additional Qualitative Evaluations On In-The-Wild Images

We show qualitative comparisons with other state-of-the-art methods on in-the-wild image sequences. We detect the faces with the state-of-the-art face detector from Hu et a. [1]. Although it is the current best performer on the WIDER Face Detection Benchmark [6], we observe significant variations in its bounding box quality when applying under these challenging settings. As shown in Figure 1, performance of the competing methods can drastically degrade due to both the large appearance variations and imperfect bounding boxes while our method still demonstrates decent robustness and accuracy.

1.2. Additional Quantitative Evaluations

1.2.1 3D Model Fitting

We evaluated our 3D model fitting on the AFLW-2000 dataset and compared to Zhu et al. [7]. We did not compare to Jourabloo et al. [2] as they do not provide 3D fittings. Our average error 4.25 percent and the average error of Zhu et al.[7] is 5.18 percent.

1.2.2 Dense 2D Correspondence

We evaluated our dense flow on the AFLW-2000 dataset [7]. On a 128x128 image, each flow had an average error of 2.54 pixels.

1.2.3 Identity Extraction

We also performed evaluations on the consistency of the extracted identity shape of the reconstructed 3DMM across multiple frames or views of the same individual. We tested our method on 3 views of the same person using data from The Wikihuman Project [4] and extracted the neutral identity meshes. The average L2 vertex distance between any two of the neutral identity meshes is 5.45 mm.

References

[1] P. Hu and D. Ramanan. Finding tiny faces. *CoRR*, abs/1612.04402, 2016. 1

[2] A. Jourabloo and X. Liu. Large-pose face alignment via cnn-based dense 3d model fitting. In *IEEE CVPR*, pages 4188–4196, 2016. 1, 2

[3] V. Kazemi and J. Sullivan. One millisecond face alignment with an ensemble of regression trees. In *IEEE CVPR*, pages 1867–1874, 2014. 2

[4] D. H. League. The wikihuman project. http://gl.ict.usc.edu/Research/DigitalEmily2/, 2015. Accessed: 2015-12-01. 1

[5] J. Saragih and R. Goecke. A nonlinear discriminative approach to aam fitting. In *IEEE ICCV*, pages 1–8, 2007. 2

[6] S. Yang, P. Luo, C.-C. Loy, and X. Tang. Wider face: A face detection benchmark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5525–5533, 2016. 1

[7] X. Zhu, Z. Lei, X. Liu, H. Shi, and S. Z. Li. Face alignment across large poses: A 3d solution. In *IEEE CVPR*, June 2016. 1, 2

*ronaldyu@usc.edu
†shunsuke.saito16@gmail.com
‡haoxli@adobe.com
§ceylan@adobe.com
¶hao@hao-li.com

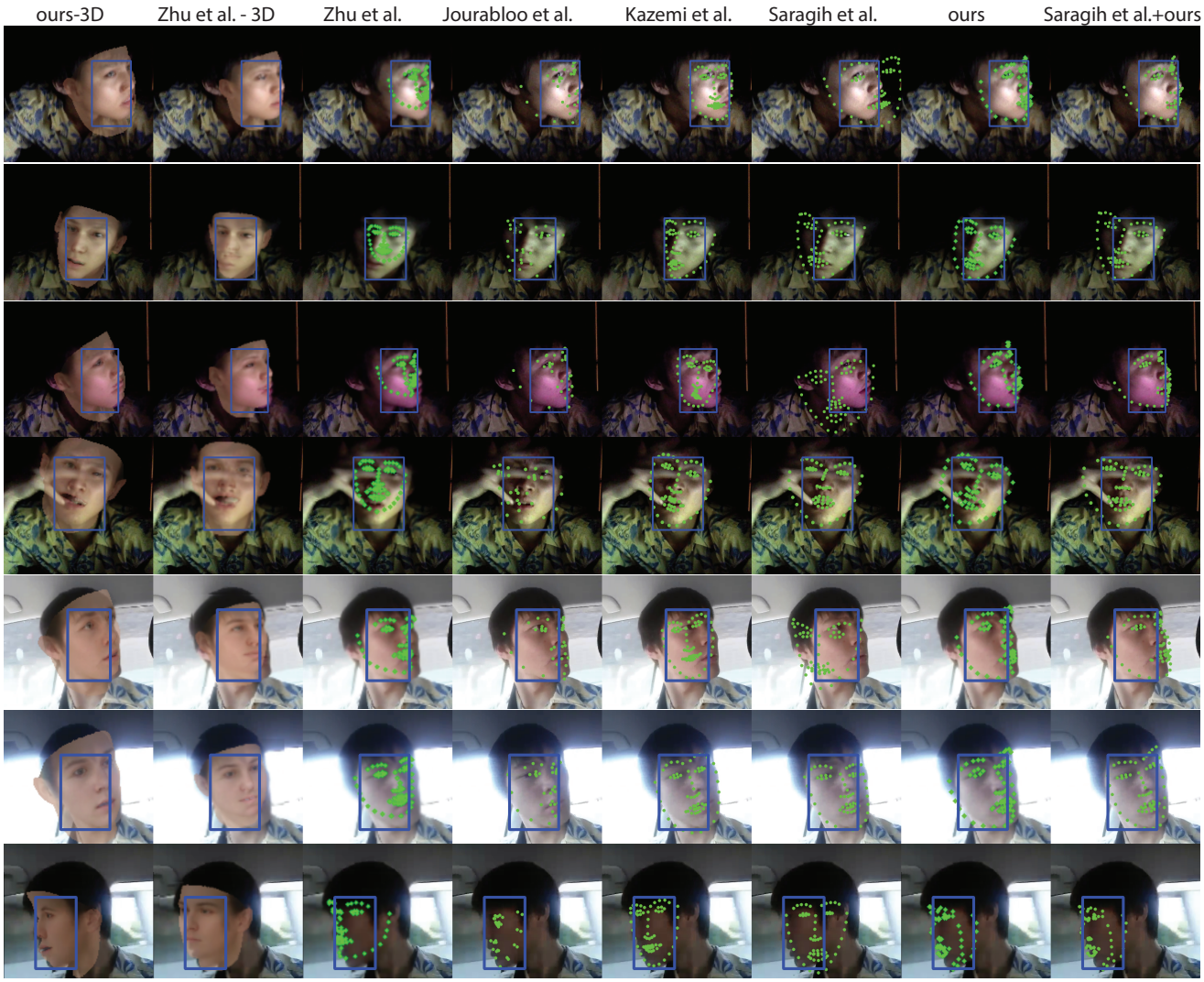

Figure 1. We provide visual 2D and 3D facial alignment results on challenging in-the-wild images using our method, the method of Zhu et al. [7], Jourabloo et al. [2], Kazemi et al. [3], and Saragih et al. [5]. We also show the results obtained by [5] when initialized with our predictions. We note that these results do not have perfect bounding boxes, but our method is more robust than others