Study on the interest of hybrid fundamental matrix for head mounted eye tracker modeling

Sacha Bernet[1]
sacha.bernet@uha.fr
Peter Sturm[2]
peter.sturm@inria.fr
Christophe Cudel[1]
christophe.cudel@uha.fr
Michel Basset[1]
michel.basset@uha.fr

[1] MIPS Université de Haute Alsace
12, rue des frères Lumière
68093 Mulhouse cedex, France
[2] INRIA Rhône Alpes
Inovallée, 655 Avenue de l'Europe
38334 Saint Ismier cedex, France

In this paper, we propose a new approach for calculating a user's real point of regard (POR) using a head mounted binocular eye tracking system. We show that the coupling of one of the user's eyes and the scene camera can be considered as a stereovision system. We then propose a new description for binocular eye trackers, in which we model them via pairs of stereovision systems having the scene camera in common. In this model, we use a hybrid fundamental matrix which allows to take into account the spherical shape of the eyeball model we use. Whatever the distance our model does not require any a priori information on the user's morphology to calculate the POR.

We first described eye tracking system via a stereovision model. This model reveals that there is no longer a point-to-point relationship with a monocular eye tracker, but a point-to-line relation, i.e. to each position of the pupil will correspond a line in the scene camera (an epipolar line). Therefore, we suggest using a binocular eye tracker to determine the POR by the intersection of two epipolar lines.

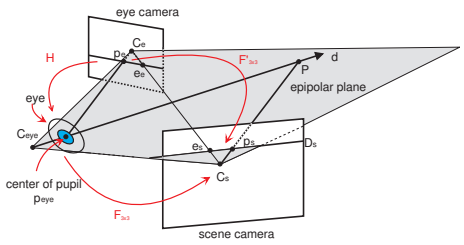

Figure 1: Model of the Eye-tracker system and the eye. Here, the eye's surface is considered to be planar.

The fundamental matrix noted F is the algebraic representation of the epipolar geometry. It only depends on the camera configuration (intrinsic and extrinsic parameters), not on the objects in the scene. For every point p_e (cf. figure 1) in the eye camera image the corresponding epipolar line is:

$$D_s = F\ p_e \tag{1}$$

The dimensions of the matrix F is 3×3. Knowing one of the two points p_e or p_s, the following relation is equivalent to the corresponding epipolar line and constrains the matching of points:

$$p_s^t\ F\ p_e = 0 \tag{2}$$

The relation between the eye plane and the eye camera can be defined as a homography as follows:

$$p_{eye} = H\ p_e \tag{3}$$

If F denotes the fundamental matrix between the scene camera and the eye, then the epipolar geometry between the eye camera and the scene camera, can be obtained by replacing p_e in equation (2), by the definition of p_{eye} as of equation (3):

$$p_s^t \underbrace{F\ H}_{F'}\ p_e = 0 \tag{4}$$

We propose to extend the previous stereo model by assuming a spherical eye shape. Although this remains an approximation to the true shape, it is much better than a planar model. Also, we know from previous works [2] that the pupil's displacement is well described by the eyeball model. It means that in any situation, the pupil center is located on a sphere.

In [1], it was shown that the epipolar geometry between a central catadioptric camera and a perspective one, can be described by a 6×6 "hybrid" fundamental matrix. We can transpose this directly to our case. Let us first write down the associated expressions and then explain their meaning and application. Let P be the POR and $p_s = (x_s, y_s, 1)^t$ its image in the scene camera (we sometimes call this POR too). Let $p_e = (x_e, y_e, 1)^t$ be the image of the pupil center in the eye camera. From the above observations and [1], it follows that there exists a fundamental matrix F of size 6×6 such that the epipolar constraint between p_s and p_e can be written as:

$$\begin{pmatrix} x_s^2 & x_s y_s & y_s^2 & x_s & y_s & 1 \end{pmatrix} F_{6\times 6} \begin{pmatrix} x_e^2 \\ x_e y_e \\ y_e^2 \\ x_e \\ y_e \\ 1 \end{pmatrix} = 0 \tag{5}$$

This expression can be interpreted as follows. For a given point p_s in the scene camera, the possible matching points p_e in the eye camera, must lie on a conic (the above equation is quadratic in the coordinates of p_e), the epipolar conic. The other way round is similar; however, the epipolar conic in the scene camera, degenerates into a pair of lines (cf. figure 2). The reason for this is as follows: when back-projecting a point p_e to the sphere, there are two mathematical solutions for the intersection of the sphere and the back-projection line, the true one and a "parasite" solution. To each of these, corresponds a line of sight of the eye. The epipolar curve is thus the union of the images of two lines i.e. a pair of lines or a degenerate conic.

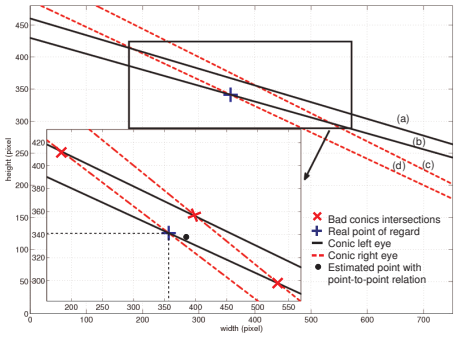

Figure 2: Calculation's simulation POR using a matrix F. Two lines (a) and (b) correspond to the conic of the left eye. Same for the right eye. The lines intersections give the 4 solutions for the POR.

[1] K. Barreto, J.and Daniilidis. Epipolar geometry of central projection systems using veronese maps. volume 1, page 1258Ű1265, 2006.

[2] Arantxa Villanueva, Rafael Cabeza, and Sonia Porta. Eye tracking: Pupil orientation geometrical modeling. *image and vision computing*, 24 (7):663–679, 2005.