Measuring and tracking eye movements of a behaving archer fish by real-time stereo vision

Avi Ben-Simon [a,c], Ohad Ben-Shahar [b,c], Ronen Segev [a,c,*]

[a] Department of Life Sciences, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel
[b] Department of Computer Science, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel
[c] Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

ARTICLE INFO

ABSTRACT

The archer fish (*Toxotes chatareus*) exhibits unique visual behavior in that it is able to aim at and shoot down with a squirt of water insects resting on the foliage above water level and then feed on them. This extreme behavior requires excellent visual acuity, learning, and tight synchronization between the visual system and body motion. This behavior also raises many important questions, such as the fish's ability to compensate for air–water refraction and the neural mechanisms underlying target acquisition. While many such questions remain open, significant insights towards solving them can be obtained by tracking the eye and body movements of freely behaving fish. Unfortunately, existing tracking methods suffer from either a high level of invasiveness or low resolution. Here, we present a video-based eye tracking method for accurately and remotely measuring the eye and body movements of a freely moving behaving fish. Based on a stereo vision system and a unique triangulation method that corrects for air–glass–water refraction, we are able to measure a full three-dimensional pose of the fish eye and body with high temporal and spatial resolution. Our method, being generic, can be applied to studying the behavior of marine animals in general. We demonstrate how data collected by our method may be used to show that the hunting behavior of the archer fish is composed of surfacing concomitant with rotating the body around the direction of the fish's fixed gaze towards the target, until the snout reaches in the correct shooting position at water level.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Archer fish (*Toxotes chatareus*) are well known for their ability to shoot down insects resting on foliage above the water level by spitting a jet of water at them. The fish then feed on the insects that have fallen into the water. Accurate shooting can be achieved for targets as small as 3–4 mm at distances of up to 90 cm (Timmermans, 2000, 2001). The fish's eyes, which have an anatomy similar to that of the eyes of most vertebrates, are located near the mouth in a way that allows the fish to use binocular vision (Luling, 1958, 1963). The fish's eye movements are similar to those of terrestrial vertebrates, including man, i.e., they consist of rapid saccade movements with short fixations between them. The process of shooting at an insect may last no longer than a second, during which time the fish identifies its prey, positions its body such that only the tip of its mouth sticks out of the water, and shoots accurately at the target. During this process, the fish has to identify the target as rewardable,

compensate for the air–water refraction, and assess the exact location of the target in space. This extreme visually guided behavior raises numerous questions regarding key issues such as the mechanism(s) for target identification and for learning target identities (Ewert, 1974, 1997; Schuster et al., 2007, 2006), refraction compensation (Barta and Horvath, 2003; Dill, 1977; Schuster et al., 2004; Temple, 2007; Timmermans and Vossen, 2000); the contribution of binocular vision (Dill, 1977; Gonzalez and Perez, 1998; Luling, 1963); and the effect of eye movements on encoding visual stimuli by the fish's visual system (Martinez-Conde et al., 2004; Rucci et al., 2007; Segev et al., 2007). This diverse set of questions can be examined by using the archer fish as a model platform.

To address many of the above questions more rigorously, a better understanding is needed of the actual stimulus that falls on the fish retina during a behavioral task. To facilitate such an understanding, it is necessary to have the ability to measure eye movements of a behaving fish performing controlled behavioral tasks. These eye movement measurements can then be combined with behavioral and electrophysiological measurements or used in isolation to study the fish brain processing mechanism(s). However, measuring archer fish eye movements constitutes a significant technical challenge due to the small size of the fish, the type of

* Corresponding author at: Department of Life Sciences, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel. Tel.: +972 8 647 9226; fax: +972 8 646 1710.
E-mail address: ronensgv@bgu.ac.il (R. Segev).

habitat (aqueous medium), and the fact that the fish rarely exhibits shooting behavior when under stress.

A method commonly used for measuring eye movements in behaving animals is based on a search coil (Robinson, 1963). When the coil is placed at the center of three orthogonal magnetic fields, each oscillating at a different frequency, the voltage induced in the coil is a function of the coil orientation with respect to the field directions in space, and the coil orientation can be measured on the basis of this relationship. Although we have indeed used search coils for recordings of eye movements in a behaving archer fish (Plotkin et al., 2008), two main drawbacks prevent this approach from being truly useful. Firstly, the presence of the coil on the fish eye and their external wiring induce a high level of stress, which significantly affects the archer fish behavior. Secondly, the momentum and friction of the coil in the water significantly affect the fish eye movements (due to the size and mass of the coil).

Another way to obtain measurements of eye movements in animals is to use video-based methods. In general, these methods require a fixed position of the animal's head relative to a video camera. Measurements of horizontal and vertical eye movements can then be achieved by tracking the pupil and/or corneal reflection (Stahl et al., 2000). Measurements of torsional eye movements require the attachment of two or more markers on the animal eye (Migliaccio et al., 2005; Nakayama, 1974). Although such a method has previously been applied in freely moving monkeys (Shepherd and Platt, 2006), it is not suitable for freely moving smaller animals due to the relatively large size of the eye tracking system. For small marine animals, the problem is exacerbated by the nature of their habitat—water, i.e., measurements have to be acquired in an aqueous medium. For a freely moving fish in a water tank, these constraints can be solved by placing two cameras outside the tank for measurement of the eye movements. However, such a set-up introduces the problem of air–glass–water refraction.

To overcome these difficulties in measuring the eye movements of a behaving archer fish, we have developed a novel, non-invasive, remote video-based method that exploits recent developments in computer vision. With this method, we detect and track three reference points on the fish eye by using two high-frequency video cameras positioned outside the water tank. A novel multi-refraction triangulation algorithm is then used to infer both angular and translational eye movements in three dimensions.

Over the past few years, a number of studies have addressed the problem of video-based reconstruction of marine animals inside water (Treibitz et al., 2008; Yekutieli et al., 2007). However, to our best knowledge, our approach is the first to apply such a video-based eye tracking system that can remotely reconstruct full eye and body positions and orientation of a fish. Evaluation of our method's spatial accuracy indicates an estimation precision of ~0.05°, an error that corresponds to an archer fish retinal offset as small as the radius of a photoreceptor (Segev et al., 2007), which is far superior to previously reported methods for non-aquatic animals (Migliaccio et al., 2005). Evaluation of our method's temporal accuracy indicates its ability to detect both fixational eye movements and small saccades. Moreover, we show that the accuracy of the method is limited only by the properties of the cameras and that sub-photoreceptor measurement accuracy can be obtained simply by using cameras with better spatial and temporal resolutions.

In the results section, we also present measurements of the eye movement of a freely moving fish during a behavioral shooting task. These measurements, which were based on analyzing the images of three markers on the fish eye, represent the fish gaze direction and include both eye and body movements. It should be stressed that our method, being generic, can easily be extended to separately measure body and eye movements by attaching three more markers on the fish body. Such procedure would allow a subtraction of the fish body movement from its eye movements in order to obtain *relative* eye movements.

2. Methods

Estimation of eye pose in space requires the three-dimensional (3D) reconstruction of both eye orientation and position. To facilitate this computation using only the image of the eye, we attached three small circular plastic markers to the eye of an archer fish. These markers were then imaged by two video cameras (Flea 2, Point Grey Research, Netcom, Israel) during a behavioral task of the fish. Thereafter, multi-refraction triangulation was applied to resolve both translational and angular fish eye movements. The details of this process and the calibration procedure that enables the transformation between coordinate systems of the cameras and the "world" (i.e., the water tank) are presented below.

2.1. Video capturing the fish during a behavioral task

Capturing the fish eye during a behavioral task through video cameras introduce several constrains regarding the fish location and pose in the water tank during the task. Since the fish eyes are relatively laterally to its head, the cameras must capture its profile; second, the fish behavior should be captured in the smallest field of view possible in order to maximize the resolution of the fish eye image; third, if the cameras have limited focal depth, the fish distance from the cameras should remain in a relatively narrow range. To address these constraints we added a glass panel inside the water tank, parallel to the water tank walls, 3 cm apart, and placed the target between this panel and wall. This configuration enforces the fish to identify and shoot at the target in the small space between the panel and the wall. Therefore, it enables capturing the fish's eye face on, in small volume and in focus, by the two cameras fields of view during each shooting task. An overview of the video system and is presented in Fig. 1.

2.2. Coordinate frames and transformations

The fish's eye (or for that matter, any rigid object) can be considered as a set of physical points in a local eye coordinate frame (F_{eye}). The pose of the eye in 3D space can therefore be regarded as the relative configuration of that local frame in some global and fixed world coordinate frame (F_{world}). The description of this relative configuration is simply the Euclidean transformation that would align F_{eye} with F_{world}. In our video-based system, this transformation is inferred from the projection of the fish eye on the image plane of the camera i.e., it is based on the perspective projection of the fish eye to a local image (or camera) coordinate frame. In our algorithm, *two* local image coordinate frames (F_{l_image} and F_{r_image}) are used, and the transformation from F_{world} to these frames is approximated by a *pinhole camera model* (Hartley and Zisserman, 2004).

We used the following notations in our calculations: vector quantities (points or vectors) are denoted by boldface lowercase letters (e.g., a) and matrix quantities by boldface uppercase letters (e.g., A). Scalar quantities are denoted by Roman lowercase letters. The subscripts e, w, l, and r are used for quantities that correspond to the eye, the world, and the left and right image coordinate frames, respectively. The superscripts w, l, and r are used to describe the coordinates of points and vectors in the coordinate frames of the world, and the left and right images, respectively. To simplify the description of the different coordinate systems and the transformation between them, the following additional conventions are applied (Fig. 2):

- F_{eye}—The eye coordinate frame is derived by using the three reconstructed reference points on the fish eye. We denote these
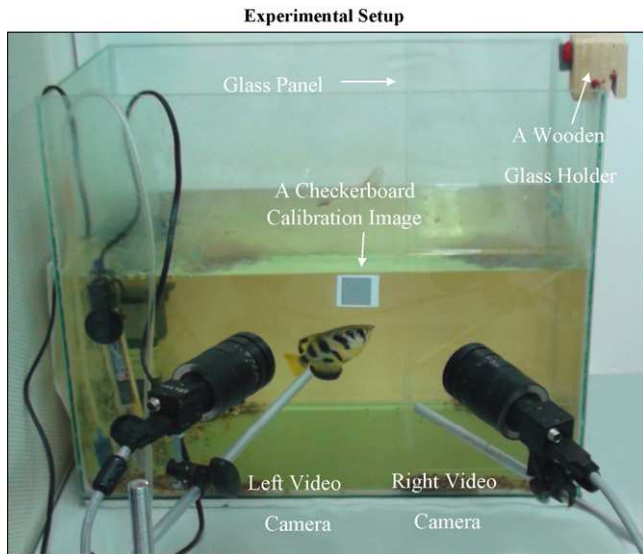
Fig. 1. Configuration of the eye movement measurement system.
Two CCD cameras are directed at the water tank with an overlapping field of view. Each point on the fish eye emits light rays, which propagate through three media – water, glass and air – and are captured by the cameras' CCDs. Projecting the rays backwards from the point of incidence at the camera's CCD through the three media provides a way to locate the 3D position of each point on the fish eye. Each experiment starts by calibrating the cameras, using a checkerboard panel of known dimensions. Before calibration this panel is attached to a surface on the water tank glass wall closest to the cameras, and after calibration the panel is removed. A glass panel is used to restrict the fish movement during shooting task to the vicinity of the checkerboard panel. This restriction ensures that the fish profile remains in a relatively constant distance from the side glass, by preventing it from turning away or towards the cameras during the task. To obtain maximal accuracy in detecting eye movements during a behavioral task, the image of the fish eye should capture as much portion of the cameras field of view as possible. We found that choosing a field of view at each axis of approximately 4 cm on the water tank panel, which is half the mean length of an archer fish, is sufficient for this purpose.

points as **m1**, **m2**, and **m3**, and we use them to define two coplanar vectors: **e1** = **m2** – **m1**, and **e2** = **m3** – **m1**. Since these vectors are not necessarily orthogonal, the eye reference frame (i.e., a set of three orthonormal vectors) is constructed as follows:

$$\mathbf{x}_e = \frac{\mathbf{e2} \times (\mathbf{e2} \times \mathbf{e1})}{\left\| \mathbf{e2} \times (\mathbf{e2} \times \mathbf{e1}) \right\|}; \quad \mathbf{y}_e = \frac{\mathbf{e2}}{\left\| \mathbf{e2} \right\|}; \quad \mathbf{z}_e = \frac{\mathbf{e2} \times \mathbf{e1}}{\left\| \mathbf{e2} \times \mathbf{e1} \right\|} \quad (1)$$

where $\mathbf{x}_e$, $\mathbf{y}_e$, and $\mathbf{z}_e$ are three orthonormal vectors that represent $\mathbf{F}_{eye}$.

- $\mathbf{F}_{l\_image}$—The x and y axes of $\mathbf{F}_{l\_image}$ are set parallel to the x and y directions of the left camera's charge-coupled device (CCD), respectively, while its z axis is set perpendicularly to the x and y axes. The origin of $\mathbf{F}_{l\_image}$ is denoted by $\mathbf{o}_l$.
- $\mathbf{F}_{r\_image}$—The construction of the right image frame is similar to that of $\mathbf{F}_{l\_image}$, but for the right camera's CCD. The origin of $\mathbf{F}_{r\_image}$ is denoted by $\mathbf{o}_r$.
- $\mathbf{F}_{world}$—The origin of $\mathbf{F}_{world}$ is set at the corner of the water tank, while its x and y axes are set parallel to the glass along the principal directions of a rectangular calibration plate fixed to the water tank as shown in Fig. 2 (the calibration procedure is described below). The z axis of $\mathbf{F}_{world}$ is set orthogonal to its x and y axes in the direction away from the water tank towards the cameras.

The relationships among all these coordinate frames and how their relative transformations are inferred or calibrated are described below and in Appendix A.
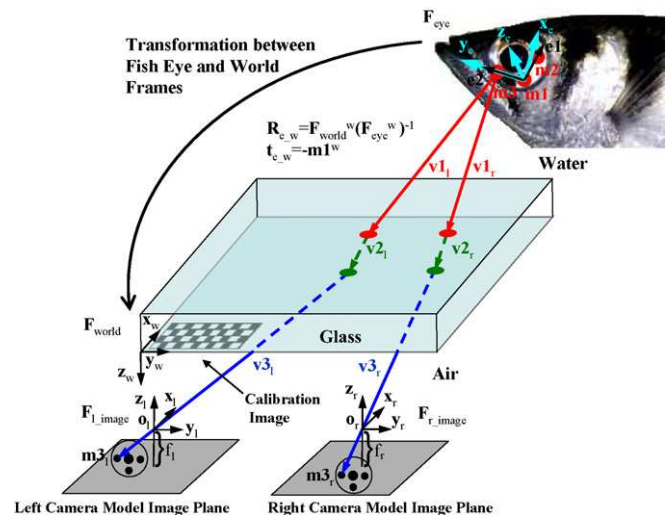

Fig. 2. Schematic drawing of the video-based eye tracking triangulation method. Our method is based on fixing three reference points on the fish eye **m1**, **m2**, and **m3**. From these points, we can define a reference frame, $\mathbf{F}_{eye}$ with axes $\mathbf{x}_e$, $\mathbf{y}_e$ and $\mathbf{z}_e$, that represents the fish eye. From the transformation between $\mathbf{F}_{eye}$ and the water tank frame ($\mathbf{F}_{world}$), the fish eye rotation angles and translation are determined. For this purpose, we need to determine, by triangulation from the two cameras, the exact 3D position of each of the points **m1**, **m2**, and **m3**. The triangulation process starts by considering two light rays that are emitted from a point (e.g., **m3**) on the fish eye inside the water tank and pass through the water (represented by vectors $\mathbf{b}_l$ and $\mathbf{b}_r$), glass ($\mathbf{g}_l$ and $\mathbf{g}_r$) and air ($\mathbf{a}_l$ and $\mathbf{a}_r$) media. In the air, the rays arrive at the two cameras. Each camera is characterized by a pinhole camera model where the camera is represented by a projection center (labeled $\mathbf{o}_l$, and $\mathbf{o}_r$ for the left and right cameras, respectively) and image plane (representing the camera's CCD). The rays pass through the projection centers of the left and right camera models to reach the image planes. The distances between the projection centers and the image planes are the focal lengths of the left and right cameras, designed as $\mathbf{f}_l$, and $\mathbf{f}_r$, respectively. Note that we use three sets of coordinate systems in our analysis: that of the water tank $\mathbf{x}_w$, $\mathbf{y}_w$, $\mathbf{z}_w$, that of the left camera model $\mathbf{x}_l$, $\mathbf{y}_l$, $\mathbf{z}_l$, and that of the right camera model, $\mathbf{x}_r$, $\mathbf{y}_r$, $\mathbf{z}_r$. For reconstruction of the path of the rays, we first need to perform a calibration process for the cameras' parameters. The calibration is performed by projection of a checkerboard calibration image with known dimensions (attached to the water tank outer surface) on the image plane of each camera. The parameters of each camera are then determined from the relation between the known location of the checkerboard grid corners in the water tank coordinate system and the projected grid corners on the camera's pixel space. This calibration is performed separately for each camera (see Appendix A for details). Note that this schematic drawing can be visualized as an upper view of Fig. 1.

2.3. Multi-refraction triangulation

Fig. 2 presents a schematic drawing of our video triangulation method. Each of the two cameras captures a light ray originating from a reference point on the fish eye. According to Snell's law, the two light rays are refracted twice along their path—first at the water–glass interface and then at the glass–air interface. By knowing the pixel coordinates of the projected points in the two camera planes and the configuration of the camera models with regard to the water tank, we can trace the two twice-refracted rays back to their origins and obtain the 3D position of the reference point by estimating the point of intersection of the rays.

As shown in Fig. 2, the ray outside the water tank that reaches the left camera model passes through the projection point of the left camera $\mathbf{o}_l$ and reaches the left image plane at point $\mathbf{m3}_l$. Assuming that we know these two points in the world coordinate system, we can define a vector $\mathbf{a}_l{}^w$ that passes through both of them. Vectors $\mathbf{g}_l{}^w$ and $\mathbf{b}_l{}^w$ depict the trajectory of the subsequent sections of the ray in the glass and in the water, respectively. Based on our convention for $\mathbf{F}_{world}{}^w$, refractions of the rays affect only their z components according to Snell's law. Therefore, calculating the intersection point of $\mathbf{a}_l{}^w$ with the glass and applying Snell's law,

we can easily calculate $\mathbf{g}_l{}^w$. The calculation of $\mathbf{b}_l{}^w$ is performed in a similar fashion, but in this case the refraction taken into consideration is that between the glass and the water at the intersection point of $\mathbf{g}_l{}^w$ with the inner plane of the glass.

The same tracing procedure is performed for the right camera ray whose components are denoted by vectors $\mathbf{b}_r{}^w$, $\mathbf{g}_r{}^w$ and $\mathbf{a}_r{}^w$. The triangulation of the reference points is determined by the point of shortest distance between the two vectors $\mathbf{b}_l{}^w$ and $\mathbf{b}_r{}^w$.

2.4. Rotation angles and translation vector

We determine the fish eye pose in space by using the rotation that is required to align $\mathbf{F}_{world}{}^w$ with $\mathbf{F}_{eye}{}^w$. In technical terms, this requires the recovery of a rotation matrix $\mathbf{R}_{e\_w}$ and a translation vector $\mathbf{t}_{e\_w}$ that solve the equation:

$$\mathbf{F}_{world}^{w} = \mathbf{R}_{ew}\mathbf{F}_{eye}^{w} + \mathbf{t}_{ew} \tag{2}$$

The rotation matrix represents the three rotations around the x, y and z axes of $\mathbf{F}_{world}{}^w$ according to a standard Euler angle rotation matrix $\mathbf{R} = \mathbf{R}_x\mathbf{R}_y\mathbf{R}_z$, where:

$$\mathbf{R}_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\psi) & \sin(\psi) \\ 0 & -\sin(\psi) & \cos(\psi) \end{pmatrix}; \quad \mathbf{R}_y = \begin{pmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{pmatrix};$$

$$\mathbf{R}_Z = \begin{pmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & \sin(\psi) \\ 0 & 0 & 1 \end{pmatrix} \tag{3}$$

and ψ, θ, and ϕ are the angles of rotation around the x, y and z axes of $\mathbf{F}_{eye}{}^w$, respectively. Following our convention for $\mathbf{F}_{eye}{}^w$, the translation vector $\mathbf{t}_{e\_w}$ between these frames is simply $\mathbf{m1}^w$.

2.5. Calibration

The reconstruction process described above requires knowledge of the coordinates of the two camera projection centers, $\mathbf{o}_l$ and $\mathbf{o}_r$, and the coordinates of the image points $\mathbf{m3}_l$ and $\mathbf{m3}_r$, all given in the world coordinate system $\mathbf{F}_{world}$ (Fig. 2). Since all these points are measured (or known) in their respective camera frames (e.g., for the left camera, $\mathbf{m3}_l{}^1$ and $\mathbf{o}_l{}^1$ are known), we must first transform them to the world coordinate system. The parameters of these transformations – the camera's *intrinsic* and *extrinsic* parameters (Tsai, 1987) – are determined via a calibration process that determines the relationship between a set of known reference points and their measured image projections. We accomplished this calibration by using a calibration image of a checkerboard panel attached to the outer surface of the water tank as shown in Fig. 2. For purposes of convenience, one of the corners of this panel is defined as the origin of $\mathbf{F}_{world}$, and since the size of the checkerboard squares is known, the coordinates of all other grid points are known and can be used for calibration. This calibration process is based on a direct linear transformation (Abdel-Aziz and Karara, 1971; Tsai, 1987), the technical details of which are presented in Appendix A. The computation itself was performed using a public domain Matlab toolbox (Bouguet, 2008; Heikkila and Silven, 1997; Zhang, 1999).

2.6. Detection of the fish eye reference points

The three reference points on the fish eye are extracted from three circular markers that are attached to the fish eye with superglue. These circular markers are projected as ellipses on the camera's image plane. Since projective geometry dictates that the center of an ellipse is invariant, these points can be used as reference points for triangulation. Below we explain how the projections of these points are detected in the image planes.
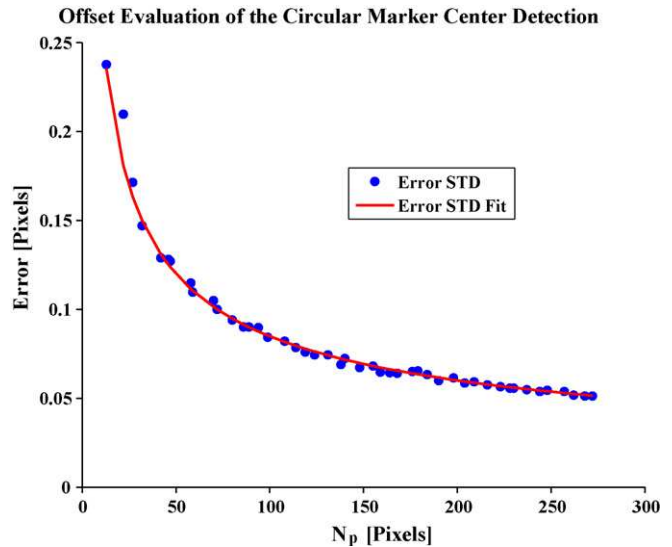

Fig. 3. Evaluation of the error in detecting a circle center through our tracking system.

In this figure we set to determine the connection between the number of pixels, which represent each circle and the accuracy in detecting these circles center. For analysis we generated 46 circles, with random radii ranging from 3.0 to 48.0 pixels (including fraction of pixels). The center point of each circle was randomly jittered inside the center pixel. These circles centers coordinates were then measured from the circles images using our program and were compared to the known circles centers coordinates. For each circle, the standard deviation of the error between its known and measured coordinates is shown. Since our program uses the circle edge for its center detection, the standard deviation is shown as a function of the number of pixels (N_p) that were used by the program for the center detection. The fit to the standard deviation of the error is $\sim 0.85 N_p^{(-0.5)}$, and in our measurements we used $N_p > 100$, which yields a standard deviation of the error of <0.085 pixels.

Each marker comprises a black ring (0.1 mm thick, 2 mm inner diameter) glued on top of a white plastic disk (0.2 mm thick, 2.5 mm in diameter). The center of the image of this marker is then detected by fitting an ellipse to its edge map (Halif and Flusser, 2000). Therefore, the detection accuracy of the projected marker center depends on the length of the marker's elliptic edge. In our experiments, the edge of each marker is represented by ≥ 100 pixels, which facilitates a sub-pixel accuracy in the estimation of its center. To quantify this dependence, we simulated the relation between the error in detecting a circle center and the circumference length of that circle. We generated circles with random radii, and placed each of their center points in random locations inside the center pixel. Using a Matlab program (see following paragraph), we then estimated the coordinates of the centers of these circles and compared them to the ground truth. Fig. 3A shows that we can expect the standard deviation in the error in detecting the circle center to be less than 0.085 pixels.

Finding and tracking the marker centers in the two image frames is performed automatically by a Matlab program. One of the markers is chosen in the first frame, and a bounding box that contains only the marker image is defined as a mask. The masked image edge is then fitted to an ellipse whose center is determined by the intersection of its two main axes. Around this ellipse center, a new bounding box whose side length is 50% larger than the major axis of the ellipse is used on the following image frame for fitting an ellipse for the next time step. This tracking-like process is possible due to the small changes in marker centers between two consecutive images, and it is repeated until the marker centers in pixel coordinates are determined for all frames.

3. Results

3.1. System accuracy

To evaluate the accuracy of our video-based eye tracking system, we compared its performance with that of a search coil system (Coil System CS681, Primlec, Regensdorf, Switzerland) attached to a phantom (dummy) that provides a controlled platform for 3D rotations and translations. We attached to the phantom a set of two perpendicular search coils and, next to it, a set of the three visual markers. The phantom was immersed in water and manually rotated and translated so as to mimic fish eye movements. The phantom pose was evaluated using the two approaches. Since the precision of the search coil system is 0.005°, its results were considered "ground truth" for the 3D pose of the phantom in space.

The comparison of the search coil and video measurements was made on the premise that both the set of coils and the three marker centers, separately, define a plane. Each of these planes can be represented by its unit normal vector. Since the two planes need not necessarily be identical, the angle between the two normal vectors is constant and should ideally remain so. Thus, the fluctuations of this angle around its mean represent the measurement accuracy up to the precision of the search coil system.

Fig. 4 presents the results of this analysis on a 3 s sequence of phantom rotation measurements. Fig. 4A and B present the elevation and horizontal rotation angles of the two normal vectors. Examining our measurement error (Fig. 4C), we found that the devi-
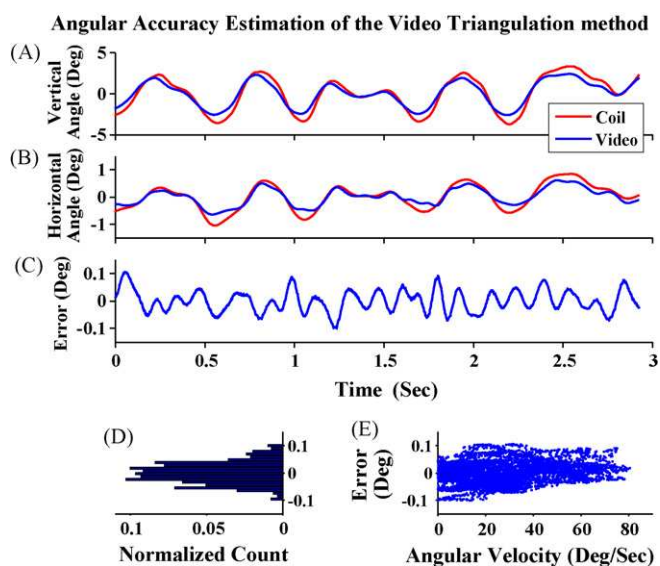
ation between the two measurement systems does indeed remain low up to a standard deviation of 0.047°. In addition, the histogram of angular errors (Fig. 4D) reveals that the maximal absolute value of the error is less than 0.1°.

Finally, since our video system acquires single frames in a finite time (approximately 8 ms for a frame), it was necessary to verify that pixel smearing due to eye movement does not dramatically influence the accuracy of the measurement. Fig. 4E shows the video system angular error as a function of the phantom's angular velocity, which was calculated from the search coil system (by dividing every two consecutive angular measurements by the system's sampling time). It was thus shown that the error is hardly affected at all by the angular velocity for most velocities that characterize the fish's eye movements (in the range 0–80° s^{-1}).

3.2. Simple distance-based accuracy evaluation

The accuracy of our system depends on many parameters that may differ slightly in each experiment (e.g., camera configurations, field of view, etc.). Thus, it is useful to have an additional method to obtain a rough and fast evaluation of the accuracy without using the search coil so as to ensure the functionality of the systems prior to each behavioral experiment. Here, we present such a method for a coarse accuracy evaluation that does not require a search coil.

For this alternative evaluation approach, we assume that the three markers are attached to a rigid body and therefore their spacing should remain constant. This assumption is reasonable due to the stiffness of the sclera and the intra-ocular pressure that maintains the mechanical stability of the eye. To validate our measurement accuracy, we measured the variance in the estimated spacing of the markers, which under the assumption of a rigid body, represents the noise in the measurement. Fig. 5 shows that the standard deviation of the estimated distance between the markers is about 13 μm. Taking the ratio of this deviation and the distance between the markers (approximately 1 cm) indicates the angular accuracy of our system is approximately 0.070°. Note that this value is comparable with the more accurate angular error evaluation of 0.047° that we obtained through the search coil comparison.

3.3. Measuring eye movements of freely moving fish

As a first application of our novel technique, we examined the fish eye movement during shooting behavior to reveal the relation between body posture and gaze direction. Previous anatomical observations (Timmermans and Souren, 2004) have indicated that to tune the squirt direction while shooting, the archer fish can rotate only its entire body. Therefore, prior to changing its body direction, the fish needs to perform a computation to compensate for the image refraction and adjust its body orientation in space accordingly.

During target acquisition, the fish can adopt one of several possible strategies of body and eye movement in space. For example, the fish can swim with its gaze orientation fixed in space (on target) while it continuously adjusts the shooting direction by rotating its body around its eyes; it can swim along the shooting direction with a fixed body direction and rotate its body at the water level; or it can swim directly upwards while its eyes perform smooth pursuit of the target. In the last case, the target refraction angle changes continuously.

We found that during shooting at a prey the fish mainly adopts the first strategy, i.e., fixating its gaze and rotating its body around its eyes. This hunting behavior was investigated in detail in two similar behavioral experiments (Fig. 6).

In these experiments the fish swam along the water tank glass wall (see Fig. 1) towards a target (a small piece of salmon) placed in front of it on the glass wall, 15 cm above water level. Approximately
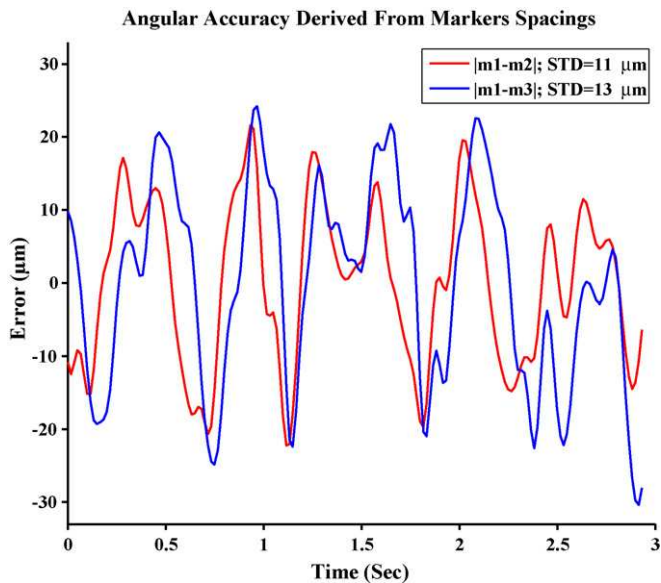

Fig. 4. Video system accuracy evaluation through the use of a search coil system as a reference.
Both our novel video-based eye tracking system and a search coil system simultaneously measured rotation angles of two planes lying on a phantom that was immersed in water and rotated by hand. One plane was comprised of two coils and the other one of the three dots. **A**. A 3 s time trace example of the vertical rotation angles of the coils and markers planes normal, as measured by the two systems. **B**. The same as in **A**, but for the horizontal rotation angles. **C**. The difference in measurement between the video and search coil systems. Since the accuracy of the search coil is superior to that of the video method, we can regard the difference in measurements as the error in the video measurement. **D**. The histogram of the error angles in panel **B** reveals that the standard deviation of the error (∼0.047°) is less than the diameter of a single photoreceptor on the retina (∼0.1°). **E**. Each dot represents one measurement of the error in angles as a function of angular velocity. The dependence of the error on the angular velocity shows that in the range of angular velocities that characterize the archer fish eye (typical angular velocities during saccades are in the range of 50–100° s^{-1} (Segev et al., 2007)) the error does not increase. One should expect, however, errors to increase due to pixel smearing during measurements of faster eye movements, a problem that can be handled by using cameras with faster frame rates.

Fig. 5. An independent angular accuracy measurement.
From the spacings of the markers and from their fluctuations, we derived an independent accuracy measurement. We used the same data that were used for comparing the accuracy of the two systems (Fig. 4B). **A**. The red and blue lines represent the deviation from the mean of the spacings between markers **m1** and **m2**, and markers **m1** and **m3**, respectively. The spacings between the markers are expected to remain constant during the experiment since the markers are attached to a rigid body. Therefore, the standard deviation of the fluctuations in distance between these spacings is a rough indication of the system's translational accuracy. By dividing the standard deviations of the red and blue lines by their spacing distances, which were approximately 1 cm, we obtained a rough estimation of angular accuracy of 0.06° and 0.07°, respectively. This angular accuracy estimation is twice in size than the angular accuracy estimation we received by comparing the video-based and the search coil systems (Fig. 4C). These deviations might be the result of correlations in the fluctuations of the markers spacings, as expected from markers which are attached to the same rigid body.

10 cm from the glass wall the fish's snout emerged out of the water, the fish shot down the target, and then fed on it.

Fig. 6A presents a sequence of snapshots showing shooting behavior in one of the experiments. An examination of the eye rotation angles (Fig. 6C and D) and the fish body postures (Fig. 6E) during the experiments shows that target acquisition starts when the fish makes a saccade towards the target (Figs. 6C and D) and fixates on it. Thereafter, the fish surfs upwards with its eyes fixed on the target, while its body rotates around the eyes (Fig. 6E) until its snout reaches the water level. Note, that the fish body rotations were measured qualitatively through examining the differences in the fish's body outlines. These differences are sufficient to clearly show the fish's body rotations, though quantitative measurement of this behavior could be obtain by duplicating our system with body markers also.

From examining the fixed orientation of the fish's eye in space (Fig. 6C and D), we conclude that during surfacing the fish eye moves in a straight trajectory along the refracted image of the target. We observed that the water depth range during this straight trajectory varied in the different experiments between practically zero, when the fish shot from close to the water level, to more than 10 cm. We conjecture that the fish chooses this trajectory since it is the optimal path in terms of accuracy and time between the saccade and the squirt. This complex behavior of the fish adjusting its body direction while surfacing with its eyes in the target direction implies that there is only one degree of freedom for the squirt direction, when the snout of the fish reaches the water level.

In these behavioral experiments, the distances between the centers of the markers were 6 and 5 mm, and the standard devi-

ation in their estimation was 5 and 3 μm, respectively. Applying our simple accuracy evaluation approach to these data (see section *Simple distance-based accuracy evaluation*), we expected angle errors of 0.047° and 0.034°, respectively. In these experiments, such angle errors bound the accuracy of the three angles of rotation at 0.047°.

4. Discussion

Here we present a novel 3D eye tracking method and system for the archer fish, which is based on multi-refractive stereo geometry for reconstruction of three reference points on the fish eye. These points are used to build a reference frame on the fish eye, from which the fish's rotational and translational eye movements are extracted. The system's accuracy was evaluated using a search coil as the reference. We found a standard deviation in the angular accuracy of 0.047° and a maximal absolute error of 0.1° for our video-based measurements. This accuracy enables the recoding of eye movement with a resolution that is comparable to the movement of the diameter of a single photoreceptor on the retina. We demonstrated the system's ability to measure eye movements of a freely moving archer fish in the process of acquiring and shooting a target. By adding an additional three markers on the fish body, this system can be extended to independently measure the fish eye and body movements. Our method, which facilitates the study of the extreme visual behavior of the archer fish, can also be applied for other small and freely moving animals.

Our system has several advantages over other eye tracking systems. First, due to the use of multiple views geometry, it can detect translational eye movements. Second, the use of markers instead of wired coils is less stressful to the fish. Although the attachment of the three markers initially induces some stress to the fish, which results in a tendency to hide in the water tank, after a few hours the fish seems to resume to its normal behavior and shoots at targets normally. In contrast, with the search coil system, the fish is continuously under high stress, and friction of the coils in the water may significantly alter the fish eye movements. Indeed, as we have observed, the shooting accuracy decreases when a coil is attached to the eye.

To evaluate the factors determining the accuracy of the method, we must examine both the temporal and the spatial accuracy. With regard to the former, it was previously shown that the frequencies of fixational eye movements are <15 Hz (Segev et al., 2007). Our system is thus suitable for measuring these eye movements in this frequency band, but some saccade frequencies might be higher, and their detection would thus require cameras with a higher sampling rate.

There are two main factors that determine the spatial accuracy of our method: the size of the markers in each camera's pixel space (the more pixels, the higher the accuracy) and the distance between the markers (the greater the distance, the higher the accuracy). While the physical size of each marker and the distance between the markers is dictated by the eye size, the pixel space size of each marker is determined by the camera (the higher the resolution of the camera, the better the accuracy). We therefore expect the accuracy to scale linearly with the camera resolution. For example, from Fig. 3 it can be seen that we obtained a standard deviation of the pixel error of <0.085 pixels for our cameras. Therefore, taking a field of view of 6.4 cm × 4.8 cm, we obtain each pixel side length distance to be 100 μm, with an error of <8.5 μm. Thus, we expect an angular error of <0.05° for a distance between markers of 1 cm. For cameras with double the number of pixels in each axis, we expect the spatial accuracy to improve by a factor of 2, and a corresponding improvement in the system's angular accuracy into the sub-photoreceptor resolution.
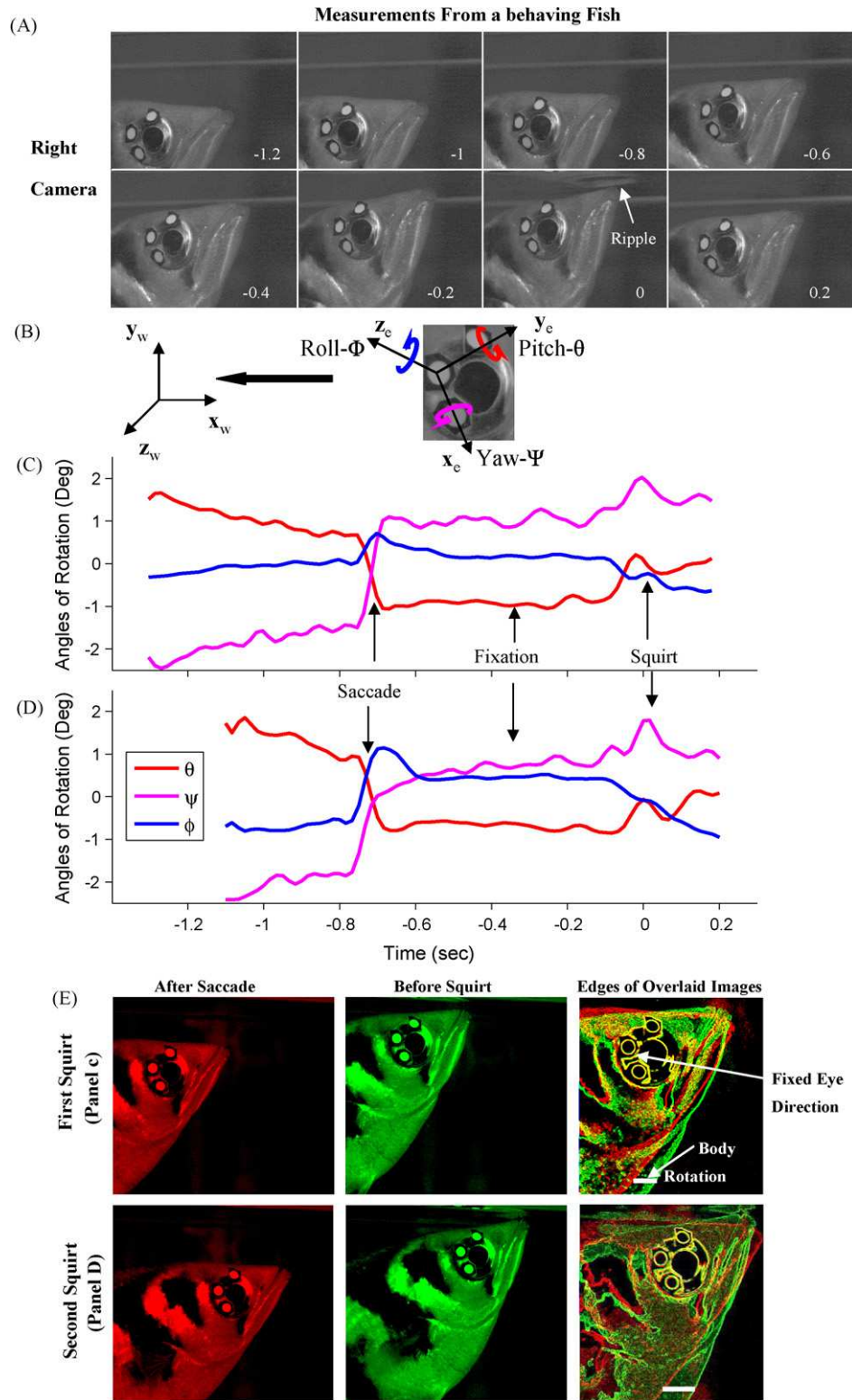
Fig. 6. Measurements of archer fish eye movement while locating and shooting at a target. **A**. Eight images taken by the right camera, each with the elapsed time in seconds indicated. Time zero is taken as the moment of shot. **B**. Fish eye reference frame rotation angles ψ, θ, and ϕ (see *Rotation angles and translation vector section*) that are used to describe the relation between this reference frame and the water tank reference frame. **C and D**. Two different shooting experiments in which the three angles of rotation were measured by the video method. During each experiment the fish shot down, by squirting a jet of water, a piece of salmon attached to one of the water tank glass walls, 15 cm above water level, and then fed on it. Saccades and squirts are indicated. **E**. During target acquisition, the fish performs a saccade towards the target and fixes its eyes on the target until the squirt, as indicated by the plateau in eye movements in panels **C** and **D**. By comparing images from the two different shots in **C** and **D**, we found that during target acquisition, the fish also rotates its body to the correct shooting posture (right panels). This may be seen by examining the video frames from the two shots after the saccade and before the squirt. By overlaying the two images while aligning the two eyes on the right panels and keeping the same color scheme of red and green, we found the regions of the fish that remain at the same orientation; these are colored yellow. The red or green colors indicate body and eye regions that rotate in space. The yellow markers on the fish eye indicate that while the fish surfaces it maintains an almost constant gaze direction in space while fixating on the target. The angle difference in the fish posture (indicated with an arrow) indicates that during the targeting process the body rotates around the eye so that the fish emerges with its snout in an accurate firing position.

Appendix A.

Here we derive the extrinsic and intrinsic parameters of the left camera (Abdel-Aziz and Karara, 1971; Heikkila and Silven, 1997). The procedure for the derivation of these parameters for the right camera is the same. We use the rotation matrix $\mathbf{R}$ to denote the camera's extrinsic parameters and $\mathbf{t}$ to denote the translation vectors that are used for transformation between $\mathbf{F}_{world}^w$ and the $\mathbf{F}_{l\_image}^1$ (Fig. 2). For n points on the checkerboard, denoted as $\mathbf{X}(1,\ldots,n)^w$, the transformation of these points from the water tank to the camera pinhole model coordinate systems is given by:

$$\mathbf{X}(1,\ldots,n)^1 = \mathbf{R} \cdot \mathbf{X}(1,\ldots,n)^1 + \mathbf{t} \tag{4}$$

where $\mathbf{X}(1,\ldots,n)^1$ denotes the n points in the camera model coordinate system. After the transformation to the camera model coordinate system, we can perform the transformation to the two-dimensional camera pixel coordinate system.

We define a point i on the checkerboard, in the camera model coordinate system, denoted as $\mathbf{X}(i)^1 = (x^1, y^1, z^1)$. Then, we normalize i by z^1:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^1/z^1 \\ y^1/z^1 \end{pmatrix} \tag{5}$$

where x and y are the normalized camera model coordinates. Now, according to the pinhole camera model, by using the normalized coordinates, we can transfer x and y to the pixel space:

$$\begin{pmatrix} x_P \\ y_p \end{pmatrix} = \begin{pmatrix} f_x x + CC_x \\ f_y y + CC_y \end{pmatrix} \tag{6}$$

where f_x and f_y are the focal lengths of the camera in pixels along the camera x and y axes, respectively, and CC_x and CC_y represent the center of the camera in pixels along the camera x and y axes, respectively. The relationship between the pixel coordinates and the normalized camera coordinates is given by:

$$\begin{pmatrix} x_P \\ y_p \\ 1 \end{pmatrix} = \mathbf{KK} \begin{pmatrix} x \\ y \\ 1 \\ 1 \end{pmatrix} \tag{7}$$

where:

$$\mathbf{KK} = \begin{pmatrix} f_x & 0 & CC_x & 0 \\ 0 & f_y & CC_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{8}$$

$\mathbf{KK}$ is thus known as the camera's intrinsic parameters matrix.

Now, we can write the complete transformation of point i, denoted as $\mathbf{X}(i)^w = (x^w, y^w, z^w)$, from the checkerboard coordinate system to its camera pixel space point (x_p, y_p):

$$\begin{pmatrix} x_P \\ y_p \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & CC_x & 0 \\ 0 & f_y & CC_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \frac{1}{z^1} \begin{pmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_x \\ r_{2,1} & r_{2,2} & r_{2,3} & t_y \\ r_{3,1} & r_{3,2} & r_{3,3} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^w \\ y^w \\ z^w \\ 1 \end{pmatrix} \tag{9}$$

where $r_{i,j}$ denotes the ith row and the jth column element of the rotation matrix, $\mathbf{t}_i$ is the ith element of the translation vector, and

$$z^1 = r_{3,1}x^w + r_{3,2}y^w + r_{3,3}z^w + \mathbf{t}_z \tag{10}$$

The first matrix in the right-hand side of Eq. (9) represents the camera's intrinsic parameters, and the second matrix represents the camera's extrinsic parameters. A solution for these parameters requires at least 10 grid corner points: 6 points for the extrinsic parameters of rotation and translation and 4 points for the intrinsic parameters. (We should keep in mind that by using more grid points we should expect higher accuracy of the calibration procedure.) To solve Eq. (9), we multiply both sides by z^1 and then multiply the intrinsic parameters matrix by the extrinsic parameters matrix and define a matrix $\mathbf{A1}$:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3}a & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \end{pmatrix}$$

$$= \begin{pmatrix} f_x & 0 & CC_x & 0 \\ 0 & f_y & CC_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} r_{1,1} & r_{1,2} & r_{1,3} & \mathbf{t}_x \\ r_{2,1} & r_{2,2} & r_{2,3} & \mathbf{t}_Y \\ r_{3,1} & r_{3,2} & r_{3,3} & \mathbf{t}_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{11}$$

Now Eq. (8) can be rewritten as:

$$\begin{pmatrix} x_P z^1 \\ y_p z^1 \\ z^1 \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3}a & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \end{pmatrix} \begin{pmatrix} x^w \\ y^w \\ z^w \\ 1 \end{pmatrix} \tag{12}$$

We can rearrange Eq. (11) to yield:

$$\begin{pmatrix} x^w & y^w & z^w & 1 & 0 & 0 & 0 & 0 & -x^w x_p & -y^w x_p & -z^w x_p & -x_p \\ 0 & 0 & 0 & 0 & x^w & y^w & z^w & 1 & -x^w y_p & -y^w y_p & -z^w y_p & -y_p \end{pmatrix} \mathbf{A2} = 0 \tag{13}$$

where $\mathbf{A2} = (a_{1,1}\ a_{1,2}\ a_{1,3}\ a_{1,4}\ a_{2,1}\ a_{2,2}\ a_{2,3}\ a_{2,4}\ a_{3,1}\ a_{3,2}\ a_{3,3}\ a_{3,4})^T$.

Since $\mathbf{A2}$ is over determined and is known up to a scaling factor, we can set $\mathbf{A1}_{3,4} = 1$, and similarly to Abdel-Aziz and Karara (1971) and Heikkila and Silven (1997), we obtain the following equation:

$$\begin{pmatrix} x_1^w & y_1^w & z_1^w & 1 & 0 & 0 & 0 & 0 & -x_1^w x_{p,1} & -y_1^w x_{p,1} & -z_1^w x_{p,1} \\ 0 & 0 & 0 & 0 & x_1^w & y_1^w & z_1^w & 1 & -x_1^w y_{p,1} & -y_1^w y_{p,1} & -z_1^w y_{p,1} \\ & & & & & \cdot & & & \\ & & & & & \cdot & & & \\ & & & & & \cdot & & & \\ x_N^w & y_N^w & z_N^w & 1 & 0 & 0 & 0 & 0 & -x_N^w x_{p,N} & -y_N^w x_{p,N} & -z_N^w x_{p,N} \\ 0 & 0 & 0 & 0 & x_N^w & y_N^w & z_N^w & 1 & -x_N^w y_{p,N} & -y_N^w y_{p,N} & -z_N^w y_{p,N} \end{pmatrix}$$

$$\times \begin{pmatrix} a_{1,1} \\ a_{1,2} \\ a_{1,3} \\ a_{1,4} \\ a_{2,1} \\ a_{2,2} \\ a_{2,3} \\ a_{2,4} \\ a_{3,1} \\ a_{3,2} \\ a_{3,3} \end{pmatrix} = \begin{pmatrix} x_{p,1} \\ y_{p,1} \\ \cdot \\ \cdot \\ \cdot \\ x_{p,N} \\ y_{p,N} \end{pmatrix} \tag{14}$$

and the variables of vector $\mathbf{A2}$ can be solved through the inverse matrix technique.

The matrix of the intrinsic and extrinsic parameters $\mathbf{A1}$ can also be written as:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \end{pmatrix} = \begin{pmatrix} f_x r_1 + CC_x r_3 & f_x \mathbf{t}_x + CC_x \mathbf{t}_z \\ f_y r_2 + CC_y r_3 & f_y \mathbf{t}_y + CC_y \mathbf{t}_z \\ r_3 & \mathbf{t}_z \end{pmatrix} \tag{15}$$

where r_i denotes the ith row of the rotation matrix. Note that the rotations in the rotation matrix are diagonal, therefore, $r_i r_i = 1$ and

$r_i r_j = 0$. The intrinsic and extrinsic parameters can now be represented in terms of the variables of vector **A2** as follows:

$$r_3 = (a_{3,1}, a_{3,2}, a_{33})$$
$$CC_x = (a_{3,1}, a_{3,2}, a_{33})(a_{1,1}, a_{1,2}, a_{1,3})^T$$
$$CC_y = (a_{3,1}, a_{3,2}, a_{33})(a_{2,1}, a_{2,2}, a_{2,3})^T$$
$$f_x = (a_{1,1}, a_{1,2}, a_{13})(a_{1,1}, a_{1,2}, a_{1,3})^T - CC_x$$
$$f_y = (a_{2,1}, a_{2,2}, a_{23})(a_{2,1}, a_{2,2}, a_{2,3})^T - CC_y$$
$$r_1 = \frac{((a_{1,1}, a_{1,2}, a_{13}) - CC_x r_3)}{f_x}$$
$$r_2 = \frac{((a_{2,1}, a_{2,2}, a_{23}) - CC_y r_3)}{f_y} \tag{16}$$
$$t_z = a_{3,4}$$
$$t_x = \frac{(a_{1,4} - CC_x t_z)}{f_x}$$
$$t_y = \frac{(a_{2,4} - CC_y t_z)}{f_y}$$

References

Abdel-Aziz YI, Karara HM. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. In: Proc. ASP Symposium on Close-Range Photogrammetry; 1971. p. 1–18.

Barta A, Horvath G. Underwater binocular imaging of aerial objects versus the position of eyes relative to the flat water surface. J Opt Soc Am A Opt Image Sci Vis 2003;20:2370–7.

Bouguet JY. Camera calibration tutorial and Matlab code available at: http://www.vision.caltech.edu/bouguetj/calib_doc/, 2008.

Dill LM. Refraction and spitting behavior of archerfish (*Toxotes chatareus*). Behav Ecol Sociobiol 1977;2:169–84.

Ewert JP. The neural basis of visually guided behavior. Sci Am 1974;230:34–42.

Ewert JP. Neural correlates of key stimulus and releasing mechanism: a case study and two concepts. Trends Neurosci 1997;20:332–9.

Gonzalez F, Perez R. Neural mechanisms underlying stereoscopic vision. Prog Neurobiol 1998;55:191–224.

Halif R, Flusser J. Numerically stable direct least squares fitting of ellipses. Czech Republic: Department of Software Engineering, Charles University; 2000.

Hartley R, Zisserman A. Multiple view geometry in computer vision. Cambridge: Cambridge University Press; 2004.

Heikkila J, Silven O. A four-step camera calibration procedure with implicit image correction. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Proceedings; 1997. p. 1106–12.

Luling K. Morphologisch-anatomische und histologische untersuchungen am auge des schutzenfisches *Toxotes jaculatrix* (Pallas, 1766) nebst bemerkungen zum spuckgehahaben. Zeitschrift für Morphologie und Ökologie der Tiere 1958;47:529–610.

Luling K. Archer Fish Sci Am 1963;209:100–29.

Martinez-Conde S, Macknik S, Hubel D. The role of fixational eye movements in visual perception. Nat Rev Neurosci 2004;5:229–40.

Migliaccio AA, Macdougall HG, Minor LB, Della Santina CC. Inexpensive system for real-time 3-dimensional video-oculography using a fluorescent marker array. J Neurosci Methods 2005;14:141–50.

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

Plotkin A, Paperno E, Vasserman G, Segev R. Magnetic tracking of eye motion in small, fast-moving animals. IEEE Trans Magnetics 2008;44:4492–5.

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.

Rucci M, Iovin R, Poletti M, Santini F. Miniature eye movements enhance fine spatial detail. Nature 2007;447:851–4.

Schuster S. Archerfish Curr Biol 2007;17:R494–5.

Schuster S, Rossel S, Schmidtmann A, Jager I, Poralla J. Archer fish learn to compensate for complex optical distortions to determine the absolute size of their aerial prey. Curr Biol 2004;14:1565–8.

Schuster S, Wohl S, Griebsch M, Klostermeier I. Animal cognition: how archer fish learn to down rapidly moving targets. Curr Biol 2006;16:378–83.

Segev R, Schneidman E, Goodhouse J, Berry 2nd MJ. Role of eye movements in the retinal code for a size discrimination task. J Neurophysiol 2007;98:1380–9.

Shepherd SV, Platt ML. Noninvasive telemetric gaze tracking in freely moving socially housed prosimian primates. Methods 2006;38:185–94.

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.

Temple SE. Effect of salinity on the refractive index of water considerations for archer fish aerial vision. J Fish Biol 2007;70:1626–9.

Timmermans P. Prey catching in the archer fish: marksmanship, and endurance of squirting at an aerial target. Neth J Zool 2000;50:411–23.

Timmermans P. Prey catching in the archer fish: angles and probability of hitting an aerial target. Behav Process 2001;55:93–105.

Timmermans P, Vossen J. Prey catching in the archer fish: does the fish use a learned correction for refraction? Behav. Process 2000;52:21–34.

Timmermans P, Souren PM. Prey catching in archer fish: the role of posture and morphology in aiming behavior. Physiol Behav 2004;81:101–10.

Treibitz T, Schechner YY, Singh H. Flat refractive geometry. In: IEEE Conference on Computer Vision and Pattern Recognition; 2008. p. 3903–10.

Tsai RY. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. IEEE J Robotic Autom 1987;3:323–44.

Yekutieli Y, Mitelman R, Hochner B, Flash T. Analyzing octopus movements using three-dimensional reconstruction. J Neurophysiol 2007;98:1775–90.

Zhang Z. Flexible camera calibration by viewing a plane from unknown orientation. In: IEEE International Conference on Computer Vision; 1999. p. 666–73.