

# A novel method for measuring gaze orientation in space in unrestrained head conditions

**Benedetta Cesqui**

Laboratory of Neuromotor Physiology,  
IRCCS Santa Lucia Foundation, Rome, Italy



**Rolf van de Langenberg**

Laboratory of Neuromotor Physiology,  
IRCCS Santa Lucia Foundation, Rome, Italy  
Institute of Human Movement Sciences and Sport,  
Zürich, Switzerland



**Francesco Lacquaniti**

Laboratory of Neuromotor Physiology,  
IRCCS Santa Lucia Foundation, Rome, Italy  
Department of Systems Medicine–Neuroscience Section  
and Center of Space Biomedicine, University of Rome,  
Tor Vergata, Rome, Italy



**Andrea d'Avella**

Laboratory of Neuromotor Physiology,  
IRCCS Santa Lucia Foundation, Rome, Italy



Investigation of eye movement strategies often requires the measurement of gaze orientation without restraining the head. However, most commercial eye-trackers have low tolerance for head movements. Here we present a novel geometry-based method to estimate gaze orientation in space in unrestricted head conditions. The method combines the measurement of eye-in-head orientation—provided by a head-mounted video-based eye-tracker—and head-in-space position and orientation—provided by a motion capture system. The method does not rely on specific assumptions on the configuration of the eye-tracker camera with respect to the eye and uses a central projection to estimate the pupil position from the camera image, thus improving upon previously proposed geometry-based procedures. The geometrical parameters for the mapping between pupil image and gaze orientation are derived with a calibration procedure based on nonlinear constrained optimization. Additionally, the method includes a procedure to correct for possible slippages of the tracker helmet based on a geometrical representation of the pupil-to-gaze mapping. We tested and validated our method on seven subjects in the context of a one-handed catching experiment. We obtained accuracy better than  $0.8^\circ$  and precision better than  $0.5^\circ$  in the measurement of gaze orientation. Our method can be used with any video-based eye-tracking system to

investigate eye movement strategies in a broad range of naturalistic experimental scenarios.

## Introduction

Oculography has been extensively used to describe eye movement strategies during different human behaviors and to obtain insight into the mechanisms underlying visuo- or vestibulo-oculomotor coordination. For example, eye movement characteristics have been used to reveal subject intentions (Bekkering & Neggers, 2002; Land & Hayhoe, 2001; Snyder, Batista, & Andersen, 2000), visual strategies when intercepting moving objects or when exploring the environment (Zago, McIntyre, Senot, & Lacquaniti, 2009), and to characterize the ability to predict target motion features during pursuit (Land, 2012; Spering & Montagnini, 2011). Oculography has also been used for diagnosis of neurological and vestibular disorders (Anastasopoulos, Kimmig, Mergner, & Psilas, 1996; Jaafari et al., 2011; Warabi, Kase, & Kato, 1984) and in robotic and virtual reality interfaces (Abbott & Faisal, 2012; Lee, Woo, Kim, Whang, & Park, 2010).

Citation: Cesqui, B., van de Langenberg, R., Lacquaniti, F., & d'Avella, A. (2013). A novel method for measuring gaze orientation in space in unrestrained head conditions. *Journal of Vision*, 13(8):28, 1–22, <http://www.journalofvision.org/content/13/8/28>, doi:10.1167/13.8.28.

Most commercial eye trackers—either based on dual search coil technique (Collewijn, Van der Steen, Ferman, & Jansen, 1985; Robinson, 1963) or video image processing—only measure gaze orientation in space under restrained head conditions or with low tolerance for head movements, hence preventing their use in naturalistic experimental scenarios. The dual search coil technique, for example, measures eye rotations by means of small coils embedded in a modified contact lens and in the magnetic fields generated by two larger coils oriented horizontally and vertically in space. The head is located in the center of those two magnetic fields, so that when the eye moves, the orientation of the coil changes with respect to the fields. This method takes into account small head rotations but not head translations in the estimation of gaze-in-space orientation. In addition to its restrictions on head movement, the dual search coil technique might be impractical in many contexts due to its invasive nature and rather large space requirements.

Video-based eye trackers, on the other hand, measure gaze orientation from the image of the pupil recorded by a camera pointing at the eye. Each camera senses the light, typically infrared, reflected by the eye, and uses the contrast of the pupil image with respect to that of the iris to locate the pupil center. When the cameras and the infrared emitters are mounted remotely (Duchowski, 2007), only limited head motion is tolerated, as several studies reported a considerable influence of small head motion on the gaze estimation accuracy (Morimoto & Mimica, 2005). For this reason the head is often restricted with a bite bar or a chin rest, which might be unpractical in several behavioral studies. Thus, head-mounted configurations are better suited for experimental conditions that require unconstrained head motion. In these systems the cameras and the infrared LED emitters are mounted on a helmet worn by the subject, and head position, measured with respect to the stimulus display by means of an optical tracking system, is combined with the eye-in-head coordinates to determine gaze direction (Hayhoe, McKinney, Chajka, & Pelz, 2012; Pelz, Hayhoe, & Loeber, 2001). Some systems include an additional camera to capture the scene as seen by the subject, allowing to track gaze target in the scene (Land & McLeod, 2000; Land & Tatler, 2001). However, as head-mounted systems require a tight fixation of a helmet on the subject's head, they cannot be used continuously for long periods. A recent study has reported comparable performance with the two search coil and video-based techniques (Kimmel, Mammo, & Newsome, 2012; van der Geest & Frens, 2002). However, given its mobility and less invasive nature, head-mounted video-based systems are often preferred over search coils and video-based remote systems. Despite the broad applicability of these eye-tracking

systems, the mapping of eye-in-head orientation to gaze-in-space is still a delicate issue.

Recently, Ronsse, White, and Lefevre (2007) extended the geometry-based approach developed by Moore, Haslwanter, Curthoys, and Smith (1996) for the measurement of gaze orientation estimation under head-fixed condition, to the case of unrestricted head movements. The method was developed using a video-based eye tracker and requires the integration of head position and orientation in an earth-fixed reference frame measured with a motion tracking device and eye orientation in a head-fixed reference frame measured with a head-mounted eye tracker. After deriving the geometric relationships expressing gaze as a function head pose and eye orientation in the head, a calibration procedure was developed to estimate the underlying geometrical parameters. Three important assumptions were made. First, a simpler orthographic projection instead of a perspective projection was used to determine the pupil position in space from its image on the eye tracker focal plane. The advantage of this assumption is a reduction in the number of variables that need to be computed, yet it holds only when the distance between the eye center and the camera focal plane is several order of magnitude larger than the eye radius, or for smaller eye movements (Moore et al., 1996; Nakayama, 1974). Second, the optical axis of the camera was assumed to be aligned with gaze direction when subjects were in *primary position*, i.e., looking straight ahead, and to pass through the eye center. As a consequence, the rotational matrix between a reference frame attached to the focal plane of the camera and the head-fixed reference frame reduces to a rotation of the camera around its optical axis, decreasing the number of parameters and thus reducing the computational cost of the procedure. Ronsse and colleagues stated that this assumption could be satisfied with a proper adjustment of the eye tracker cameras. Indeed, Moore and colleagues demonstrated that horizontal and vertical components of the rotational matrix within  $5^\circ$  could be ignored, their effect being captured by other calibration parameters (Moore et al., 1996). While this approach is reasonable in the case of eye trackers that use a mirror to project the eye image onto the camera plane, other systems would require mounting the cameras almost at eye height, with a dramatic reduction of subject's field of view. A third assumption of Ronsse and colleagues was that the mapping of pupil pixel coordinates to eye rotation angles was the same for  $x$  and  $y$  axes of the camera's charged-coupled device (CCD). This assumption is problematic for two reasons. First, pixels are not square in most commercial digital cameras so that identical  $x$  and  $y$  pixel displacements correspond to different actual displacements of the pupil image. Second, eye kinematics cannot be accounted for by a purely rotational model

(Fry & Hill, 1962), which assumes the eye to be a perfect sphere and the pupil to rotate about the sphere's center as in Moore et al. (1996) and Ronsse et al. (2007). A better approximation is achieved by decoupling the vertical and horizontal rotational axes (Schreiber & Haslwanter, 2004), i.e., by taking into account the relative shift between the two. This implies that identical yaw and pitch rotations of the eye inside the orbit may produce different  $x$  and  $y$  displacements of the pupil image on the camera's CCD. For both these reasons, introducing an additional  $y$ -gain parameter should markedly improve the calibration outcome.

Here we present a novel geometry-based procedure to estimate gaze orientation in space under unrestricted head movement conditions from the eye tracker pupil position and head pose recordings as a function of an unconstrained configuration of the eye tracker cameras with respect to subject's head and eyes, hence relaxing the first two constraints imposed by Ronsse et al. (2007). The geometrical parameters required by the procedures are derived with a calibration procedure based on nonlinear constrained optimization. The eye tracker employed in the present study to validate the proposed method was a video-based system with two cameras mounted on a wearable helmet to measure pupil position. The advantage with respect to the use of ground-based video eye-trackers is that subjects can move freely in space. However, an important disadvantage of such head-mounted devices is that the helmet—and hence the eye tracking cameras—may move with respect to the head and thereby invalidate system calibration. To account for such an occurrence, we developed a *drift correction* procedure to adjust all model parameters potentially altered by helmet displacement. This approach guarantees accurate gaze estimation throughout an experimental session without the need to repeat the entire calibration procedure. This approach could be helpful also in the case of remotely mounted cameras to alleviate the need for bite bars, hence allowing more comfortable chin rests. The method was tested and validated during a one-handed catching experiment similar to those reported previously (Cesqui, d'Avella, Portone, & Lacquaniti, 2012). A second experimental test in which the helmet slippage was manually induced by the experimenter was also carried out to systematically evaluate the *drift correction* procedure.

## Methods

The section is organized as follows. In the “Theoretical background” subsection we present the procedure used to estimate gaze orientation in space from the

geometrical relationships between the three-dimensional (3-D) coordinates of gaze expressed in a world reference system and the corresponding two-dimensional (2-D) image coordinates of the projection of the pupil-center on the tracker's camera. In the same subsection we also describe the geometrical relationships underlying the drift correction procedure. In the “Parameter estimation procedure” subsection we describe the procedures used to initialize both the calibration and drift correction algorithms. The experimental procedures used to test the proposed method are presented in the “Experimental procedure” section. Finally the tests performed to validate the method are described in the “Calibration and drift correction validation” subsection.

## Theoretical background

The method combines eye-in-head orientation measured by an eye tracker with head position in space measured with a motion tracking system to estimate gaze-in-space orientation. The algorithm steps are reported in the flow chart on the left of Figure 1. Following the black arrow downward, the  $(x, y)$  coordinates of the pupil position on the camera focal plane measured by a video-based eye-tracking system are first mapped into the 3-D coordinates of the pupil center in a reference frame centered in the eye tracker camera plane and then transformed into a second coordinate system fixed with respect to the head and centered in the eye. Finally, the resulting eye-in-head orientation vector is transformed into a ground-based coordinate system using the information on head-in-space orientation provided by a motion-tracking system. To perform these transformations, the relative positions and orientations of the tracker camera with respect to the eye, and of the eye reference frame with respect to the head, must be known. To this aim, we developed a calibration procedure to estimate all the parameters underlying the geometrical configuration system.

### ***Geometrical relationships between gaze orientation in space and pupil coordinates in the camera image plane***

Five different reference frames are involved in the problem. In the following text the versors (unit length vectors) of the axes defining a reference frame are indicated with bold lowercase letters. Vectors representing the coordinates of points in 3-D space are indicated with underlined uppercase letters, with a superscript specifying the reference frame used to define those coordinates (i.e.,  $\underline{P}^C$ ).

1.  $[e_1, e_2, e_3]$  is the reference frame attached to the camera with the origin at the intersection of the

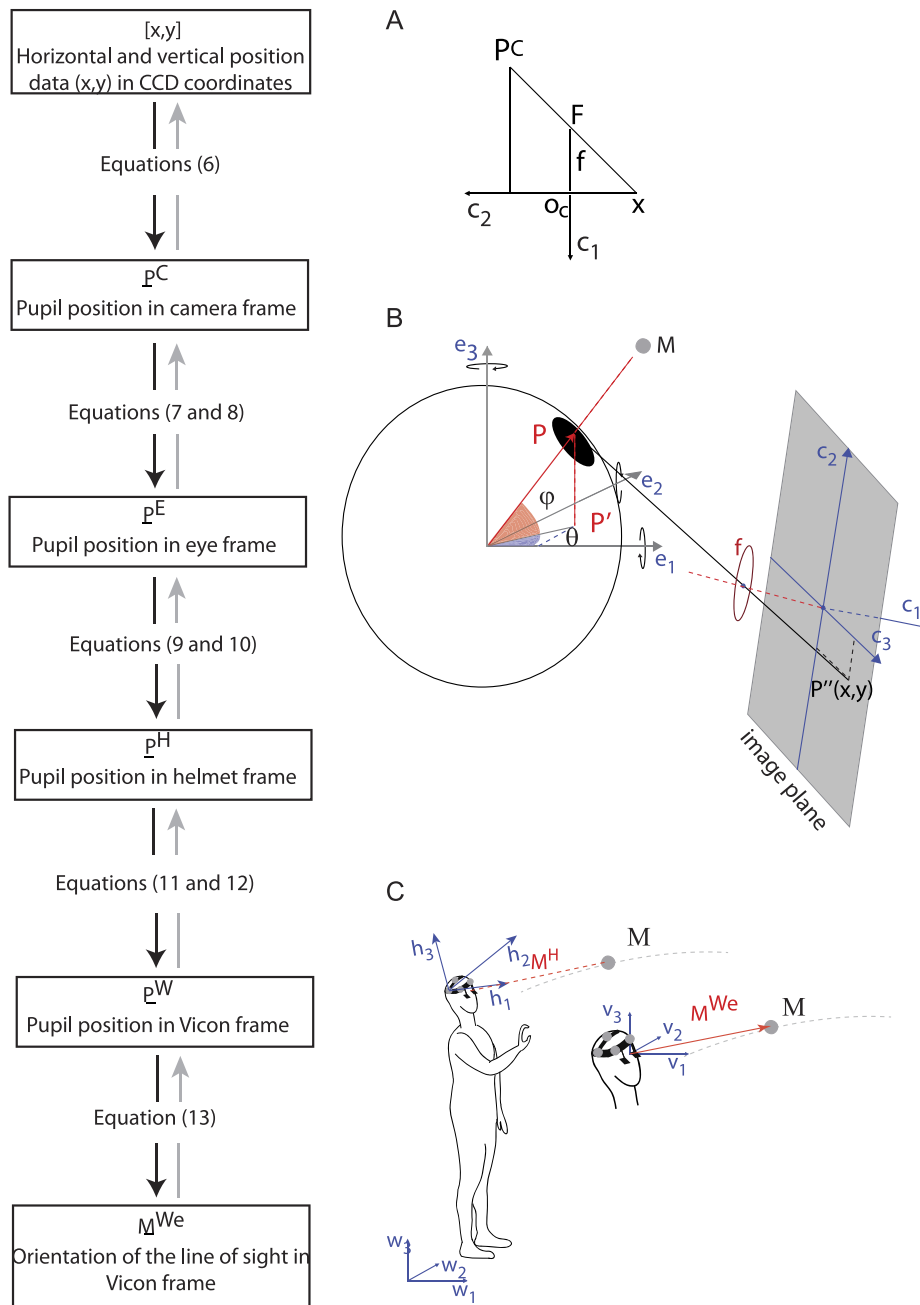


Figure 1. Coordinate system transformations. Left: Flow chart of the mapping between camera coordinates and gaze orientation. The gray arrows (upward pathway) describe the transformation of the recorded target marker position in world coordinate frame into the position of the projection of center of the pupil in the camera reference frame; the black arrows (downward pathway) describe the inverse process. Right: Illustration of the different reference frames involved in the transformations: (A) schematic representation of the perspective projection of the pupil position on the camera x-axis; (B) camera reference frame  $[c_1, c_2, c_3]$  configuration with respect to the eye reference frame; (C) target marker position vector defined with respect to a reference frame oriented as the world reference frame  $[w_1, w_2, w_3]$  and centered in the eye (i.e.,  $M^{We}$ ), and with respect to the helmet reference frame (i.e.,  $M^H$ ).

camera focal plane with the optical axis of the camera lens. The  $c_1$  axis runs along the optical axis, pointing away from the lens, assuming that the camera plane is perpendicular to the optical axis of the lens. The  $c_2$  and  $c_3$  axes are respectively the  $x$  and  $y$  axis of the camera image plane.

2.  $[e_1, e_2, e_3]$  is the reference frame attached to the eye orbit with the origin in the center of the eye. The orientation of the axes are defined while the subject looks straight ahead at a far target at eye height (*the primary position*):  $e_1$  points out of the face and is parallel to the line of sight,  $e_2$  is parallel to the

intraocular axis, and the  $e_2$ - $e_3$  plane is parallel to the subject's frontal plane.

3.  $[\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3]$  is a reference frame attached to the eye-tracker's *helmet* and is defined by the position of three noncollinear points on the helmet, indicated as  $\underline{\mathbf{M}}_1$ ,  $\underline{\mathbf{M}}_2$ , and  $\underline{\mathbf{M}}_3$  (Figure 1B). In particular, markers were applied on the eye tracker such that the vector between  $\underline{\mathbf{M}}_1$  and  $\underline{\mathbf{M}}_2$  was roughly collinear with  $e_1$ , and the vector between  $\underline{\mathbf{M}}_2$  and  $\underline{\mathbf{M}}_3$  was roughly collinear with  $e_2$ . The origin of the frame is coincident with  $\underline{\mathbf{M}}_1$ . Overall, the head frame axes are defined as follows:

$$\begin{aligned} \mathbf{h}_1 &= (\underline{\mathbf{M}}_2 - \underline{\mathbf{M}}_1) / \|\underline{\mathbf{M}}_2 - \underline{\mathbf{M}}_1\| \\ \mathbf{h}_3 &= \left( \mathbf{h}_1 \times (\underline{\mathbf{M}}_3 - \underline{\mathbf{M}}_1) \right) / \|\mathbf{h}_1 \times (\underline{\mathbf{M}}_3 - \underline{\mathbf{M}}_1)\| \\ \mathbf{h}_2 &= \mathbf{h}_3 \times \mathbf{h}_1 \end{aligned} \quad (1)$$

where  $\times$  is the vector product and  $\|\cdot\|$  is the Euclidian norm. Under the assumption that the helmet is strapped to the skull, and hence is perfectly stable and moved exactly with the subject head, the position and the orientation of the  $[\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3]$  coincide with also the pose of the head in space.

4.  $[\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3]$  is the world reference frame, i.e., the reference frame of the tracker system (a *Vicon* system in our experimental tests; see below).
5.  $[\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3]$  is the skull reference frame with origin in the center of the skull and oriented as  $[\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3]$ , (i.e.,  ${}^H\mathbf{R}_S = \mathbf{I}$ ) and used to derive the drift correction procedure.

Hereafter, we indicate with the C, E, H, W, and S superscripts respectively the camera, eye, helmet, world, and the skull reference frames. The  $3 \times 3$  rotation matrices and the  $3 \times 1$  translation vectors are represented with indexed uppercase bold characters. For example,  ${}^E\mathbf{T}_C$  is the translation vector that brings the origin of the camera reference frame ( $\underline{\mathbf{Q}}_C$ ) into the origin of the eye reference frame, that is, the coordinates of  $\underline{\mathbf{Q}}_C$  in the  $[e_1, e_2, e_3]$  reference frame. In general, a 3-D vector representing a point in space,  $\underline{\mathbf{P}}^A$ , whose components are defined with respect to a generic coordinate system  $A$ , can be represented in a different coordinate system  $B$ , i.e.,  $\underline{\mathbf{P}}^B$ , according to the following linear transformation:

$$\underline{\mathbf{P}}^B = {}^B\mathbf{R}_A \cdot \underline{\mathbf{P}}^A + {}^B\mathbf{T}_A, \quad (2)$$

where  ${}^B\mathbf{R}_A$  is the rotation matrix constructed with the versors of the  $A$  frame expressed in the coordinates of the  $B$  frame, and  ${}^B\mathbf{T}_A$  is the translation vector expressing the position of the center of the  $A$  frame in the coordinates of the  $B$  frame. Fick angles were used to define each rotation matrix (Haslwanter & Moore,

1995). According to this convention, the orientation of a vector in space is defined by the composition of a horizontal rotation by an angle  $\theta$ , followed by a vertical rotation by an angle  $\varphi$ , followed by torsion by an angle  $\psi$ . The expanded form of the rotation matrix defined by the Fick angles can be found elsewhere (Haslwanter & Moore, 1995; Moore et al., 1996; Ronsse et al., 2007). The transpose of the matrix  $\mathbf{R}$  is indicated by  $\mathbf{R}'$ .

### **From pupil coordinates recorded in the camera image plane to the orientation of the gaze vector in space**

The 3-D coordinates of the pupil center in the  $[\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3]$  frame, i.e.,  $\underline{\mathbf{P}}^W$ , are estimated for each eye from the coordinates  $(x_{raw}, y_{raw})$  provided by an eye-tracking system and derived from the position of the projection of center of the subject's pupil on the eye tracker's camera image plane  $(x, y)$  according to the flow chart of Figure 1 (leftmost panel, downward black arrows' direction). The  $(x_{raw}, y_{raw})$  coordinates saved in the data records are usually the result of some internal linear transformations applied by the eye-tracker data acquisition software to the original horizontal and vertical coordinates  $(x_{meas}, y_{meas})$  of the center of the pupil on the camera image plane measured by the eye tracker. In general, the pupil position (P), i.e., the coordinates of P' in Figure 1, *panel B*, is referred to an arbitrary reference frame system centered in some point of the camera sensor (typically a CCD sensor), according to an initial calibration of the eye-tracker. If the camera does not move with respect to the eye, any change of the pupil position on the CCD represents a pure rotation of the eye. However, most of the available eye-trackers in commerce rely on the pupil-corneal reflection (P-CR) technique (Morimoto & Mimica, 2005), which determines the  $(x_{meas}, y_{meas})$  coordinates as the difference vector between the pupil position on the CCD and the first surface corneal reflection (CR) of an illumination source. Notably, when used in the P-CR mode, the rotational gain, i.e., the CR displacement for a unit displacement of the pupil, is approximately 0.5, since the CR moves half the distance as the pupil center during an eye rotational movement (Hua, Krishnaswamy, & Rolland, 2006; Li, Munn, & Pelz, 2008). Accordingly, the calibration parameters transforming the  $(x_{meas}, y_{meas})$  coordinates into the  $(x_{raw}, y_{raw})$  data output might be different in the P and P-CR recording modes. Moreover, other additional scaling and translations might be performed on the measured data depending on the specific eye tracker raw data processing. For instance, in the case of the EyeLink-II system (SR Research, Ltd., Mississauga, Ontario, Canada) used in the present study, each data are linearly transformed in order to always be a positive integer ranging from 0 to 30,000 camera units (SR Research, personal communication, February,

2009 and November, 2012). Thus, the relationships between the pupil coordinates on the CCD and the data provided by the eye-tracking system are expressed by:

$$\begin{aligned} x_{raw} &= \alpha_x x_{meas} + x_{off} \\ y_{raw} &= \alpha_y y_{meas} + y_{off} \end{aligned} \quad (3)$$

where the  $\alpha_x$  and  $\alpha_y$  are scaling factors and the  $x_{off}$  and  $y_{off}$  are offset parameters corresponding to the origin of the camera frame (i.e., the intersection of the image plane with the optical axes) on the CCD. Their values might be different across recording modes, and likely to change every time the system proprietary calibration procedure is carried out, but unfortunately the algorithm code is often not available to users. Hence we need to estimate their value with dedicated procedures as described in next sections.

Coordinates measured on the CCD ( $x_{meas}$ ,  $y_{meas}$ ) are expressed in camera units (c.u.), which depend on the size and number of pixels and on the eye-tracker algorithm used to estimate the pupil center. Since coordinates in the camera reference frame introduced above are estimated in meters, we need to know a factor  $u$ , converting camera units into meters, to compute the pupil position ( $x$ ,  $y$ ) with respect to the camera coordinates frame:  $x[m] = x_{meas} [c.u.] \times u$ . Such conversion factor  $u$  is often not reported by manufacturers. If  $f$  is the known focal length of the camera lens in meters, which we assume to be provided by the manufacturer (7.5 mm in our system), and  $k$ , the camera focal length expressed in camera units, can be estimated with a specific procedure (see below), then  $u$  can also be estimated as  $f/k$ .

The relation between the projection on the image plane ( $\underline{P}''$  in Figure 1B) measured in meters and its relative ( $x_{raw}$ ,  $y_{raw}$ ) coordinate could be formulated from Equation 3 as:

$$\begin{aligned} x &= \frac{(x_{raw} - x_{off})}{\alpha} u \\ y &= \frac{(y_{raw} - y_{off})}{\alpha g} u \end{aligned} \quad (4)$$

where  $g$  is the y-gain parameter, and  $\alpha$  is the scaling factor parameter applied to the ( $x_{meas}$ ,  $y_{meas}$ ) coordinates, so that:  $g = \alpha_y/\alpha_x$ , and  $\alpha = \alpha_x$ , with  $\alpha_x$  and  $\alpha_y$  representing scaling factors in Equation 3.

According to the perspective projection on the camera plane shown in Figure 1A, the relation between the pupil position expressed in the camera reference frame ( $\underline{P}^C$ ) and  $\underline{P}''$ , is given for the x-axis (similar equations apply to y-axis), by:

$$\frac{x}{f} = \frac{x - P_2^C}{P_1^C} \quad (5)$$

where  $P_1^C$  and  $P_2^C$  are the coordinates of the pupil position in camera coordinates along the  $\mathbf{c}_1$  and  $\mathbf{c}_2$  axis, and  $x$  is the coordinate of  $\underline{P}''$  along the x-axis defined in Equation 4). Thus, the pupil coordinates expressed with respect to the camera frame  $\underline{P}^C$  can be computed from  $\underline{P}''$  (assuming that the eye is a sphere of radius  $r_e$ ) by solving the following system of equations:

$$\begin{cases} P_2^C = \frac{x}{f}(f - P_1^C) \\ P_3^C = \frac{y}{f}(f - P_1^C) \\ (P_1^C - T_1)^2 + (P_2^C - T_2)^2 + (P_3^C - T_3)^2 = r_e^2 \end{cases} \quad (6)$$

where  $\underline{T}$  is the vector expressing the center of the camera frame in the eye coordinate system, i.e.,  ${}^E T_C$ .

According to Equation 2, the pupil center in the [ $\mathbf{e}_1$ ,  $\mathbf{e}_2$ ,  $\mathbf{e}_3$ ] reference frame, i.e.,  $\underline{P}^E$ , is given by:

$$\underline{P}^E = \mathbf{R}_C \cdot \underline{P}^C + {}^E T_C \quad (7)$$

where  ${}^E \mathbf{R}_C$  is the rotation matrix of the camera frame with respect to the eye frame and  ${}^E T_C$  is the translation vector of the center of the camera reference frame with respect to the center of the eye reference frame (i.e., the center of the eye). The eye horizontal (i.e., azimuth),  $\theta$ , and vertical (i.e., elevation),  $\varphi$ , orientation angles, are given by:

$$\begin{aligned} \vartheta^E &= \arctan\left(\frac{P_2^E}{P_1^E}\right), \\ \varphi^E &= \arctan\left(\frac{P_3^E}{\sqrt{(P_1^E)^2 + (P_2^E)^2}}\right). \end{aligned} \quad (8)$$

Similarly, the gaze orientation vector  $\underline{P}^H$  and the corresponding orientation angles,  $\vartheta^H$  and  $\varphi^H$ , expressed with respect to the helmet reference frame are given by:

$$\begin{aligned} \underline{P}^H &= \underline{P}^{He} + {}^H T_E = {}^H \mathbf{R}_E \cdot \underline{P}^E + {}^H T_E \quad (9) \\ \theta^H &= \arctan\left(\frac{P_2^{He}}{P_1^{He}}\right), \\ \varphi^H &= \arctan\left(\frac{P_3^{He}}{\sqrt{(P_1^{He})^2 + (P_2^{He})^2}}\right) \end{aligned} \quad (10)$$

where  $\underline{P}^{He}$  is the vector expressing the position of the pupil in eye coordinates,  $\underline{P}^E$ , in a reference system with the same origin but with the axes rotated as those of the helmet reference frame;  ${}^H \mathbf{R}_E$  and  ${}^H T_E$  are respectively the rotation matrix and the translation vector of the eye reference frame with respect to helmet reference frame.

Finally the gaze orientation angles expressed with respect to the  $W$  reference frame system are given by:

$$\underline{P}^W = \underline{P}^{We} + {}^W T_H = {}^W R_H \cdot \underline{P}^H + {}^W T_H \quad (11)$$

$$\vartheta^{We} = \arctan\left(\frac{P_2^{We}}{P_1^{We}}\right),$$

$$\varphi^{Ve} = \arctan\left(\frac{P_3^{We}}{\sqrt{(P_1^{We})^2 + (P_2^{We})^2}}\right) \quad (12)$$

where  $\underline{P}^{We}$  is the orientation of the line of sight in world coordinates and  ${}^W R_H$  and  ${}^W T_H$  are respectively the rotation matrix and the translation vector of the helmet reference frame with respect to world reference frame. In particular, the rotation matrix  ${}^W R_H$  is computed from the position vectors of  $\underline{M}_1$ ,  $\underline{M}_2$ , and  $\underline{M}_3$  points as specified in Equation 1.

#### From target position in space to pupil coordinates in the camera image plane

According to the flow chart on the left of Figure 1 (upward gray arrows' direction), the projection of the pupil center in the camera image plane can be estimated from the target marker position with respect to the world reference frame,  $\underline{M}^W$ , as follows. The vector from the eye center to the target marker in the world frame reference frame centered on the eye,  $\underline{M}^{We}$ , and its orientation angles (i.e., the orientation of the line of sight) are given by:

$$\underline{M}^{We} = \underline{M}^W - {}^W T_E = \underline{M}^W - ({}^W R_H \cdot {}^H T_E + {}^W T_H) \quad (13)$$

$$\vartheta^W = \arctan\left(\frac{M_2^{We}}{M_1^{We}}\right),$$

$$\varphi^W = \arctan\left(\frac{M_3^{We}}{\sqrt{(M_1^{We})^2 + (M_2^{We})^2}}\right) \quad (14)$$

Next,  $\underline{M}^E$ , i.e., the target marker expressed in eye coordinates, is computed by means of the inverse of functions Equations 7 and 9:

$$\underline{M}^H = {}^W R'_H \cdot \underline{M}^{We} - ({}^W R'_H \cdot {}^W T_H), \quad (15)$$

and

$$\underline{M}^E = {}^H R'_E \cdot \underline{M}^H - ({}^H R'_E \cdot {}^H T_E). \quad (16)$$

By scaling the  $\underline{I}^E$  vector to the length of the eye radius, the position of the pupil center with respect to eye reference frame,  $\underline{P}^E$  is estimated as:

$$\underline{P}^E = \frac{\underline{I}^E}{\|\underline{I}^E\|} r_e. \quad (17)$$

Thus, by inverting the functions of Equations 5 and 3, the  $\underline{P}^E$  vector is first transformed into the camera reference frame, returning the estimate of the pupil center in the camera coordinates, i.e., the  $\underline{P}^C$  vector, and finally projected onto the camera's image plane through a perspective projection, yielding the estimate of the coordinates of the pupil image center  $(\hat{x}, \hat{y})$ .

#### Drift correction geometry

Video-based eye-tracking systems rely on locating the pupil center on the camera image plane. Once the system is calibrated, the position of the pupil can be transformed into eye-in-head orientation. However, this measurement suffers for potential drawbacks. For instance the tracker accuracy drops dramatically if the camera moves with respect to the eye after the system calibration. It has been reported that a 0.1-mm displacement introduces an artifact of about  $1^\circ$  in the output gaze orientation (Li et al., 2008). Even if the cameras are securely attached on a head-mounted tracker, the helmet may move with respect to the subject's head and eyes. One option is to use the tracker in the P-CR recording mode, and hence to refer the pupil position to the position of the CR of a dedicated illumination source. Since the CR and P images on the CCD change in unison when the camera moves with respect to the eye, any change in the vector difference between the center of the pupil and the center of the corneal reflection should be solely ascribed to eye rotational movement. For this reason the P-CR mode is often preferred to the P mode in most of the applications (Morimoto & Mimica, 2005). Nevertheless this approach cannot compensate for an accidental helmet slippage, which—in addition to a camera translation—also involves a displacement of the illumination source with respect to the eye and hence a change of the CR position on the CCD. Moreover, small changes of the P-CR vector occur even in absence of eye movements when the camera moves.

To overcome these limits, and to minimize possible errors due to helmet slippage, a further procedure was developed to correct the rotation matrices and translation vectors parameterizing the transformation of the coordinates of the pupil center into gaze orientation in the helmet frame (i.e.,  ${}^H R_E$ ,  ${}^H T_E$ ,  ${}^E R_C$ ,  ${}^E T_C$ ), taking into account helmet slippage. In the following derivation we will first consider the situation in which eye position is tracked in pupil mode, and we will then discuss the implications of using the P-CR recording mode.

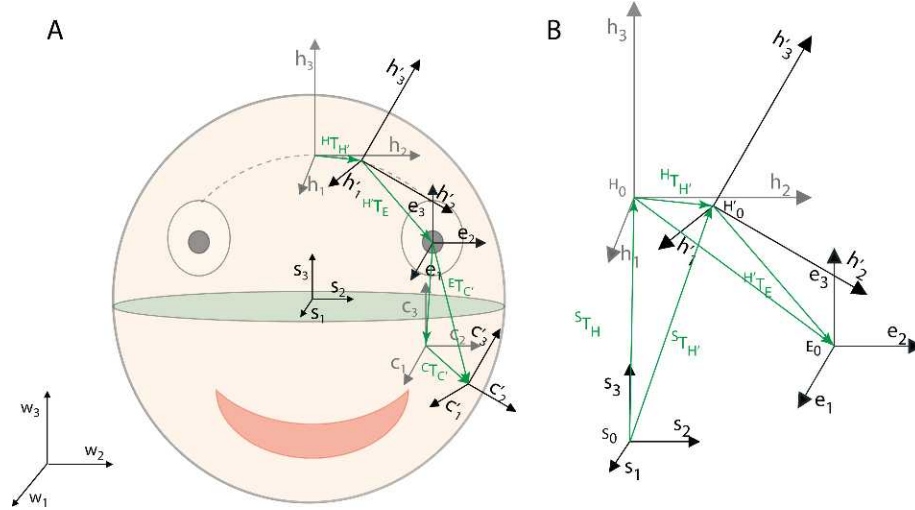


Figure 2. Drift correction geometry. (A) Schematic representation of the configuration of all the reference frames involved in the problem:  $[h_1', h_2', h_3']$  is the helmet drifted reference frame after helmet displacement,  $[c_1', c_2', c_3']$  is the camera drifted reference frame after the same displacement (head-mounted system slippage),  $[s_1, s_2, s_3]$  is the face reference frame centered in the skull center and oriented as  $[h_1, h_2, h_3]$ ; (B)  $[h_1', h_2', h_3']$  reference frame with respect to the  $[h_1, h_2, h_3]$  and the  $[s_1, s_2, s_3]$  reference frames. Helmet displacement was assumed to be only due to a rotation with respect to the center of the skull.

Two additional reference frames, derived from those defined above, are required for this derivation (see Figure 2):

1.  $[h_1', h_2', h_3']$  is the helmet drifted reference frame after the helmet displacement.
2.  $[c_1', c_2', c_3']$  is the camera drifted reference frame after the helmet displacement.

The helmet displacement was assumed to be described by a rotation with respect to the center of the skull. It follows that (Figure 2):

$${}^H\mathbf{R}_E = {}^H\mathbf{R}_H \cdot {}^H\mathbf{R}_E. \quad (18)$$

According to Equation 2, the origin of the head frame in the original head reference frame, i.e.,  ${}^H\mathbf{T}_{H'}$ , is given by:

$${}^H\mathbf{T}_{H'} = {}^H\mathbf{R}_S \cdot {}^S\mathbf{T}_{H'} + {}^H\mathbf{T}_S = {}^S\mathbf{T}_{H'} + {}^H\mathbf{T}_S. \quad (19)$$

where  ${}^H\mathbf{T}_S$  is the origin of the skull reference frame in head coordinates and  ${}^S\mathbf{T}_{H'}$  is the origin of the drifted head frame in skull coordinates. Moreover,

$${}^H\mathbf{T}_S = {}^H\mathbf{R}_{H'} \cdot {}^H\mathbf{T}_S + {}^H\mathbf{T}_{H'} \quad (20)$$

Thus, combining Equations 19 and 20, we obtain:

$${}^H\mathbf{T}_{H'} = {}^H\mathbf{T}_S - {}^H\mathbf{R}_{H'} \cdot {}^H\mathbf{T}_S. \quad (21)$$

The new  ${}^H\mathbf{T}_E$  translation vector is then given by:

$${}^H\mathbf{T}_E = {}^H\mathbf{R}'_{H'} \cdot ({}^H\mathbf{T}_E \cdot {}^H\mathbf{T}_{H'}). \quad (22)$$

By substituting Equation 21 into Equation 22:

$${}^H\mathbf{T}_E = {}^H\mathbf{R}_H \cdot ({}^H\mathbf{T}_E - {}^H\mathbf{T}_S) + {}^H\mathbf{T}_S \quad (23)$$

Since the helmet is a rigid body, after the slippage:  ${}^C\mathbf{R}_C = {}^H\mathbf{R}_H$ . Thus:

$${}^E\mathbf{R}_{C'} = {}^E\mathbf{R}_C \cdot {}^H\mathbf{R}_H \quad (24)$$

$$\begin{aligned} {}^E\mathbf{T}_{C'} &= {}^H\mathbf{R}'_E \cdot {}^H\mathbf{T}_{C'} - {}^H\mathbf{T}_E \\ &= {}^H\mathbf{R}'_E \cdot ({}^H\mathbf{R}_E \cdot {}^E\mathbf{T}_C + {}^H\mathbf{T}_E) - {}^H\mathbf{T}_E. \end{aligned} \quad (25)$$

According to the considerations presented above, when eye tracking is performed in P-CR recording mode, the error introduced by the camera translation with respect to the eye associated with the helmet rotation over the head is already partially compensated. However, the drift correction procedure may be used to modify the calibration parameters compensating for the error associated with the displacement of the illumination source with respect to the eye.

## Parameter estimation procedures

### Calibration procedure for the gaze-to-pupil parameters

A calibration procedure is required to identify all the 21 parameters introduced in the previous section necessary to map a gaze target into camera coordinates of the center of the pupil:

- Four camera gains:  $g^L g^R$  and  $\alpha^L \alpha^R$  parameters for the two eyes;
- Two anthropometric parameters: eye radius ( $r_e$ ) and inter ocular distance ( $IOD$ ).
- Twelve eye-to-camera tracker configuration parameters:



- ${}^E\mathbf{R}_C^L$ ,  ${}^E\mathbf{R}_C^R$ ,  ${}^E\mathbf{T}_C^L$ ,  ${}^E\mathbf{T}_C^R$ : the rotation matrices (each defined by the three Fick rotation angles) and the translation vectors of the eyes' reference frames with respect to the cameras reference frames for both the left (L) and the right (R) eye, indicated with the uppercase superscript letters;
- Three head to eye configuration parameters:
  - ${}^H\mathbf{T}_E$ : the translation vector of the midpoint between the eyes with respect to the center of the head reference frame. The two translation matrices,  ${}^H\mathbf{T}_E^L$  and  ${}^H\mathbf{T}_E^R$ , relative to the left and the right eye, are computed from the  ${}^H\mathbf{T}_E$  vector, by knowing the interocular distance, i.e., the *IOD* parameter. We assumed that the eyes lie along the interocular axis, parallel to the  $\mathbf{h}_2$  axis, so that:

$${}^H\mathbf{T}_E^L = {}^H\mathbf{T}_E + \frac{1}{2}IOD\mathbf{h}_2$$

$${}^H\mathbf{T}_E^R = {}^H\mathbf{T}_E - \frac{1}{2}IOD\mathbf{h}_2$$

The rotation matrix  ${}^H\mathbf{R}_E$  is assumed to be the same for both left and right eyes; it is initialized with data from the static trial as described in the next section and not optimized.

The advantage of this approach is that only four parameters (instead of six) must be estimated with a reduction of computational cost.

The protocol for a *calibration trial* carried out to estimate these parameters is similar to that used in a previous study (Ronsse et al., 2007). The subject is required to gaze at a marker slowly moved by the experimenter within a workspace of interest. The position of the target marker and the head pose extracted from the position of three markers placed on the eye-tracker helmet worn by the subject are recorded by a motion-tracking system. The pupil coordinates on the camera image plane are measured by a video-based eye tracker.

Data are processed in Matlab (Matworks Inc., Natick, MA) using a nonlinear optimization algorithm with constraints (function *fmincon*) that determines the required calibration parameters iteratively by minimizing the error between the coordinates of the center of the pupil image estimated from the spatial position of the target marker ( $\hat{x}$ ,  $\hat{y}$ ) (see previous sections) and those measured from the eye tracker camera images ( $x$ ,  $y$ ):

$$Err = |\hat{x} - x|^2 + |\hat{y} - y|^2. \quad (26)$$

Hereafter, we will refer to this calibration approach as the standard (SND) procedure.

An alternative calibration approach can be derived if the eye tracker provides an estimate of the eye-in-head angles using a proprietary calibration algorithm to map

the pupil position on the CCD into eye-in-head coordinates defined with respect to a head-referenced reference frame. Such eye-in-head coordinates can then be combined with head pose measurement to estimate gaze in space. In the case of the eye tracker used in the present analysis, the EyeLink-II (see “Experimental procedures” below), the  $(x, y)$  head-referenced (HREF) coordinates are recorded in camera units and represent the position of a point on a plane at a given distance (15,000 c.u.) from the eye. These coordinates are independent of the display distance and its resolution. If the eye-in-head angles are provided by the eye-tracker software, the calibration algorithm only needs to determine 10 parameters of the mapping between the helmet reference frame and the eye-in-head reference frame ( ${}^H\mathbf{T}_E$  helmet-to-eye translation vector, the left and right  ${}^H\mathbf{R}_E^L$  and  ${}^H\mathbf{R}_E^R$  helmet-to-eye rotations Fick angles, and *IOD*). Thus, the error between the eye-in-head angles provided by the eye tracker and their estimation derived from the target position according to Equations 13 through 16 is computed first. Then, the required calibration parameters are determined through a nonlinear iterative minimization of such error as for the procedure described above. Hereafter, we will refer to this second calibration approach as HREF. This procedure is less computationally demanding than the SND procedure, since it optimizes a smaller number of parameters and it does not require any assumptions with respect to the camera and the eye-to-camera parameters. However, it relies on a proprietary eye-in-head calibration procedure that may have limitations that are not under the experimenter's control. We performed a dedicated analysis to compare the two approaches.

### Parameter initial values and constraints

The optimization algorithm requires a choice of the solution space for each parameter, that is, a choice of an initial value and of the maximum and minimum of the interval of allowed values. In particular, if the initial values are close to the real values, the optimization is more likely to find the correct solution, i.e., the global minimum of the error function. Below we summarize the procedure applied to initialize all the parameters listed in the previous section, as well as the procedure to select the rotation matrix  ${}^H\mathbf{R}_E$ , which defines the orientation of the eyes in the *primary position* (not optimized by the algorithm).

### Camera parameters

1. The initial value of the focal length parameter expressed in camera unit (i.e.,  $k$ ), used to compute the conversion factor  $u$  (i.e.,  $u = f / k$ ) in Equation 4,

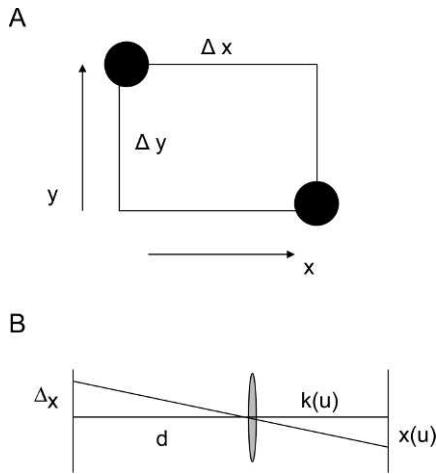


Figure 3. Initial estimation of camera parameters. (A) The focal length in camera units is estimated by focusing each camera on each of the two high-contrast discs sited at known distance  $\Delta_x = \Delta_y$  and known position  $d$  with respect to the camera focal lens; (B) central projection geometry (only the  $x$ -axis is shown) used to extract  $k$  from the EyeLink-II discs recorded positions expressed in camera coordinates  $(x, y)$ .

was estimated as follows. Each camera of the eye tracker was focused on two high-contrast discs (which simulate the eyes pupil), located at a known distance with respect to the camera focal point. The coordinates of the discs centers in the eye tracker camera plane were recorded by the EyeLink-II in P mode. In particular, we used two discs at the opposite corners of a square (Figure 3A). Two different square sizes ( $\Delta$ ) and distances ( $d$ ),  $\Delta = 2$  cm at  $d = 13.2$  cm and  $\Delta = 4$  cm at  $d = 15.3$  cm, were tested, and the results were averaged. According to central projection geometry (Figure 3B):

$$\frac{\Delta}{d} = \frac{x}{k}, \quad (27)$$

where  $x$  is the coordinate of pupil position on the CCD in camera units and  $k$  is the focal length in camera units. It follows that:

$$k = \frac{x \cdot d}{\Delta}. \quad (28)$$

Similarly,

$$k_y = \frac{y \cdot d}{\Delta}. \quad (29)$$

The  $g$  parameter was initialized as  $g = k / k_y$ . For our experimental setup (see below), we found  $k = 75,000$  and set  $g^L = g^R = 0.85$  (retaining two significant digits). The  $\alpha^L$   $\alpha^R$  gain parameters were set to 1.

When using the tracker in the P-CR recording mode (see the “Theoretical background” section), we considered that the rotational gain is different from 1, since the angle measured in P-CR mode is almost half the angle seen in the P mode (Li et al., 2008). To this aim the  $k$  parameter, that is the estimate distance between the camera focal point and its relative image plane expressed in camera units, was initialized as a fraction of the value used in the case of the P mode (i.e., 50,000 c.u.). However, an error in the estimation of the  $k$  parameter could be compensated by the optimization of two gain parameters, i.e., the  $\alpha$  and  $g$ .

Finally the  $(x_{off}, y_{off})$  parameters in Equation 4 should be initialized recording the pupil position when the eye was in the center of the CCD. However, any error in the estimation of these offsets will be compensated by the optimization of the eye-to-camera geometry configuration parameters (i.e.,  ${}^E R_C^L$ ,  ${}^E R_C^R$ ,  ${}^E T_C^L$ ,  ${}^E T_C^R$  parameters).

2.  $r_e$  was initialized to 0.012 m (Marieb, 2001).
3.  ${}^H R_E$  rotation matrix was defined for each subject using the data recorded during a trial (*static trial*) carried out prior to the experimental session, with the subject in the *primary position*, i.e., looking straight ahead at a static distant target of known position. The head pose, extracted from the position of the markers placed on the helmet, and the eye pupil position were recorded. By definition:

$${}^H R_E = {}^W R'_H \cdot {}^W R_E, \quad (30)$$

where:

${}^W R_H$  is the orientation of the head in the reference frame  $[h_1, h_2, h_3]$  with respect to the world reference frame  $[w_1, w_2, w_3]$ , computed from the helmet markers as specified in Equation 1.

${}^W R_E$  is the orientation of the eye reference frame with respect to the world reference frame. In the primary position,  $e_1$  is horizontal (i.e.,  $e_1[3] = 0$ ) and oriented as the vector between the gaze target and the middle point between the eyes lying along the *IOD* axis; the  $e_3$  axis is vertical ( $e_3 = [0 \ 0 \ 1]$ ) and  $e_2$  is given by the vector product  $e_3 \times e_1$ .

The other parameters were measured for each subject after the helmet and the cameras were properly positioned and oriented:

4. *IOD* was measured with a ruler.
5.  ${}^H T_E$  vector was estimated by measuring with a ruler the horizontal, vertical, and lateral distances between the  $\underline{M}_1$  marker and the right eyeball center.
6.  ${}^E R_C^L$ ,  ${}^E R_C^R$ ,  ${}^E T_C^L$ ,  ${}^E T_C^R$  initial values were estimated by measuring the horizontal vertical and lateral distance between the center of the eye and the

Calibration parameters	Variation ranges
$\alpha^L, \alpha^R$ scaling factor	$\pm [0.5 \ 0.5] \text{ u}$
$g^L, g^R$ gain of the camera	$\pm [0.5 \ 0.5] \text{ u}$
Eye radius	$\pm 0.003 \text{ m}$
Intraocular distance	$\pm 0.005 \text{ m}$
$[\theta \ \varphi \ \psi]$ of ${}^E R_C^L$	$\pm [20 \ 20 \ 20]^\circ$
$[\theta \ \varphi \ \psi]$ of ${}^E R_C^R$	$\pm [20 \ 20 \ 20]^\circ$
${}^E T_C^L$	$\pm [0.02 \ 0.02 \ 0.02] \text{ m}$
${}^E T_C^R$	$\pm [0.02 \ 0.02 \ 0.02] \text{ m}$
${}^H T_E$	$\pm [0.01 \ 0.01 \ 0.01] \text{ m}$
$[\theta \ \varphi \ \psi]$ of ${}^H R_H$	$\pm [5 \ 5 \ 5]^\circ$
${}^H T_S$	$\pm [0.01 \ 0.01 \ 0.01] \text{ m}$

Table 1. Ranges of variability admitted by the calibration and drift correction procedures for each estimated parameter and transformation matrices introduced in the Methods. *Note:* In the case of rotational matrices, the ranges are relative to the  $\theta$ ,  $\varphi$ , and  $\psi$  Fick angles—respectively the horizontal, the vertical, and torsional components of the rotations.

approximate position of the center of the focal plane of the eye-tracker camera. Rotational matrix for the left and right eye were then computed from the Fick angles according to the procedure specified in Haslwanter and Moore (1995) and Moore et al. (1996), and initialized so that  $(0, \arctan(d_3/d_1), 180^\circ)$ .

Upper and lower boundaries of variation for each parameter are reported in Table 1.

### Estimate of drift correction parameters

This procedure aims at estimating the rotation matrix  ${}^H R_H$  and the position of the origin of the skull reference frame with respect to the head reference frame,  ${}^H T_S$ , in order to compute, according to Equations 18 and 25, the new rotation matrices and transformation vectors  ${}^H R_E$ ,  ${}^H T_E$ ,  ${}^E R_C$ , and  ${}^E T_C$ .

Similar to the *static trial* described above, during the *drift correction* trial, the subject was asked to remain in the primary position and gaze for a few seconds at a point in space of known position, while both the head pose and the eye pupil coordinates were recorded.

A nonlinear optimization algorithm (Matlab function *fmincon*) was then used to find the solution for the transformation matrix,  ${}^H R_H$ , and the translation vector,  ${}^H T_S$ , with the minimum error starting from a number of initial conditions randomly selected in the solution space. In particular, the Fick angles defining the helmet rotation were initially assumed to be null and were allowed to vary within a range of  $\pm 5^\circ$ ; the center of the skull reference frame was initialized according to an estimate value of the center of the skull with respect to the head reference frame origin.

Upper and lower boundaries of variation are reported in Table 1.

## Experimental procedures

The proposed method was developed and tested in the context of an experimental investigation of the visuomotor control strategies adopted for intercepting a flying ball. In the following sections we report an overview of the experimental set up and a description of the eye-tracker calibration and drift correction procedures carried out for testing the method with data collected from standing, unrestrained subjects during a ball catching experimental session. A further calibration session was carried out for evaluating the drift correction procedure in controlled conditions with data collected from a sitting subject with the head restrained by a chin rest. In these conditions, helmet slippage was simulated by manually displacing the eye tracker over the subject's head.

### Participants and experimental protocol

Seven right-handed subjects (six males and one female, labeled S<sub>1</sub> through S<sub>7</sub>), between 22 and 42 years old ( $30 \pm 6$ , mean  $\pm SD$ ) participated to the study. Two of them were authors of this study. They all had normal or corrected-to-normal vision, were informed about the procedure and the aims of the study, and gave their written informed consent to participate in the experiment. All procedures were approved by the Ethical Review Board of Santa Lucia Foundation and adhered to the tenets of the Declaration of Helsinki.

The experimental protocol was similar to that reported in a previous study (Cesqui et al., 2012). Briefly, participants were asked to stand, look straight ahead (i.e., in the *primary position*), and be ready to catch a ball projected by a dedicated launching apparatus, Figure 4A).

Eight different conditions of ball flight were tested. During the launch session, for each flight condition subjects performed one block of at least 10 trials, for a total of eight blocks. Prior to the beginning of experiment, subjects performed a *static trial* followed by two *calibration trials*, aimed at the extraction of all parameters required for gaze estimation according to the procedures described above. A third *calibration trial* was also performed at the end of the experiment. Finally, a *drift correction trial* was performed at the end of each block. Details of the experimental sequence are reported in Table 2.

### Data acquisition

During the experiment, the eye movements were recorded using the EyeLink-II video-based eye tracker

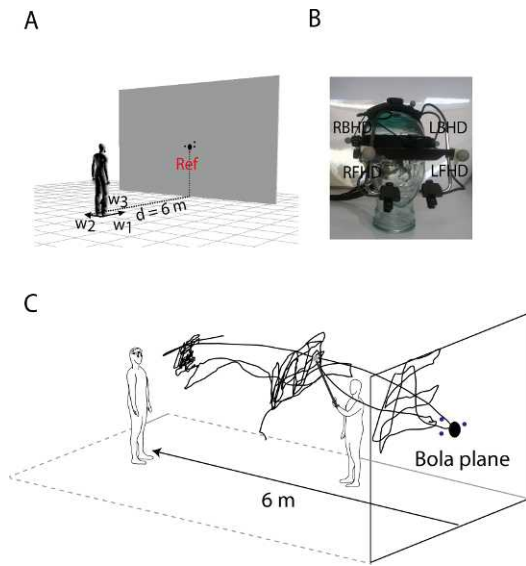


Figure 4. Calibration experimental set up and procedure. (A) Schematic representation of the experiment carried out in our laboratories: a subject stood in front of a large screen (Bola plane) with a small hole through which balls are projected. (B) Vicon retro-reflective markers placement on the EyeLink-II tracker to measure head position. (C) Schematic representation of a typical *calibration trial* (subject  $S_4$ ). The subject was required to gaze at a Vicon marker placed at the end of a bar slowly moved by the experimenter within the entire region of space covered by the ball trajectories in the catching experiment.

(SR Research, Ltd., Mississauga, Ontario, Canada). The spatial position of the markers placed on the eye-tracker helmet mounted on the subject's head, the position of the ball throughout its entire flight, and the spatial position of markers used in the calibration and in the drift correction procedures were tracked at 100 Hz using a Vicon-612 motion capture system (Vicon, Oxford, UK). A large tracking volume ( $6 \times 3 \times 3 \text{ m}^3$ ) was required to capture the motion of both the ball and the subject's upper limb. The marker reconstruction residuals, averaged over the nine cameras, obtained in such volume with the standard Vicon calibration procedure, ranged across subjects between 0.91 and 0.99 mm (mean 0.96 mm). Head movements were recorded by means of several retro-reflective markers attached to the surface of the EyeLink-II helmet (Figure 4B): left front head (LFHD, i.e.,  $\underline{M}_3$ ); left back head (LBHD), right front head (RFHD, i.e.,  $\underline{M}_2$ ); right back head (RBHD, i.e.,  $\underline{M}_1$ ). Marker coordinates were referred to a right-handed calibration frame placed on the floor at 6-m distance from the launch plane. This reference frame represents the  $[\underline{w}_1, \underline{w}_2, \underline{w}_3]$  world reference frame introduced above, and it was oriented such that the  $\underline{w}_1$  axis was horizontal and pointed from the subject to the launch location, and the  $\underline{w}_3$  axis was vertical and pointed upward (Figure 4A). A consumer-

Trial type	Number of trials
Static trial	1
Calibration trial	2
Block I	10
Drift correction trial	1
Block II	10
Drift correction trial	1
...	...
Block VIII	10
Drift correction trial	1
Static trial	1
Calibration trial	1

Table 2. Blocks sequence schema of the experiment.

grade video camera was used to film the entire experimental session.

### Calibration procedure

Prior to the onset of the experimental session, the eye tracker was fitted to the subject and calibrated according to the standard procedure specified in the manufacturer's user manual. While this procedure was not necessary for our procedure, it allowed us to assess the correct adjustment of position and orientation of the cameras. To this aim, the EyeLink-II cameras and headband were adjusted in order to reliably track the pupil position. Participants were seated in front of a computer monitor and were required to keep their head still while fixating several points of known position on the monitor. We used the P-CR mode for tracking both eyes at 250 Hz, and we recorded both the raw and the HREF data. Once the standard calibration procedure of the eye tracker was successfully completed, the subject performed a *static trial*, standing in *primary position* at 6 m from the launcher and looking straight ahead at a static target placed in correspondence of the lower edge of the exit hole on the screen from which the ball was projected, i.e., the REF marker in Figure 4A. The subject was then required to gaze at a Vicon marker located on the edge of a stick that was slowly moved by the experimenter within the subject's field of view (i.e. *the calibration trial*), as shown in Figure 4C. Overall, each *calibration trial* lasted approximately 5 min. The experimenter paid great attention to probing the entire region of space where the subject would have possibly directed his/her gaze for tracking the motion of the ball throughout its flight. Also, it was important to train the algorithm with a large set of eye-head coordination configurations. To this aim, during the first of the two initial *calibration trials*, subjects were required to gaze at the target mainly exploiting eye movements, i.e., by minimizing head movements, thus covering a large range of orientation angles in eye

coordinates. Indeed, during pilot experiments with unrestrained head conditions, subjects tended to pursue the target by increasing the head contribution to gaze while leaving the eye pupils in the center of the orbits. During the second calibration trial, subjects were instead asked to track the target using their preferred eye/head coordination. As the helmet was tightened on the head to avoid undesired displacements, subjects were instructed to pause during the experiment and take off the helmet whenever they felt uncomfortable. In these cases, as for Subjects 1 and 2, the calibration procedure was repeated before restarting the experimental session. Eye pupil coordinates on the camera plane recorded by the EyeLink-II system and positions of the target and head markers collected with the Vicon system during static and calibration trials were digitally low-pass filtered at 25 Hz cutoff frequency for EyeLink-II data, and 15 Hz cutoff frequency for Vicon data; (FIR filter; Matlab `filtfilt` function). Low-pass filtering was preferred to reduce signal noise due to suboptimal CR detection, which could affect the calibration algorithm outcomes, hence the extracted geometrical configuration of the system. However, once that the mapping between 3-D gaze coordinates and 2-D image coordinates of the projection of the pupil on the camera CCD was determined, the algorithm could be used to reconstruct any data, regardless of the filter applied. Data segments corrupted by high noise due to poor tracking were removed according to the following procedure. First, data were differentiated to obtain first and second derivatives. Then, time intervals of  $\pm 20$  ms from the instant in which the acceleration exceeded its mean value  $\pm 3$  *SD*, were considered as outliers, and eliminated. Visual inspection was also carried out to manually detect and remove outliers.

### Drift correction procedure

Subjects performed *drift-correction trials* throughout the experimental session at the end of each block of 10 ball launches. They were asked to remain in the primary position and gaze at a point in space, the REF marker of Figure 3A, for 5 s. Similarly to the *static trial*, the positions of the three head markers (i.e.,  $\underline{M}_1$ ,  $\underline{M}_2$  and  $\underline{M}_3$ ) were recorded by the Vicon motion capture system, and left and right eye pupil positions were tracked with the EyeLink-II system. The corrected transformation matrices were then estimated as described above.

A calibration session in which we manually induced helmet slippage was conducted to evaluate the performance of the drift correction procedure. During the experiment one subject sat in front of a panel ( $80 \times 115$  cm), with the head immobilized by a chin rest. This was done to avoid accidental helmet and camera displacement and to make sure that any slippage occurred due

to controlled manual movement. The experiment was subdivided in 10 blocks. Each block consisted of one *static trial* and two *calibration trials*. During the *static trial* the subject was asked to look straight ahead toward a reference position (REF) at the center of the panel. During the *calibration trial* the subject was instructed to fixate each of the 35 ( $5 \times 7$ ) targets located over the entire panel surface for a few seconds. The targets were equally spaced horizontally and vertically. The panel height and horizontal positions were regulated in order to have the REF marker aligned with the eyes midpoint. The distance between the panel and the subject was adjusted to cover a gaze ranging from  $-15^\circ$  to  $15^\circ$  in elevation and azimuth angles with respect to the reference position. The first calibration session was carried out after the EyeLink-II proprietary calibration procedure was completed (i.e., *cal*). Between each of the subsequent blocks, the helmet was manually displaced over the subject's head by the experimenter by opening the rear clamp, twisting the helmet over the subject's head, and then closing the clamp again. The displacements were intentionally exaggerated to induce a large error in the algorithm reconstruction accuracy. The procedure is comparable to the actual removal and subsequent—very inaccurate—replacement of the helmet on the participant's head, so that the performance of the drift correction procedure after manual displacement can be regarded as the lower bound of its performance after actual helmet removal and replacement, which might be necessary in some experimental situations. All data were collected and analyzed as reported in previous sections.

### Calibration and drift-correction validation

To validate the proposed method, we carried out five different analyses. The first analysis was aimed at evaluating the error in the estimate of gaze orientation on data not used for the SND calibration procedure. In particular, we used data collected during the *calibration trials* carried out at the beginning of the experimental session. The *gaze orientation angles* extracted from the position of the target captured by the motion tracker system ( $\theta$ , azimuth, and  $\varphi$ , elevation) were compared to those estimated from the data recorded with the eye tracker ( $\hat{\theta}$  and  $\hat{\varphi}$ ). The procedure was carried out for both the eyes. Hereafter we define the *accuracy* as the mean value of errors across samples, and the *precision* as the corresponding *SD*. For each subject, the total duration of the two *calibration trials* carried out at the beginning of the experiment was divided in three time intervals of the same number of samples. We repeated the calibration procedure three times. At each iteration, the data from two of the three time intervals (i.e.,  $2/3$  of the data) were used to calibrate the system, and the

data from the third time interval (1/3 of the data) were used to estimate the error. By repeating the procedure three times, all of the possible intervals combinations of calibration data and test data were considered.

Residuals from all three repetitions were pooled together and used to compute the mean and *SD* of the error of the azimuth and elevation angles. In addition to the separate quantification of azimuth and elevation errors, we also computed a visual angle error as (Ronsse et al., 2007):

$$\Delta\alpha = \sqrt{(\vartheta - \hat{\vartheta})^2 + (\varphi - \hat{\varphi})^2}. \quad (31)$$

The second analysis compared the performance of the HREF and the SND calibration approaches. As HREF coordinates were stored by the EyeLink-II system in the original files together with the raw data used by our SND procedure, we could directly compare the performance of the two procedures. To this aim we ran the same analysis as described above, but we used the HREF procedure and data. Paired *t* tests were performed to compare accuracy and precision of the estimated gaze orientation achieved with the SND and the HREF procedures. In particular, separate tests were run for the azimuth, elevation, and visual angle errors and data from all subjects, and left and right were pooled together. The significance level was set to 0.05.

The third and fourth analyses evaluated the drift correction procedure. Data both from the control experiment in which the slippage was manually induced and from the catching experiment were used. In the first case, bias in gaze orientation estimate was quantified using the calibration parameters extracted with the data of one block, i.e., *BP block* (before perturbation), and computing the error on the *calibration trials* of the following block after the helmet perturbation was applied, i.e., *AP block* (after perturbation). This error was then compared with that obtained applying the drift correction. To this aim the *static trial* recorded at the beginning of the *AP block* was used as a *drift correction trial* of the calibration parameters extracted in the *BP block*. The data relative to the second block (i.e., Block II) were not used due to noise in the CR signal. We run a total of nine tests, relative to each pair of subsequent blocks. The last test was done using the last block (i.e., Block IX) as *BP block*, and the first calibration session carried out after the EyeLink-II proprietary calibration procedure was completed (i.e., *cal*) as *AP block*. In the second case, bias in gaze orientation estimate, possibly accumulated throughout the experimental session due to helmet slippage, was quantified by computing the error on the last *calibration trial* using the calibration parameters extracted at the beginning of the session. This error was compared with the error obtained after performing the drift

Subjects	${}^E R_C$ Left eye [ $\theta \ \varphi \ \psi$ ] $^\circ$	${}^E R_C$ Right eye [ $\theta \ \varphi \ \psi$ ] $^\circ$
S <sub>1</sub>	[12 31 -179]	[-8 20 179]
S <sub>2</sub>	[6 28 -176]	[-2 29 179]
S <sub>3</sub>	[-3 27 -178]	[3 18 -178]
S <sub>4</sub>	[6 26 -175]	[1 29 179]
S <sub>5</sub>	[10 19 -179]	[2 14 -180]
S <sub>6</sub>	[2 24 -180]	[-7 28 179]
S <sub>7</sub>	[9 17 -176]	[12 28 -176]

Table 3. Fick angles (in degrees) of the eye-to-camera rotation matrices (left and right eyes) estimated with the calibration procedure object of the present study. *Note:* For each participant,  $\theta$ ,  $\varphi$ , and  $\psi$  are respectively the horizontal, vertical, and torsional (i.e., rotation around the optical axis) components of rotational matrix.

correction. In particular, the data relative to the *drift correction trial* recorded at the end of the last block were used. In the case of Subjects 1 and 2, who paused about midway through the experimental session (see above), the *calibration trial* collected at the restart of the session was used.

Finally, the evolution bias in gaze orientation estimate throughout the experiment was also evaluated by computing the reconstruction error on the *drift correction trial*, recorded at the end of each block.

## Results

### Camera parameters

Table 3 reports the Fick angles of the rotation matrix between the eye reference frame and the camera reference frame, for each subject and eye, extracted by the calibration algorithm. These values show that, in most cases, the camera optical axes were oriented at angles larger than  $5^\circ$  with respect to the optical axes of the eyes in primary position. Indeed, one of the main concerns regarding the initial configuration of the camera on the EyeLink-II helmet system was to guarantee both a large field of view by positioning the cameras sufficiently below the eye and, at the same time, to center and to maximize the size of the pupil image on the camera field of view to have a good tracking range and resolution. These results show that the offset matrix between the eye and camera system could not be ascribed only to a rotation of the camera around its optical axis as assumed by Moore and Ronsse (Moore et al., 1996; Ronsse et al., 2007), and further support our approach. Notably, the orientation of the camera was different for the left and the right eye.

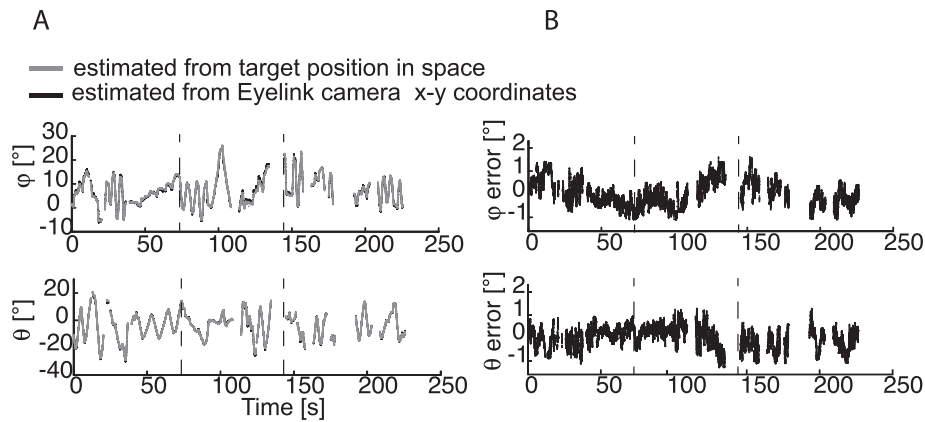


Figure 5. Example of gaze orientation angles estimation. Data from subject  $S_4$  are shown. (A) Gaze orientation angles estimated from the position of the target captured by the motion tracker system are represented by the black lines. Gaze orientation angles estimated from the data recorded with the eye tracker are represented by the gray lines. (B). The angles estimation error is defined as the difference between the gaze orientation estimated from the EyeLink-II camera coordinates and that from the target position in space. Top panels: elevation angle ( $\phi$ ); Bottom panels: azimuth angle ( $\theta$ ). Gaps represent the time intervals excluded by the analysis as specified in the Methods. The dashed lines represent the different time intervals in which the algorithm was evaluated as specified in the Methods.

## Gaze orientation error

An example of the gaze orientation angles estimated during a calibration trial in one subject ( $S_4$ ) is shown in Figure 5. The error in gaze orientation estimates achieved with the SND and the HREF calibration procedures are reported in Table 4. With the SND procedure, across subjects the gaze azimuth angle was estimated on average with an accuracy of  $0.18^\circ$  (range  $0.01^\circ$ – $0.39^\circ$ ) and a precision of  $0.48^\circ$  (range  $0.4^\circ$ – $0.72^\circ$ ), the gaze elevation angle with an accuracy of  $0.12^\circ$  (range  $0.01^\circ$ – $0.31^\circ$ ) and a precision of  $0.49^\circ$  (range  $0.34^\circ$ – $0.81^\circ$ ), and the visual angle with an accuracy of  $0.56^\circ$  (range  $0.36^\circ$ – $0.78^\circ$ ) and a precision of  $0.37^\circ$  (range  $0.27^\circ$ – $0.58^\circ$ ).

With the HREF procedure, instead gaze azimuth angle was estimated on average with an accuracy of  $0.04^\circ$  (range  $0.01^\circ$ – $0.09^\circ$ ) and a precision of  $1.41^\circ$  (range  $1.12^\circ$ – $1.91^\circ$ ), the gaze elevation angle with accuracy of  $0.13^\circ$  (range  $0.01^\circ$ – $0.32^\circ$ ) and precision of  $0.91^\circ$  (range  $0.45^\circ$ – $1.35^\circ$ ), and the visual angle with accuracy of  $1.5^\circ$  (range  $1.1^\circ$ – $1.83^\circ$ ) and a precision of  $0.8^\circ$  (range  $0.68^\circ$ – $1.1^\circ$ ). A  $t$  test analysis conducted to compare the performance of the two approaches in the gaze angles estimation showed that there was a significant difference in the SND and HREF accuracy of the azimuth and visual angles estimation and in the precision of both the azimuth, the elevation, and the visual angles ( $p < 0.001$ ). No significant difference was found in the accuracy in the estimation of the elevation angle ( $p = 0.41$ ). Overall these results confirmed that the SND estimation of azimuth was less accurate but more precise than the HREF estimation; the SND estimation

of elevation was more precise than the HREF estimation; and the SND performed much better than the HREF in the estimating of the gaze visual angle.

Our method allowed very accurate and precise measurements of gaze orientation under unrestrained head movement conditions. Moreover, it allowed the subject to be distant from the eye-tracker station during the experiment and potentially to move/walk within the room. In the case of our experiments, for example, the head rotation angles spanned by our calibration data varied from  $-13^\circ$  to  $40^\circ$  in azimuth (measures are referred with respect to the *primary position*), i.e., maximum excursion observed  $50^\circ$ , and ranging from  $-29^\circ$  to  $31^\circ$  in elevation angle, i.e., maximum excursion observed  $60^\circ$ . These intervals were larger than those reported in most of eye tracker technical specifications sheet. For instance, the EyeLink-II system used in the present study allows for head rotations between  $\pm 15^\circ$ . Moreover, it requires the subject to be positioned directly in front of the display monitor at a distance of 40 to 140 cm. According to manufacturer specifications, the head tracker can tolerate changes  $\pm 30\%$  of the display-to-head distance at calibration, while the admissible head horizontal and vertical movements are less than the width and the height of the monitor.

In summary, these results indicate that the SND approach, which makes explicit use of the position of the eye-tracker cameras with respect to the eye to estimate the eye-in-head orientation, allows for a higher tracker accuracy precision in the estimation of the gaze visual angle than the HREF approach using the eye-in-head orientation provided by the EyeLink-II system.

Subject	Left eye			Right eye		
	$\Delta\theta$	$\Delta\varphi$	$\Delta\alpha$	$\Delta\theta$	$\Delta\varphi$	$\Delta\alpha$
<b>SND</b>						
S <sub>1</sub>	0.01 ± 0.42	0.01 ± 0.75	0.45 ± 0.49	0.17 ± 0.39	−0.31 ± 0.65	0.43 ± 0.45
S <sub>2</sub>	−0.32 ± 0.60	−0.05 ± 0.80	0.36 ± 0.42	−0.11 ± 0.36	−0.10 ± 0.46	0.52 ± 0.30
S <sub>3</sub>	0.14 ± 0.72	−0.01 ± 0.39	0.69 ± 0.44	−0.21 ± 0.58	−0.24 ± 0.43	0.68 ± 0.37
S <sub>4</sub>	−0.15 ± 0.44	−0.06 ± 0.60	0.67 ± 0.36	−0.13 ± 0.44	−0.02 ± 0.53	0.58 ± 0.38
S <sub>5</sub>	−0.39 ± 0.66	0.19 ± 0.35	0.78 ± 0.35	−0.38 ± 0.65	−0.08 ± 0.44	0.78 ± 0.38
S <sub>6</sub>	0.05 ± 0.38	−0.11 ± 0.35	0.45 ± 0.27	0.12 ± 0.41	0.08 ± 0.34	0.46 ± 0.29
S <sub>7</sub>	−0.10 ± 0.41	−0.09 ± 0.35	0.48 ± 0.27	0.18 ± 0.38	−0.09 ± 0.47	0.52 ± 0.37
<b>HREF</b>						
S <sub>1</sub>	−0.06 ± 1.34	−0.25 ± 0.87	1.44 ± 0.71	−0.03 ± 1.12	−0.16 ± 1.18	1.45 ± 0.71
S <sub>2</sub>	−0.05 ± 1.45	0.32 ± 0.91	1.54 ± 0.76	0.06 ± 1.29	−0.02 ± 0.82	1.31 ± 0.73
S <sub>3</sub>	0.09 ± 1.65	0.13 ± 1.35	1.83 ± 1.07	0.06 ± 1.34	0.08 ± 1.22	1.54 ± 0.93
S <sub>4</sub>	0.01 ± 1.91	−0.13 ± 0.85	1.80 ± 0.99	0.02 ± 1.35	−0.18 ± 0.97	1.43 ± 0.79
S <sub>5</sub>	−0.01 ± 1.41	−0.01 ± 0.74	1.39 ± 0.68	0.00 ± 1.50	0.02 ± 0.53	1.35 ± 0.78
S <sub>6</sub>	−0.03 ± 1.36	−0.29 ± 1.14	1.54 ± 0.90	0.03 ± 1.44	−0.20 ± 1.17	1.63 ± 0.85
S <sub>7</sub>	−0.01 ± 1.21	−0.02 ± 0.45	1.08 ± 0.69	0.06 ± 1.23	−0.06 ± 0.58	1.15 ± 0.70

Table 4. Gaze orientation error after calibration. *Note:* Mean and *SD* of azimuth ( $\theta$ ), elevation ( $\varphi$ ), and visual ( $\alpha$ ) angles obtained with the SND and HREF calibration procedures are reported for each subject and eyes.

## Drift correction

In an experimental test in which the helmet slippage was manually induced by the experimenter, we evaluated the outcomes of the drift correction procedure by first computing the gaze angles error on the *calibration trials* of one block using the calibration parameters extracted on the *calibration trials* of the previous block, and then comparing this error with that obtained after the drift correction was applied. In Figure 6A shows screenshots of the camera images of the left eye taken from the EyeLink-II system display during the *static trial* carried out at the beginning of each block. The displacement of the helmet over the head induced a shift of the eye image with respect to center of the camera plane. Notably, the relative position between the CR (indicated by a yellow dot and an unboxed cross) and the center of the pupil (at the center of a blue disk indicated by a boxed cross) changed over the blocks. In Figure 6 (Panel B), black bars represent the mean reconstruction error ( $\pm SD$ ) when the initial calibration parameters were used to estimate gaze coordinates; white bars represent the error ( $\pm SD$ ) achieved after the drift correction was applied. Overall, the drift correction conspicuously improved the algorithm accuracy in gaze angles estimation. In particular, the error decreased up to 96% in the estimate of the azimuth angle, passing from 2.71° to 0.11° (i.e., helmet slippage number 2 in Figure 6, right eye; measures expressed in absolute value), as following the perturbation between Blocks II and III. Similarly the gaze estimation error in elevation angle was reduced up to 97%, passing from 8.42° to 0.27° (i.e.,

helmet slippage number 3 in Figure 6, left eye), as following the perturbation between Blocks III and IV.

We validated the drift correction method during the performance of a one-handed catching experiment. We compared the error on a calibration trial performed at the end of the experiment obtained with the calibration parameters obtained in a calibration trial performed at the beginning of the experiment with and without the application of the drift correction. Subjects were involved in a challenging experimental condition for eye-tracking due to the large head excursions required to track the ball flying at high speed, likely causing small displacements of the helmet over the head. Head orientation angles with respect to the *primary position* varied, across subjects and experimental conditions, ranging from 5° to 47° in azimuth and ranging from −4° to 48° in the elevation. Figure 7A shows an example of the right eye elevation angle measured during the last calibration trial for subject S<sub>2</sub>, before (top panel) and after (bottom panel) the correction was applied. The error was considerably reduced in the second case. Results for all subjects are presented in Figure 7B, showing that when the error was large, as for subjects S<sub>1</sub> and S<sub>2</sub>, the drift correction improved the estimate. In particular, for subject S<sub>1</sub> the error in the estimate of the right eye elevation angle decreased by 85%, from 4.79° to 0.73°, and the error in the estimate of the right eye azimuth angle decreased by 94%, from 1.94° to 0.11°. Similarly, for subject S<sub>2</sub> the error in the estimate of the right eye elevation angle decreased by 97%, from 3° to 0.06°. Notably, in the case of S<sub>2</sub> a large error was only present in the right eye elevation angle, suggesting that



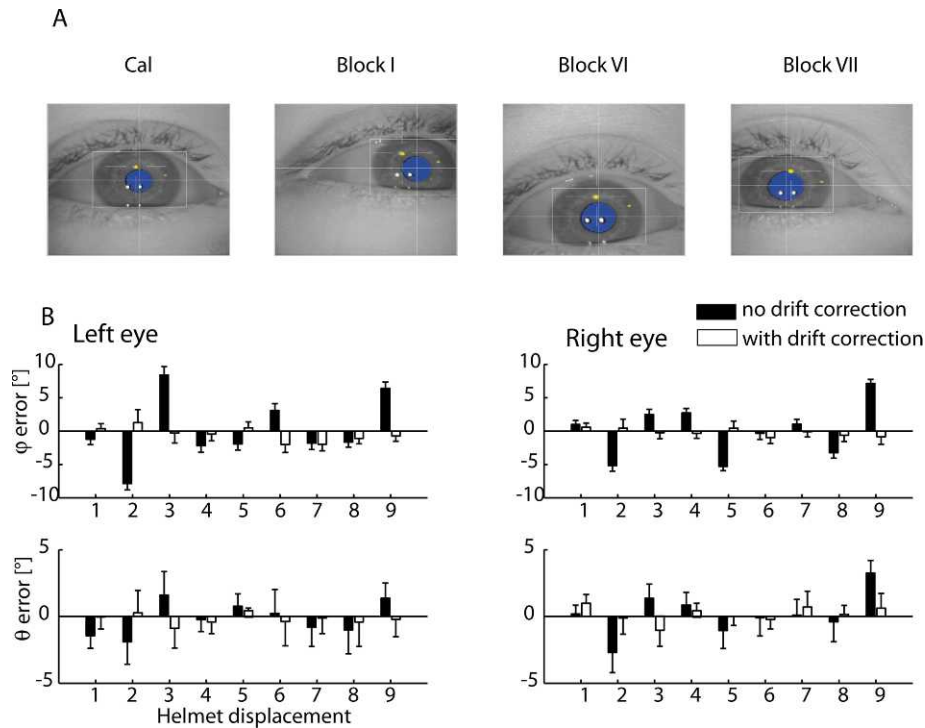


Figure 6. Drift correction evaluation: the control experiment case. (A) Left eye images during the static trial recorded at the beginning of the first calibration session carried out after the EyeLink-II proprietary calibration procedure was completed (*cal*), and in following blocks after the helmet slippage was manually induced (only three blocks are shown as an example). Notably, the relative position between the CR (yellow dot and unboxed cross) and the pupil (boxed cross) did not remain stationary. (B) Summary of the results on the gaze error relative to all the nine helmet displacements carried out in the test experiment. Black bars represent the mean error  $\pm$  *SD* when the calibration parameters extracted with the data of one block were used to estimate gaze orientation angles of the following calibration trial; white bars represent the error  $\pm$  *SD* achieved when the drift correction was applied.

an accidental displacement of the right camera had occurred rather than helmet slippage.

The progression of the error in the estimation of gaze throughout the catching experiment is shown in Figure 8. For subjects  $S_1$ ,  $S_2$ , and  $S_7$ , helmet slippage or an accidental displacement of the camera had likely occurred during the experiment (in the fourth block for  $S_7$ , in the fifth block for  $S_7$ , and in the last block for  $S_1$ ). For the remaining participants, the gaze estimation error did not increase dramatically throughout the experiment, as shown by the small error observed on the fixation trials carried out at the end of each block (Figure 8).

## Discussion

We introduced a novel method to estimate gaze orientation in space under unrestrained head movement conditions. The method integrates the measurement of head pose in space obtained with a motion tracker system and the measurement of eye orientation in the head measured with a video-based eye-tracker

system. We applied a geometry-based approach in order to derive the nonlinear equations relating gaze orientation to the position of the center of the pupil image on the focal plane of the tracker's camera. A nonlinear optimization algorithm was then used to estimate all the parameters after initialization with approximate estimates based on geometrical measurements. Finally the method was validated during an experimental session in which subjects performed a catching task, which required fast arm and head movements and tracking of a ball flying at high speed, i.e., a challenging experimental condition for eye tracking. Our method was able to achieve an accuracy on average of  $0.56^\circ$ , and better than  $0.78^\circ$ , and a precision on average of  $0.37^\circ$ , and better than  $0.49^\circ$ , in the measure of the gaze visual angle (see Table 4). Moreover, the estimate of the azimuth and elevation angles had an accuracy of  $0.07^\circ$  and better than  $0.39^\circ$ , and a precision of  $0.49^\circ$  better than  $0.80^\circ$ . We compared the gaze angle reconstruction error achieved with our method to the error achieved with a simpler procedure which used the EyeLink-II proprietary calibration algorithm to extract the rotation angle of the eye relative to the head (HREF mode) in combination with

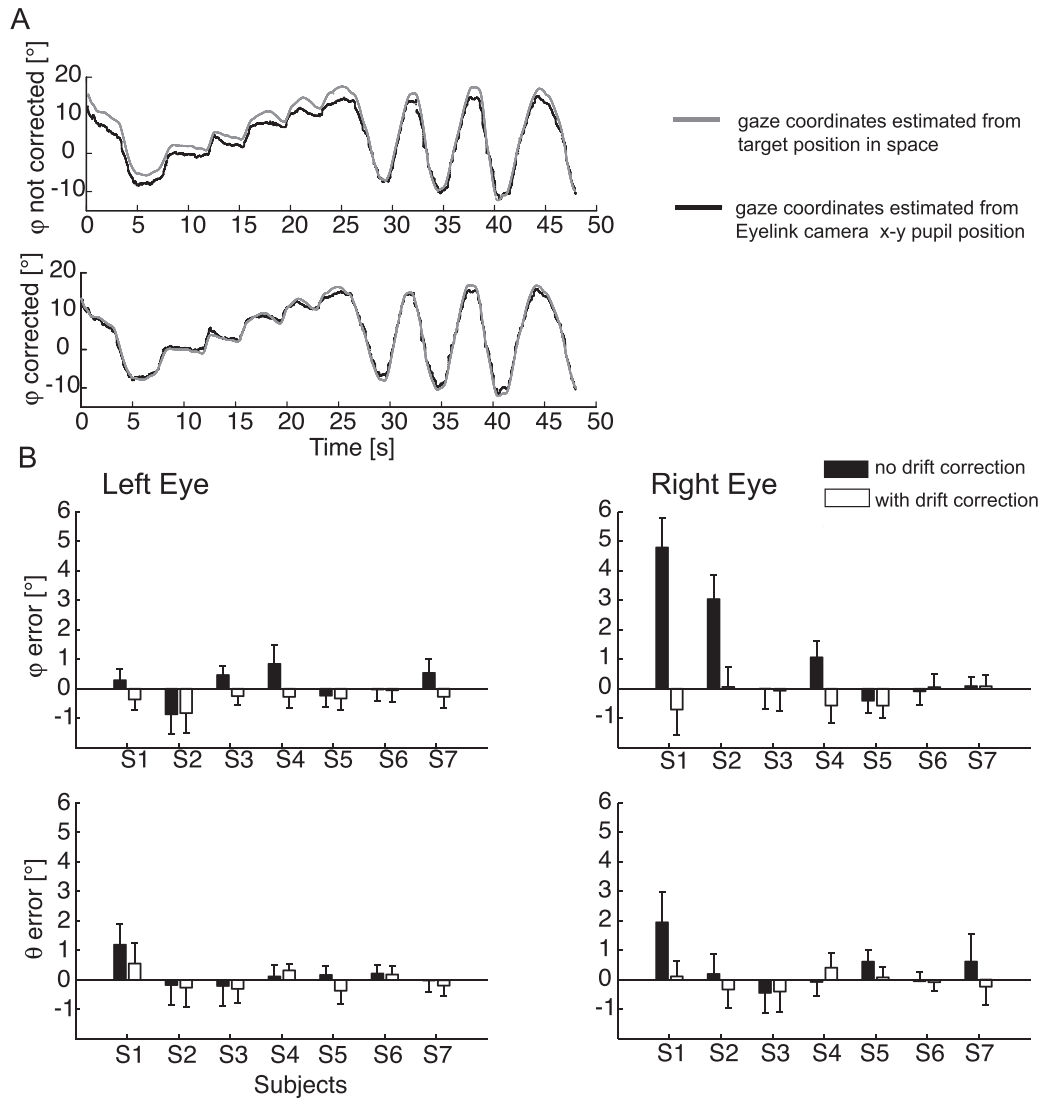


Figure 7. Drift correction evaluation during a catching experiment. (A) Example of gaze elevation angle reconstruction for subject  $S_2$  during the last calibration trial collected at the end of the experimental session. Top: Angle estimated from the position of the target captured by the motion tracker system (black line) and from pupil position recorded by the eye tracker (gray line) without drift correction. Bottom: same as Top but with the drift correction applied to the estimation of the angle from pupil position. (B) Summary of results for all subjects. Black bars represent the mean error  $\pm$  SD when the calibration parameters were used to estimate gaze orientation angles of the last calibration trial; white bars represent the error  $\pm$  SD achieved when the drift correction was applied.

the measurement of the head pose to compute gaze direction in space. Our procedure based on the raw data collected in P-CR mode achieved a higher tracking precision with respect to the HREF approach (i.e., on average respectively  $0.48^\circ$  and  $1.41^\circ$  for the azimuth angle;  $0.49^\circ$  and  $0.91^\circ$  for the elevation angle).

However, the SND procedure was less accurate than the HREF one in the case of the azimuth angle estimation (i.e., average reconstruction errors were respectively  $0.18^\circ$  and  $0.04^\circ$ ).

Our approach outperforms previously proposed methods (Moore et al., 1996; Ronsse et al., 2007) which reported an accuracy between  $0.83^\circ$  and  $2.35^\circ$  and a precision between  $0.52^\circ$  and  $3.47^\circ$  in the estimate of the

visual angle (see table 2 in Ronsse et al., 2007). With respect to those previous approaches, the proposed model used more parameters to describe the problem geometry, which could explain in part the better accuracy achieved in the gaze angle estimation. However, the additional geometrical parameters explicitly described the eye-to-camera transformations and hence allowed us to develop the drift correction procedure as described in the Methods section.

Our model does not require any assumption on the configuration of the eye-tracker camera with respect to the subject's eye and head. In particular, the camera optical axis was not assumed to pass through the center of the eye and to be aligned with the line of sight when

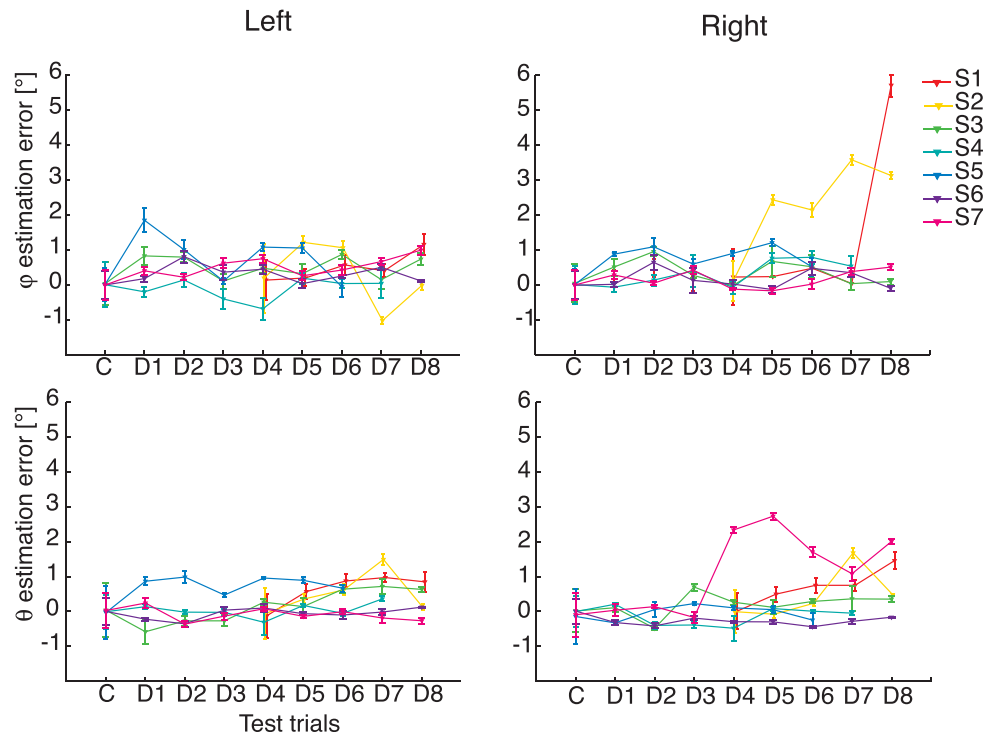


Figure 8. Error progression in the gaze estimation throughout the experiment. Gaze orientation angle error (mean and *SD*) of *drift correction* fixation trials recorded at the end of each experiment block. The calibration parameters used for the reconstruction were extracted with data relative to the first two *calibration trials* carried out at the beginning of the experimental session, and no drift correction was applied. Each participant is coded with a different color, and results for both left and right eyes are shown. Top panels: elevation angles; Bottom panels: azimuth angles.

in the primary position (i.e., when fixating a distant point straight ahead at eye height). Furthermore, it uses a perspective projection instead of an orthographic projection to estimate the pupil position from its image on the eye-tracker image plane; hence no constraints on the distance between the camera lens plane and the eye center were necessary (Moore et al., 1996; Nakayama, 1974). Our approach can be used when the tracker cameras are positioned much below the eye center, as with the eye-tracker system we used for the experimental validation of the approach. In fact in our experiments, as indicated by the results of the calibration procedure (Table 3), the rotation matrices expressing the orientation of the camera frame with respect to the eye frame presented azimuth and elevation angles larger than  $5^\circ$ , a value above which it has been previously suggested that the effect of an eye-camera misalignment cannot be neglected (Moore et al., 1996). This result hence confirmed that the assumption made by Ronsse et al. (2007) did not hold in our case. Small misalignments may be possible only when using a video-based eye-tracker system that measures the pupil position by reflecting the eye image on a small mirror. Under these conditions the camera can be mounted at a right angle with respect to the line of sight, allowing for a wide field of view, while

maintaining the camera reference frame parallel to the eye reference frame. In contrast, our method can be used with any video-based eye-tracker system, independently of the camera geometry and the pupil detection methods.

We achieved gaze orientation estimation performance in unrestrained head movement conditions comparable with that obtained in the typical usage conditions of the EyeLink-II system with subjects sitting in front of a computer monitor. In these conditions, the subject head is fixed or it can be moved in a very limited range of positions and orientations while it is tracked using infrared emitters mounted at the corners of the monitor and an infrared camera attached to the helmet. For the subjects enrolled in the present study, the average error of gaze visual angle (i.e., the accuracy parameter defined in our study), evaluated with the EyeLink-II proprietary calibration algorithm and reported in the data sheet available to user ranged within a  $0.22^\circ$ – $0.65^\circ$  interval, while the *SD* (i.e., the precision parameter defined in our study) ranged within a  $0.04^\circ$ – $0.81^\circ$  interval.

A key advantage of the present approach with respect to previously proposed approaches (Johnson, Liu, Thomas, & Spencer, 2007; Moore et al., 1996; Ronsse et al., 2007) is its ability to correct the system

calibration parameters upon accidental displacement (or deliberate removal and replacement) of the eye tracker. In this respect, the evaluation of the *drift correction* procedure showed that the error in the gaze orientation estimation could be improved by up to 97%, (see Figures 6 and 7). Notably, in our experiment, eye movements were recorded in the P-CR mode. This approach is preferable to the simpler tracking of the pupil only because it partially compensates for the displacement of the camera with respect to the eye, which could be related to an accidental bump on the eye tracker, as well as tremor or simply subject's facial expression during the experiment. In fact, the relative position of pupil and CR on the camera image remains stationary under the assumption of a flat corneal surface (Li et al., 2008) and is affected much less than the absolute position of the pupil. In our tests, we observed that when the eye movements were recorded in pupil mode, the gaze reconstruction error was larger than in the P-CR mode (data not shown). However, if the illuminator is on a head-mounted helmet, as in our case, the P-CR measurement is affected by a displacement of the CR-illuminating source that may occur together with a camera displacement. In fact, even in the absence of eye movements, the CR position on the camera image—and hence the P-CR coordinates—may change (see Figure 6). The drift correction results thus indicate that our procedure provides a valid tool to improve gaze estimation after an accidental helmet slippage, even when using the P-CR recording mode. In these conditions the correction of the calibration parameters presumably reduced the residual error not already compensated by the use of P-CR recording mode. However, explicitly taking into account the detailed geometry of the drift correction in P-CR mode would require modeling the position of the illuminating source with respect to the eye, which falls outside the scope of the present study and may prove problematic without access to the separate CR and P coordinates, as with the EyeLink-II system. On the other hand, we should expect a straightforward relation between the helmet slippage estimated with the drift correction algorithm and that measured by tracking three Vicon markers applied on the subject's face if the eye movements were recorded in the P mode. To investigate this issue we ran the same drift correction validation test reported in the present study using the EyeLink-II system in pupil-tracking mode instead of P-CR mode (data not shown). We found that the torsional components (i.e., the  $\psi$  angle of the  ${}^H R_H$  matrix, also representing the rotation around the camera optical axis in the case of the camera frame system) of the slippage rotation matrices evaluated by the two different approaches matched, as expected. However, for the horizontal and the vertical components it was not possible to find a clear relationship. Such discrep-

ancies might be due to some of the simplifying assumptions made for deriving the drift correction algorithm, such as considering the helmet slippage as a pure rotation with respect to the center of the skull. While those issues will need additional investigations in the future, yet the analysis on the gaze angle error estimated with eye-tracking data collected in P-CR mode confirmed that the procedure provides a valid tool to correct the system calibration parameters, and overall we observed a reduced mean value of the error after the correction was applied (up to 98%). This result is in line with Moore and colleagues (Moore et al., 1996) who showed that the torsional offset contributes more to the error in the determination of the eye position than the horizontal and vertical offsets. Our algorithm would then mainly compensate for this effect caused by the slippage, ignoring those that are less relevant. Overall, these results suggest that the drift correction procedure could be used in several experiments carried out in naturalistic condition to reduce the possibility for error accumulation in the gaze angle estimation throughout the session.

Some considerations on the use of the drift correction procedure are needed at this point. For instance, the analysis of the error progression throughout the catching experiment reported in Figure 8 suggested that helmet movements occurred in three out of the seven participants in our study (i.e.,  $S_1$ ,  $S_2$ , and  $S_7$ ). As expected, due to the large head rotations required to pursue the ball (i.e., up to almost  $50^\circ$  in both azimuth and elevation angles across subjects and ball flight conditions), helmet movements during these type of applications are likely. Practical use of the drift correction procedure would then require the running of a fixation trial several times during the experiment, and then correct the calibration parameters according to the outcomes of an offline analysis of the gaze error. Alternatively one could ask the subject to fixate a specific point prior to the beginning of each trial, which could be used to correct for the drift. As stated previously, the use of the drift correction procedure would allow helmet removal and replacement within an experimental session, provided that the camera position with respect to the helmet remains unchanged. In this case, the experimenter should take care to replace the helmet in the same configuration with respect to the head.

In conclusion, our method is robust and provides an accurate and precise estimate of gaze orientation in space when the head is free to move. It can be used in any experimental scenario that requires unrestrained head movements and subject displacements within a large workspace. The only constraints we see are represented by the region of space tracked by the motion capture systems and the length of the eye-tracker cable. Future improvements of our method may

be obtained by including additional parameters in the model. For instance, the eye eccentricity could be taken into account when computing the pupil position and orientation in eye coordinates (Equations 5 and 6). Similarly, a further camera parameter such as the radial distortion of the camera lens could also be optimized.

*Keywords:* oculography, gaze orientation in space, calibration, nonlinear optimization, drift correction

## Acknowledgments

Supported by the Italian Ministry of Health, the Italian Space Agency (DCMC and CRUSOE projects), and the EU Seventh Framework Program (FP7-ICT No. 248311 AMARSi).

Commercial relationships: none.

Corresponding author: Benedetta Cesqui.

Email: b.cesqui@hsantalucia.it.

Address: Laboratory of Neuromotor Physiology, IRCCS Santa Lucia Foundation, Rome, Italy.

## References

- Abbott, W. W., & Faisal, A. A. (2012). Ultra-low-cost 3D gaze estimation: An intuitive high information throughput compliment to direct brain-machine interfaces. *Journal of Neural Engineering*, *9*(4), 046016.
- Anastasopoulos, D., Kimmig, H., Mergner, T., & Psilas, K. (1996). Abnormalities of ocular motility in myotonic dystrophy. *Brain*, *119*(Pt 6), 1923–1932.
- Bekkering, H., & Neggers, S. F. (2002). Visual search is modulated by action intentions. *Psychological Science*, *13*(4), 370–374.
- Cesqui, B., d'Avella, A., Portone, A., & Lacquaniti, F. (2012). Catching a ball at the right time and place: Individual factors matter. *PLoS One*, *7*(2), e31770.
- Collewijn, H., Van der Steen, J., Ferman, L., & Jansen, T. C. (1985). Human ocular counterroll: Assessment of static and dynamic properties from electromagnetic scleral coil recordings. *Experimental Brain Research*, *59*(1), 185–196.
- Duchowski, A. T. (2007). *Eye tracking methodology* (2nd ed.). London: Verlag.
- Fry, G. A., & Hill, W. W. (1962). The center of rotation of the eye. *American Journal of Optometry & Archives of American Academy of Optometry*, *39*, 581–595.
- Haslwanter, T., & Moore, S. T. (1995). A theoretical analysis of three-dimensional eye position measurement using polar cross-correlation. *IEEE Transactions on Biomedical Engineering*, *42*(11), 1053–1061.
- Hayhoe, M. M., McKinney, T., Chajka, K., & Pelz, J. B. (2012). Predictive eye movements in natural vision. *Experimental Brain Research*, *217*(1), 125–136.
- Hua, H., Krishnaswamy, P., & Rolland, J. P. (2006). Video-based eyetracking methods and algorithms in head-mounted displays. *Optics Express*, *14*(10), 4328–4350.
- Jaafari, N., Rigalleau, F., Rachid, F., Delamillieure, P., Millet, B., Olie, J. P., et al. (2011). A critical review of the contribution of eye movement recordings to the neuropsychology of obsessive compulsive disorder. *Acta Psychiatrica Scandinavica*, *124*(2), 87–101.
- Johnson, J. S., Liu, L., Thomas, G., & Spencer, J. P. (2007). Calibration algorithm for eyetracking with unrestricted head movement. *Behavior Research Methods*, *39*(1), 123–132.
- Kimmel, D. L., Mammo, D., & Newsome, W. T. (2012). Tracking the eye non-invasively: Simultaneous comparison of the scleral search coil and optical tracking techniques in the macaque monkey. *Frontiers in Behavioral Neuroscience*, *6*, 49.
- Land, M. F. (2012). The operation of the visual system in relation to action. *Current Biology*, *22*(18), R811–R817.
- Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*(25–26), 3559–3565.
- Land, M. F., & McLeod, P. (2000). From eye movements to actions: How batsmen hit the ball. *Nature Neuroscience*, *3*(12), 1340–1345.
- Land, M. F., & Tatler, B. W. (2001). Steering with the head: The visual strategy of a racing driver. *Current Biology*, *11*(15), 1215–1220.
- Lee, E. C., Woo, J. C., Kim, J. H., Whang, M., & Park, K. R. (2010). A brain-computer interface method combined with eye tracking for 3D interaction. *Journal of Neuroscience Methods*, *190*(2), 289–298.
- Li, F., Munn, S., & Pelz, J. (2008). A model-based approach to video based eye tracking. *Journal of Modern Optics*, *55*(4–5), 503–531.
- Marieb, E. N. (2001). *Human anatomy and physiology* (1st ed.). San Francisco: Benjamin Cummings.
- Moore, S. T., Haslwanter, T., Curthoys, I. S., & Smith, S. T. (1996). A geometric basis for measurement of

- three-dimensional eye position using image processing. *Vision Research*, 36(3), 445–459.
- Morimoto, C. H., & Mimica, M. R. M. (2005). Eye gaze tracking techniques for interactive applications. *Computer Vision & Image Understanding*, 98(1), 4–24.
- Nakayama, K. (1974). Photographic determination of the rotational state of the eye using matrices. *American Journal of Optometry & Physiological Optics*, 51(10), 736–742.
- Pelz, J., Hayhoe, M., & Loeber, R. (2001). The coordination of eye, head, and hand movements in a natural task. *Experimental Brain Research*, 139(3), 266–277.
- Robinson, D. A. (1963). A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Transactions in Biomedical Engineering*, 10, 137–145.
- Ronsse, R., White, O., & Lefevre, P. (2007). Computation of gaze orientation under unrestrained head movements. *Journal of Neuroscience Methods*, 159(1), 158–169.
- Schreiber, K., & Haslwanter, T. (2004). Improving calibration of 3-D video oculography systems. *IEEE Transactions in Biomedical Engineering*, 51(4), 676–679.
- Snyder, L. H., Batista, A. P., & Andersen, R. A. (2000). Intention-related activity in the posterior parietal cortex: a review. *Vision Research*, 40(10–12), 1433–1441.
- Spering, M., & Montagnini, A. (2011). Do we track what we see? Common versus independent processing for motion perception and smooth pursuit eye movements: A review. *Vision Research*, 51(8), 836–852.
- van der Geest, J. N., & Frens, M. A. (2002). Recording eye movements with video-oculography and scleral search coils: A direct comparison of two methods. *Journal of Neuroscience Methods*, 114(2), 185–195.
- Warabi, T., Kase, M., & Kato, T. (1984). Effect of aging on the accuracy of visually guided saccadic eye movement. *Annals of Neurology*, 16(4), 449–454.
- Zago, M., McIntyre, J., Senot, P., & Lacquaniti, F. (2009). Visuo-motor coordination and internal models for object interception. *Experimental Brain Research*, 192(4), 571–604.