A Brief Comparison of Implementations

groveh

April 2019

1 Introduction

The article from the previous report, "In-plane face orientation estimation in still images" [1], presented a methodology where orientation estimation was a direct result of facial detection. In relation to methods that I use for determining rotation estimation within natural images, this approach is distant because of an absence in frontal face detectors in my implementation. What our two approaches do share, however, is the multi-stage process of determining orientation estimation from varying angle intervals.

2 Discrete vs. Continuous Estimation

Course estimation is usually performed when the true orientation angle of an image resides within a large range of angles, such as the full 360 degree spectrum. A discrete approach estimates various selected angles within the full range of possibilities, typically a set interval apart. Continuous estimation handles more precise rotation estimation within a narrower range of possible ground truth angles. This fine estimation is typically found in the later stages of algorithms, and is responsible for a final output. My implementation of this process begins with discrete estimation of the course angle over a full 360 degrees within $\pm 45^\circ$, using an approach that uses 8 discrete rotation angles that are 45° apart. This network then feeds the image into a network trained to expect an image to be no more than 45° in either direction from true orientation. Using continuous estimation within this range, a final angle estimate of the original image is the resulting output. This is similar to the approach in Danisman, Bilasco [1] where they perform discrete estimations over intervals of 45° except within common upright angles, namely $\pm 50^\circ$, where they use a 3° interval. This network then feeds the image into a network with a narrower range of estimated angles, estimating no more than $\pm 40^\circ$ with an interval of 3°. In addition to this second network, they measure the range of angles at which a frontal face detector finds a face as a form of validation. While my implementation uses a similar methodology, the notable differences are *i.* My course estimation does not infer common upright angles, and therefore the interval is consistent, *ii.* I employ a continuous regression algorithm with zero interval for my precise estimation stage, contrast

to their narrower network, and *iii.* the final output of my precise network is the final output, with no form of validation in contrast to their implementation where they require a range of angles of true positive detection to validate their output, and feed it back into the network when validation fails.

References

[1] Taner Danisman and Ioan Marius Bilasco. In-plane face orientation estimation in still images. *Multimedia Tools and Applications, Springer Verlag*, 2016.