Face Alignment in Full Pose Range: A 3D Total Solution Supplemental Material

Xiangyu Zhu, Xiaoming Liu, *Member, IEEE,* Zhen Lei, *Senior Member, IEEE,* and Stan Z. Li, *Fellow, IEEE*

1 Results Demo

In this section, we demonstrate some alignment results of 3DDFA in Fig. 1. Since the landmark visibility can be easily computed from the fitted dense 3D model [1], we also demonstrate the landmark visibility. In Fig. 1, the first and second rows demonstrate the samples with large yaw angles, the third row shows the samples with both large yaw and pitch angles, the last row shows the faces with both extreme expressions and poses.

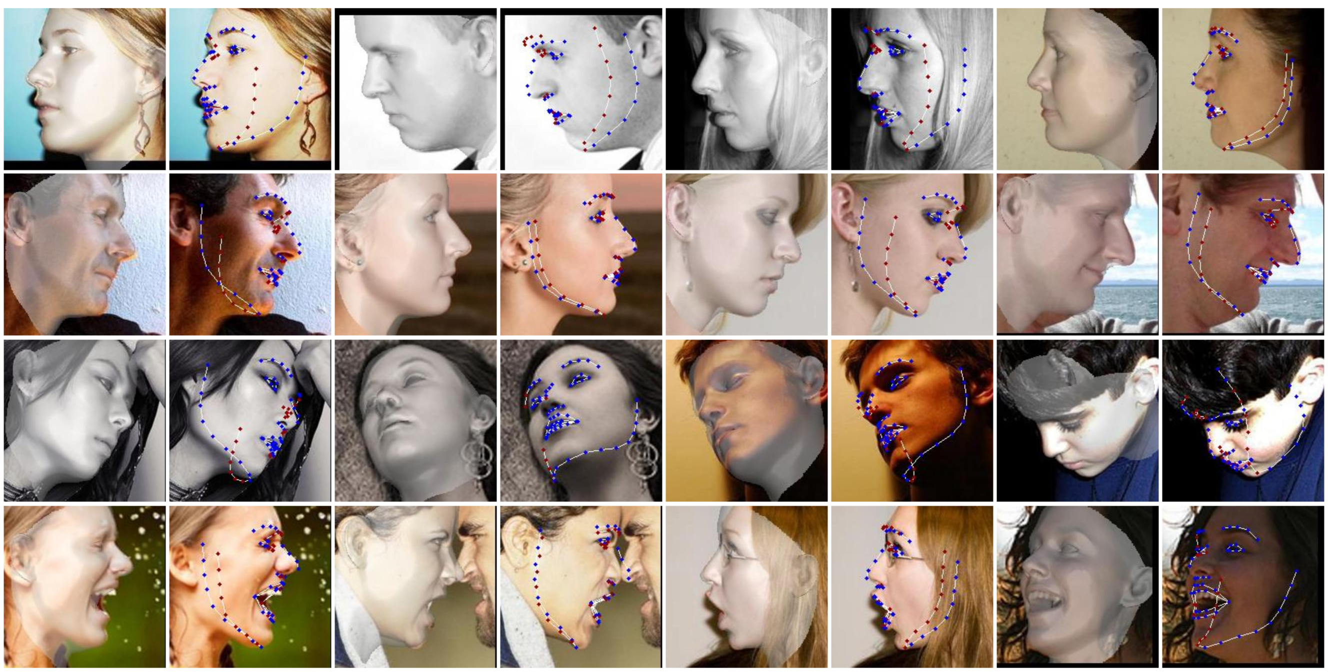


Fig. 1. The results of 3DDFA in AFLW. For each pair, the left one overlaps the fitted 3D face on the image. The right one shows the landmarks overlayed on the 3D face model. The blue/red ones indicate visible/invisible landmarks.

2 Database

2.1 300W-3D

This database contains the images in 300W [3] and their ground truth 3D faces. It is constructed by fitting the 3D Morphable Model with the Multi-Features Framework [4]. Different from the original algorithm, the 3D landmarks are adjusted by landmark marching [5] and the 68-landmarks constraint is adopted throughout the fitting process. Each sample is checked, few bad samples are adjusted manually. Fig. 2 demonstrates some samples in the database.

2.2 300W-LP

The 300W across Large Poses (300W-LP) database contains the synthesized face images from the face profiling algorithm described in Section 4. Some samples are demonstrated in Fig. 3. Besides, Fig. 4(a) and Fig. 4(b) demonstrate the yaw angle distribution before and after face profiling respectively. The distribution is reversed since face profiling only enlarges the yaw angle without shrinking it. Considering that the fidelity of a synthesized sample is negatively related with the Δyaw, we augment each training sample with the times of $\lceil (5 * 0.8^{\Delta yaw/5^\circ}) \rceil$. Fig. 4(c) shows the yaw distribution after augmentation.

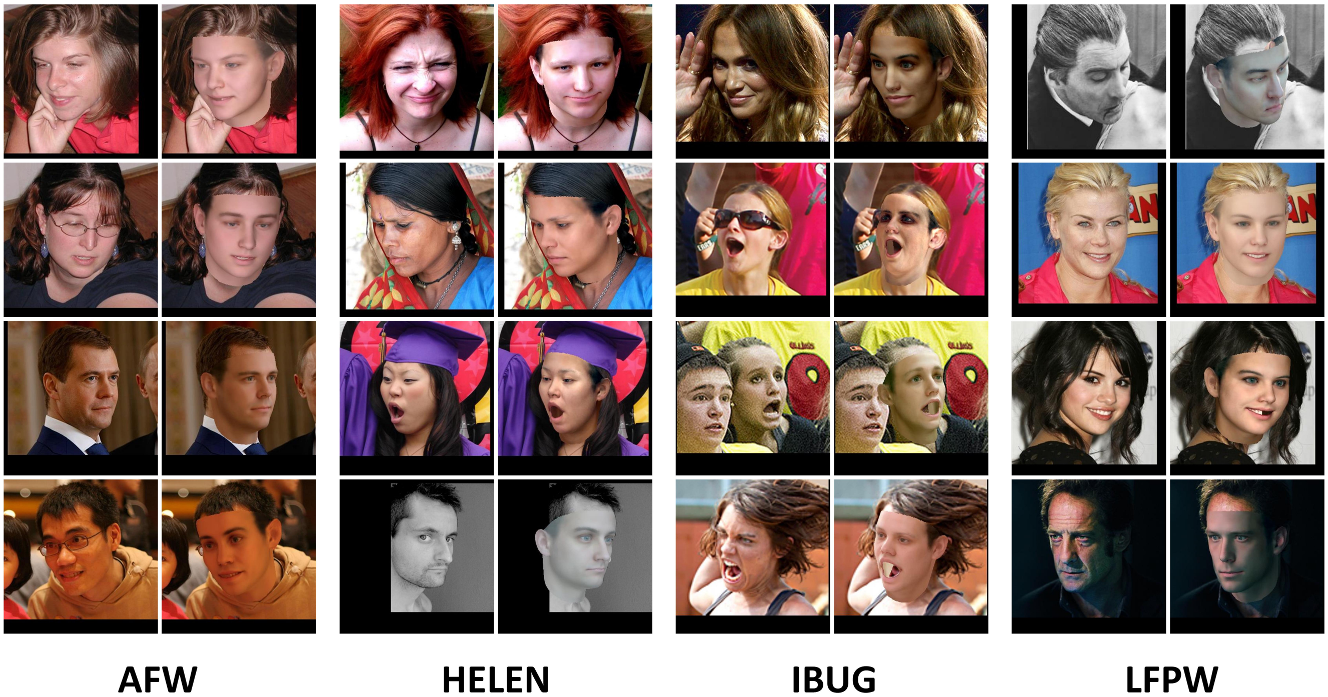


Fig. 2. The **300W-3D** database. For each sample, the left one is the original image, the right one is the fitted 3DMM.

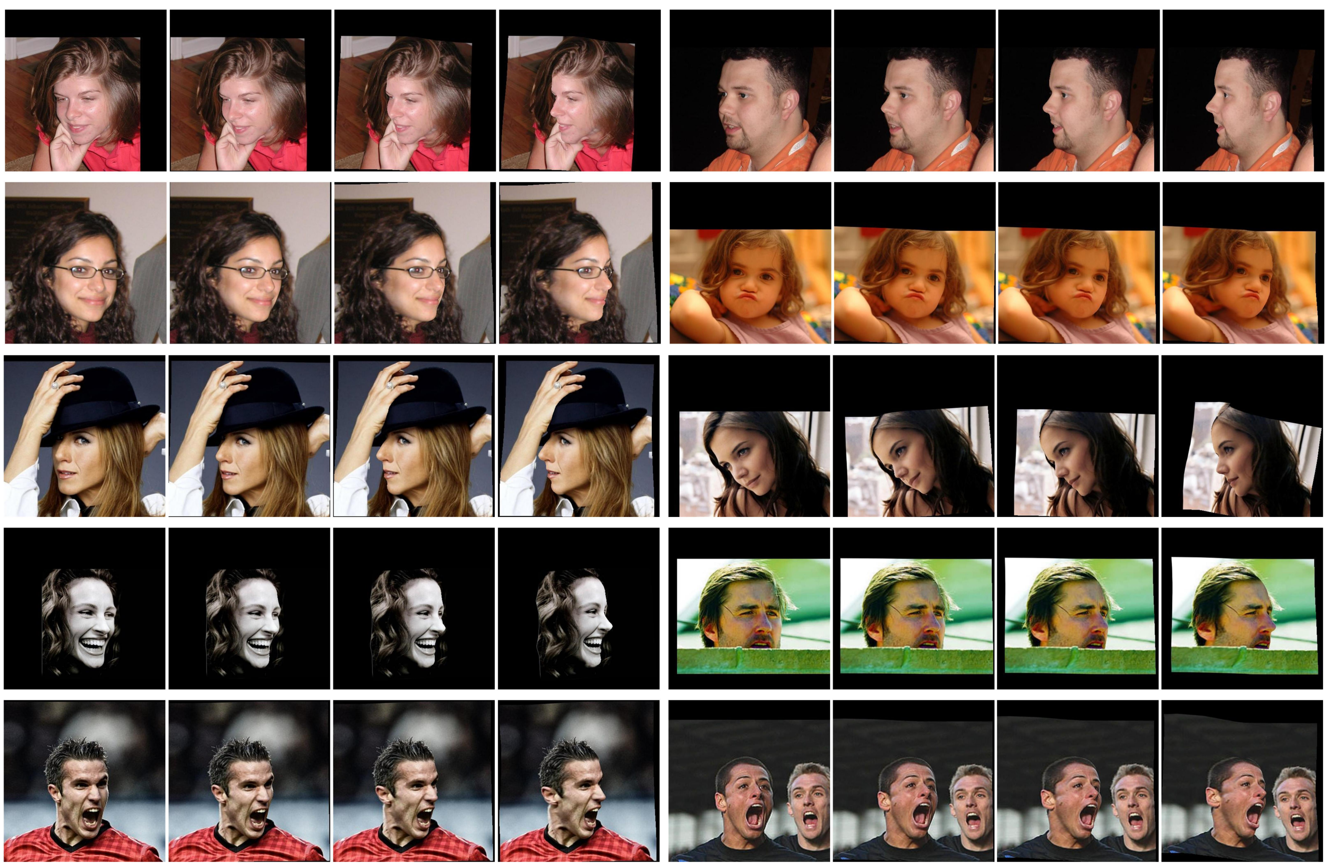
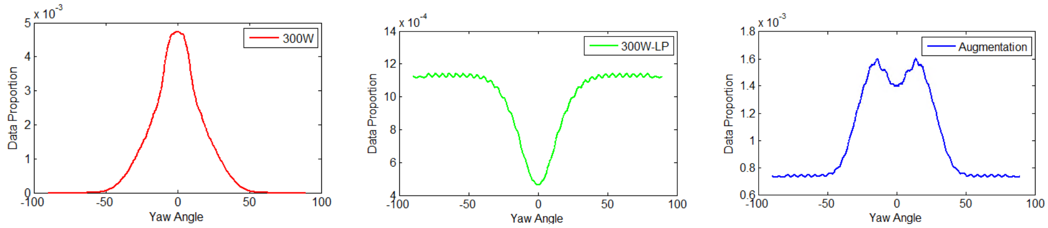


Fig. 3. The **300W-LP** database. For each sample, the first is the original image, followed by synthesized larger-pose faces, each with increased 10 degree yaw.



(a) (b) (c)

Fig. 4. The yaw angle distribution of (a) 300W, (b) 300W-LP, (c) After augmentation.

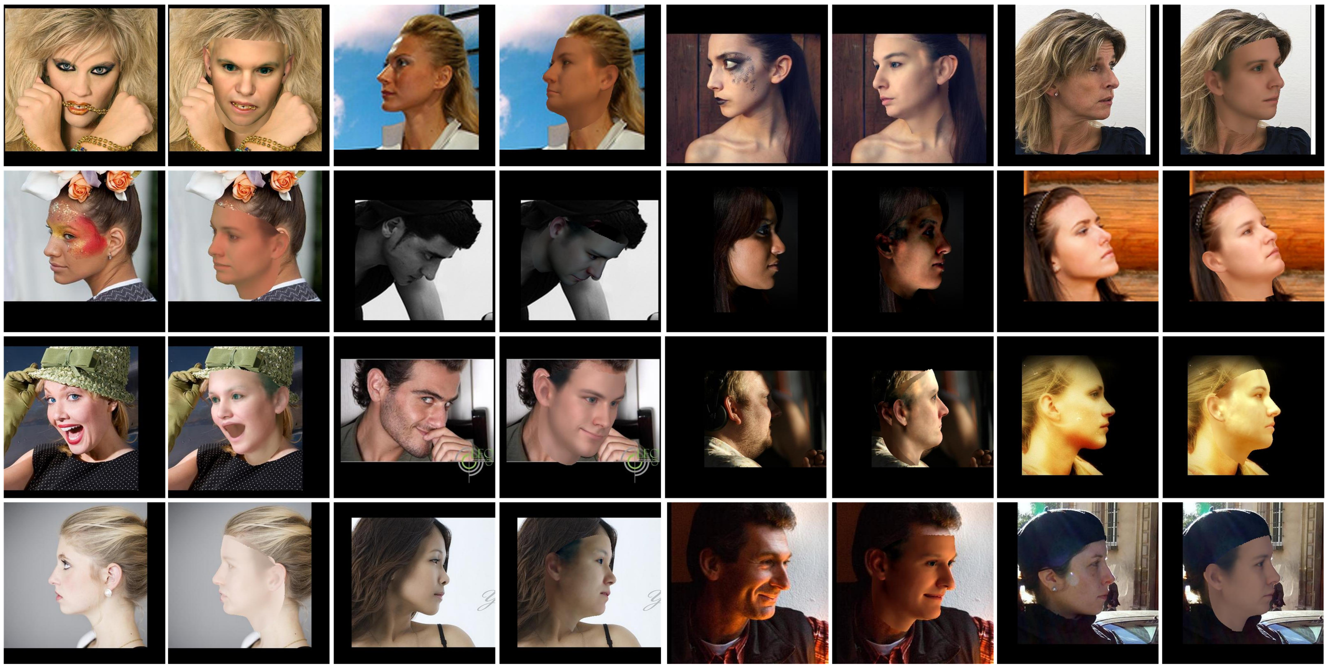


Fig. 5. The **AFLW2000-3D** database. For each sample, the left one is the original image, the right one is the fitted 3DMM.

2.3 AFLW2000-3D

The AFLW2000-3D database contains the first 2000 samples in AFLW [6] with their ground truth 3D faces. This database is more challenging to construct than 300W-3D because the AFLW ignores the occluded landmarks (both occluded and self-occluded) and the landmarks do not contain expression information (without lip landmarks). As a result we need a specifical method to deal with the unstandardised landmarks. We adopt 6 steps to construct AFLW2000-3D:

Step 1. We train pose-independent SDM models on 300W-LP to detect 68 landmarks. Specifically, we train 7 alignment models with the samples in 7 yaw intervals including $[-90°, -75°]$, $[-75°, -45°]$, $[-45°, -15°]$, $[-15°, 15°]$, $[15°, 45°]$, $[45°, 75°]$ and $[75°, 90°]$.

Step 2. According to the labelled yaw angles, we coarsely align each AFLW sample with the corresponding SDM model, getting 68 landmarks.

Step 3. We fit 3DMM with Multi-Features Framework (MFF) [4], where the 68 landmarks are used for initialization and the labelled 21 visible landmarks are used as a fitting constraint.

Step 4. We manually check the results and filter the failure results. There are 389 failure results.

Step 5. For the 389 samples, we manually label additional landmarks using the 68-landmark makeup according to the following criterion: If eyebrow fitting fails, we label the $18, 20, 22$ landmarks for the left eyebrow and the $23, 25, 27$ landmarks for the right eyebrow. If eye fitting fails, we label the $37, 38, 40, 42$ landmarks for the left eye and the $43, 44, 46, 47$ landmarks for the right eye. If nose fitting fails, we label the $28, 31, 32, 34, 36$ landmarks. If mouth fitting fails, we label the $61, 63, 65, 67$ landmarks, which is the most common case since AFLW dose not provide upper and lower lip landmarks. If face contour fitting fails, we label the $1, 5, 9, 13, 17$ landmarks.

Step 6. We add the additional labelled landmarks and re-fit the 389 samples with MFF. Fig. 5 demonstrates some samples in AFLW2000-3D.

REFERENCES

[1] T. Hassner, S. Harel, E. Paz, and R. Enbar, “Effective face frontalization in unconstrained images,” in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015. 1

[2] P. Hu and D. Ramanan, “Finding tiny faces,” *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*, 2017.

[3] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, “300 faces in-the-wild challenge: The first facial landmark localization challenge,” in *Computer Vision Workshops (ICCVW), 2013 IEEE International Conference on*. IEEE, 2013, pp. 397–403. 1

[4] S. Romdhani and T. Vetter, “Estimating 3D shape and texture using pixel intensity, edges, specular highlights, texture constraints and a prior,” in *Computer Vision and Pattern Recognition (CVPR), 2005 IEEE Conference on*, vol. 2. IEEE, 2005, pp. 986–993. 1, 3

[5] X. Zhu, Z. Lei, J. Yan, D. Yi, and S. Z. Li, “High-fidelity pose and expression normalization for face recognition in the wild,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 787–796. 1

[6] M. Köstinger, P. Wohlhart, P. M. Roth, and H. Bischof, “Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization,” in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2144–2151. 3