ELSEVIER

Journal of Neuroscience Methods 159 (2007) 158–169

JOURNAL OF NEUROSCIENCE METHODS

www.elsevier.com/locate/jneumeth

Computation of gaze orientation under unrestrained head movements

Renaud Ronsse [a,*], Olivier White [b,c], Philippe Lefèvre [c,d]

[a] *Department of Electrical Engineering and Computer Science (Montefiore Institute), Université de Liège, Grande Traverse 10 (B28), B-4000 Liège, Belgium*
[b] *Unité de Réadaptation et de médecine physique, Université catholique de Louvain, Avenue Mounier 53, B-1200 Bruxelles, Belgium*
[c] *Centre for Systems Engineering and Applied Mechanics (CESAME), Université catholique de Louvain, Avenue Georges Lemaître 4, B-1348 Louvain-la-Neuve, Belgium*
[d] *Laboratory of Neurophysiology, Université catholique de Louvain, Avenue Hippocrate 54, B-1200 Bruxelles, Belgium*

Received 12 March 2006; received in revised form 7 June 2006; accepted 22 June 2006

Abstract

Given the high relevance of visual input to human behavior, it is often important to precisely monitor the spatial orientation of the visual axis. One popular and accurate technique for measuring gaze orientation is based on the dual search coil. This technique does not allow for very large displacements of the subject, however, and is not robust with respect to translations of the head. More recently, less invasive procedures have been developed that record eye movements with camera-based systems attached to a helmet worn by the subject. Computational algorithms have also been developed that can calibrate eye orientation when the head's position is fixed. Given that camera-based systems measure the eye's position in its orbit, however, the reconstruction of gaze orientation is not as straightforward when the head is allowed to move. In this paper, we propose a new algorithm and calibration method to compute gaze orientation under unrestrained head conditions. Our method requires only the accurate measurement of orbital eye position (for instance, with a camera-based system), and the position of three points on the head. The calculations are expressed in terms of linear algebra, so can easily be interpreted and related to the geometry of the human body. Our calibration method has been tested experimentally and validated against independent data, proving that is it robust even under large translations, rotations, and torsions of the head.

Keywords: Eye–head coordination; Gaze orientation; Gaze measurement; Calibration; Unrestrained head

1. Introduction

The accurate measurement of eye movements is crucial to oculomotor research. These movements are commonly expressed in terms of their horizontal, vertical, and torsional components, assuming a ball-in-socket model for the eye with 3 degrees of freedom (DOF). There is an enormous body of literature investigating eye movements under the head fixed condition, and describing accurate methods of measuring these rotations. Such methods are mainly based on either the *dual search coil technique* (Robinson, 1963; Collewijn et al., 1985) or *video image processing* devices (e.g., Nakayama, 1974; Haslwanter, 1995; Moore et al., 1996). Both coil- and video-based techniques are widely used by the oculomotor community in behavioral and clinical studies (e.g., Orban de Xivry et al., 2006; Yuksel et al., 2005).

The dual search coil technique is based on the measurement of electric fields induced in a coil placed directly on the subject's eye. The intensity of the electric field in the coil depends on the orientation of the coil (i.e., of the eye) with respect to an alternating magnetic field. Video image processing devices are designed to measure eye movements with camera-based systems. These techniques are more difficult to apply, however, when gaze orientation arises from a combination of head and eye movements. Video-based systems are usually based on processing images from a camera fixed to the head. It follows that they do not capture any head movement, but only the eye-in-head component of gaze orientation. The coil technique captures the true eye-in-space orientation, on the other hand, but cannot cope with *translations* of the head since the magnetic field in the recording chamber is uniform. In addition, *rotations* of the head always induce a translation of the eye center that is not captured by search coil algorithms. This induces a non-linear bias in the computed signals.

In this study, we measure the eye-in-head and head-in-space positions independently. Our technique is therefore valid for

* Corresponding author. Tel.: +32 4 366 26 97; fax: +32 4 366 29 89.
E-mail address: R.Ronsse@ulg.ac.be (R. Ronsse).

0165-0270/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2006.06.016

very large head movements, or even displacements of the body. Using the coil technique with an unrestrained head would require the measurement of head position, and the integration of this information using an algorithm similar to that presented in this paper. This study provides a robust geometrical basis for computing the gaze orientation with no restrictions on head movement.

The mathematical developments of this paper are based on common linear algebra operations. Head rotations are represented as a 3×3 matrix, according to the well-known sequence of Fick angles (Fick, 1874; Haslwanter, 1995), and head translations as a 3×1 position vector. The combination of head position and orientation, hereafter referred to as the head *pose*, therefore has 6 DOF. The eye-in-head orientation is similarly represented by a 3×3 matrix with 3 DOF, since the center of the eye is assumed to be fixed with respect to the head. Similar formalisms have been used by Allison et al. (1996) in testing the vestibular system, and by Wang and Sung (2002) to estimate gaze orientation in front of a computer screen. These translations and rotations could be represented in other ways, e.g. dual-number quaternions. There are even some papers dedicated to comparing these methods, originally from the perspective of robot kinematics (e.g., the survey by Aspragathos and Dimitros, 1998) and later with respect to the computation of eye rotations (Tweed et al., 1990; Haslwanter, 1995). Note that for the sake of simplicity this paper does not take into account eye-in-head *torsion*, which corresponds to the third Fick angle and captures eye rotation around the optical axis. This angle can be measured by both search coil (e.g. Tweed et al., 1990) and video-based devices (e.g. Moore et al., 1996), however, and can easily be integrated into the eye-in-head orientation matrix. Eye torsion does not change the line of sight.

This paper also addresses the issue of calibration. We present an efficient calibration protocol based on gaze fixation during self-paced, smooth head movements. This protocol can be easily adapted to a broad range of environments, since it only requires knowing the location of the fixation target in a 3D, ground-based coordinate system.

We will particularly stress the algorithmic sequence followed in computing gaze orientation from the head-in-space and eye-in-head components, via translations and rotations that are simply described using linear algebra. The method is therefore simple, robust, and computationally efficient. Its main hardware requirement is a device that can measure the position of three points on the subject's head in a ground-based coordinate system. This paper also describes experimental results validating our algorithm, obtained by using a video-based device to measure the eye-in-head position. The method essentially consists of two steps:

(1) determining the eye orientation in the head coordinate system;
(2) rotating the eye orientation vector into a ground-based coordinate system, using information on the head orientation provided by the head measurement device.

The rest of this paper is organized as follows. Section 2.1 describes the geometrical relationships required to compute the eye-in-head and head components of gaze, and describes their mutual interaction. Section 2.2 discusses the calibration required to integrate measurements from both acquisition devices. Section 2.3 describes the validation tasks. The results of validation are summarized in Section 3, and the characteristics of this new method are discussed in Section 4.

2. Materials and methods

2.1. Geometrical developments

A geometric basis for measurement of eye position under the head fixed condition have been derived by Moore et al. (1996). The center of the eye is assumed to be fixed with respect to the head throughout the task. This section explains how Moore's method can be extended to unrestrained head conditions, when the subject's head is free to move without constraint. This technique provides *separate* signals for the head and eye orientations, and therefore distinguishes the relative contributions of each component to gaze orientation.

Gaze is treated as a vector in space, with its origin at the eye center and its direction following from both the head's orientation in space and the eye's orientation in the head. According to this method, gaze orientation has to be derived independently for each eye. This paper therefore focuses on derivations involving one eye, except for a short section dedicated to the calculation of vergence (Section 2.1.4). In the following discussion, matrices are represented by bold uppercase characters (e.g. $\mathbf{R}$), vectors and points in 3D space by normal uppercase characters (e.g. P), and scalar quantities by lowercase italic characters (e.g. x).

2.1.1. Eye position through image processing

According to Moore et al. (1996), the vertical and horizontal Fick angles of eye-in-head orientation as measured by a video-

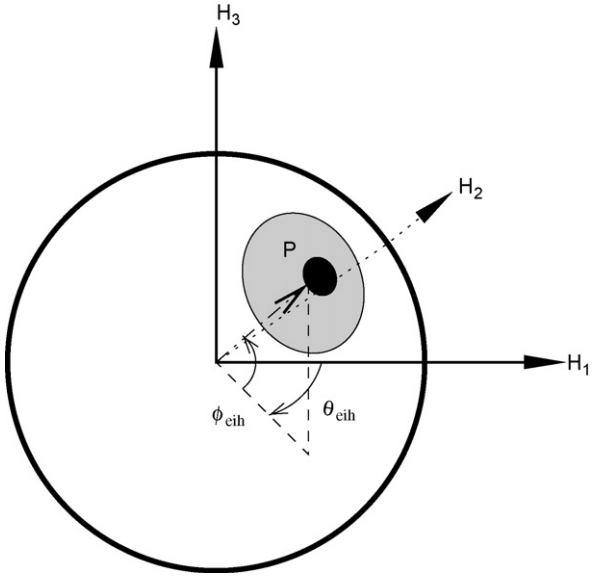

Fig. 1. Diagram of the eye's orientation in the head coordinate system [H_1, H_2, H_3]. P denotes the pupil center, and (θ_{eih}, ϕ_{eih}) are the horizontal and vertical Fick angles of the eye's orientation.

based device are given by

$$\phi_{\text{eih}} = \arcsin\left(-a'_{21}x - a'_{22}y - a'_{23}\right) \tag{1}$$

$$\theta_{\text{eih}} = \arcsin\left(\frac{a'_{11}x + a'_{12}y + a'_{13}}{\cos\left(\phi_{\text{eih}}\right)}\right), \tag{2}$$

where the coefficients ('gains') a'_{ij} are determined by calibration (see Section 2.2) and (x, y) are the coordinates of the pupil's center in the camera image. Both angles are depicted in Fig. 1.

The main geometrical developments of Moore et al. (1996) are summarized in Appendix A. The eye orientation *vector* in the head coordinate system follows directly from (A.6).

2.1.2. Head pose through image processing

Head pose is defined in terms of a *ground-based* (i.e., motionless with respect to the laboratory) coordinate system $[\mathsf{G}_1, \mathsf{G}_2, \mathsf{G}_3]$ (see Fig. 2). To efficiently compute the head pose one must measure the position of three points on the head, which must not be collinear. Let us denote these points by T_a, T_b and T_c. They are represented by grey dots in Figs. 2 and A.1, and define a plane parallel to the frontal plane H_2–H_3. Since the head is assumed to be a rigid body, the position of these points completely determines the head pose. It is of particular interest to determine the position of the eye center, i.e., the origin of the gaze. We assume that the position of this point can be deduced from the positions of $\mathsf{T}_{\{a,b,c\}}$ and prior knowledge of the head's anthropomorphic characteristics. The position of the eye center, $\mathsf{E} = (e_1, e_2, e_3)^{\mathrm{T}}$, is taken as the origin of the $[\mathsf{H}_1, \mathsf{H}_2, \mathsf{H}_3]$ coordinate system as depicted in Fig. 2.

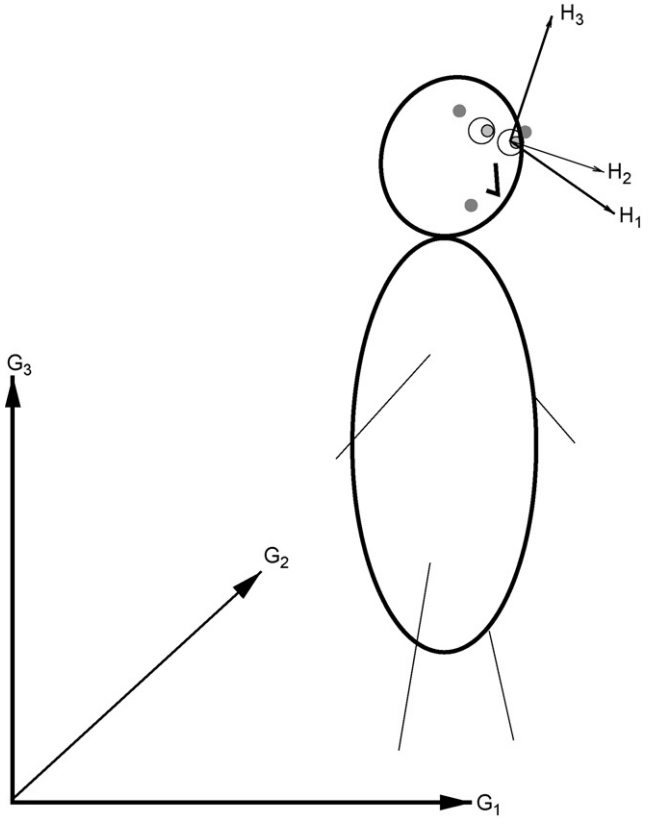

Fig. 2. Diagram of the whole body. This figure emphasizes the components of the gaze orientation: the head component is due to the head pose in the ground-based coordinate system $[\mathsf{G}_1, \mathsf{G}_2, \mathsf{G}_3]$, and the eye-in-head component is due to the eye's orientation in the head coordinate system $[\mathsf{H}_1, \mathsf{H}_2, \mathsf{H}_3]$. The grey dots denote the points T_a, T_b and T_c (see Fig. A.1 for more details), which are measured to determine the head pose.

The head orientation is defined as the orientation of the vector H_1 with respect to the coordinate system $[\mathsf{G}_1, \mathsf{G}_2, \mathsf{G}_3]$. This unit vector is computed using the cross product (hereafter denoted by $\times$) of two vectors between different pairs of points in $\mathsf{T}_{\{a,b,c\}}$, e.g.

$$\mathsf{H}_1 = \frac{(\mathsf{T}_c - \mathsf{T}_a) \times (\mathsf{T}_b - \mathsf{T}_a)}{|(\mathsf{T}_c - \mathsf{T}_a) \times (\mathsf{T}_b - \mathsf{T}_a)|}. \tag{3}$$

The head orientation angles follow from this vector in a straightforward manner:

$$\theta_{\text{h}} = \arctan\left(\frac{h_{12}}{h_{11}}\right) \tag{4}$$

$$\phi_{\text{h}} = -\arcsin\left(h_{13}\right), \tag{5}$$

where $(h_{11}, h_{12}, h_{13})^{\mathrm{T}} = \mathsf{H}_1$.

The torsional component of the head orientation must also be computed. This does not influence the line of sight, but it does modify the relationship between the eye-in-head $[\mathsf{H}_1, \mathsf{H}_2, \mathsf{H}_3]$ and ground-based $[\mathsf{G}_1, \mathsf{G}_2, \mathsf{G}_3]$ coordinate systems. The head rotation must be computed to know how the two gaze components should be combined. The head axis H_1 is made parallel to G_1 by left-multiplying the vectors T with the following orthogonal rotation matrix:

$$\begin{pmatrix} \cos(\theta_{\text{h}})\cos(\phi_{\text{h}}) & -\sin(\theta_{\text{h}}) & \cos(\theta_{\text{h}})\sin(\phi_h) \\ \sin(\theta_{\text{h}})\cos(\phi_{\text{h}}) & \cos(\theta_{\text{h}}) & \sin(\theta_{\text{h}})\sin(\phi_{\text{h}}) \\ -\sin(\phi_{\text{h}}) & 0 & \cos(\phi_{\text{h}}) \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} \cos(\theta_{\text{h}})\cos(\phi_{\text{h}}) & -\sin(\theta_{\text{h}}) & \cos(\theta_{\text{h}})\sin(\phi_{\text{h}}) \\ \sin(\theta_{\text{h}})\cos(\phi_{\text{h}}) & \cos(\theta_{\text{h}}) & \sin(\theta_{\text{h}})\sin(\phi_{\text{h}}) \\ -\sin(\phi_{\text{h}}) & 0 & \cos(\phi_{\text{h}}) \end{pmatrix}^{\mathrm{T}}. \tag{6}$$

Geometrically, head torsion corresponds to the angle between this rotated vector $(\mathsf{T}_c - \mathsf{T}_b)$, hereafter referred to as the *forehead* vector $\mathsf{F} = (f_1, f_2, f_3)^{\mathrm{T}}$, and the G_2 axis. This angle is equal to

$$\psi_{\text{h}} = \arctan\left(\frac{\cos(\theta_{\text{h}})\sin(\phi_{\text{h}})f_1 + \sin(\theta_{\text{h}})\sin(\phi_{\text{h}})f_2 + \cos(\phi_{\text{h}})f_3}{-\sin(\phi_{\text{h}})f_1 + \cos(\theta_{\text{h}})f_2}\right). \tag{7}$$

2.1.3. The gaze in space

This section integrates the previously defined eye-in-head and head components of the gaze orientation into a single vector giving the gaze orientation in the $[\mathsf{G}_1, \mathsf{G}_2, \mathsf{G}_3]$ coordinate system. The origin of this vector is the center of the eye E. The orientation vector follows directly from the eye-in-head vector P, and can be obtained by substituting (1) and (2) into (A.6), then applying three rotations to the result: first by the horizontal

angle θ_h, then by the "meridian" angle ϕ_h, and finally by the torsional angle ψ_h obtained from (4), (5) and (7) respectively:

$$\mathsf{P}_G = \begin{pmatrix} \cos(\theta_h) & -\sin(\theta_h) & 0 \\ \sin(\theta_h) & \cos(\theta_h) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\phi_h) & 0 & \sin(\phi_h) \\ 0 & 1 & 0 \\ -\sin(\phi_h) & 0 & \cos(\phi_h) \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\psi_h) & -\sin(\psi_h) \\ 0 & \sin(\psi_h) & \cos(\psi_h) \end{pmatrix} \mathsf{P}$$

$$= \begin{pmatrix} \cos(\theta_h)\cos(\phi_h) & \cos(\theta_h)\sin(\phi_h)\sin(\psi_h) - \sin(\theta_h)\cos(\psi_h) & \cos(\theta_h)\sin(\phi_h)\cos(\psi_h) + \sin(\theta_h)\sin(\psi_h) \\ \sin(\theta_h)\cos(\phi_h) & \sin(\theta_h)\sin(\phi_h)\sin(\psi_h) + \cos(\theta_h)\cos(\psi_h) & \sin(\theta_h)\sin(\phi_h)\cos(\psi_h) - \cos(\theta_h)\sin(\psi_h) \\ -\sin(\phi_h) & \cos(\phi_h)\sin(\psi_h) & \cos(\phi_h)\cos(\psi_h) \end{pmatrix} \mathsf{P} = \mathsf{R}_h\mathsf{P} \tag{8}$$

P_G therefore denotes the vector from E to P in the ground-based coordinate system, and defines the orientation of the line of sight.

2.1.4. Vergence

In general, the lines of sight of both eyes intersect at a specific distance from the subject called the depth of focus. The angle defining the difference between the two lines of sight is called the *vergence*. This point of intersection is easily determined with our method, since the origins (E_l and E_r, for the left and right eyes respectively) and directions (P_{Gl} and P_{Gr}) of both lines are known. While the two lines of sight should belong to a common plane, measurement noise may skew their apparent orientations. In this case, the best approximation to their point of intersection is halfway along the segment connecting their two closest points.[1] This point, denoted P_{Gv}, is equal to (Goldman, 1990)

$$\mathsf{P}_{Gv} = \frac{(\mathsf{E}_l + \mathsf{P}_{Gl}s_l + \mathsf{E}_r + \mathsf{P}_{Gr}s_r)}{2}, \tag{9}$$

where s_l and s_r are given by

$$s_l = \frac{\det(\mathsf{E}_r - \mathsf{E}_l, \mathsf{P}_{Gr}, \mathsf{P}_{Gl} \times \mathsf{P}_{Gr})}{|\mathsf{P}_{Gl} \times \mathsf{P}_{Gr}|^2} \tag{10}$$

$$s_r = \frac{\det(\mathsf{E}_r - \mathsf{E}_l, \mathsf{P}_{Gl}, \mathsf{P}_{Gl} \times \mathsf{P}_{Gr})}{|\mathsf{P}_{Gl} \times \mathsf{P}_{Gr}|^2}. \tag{11}$$

P_{Gv} can therefore be interpreted as the subject's point of interest.

2.1.5. Summary of the procedure

After calibration (see Section 2.2) it is possible to compute the angular orientation of the eye in space, given the position of the pupil center $(x, y)^T$ in the head coordinate system and the positions of the points T, as follows:

(i) The head orientation angles θ_h, ϕ_h, and ψ_h are computed from the positions of T_a, T_b and T_c using Eqs. (4), (5) and (7).
(ii) The position of the eye center E is computed from the positions of T_a, T_b and T_c and known anthropomorphic parameters.
(iii) Vertical and horizontal eye angles are computed from (1) and (2), using coefficients determined through calibration.
(iv) The eye-in-head orientation vector P is then determined from these angles and (A.6).
(v) The eye orientation vector is expressed in terms of the ground coordinate system by applying the rotation matrix (8).
(vi) Optionally, the eye-in-space horizontal and vertical Fick angles of the gaze are computed by the equations

$$\theta_G = \arctan\left(\frac{p_{G_2}}{p_{G_1}}\right) \tag{12}$$

$$\phi_G = -\arcsin(p_{G_3}), \tag{13}$$

where $(p_{G_1}, p_{G_2}, p_{G_3})^T = \mathsf{P}_G$.
(vii) The point of intersection between the lines of sight and the angle of vergence can also be computed, according to the equations derived in Section 2.1.4.

2.2. Calibration method

As addressed earlier, the points T fixed to the head are assumed to be measured in the ground-based coordinate system. We also assume that they define a plane parallel to the frontal plane, such that $\theta_h = \phi_h = \psi_h = 0$ when the subject is looking straight ahead (the primary position). If the points T cannot be accurately fixed with respect to the subject's head, they can still be calibrated to ensure that $\theta_h = \phi_h = \psi_h = 0$ in the primary position by a method outlined in Appendix B.

Calibration protocols for the measurement of eye-in-head orientation via pupil detection by image processing devices are well documented in the literature (e.g., Haslwanter and Moore, 1995; Moore et al., 1996; Clarke et al., 2002; Schreiber and Haslwanter, 2004). They are generally based on a small number of fixations at *known* horizontal and vertical Fick angles in the head coordinate system. The objective of calibration is to identify the unknown coefficients a'_{ij} in (1) and (2). The torsional component does not need to be calibrated for image processing devices, since the polar cross-correlation technique provides an angular measurement directly. It is only necessary to define an iral reference signature indicating the eye's primary position in the orbit.

This paper considers tasks in a framework where the head is unrestrained, and it is inconvenient to ask a subject to keep

[1] Geometrically, this segment is *orthogonal* to both lines of sight.

their head fixed during calibration. The identification of fixation targets at known horizontal and vertical positions with respect to the head coordinate system is therefore impractical. To cope with these restrictions, we propose a calibration protocol that integrates the head pose measurement. The subject is asked to move their head while keeping their gaze fixed on a specific point. The results of the calibration have to match the eye displacements generated by this procedure, given the horizontal and vertical Fick angles of the target in the head coordinate system $[\mathsf{H}_1, \mathsf{H}_2, \mathsf{H}_3]$ (Fig. 2).

In the head coordinate system, the Fick angles of the fixation target C vary with head position as follows:

$$\theta_{\text{tar}} = \arctan\left(\frac{c_{\text{f2}}}{c_{\text{f1}}}\right) \tag{14}$$

$$\phi_{\text{tar}} = -\arcsin(c_{\text{f3}}), \tag{15}$$

where

$$(c_{\text{f1}}, c_{\text{f2}}, c_{\text{f3}})^{\text{T}} = \mathsf{C}_{\text{f}} = \mathsf{R}_{\text{h}}^{\text{T}}(\mathsf{C} - \mathsf{E}) \tag{16}$$

is the position of C in the head coordinate system whose origin is the eye center. The calibration is performed by inverting (1) and (2), and adding a third equation that represents the translation offset (see Maxwell, 1951; Denavit and Hartenberg, 1955, for references about general homogeneous coordinates):

$$\begin{pmatrix} \sin(\theta_{\text{tar}})\cos(\phi_{\text{tar}}) \\ -\sin(\phi_{\text{tar}}) \\ 1 \end{pmatrix} = \text{A}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} a'_{11} & a'_{12} & a'_{13} \\ a'_{21} & a'_{22} & a'_{23} \\ a'_{31} & a'_{32} & a'_{33} \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \tag{17}$$

The best "gain" matrix **A** for this overdetermined system can be determined by any numerical processing software. The computation gives the best-fitting solution matrix, in the least squares sense, to the series of eye signals $(x, y, 1)^{\text{T}}$ and target signals $(\sin(\theta_{\text{tar}})\cos(\phi_{\text{tar}}), -\sin(\phi_{\text{tar}}), 1)^{\text{T}}$ generated by the head movements. This problem is well-conditioned, since the horizontal and vertical eye positions generated during the calibration task are highly independent. The gains a'_{31} and a'_{32} are identified to zero, and a'_{33} to 1, to agree with the form of the vectors used in (17).

The subjects we tested reported no difficulty in maintaining gaze fixation during the calibration task, since the velocity of their head displacements was only about $25^\circ\,\text{s}^{-1}$ on average. In this range the gain of smooth pursuit is very close to 1 with negligible phase lag (see Lisberger et al., 1981), validating the accuracy of gaze fixation for the expected behavior.

2.2.1. Summary of the procedure

To calibrate the eye-in-head orientation with a video-based unit, the subject is asked to maintain gaze fixation on a known point (C) while moving their head first horizontally and then vertically. A series of pupil center displacements $(x, y)^{\text{T}}$ and head pose movements are thereby generated. The gain matrix **A** is obtained through the following procedure:

(i) If necessary, the points T are artificially corrected such that $\theta_{\text{h}} = \phi_{\text{h}} = \psi_{\text{h}} = 0$ in the primary position by asking the subject to maintain this position for a few seconds. This procedure is detailed in Appendix B.
(ii) The head Fick angles θ_{h}, ϕ_{h} and ψ_{h} are calculated from the T positions and Eqs. (4), (5) and (7).
(iii) The position of the eye center E is computed from the positions of T_a, T_b and T_c and known anthropomorphic parameters.
(iv) The target position in the head coordinate system is computed from (16).
(v) The Fick angles of the target in the head coordinate system are computed by (14) and (15).
(vi) The components a'_{ij} of (17) are calculated, using θ_{tar} and ϕ_{tar} from the previous step and $(x, y)^{\text{T}}$ values from the eye-in-head measurement device.

This calibration method could be easily adapted to clinical studies of patients with oculomotor disorders who are not able to maintain gaze fixation while moving the head. In this case, a "discretized" version of the calibration task could be proposed: for several steady head positions, the patient would be asked to stabilize their gaze towards the calibration target C. Only steady eye-head orientations could be used to compute the linear regression (17).

2.3. Experimental validation

Data have been collected on five human subjects (S1 is female, and S2–S5 are male) between 24 and 27 years of age (mean 25.4). They provided informed written consent, and reported no history of neurological or musculoskeletal disorder. All had normal vision, either natural or corrected. One subject is the first author of this paper, while the others were naive regarding the goals of the experiment. All the procedures conducted were approved by the Université catholique de Louvain Ethics Committee, in compliance with the Helsinki declaration.

A sketch of the experimental setup is given in Fig. 3. Initially, the subject was asked to stand upon a cross, marked on the ground (point A). He or she had to keep their gaze fixed on an IRED marker (point C) placed on a camera tripod 1.3 m away. Following the method described in Section 2.2, each subject was asked to move their head first with a pure horizontal movement, then with a pure vertical movement (Pattern 1 on Fig. 3). Two other patterns were used to *validate* the calibration so obtained. Pattern 2 consisted of a circular head motion, and was used to validate the *coupling* between horizontal and vertical components. Pattern 3 consisted of a ∞-shaped head motion, the subject being asked to generate significant head *torsion*. The robustness of the calibration to lateral displacement was evaluated by asking the subject to repeat these patterns while standing on positions B and C (0.5 m to the right and left of point A). The robustness of the calibration to distance was evaluated by repeating the proce-

dure at position D (0.5 m in front of point A). The task was easy to perform under all conditions, resulting in unambiguous gaze orientations. Except for pattern 1 on position A, which was used in calibration, every other pattern and position was used only to validate the method. The method's accuracy is determined by comparing the computed gaze orientation to the actual position of C. The subjects went through the following sequence of patterns: position A, pattern 1 (calibration) then patterns 2 and 3; position B, patterns 1, 2 and 3; position C, patterns 1, 2 and 3; position D, patterns 1, 2 and 3; finally position A, pattern 1 once more for a second calibration.

Two-dimensional (horizontal and vertical) recordings of both eyes were made simultaneously using a Chronos head-mounted, video-based eye tracker (Clarke et al., 2003, CHRONOS VISION GmbH, Berlin, Germany). The calculation of eye positions was based on determination of the pupil center (see Zhu et al., 1999, and references therein). The recording frame rate was 200 Hz. The system is linear in the horizontal and vertical planes for deflections up to ±20°, and has a resolution better than 5′. System noise was measured to be 0.11° for the horizontal plane and 0.09° for the vertical plane (Clarke et al., 2002, 2003). A bite-bar was mounted on the helmet frame to prevent slippage between the head and the helmet. This bite-bar was not mandatory, however, and could be removed for subject comfort. In this case the calibration task would be performed at regular intervals to compensate for any slippage between the head and the helmet. The second video-based device used was a 3D position measurement system. The positions of infrared light-emitting diodes (IREDs) on the Chronos helmet and at the fixation target (the four grey dots on Fig. 3) were measured using an OptoTrak 3020 system (Northern Digital, Ontario, Canada). The OptoTrak was mounted on the ceiling about 3 m in front of the subject. The positions of the IREDs were rotated so that they could be expressed in a coordinate system with axes parallel to the floor and centered on the fixation target. The axes G's of this ground coordinate system are shown in Fig. 2. The position of each IRED was sampled with a frequency of 200 Hz and resolution of about 0.1 mm within this working environment. The eye signals and IRED signals were filtered at 48 Hz by a zero-phased digital filter (autoregressive, forward and backward). The Chronos eye tracker and the OptoTrak were synchronized by an external computer. Each pattern was executed over a period of 20 s, and its data recorded in a separate file.

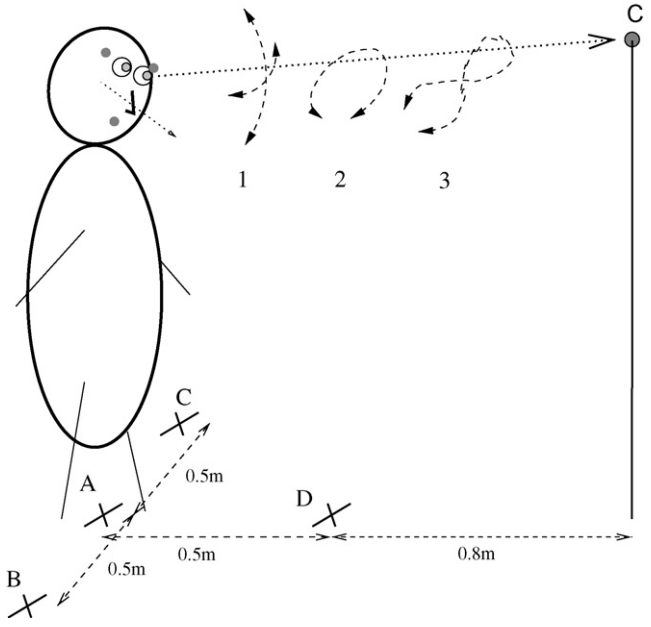

Fig. 3. Experimental setup. Standing on one of the fixed positions (crosses A, B, C and D), the subject is asked to maintain gaze fixation on the grey spot (point C) while moving their head in a cross pattern (1), a circular pattern (2), or an ∞-shaped pattern (3).

3. Results

Experimental results are presented in two separate sections. Section 3.1 describes the calibration results, and Section 3.2 validates the calibration matrices using data from the patterns described in Section 2.3.

3.1. Calibration

Calibration of the eye-in-head measurement device relies on the algorithm described in Section 2.2. It assumes that the eye position E is known perfectly with respect to the points T. To keep the method as simple as possible, we assume that this distance is similar for all subjects. Empirical observations of all the subjects wearing the helmet allow us to estimate the eye positions as

$$\mathrm{E} = \frac{2\mathrm{B_T} + \mathrm{T}_c}{3} - 0.09\,\mathrm{m}\,\mathrm{H}_1 + 0.01\,\mathrm{m}\,\mathrm{F} \tag{18}$$

for the left eye, and

$$\mathrm{E} = \frac{2\mathrm{B_T} + \mathrm{T}_b}{3} - 0.09\,\mathrm{m}\,\mathrm{H}_1 - 0.01\,\mathrm{m}\,\mathrm{F} \tag{19}$$

for the right eye. In these equations $\mathrm{B_T}$ is the barycenter of the points T (see Appendix B), which gives an interocular distance of about 0.07 m. Note that any errors induced by this approximation are corrected to the zeroth and first order by the linear gains in (17) during the calibration procedure. A more complicated calibration algorithm could treat these distances as additional unknown parameters to be determined. In this case the calibration would become non-linear, however, and in addition to being less robust would require a more complicated and computationally costly implementation.

To calibrate the helmet unit, the subject executed pattern 1 in position A (see Fig. 3). From the E position computed by (18) or (19), the *target* Fick angles are computed according to (14) and (15). The overall motion is slow enough to assume that the subjects maintained a permanent fixation on the point C. In the calibration data (a 20 s recording) the experimenter manually excluded eye blinks, small saccades, and eye movements outside the detection range to reduce signal distortion in the linear regression (17).

The real target angles and eye-in-head angles are compared after calibration in Fig. 4 (left eye, S2). This diagram empha-

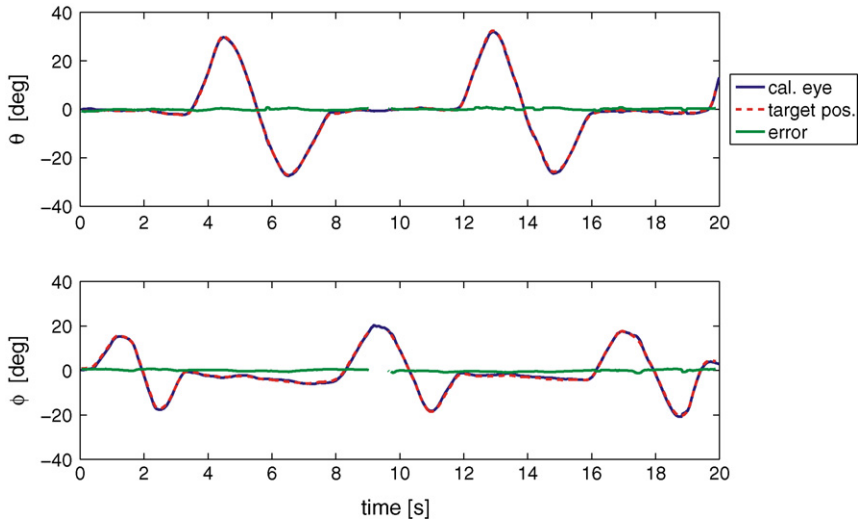

Fig. 4. The target position (dashed red lines) expressed in θ_{tar} (above) and ϕ_{tar} (below) is compared to the computed eye position after calibration (solid blue lines). The solid green lines represent the difference between the two orientations. For eccentric head orientations, in some cases one of the LEDs on the helmet could not be measured (the gaps in the green and red curves around 9 s).

sizes the sinusoidal motion (alternating pure horizontal and pure vertical movements) of the eye-in-head orientation during the smooth head displacement of the calibration task. Fig. 4 also illustrates the low frequency of the smooth head movements. The green lines denote the error, i.e., the difference between the target angles and the computed eye-in-head angles. The errors in θ over the test period have an average value of 0.01° and a standard deviation of 0.35°. For ϕ the average error is 0°, and the standard deviation is 0.43°. Well-conditioned linear regression of (17) gives the following matrix equation for the orientation of the left eye (subject S2):

$$\begin{pmatrix} \sin(\theta_{tar})\cos(\phi_{tar}) \\ -\sin(\phi_{tar}) \\ 1 \end{pmatrix} = \begin{pmatrix} 0.0076 & 0.0006 & -0.0005 \\ -0.0002 & 0.0089 & -0.0013 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (20)$$

The average errors and standard deviations of the error, for both eyes and each subject, are listed in Table 1. The mean error is never significantly different from 0, as should be the case for a linear regression method. The standard deviations, however, certainly reflect a real variability in eye positions (see Ott et al., 1990, and references therein) as well as system noise.

3.2. Validation tasks

After performing the calibration task, the subjects were asked to move their heads in other patterns and stand on different predefined points in the workspace. All these tests were to be performed while maintaining their fixation on the same point C (see Section 2.3). Under these conditions, the known position of C was compared *a posteriori* to the gaze orientation obtained by our algorithm.

Typical head motions for all three patterns are depicted in Fig. 5. All these data were obtained from subject S2, while standing on point B. The mean value of θ_h is therefore positive (the subject's gaze points to the left), while the mean value of ϕ_h and ψ_h are close to 0 for each pattern. Pattern performances were similar for all subjects, so their validation data were pooled in the analysis.

Several error parameters were computed for each eye, each pattern, and each standing position. $\Delta\theta_{left,right}$ is the difference between θ_{tar} and the gaze angle θ_G; $\Delta\phi_{left,right}$ is the difference between ϕ_{tar} and the gaze angle ϕ_G. We also define the error $\Delta|\alpha_{left,right}|$ as the absolute value of the angle between the the-

Table 1
Mean and standard deviation of the error by the calibration method

	First calibration		Second calibration	
	Left eye (°)	Right eye (°)	Left eye (°)	Right eye (°)
S1				
θ	0.06 ± 1.16	0.01 ± 0.37	0.06 ± 1.05	−0.01 ± 0.27
ϕ	0 ± 0.53	0 ± 0.44	0 ± 0.49	0 ± 0.24
S2				
θ	0.02 ± 0.48	0 ± 0.38	0.01 ± 0.35	0 ± 0.35
ϕ	0 ± 0.65	0 ± 0.73	0 ± 0.43	0 ± 0.43
S3				
θ	0.02 ± 0.65	0 ± 0.37	0.04 ± 0.75	0 ± 0.52
ϕ	0 ± 0.45	0 ± 0.59	−0.01 ± 1.32	−0.01 ± 1.38
S4				
θ	0.03 ± 0.99	−0.01 ± 0.65	0.01 ± 0.28	−0.01 ± 0.39
ϕ	0 ± 0.85	0 ± 0.58	0 ± 0.36	0 ± 0.4
S5				
θ	0.09 ± 1.72	−0.01 ± 0.96	0.02 ± 0.43	0.01 ± 0.82
ϕ	0 ± 0.38	−0.01 ± 1.08	−0.01 ± 1.16	0 ± 0.6

Values are in mean ± S.D.

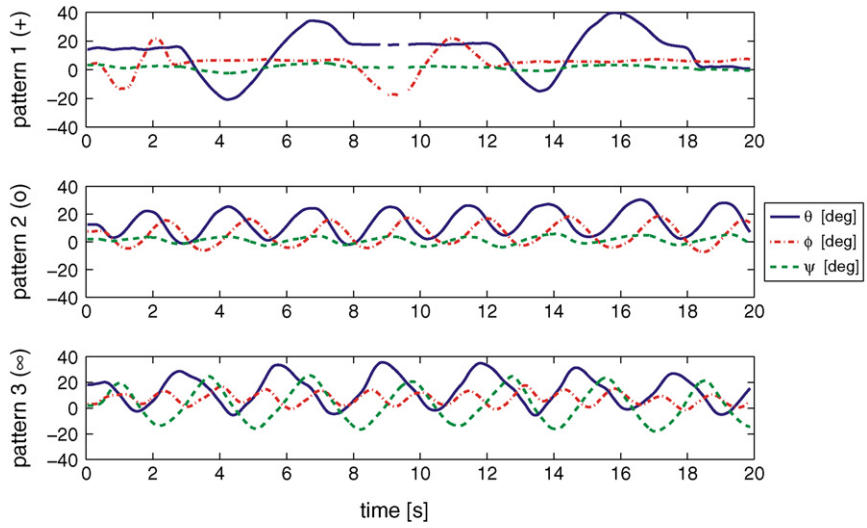

Fig. 5. Typical head motions for the three patterns. Pattern 1 (top) is characterized by alternating pure horizontal (θ, blue solid) and pure vertical (ϕ, red dash-dotted) movements, with no torsion (ψ, green dashed). Pattern 2 (middle) is characterized by coupled, sinusoidal horizontal and vertical movements with a 90° phase difference, and limited torsion. In this pattern, the head traces a circle. Pattern 3 (bottom) is characterized by harmonic horizontal and vertical movements in a 2:1 frequency ratio (the head follows an ∞-shaped path), and a significant torsion component. For some head orientations, one of the LEDs on the helmet was not measured (the gaps in the curves).

Table 2
Error parameters for each standing position {A, B, C, D} and head pattern {1, 2, 3}

	θ_h	ϕ_h	ψ_h	$\Delta\theta_{left}$		$\Delta\phi_{left}$		$\Delta\lvert\alpha_{left}\rvert$		$\Delta\theta_{right}$		$\Delta\phi_{right}$		$\Delta\lvert\alpha_{right}\rvert$	
				Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
A															
P1	±35	±24	±4	−0.06	0.75	0.02	0.62	0.83	0.52	−0.02	0.58	0.03	0.61	0.72	0.44
P2	±28	±24	±7	0.83	1.64	0.52	1.8	2.26	1.33	0.27	1.14	0.71	1	1.5	0.79
P3	±36	±23	±30	0.6	1.69	−0.64	1.78	2.35	1.11	0.09	1.5	−0.5	1.71	2.02	1.17
B															
P1	−22, 47	±22	±4	0.45	3.47	−0.36	1.08	1.69	3.23	−0.59	1.18	−0.03	0.86	1.32	0.86
P2	−17, 41	±24	±9	0.23	1.09	−0.33	1.01	1.2	0.96	−0.35	1.28	−0.35	0.88	0.97	1.25
P3	−37, 46	±22	±30	0.06	3.47	−0.5	1.56	1.61	3.47	−0.45	1.26	−0.64	1.48	1.72	1.18
C															
P1	−50, 23	±24	±4	−0.45	0.99	−0.25	1.37	1.43	1.05	−0.58	1.07	−0.12	1.51	1.61	1.09
P2	−37, 16	±23	±9	−0.31	1.2	−1.23	1.43	1.98	1.08	−0.45	1.1	−0.85	1.55	1.91	0.95
P3	−53, 23	±25	±33	−0.12	1.16	−1.15	1.44	1.95	0.98	0.05	1.45	−1.21	1.6	2.22	1.09
D															
P1	±35	±20	±4	0.11	0.82	0.4	0.88	1.04	0.72	0.37	0.78	0.45	0.84	1.08	0.69
P2	±29	±25	±7	0.56	0.91	0.72	1.16	1.59	0.71	0.76	0.66	0.73	1.39	1.75	0.66
P3	±34	±21	±32	0.49	0.98	1.03	1.54	1.96	0.88	0.81	1.02	0.41	1.35	1.77	0.75
A															
P1	±36	±24	±4	0.07	2.31	0.03	0.91	0.76	2.33	−0.04	0.48	0.01	0.62	0.58	0.52

The three first columns present the overall range of head angles for each test. The following angular errors for the left and right eyes are averaged across all subjects: $\Delta\theta_{left}$ is the difference between θ_{tar} and θ_G, $\Delta\phi_{left}$ is the difference between ϕ_{tar} and ϕ_G, and $\Delta|\alpha_{left}|$ is the total angular error as defined by (21). All these data are expressed in degrees (°). The listed ranges are the standard deviations of the angular errors.

oretical and the real gaze vectors:

$$\Delta|\alpha| = \left|\arccos\left(\frac{C_f P_G}{|C_f|}\right)\right| \geq 0. \tag{21}$$

All these parameters are given in Table 2, together with the typical ranges of θ_h, ϕ_h and ψ_h (i.e., the minimum and maximum head angles for each subject). The average absolute error is always less than 3.5°, which is satisfactory with respect to the range of head movements. In particular, the third pattern generates the most important errors. This is due to the relative difficulty of maintaining a fixed gaze while generating the significant head torsion required by this pattern.

As a supplement to this paper, several animations of the three patterns have been uploaded to the publisher's web service[2]. They are also available on the first author's web page.

[2] See Appendix C.

4. Discussion

This paper investigates the geometry of human gaze orientations. To compute gaze orientation in a behavioral context where the head has complete freedom of movement, not only head rotations but also head *translations* must be taken into account. This paper provides a simple calibration protocol that can integrate both eye-in-head and head-in-space orientations into a single coordinate system.

Our method rests on separate measurements of the head pose and the eye-in-head position. The head pose is determined by measuring the positions of three markers fixed to the head. In the present experiment the eye-in-head position is recorded with a video-based device, but any other classical eye recording method (corneal reflection methods, EOG, etc.) will do as well. The availability of independent head and eye-in-head orientation signals invites further investigation of their interaction. For example, if a search coil is used then eye-in-space angular components are recorded directly but a contribution due to head motion must be subtracted from this signal (see, e.g. Crawford and Vilis, 1991; Crawford et al., 1999, 2003; Harris et al., 2001; Han et al., 2005).

Dual search coil and video-based devices are equally popular methods of measuring eye orientation, and each has its specific advantages and drawbacks. On the one hand, *dual search coils* (Robinson, 1963; Collewijn et al., 1985) are expensive and fragile. They are also invasive, since a human subject can endure the presence of a coil on their eye for at most 40 min, and require the supervision of an ophthalmologist. Moreover, because the coil measures eye orientation in a ground-based coordinate system it is not straightforward to isolate the eye-in-head component of the gaze orientation. The main advantages of the search coil technique are its excellent spatial accuracy and the high sampling frequencies that can be obtained. With horizontal, vertical, and torsional gaze orientations available in real time, it is easy to implement closed loop protocols. On the other hand, recent *video image processing* devices have been able to accurately measure eye movements using camera-based systems (see the paper by Clarke et al., 2003, that presents the device used in this study). These techniques are currently approaching the search coil technique in terms of spatial and temporal accuracy. Most importantly, they are much less invasive.

With the development of more advanced hardware, a mathematical formalism based on 3D rotations has been established to relate facial images to gaze orientation (Nakayama, 1974; Tweed and Vilis, 1987; Tweed et al., 1990; Van Opstal, 1993; Haslwanter, 1995). Horizontal and vertical displacements of the eye are computed by tracking the pupil center (see Zhu et al., 1999, and references therein). Torsion may be computed either by tracking natural or artificial landmarks on the eye (Nakayama, 1974; Parker et al., 1985; Ott et al., 1990), or by tracking the iral intensity along a circular sampling path (the *polar cross-correlation* method) (Hatamian and Anderson, 1983; Vieville and Masse, 1987; Tweed et al., 1990; Clarke et al., 1991; Moore et al., 1991, 1996; Haslwanter, 1995, 2000; Haslwanter and Moore, 1995; Schreiber and Haslwanter, 2004).

The main limitation of video-based acquisition systems is that they do not capture head motions, since the cameras recording eye motion are fixed to a helmet worn by the subject. This paper proposes a method of integrating both translations and rotations of the head into any video-based system, based on the measurement of three points on the subject's head. This extra information allows the gaze orientation to be calculated accurately in a ground-based coordinate system. The first part of this paper (Section 2.1) covered the geometrical developments of our method. Eye-in-head orientation is first computed using classical techniques (Moore et al., 1996), then translated and rotated into the ground coordinate system once the head pose has been computed. For the sake of simplicity, this paper did not summarize the existing methods of computing eye *torsion* from video-based devices. Eye torsion does not modify the line of sight, so is not a critical issue in this paper.

In addition to the accurate video-based devices described above, we would like to discuss another fruitful aspect of gaze estimation in the literature. Several methods have been developed to estimate the gaze orientation of a head before a computer screen. The goal of such studies is to use gaze as part of the interface between humans and computerized devices, for applications both clinical and otherwise. The main difference between these methods and the technique described in this paper lies in the hardware; the cameras that track the eye pupil must be fixed to a ground-based coordinate system instead of the subject's head. This has the advantage of measuring gaze orientation directly in the ground-based coordinate system. On the other hand, it also requires continuous monitoring of the distance between the camera and the eye, which is not easy to measure accurately. Even though the most recent of these methods take head translations into account, they cannot compute gaze orientation for very large head displacements (for example, about 1 m). Furthermore, these techniques do not achieve the level of spatial and temporal resolutions required in oculomotor research. Papers of interest in this field of study include that of Newman et al. (2000), which proposes a method of reconstructing the head pose in real time (30 Hz) by tracking typical features of the head; and that of Wang and Sung (2002), which presents a similar method based on linear algebra with homogeneous coordinates. An overview of this topic can be found in recent papers (Shih and Liu, 2004; Yoo and Chung, 2005), as well as the "Computer Vision Bibliography" web page which refers extensively to this literature [3]. Finally, an alternative approach to determining the point of gaze has been developed by Yu and Eizenman (2004). This method is based on the detection of corresponding points in an image from a head-mounted scene camera and a reference image. This method is not designed to accurately locate the head and other objects in a 3D ground-based coordinate system.

In contrast with the restricted workspace of a computer screen, our method allows not only large head movements but also subject displacements as long as the head markers T remain in the recording space.

[3] http://iris.use.edu/Vision-Notes/bibliography/people911.html.

In Section 2.2, we detailed an algorithm for calibrating a video-based eye tracker unit by converting pixel outputs into angular displacements. The mathematical operations involved in this calibration rest on the rotation and translation of vectors and matrices, as was the case for computation of the gaze orientation.

The main advantages of our method can be summarized as follows:

- it is non-invasive for the subjects, since a video-based device is used to compute eye-in-head orientation;
- its algorithm is expressed in terms of linear algebra, and is therefore computationally efficient;
- it separates the eye-in-head and the head-in-space components of gaze, thus allowing the study of their mutual interaction;
- it can be used with a wide variety of acquisition devices, provided that the 3D positions of three points on the head can be independently measured;
- it is based on the geometry of the body, so is easy to understand and implement;
- it is robust under challenging experimental conditions; a preliminary version of this method has already been tested in parabolic flight campaigns by the European Space Agency.

Future improvements to our method mainly concern algorithmic issues: nonlinear estimation of the eye center position with respect to the head markers (this would, however, greatly complicate the calibration algorithm), more accurate pupil center detection (see Zhu et al., 1999), decoupling the horizontal and vertical axes of eye rotation (Schreiber and Haslwanter, 2004), etc.

Section 3 was devoted to the experimental validation of our method. The calibration process has been strongly validated by data from five different human subjects, confirming that the proposed linear algorithm accurately reflects the dynamics of gaze orientation. Using the calibration matrices so obtained, we also performed other validation tasks. As expected, the smallest angular errors were measured for the first pattern in position A; i.e., for the data using in computing the calibration matrix. The errors remained within reasonable bounds, however, for all head patterns and standing positions. In general, the angular errors obtained in the validation tasks compared favorably to the natural variability of gaze orientation in typical fixation tasks (see Ott et al., 1990, and references therein). The errors incurred by this method are also similar to those obtained by previous experiments comparing the accuracy of video-based and search coil eye tracking techniques in human experiments (e.g., Ott et al., 1990; Moore et al., 1996; Clarke et al., 2002, 2003).

This paper therefore provides a computationally efficient procedure for computing eye orientation in a ground-based coordinate system. This method can be implemented rapidly in a variety of settings, since it is based on linear transformations related to the actual body configuration.

Acknowledgements

This work was supported by the Belgian Program on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office; the 'Fonds National de la Recherche Scientifique'; the 'Fondation pour la Recherche Scientifique Médicale'; an internal research grant ('Fonds Spéciaux de Recherche') from the 'Université catholique de Louvain'; the European Space Agency; and PRODEX (Belgium).

Appendix A. Eye-in-head orientation

We define an orthogonal, right-handed, head-based coordinate system $[\mathsf{H}_1, \mathsf{H}_2, \mathsf{H}_3]$ with its origin at the center of the eyeball. The H_2 axis is parallel to the interaural axis, and the H_2–H_3 plane is parallel to the frontal plane. The H_1 axis therefore points out of the face from the occiput (see Fig. A.1). The eye's *primary position* is defined as the position where its line of sight corresponds to the H_1 axis. We also define a *camera* coordinate system $[\mathsf{C}_1, \mathsf{C}_2, \mathsf{C}_3]$, where C_2 and C_3 lie within the image plane and C_1 corresponds to the optical axis of the camera.

According to the definition of Moore et al. (1996), the coordinates of the pupil center with respect to the head coordinate system $\mathsf{P} = (p_1, p_2, p_3)^{\mathrm{T}}$ and the the camera coordinate system $\mathsf{P}' = (p_1', p_2', p_3')^{\mathrm{T}}$ are related by

$$\mathsf{P}' = \mathsf{R}_{\mathrm{cam}}\mathsf{P} + \mathsf{T}_{\mathrm{cam}}, \tag{A.1}$$

where $\mathsf{T}_{\mathrm{cam}}$ and $\mathbf{R}_{\mathrm{cam}}$ are the translation vector and rotation matrix relating the head coordinate system to the camera coordinate system. Angular positions of the eye are expressed in terms of Fick angles (Fick, 1874; Haslwanter, 1995), which are commonly used in oculomotor research. If the space around the

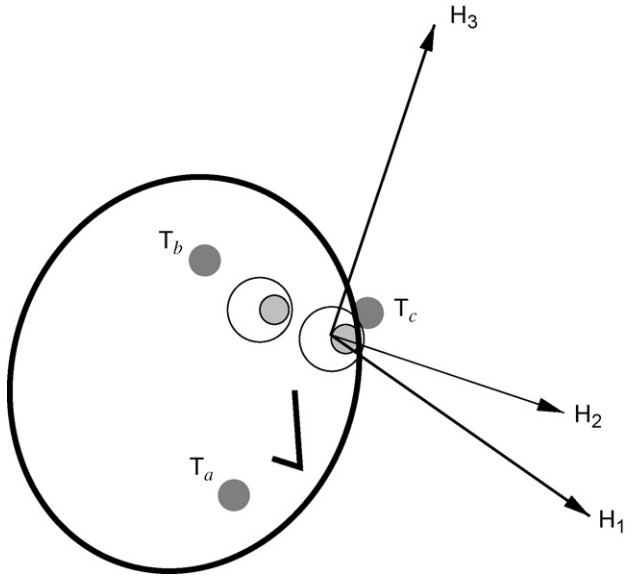

Fig. A.1. Diagram of the head in space. The coordinate system $[\mathsf{H}_1, \mathsf{H}_2, \mathsf{H}_3]$ is fixed to the head, with its origin at the center of the (arbitrarily chosen) left eye. This point is also the origin of the gaze orientation vector. The points $\mathsf{T}_{\{a,b,c\}}$ are fixed on the head, and discussed in Section 2.1.2.

eye is viewed as a sphere marked with parallels and meridians, then the sequence of Fick angles defining an orientation is as follows: first a horizontal rotation θ along the equator, then a vertical rotation ϕ along a meridian, and finally a torsional rotation ψ about the optical axis. According to the right-hand rule, eye movements to the left, down, and clockwise are positive from the subject's point of view.

Moore et al. (1996) showed that the horizontal (θ_{cam}) and vertical (ϕ_{cam}) components of the offset matrix $\mathbf{R}_{\text{cam}}$ can be captured by calibration gains and do not influence the accuracy of measurement if they remain bounded within 5°. This assumption is reasonable, provided the camera is properly fixed to the acquiring device. Eq. (A.1) therefore reduces to

$$\mathsf{P}' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\psi_{\text{cam}}) & -\sin(\psi_{\text{cam}}) \\ 0 & \sin(\psi_{\text{cam}}) & \cos(\psi_{\text{cam}}) \end{pmatrix} \mathsf{P} + \mathsf{T}_{\text{cam}}, \tag{A.2}$$

where ψ_{cam} denotes an offset rotation of the camera around its optical axis. If the distance between the lens plane and the center of the eye is large compared to the distance between the lens plane and the image plane of the camera (Nakayama, 1974; Moore et al., 1996), then the projection P'' of P' onto the image plane is given by

$$\mathsf{P}' = \begin{pmatrix} 0 \\ x \\ y \end{pmatrix} = k \begin{pmatrix} 0 \\ p'_2 \\ p'_3 \end{pmatrix} \tag{A.3}$$

where k is a scaling factor related to the image magnification and p'_i are the individual components of P'. By inverting (A.2) and (A.3), we find the following relation between the actual eye position P and the pupil center $(x, y)^T$ as measured by the camera:

$$p_2 = a_{11}x + a_{12}y + a_{13} \tag{A.4}$$

$$p_3 = a_{21}x + a_{22}y + a_{23}, \tag{A.5}$$

where the coefficients $a_{11} = a_{22} = \cos(\psi_{\text{cam}})/k$, $a_{12} = -a_{21} = \sin(\psi_{\text{cam}})/\text{k}$, $a_{13} = -\cos(\psi_{\text{cam}})t_2 - \sin(\psi_{\text{cam}})t_3$, and $a_{23} = \sin(\psi_{\text{cam}})t_2 - \cos(\psi_{\text{cam}})t_3$ all have to be determined by an appropriate calibration. The scalar variables p_i (t_i, etc.) denote the individual components of P (T_{cam}, etc.) respectively.

The pupil center P follows from applying the Fick rotation (θ_{eih}, ϕ_{eih}) to the eye's primary position $(r_p,0,0)^T$, where r_p is the distance between the center of the eye and the center of the pupil.

$$\mathsf{P} = \begin{pmatrix} \cos(\theta_{\text{eih}})\cos(\phi_{\text{eih}}) & -\sin(\theta_{\text{eih}}) & \cos(\theta_{\text{eih}})\sin(\phi_{\text{eih}}) \\ \sin(\theta_{\text{eih}})\cos(\phi_{eih}) & \cos(\theta_{\text{eih}}) & \sin(\theta_{\text{eih}})\sin(\phi_{\text{eih}}) \\ -\sin(\phi_{\text{eih}}) & 0 & \cos(\phi_{\text{eih}}) \end{pmatrix} \begin{pmatrix} r_p \\ 0 \\ 0 \end{pmatrix} \tag{A.6}$$

These rotations are depicted in Fig. 1. Substituting this result into (A.4) and (A.5), we arrive at the eye-in-head orientation:

$$\phi_{\text{eih}} = \arcsin\left(-\frac{a_{21}x + a_{22}y + a_{23}}{r_p}\right) \tag{A.7}$$

$$\theta_{\text{eih}} = \arcsin\left(\frac{a_{11}x + a_{12}y + a_{13}}{r_p \cos(\phi_{\text{eih}})}\right) \tag{A.8}$$

This result gives (1) and (2), which define $a'_{ij} \triangleq a_{ij}/r_p$.

Appendix B. Calibration of the points T

Our gaze estimation method rests on the fact that the points $\mathsf{T}_{\{a,b,c\}}$ define a plane parallel to the frontal plane $\mathsf{H}_2 - \mathsf{H}_3$ (Fig. A.1), while the forehead vector $\mathsf{F} = (\mathsf{T}_b - \mathsf{T}_c)$ is parallel to H_2. In theory this ensures that $\theta_{\text{h}} = \phi_{\text{h}} = \psi_{\text{h}} = 0$ whenever the subject stands up and looks straight ahead (the primary position). The placement of these points on the subject's head, however, cannot be accurate enough to validate this assumption. Offset yaw (θ_{off}), pitch (ϕ_{off}), and torsion (ψ_{off}) angles must therefore be measured while a subject is maintaining the primary position. This appendix describes a calibration protocol that compensates for these errors by virtually moving the points T.

Prior to the calibration pattern, the subject is asked to maintain the primary position by looking straight ahead while their head orientation angles, i.e. the offset angles, are recorded. In this reference posture, the real points $\mathsf{T}_{\{a,b,c\}}$ are related to the so-called corrected points $\mathsf{T}'_{\{a,b,c\}}$ by the transformation $(\mathsf{T}_a, \mathsf{T}_b, \mathsf{T}_c) = \mathbf{R}_{\text{off}}(\mathsf{T}'_a, \mathsf{T}'_b, \mathsf{T}'_c)$, where $\mathbf{R}_{\text{off}}$ is a rotation matrix with exactly the same structure as $\mathbf{R}_{\text{h}}$ in (8). However, this last relation is only valid when the Fick angles of $\mathsf{T}'_{\{a,b,c\}}$ are equal to zero. To put it another way, when the points $\mathsf{T}_{\{a,b,c\}}$ are such that their *measured* Fick angles are equal to zero, we have $(\mathsf{T}'_a, \mathsf{T}'_b, \mathsf{T}'_c) = \mathsf{R}_{-\text{off}}(\mathsf{T}_a, \mathsf{T}_b, \mathsf{T}_c)$. In this case, $\mathbf{R}_{-\text{off}}$ is the Fick rotation matrix with angles $-\theta_{\text{off}}$, $-\phi_{\text{off}}$ and $-\psi_{\text{off}}$. In order to make this relation valid for any head orientation, the sets of points $\mathsf{T}_{\{a,b,c\}}$ and $\mathsf{T}'_{\{a,b,c\}}$ must be pre-multiplied by $\mathsf{R}_{\text{h}}^{-1} = \mathsf{R}_{\text{h}}^{\text{T}}$ ($\mathbf{R}_{\text{h}}$ is an orthogonal rotation matrix):

$$\mathsf{R}_{\text{h}}^{\text{T}}(\mathsf{T}'_a, \mathsf{T}'_b, \mathsf{T}'_c) = \mathsf{R}_{-\text{off}}\mathsf{R}_{\text{h}}^{\text{T}}(\mathsf{T}_a, \mathsf{T}_b, \mathsf{T}_c). \tag{B.1}$$

The new set of head points is therefore

$$(\mathsf{T}'_a, \mathsf{T}'_b, \mathsf{T}'_c) = \mathsf{R}_{\text{h}}\mathsf{R}_{-\text{off}}\mathsf{R}_{\text{h}}^{\text{T}}(\mathsf{T}_a, \mathsf{T}_b, \mathsf{T}_c). \tag{B.2}$$

Finally, for convenience this rotation is applied around the *barycenter* B_{T} of $\mathsf{T}_{\{a,b,c\}}$, $((\mathsf{T}_a + \mathsf{T}_b + \mathsf{T}_c)/3)$ which is then conserved.

$$(\mathsf{T}'_a, \mathsf{T}'_b, \mathsf{T}'_c) = \mathsf{R}_{\text{h}}\mathsf{R}_{-\text{off}}\mathsf{R}_{\text{h}}^{\text{T}}(\mathsf{T}_a - \mathsf{B}_{\mathsf{T}}, \mathsf{T}_b - \mathsf{B}_{\mathsf{T}}, \mathsf{T}_c - \mathsf{B}_{\mathsf{T}}) + (\mathsf{B}_{\mathsf{T}}, \mathsf{B}_{\mathsf{T}}, \mathsf{B}_{\mathsf{T}}) \tag{B.3}$$

In summary, the set of points $\mathsf{T}'_{\{a,b,c\}}$ can be simply derived from the measured points $\mathsf{T}_{\{a,b,c\}}$. Their Fick angles are equal to zero when the subject is in the primary position; the plane they

define is parallel to $H_2 - H_3$, and the forehead vector F is parallel to H_2. The real Fick angles for other head poses are computed using these corrected points, according to (4), (5) and (7). For the sake of simplicity, however, the notation ($'$) specifying the use of corrected points is not used elsewhere in this paper.

Appendix C. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jneumeth.2006.06.016.

References

Allison RS, Eizenman M, Cheung BSK. Combined head and eye tracking system for dynamic testing of the vestibular system. IEEE Trans Biomed Eng 1996;43(11):1073–82.

Aspragathos NA, Dimitros JK. A comparative study of three methods for robot kinematics. IEEE Trans Syst Man Cybern B Cybern 1998;28(2):135–45.

Clarke AH, Ditterich J, Druen K, Schonfeld U, Steineke C. Using high frame rate cmos sensors for three-dimensional eye tracking. Behav Res Methods Instrum Comput 2002;34(4):549–60.

Clarke, AH, Ditterich, J, Druen, K, Schonfeld, U, Steineke, C. The chronos eye tracker: description and verification study; 2003.

Clarke AH, Teiwes W, Schrerer H. Videooculography—an alternative method for measurement of three-dimensional eye movements. In: Schmid R, Zambarbieri D, editors. Oculomotor control and cognitive processes. Amsterdam: Elsevier; 1991. p. 431–43.

Collewijn H, Van der Steen J, Ferman L, Jansen TC. Human ocular counterroll: assessment of static and dynamic properties from electromagnetic scleral coil recordings. Exp Brain Res 1985;59(1):185–96.

Crawford JD, Ceylan MZ, Klier EM, Guitton D. Three-dimensional eye-head coordination during gaze saccades in the primate. J Neurophysiol 1999;81(4):1760–82.

Crawford JD, Martinez-Trujillo J, Klier EM. Neural control of three-dimensional eye and head movements. Curr Opin Neurobiol 2003;13(6):655–62.

Crawford JD, Vilis T. Axes of eye rotation and listing's law during rotations of the head. J Neurophysiol 1991;65(3):407–23.

Denavit J, Hartenberg RS. A kinematic notation for the lower-pair mechanism based on matrices. ASME J Appl Mech 1955:215–21.

Fick A. Die bewegungen des menschlichen augapfels. Zeitschrift für rationelle Medizin 1874;4:109–28.

Goldman R. Intersection of two lines in three-space. In: Glassner AS, editor. Graphics Gems, vol. I. San Diego: San Diego Academic Press; 1990. p. 304.

Han YH, Kumar AN, Reschke MF, Somers JT, Dell'Osso LF, Leigh RJ. Vestibular and non-vestibular contributions to eye movements that compensate for head rotations during viewing of near targets. Exp Brain Res 2005;165(3):294–304.

Harris L, Beykirch K, Fetter M. The visual consequences of deviations in the orientation of the axis of rotation of the human vestibulo-ocular reflex. Vision Res 2001;41(25–26):3271–81.

Haslwanter T. Mathematics of three-dimensional eye rotations. Vision Res 1995;35(12):1727–39.

Haslwanter, T. Computational and experimental aspects of rotatory eye movements in three dimensions. Habilitationsschrift; 2000.

Haslwanter T, Moore ST. A theoretical analysis of three-dimensional eye position measurement using polar cross-correlation. IEEE Trans Biomed Eng 1995;42(11):1053–61.

Hatamian M, Anderson DJ. Design considerations for a real-time ocular counterroll instrument. IEEE Trans Biomed Eng 1983;30(5):278–88.

Lisberger SG, Evinger C, Johanson GW, Fuchs AF. Relationship between eye acceleration and retinal image velocity during foveal smooth pursuit in man and monkey. J Neurophysiol 1981;46(2):229–49.

Maxwell EA. General homogeneous coordinates in space of three dimensions. Cambridge, UK: Cambridge University Press; 1951.

Moore ST, Curthoys IS, McCoy SG. Vtm-an image-processing system for measuring ocular torsion. Comput Methods Programs Biomed 1991;35(3):219–30.

Moore ST, Haslwanter T, Curthoys IS, Smith ST. A geometric basis for measurement of three-dimensional eye position using image processing. Vision Res 1996;36(3):445–59.

Nakayama K. Photographic determination of the rotational state of the eye using matrices. Am J Optom Physiol Opt 1974;51(10):736–42.

Newman R, Matsumoto Y, Rougeaux S, Zelinsky A. Real-time stereo tracking for head pose and gaze estimation. In: Crowley JL, editor. 4th IEEE international conference on automatic face and gesture recognition; 2000.

Orban de Xivry JJ, Bennett SJ, Lefevre P, Barnes GR. Evidence for synergy between saccades and smooth pursuit during transient target disappearance. J Neurophysiol 2006;95(1):418–27.

Ott D, Gehle F, Eckmiller R. Video-oculographic measurement of 3-dimensional eye rotations. J Neurosci Methods 1990;35(3):229–34.

Parker JA, Kenyon RV, Young LR. Measurement of torsion from multitemporal images of the eye using digital signal processing techniques. IEEE Trans Biomed Eng 1985;32(1):28–36.

Robinson DA. A method of measuring eye movement using a scleral search coil in a magnetic field. IEEE Trans Biomed Eng 1963;10:137–45.

Schreiber K, Haslwanter T. Improving calibration of 3-d video oculography systems. IEEE Trans Biomed Eng 2004;51(4):676–9.

Shih SW, Liu J. A novel approach to 3-d gaze tracking using stereo cameras. IEEE Trans Syst Man Cybern B Cybern 2004;34(1):234–45.

Tweed D, Cadera W, Vilis T. Computing three-dimensional eye position quaternions and eye velocity from search coil signals. Vision Res 1990;30(1):97–110.

Tweed D, Vilis T. Implications of rotational kinematics for the oculomotor system in three dimensions. J Neurophysiol 1987;58(4):832–49.

Van Opstal J. Representation of eye position in three dimensions. In: Berthoz A, editor. Multisensory control of movement. Oxford: Oxford University Press; 1993. p. 27–41.

Vieville T, Masse D. Ocular counter-rolling during active head tilting in humans. Acta Otolaryngol 1987;103(3–4):280–90.

Wang JG, Sung E. Study on eye gaze estimation. IEEE Trans Syst Man Cybern B Cybern 2002;32(3):332–50.

Yoo DH, Chung MJ. A novel non-intrusive eye gaze estimation using cross-ratio under large head motion. Comput Vision Image Understand 2005;98(1):25–51.

Yu LH, Eizenman M. A new methodology for determining point-of-gaze in head-mounted eye tracking systems. IEEE Trans Biomed Eng 2004;51(10):1765–73.

Yuksel D, Optican LM, Lefevre P. Properties of saccades in duane retraction syndrome. Invest Ophthalmol Vis Sci 2005;46(9):3144–51.

Zhu D, Moore ST, Raphan T. Robust pupil center detection using a curvature algorithm. Comput Methods Programs Biomed 1999;59(3):145–57.