GAZE DIRECTION ESTIMATION TOOL BASED ON HEAD MOTION ANALYSIS OR IRIS POSITION ESTIMATION

Alexandre Benoit, Alice Caplier, and Laurent Bonnaud

Laboratoire des Images et des Signaux,
46 avenue Félix Viallet, Grenoble, France
email: name.surname@.lis.inpg.fr
web: www.lis.inpg.fr

ABSTRACT

In this paper, we present two versions of a demonstrator for gaze direction estimation. The first version is based on the analysis and the interpretation of head movements, especially in case of important head rotations. In such a case, we suppose that the gaze direction coincides with the head orientation. The second version of the demonstrator is based on the analysis of the iris centres positions (with respect to the eyes corners) and it supposes that the head remains quasi static or undergoes small translations. This version supposes also that the size of the head in the processed frame is big enough to be able to extract the iris centre position with enough accuracy. A typical example of use is the analysis of the gaze direction of a user working in front of a PC.

In both cases, the result of the demonstrator is a characterization of the gaze direction in term of looking left, right, up or down. An evaluation of the accuracy of the estimated gaze direction is beyond the scope of the present work.

1. INTRODUCTION

This paper presents a demonstrator for gaze direction estimation based on head movement or on iris centre position. Such a demonstrator could be useful for dumb people and for paraplegic people who lost the use of their hands. Indeed, it is possible to associate specific task to some particular head movements.

In this demonstrator, we suppose that the gaze direction is related to the head orientation if the head is moving and that it is related to the iris position is case of quasi-static head. The data acquisition system is a single digital camera. The main advantage of this solution is that iris position and head motion can be extracted from the same video data. Other systems for gaze direction estimation usually use more sophisticated acquisition systems such as infrared cameras, multiple digital cameras or optical zooming [1,2,3].

In sections 2 and 3, the methods for head motion analysis and iris position estimation are described. Section 4 gives a description of the demonstrator.

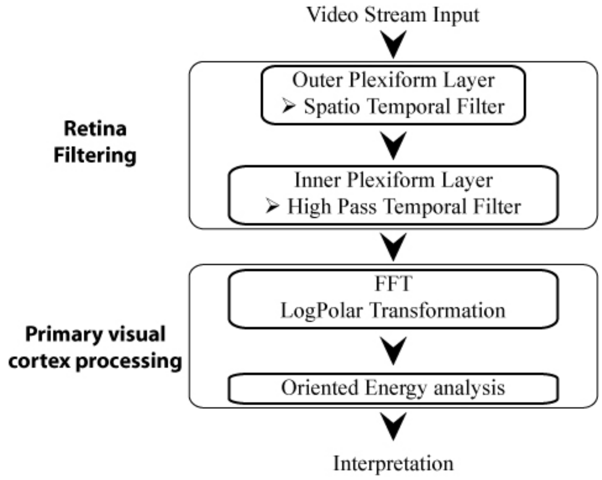

Figure 1: Algorithm for head motion analysis.

2. GAZE ESTIMATION BASED ON HEAD MOTION ANALYSIS

In case of important head rotations, we suppose that the gaze direction is mainly related to the head orientation. The first version of the proposed demonstrator is related to head motion analysis.

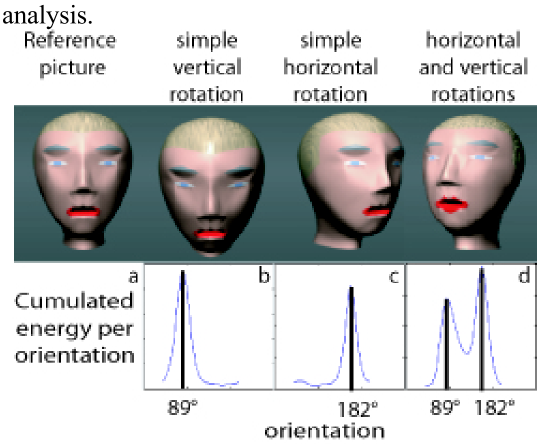

Figure 2: Simple and mixed head motion estimation

In [4] is described a method based on biological approach for head motion analysis. The method is based on the processing achieved by the human visual system (see figure 1). In a first step, a filter coming from the modelling of the human retina is applied. This filter enhances moving contours and cancel static ones. In a second step, the FFT of the filtered image is computed in the log polar domain as a modelling of the primary visual cortex. Head movements are related to the variation of the spectrum energy. They induce variations of the global spectrum energy of each image: this energy increases in case of moving head and is minimum in case of static head. This yields to an easy motion analysis: motion direction is related to the orientation of the maximum of the image spectrum energy and motion amplitude is related to the amplitude of the total energy spectrum. Figure 2 gives an example of the log-polar energy of the spectrum in case of different simple or mixed head motions. The abscissa of maximum of the log-polar energy gives the motion orientation.

Figure 3: (top) face extraction without temporal stabilisation of the square dimensions, **(bottom)** face extraction and median temporal filtering.

3. GAZE DIRECTION BASED ON IRIS POSITION

When the head is quasi static or if the head undergoes motion of translation only (for example, in the case of a person working in front of a PC), we suppose that the gaze direction is mainly related to the iris position with respect to the eyes corners. The second version of the proposed demonstrator is based on the detection of the eyes corners and iris centre based on the extraction of the iris contours.

3.1. Face detection and tracking

First of all, a bounding box around the face has to be extracted. Face extraction has been extensively studied. [5,6] present two recent surveys about this topic. For the proposed demonstrator, we chose to use the face detector developed by Viola and Jones [7]. This detector is efficient and fast and the code is free accessible [8]. As a result, the algorithm provides a square bounding box around the face and a point belonging to each eye (these points being not necessary the iris centres). Figure 3-top gives examples of the detected face bounding box with the MPT algorithm.

The main drawback of this algorithm is that the extracted face bounding box is not stable from frame to frame. In order to cope with this problem, we add a median temporal filter working on 5 consecutive frames. As a result, the dimensions and the position of the face bounding box from frame to frame remain stable (see figure 3-bottom).

3.2. Detection of the iris centre

The method used for the extraction of the iris contour and centre is described in [9]. Iris contour being the frontier between the dark area of iris and the eye white, it is supposed to be a circle made of points of maximum of luminance gradient. Since the eyes could be slightly closed, the upper part of the iris could be occluded. So for each iris, we are looking for the lower part of the iris circle. We suppose that the iris radius is known.

With the face bounding box dimension H_{face}, a bounding box (H_{eye}, L_{eye}) around each eye has been defined with the following relations:

$$H_{eye} = 0.1 * H_{face} \quad and \quad L_{eye} = 0.2 * H_{face}$$

These relations have been learnt on the 400 images of the ORL database [10]. Each bounding box is centred on the eye point given by the MPT algorithm.

In each eye bounding box, each semi-circle of iris maximizes the normalized flow of luminance gradient (*NFLG*):

$$NFLG = \frac{1}{length\ SC} \sum_{p \in SC} \vec{\nabla} I(p).\vec{n}$$

3.3. Detection of eyes corners

A tracking process of points with maximum value of luminance gradient is used in order to localize eyes corners.

The centre of eye and brow bounding box coincides with the iris detected centre. This allows an efficient delimitation of search area for eye and brow.

Figure 4 gives an illustration of the corners detection method: starting with the points X_1 and X_2, pixels of maximal gradient of luminance located at two pixels from the vertical of the iris circle limits, a tracking process, to the left direction, of pixels with highest value of luminance gradient yields to the detection of the first corner C_1. Since the initial point X_1 is located below the corner C_1, only the three (black) points $\begin{pmatrix} \bullet & \bullet \\ \bullet & X_1 \end{pmatrix}$ located above and on the left of X_1 are tested. The curve between X_1 and C_1 (resp. X_2 and C_2) is made of pixels with local maximum of luminance gradient. The tracking stops when the luminance gradient becomes negative since a skin pixel is clearer than an eye corner pixel.

A similar tracking process to the right direction yields to the detection of the second corner C_2.

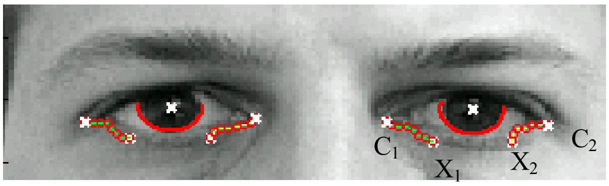

Figure 4: tracking process for eyes corners detection.

3.4. Iris position estimation

Once eyes corners and iris centres have been detected, the iris positions are estimated by computing the distances between the centres and the corners. A phase of initialisation with the person looking straight forwards is necessary.
In order to make the point's detector faster, the process of points tracking described in [11] is used.

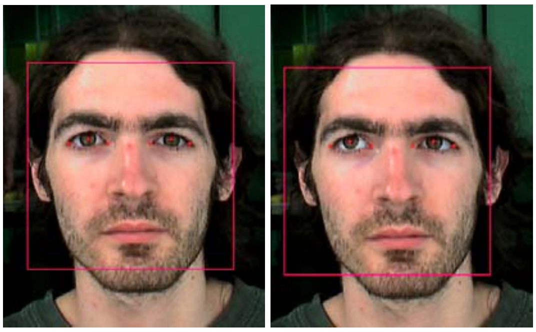

Figure 5: Different positions of iris centres **(a)** straight forwards, **(b)** on the right

4. DEMONSTRATOR

We propose two versions of the gaze direction estimation demonstrator. For both versions, data are acquired with a single digital camera. Images are processed on a standard PC, Pentium IV, 2GHz (????).
Figure 5 illustrates the detection of gaze direction based on the analysis of head motion. It is possible to know if a person is looking on the left, on the right, up or down. The precision of the gaze direction estimation is not yet available. This version of the demonstrator is running in real time.

Figure 5: Presentation of the demonstrator interface; in this case, the user is looking on the left.

Figure 6 gives an illustration of the estimated gaze direction based on the iris centres positions. The person in front of the computer screen looks at the different icons. The (red) points represent the fixation points that have been automatically detected. In this particular case, the head remains fixed and the complex calibration step described in [12] has been used. In case of use with no strong constraints on the head movements, such accuracy in gaze direction estimation is not possible. Only coarse information is available such as iris on the left, iris on the right, iris up or iris down.

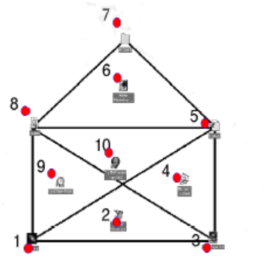

Figure 6: Gaze direction estimation

In the next version of the demonstrator, the virtual clone developed by the Institut de la Communication Parlée and presented on figure 7 is going to be integrated and will be used as a magic mirror.

Figure 7: Virtual clone and magic mirror

5. CONCLUSION

We presented a preliminary version of a demonstrator for gaze direction estimation. Two versions are available mainly depending on the head motion and on the size of the face in the current image. The fusion of both versions of the demonstrator in order to be able to cope with all the cases is under study.

ACKNOWLEDGMENTS :

This work was supported by the Similar network of excellence.

REFERENCES

[1] D. Hansen, A. Pece – Eye Tracking in the Wild – *Computer Vison and Image Understanding*, vol. 98, pp. 236-274, 2005.

[2] B. Noureddin, P. Lawrence, C. Man – A Non Contact System for Tracking Gaze in a Human Computer Interface – *Computer Vison and Image Understanding*, vol. 98, pp. 52-82, 2005.

[3] J. Wang, E. Sung, R. Venkateswarlu – Estimating the Eye Gaze from one Eye – *Computer Vison and Image Understanding*, vol. 98, pp. 83-103, 2005.

[4] A. Benoit., A. Caplier. –Motion Estimator Inspired from Biological Model for Head Motion Interpretation – *WIAMIS05*, Montreux, Suisse, April 2005.

[5] H. Hjelmäs, B. Low – Face detection: a survey - *Computer Vison and Image Understanding*, vol. 83, pp. 236-274, 2001.

[6] M.H. Yang, D. Kriegman, N. Ahuja - Detecting face in images : a survey - *IEEE Trans on PAMI*, vol.24, n°1, pp. 34-58, January, 2002

[7] P. Viola, J. Jones – Robust real time face detection - *International Journal of Computer Vision*, 57(2), pp. 137-154, May, 2004.

[8] Machine Perception Toolbox, face detection algorithm: http://mplab.ucsd.edu/grants/project1/free-software/MPTWebSite/introductionframe.html

[9] Z. Hammal, N. Eveno, A. Caplier, P.Y. Coulon - Parametric models for facial features segmentation- To be published in *Eurasip Journal of Signal Processing*

[10] The ORL Database of Faces, Cambridge University Department, http://www.uk.research.att.com/facedatabase.html

[11] Z. Hammal, C.Massot, A. Caplier - Eyes segmentation applied to gaze direction and vigilance estimation- *3rd International Conference on Advances in Pattern Recognition*, Bath, United Kingdom, August 2005