Orientation Estimation Using Facial Detection

groveh

April 2019

1 Introduction

Face detection is a relatively recent task that has now become one of the most widely known machine learning feats. This is due to its accuracy as well as because major companies have prioritized this problem with a significant portion of their research and development departments. The general public has been made aware of this issue from companies like Facebook, integrating this technology even when it was at an early stage of development. While advancements in detection have improved over the years as a result of improvements in numerous stages, orientation detection can be credited for some of this fame. The academic paper In-plane face orientation estimation in still images observes the importance of facial features being upright, as well as take a unique approach to determining the rotation within two dimensional images of faces.

2 The Importance of Orientation

Face detection algorithms are very advanced in locating and utilizing facial features to determine both if there is a face present, as well as whose face it is in the image. They can pick up multiple faces, and some at real time speed. Most facial algorithms, whether gender, expression, or recognition, require upright images of faces in order to maintain high levels of accuracy. Natural directions for orientation movement consist of in-plane (roll) and out of plane rotations (pitch and yaw).

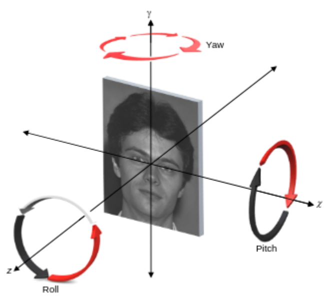

Figure 1: In-plane rotation (roll) versus out-of-plane rotation (pitch and yaw)

Many recognition algorithms initially assume that both in and out of plane rotation is correct, ignoring possible deviations in orientation. A neglect to rotation limits the use of algorithms in a more natural setting where any of the possible rotations may be present. Roll rotation estimation can be calculated either prior to or after the process of detection. Because of this, many frameworks have included multiple stages of detection and orientation estimation. Initial estimation is typically run on rotation invariant datasets, or instances with multiple views. These are too complex to be applied to naturally found photographs, however the multi-level approach has been shown to be problematic when two classifiers (before and after orientation estimation) disagree on the detection of a face, and post detection estimation usually requires feature detection or symmetry computation. This paper aims to address facial orientation estimation by exploiting the limits that algorithms can detect faces, called the rotation disadvantage.

3 Viola Jones Detection

This paper built their in-plane orientation estimation algorithm on top of a Viola-Jones (VJ) frontal face detector, which boasts a high true positive rate, paired with a low false positive rate and it runs in real time because it is a face detection only algorithm with no recognition. The goal of this paper is to determine faces from non faces, so recognition is irrelevant to the task at hand. The VJ detection algorithm is a strong binary classifier built from numerous weak detectors, referred to in machine learning as boosting. For every weak model throughout the chain, the following model is created to attempt to correct at least some of the errors the the previous model let through. VJ uses AdaBoost (Adaptive Boosting) which is adaptive in the sense that subsequent weak models are tweaked in favor of instances misclassified as a result of previous weak classifiers.

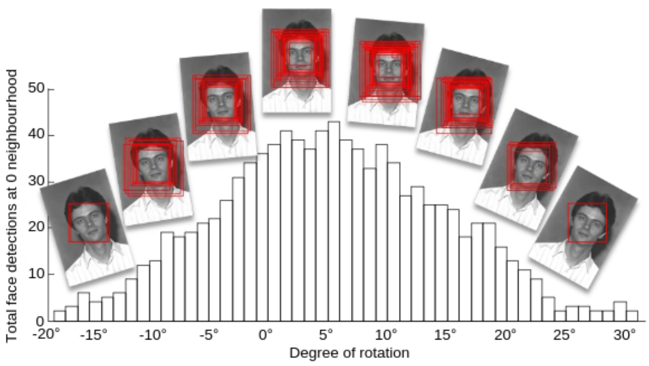

Figure 2: Number of VJ detections obtained at different degrees of roll rotation.

4 Implementation

Rotation disadvantage is the point at which the VJ algorithm cannot detect a face anymore due to in-plane rotation. This is precisely the behavior that this paper aims to utilize in order to estimate the orientation of images. True positive rates have been measured to pin this limit at around ± 15 degrees of optimal orientation. Figure 2 shows the distribution of faces detected at each angle in a specific range around an images true orientation. With an upper and lower bound of where the algorithm can detect a face, a mean angle can be calculated to infer the natural rotation. The implementation has a two levels in the detection process, shown in Fig.

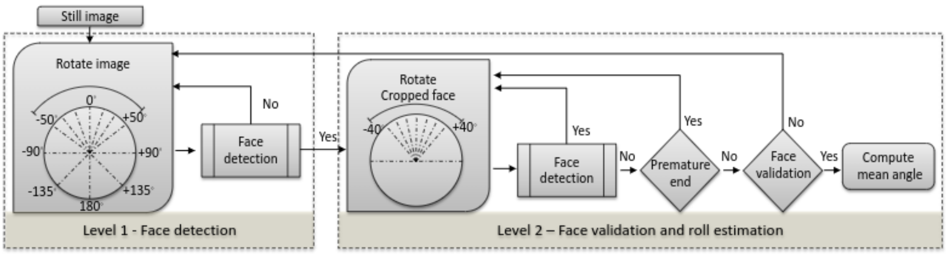

Figure 3: The two levels of facial detection.

3. The first involves repeating frontal face detection at various orientation angles, and the second uses these images to compute the minimum and maximum angles that the detection algorithm still finds a face. The mean value of these detections from level 2 is used to determine the roll rotation of the face. Fig. 3 shows a more detailed diagram of the process.

More specifically, a still image is fed into the first level of the algorithm where it estimates the course level orientation at angles at a 45 degree interval within the whole 360 degree spectrum, as well as more common head pose angles between $\pm 50°$ in intervals of 3 degrees. If a face is found during any of the intervals, the image is sent to the second level of detection where they undergo a relatively similar process, but this stage is a more specific detection, repeating frontal face detection between $\pm 40°$ at a 3 degree interval. For each detection in level 2, the image is fed back into the algorithm in order to capture every detection possible within the range. The premature ending stage ensures that an ample range of angles is used in order to capture the mean. The final stage is face validation where the face if confirmed if there exists an angular confidence of some C degree of successive detections, in the case of this paper $C = 21°$. The level 2 sequence can be better visualized in Fig. 4.

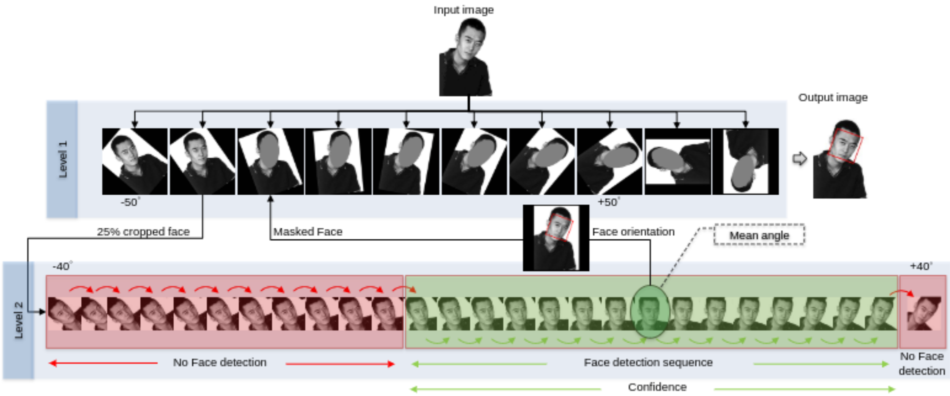

Figure 4: The process of level 2 and validating the face through a sequence of detections.

5 Results

The algorithm in this study results in accuracies within a range of $\pm 1.1°$ to $\pm 3.5°$ mean error on public datasets (CMU Rotated, Boston University Face Tracking (BUFT), Caltech, FG-NET Aging, BioID and manchester Talking-Face), shown in Fig. 5. A major contribution of this study is that orientation estimation is an immediate result of facial detection with no added complexity. The same algorithm that is used to detect faces within an image is also used to find the upright position of the face from mean angle calculated with the rotation disadvantage principle. They stress the necessity for a simple solution to orientation estimation referencing the universal nature of the fundamental problem. Using repetitive frontal face detections in two different intervals along with an angular confidence for validation, the method described in this paper can be applied to any 2D image in the wild.

Dataset	RMSE	MAE	STD
BioID	2.73 °	2.20 °	1.61 °
BUFT	2.68 °	1.93 °	1.83 °
Caltech	1.46 °	1.13 °	0.91 °
CMU	3.75 °	2.87 °	2.42 °
FG-NET	2.63 °	1.86 °	1.86 °
Manchester	2.07 °	1.69 °	1.18 °
YouTube C	5.59 °	3.50 °	4.36 °

Figure 5: Results on well known datasets.