Face Alignment Post-Processing

Facial Contour Landmark Refinement

Andy Zeng

Departments of Computer Science, Mathematics, University of California, Berkeley

CMU Computer Vision Group

The Robotics Institute

Introduction

Problem

State-of-the-art face alignment performance:

Comprehensive (high-level) facial contour accuracy: excellent

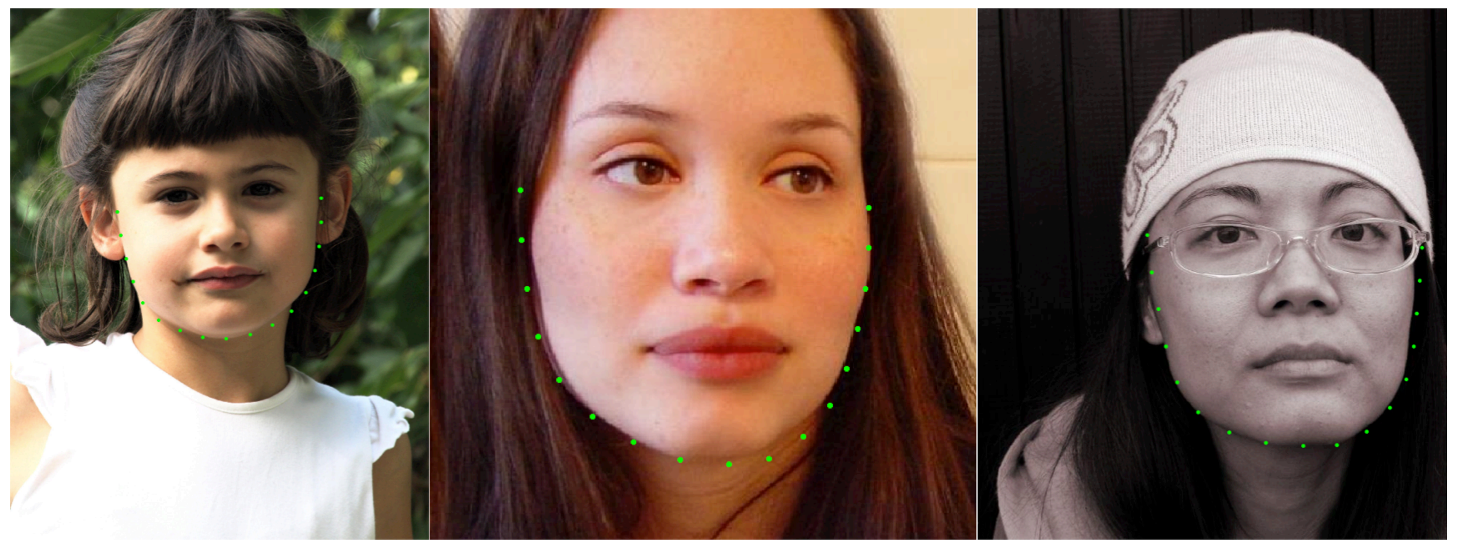
Local landmark (low-level) facial contour accuracy: poor

Problem Source

Algorithmic weaknesses of machine learning:

- Deficient optimization and regression
- Insufficient training
- Erroneous feature detection

Examples



Motivation

Poor local accuracy compromises quality performance for facial contour accuracy-sensitive applications such as 3D face modeling and face animation.

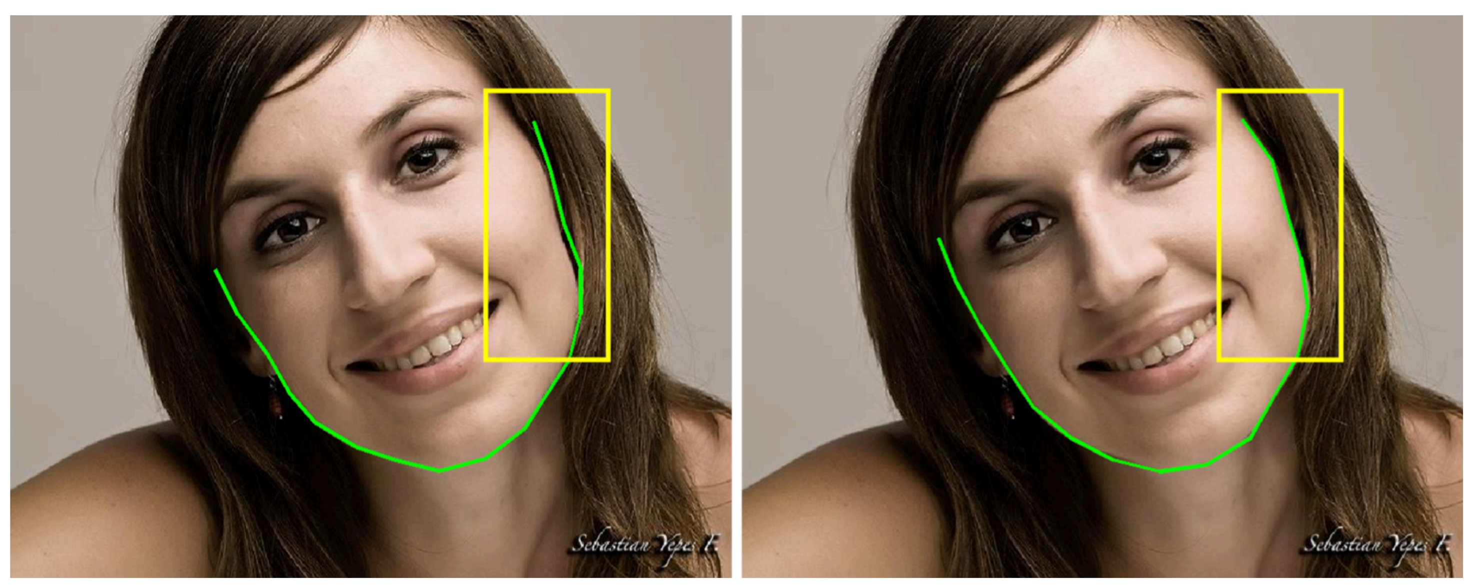
Applications

The popularity of face alignment is ever increasing in industry today, and many applications continue to suffer from local contour inaccuracies. Here are some examples:

- **TAAZ.com** - online virtual facial makeover application.
- **Face Switching** – special effect application.
- **Face De-Identification** – facial stock replacements to preserve online photograph privacy.
- **Burst-mode Facial Replacement** – "the perfect picture."

Goal

Create an efficient edge detection based facial contour refinement algorithm for face alignment post-processing, and demonstrate its effectiveness in fixing such contour inaccuracies.



Results

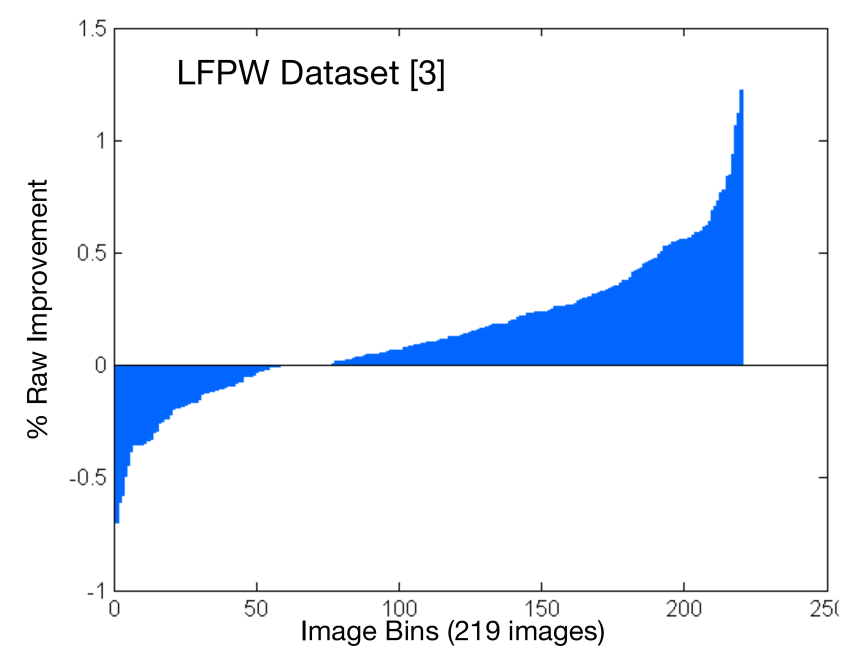
A selection of sample post-refinement results from the Helen image dataset [4]. Landmark point color codes: ground truth labels, occluded alignment labels, non-occluded alignment labels, post-refinement labels.

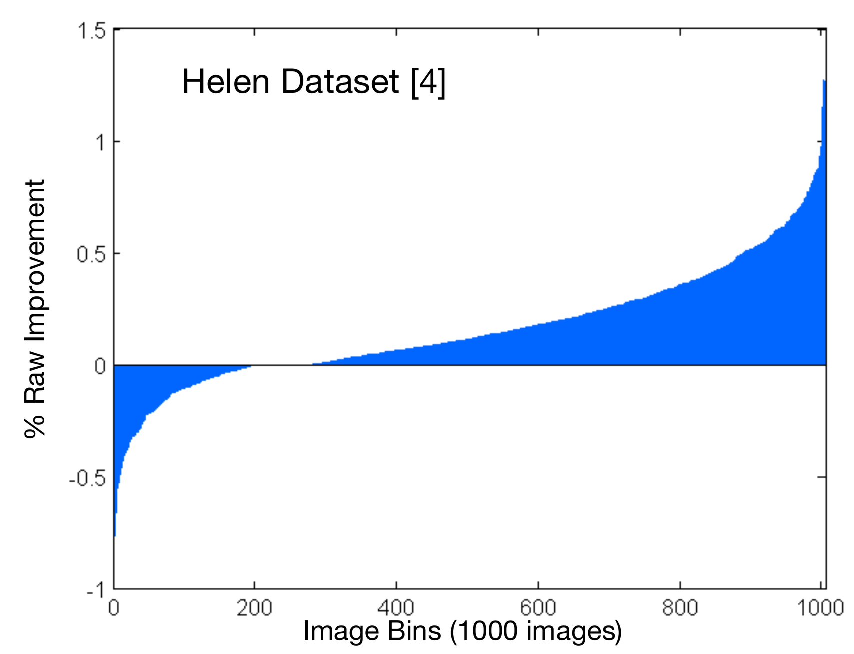
Experiments

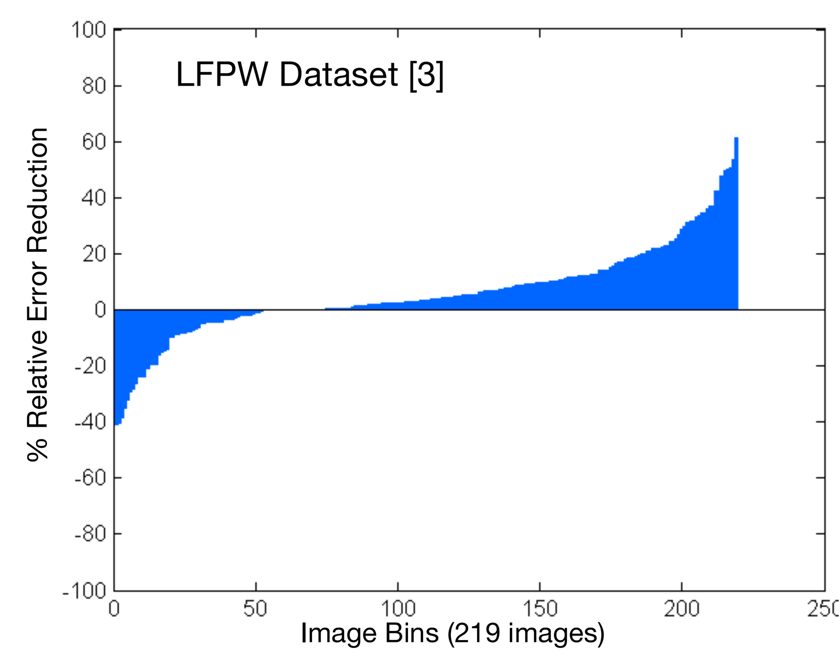
Quantitative Performance

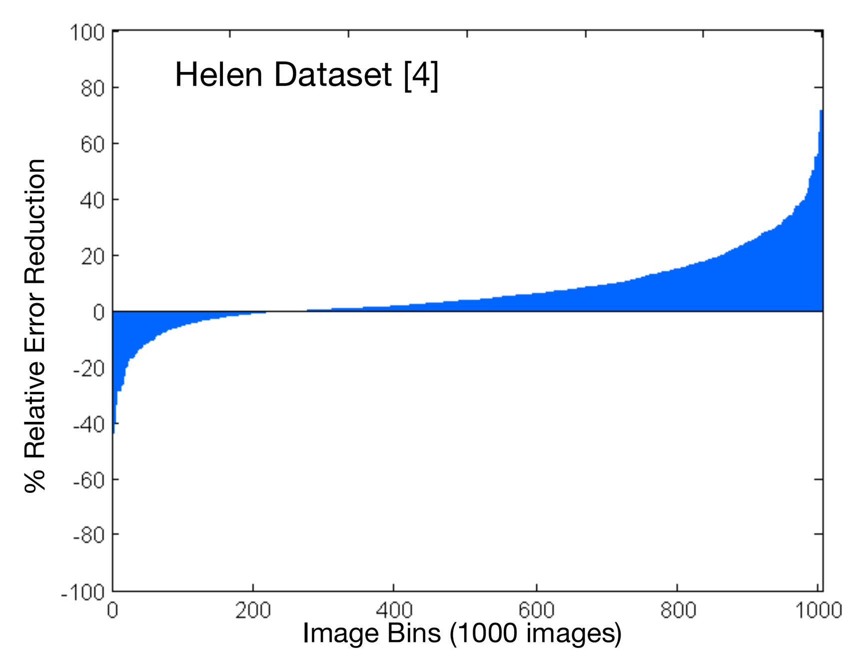
Measurement logistics:

- Alignment algorithm used for pre-refinement: All Pose Face Alignment Robust to Occlusion [1]
- Error metric: average perpendicular distance to the ground truth contour normalized with respect to the inter-pupillary distance.



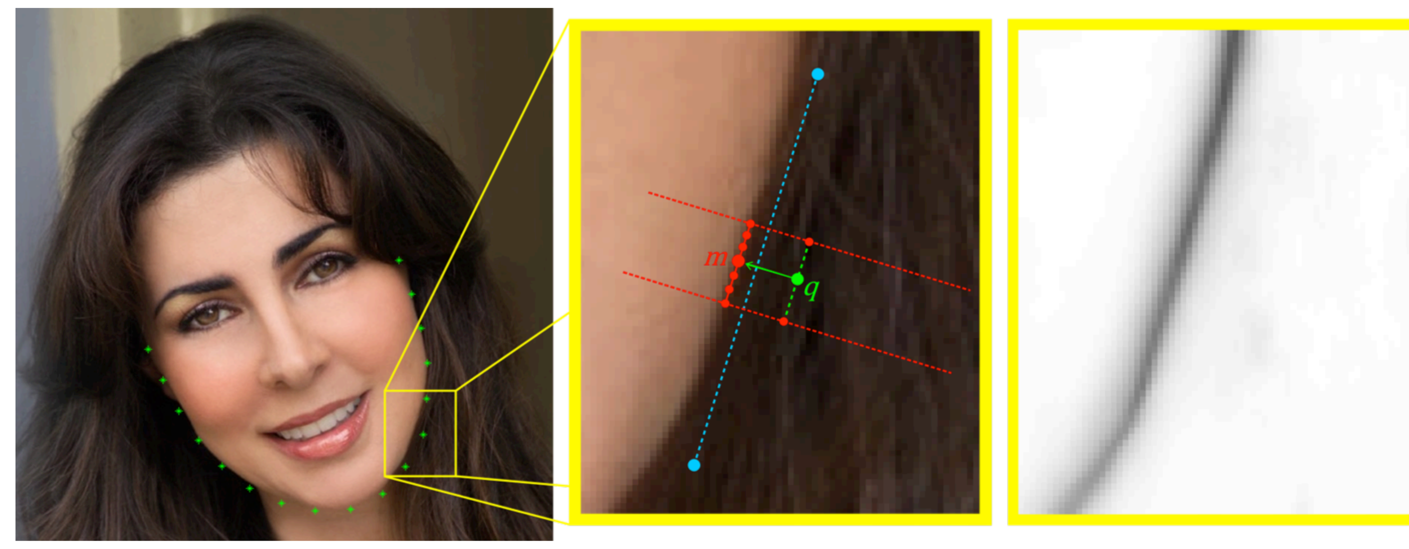






Method

Algorithm



For a cascading series of line segments parallel to the contour generated by the alignment result, shift the original alignment landmark q to the midpoint m of the line segment l_i with the highest score determined by function f

$$l_i = \{p_1, p_2, ..., m, ..., p_n\} \qquad f(l_i) = \frac{\sum_{k=1}^{n} E_{PD}(p_k)}{n} \cdot \left(1 - \frac{4||m-q||_2}{\varphi}\right)$$

where φ is the detected inter-pupillary distance and E_{PD} is the response function returned from multi-scale Piotr Dollar's edge detection [2].

Discussion

Observations

Remaining issues to be addressed:

- An edge detection based approach to contour adhesion is naturally weak and sensitive to shadows and hair.
- The local accuracy of human-labeled ground truth for many of the images from the datasets is questionable. A qualitative assessment is needed for a more accurate representation of visual performance.

Future Work [in progress]

Use machine-learning algorithms to train a classifier(s) to prematurely reject pre-refined landmark points that could potentially suffer from negative improvement after refinement.

References

[1] V. N. Boddetti, M. Roh, J. J. Shin, T. Oguri, T.Kanade "All Pose Face Alignment Robust to Occlusion," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014.

[2] P. Dollár and C. Zitnick, "Structured Forests for Fast Edge Detection", *International Conference on Computer Vision*, 2013.

[3] P. N. Belhumeur, D. W. Jacobs, D. J. Kriegman, and N. Kumar, "Localizing parts of faces using a consensus of exemplars," in *IEEE Conf. Computer Vision and Pattern Recognition*, 2011.

[4] V. Le, J. Brandt, Z. Lin, L. Bourdev, and T. S. Huang, "Interactive facial feature localization," in *European Conference on Computer Vision*, 2012.