

Inexpensive system for real-time 3-dimensional video-oculography using a fluorescent marker array

Americo A. Migliaccio^a, Hamish G. MacDougall^d, Lloyd B. Minor^{a,b,c},
Charles C. Della Santina^{a,b,*}

^a Department of Otolaryngology-Head and Neck Surgery, Laboratory of Vestibular Neurophysiology,
Johns Hopkins University School of Medicine, 601 N. Caroline Street, Room 6253, Baltimore, MD 21287, USA

^b Biomedical Engineering, Johns Hopkins University School of Medicine, USA

^c Neuroscience, Johns Hopkins University School of Medicine, USA

^d School of Psychology-Vestibular Research Laboratory, University of Sydney, Australia

Received 30 January 2004; received in revised form 28 September 2004; accepted 30 September 2004

Abstract

We describe a novel, inexpensive method for real-time measurement of binocular three-dimensional eye position. The method employs consumer-grade digital video cameras (“webcams”) to track an array of three fluorescent non-collinear markers affixed to each eye. The instantaneous position of the marker array relative to a reference position is used to construct a rotation matrix describing the eye rotation. The mathematical computation used to determine the rotation matrix is conceptually simpler and computationally more efficient than methods previously described, allowing generation of binocular three-dimensional eye position in real-time during image acquisition. The fluorescent marker is illuminated using a UV-A light source. The light source and reflective artifacts are filtered out to improve the signal to noise ratio. In addition, we present a method to align the camera with the center of eye rotation. When tested in vitro, the video-oculography (VOG) method had a <2.9% positional error (in each component of 3-D eye position) for eye positions within 20° of center. We directly compared this method of VOG to the search coil technique by measuring three-dimensional eye position simultaneously using search coils and VOG in a chinchilla (*C. laniger*). The in vivo positional difference between the two methods was <3.1% for each component of 3-D eye position.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Three-dimensional; 3-D; Video-oculography; Eye position; UV; Fluorescent; Vestibular; Vestibulo-ocular reflex; VOR

1. Introduction

Precise and accurate measurement of eye rotation is central to vestibulo-ocular research. The eye can rotate in three dimensions: horizontally, vertically and torsionally (about the line of sight). The “gold standard” method for measuring three-dimensional eye position is the scleral search coil technique (Robinson, 1963; Collewijn et al., 1985). In animal experiments, the search coils can either be implanted (Paige and Tomko, 1991; Minor et al., 1999) or glued to the eye (Gilchrist et al., 1998). Implantation of search coils

can restrict eye movements due to conjunctival scarring and inflammation or coil lead tension. These problems become increasingly significant when coils are implanted in rodents and other species with small eyes (Stahl et al., 2000). Depending on surgical technique, coil implantation can also damage extraocular muscles and their pulleys, further distorting eye movements (Demer et al., 1995). Temporarily gluing coils to the frontal surface of the eye minimizes the risk of restriction due to scarring; however, glued coils and their leads can impede eye movement range by impacting the lids and canthi. Repeated impact of the coils against lid margins can dislodge the coils, limiting the duration of experiments. Whether implanted or glued, search coils require uniform, stable magnetic fields for transduction of eye rotation. These fields can

* Corresponding author. Tel.: +1 410 955 7381; fax: +1 410 614 7222.
E-mail address: charley.dellasantina@jhu.edu (C.C. Della Santina).

be distorted by metallic objects and by currents flowing in equipment nearby.

The drawbacks of the search coil technique have prompted efforts to develop video-oculographic (VOG) systems for measurement of three-dimensional eye position. Such systems are gaining wider acceptance. Horizontal and vertical eye position can be determined by tracking the pupil and/or a corneal reflection (Stahl et al., 2000; Kaufman, 2002). To determine ocular torsion, most currently available VOG systems either track two or more landmarks on the eye, which involve the use of vector cross products to calculate the axis of rotation (Nakayama, 1974; Parker et al., 1985; Yamanobe et al., 1990; Ott et al., 1990), or employ some variation of the polar cross correlation method, which involves measuring and tracking changes in iral contrast along a circular sampling path (Anderson and Hatamian, 1983; Vieville and Masse, 1987; Clarke et al., 1991; Moore et al., 1991). In humans, pronounced iral striations make the polar cross correlation method practical. In animals that do not have pronounced iral striations, such as the chinchilla, rabbit, guinea pig and mouse, it is more practical to track attached landmarks.

We describe an inexpensive technique for real-time measurement of three-dimensional eye position using consumer-grade digital video cameras to track an array of three 1 mm × 1 mm markers on a piece of plastic film affixed to the cornea. To increase contrast between the markers and unwanted corneal reflections, the markers were fluorescent and illuminated with a UV-A light source outside the camera's range of spectral sensitivity. The marker array was glued to the frontal surface of each cornea and typically remained attached for >4 h. The three-dimensional eye rotation necessary to move a marker array from a reference position to a final position was calculated using a mathematical method that is simpler and more efficient than others described in the literature (Nakayama, 1974; Ott et al., 1990). We used inexpensive, consumer-grade "webcam" digital video cameras with modified lenses and the LabVIEW G programming language (National Instruments, Austin, TX) to simplify software development. Binocular three-dimensional eye positions were computed and displayed in real-time using an intuitive graphical user interface. We validated the system *in vitro* using a Fick gimbal and *in vivo* by comparing eye movement measurements made simultaneously using both search coils and VOG.

2. Materials and methods

2.1. Fluorescent marker array and light source

We fabricated fluorescent marker arrays using plastic film laminated on paper saturated with fluorescent yellow ink. The film was opaque except for three transparent 1 mm × 1 mm windows separated by 1 mm and arranged in a 45° right triangle. (This pattern was formed using a Casio KL-750



Fig. 1. Fluorescent marker array on the eye of a chinchilla. For this photograph, the eye was illuminated with visible light in addition to UV. Under normal testing conditions, only UV illumination is used, increasing the relative brightness of fluorescent markers.

film-based label maker to print a colon followed by a period [:.]. The distance between the windows of the plastic ink backing tape was fixed at 1 mm.) A small amount of cyanoacrylate was used to improve adhesion of the marker array to the eye, as used in previous animal studies (Hess and Dieringer, 1991). In order to measure vestibular-mediated eye movements in the absence of vision, we deliberately covered the pupil with a thin film of cyanoacrylate. For experiments requiring intact vision, the marker array could be placed away from the pupil, or on an annular contact lens. A diffuse ultraviolet (UV-A) light source (360 nm peak, 9 W, FPX7BLB, Ushio Inc., Japan) illuminated the array (Fig. 1). Alternatively, 380 nm UV-A light-emitting diodes (LEDs) (SSL-LX5093SUVC, Lumex Inc.) are also suitable. Depending on the spectral sensitivity of the specific camera used, a UV cut filter (SKYLIGHT 1B Hoya, Japan) or a yellow pass filter (K2 yellow filter Hoya, Japan) can be used to improve contrast. No filter was necessary with the webcams we used, because their color CCD is already less sensitive to UV than most monochrome image sensors.

The maximum allowable exposure of UV-A (320–400 nm) that will not harm the eye (cornea and lens) for human use is 1 J/cm² (ICNIRP, 1996). The "black light" or UV-A lamp (sometimes called a "Wood's Lamp") that was used in this study is not considered hazardous because the UV-A radiance at the lamp surface is only about 3 mW/cm². At 30 cm distance, the UV-A radiance at the eye surface is about 50 μW/cm² and would require >5 h exposure to reach 1 J/cm². The UV LED light source generates about 3 mW, the beam angle is 30° so at 20 cm, the UV-A radiance at the eye surface is about 33 μW/cm² and would require >8 h to reach 1 J/cm².

2.2. Camera translational alignment with the eye

Translational misalignment occurs when the camera is not pointing through the center of eye rotation. In chinchillas and other animals with protuberant eyes, one way to ensure alignment is by retracting the eyelids sufficiently to view the full extent of the globe in the camera image, then ensuring the globe is centered in the image frame. To achieve the same goal in a less invasive manner, we employed four 3-mm diameter white LEDs (NSPW300BS, Nichia Tokushima, Japan) positioned around each camera in the form of a right-angle cross, with LEDs facing the eye and equidistant from the center of the lens. When the camera's optical axis is perfectly orthogonal to a flat mirror, the four LED reflections seen in the video image are centered on and equidistant from the center pixel of the image. If the camera is tilted with respect to the mirror, the LED reflections are no longer equidistant from each other, and the center of the cross they define is no longer aligned with the center pixel of the image. When this system is used to align the camera with an approximately spherical, symmetric, convex, reflective surface (such as a chinchilla eyeball), the camera's optic axis is directed through the surface's center of curvature when the LED reflections are equidistant from the center pixel of the image. The greater the curvature of the reflective surface, the more sensitive this method is at detecting misalignment.

Fourteen chinchilla eyes were dissected after euthanasia and measured using calipers, revealing a mean globe diameter of 13.6 ± 0.6 mm. The camera/lens system employed has a resolution at the eye surface of 59 pixels/mm. For these parameters, when the LEDs are positioned equidistant (34 mm) from the center of the lens, a minimum detectable 1-pixel shift of the LED-cross image center occurs for a 0.32 mm translation of the camera perpendicular to its optic axis.

2.3. Camera angular alignment with the eye

Pure angular misalignment occurs when the camera is pointing through the center of the eye, but the camera coordinate system is not rotationally aligned with the reference coordinate system. In our experiments, the reference coordinate system aligns with the animal's head (specifically, with the plane through the animal's horizontal semicircular canals, the midsagittal plane and the coronal plane orthogonal to these), which in turn is aligned by virtue of a rigid head mount with the Fick gimbal superstructure device holding the animal. To ensure correct angular orientation of the cameras with respect to the gimbal (and thus to the animal's head), each camera is mounted on a rotating turret able to adjust only in azimuth and elevation. Each turret maintains the torsional position of the camera at 0° with respect to the superstructure (and head) coordinate system. The azimuth and elevation of the camera with respect to the superstructure are measured with a protractor to an accuracy of 0.25° . These angles are used to convert measured eye rotations into the superstructure coordinate frame of reference. Since conver-

sion from camera coordinates to head coordinates relies on the known orientation of the former with respect to the latter, improper measurement of camera orientation with respect to the superstructure would cause a predictable error of eye rotational positions calculated from VOG data. This error can be corrected by multiplying each rotation matrix representing instantaneous eye position by the matrix describing rotation from the incorrect frame of reference to the correct frame of reference.

Mixed angular plus translational misalignment occurs when the camera is first brought into perfect alignment and then is rotated so that the camera's optic axis changes direction and no longer points through the center of rotation of the eye. The error due to this combined misalignment can be corrected post hoc if the translational and angular deviation from ideal camera position are known (see below).

2.4. Image acquisition

The VOG system was installed on a Pentium IV 2.4 GHz, 1 GB RAM PC, running Windows 2000. IEEE 1394 Firewire webcams (PYRO1394 WebCam, ADS Technologies, USA) retrofitted with 1/4" format 16.0 mm focal length, $f/2.0$ C-mount board lens (BL160, Allthings Inc., Australia) were used to acquire 640 pixels \times 480 pixels B&W (eight-bit) images at 30 Hz for each eye. For use with the 16 mm lens, a 5-mm plastic spacer was placed between the lens housing and the printed circuit board of the webcam, which allowed the camera with 16 mm lens to be focused on a point 50 mm away. National Instruments LabVIEW 7.0, NI-IMAQ Vision 7.0.1 and NI-IMAQ for IEEE 1394 Cameras 1.5 standard modules were used to control camera settings such as contrast and brightness. Standard NI-IMAQ modules were used to change the image threshold so that only the three bright markers were visible on a black background and to determine the center of each marker using a center of mass algorithm. Camera magnification is set so that the medial and lateral canthi are at the edges of the video image. Pixel size was calibrated using the known distance between markers and verified by using a micrometer.

2.5. Image processing

Assuming the eye is a sphere that rotates about its center (or more precisely, that as the marker array moves with the eye, it travels along a spherical surface of radius approximately equal to that of the globe), assuming the eye is centered on the camera's optical axis, and defining the center of the eye as the origin of a coordinate system (i, j, k), one can calculate the position in space of each marker. The marker array can be positioned anywhere on the eye as long as it remains visible during eye rotations. The i, j and k axes measure *translation* of each marker in space, with j and k equaling the horizontal and vertical positions of each marker (measured in pixels) from the center of the video image, and i being the distance from the globe center to the marker along the optic axis of

the camera. The i coordinate is calculated from j and k and the known radius of the eye as follows:

$$i = \sqrt{((\text{eye_radius_in_pixels})^2 - (j^2 + k^2))}$$

The rotation matrix \mathfrak{R} uniquely describing the eye rotation required to move the three markers from one position to another is:

$$\mathfrak{R} = \begin{bmatrix} i_0 & i_1 & i_2 \\ j_0 & j_1 & j_2 \\ k_0 & k_1 & k_2 \end{bmatrix}_{\text{current}} \times \begin{bmatrix} i_0 & i_1 & i_2 \\ j_0 & j_1 & j_2 \\ k_0 & k_1 & k_2 \end{bmatrix}_{\text{ref}}^{-1}$$

The subscript ‘ref’ refers to the marker position before a rotation and defines the reference or zero *rotational* position of the eye. The subscript ‘current’ refers to the marker positions in 3-D space after a rotation. Euler angles, rotation vectors and quaternions are calculated directly from the rotation matrix (Haslwanter, 1995; Migliaccio and Todd, 1999). Using 640 pixels \times 480 pixels cameras and magnification optimized so that the marker array range of motion filled the camera image frame, the absolute resolution of the system (assuming minimum detectable image shift of one pixel) is $<0.2^\circ$.

The above algorithm depends on alignment of the camera center with the center of the eye, which we ensure using the techniques described above. If unusual circumstances dictate that the camera cannot be precisely aligned with eye (e.g. in an experimental apparatus with limited space for camera placement), translational and angular misalignment can be corrected post hoc if the misalignment is known. The correction must be in the following sequence. First, translational misalignment can be corrected by redefining the image center so that rather than using the default (center pixel of the camera image), the image pixel that aligns with the center of rotation of the eye is defined as the origin. Second, angular misalignment can be corrected by multiplying each eye rotation matrix by the inverse rotation matrix describing camera rotational position in the reference coordinate frame (whether head or superstructure/test-apparatus coordinates).

Tracking of one or more markers may be transiently lost during a blink, burst of saccades or extreme eye deviation that moves the markers behind the retracted eyelid or into a poorly illuminated region. Upon subsequent reacquisition of the marker image, potential uncertainty exists regarding which marker is which. One approach to resolving this ambiguity would be to separately track each marker using different colors, shapes, sizes or relative positioning in the marker array. Alternatively, if the eye is assumed to stay within the $\sim 45^\circ$ oculomotor range, then only one of six possible permutations gives the correct pairing of all three markers from one image to the next, even if intervening images have been lost. We determined the correct pairing by calculating the summed square of marker travel distances for each permutation and accepting the permutation that resulted in the smallest value. The correct permutation was always selected, regardless of the cause or duration of transient marker image loss.

Binocular video analysis is computationally intensive, so we optimized our code to use NI-IEEE 1394 software interrupts, freeing the CPU for other processing until a new video image was acquired. Camera output was time-shifted to account for the delay between image acquisition at the CCD and arrival of the new image in PC memory. This delay was found to be 33.2 ± 0.1 ms by measuring the time delay between actual and VOG-derived motion of a digitally controlled motor turning a simulated eye tracked by the VOG system continuously for 5 min.

To maximize angular resolution during in vivo testing, we used a 16 mm focal length lens (All Things Inc., Australia) in place of the standard 4 mm webcam lens. This provided sufficient magnification (for a camera-to-eye center distance of 50 mm) such that the palpebral fissure filled the camera image. The operating range (over which images can be focused) for the 16 mm lens was 40–300 mm when a 5 mm plastic spacer was placed between the webcam’s C-mount and the printed circuit board to which it normally attaches. To faithfully replicate this arrangement during in vitro calibration tests using a precision-machined Fick gimbal with an effective marker-travel radius of curvature of 31 mm ($4.5\times$ that of the chinchilla eye), we used the standard 4 mm focal length lens without a spacer and positioned the camera 242 mm from the center of rotation of the gimbal.

We used the NI-IMAQ “Learn Calibration Template” standard module to quantify and correct pincushion, barrel, perspective and non-radial distortion due to the lens. We constructed a calibration grid composed of black filled circles on a white background. The circles were positioned at each intersection on a square grid. The saved video image of the calibration grid, in addition to the known distance between circles, were input to the NI-IMAQ “Learn Calibration Template” module to calibrate each lens. The 4 mm (calibration grid at 242 mm, grid spacing: 9 mm, circle radius: 2 mm) and 16 mm (calibration grid at 50 mm, grid spacing: 2 mm, circle radius: 0.5 mm) lenses produced maximum pixel location deviations of <1.8 and $<0.1\%$, respectively. Distortion due to the 16 mm lens was negligible, so that recalibration of the 16 mm lens, unlike the 4 mm lens, had little or no effect on the VOG system output.

2.6. Search coil system

For each eye, two insulated copper watch coils (2 mm diameter insulated copper, 80 turns, ~ 0.003 mg) were glued together with cyanoacrylate and positioned orthogonal to one another (as in Gilchrist et al., 1998). Lead wires from the coil pair were tightly twisted to minimize artifacts generated by the movement of loops in the lead wire within the magnetic fields. Lead connectors were kept fixed with respect to fields during calibration and during the experiment.

Three pairs of field-generating coils, each with a side length of 45 cm, were rigidly attached to a superstructure that moved with the animal. The three magnetic fields were mutually orthogonal and aligned with the X (naso-occipital), Y

(interaural) and Z (superior–inferior) coordinate axes. The X, Y and Z fields oscillated at 100, 50 and 75 kHz, respectively. The three amplitude-modulated frequency signals induced in each coil were demodulated to produce three voltages proportional to the angles between each coil and each magnetic field (Remmel, 1984). A full description of the calibration and signal processing methods applied to these signals has been described elsewhere (Straumann et al., 1995; Migliaccio et al., 2003). The peak-to-peak noise at the output of the circuit was equivalent to an eye movement of 0.02° . All signals transducing motion of the head or the eye were passed through eight-pole Butterworth anti-aliasing filters with a corner frequency of 100 Hz. Coil signals were digitized at 16-bit resolution at a sampling rate of 1 kHz.

2.7. *In vitro* experiment: Fick gimbal and VOG

Although precision machined and accurate to 0.1° , the Fick gimbal we employed for calibration dictated that during *in vitro* calibration, the marker array was 31 mm away from the center of rotation of the Fick gimbal (equivalent to $4.5\times$ the radius of a chinchilla eye). In order to faithfully simulate the *in vivo* case, the *in vitro* marker array was scaled to $4.5\times$ larger than the *in vivo* marker array, and we used a 4 mm focal length lens (instead of the 16 mm lens used *in vivo*) to make the pixel size of the marker array *in vitro* the same as *in vivo* (each marker was ~ 64 pixels \times 64 pixels). The camera was placed 242 mm directly in front of the Fick gimbal center, with the image center aligned with the center of the gimbal. The gimbal was rotated in 10° steps from -20 to $+20^\circ$ in all combinations of yaw, pitch and roll.

To quantify effects of *translational* misalignment, measurements were then repeated with the camera deliberately translated horizontally so that the difference between the gimbal center and image center was either 17 or 34% of the gimbal radius (equivalent to 1.1 or 2.2 mm off the center of a chinchilla eye, both much larger than the 0.32 mm precision of the 4-LED aligning technique). Measurements were again repeated with each camera intentionally rotated 1.3 and 2.5° in yaw so as to create combined *angular* horizontal misalignments of 1.3 and 2.5° in addition to 1.1 or 2.2 mm translational misalignment.

2.8. *In vivo* experiment: combined search coils and VOG

We simultaneously measured eye movements mediated by the vestibulo-ocular reflex in response to head rotations in a healthy normal chinchilla (*Chinchilla laniger*) with both VOG and search coils to verify the accuracy of the VOG system. A head bolt was placed surgically under sevoflurane anesthesia for animal restraint during subsequent eye movement testing. The dorsal surface of the skull was exposed, both bullae were opened using an otologic drill, and a lightweight ceramic rod was positioned on the skull in the midline and embedded in a dental acrylic cap extending from one bulla to the other. The ceramic rod was oriented so that

when the animal was placed into a plastic restraining device contained within a gimbal superstructure, the animal's horizontal semicircular canal plane aligned with the horizontal plane of the gimbal (and of the VOG cameras), which could then be rotated so as to bring any of the vestibular semicircular canals into the earth-horizontal plane for rotational testing. The gimbal structure was mounted on a servo-controlled motor (model 130-80/ACT2000; Acutronic USA Inc, Pittsburgh, PA) positioned so that the center of the animal's skull rotated about the motor's earth-vertical axis. All surgical and other animal care procedures were done in accordance with a protocol approved by the Animal Care and Use Committee of the Johns Hopkins University School of Medicine.

Cameras were rigidly fixed to the test apparatus and positioned along the corneoretinal axis of each eye in its resting position ($\sim 70^\circ$ off the naso-occipital axis in the horizontal plane). Magnification was set so that the medial and lateral canthi were both visible in the video image. Alignment was achieved by: (a) positioning the center of the cross defined by the four LED reflections so that it aligned with the center pixel of the image and, (b) ensuring that the LED reflections were equidistant from the image center and aligned on the vertical and horizontal image midlines. The horizontal and vertical angles between the camera and the test apparatus were measured with a protractor and recorded. Eyelids were retracted by the minimum amount necessary to image the marker array throughout the oculomotor range. Each eye received two drops of anesthetic (proparacaine hydrochloride 0.5%) before the marker arrays and search coils were affixed approximately 2 mm apart. The pupil was occluded by a thin film of cyanoacrylate. The testing chamber was dark apart from the UV source.

The animal was rotated at 0.5, 1, 2 and 5 Hz with peak velocity 20 or $50^\circ/\text{s}$ in the plane of each semicircular canal pair: yaw (horizontal canals), LARP (left anterior and right posterior canals) and RALP (right anterior and left posterior canals). The animal was also rotated about its naso-occipital axis (roll) and its interaural axis (pitch). At these frequencies and velocities, the marker array was at no time obscured by the 3-D coil; however, at 0.5 and 1 Hz, the 3-D coil did at times impact the eyelids and/or canthi. The eye movements generated by the vestibulo-ocular reflex should be approximately equal in magnitude, but opposite in direction to the head rotation (Ewald, 1892; Merwin et al., 1989). In rodents, the reported compensatory eye movement is ~ 60 – 80% of the head rotation (Stahl et al., 2000; Kaufman, 2002).

3. Results

3.1. *In vitro* experiments: Fick gimbal and VOG

Table 1 shows the difference between the actual gimbal angle and the angle as measured by VOG. For horizontal, vertical and torsional Fick angles from -20 to 20° , the root mean square positional VOG error was 2.4, 2.6 and 2.9%,

Table 1

VOG angle (in degrees) measurements (horizontal, vertical, torsional) for Fick gimbal angles from -20 to 20° in 10° steps, after geometric compensation due to distortion produced by the 4 mm standard webcam lens

Degrees		20 (left)	10 (left)	0	-10 (right)	-20 (right)
20 (down)	20 (CW)	20.36, 19.76, 19.61	10.21, 19.72, 20.17	0.35, 19.70, 20.42	-9.90, 19.75, 20.43	-19.84, 19.83, 20.49
	10 (CW)	20.18, 19.72, 9.68	10.01, 19.59, 9.69	-0.01, 19.57, 9.94	-10.07, 19.62, 9.78	-20.21, 19.68, 9.59
	0	20.12, 19.68, -0.97	9.95, 19.54, -0.99	-0.04, 19.53, -0.33	-10.13, 19.58, 0.18	-20.31, 19.66, 0.11
	-10 (CCW)	19.77, 20.44, -10.52	9.89, 19.41, -10.34	-0.11, 19.39, -10.83	-10.49, 19.44, -10.49	-20.46, 19.51, -10.40
	-20 (CCW)	19.66, 20.44, -20.54	9.82, 20.32, -20.28	-0.23, 20.32, -20.63	-10.17, 20.37, -20.11	-20.28, 20.34, -20.43
10 (down)	20 (CW)	20.41, 9.81, 19.91	10.27, 9.82, 20.05	0.18, 9.90, 20.39	-9.81, 9.89, 20.39	-19.83, 9.82, 20.96
	10 (CW)	20.08, 9.57, 9.54	9.99, 9.49, 9.87	0.00, 9.46, 10.18	-9.98, 9.45, 10.36	-19.98, 9.48, 10.64
	0	19.73, 10.47, -0.22	9.83, 10.40, -0.50	-0.01, 9.67, -0.23	-10.05, 10.38, 0.02	-20.23, 10.37, -0.08
	-10 (CCW)	19.75, 9.36, -10.87	9.88, 9.29, -10.75	-0.16, 9.26, -10.54	-10.27, 9.26, -10.33	-20.53, 9.28, -10.27
	-20 (CCW)	19.72, 9.94, -20.83	9.83, 9.88, -20.41	-0.10, 9.86, -20.29	-10.29, 9.87, -20.97	-20.35, 9.88, -20.92
0	20 (CW)	20.48, -0.22, 20.49	10.29, -0.25, 20.26	0.21, -0.28, 20.11	-9.87, -0.23, 20.13	-19.86, -0.28, 20.43
	10 (CW)	20.12, -0.22, 10.07	9.95, -0.24, 10.07	0.10, -0.25, 9.99	-9.96, -0.22, 10.02	-19.90, -0.27, 10.11
	0	19.84, 0.03, -0.27	9.83, 0.02, -0.30	0.00, 0.00, 0.00	-10.01, -0.05, 0.03	-20.17, -0.12, -0.19
	-10 (CCW)	19.79, -0.21, -10.36	9.81, -0.13, -10.24	-0.17, -0.16, -10.14	-10.39, -0.10, -10.36	-20.53, -0.28, -10.46
	-20 (CCW)	19.60, -0.04, -21.01	9.76, -0.04, -20.82	-0.17, -0.07, -20.31	-10.27, -0.12, -20.84	-20.63, -0.17, -20.04
-10 (up)	20 (CW)	20.39, -10.32, 20.26	10.16, -10.25, 20.94	0.15, -10.26, 20.41	-9.81, -10.33, 20.14	-19.90, -10.38, 19.92
	10 (CW)	20.31, -10.29, 10.97	10.08, -10.23, 10.66	0.03, -10.26, 10.18	-10.03, -10.22, 9.95	-19.96, -10.33, 9.95
	0	20.02, -9.62, 0.66	9.76, -9.59, 0.46	-0.01, -9.59, -0.05	-10.18, -9.69, -0.31	-20.34, -9.81, -0.59
	-10 (CCW)	19.73, -10.17, -9.40	9.85, -10.22, -9.47	-0.20, -10.24, -10.04	-10.21, -10.11, -10.36	-20.36, -10.23, -10.65
	-20 (CCW)	19.64, -10.29, -20.09	9.80, -10.23, -20.14	-0.13, -10.14, -20.45	-10.25, -10.12, -20.97	-20.36, -10.26, -20.39
-20 (up)	20 (CW)	20.25, -20.40, 20.60	10.17, -20.32, 20.63	0.10, -20.30, 20.33	-9.90, -20.36, 20.59	-19.88, -20.28, 20.12
	10 (CW)	20.66, -20.33, 10.08	10.13, -20.05, 10.23	0.03, -20.26, 10.88	-9.89, -20.30, 10.44	-20.33, -20.52, 9.73
	0	20.26, -20.47, 0.48	9.96, -20.42, 0.33	0.03, -20.24, 0.06	-10.18, -20.45, -0.13	-20.44, -19.69, -0.93
	-10 (CCW)	20.01, -20.61, -9.08	9.89, -20.52, -9.65	-0.06, -20.35, -9.88	-10.20, -20.36, -10.48	-20.34, -20.37, -10.26
	-20 (CCW)	19.65, -20.12, -19.66	9.81, -20.26, -20.37	-0.17, -20.47, -19.68	-10.20, -20.11, -20.11	-20.30, -20.32, -20.05

The positional error for each of the three dimensions is $<2.9\%$.

respectively. For horizontal, vertical and torsional Fick angles from -30 to 30° , the root mean square positional VOG error was 2.6, 2.7 and 3.3%, respectively. The resolution of the VOG system (due to image pixelization) was $<0.2^\circ$. With the gimbal stationary, the velocity noise was $<3^\circ/\text{s}$ RMS ($<5^\circ/\text{s}$ peak) in each dimension.

For excursions within -20 to 20° of center with *translational* horizontal misalignment of the camera equal to 17% of the gimbal radius (equivalent to a 1.1 mm translational offset in vivo, or about 3.6 times the error of the camera aligning system), the error in positional horizontal, vertical and torsional Fick components measured by VOG was 3.5, 4.1 and 5.4%, respectively. For translational misalignment equal to 34% of the gimbal radius, the VOG errors were 6.9, 8.3 and 10.7%, respectively. Translational misalignment of 32% of the gimbal radius was required to elicit positional errors $>10\%$.

For excursions within -20 to 20° of center with *combined angular and translational* horizontal misalignment of the camera equal to 1.3° and 1.1 mm, the error in positional horizontal, vertical and torsional Fick components measured by VOG was 3.8, 5.9 and 6.3%, respectively. For combined misalignment of 2.5° and 2.2 mm, the positional VOG errors were 7.7, 10.6 and 12.3%, respectively. Combined angular and translational misalignment of 2.0° and 1.6 mm was required to reach positional errors $>10\%$.

Errors created by deliberate horizontal misalignment are greater in vertical and torsional directions than in the horizontal direction. This occurs because after the camera is translated or angled horizontally, purely vertical movements seem to have torsional components and vice versa. Conversely, when the camera is translated or angled vertically, purely horizontal movements seem to have torsional components and vice versa. Similarly, when the camera is deliberately misaligned torsionally, purely horizontal movements seem to have vertical components and vice versa.

3.2. In vivo experiment: combined search coils and VOG

Measured eye movement responses to yaw and roll head rotations at 2 Hz with peak velocity of $50^\circ/\text{s}$ in head coordinates are shown in Fig. 2A and B. To facilitate comparison between coil and VOG measurements, the horizontal, vertical and torsional components of eye velocity are plotted in separate panels. Despite the fact that derivation of eye positions from search coil signals and from VOG data rely on completely independent computations, 3-D eye velocity measured using both methods matched closely. During head rotational stimuli with frequencies ≤ 5 Hz, the average RMS *position* difference between the two methods was 2.1, 2.6 and 3.1% for the horizontal, vertical and torsional components of eye position, respectively. The average RMS *velocity* difference between the two methods was 2.2, 3.7 and 4.1% for the horizontal, vertical and torsional components of eye velocity, respectively. The phase difference of the eye movement response measured using coils and VOG was $<2^\circ$ for eye

movements ≤ 2 Hz, and $<7^\circ$ for eye movements ≤ 5 Hz. A sinusoidal fit to the data revealed no statistical difference in the offset ($P=0.83$) or amplitude ($P=0.77$) between the coil and VOG data. A half cycle analysis of the VOG data revealed no statistical difference ($P=0.63$) in phase between the rising and falling portion of the sinusoidal response.

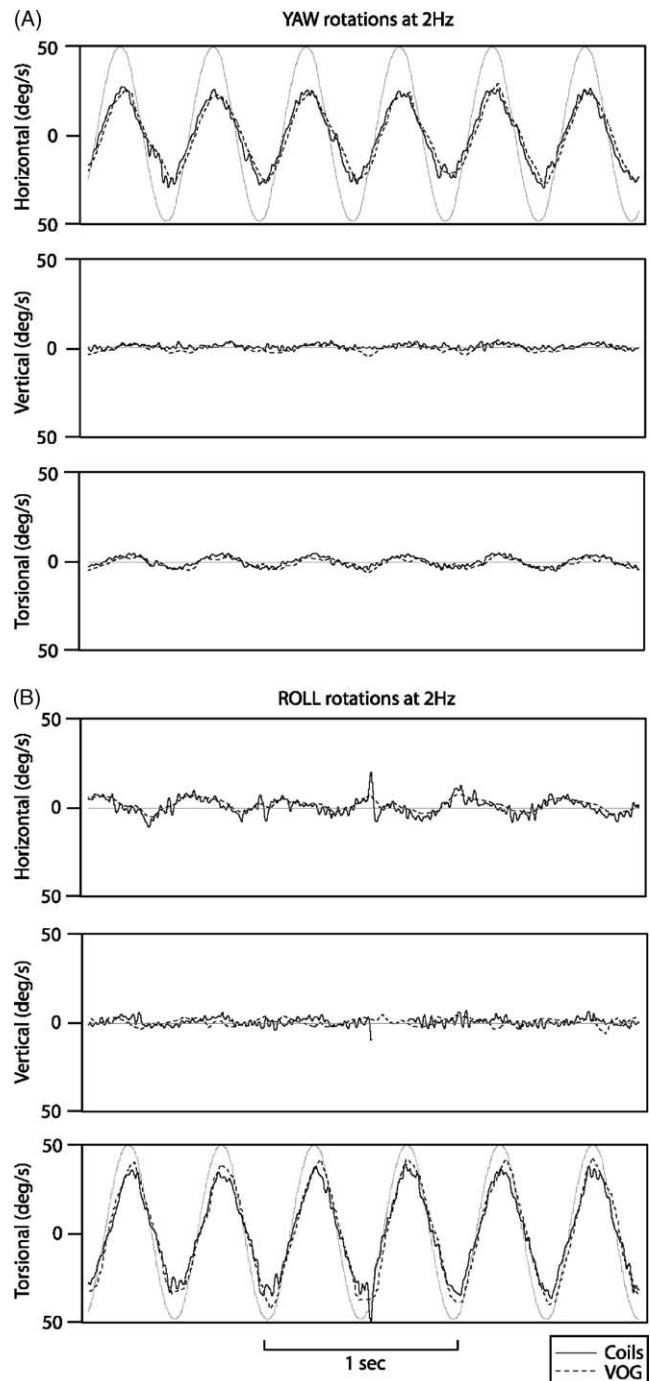


Fig. 2. The eye movement responses to yaw (A) and roll (B) head rotations at 2 Hz with peak velocity of $50^\circ/\text{s}$, in head coordinates. The horizontal, vertical and torsional components of eye velocity are shown in separate panels. The peak difference in eye velocity measured using the coil and VOG method is $<4.1\%$ for each component.

Being a lateral-eyed animal, the chinchilla's resting eye position is $\sim 70^\circ$ off the naso-occipital axis, so during purely roll *head* rotations, *eye* velocity had large vertical and torsional components in *eye* coordinates (camera coordinates). In one eye, the vertical and torsional eye velocity components were in phase with each other, whereas in the other eye they were in anti-phase. When the frame of reference was changed from eye to head coordinates, equivalence of the eyes' movement became apparent, as expected for eye movements mediated by the vestibulo-ocular reflex (Ewald, 1892). The roll (torsion in head coordinates) eye velocity was large and comparable to head roll velocity, while horizontal and vertical eye velocity components were small.

4. Discussion

Using inexpensive consumer grade webcams sampling at 30 Hz, the VOG method we present accurately measures eye position during eye rotations up to 2 Hz. For horizontal, vertical and torsional Fick angles from -20 to 20° , the root mean square positional VOG error was 2.4, 2.6 and 2.9%, respectively. The angular position resolution of the system is $<0.2^\circ$ for each component and the velocity noise is $<5^\circ/\text{s}$ peak ($3^\circ/\text{s}$ RMS) for each component.

VOG and search coil techniques possess distinct sets of advantages and disadvantages. Implanting scleral search coils requires delicate surgery and can alter eye movements due to surgical trauma. Very fine wire can be used to make coils small, but then the experimental window of opportunity is limited, because the coils tend to break a few weeks after implantation. Apart from placement of a marker array on the eye (which could instead be incorporated into a standard contact lens for humans), VOG techniques like that described here are less invasive or completely non-invasive, and there is less likelihood of restriction of normal eye movements (Stahl et al., 2000). Although glue-on marker arrays share some of the same problems as glue-on 3-D coils, the marker arrays we use are approximately one tenth the thickness of our 3-D glue-on coils (0.2 mm versus 2 mm). The marker array can slip under the eyelids and canthi without altering the eye movement, whereas the glue-on 3-D coil cannot. Although still very light, the glue-on 3-D coil weighs approximately 16 times more than the marker array (4.03 mg versus 0.28 mg) and it has a greater moment arm. In addition, the glue-on 3-D coil lead may also impede eye movements.

The VOG technique has a different set of disadvantages. The camera must be fixed with respect to the head, making head-free experiments on small animals difficult (although head-fixed VOG is well developed for humans (Moore et al., 1996)). Spatial resolution is limited by the size of the video pixel, and as a consequence, quantization noise may be higher than with coils. Resolution can be improved somewhat by enlarging the percentage of the video image occupied by the marker array, but only to a limit, because the field of view must remain large enough to contain the marker

array for extremes of eye position. Increased CCD resolution is also possible, though at increased cost and with increased computational overhead. Temporal resolution is also limited in the inexpensive system described here, which samples video images at 30 Hz for each of two cameras (one per eye). We have also implemented a higher performance alternative system using high-speed industrial-grade CMOS IEEE1394/Firewire cameras (Scorpion II, Point Grey Research, USA), at a cost ~ 10 times that of the consumer-grade "webcams" described above. Due to the "plug and play" nature of the IEEE1394/Firewire bus standard, binocular real-time VOG was achieved at 120 Hz sample rate on the personal computer described earlier without change in spatial resolution or interface software. Depending on the particular application, these more expensive cameras also permit trade-off of temporal and spatial resolution. For example, a sample rate of 240 Hz was made possible by decreasing the image resolution (or equivalently, the image area) by 50%.

In theory, only two markers should be required to determine 3-D eye rotational position. Early versions of this VOG system calculated eye position using a marker array with only two markers. However, we found that two-marker methods required more complex and CPU-intensive algorithms than the three-marker technique described here. In addition, two-marker methods were more susceptible to errors when real world geometry deviated from ideal. The addition of one redundant marker resulted in less noisy, more accurate and more robust 3-D eye position measurements.

Like all VOG techniques, the technique we present relies upon a number of simplifications, including the assumption that the eye rotates about a single point in space that does not translate. A detailed discussion of this and other related potential errors of VOG has been presented by Moore et al. (1996). Our technique does not rely upon an assumption that the eye is a sphere, even though our examination of dissected chinchilla eyes revealed that this would be a very good approximation. Instead, we assume that the marker array remains a fixed distance from a fixed center of rotation, and that the center of rotation is coincident with the center of rotation of the eye. While the distinction is moot for chinchillas (which have relatively spherical eyes), it would be important for species in which the cornea bulges beyond the approximately spherical surface defined by the rest of the globe (e.g., humans).

In humans and non-human primates, the subject can be taught to look at the center of the camera lens, thereby aligning the centers of curvature of the cornea and globe on the camera's optic axis when the 4-LED technique is employed. In species with prominent corneal eccentricity, the 4-LED could still be employed by manually rotating the eye so that the four LED reflections appear on the sclera, rather than cornea. The fact that corneal eccentricity was negligible in the animals we used was corroborated by the fact that while the positions of LED images were very sensitive to angular or translational movement of the camera with respect to the eye, they were insensitive to rotation of the eye itself, regardless

of whether the camera center pixel was imaging the cornea or sclera.

Most of our vestibular eye movement experiments require absence of visual input, so we intentionally position the opaque marker array and cyanoacrylate film over the pupil and iris to obstruct vision. If necessary for applications requiring unobstructed vision, the marker array may be positioned to the side of the pupil. This is straightforward using our current marker arrays in chinchillas. It could be more difficult in much smaller animals (e.g., mice), but could be accomplished by scaling down marker array sizes and scaling up camera magnification. Unobstructed vision can be confirmed by examination of the dilated pupil under infrared illumination and testing responses to optokinetic stimuli.

While one could forego the use of an affixed marker array and instead tattoo or etch markers onto the scleral conjunctiva, a method to increase marker signal-to-noise ratio above that of corneal reflections, such as UV-excited fluorescence used here, is helpful. Alternatively a marker array using anti-Stokes (or ‘up-converter’) fluorescent pigments may be used with infrared illumination (Watts, 1974; Wright, 1976). In contrast to most fluorescent pigments, which can only emit wavelengths longer than those they absorb, anti-Stokes pigments emit visible light under infrared illumination (inorganic oxisulphide emits at 550 nm). As with the UV-fluorescent markers detailed here, the spectral shift between the infrared illumination source and the marker emissions still allows the use of optical filters to remove corneal reflection artifacts. This approach has the additional benefit that the marker array need not be positioned over the pupil for experiments that require absence of visual input, because the infrared light sources are not visible. Another possibility is the construction of a silicone scleral “contact lens,” similar to those used in human studies (Robinson, 1963), with fluorescent markers embedded into it. With such an approach, this accurate, inexpensive VOG technique could be used to measure real-time binocular three-dimensional eye position in humans.

Acknowledgements

We gratefully acknowledge Professor Ian S. Curthoys of the Vestibular Research Laboratory, Sydney University, Australia, for his inspiration and pioneering efforts in video-oculography development. Supported by the National Institute on Deafness and other Communication Disorders (K08-DC006216 to CCDS, R01-DC2390 to LBM) and by an American Otological Society Clinician-Scientist Award to CCDS.

References

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

- Clarke AH, Teiwes W, Scherer SH. Videooculography: an alternative method for measurement of three-dimensional eye movements. In: Schmid R, Zambbarbieri D, editors. *Oculomotor control and cognitive processes*. Amsterdam: Elsevier; 1991. p. 431–43.
- Collewijn H, Van der Steen J, Ferman L, Jansen TC. Human ocular counterroll: assessment of static and dynamic properties from electromagnetic scleral coil recordings. *Exp Brain Res* 1985;59:185–96.
- Demer JL, Miller JM, Poukens V, Vinters HV, Glasgow BJ. Evidence for fibromuscular pulleys of the recti extraocular muscles. *Invest Ophthalmol Vis Sci* 1995;36:1125–36.
- Ewald JR. *Physiologische untersuchungen uber das endorgan des nervus octavus*. Wiesbaden, Germany: Bergmann; 1892.
- Gilchrist DP, Curthoys IS, Cartwright AD, Burgess AM, Topple AN, Halmagyi M. High acceleration impulsive rotations reveal severe long-term deficits of the horizontal vestibulo-ocular reflex in the guinea pig. *Exp Brain Res* 1998;123:242–54.
- Haslwanter T. Mathematics of three-dimensional eye rotations. *Vis Res* 1995;35:1727–39.
- Hess BJ, Dieringer N. Spatial organization of linear vestibuloocular reflexes of the rat: responses during horizontal and vertical linear acceleration. *J Neurophysiol* 1991;66:1805–18.
- International Commission on Non-Ionizing Radiation Protection. Guidelines on UV radiation exposure limits. *Health Phys* 1996;71:978.
- Kaufman GD. Video-oculography in the gerbil. *Brain Res* 2002;958:472–87.
- Merwin Jr WH, Wall III C, Tomko DL. The chinchilla’s vestibulo-ocular reflex. *Acta Otolaryngol* 1989;108:161–7.
- Migliaccio AA, Todd MJ. Real-time rotation vectors. *Australas Phys Eng Sci Med* 1999;22:73–80.
- Migliaccio AA, Cremer PD, Aw ST, Halmagyi GM, Curthoys IS, Minor LB, Todd MJ. Vergence-mediated changes in the axis of eye rotation during the human vestibulo-ocular reflex can occur independent of eye position. *Exp Brain Res* 2003;151:238–48.
- Minor LB, Lasker DM, Backous DD, Hullar TE. Horizontal vestibuloocular reflex evoked by high-acceleration rotations in the squirrel monkey. I. Normal responses. *J Neurophysiol* 1999;82:1254–70.
- Moore ST, Curthoys IS, McCoy SG. VTM: an image-processing system for measuring ocular torsion. *Comput Methods Prog Biomed* 1991;35:219–30.
- Moore ST, Haslwanter T, Curthoys IS, Smith ST. A geometric basis for measurement of three-dimensional eye position using image processing. *Vis Res* 1996;36:445–59.
- Nakayama K. Photographic determination of the rotational state of the eye using matrices. *Am J Optom Physiol Opt* 1974;51:736–42.
- Ott D, Gehle F, Eckmiller R. Video-oculographic measurement of 3-dimensional eye rotations. *J Neurosci Methods* 1990;35:229–34.
- Paige GD, Tomko DL. Eye movement responses to linear head motion in the squirrel monkey. I. Basic characteristics. *J Neurophysiol* 1991;65:1170–82.
- Parker JA, Kenyon RV, Young LR. Measurement of torsion from multi-temporal images of the eye using digital signal processing techniques. *IEEE Trans Biomed Eng BME* 1985;32:28–35.
- Rommel RS. An inexpensive eye movement monitor using the scleral search coil technique. *IEEE Trans Biomed Eng* 1984;31:388–90.
- Robinson DA. A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Trans Biomed Eng* 1963;10:137–45.
- Stahl JS, van Alphen AM, De Zeeuw CI. A comparison of video and magnetic search coil recordings of mouse eye movements. *J Neurosci Methods* 2000;99:101–10.

- Straumann D, Zee DS, Solomon D, Lasker AG, Roberts DC. Transient torsion during and after saccades. *Vis Res* 1995;35:3321–34.
- Vieville T, Masse D. Ocular counter-rolling during active head tilting in humans. *Acta Otolaryngol* 1987;103:280–90.
- Watts RK. Optical properties of ions in solids. In: Di Bartolo B, editor. *Nato Advanced Study Institutes Series*, vol. 8. New York: Plenum; 1974. p. 337.
- Wright JC. Radiation processes in molecules and condensed phases. In: Fong FK, editor. *Topics in applied physics*, vol. 15. New York: Springer; 1976. p. 239.
- Yamanobe S, Taira S, Morizono T, Yagi T, Kamio T. Eye movement analysis system using computerized image recognition. *Arch Otolaryngol Head Neck Surg* 1990;116:338–41.