Nonparametric Context Modeling of Local Appearance
for Pose- and Expression-Robust Facial Landmark Localization

Brandon M. Smith[1]   Jonathan Brandt[2]   Zhe Lin[2]   Li Zhang[1]

[1]University of Wisconsin–Madison   [2]Adobe Research

http://www.cs.wisc.edu/~lizhang/projects/cvpr-face/

Figure 1 shows a visual overview of our pipeline, and supplements Section 3 of our paper. Figures 2, 3, and 4 supplement Figure 3 in our paper. The input images in Figure 2 are from the AFW dataset [3], and the input images in Figure 3 are from the IBUG dataset [2]. We note that our algorithm generally produces accurate results, even on faces with extreme head poses and expression. However, our algorithm is not perfect and makes mistakes on some especially challenging face, as shown in Figure 4.

In our judgment, faces with unusual mouth shapes and/or significant yaw+pitch head rotation are the most challening for our implementation. We speculate that this is largely due to the limited head pitch rotation variation and mouth shape variation in the MultiPIE Face Database [1], which we use as our sole exemplar database. Head yaw and pitch combinations like those shown in the left two columns of Figure 4, and the mouth shapes shown in the right two columns of Figure 4, are not represented well in our database, which results in poor landmark votes from the top exemplar images to the test image. Despite these challenging, our algorithm generally performs well on mouths, and on faces with significant yaw+pitch head rotation.

References

[1] R. Gross, I. Matthews, J. F. Cohn, T. Kanade, and S. Baker. Multi-PIE. In *IEEE International Conference on Automatic Face and Gesture Recognition*, 2008.

[2] G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. 300 Faces-In-The-Wild Challenge (300-W), 2013. http://ibug.doc.ic.ac.uk/resources/300-w/, 2013.

[3] X. Zhu and D. Ramanan. Face detection, pose estimation, and landmark localization in the wild. In *CVPR*, 2012.
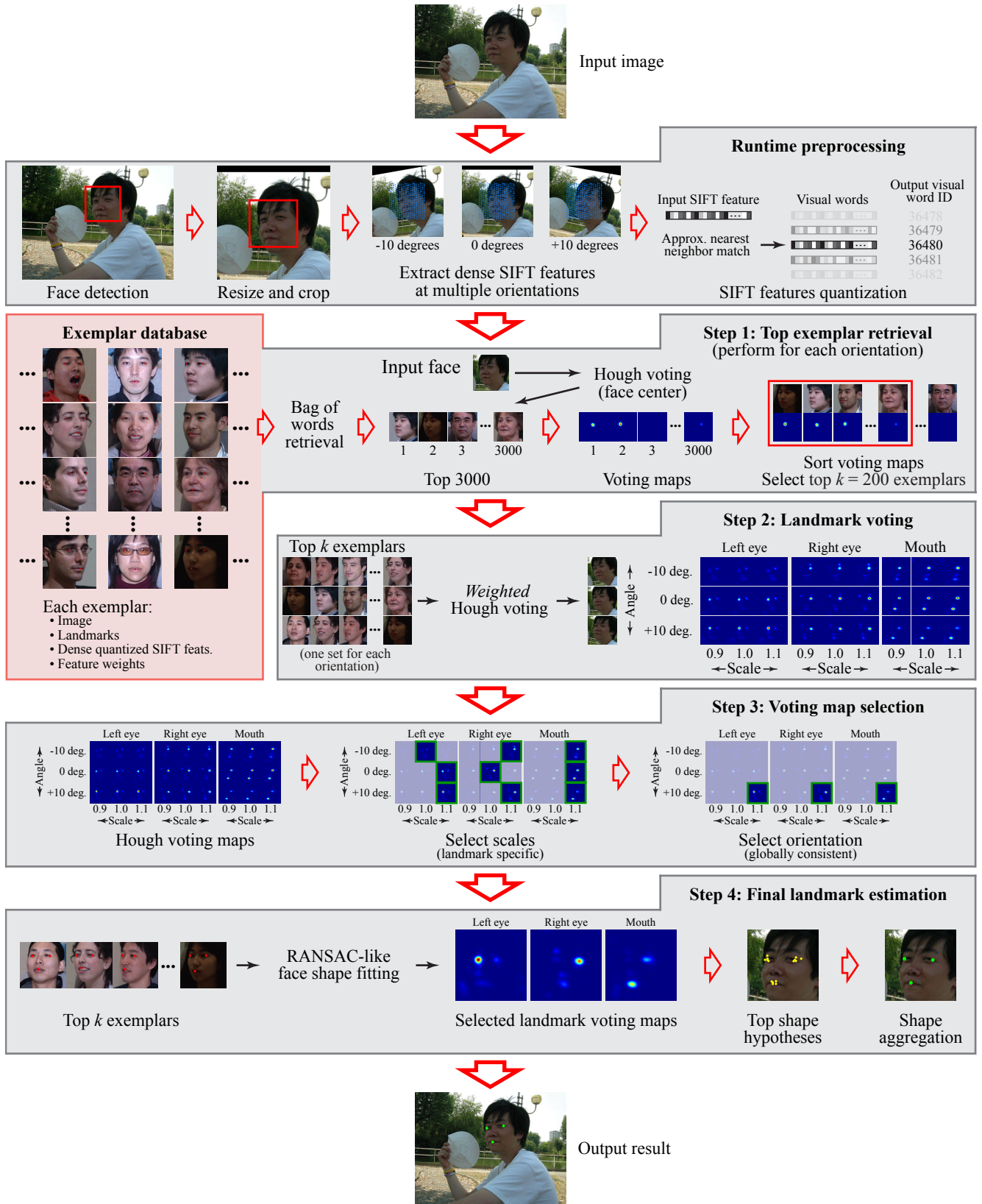
Figure 1. Overview of our pipeline. Above, for simplicity, we search over 3 face orientations, 3 face scales, and we compute 3 landmarks. However, in practice, we search over 7 orientations, 7 scales, and we compute 68 landmarks. Please see Section 3 of our paper for more details.

Figure 2. **Selected qualitative results on AFW [3].** Our method can handle a wide variety of very challenging conditions, including significant image noise and blur, occlusions, and extreme expressions and head poses. **Best viewed electronically in color.**

Figure 3. **Selected qualitative results on IBUG [2].** Our method can handle a wide variety of very challenging conditions, including significant image noise and blur, occlusions, and extreme expressions and head poses. **Best viewed electronically in color.**
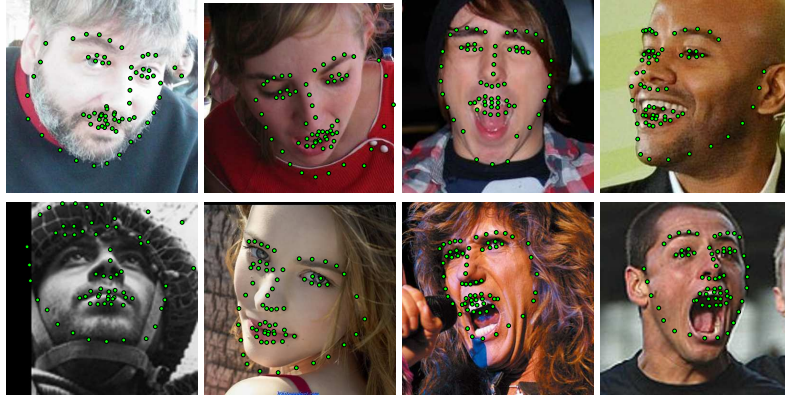

Figure 4. **Failure cases on faces with significant yaw+pitch rotation (left two columns) and on mouths (right two columns).** The input images on the top row are from AFW [3], and the input images on the bottom row are from IBUG [2]. Large errors occur infrequently, but when they do occur they are usually localized to the eyes (in the case of significant yaw and pitch head rotation) and the mouth (in the case of unusual mouth expressions). Yaw and pitch combinations and unusual mouth expressions like those shown above are not represented well in our exemplar database, which results in poor landmark votes from the top exemplar images to the test image. **Best viewed electronically in color.**