EYE DIRECTION BY STEREO IMAGE PROCESSING

KUMIKO TSUJI
Fukuoka Medical technology, Teikyo University,
4-3-124 Shin-Katsutachi-Machi,
Ohmuta, 836 ,Japan
E-mail c74206g@wisdom.cc.kyushu-u.ac.jp

MIKI AOYAGI
Department of Mathematics, Sophia University,
7-1 Kioi-cho,Chiyoda-ku, Tokyo, Japan, 102-8554,
E-mail miki-a@sophia.ac.jp

The measurement system of a movement of an eye direction is important in ophthalmology. A human inspector usually examines the movement. The purpose of this paper is to analyze two eyes of a human inspector and to show the possibility of inspection by an image processing in stereo cameras with computer programming. In order to have the same functions performed by human eyes, two cameras are used, whose distance and direction to the center of a pupil circle are the same ones of human eyes.

Key words: Eye direction, stereo image processing, pattern recognition.
Mathematic subject Classifications: 68T10, 92C50

1. Introduction

Types of disability in the vision are now detected by classifying manners of nystagmus in ophthalmology [5]. A relationship between smooth pursuit of an eye movement and visual perception has been studied [4], by observing disabled children. In the paper [4], they used an electrooculogram and a video at the same time. The electrooculogram measures electric potential differences between a nose and an ear, by applying voltage to the nose and the ear. The video records an eye movement. By this method, only a movement is known but a detailed eye direction is not yet measured. In addition, there are more information when an inspector sees children eyes by his own eyes. So, it is important to develop a machine to determine an eye direction, which plays a role of inspector's eyes. The purpose of this paper is to understand the functions of inspection with two eyes and to show the possibility of inspection by an image processing in stereo cameras with computer programming. To understand the role of inspection with two eyes, movements of stereo cameras are controlled by angle sensors as follows: (1) the distance between two cameras is 60mm as the average distance between human eyes; (2) an attitude of two cameras is controlled to direct to the center of a pupil circle in the initial procedure.

An eye direction algorithm for an eye tracking has been developed using many frames of observed eye sphere images by CCD cameras [3]. The camera projection has been considered as affine. A model of an image pupil has been constructed by assuming that a contour image is an ellipse by an affine projection [3]. However these affine projections do not give sufficient accuracy in the case when the camera and the eye sphere are near each other.

In this paper, we introduce the method to calculate an eye direction, using a perspective projection. The problem in general is that the center of a pupil does not map to the center of an image ellipse in the perspective projection. In order to settle this problem, we propose an algorithm to control attitudes of cameras. If the attitudes of the two cameras are controlled to direct toward the center of the pupil, then the eye direction is obtained, since the image of the center of the pupil attains the origin in each camera coordinate. Furthermore we propose the procedure, called an 'epipolar radius procedure', by which the attitudes of the two cameras are correctly obtained, even when the two cameras do not direct correctly to the center of the pupil.

2. Coordinate systems for stereo cameras and object eye

In this section, the eye sphere, the pupil circle, and the locations and the angles of the cameras are defined, for coordinate systems of stereo cameras and an object eye.

2.1. *World coordinate system*

An object point in 'a world right hand coordinate system' is denoted as $\mathbf{X} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$.
Its origin lies at the center of an eye sphere. X axis is the horizontal axis, Y axis is the vertical axis and Z axis directs to the midpoint of the two camera locations.

2.2. *Eye sphere*

Let r_0 be the radius of an eye sphere and a_0 be the radius of a pupil circle. The average values in human eyes are $r_0 = 12$mm and $a_0 = 4$mm.

A half sphere point $\mathbf{q}(\beta, \gamma, r_0)$ in $Z \geq 0$ is expressed as

$$\mathbf{q}(\beta, \gamma, r_0) = r_0 \mathbf{u}_3 = r_0\ {}^t(\cos\gamma \sin\beta, \sin\gamma \sin\beta, \cos\beta),$$

where r_0 is the radius of the sphere. $\mathbf{u}_3$ is the vector given by rotating the point ${}^t(0, 0, r_0)$ on Z-axis around Y axis by the angle β in $0 \leq \beta < \frac{\pi}{2}$ and next rotating around Z axis by the angle γ in $0 \leq \gamma < 2\pi$. That is, $\mathbf{u}_3$ is the third column vector in the rotation matrix T, which is defined as

$$T = \begin{pmatrix} \cos\gamma\cos\beta & -\sin\gamma & \cos\gamma\sin\beta \\ \sin\gamma\cos\beta & \cos\gamma & \sin\gamma\sin\beta \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3).$$

An eye direction (θ, ϕ) is defined as $\theta = \arctan(\cos\gamma\tan\beta)$ and $\phi = \arctan(\sin\gamma\tan\beta)$ in [3]. In this paper, the eye direction is defined by $\mathbf{u}_3$, since $\mathbf{u}_3$ is obtained if and only if θ and ϕ are obtained:

$$\tan\gamma = \frac{\tan\phi}{\tan\theta}, \qquad \tan^2\beta = \tan^2\theta + \tan^2\phi.$$

2.3. *Pupil circle*

The pupil circle with the center $\mathbf{p} = \sqrt{r_0^2 - a_0^2}\mathbf{u}_3$ is defined as

$$a_0\cos\psi\mathbf{u}_1 + a_0\sin\psi\mathbf{u}_2 + \sqrt{r_0^2 - a_0^2}\mathbf{u}_3,$$

for an arbitrary ψ. Note that the pupil circle lies on the eye sphere and the center of the pupil disc lies on the same direction with $\mathbf{q}$. Its distance $\sqrt{r_0^2 - a_0^2}$ is shorter than r_0.

2.4. *Rotation matrices for camera angle sensors*

In this section, the rotation matrices between the two camera coordinates and the world coordinate are defined, as cameras are controlled to direct to the center of the pupil. Let $\mathbf{O}_r = \begin{pmatrix} d \\ 0 \\ O_Z \end{pmatrix}$ and $\mathbf{O}_l = \begin{pmatrix} -d \\ 0 \\ O_Z \end{pmatrix}$ be the locations of the right and left cameras, respectively, where O_Z is the distance from the center of the sphere to the midpoint of the segment with the two camera locations, and $2d$ is the distance of two cameras.

Let the right and left coordinates be attached to the lenses of the cameras as $\mathbf{X}_r = \begin{pmatrix} X_r \\ Y_r \\ Z_r \end{pmatrix}$ and $\mathbf{X}_l = \begin{pmatrix} X_l \\ Y_l \\ Z_l \end{pmatrix}$ with the origins at the centers of the right and left cameras. Here Z_r and Z_l direct to the center of the pupil. Y_r and Y_l are axes perpendicular to the space constructed from the vectors $\mathbf{p} - \mathbf{O}_r$ and $\mathbf{p} - \mathbf{O}_l$. That is, $Y = Y_r = Y_l$. X_r is an axis perpendicular to (Z_r, Y_r) plane and X_l an axis perpendicular to (Z_l, Y_l) plane.

Then an object point $\mathbf{X}$ satisfies

$$\mathbf{X} = R_r\mathbf{X}_r + \mathbf{O}_r, \qquad \mathbf{X} = R_l\mathbf{X}_l + \mathbf{O}_l.$$

Here R_r and R_l are rotation matrices such that $R_r = [\mathbf{r}_1\mathbf{r}_2\mathbf{r}_3]$ and $R_l = [\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3]$, where

$$\mathbf{r}_1 = \mathbf{r}_2 \times \mathbf{r}_3 = \frac{\mathbf{l}_3 \times \mathbf{r}_3}{\|\mathbf{l}_3 \times \mathbf{r}_3\|} \times \mathbf{r}_3, \quad \mathbf{r}_2 = \mathbf{l}_2 = \frac{\mathbf{l}_3 \times \mathbf{r}_3}{\|\mathbf{l}_3 \times \mathbf{r}_3\|}, \quad \mathbf{r}_3 = \frac{\mathbf{p} - \mathbf{O}_r}{\|\mathbf{p} - \mathbf{O}_r\|},$$

$$\mathbf{l}_1 = \frac{\mathbf{l}_3 \times \mathbf{r}_3}{\|\mathbf{l}_3 \times \mathbf{r}_3\|} \times \mathbf{l}_3, \quad \mathbf{l}_3 = \frac{\mathbf{p} - \mathbf{O}_l}{\|\mathbf{p} - \mathbf{O}_l\|}.$$

The rotation matrices are determined by information of the camera angle sensors. Then the attitudes of the camera angles are controlled by the rotation matrices.

3. Calculation of an object point X and relation between right and left image points

'A perspective projection' is defined by the two equations

$$\mathbf{x}_r = \frac{1}{Z_r}\begin{pmatrix} X_r \\ Y_r \end{pmatrix}, \qquad \mathbf{x}_l = \frac{1}{Z_l}\begin{pmatrix} X_l \\ Y_l \end{pmatrix},$$

where $\mathbf{x}_r$ and $\mathbf{x}_l$ are the right and left image points, respectively.

The above equations with an unknown $\mathbf{X}$ are rewritten as follows. The first equation is the vector formula of the projection for the right image:

$$Z_r \mathbf{x}_r = {}^t\tilde{R}_r(\mathbf{X} - \mathbf{O}_r), \tag{1}$$

where $Z_r =^t \mathbf{r}_3(\mathbf{X} - \mathbf{O}_r)$, ${}^t\tilde{R}_r$ is the first and second rows in ${}^tR_r$ and ${}^t\mathbf{r}_3$ is the third row in ${}^tR_r$, i.e., $\binom{{}^t\tilde{R}_r}{{}^t\mathbf{r}_3} =^t R_r$. The second equation is the projection for the left image:

$$Z_l x_l = {}^t\mathbf{l}_1(\mathbf{X} - \mathbf{O}_l), \tag{2}$$

$$Z_l y_l = {}^t\mathbf{l}_2(\mathbf{X} - \mathbf{O}_l), \tag{3}$$

where $Z_l = {}^t\mathbf{l}_3(\mathbf{X} - \mathbf{O}_l)$. Here ${}^t\mathbf{l}_1$, ${}^t\mathbf{l}_2$ and ${}^t\mathbf{l}_3$ be three row vectors in ${}^tR_l$, i.e., $\begin{pmatrix} {}^t\mathbf{l}_1 \\ {}^t\mathbf{l}_2 \\ {}^t\mathbf{l}_3 \end{pmatrix} = {}^tR_l$. Then we have the following theorem, which is about (a) the formula determining the object point $\mathbf{X}$ from the image pairs and (b) the Longuet-Higgin's relation.

Theorem 3.1. *Denote the determinant of a matrix M by $\mid M \mid$ and the cofactor matrix by M^c.*

- *(a) The object point $\mathbf{X}$ is determined by $\mathbf{X} = B^{-1}\mathbf{d}$ where $B = \binom{\mathbf{x}_r{}^t\mathbf{r}_3 - {}^t\tilde{R}_r}{{}^t\mathbf{l}_1 - {}^t\mathbf{l}_3 x_l}$ and $\mathbf{d} = \binom{-{}^t\tilde{R}_r\mathbf{O}_r + {}^t\mathbf{r}_3\mathbf{O}_r\mathbf{x}_r}{({}^t\mathbf{l}_1 - {}^t\mathbf{l}_3 x_l)\mathbf{O}_l}$, by using (1) and (2).*
- *(b) The Longuet-Higgin's relation between the image pairs $\mathbf{x}_l$ and $\mathbf{x}_r$, is obtained from (3):*

$$y_l = \frac{a_0(\mathbf{x}_r)x_l + a_1(\mathbf{x}_r)}{a_2(\mathbf{x}_r)x_l + a_3(\mathbf{x}_r)}.$$

Coefficients $a_i(\mathbf{x}_r), (i = 0, 1, 2, 3)$ are

$$a_0(\mathbf{x}_r) = {}^t\mathbf{l}_2(\mathbf{b}_2 - \mid B_2 \mid \mathbf{O}_l),$$
$$a_1(\mathbf{x}_r) = {}^t\mathbf{l}_2(\mathbf{b}_1 - \mid B_1 \mid \mathbf{O}_l),$$
$$a_2(\mathbf{x}_r) = {}^t\mathbf{l}_3(\mathbf{b}_2 - \mid B_2 \mid \mathbf{O}_l),$$
$$a_3(\mathbf{x}_r) = {}^t\mathbf{l}_3(\mathbf{b}_1 - \mid B_1 \mid \mathbf{O}_l),$$

where

$$\mathbf{b}_1 = B_1^c(\mathbf{d}_1 + \mathbf{d}_2), \qquad \mathbf{b}_2 = B_1^c\mathbf{d}_3 + B_2^c\mathbf{d}_1,$$

$$B_1 = \begin{pmatrix} \mathbf{x}_r{}^t\mathbf{r}_3 - {}^t\tilde{R}_r \\ {}^t\mathbf{l}_1 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} \mathbf{x}_r{}^t\mathbf{r}_3 - {}^t\tilde{R}_r \\ -{}^t\mathbf{l}_3 \end{pmatrix},$$

$$\mathbf{d}_1 = \begin{pmatrix} -{}^t\tilde{R}_r\mathbf{O}_r + {}^t\mathbf{r}_3\mathbf{O}_r\mathbf{x}_r \\ 0 \end{pmatrix}, \qquad \mathbf{d}_2 = \begin{pmatrix} \mathbf{0} \\ {}^t\mathbf{l}_1\mathbf{O}_l \end{pmatrix}, \qquad \mathbf{d}_3 = -\begin{pmatrix} \mathbf{0} \\ {}^t\mathbf{l}_3\mathbf{O}_l \end{pmatrix}.$$

Proof. (a) Substituting $Z_r = {}^t\mathbf{r}_3(\mathbf{X} - \mathbf{O}_r)$ into (1), we have ${}^t\mathbf{r}_3(\mathbf{X} - \mathbf{O}_r)\mathbf{x}_r = {}^t\tilde{R}_r(\mathbf{X} - \mathbf{O}_r)$. Again by substituting $Z_l = {}^t\mathbf{l}_3(\mathbf{X} - \mathbf{O}_l)$ into (2), we obtain ${}^t\mathbf{l}_3(\mathbf{X} - \mathbf{O}_l)x_l = {}^t\mathbf{l}_1(\mathbf{X} - \mathbf{O}_l)$.
(b) From Eq.(3), we have ${}^t\mathbf{l}_3(\mathbf{X} - \mathbf{O}_l)y_l = {}^t\mathbf{l}_2(\mathbf{X} - \mathbf{O}_l)$. Therefore, we have the relation $({}^t\mathbf{l}_2 - y_l{}^t\mathbf{l}_3)B^{-1}\mathbf{d} = -{}^t\mathbf{l}_3\mathbf{O}_l y_l + {}^t\mathbf{l}_2\mathbf{O}_l$. That is,

$$({}^t\mathbf{l}_2 - y_l{}^t\mathbf{l}_3)B^c\mathbf{d} + \mid B \mid ({}^t\mathbf{l}_3\mathbf{O}_l y_l - {}^t\mathbf{l}_2\mathbf{O}_l) = 0.$$

Since $\mathbf{d} = \mathbf{d}_1 + \mathbf{d}_2 + \mathbf{d}_3 x_l$, we have $B^c\mathbf{d} = \mathbf{b}_1 + \mathbf{b}_2 x_l$.

The fact that $\mid B \mid = \mid B_1 \mid + \mid B_2 \mid x_l$ and $B^c = B_1^c + B_2^c \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} x_l$, completes the proof.

q.e.d.

Remark In [1], Theorem 3.1 is shown in case that one of cameras is translated, that is, ${}^t\mathbf{O}_r = (0, 0, 0)$, and $R_r = R_l = E$, where E is the unit matrix and $\mathbf{O}_l$ is any vector. So Theorem 3.1 is an extended result of [1] since we assume any camera direction and any camera location.

4. Initial procedure

The images of the pupil centers are given by $\mathbf{x}_r = \mathbf{0}$ and $\mathbf{x}_l = \mathbf{0}$. By putting $\mathbf{x}_r = \mathbf{0}$ and $\mathbf{x}_l = \mathbf{0}$ in Theorem 3.1 (a), the eye direction $\mathbf{u}_3$ is obtained by the formula depending on camera angles: $\mathbf{u}_3 = \frac{B_0^{-1}\mathbf{d_0}}{\|B_0^{-1}\mathbf{d_0}\|}$, where $B_0 = \begin{pmatrix} -{}^t\tilde{R}_r \\ {}^t\mathbf{l}_1 \end{pmatrix}$, $\quad \mathbf{d_0} = \begin{pmatrix} -{}^t\tilde{R}_r\mathbf{O}_r \\ {}^t\mathbf{l}_1\mathbf{O}_l \end{pmatrix}$. This is proved from the fact that we have

$${}^t\tilde{R}_r(\mathbf{p} - \mathbf{O}_r) = \mathbf{0}, \qquad {}^t\mathbf{l}_1(\mathbf{p} - \mathbf{O}_l) = 0,$$

if $\mathbf{x}_r = \mathbf{0}$ and $\mathbf{x}_l = \mathbf{0}$, and that $\mathbf{p} = \sqrt{r_0^2 - a_0^2}\mathbf{u}_3$ is the unique solution of the above equations.

The initial procedure is to calculate the eye direction $\mathbf{u}_3$ by $\frac{B_0^{-1}\mathbf{d_0}}{\|B_0^{-1}\mathbf{d_0}\|}$.

If the camera directs to the center of the pupil exactly, the correct eye direction is calculated by the initial procedure. However, since the cameras are controlled to direct to the center of the pupil by human hands, not knowing the exact location of the pupil center, the camera only directs in the neighborhood of the center of the pupil. So the outward normal $\mathbf{u}_3$ at the wrong point of $\mathbf{p}$ is calculated by the initial procedure. In the next procedure, the camera angle sensor is adjusted to direct to the correct $\mathbf{p}$, by putting $\mathbf{p} = \sqrt{r_0^2 - a_0^2}\mathbf{u}_3$ and by calculating correct β and γ, using the epipolar radius procedure in the following sections.

5. Image of the pupil circle

Substituting $\mathbf{X} = (\mathbf{u}_1, \mathbf{u}_2)\mathbf{v} + \mathbf{p}$ into $\mathbf{X}_r = {}^tR_r(\mathbf{X} - \mathbf{O}_r)$, we have

$$\mathbf{X}_r = \begin{pmatrix} {}^t\mathbf{r}_1(\mathbf{u}_1, \mathbf{u}_2)\mathbf{v} \\ {}^t\mathbf{r}_2(\mathbf{u}_1, \mathbf{u}_2)\mathbf{v} \\ {}^t\mathbf{r}_3(\mathbf{u}_1, \mathbf{u}_2)\mathbf{v} + {}^t\mathbf{r}_3(\mathbf{p} - \mathbf{O}_r) \end{pmatrix} = \begin{pmatrix} A_r\mathbf{v} \\ {}^t\mathbf{a}_r\mathbf{v} + \zeta_r \end{pmatrix}.$$

The transformation from the pupil circle to the right image is expressed as $\mathbf{x}_r = \frac{A_r\mathbf{v}}{{}^t\mathbf{a}_r\mathbf{v}+\zeta_r}$, where $\mathbf{v} = a_0\binom{\cos\psi}{\sin\psi}$, $A_r = \binom{{}^t\mathbf{r}_1(\mathbf{u}_1\mathbf{u}_2)}{{}^t\mathbf{r}_2(\mathbf{u}_1\mathbf{u}_2)}$, ${}^t\mathbf{a}_r = {}^t\mathbf{r}_3(\mathbf{u}_1\mathbf{u}_2)$, $\zeta_r = {}^t\mathbf{r}_3(\mathbf{p} - \mathbf{O}_r) = \|\mathbf{p} - \mathbf{O}_r\|$. Similarly, the transformation to the left image is expressed as $\mathbf{x}_l = \frac{A_l\mathbf{v}}{{}^t\mathbf{a}_l\mathbf{v}+\zeta_l}$, where $A_l = \binom{{}^t\mathbf{l}_1(\mathbf{u}_1\mathbf{u}_2)}{{}^t\mathbf{l}_2(\mathbf{u}_1\mathbf{u}_2)}$, ${}^t\mathbf{a}_l = {}^t\mathbf{r}_3(\mathbf{u}_1\mathbf{u}_2)$, $\zeta_l = \|\mathbf{p} - \mathbf{O}_l\|$.

5.1. *Center of image ellipse*

The center in the image ellipse of the pupil circle is obtained in the following theorem for the perspective projection. By the perspective projection, the image of the center $\sqrt{r_0^2 - a_0^2}\mathbf{u}_3$ is not the center of the image ellipse as shown in the following theorem.

Theorem 5.1.
Let $\Phi = \frac{\zeta_r^2}{a_0^2}\,{}^tA_r^{-1}A_r^{-1} - {}^tA_r^{-1}\mathbf{a}_r\,{}^t\mathbf{a}_rA_r^{-1}$. *Assume that* $|\,\Phi\,| > 0$ *and* $\phi_{11}(1+c) > 0$, *where* ϕ_{11} *is the* $(1,1)$ *element in* Φ *and* $c = \frac{a_0^2}{\zeta^2 - a_0^2}$. *Then the image of the circle is the ellipse:*

$${}^t(\mathbf{x} - \Psi)\Phi(\mathbf{x} - \Psi) = 1 + c.$$

The ellipse center is $\Psi = \frac{-A_r\mathbf{a}_r}{(\zeta_r/a_0)^2 - 1}$. *We have* $\Psi \neq \mathbf{0}$, *except that both* $\binom{{}^t\mathbf{u}_1}{{}^t\mathbf{u}_2}\mathbf{r}_1$ *and* $\binom{{}^t\mathbf{u}_1}{{}^t\mathbf{u}_2}\mathbf{r}_2$ *are perpendicular to* $\binom{{}^t\mathbf{u}_1}{{}^t\mathbf{u}_2}\mathbf{r}_3$.

Proof. The inverse mapping of $\frac{A_r\mathbf{v}}{{}^t\mathbf{a}_r\mathbf{v}+\zeta_r}$ is $\mathbf{v} = \frac{\zeta_rA_r^{-1}\mathbf{x}}{1-{}^t\mathbf{a}_rA_r^{-1}\mathbf{x}}$, since $\mathbf{x}({}^t\mathbf{a}_r\mathbf{v} + \zeta_r) - A_r\mathbf{v} = 0$ and $(\mathbf{x}{}^t\mathbf{a}_rA_r^{-1} - E)A_r\mathbf{v} = -\zeta_r\mathbf{x}$. Put ${}^t\mathbf{a}_A = {}^t\mathbf{a}_rA_r^{-1}$. Then

$$(\mathbf{x}\mathbf{a}_A - E)^{-1}\mathbf{x} = -\mathbf{x}/(1 - {}^t\mathbf{a}_A\mathbf{x}),$$

since $|\,\mathbf{x}{}^t\mathbf{a}_A - E\,| = 1 - {}^t\mathbf{a}_A\mathbf{x}$ and $(\mathbf{x}{}^t\mathbf{a}_A - E)^c\mathbf{x} = -\mathbf{x}$. Here E is the 2×2 unit matrix. q.e.d.

6. Epipolar plane and epipolar radius

Epipolar geometry is defined as follows in [2].

Definition 6.1. [2] A plane generated from the three points which consist of the two camera locations and the object point $\mathbf{X}$, is called an 'epipolar plane'. An intersection line of the epipolar plane and the image plane by the camera is called an 'epipolar line'. The epipolar lines for several objects intersect at one point. This point is called an 'epipole'. Epipolar lines and epipoles constitute a fan-shaped figure. This particular geometry such as the fan-shaped figure is called an 'epipolar geometry'.

The epipole in the image by the right (left) camera is the image of the left (right) camera location. This is denoted as $\mathbf{e}_r$ (or $\mathbf{e}_l$) [2].

Applying this definition to the object point $\mathbf{X}$, which is the center of the pupil circle, a special epipolar radius is defined as follows, which lies on the special epipolar line passing through the image of the center.

Consider an epipolar plane, which is constructed from the three points, i.e., two camera locations $\mathbf{O}_r$,$\mathbf{O}_l$ and the center $\mathbf{p}$ of the pupil circle: a point in the epipolar plane is expressed as $\mathbf{Q} = \alpha_1(\mathbf{O}_l - \mathbf{p}) + \alpha_2(\mathbf{O}_r - \mathbf{p}) + \mathbf{p}$, where α_1 and α_2 take real scalar values.

Definition 6.2. 'An epipolar radius' is defined as a segment $[\mathbf{f}_r, \mathbf{p}_r]$ (or $[\mathbf{f}_l, \mathbf{p}_l]$) projected from the specified radius $[\mathbf{f}, \mathbf{p}]$ on the intersection of the epipolar plane and the pupil disc.

Then the epipolar radius lies on the epipolar line which connects the image of the center $\mathbf{p}_r$ (or $\mathbf{p}_l$) of the pupil circle and the epipole $\mathbf{e}_r$ (or $\mathbf{e}_l$).

If a point on the object pupil circle is not on the epipole radius and lies on the epipolar plane, then the image of the pupil circle concentrates into the epipole line and the 'epipolar radius procedure' does not work.

The epipolar $\mathbf{e}_r$ (or $\mathbf{e}_l$) is obtained in the following lemma. The intersection line of the epipolar plane and the plane of the pupil disc is also given in the following lemma. Proof is performed by *Cramér's* rule and is omitted here.

Lemma 6.1.

(1) *The epipoles are determined by* $\mathbf{e}_r = \binom{r_1^1}{r_2^1}/r_3^1$, $\mathbf{e}_l = \binom{l_1^1}{l_2^1}/l_3^1$, *where* r_i^1 *is a first element in* $\mathbf{r}_i$ *and* l_i^1 *is a first element in* $\mathbf{l}_i$ *for* $i = 1, 2, 3$. *These definitions depend only the camera angles.*

(2) *Let* $\mathbf{Q}$ *be an epipolar plane* $\mathbf{Q} = \alpha(\mathbf{O}_l - \mathbf{p}) + \beta(\mathbf{O}_r - \mathbf{p}) + \mathbf{p}$, *where* $\mathbf{p}$ *is the center of the pupil circle. Let* $v_1\mathbf{u}_1 + v_2\mathbf{u}_2 + \mathbf{p}$ *be the pupil circle on the eye sphere at* $\mathbf{p}$, *where* $\mathbf{p} = \sqrt{r_0^2 - a_0^2}\mathbf{u}_3$. *Here* $\mathbf{u}_1$ *and* $\mathbf{u}_2$ *are tangent vectors at* $r_0\mathbf{u}_3$ *on the sphere,* $v_1 = a_0 \cos\psi$ *and* $v_2 = a_0 \sin\psi$. *Then the specified radius* $[\mathbf{f}, \mathbf{p}]$ *on the intersection of the epipolar plane and the pupil disc is given by the angle* ψ *of* $\mathbf{v} =^t (v_1, v_2)$ *on the pupil circle, i.e.,*

$$\psi = \arctan(-\frac{\sin\gamma(\sqrt{r_0^2 - a_0^2} - \cos\beta O_Z)}{\cos\gamma(\sqrt{r_0^2 - a_0^2}\cos\beta - O_Z)}).$$

7. Three pairs of image points $\mathbf{x}_r$ and $\mathbf{x}_l$

7.1. *Independent selection of object points $\mathbf{X}_i$ on pupil circle*

In the following theorem, it is shown that the eye direction is obtained from the three points on the pupil circle, if these points on the pupil circle are independent. The three points are calculated from three pairs of image points $\mathbf{x}_r$ and $\mathbf{x}_l$.

Theorem 7.1.
The following conditions are assumed.

(1) The three pairs of image points $(\mathbf{x}_r,\mathbf{x}_l)_i$, $(i = 0, 1, 2)$ are known, where each pair is constituted of the right image point $\mathbf{x}_r$ and the left image point $\mathbf{x}_l$.
(2) The angles and the locations of the cameras are known.
(3) Define a point $\mathbf{X}_i$ on the pupil circle by substituting $(\mathbf{x}_r,\mathbf{x}_l)_i$ into $\mathbf{X}_i = B^{-1}\mathbf{d}$. $\mathbf{X}_i$ are arranged in counterclockwise for $i = 0, 1, 2$ and $\mathbf{X}_i$ are independent.

Then the exterior product is the same direction with the eye direction $\mathbf{u}_3$: $\mathbf{u}_3 = \frac{\mathbf{X}_2-\mathbf{X}_1}{\|\mathbf{X}_2-\mathbf{X}_1\|} \times \frac{\mathbf{X}_0-\mathbf{X}_1}{\|\mathbf{X}_0-\mathbf{X}_1\|}$.

Proof. Put $\mathbf{v}_i = a_0\binom{\cos\psi_i}{\sin\psi_i}$ and put $\mathbf{X}_i = (\mathbf{u}_1, \mathbf{u}_2)\mathbf{v}_i + \mathbf{p}$. Since $\mathbf{u}_3 = \mathbf{u}_1 \times \mathbf{u}_2$ and $\mathbf{X}_i$ are arranged in counterclockwise for $i = 0, 1, 2$, it holds that ψ_i increases and that $(\mathbf{X}_2 - \mathbf{X}_1) \times (\mathbf{X}_0 - \mathbf{X}_1) = 4a_0^2 S\mathbf{u}_3$, where $S = \sin\frac{\psi_2-\psi_1}{2}\sin\frac{\psi_1-\psi_0}{2}\sin\frac{\psi_2-\psi_0}{2}$.
q.e.d.

7.2. *Optimal selection of $\mathbf{X}_i$*

The pattern matching method to recognize the planer curve with characteristic points has been introduced in [2]. A characteristic point is a bitangent point, a inflection point, or so on. A bitangent point is a point where its tangent line is tangent to the curve at more than two points. That is, if two points on the curve has the same tangent line, then we call each point a bitangent point. A point where its planer curve's curvature is 0, is called a inflection point. If there exist more than five characteristic points on the planer curve, then the projective invariant value of the cross ratio is calculated and the pattern of the planer curve is recognized by the projective invariant value. However there exit no characteristic points on the pupil circle. Therefore the idea of the cross ratio might not be applicable.

In this paper, instead of these characteristic points, we take independent three points $\mathbf{X}^{(i)}, i = 0, 1, 2$ on the object pupil circle. In the proposed algorithm, we call three points $\mathbf{X}^{(i)}$ optimal if their magnitude of the exterior product $\|(\mathbf{X}^{(2)} - \mathbf{X}^{(1)}) \times (\mathbf{X}^{(0)} - \mathbf{X}^{(1)})\|$ takes the maximum value. The maximum value is determined by only the radius of the pupil circle. The fact that the maximum value is attained means that the calculated center of the pupil is correct. So the cameras direct to the center of the pupil and the eye direction is obtained accurately.

Corollary 7.1. *The norm* $\|(\mathbf{X}_2 - \mathbf{X}_1) \times (\mathbf{X}_0 - \mathbf{X}_1)\|$ *takes the maximum value* $\frac{3\sqrt{3}}{2}a_0^2$ *where the three points are corresponding to* $\psi_i = \psi_0 + \frac{2\pi i}{3}$.

Proof. Consider the angles such that $\frac{\partial S}{\partial \psi_1} = 0$ and $\frac{\partial S}{\partial \psi_2} = 0$. Then $3(\psi_1 - \psi_0) = 2\pi$ by using $\sin\frac{\psi_2 - 2\psi_1 + \psi_0}{2} = 0$ and $\sin\frac{2\psi_2 - \psi_1 - \psi_0}{2} = 0$. q.e.d.

8. Epipolar radius procedure

In this section, we propose the epipolar radius procedure when the images of the cameras are pixel. Assume that the distance $2d$ of two cameras, the distance O_z between the midpoint of two camera locations and the center of the eye sphere, the radius r_0 of the eye sphere and the radius a_0 of the pupil circle are known. Let the right and left pixel image points (say 100 discrete points in each) be given on the right and the left contour image of the pupil circle. Let $\mathbf{y}_r$ denote the variable of the right pixel image points and $\mathbf{y}_l$ the variable of the left pixel image points.

The eye direction $\mathbf{u}_3$ is calculated by the following epipolar radius procedure for the pixel image points after calculating (β, γ) by the initial procedure.

(1) *Select the optimal points* $\mathbf{f}_{ri}$ *and* $\mathbf{f}_{li}$. — The angle of the right epipolar radius is given by $\psi_0 = \arctan(\frac{\sin\gamma(\sqrt{r_0^2 - a_0^2} - \cos\beta O_Z)}{-\cos\gamma(\sqrt{r_0^2 - a_0^2}\cos\beta - O_Z)})$ for β and γ. Calculate the right and the left image points $\mathbf{f}_r^{(i)} = \frac{A_r \mathbf{v}_i}{{}^t\mathbf{a}_r \mathbf{v}_i + \zeta_r}$ and $\mathbf{f}_l^{(i)} = \frac{A_l \mathbf{v}_i}{{}^t\mathbf{a}_l \mathbf{v}_i + \zeta_l}$, by using the points $\mathbf{v}_i = a_0\binom{\cos(\psi_0 + \frac{2\pi i}{3})}{\sin(\psi_0 + \frac{2\pi i}{3})}$ for $i = 0, 1, 2$.

(2) *Select the practical image points* $\mathbf{y}_{ri}$ *nearest to the images of the optimal points* $\mathbf{f}_r^{(i)}$. — Let $\mathbf{y}_{ri}$ be the point to minimize $\mid \mathbf{y}_r - \mathbf{f}_r^{(i)} \mid$ in the contour of the image ellipse for $i = 0, 1, 2$. Here $\mid \mathbf{x} \mid$ denotes the sum of the absolute values of the elements in $\mathbf{x} = \binom{x}{y}$: $\mid \mathbf{x} \mid = \mid x \mid + \mid y \mid$.

(3) *Select the practical left image points* $\mathbf{y}_{li}$. — If the camera directs not to the center $\mathbf{p}$, then the epipolar radius $\mathbf{v}_0$ has the error depending on (β, γ). The left image points are assigned by the Longuet Higgin's relation, which has no error depending on (β, γ). This relation for each left image point may give two solutions. In order to select one of these, one point is selected, which is near to each left image point $\mathbf{f}_l^i$. Calculate the point $\mathbf{y}_{(li)}$ to minimize $\mid y_l - \frac{a_0(\mathbf{y}_{ri})x_l + a_1(\mathbf{y}_{ri})}{a_2(\mathbf{y}_{ri})x_l + a_3(\mathbf{y}_{ri})} \mid + \|\mathbf{y}_l - \mathbf{f}_l^{(i)}\|$ for $i = 0, 1, 2$ for the right image points $\mathbf{y}_{ri}$.

(4) *Three points* $\mathbf{X}_i$ *on pupil circle.* — The three points $\mathbf{X}_i$ on the object pupil circle are calculated from the three pairs of the image points $(\mathbf{y}_{ri}, \mathbf{y}_{li})$ by substituting $(\mathbf{x}_r, \mathbf{x}_l) = (\mathbf{y}_{ri}, \mathbf{y}_{li})$ into $\mathbf{X}_i = B^{-1}\mathbf{d}$, where $B = \binom{\mathbf{x}_r{}^t\mathbf{r}_3 - {}^t\tilde{R}_r}{{}^t\mathbf{l}_1 - {}^t\mathbf{l}_3 x_l}$, $\mathbf{d} = \binom{-{}^t\tilde{R}_r\mathbf{O}_r + {}^t\mathbf{r}_3\mathbf{O}_r\mathbf{x}_r}{({}^t\mathbf{l}_1 - {}^t\mathbf{l}_3 x_l)\mathbf{O}_l}$.

(5) *Eye direction* $\mathbf{u}_3$. — Calculate the exterior product $\mathbf{X} = (\mathbf{X}_2 - \mathbf{X}_1) \times (\mathbf{X}_0 - \mathbf{X}_1)$. If $\mid \frac{3\sqrt{3}}{2}a_0^2 - \|\mathbf{X}\| \mid < \epsilon$ then stop the procedure and if $\mid \frac{3\sqrt{3}}{2}a_0^2 - \|\mathbf{X}\| \mid > \epsilon$ then calculate $\mathbf{u}_3 = \frac{\mathbf{X}_2 - \mathbf{X}_1}{\|\mathbf{X}_2 - \mathbf{X}_1\|} \times \frac{\mathbf{X}_0 - \mathbf{X}_1}{\|\mathbf{X}_0 - \mathbf{X}_1\|}$. Repeat the initial procedure in the section 4 and the epipolar radius procedure in this section, by putting

$\mathbf{p} = \sqrt{r_0^2 - a_0^2}\mathbf{X}$. Here ϵ is a small value depending on the accuracy of the precision of the computer programming, the pixel accuracy in the camera image and the accuracy of the camera angles measured by the camera angle sensors.

References

1. T. Moons, L. Van Gool, M.Van Diest and E.J. Pauwels, *In J.L. Mundy, A. Zisserman, and D.A. Forsyth, editors, Applications of Invariance in Computer Vision, Springer-Verlag,* **LNCS 825**, 297-316, (1994).
2. J. Sato, *Computervision-Geometry of Vision, corona publishing,* 172 pages (1999) (Japanese).
3. T. Takegami, T. Gotoh and G. Ohyama, *IEICE D-II* **Vol.J84-D-II No. 8** , 1580, (2001) (Japanese).
4. A. Sera and Y. Sengoku, *J. of Japanese association of occupational therapists,* **VOL 21**, 307 (2002) (Japanese).
5. K. Akiyama, T. Gotoh and H. Eguchi, *OPHTHALMOLOGY, kaibashobo co.*, 302pages (2004) (Japanese).