Robustifying Gaze Direction Estimation for Emerging Psychology-Related Applications

Laura Florea
Image Processing and Analysis Laboratory, University Politehnica of Bucharest
313 Splaiul Independentei, Bucharest, Romania
laura.florea@upb.ro

Abstract

The accuracy of estimating a person's direction of gaze from remote imaging is discussed in the framework of emerging psychology-related eye tracking applications. Eye Accessing Cues (EAC) from the Neuro-Linguistic Programming (NLP) hypothesis is envisaged. For large scale investigation, we use a person independent, multistage system for landmark localization in the eye area. It is followed by eye region analysis for gaze recognition. The gaze recognition system is inspired by human perception and uses the landmarks on the eyebrows for a better accuracy.

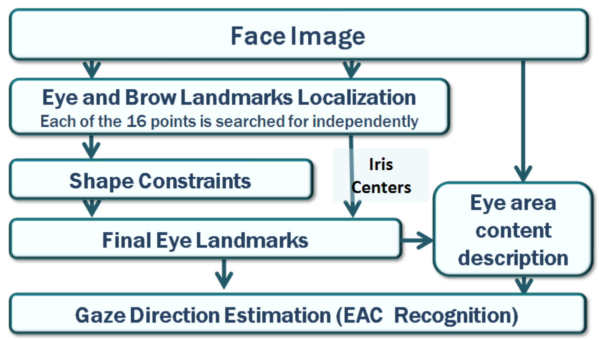

Figure 1. Algorithm flow chart.

1. Introduction

Building upon the idea that "eyes are the gateway to the soul", many hypotheses tried to correlate eye movement with cognitive processes. Neuro-Linguistic Programming (NLP) appeared as a theory in the 80' and attracted quickly many adherents. Yet the theory went directly into the commercial zone and the lack of scientific validation brought as many retractors. One of the theory's hypothesis is related to Eye-Accessing Cues (EAC) model which states that given a query, the user's way of addressing it (constructing an answer or accessing auditive, visual or kinesthetic memories) is correlated with the gaze direction. Recent psychological research [4] concluded that further testing is needed to validate it. Our experiments, employing multiple times more users than usual, showed that in a 7-case gaze direction hypothesis 35% of the subjects (compared to 14% random chance) behave according to the theory and respectively 66% for a 3-case scenario [5]. Thus, large scale experiments are required to determine if automatic separation of positive cases is possible.

The investigation system should be non-intrusive as voluntary control on subconscious eye movements is possible. We discuss a system [1] that exploits an usual digital video camera, infers eye features and upper face landmarks positions in order to robustify the gaze direction estimation. The separation of the gaze direction along the vertical axis is shown to improved by relating to brow landmarks.

2. Methodology

The proposed system for automatic recognition of gaze direction is schematically presented in figure 1. The eye and brow landmarks are detected using a multi-stage fusion approach. The resulting information is added to the position of the landmarks in order to enhance the gaze direction recognition.

Eye and brow landmarks localization. The landmark localization algorithm starts from the face detected square. The set of searched landmarks contain five points for each eye (center and four eye socket limits) and three points for each brow. For each landmark, position and intensity priors are computed on a training database. The position prior for one point is the two-dimensional histogram of the positions given the face square. The intensity prior is the probability of that point to have a certain gray-level intensity with respect to its surroundings.

A template matching algorithm is used to search in a neighborhood of an initial approximation of the landmark (found on the position prior map) and for each location the probability to have the true landmark is computed. A window centered in the investigated location is defined and represented in the descriptor space. A machine learning system estimates the likelihood of the current window to be centered in the true position of the landmark. Descriptors used are combination of integral and edge projections and we showed that a Multi-Layer Perceptron is enough for achieving good results.

The final probability of each point in the region of interest is computed as the weighted mean of the likelihoods given the matching probability, position and intensity priors. Additional shape constraints are further imposed on the relative positions of the landmarks: for each landmark, we iteratively use the global shape to construct a local constrain. The shape constraints are imposed only to landmarks that have small position variations; eye centers, which exhibit large variations, are excepted.

Gaze direction recognition. Eye landmarks are insufficient to precisely recognize gaze direction [2], thus we add information taken from the eye area in order to improve the results. The eye area is placed between the landmarks of the eye and brow (and different shapes were proposed in [2], [6], [1]). We use projections to describe the eye area, since they can be computed in real-time. The final descriptor of the eye area will be fed to a properly trained nu-SVM.

Training and testing. The eye landmark localization system was trained on 1000 annotated images from the PUT database [3]. We specifically built the Eye-Chimera database [2] that contains 1172 frontal face images, grouped according to the 7 gaze directions given by the EAC and we used it for extensive testing of the gaze direction recognition. The proposed system is person independent and the training/testing cases ratio is 1 train example to 3 testing ones.

3. Results

In table 1 we present the confusion matrix for gaze detection in a 7–case scenario. One can easily notice that most of the errors come from the looking down case. Still in our case these errors are smaller than in the older attempts.

We tested if the brow landmarks add valuable information to the overall recognition rate. When only landmarks are used as features for the gaze direction classification, the recognition rate increases with more than 10% if the eye and brow landmarks are used compared to the case where only the eye landmarks are at hand. While the use of brow landmarks may not be so intuitive, let us note that when one is looking down, is also lowering the lid, fact that humans acknowledge by relating to the eyebrow position.

In table 2 we present the confusion matrix for the up-down separation when only the eye landmarks are used and respectively when the eye and the eyebrow contribute. In the first situation the classifier tends to declare most cases as "looking down". Yet by adding the brow landmarks, the correct detection rate increases.

Table 1. Confusion matrix for the 7 cases computed on the Eye-Chimera database. $\odot$ denotes looking straight forward. The arrows in the left hand columns signals the ground truth while the top row indicates the reported direction.

Gaze	$\odot$	$\nearrow$	$\nwarrow$	$\rightarrow$	$\leftarrow$	$\searrow$	$\swarrow$
$\odot$	**0.77**	0.06	0	0.06	0	0.07	0.01
$\nearrow$	0.06	**0.73**	0.01	0.19	0	0.01	0
$\nwarrow$	0.06	0	**0.75**	0	0.18	0	0.01
$\rightarrow$	0.03	0.2	0.01	**0.59**	0	0.16	0.02
$\leftarrow$	0.06	0.01	0.32	0	**0.47**	0.01	0.14
$\searrow$	0.08	0.02	0	0.12	0.01	**0.58**	0.13
$\swarrow$	0.06	0	0.04	0.01	0.15	0.16	**0.58**

Table 2. Confusion matrix for the up-center-down separation when only eyes and respectively eyes and eyebrow landmarks are used.

Only Eye

Dir	$\odot$	$\uparrow$	$\downarrow$
$\odot$	**0.23**	0.04	0.72
$\uparrow$	0.36	**0.19**	0.45
$\downarrow$	0.22	0.02	**0.76**

Eye+EyeBrow

Dir	$\odot$	$\uparrow$	$\downarrow$
$\odot$	**0.38**	0.16	0.47
$\uparrow$	0.44	**0.37**	0.19
$\downarrow$	0.18	0.01	**0.81**

4. Conclusions

Eye Accessing Cues hypothesis from NLP theory is partially validated; results indicated large-scale and diversified tests are needed to separate positive cases from negative ones. When investigating the direction of gaze with non-mounted non-active illumination camera, by incorporating eyebrow related information, the overall accuracy increases.

References

[1] L. Florea, C. Florea, and C. Vertan. Remote recognition of the gaze direction: Anchoring with the eyebrows. *JVCI*, 35:67–77, 2016.

[2] L. Florea, C. Florea, R. Vranceanu, and C. Vertan. Can your eyes tell me how you think? A gaze directed estimation of the mental activity. In *BMVC*, 2013.

[3] A. Kasiński, A. Florek, and A. Schmidt. The PUT face database. *Image Processing & Communications*, 13(3-4):59 – 64, 2008.

[4] J. Sturt, S. Ali, W. Robertson, D. Metcalfe, A. Grove, C. Bourne, and C. Bridle. Neurolinguistic programming: systematic review of the effects on health outcomes. *Br. J. Gen Pract.*, 62(604):757–764, 2012.

[5] R. Vranceanu, C. Florea, L. Florea, and C. Vertan. NLP EAC recognition by component separation in the eye region. In *CAIP*, pages 225–232, 2013.

[6] R. Vranceanu, C. Florea, L. Florea, and C. Vertan. Gaze direction estimation by component separation for recognition of eye accessing cues. *MVAP*, 26(2-3):267–278, 2015.