ELSEVIER

Journal of Neuroscience Methods 95 (2000) 133–143

JOURNAL OF NEUROSCIENCE METHODS

www.elsevier.com/locate/jneumeth

Body modeling and model-based tracking for neuroethology

Malcolm A. MacIver [a,c], Mark E. Nelson [a,b,c,*]

[a] *The Neuroscience Program, University of Illinois, Urbana, IL 61801, USA*
[b] *The Department of Molecular and Integrative Physiology, University of Illinois, Urbana, IL 61801, USA*
[c] *The Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana, IL 61801, USA*

Received 23 June 1999; received in revised form 24 September 1999; accepted 10 October 1999

Abstract

The accurate tracking of an animal's movements and postures through time has broad applicability to questions in neuroethology and animal behavior. In this paper we describe methods for precision body modeling and model-based tracking of non-rigid animal movements without the use of external markers. We describe the process of obtaining high-fidelity urethane casts of a model organism, the weakly electric knifefish *Apteronotus albifrons*, and the use of a stylus-type 3-D digitizer to create a polygonal model of the animal from the cast. We describe the principles behind markerless model-based tracking software that allows the user to translate, rotate, and deform the polygon model to fit it to digitized video images of the animal. As an illustration of these methods, we discuss how we have used model-based tracking in the study of prey capture in nocturnal weakly electric fish to estimate sensory input during behavior. These methods may be useful for bridging between the analytical approaches of quantitative neurobiology and the synthetic approaches of integrative computer simulations and the building of biomimetic robots.

Keywords: Animal tracking; Motion capture; Casting; Moldmaking; Infrared; Camera calibration; Video digitizing; MicroScribe; Electroreception; Computational neuroethology

1. Introduction

Accurate tracking of a 3-D object from a sequence of time-varying images or sensor readings is an active topic of research in a variety of application areas. The applications are diverse, spanning animal behavior, biomechanics, real-time character animation, gesture-driven user interfaces, sign language translation, surveillance systems, and 3-D interfaces for virtual reality systems. Many of these applications employ marker-based approaches to object tracking. Marker-based approaches rely on the sensing of discrete, spatially localized points or markers on the surface of the object, such as natural body landmarks, attached reflectors, or light-emitting diodes (Spruijt et al., 1992; Winberg et al., 1993; Hughes and Kelly, 1996; Kruk, 1997; Vatine et al., 1998). In contrast, model-based approaches rely on globally fitting a surface model of the object to image or sensor data (Mochimaru and Yamazaki, 1994; Gavrila and Davis, 1996; Jung and Wohn, 1997; Tillett et al., 1997).

The model-based approach to animal tracking has not received wide application in animal behavior and neuroethological studies. However, it can provide high-resolution data on the time-varying conformation of the entire animal, and may be the best choices in situations where marker-based systems are impractical or inadequate. In our research on the electrosensory system of weakly electric fish, we use model-based tracking to accurately determine the conformation of the fish's body during prey capture behavior. Model-based tracking allows us to reconstruct electrosensory activation across the receptor array, which provides valuable insights into the neural control of sensory acquisition.

In this paper, we detail the methodology used for model-based tracking of black ghost knifefish, and discuss general considerations that may be relevant to other applications. First, we describe high-precision casting techniques and methods for the creation of a polygonal surface model based on the cast. Then, we describe how this model is used for tracking fish using

* Corresponding author.

0165-0270/00/$ - see front matter
PII: S0165-0270(99)00161-2

a two-camera infrared video system. Finally, we discuss how we link behavioral data from model-based tracking to sensory neurophysiology in our studies.

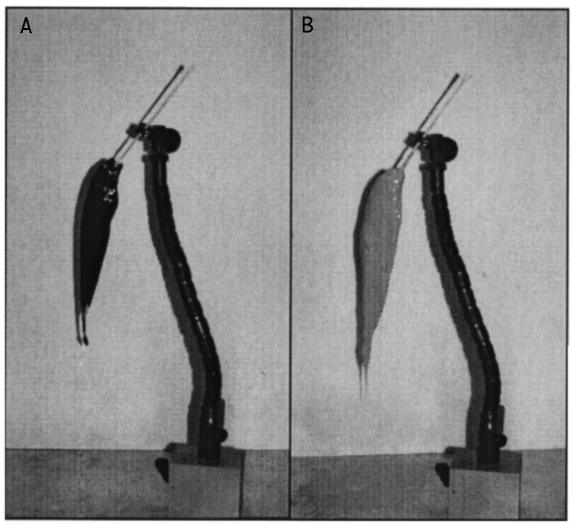

Fig. 1. Making an RTV silicone mold of a weakly electric fish. (A) The support rod is positioned so that the dorsal curvature of the fish approximates the natural posture in water. (B) During casting silicone is slowly poured on the fish until it covers the entire surface. Several layers of casting compound are added, with enough time between layers for the silicone to partially cure.

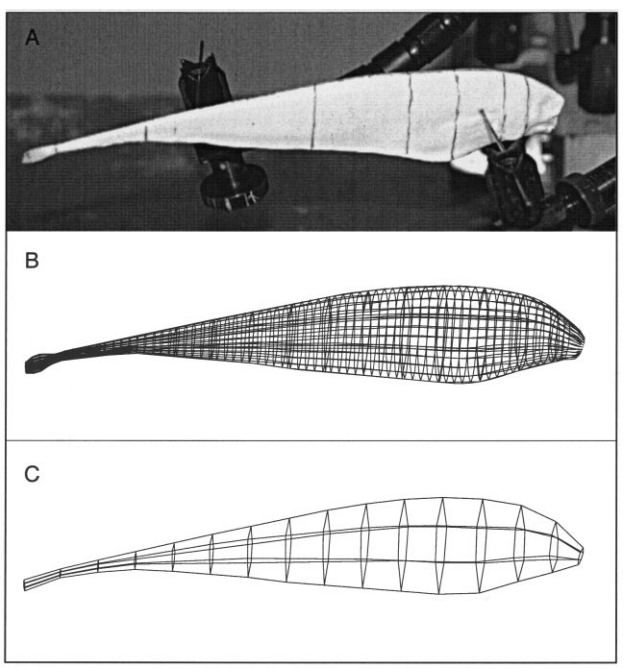

Fig. 2. (A) Illustration of the mounting of the cast for digitizing. (B) High resolution polygon surface model of *A. albifrons*, 1540 faces total, 70 longitudinal and 22 around. (C) Low resolution model, 90 faces total, 15 longitudinal and 6 around.

2. Body modeling

Model-based tracking of an animal requires an accurate quantitative representation of its surface morphology. In this section, we describe procedures for making a physical cast of the animal and creating a 3-D model from the cast. Casting objects is a well developed technical craft (Boardman, 1950; Parsons, 1973; Gardner, 1974; Waters, 1983; James, 1989). Below we describe general casting principles, as well as specific details for casting a black ghost knifefish (*Apteronotus albifrons*). The specimen shown in Figs. 1 and 2 was 190 mm long and weighed 30 g.

2.1. Preparing the specimen

Preparation for casting begins by obtaining a fresh, clean specimen. In our case, an adult *A. albifrons* was euthanized with an overdose of tricaine methanesulfonate (MS-222, Sigma, St. Louis, MO, USA). The surface of the fish was cleaned with a mild detergent and a soft brush to remove mucus. If accurate casts of the fins are desired they may be fixed with formalin prior to making the mold (McHenry et al., 1995). For our electrosensory research, the sensors of interest are not present on the fins so we were less concerned with this detail.

2.2. Posing the specimen

A typical posture of the behaving animal should be selected as the canonical posture in which it is to be cast. Prior to posing the animal for casting, we posed a recently euthanized fish by floating it on its side, directly above a reference grid in water just covering its surface. This allowed us to reproduce the natural dorsal–ventral curvature of the fish's spine. Reference photographs were then taken for correction of distortions created during cast creation (Section 2.6).

After taking a set of reference photographs, the animal is posed for creation of the mold. We approximated the natural posture of the knifefish in water by suspending it in mid-air at an appropriate angle to reproduce the natural dorsal–ventral curvature (Fig. 1A). To suspend the animal, a section of 3 mm (diameter) wooden dowel was placed into the mouth and 3 cm into the gut. A rapid curing urethane (TC 806 A/B, BJB Enterprises Inc, Tustin, CA, USA) was injected into the oral cavity to hold the support rod in place.

2.3. Selection of moldmaking and casting compounds

Two-component room temperature vulcanizing (RTV) silicone elastomer is often an excellent choice for the moldmaking material because it provides high reproduction accuracy, long mold shelf life, does not

normally require the use of mold release agents, and is compatible with a large variety of casting compounds and pouring temperatures. Tests have shown that silicone elastomer can capture surface features of 0.1–0.3 μm reliably (Bromage, 1985). There are many commercial silicone elastomer varieties and additives, giving different setting times, demolding times, pouring viscosities, cured hardnesses and elasticities. We used Rhodorsil V-1065 with Hi-Pro Blue catalyst (Rhodia Silicones VSI, Troy, NY, USA) for the flexibility, high tear strength, low shrinkage, and long shelf life of the resultant mold.

There are a larger number of potentially useful casting agents that can be poured into the finished mold to create the cast. Urethanes, silicone elastomers, and molten polyvinyl chloride (PVC) gels can be used for generating rigid and flexible casts. The stiffness of the cast can be controlled through the use of diluents and additives. Flexible PVC casts have been used for biomechanical studies on the role of body stiffness in fish swimming (McHenry et al., 1995). Because we used a contact 3-D digitizer (Section 2.6), our application required a rigid cast. We selected a particular rigid urethane casting material (TC 806 A/B, BJB Enterprises, Tustin, CA, USA) for its low uncured viscosity, which allowed it to seep into the thin sections of the mold.

2.4. Design and construction of the mold

Molds can be made in one-, two-, or multi-part configurations. When the topology of the animal permits it, a one-part mold can be constructed by simply coating the animal with several layers of the mold compound. In general, a one-part mold can be used whenever the object does not contain significant undercuts — indentations that allow the cast surface to get a locking grip on the mold (James, 1989). Because of the streamlined form of knifefish, we were able to use this type of mold.

Prior to coating the fish with the mold compound, Rhodorsil V-1065 silicone elastomer and Hi-Pro Blue catalyst were mixed in a 10:1 ratio, as specified by the manufacturer. Generally, it is recommended that the mixture be degassed to remove small bubbles trapped in the mixing process, but we did not find this necessary.

The silicone mixture was slowly poured over the fish until it was fully coated. Initially, most of the silicone ran off the surface and had to be recovered and poured over again. This process was repeated over ≈ 30 min, during which time the compound partially cured. Two additional layers were added in this way at ≈ 60-min intervals. The mold was then allowed to cure for several hours (Fig. 1B).

The resulting mold was still quite thin, and needed mechanical reinforcement prior to casting to prevent distortions due to the weight of the casting material. In some cases, a surrounding or 'mother mold' can be constructed (James, 1989) for this purpose. For our application, the mold was reinforced by wrapping a section of light cotton cloth once around the coated fish. Mold compound was applied to the cloth before it was draped around the mold.

After the reinforced silicone mold had fully cured, thin slices were cut away from the caudal end with a razor blade until the posterior tip of the caudal fin was seen. This creates a vent hole, preventing air pockets from forming in the thin end section of the mold during casting. An extraction slit was cut along the dorsal edge 2 cm from the snout that was just large enough to remove the fish without tearing the mold. The mold was washed thoroughly and dried.

2.5. Making the cast

A sprue (pour hole) must be made in the mold to allow entry of the casting compound. A 3-mm diameter sprue was cut through the mold at the caudal end of the extraction slit for the injection of the casting compound with a large syringe. The two parts of the compound were mixed in the 1:1 ratio recommended by the manufacturer. The mixture was not degassed. It was quickly injected into the mold before the material started to set (2 min). A small amount of squeezing pressure on the mold was sufficient to prevent cast mixture leakage through the extraction slit. After 4 h of curing, the cast was carefully removed from the mold, using thin wooden rods pushed along the mold–cast interface to facilitate release. A high quality rigid reproduction of the fish results (Fig. 2A). The surface quality was sufficient to see the lateral line canals and receptor pores (≈ 40 μm diameter) under a light microscope.

2.6. 3-D digitizing

The next objective is to obtain a quantitative representation of the surface of the cast. The surface representation will serve as the basis for model-based tracking in which the model surface is deformed to match video images of the behaving animal. Obtaining a quantitative representation of a surface involves measuring coordinate values of points on the surface and constructing a best fit surface model that passes near those points. We used a stylus-type contact digitizer with 0.2 mm accuracy (MicroScribe 3DX, Immersion, San Jose, CA, USA). The digitizer was operated from within a 3-D modeling software package (Rhinoceros 3D vl. 1, Robert McNeel and Associates, Seattle, WA, USA). Prior to 3-D digitizing the cast, it was securely mounted by drilling two small holes in the cast and gluing a short length of music wire in each for external clamping. The resulting setup is shown in Fig. 2A.

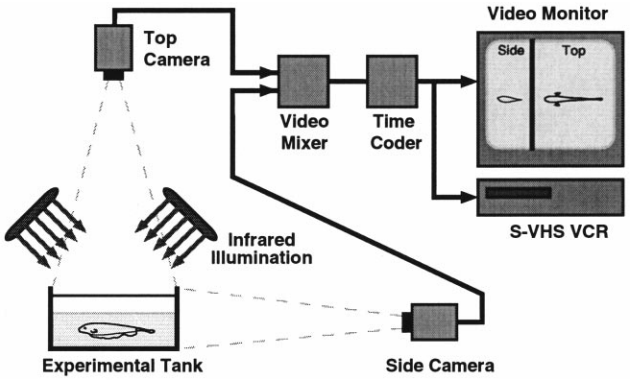

Fig. 3. Schematic diagram of two-camera infrared video setup. The experimental tank and cameras were housed within a light-tight enclosure.

When using the MicroScribe, the user has to select a set of surface points to be digitized. The selection of these points depends on the requirements of the surface generation functions available in the 3-D modeling software used with the MicroScribe. We used the surface generation function 'Sweep2' of Rhinoceros. This function requires two 'rail' curves, in our case corresponding to the dorsal and ventral edges of the fish, and multiple cross-sectional curves between the rails to define the conformation of the surface. Fifteen cross-sectional curves were hand drawn on the cast at 2–10 mm intervals depending upon the change in the surface between the cross-sections. Each of the 15 closed cross-sectional curves was entered into Rhinoceros by touching the curves with the digitizer stylus, with a point spacing of ≈ 1–4 mm depending on the local curvature of the cast. The dorsal- and ventral-edge open rail curves were entered similarly. Following entry, the curves were edited to correct minor distortions due to the moldmaking process, such as unnatural bends in the trunk and abdominal distension due to pooling of fluids. The correction process was facilitated by comparisons to a scaled reference image (Section 2.2). Information on other approaches for obtaining a quantitative representation of a surface is available at http://soma.npa.uiuc.edu/labs/nelson/model_based_tracking.html.

2.7. Creating a polygonal model

The native representation for all objects within Rhinoceros is parameterized non-uniform rational B-spline (NURBs) curves and surfaces (Piegl and Tiller, 1995). While it is possible to develop algorithms to manipulate objects in this format, it is more straightforward to manipulate polygons (Watt and Watt, 1992). We generated two polygonal models from the original parametric representation with two different resolutions. The first polygonal model consisted of 1540 quadrilateral faces, 70 longitudinal and 22 around (Fig. 2B). The second polygon model consisted of 90 faces, 15 longitudinal and six around (Fig. 2C). The low resolution model was used for tracking where the number of nodes needed to be minimized for usable screen redraw rates, and for initial electrosensory signal reconstructions (Section 4.1). The high resolution model is used in some of our electrosensory reconstructions where we need to avoid errors created by the coarse surface discretization of the low resolution model.

3. Model-based tracking

Model-based tracking consists of fitting a model of the tracked object to image or sensor data. In our application, we fit the low resolution polygonal model of the fish (Fig. 2B) to digitized video images from prey capture sequences. Several aspects of this process will be described: infrared videography, video digitizing, camera calibration, 3-D reconstruction, validation, creation of a parametric fish model, and fitting the model to images. All computations discussed below were performed using MATLAB and the Image Processing, Optimization, and Signal Processing toolboxes (The Mathworks, Natick, MA, USA), running on a Sun UltraSparc 2 Unix workstation (Sun Microsystems, Palo Alto, CA, USA).

3.1. Infrared videography

In this section we detail the methods used in video recording the behavior of a nocturnal weakly electric knifefish (*A. albifrons*) as it hunts for small prey (*Daphnia magna*) in the dark. Fish behavior was observed in an aquarium ($383 \times 293 \times 186$ mm) housed within a light-tight enclosure. The aquarium was illuminated with two arrays of 100 high power infrared light emitting gallium arsenide diodes (SIR 333, Everlight Electronics, Taipei, Taiwan). Each diode provides 35 mW of radiant power at a wavelength of 880 nM, which is above the wavelength cutoff for teleost photoreceptors (Fernald, 1988). Fig. 3 shows the configuration of the behavioral recording setup. Some aspects of this system are similar to those described by Rasnow et al. (1997).

Activity of the fish and prey was imaged with two black-and-white CCD cameras with infrared blocking filters removed (VDC 2624, Sanyo Fisher, Chatsworth, CA, USA; XC-77, Sony Electronics, Park Ridge, NJ, USA). These were synchronized with an external signal from a camera adapter (Sony DC-77RR). A video splitter was used to merge the two signals into one split-screen image (AD1470A, American Dynamics, Pearl River, NY, USA). A longitudinal time code generator was used to dub a time code display window onto the video (TC-3, Burst Electronics Inc, Corrales,

NM, USA). This provides time-stamping of each field of the behavioral sequence (two fields drawn 16.7 ms apart comprise one video frame). The video signal was recorded on a S-VHS format videocassette recorder (VCR) using S-VHS videotape (AG-7350-P, Panasonic Communications and Systems, Secaucus, NJ, USA; ST-126 videotape, Maxell, NJ, USA).

3.2. Video digitizing and image processing

Recorded video of animal behavior was played back on an S-VHS player and input to a video digitizing system (Avid Media Composer 1000-7, Avid Technology, Tewksbury, MA, USA). Video signals were digitized using the monochrome AVR 77 format and exported as 720×486 pixel 8-bit grayscale TIFF files.

After digitization, a number of image manipulations were performed in order to eliminate motion interlace blur, increase small-object contrast, and resize the image. Motion interlace blur refers to an image distortion created by the way video images are displayed. To reduce flicker, the horizontal scan lines of a video image are drawn in two sets, the first set consisting of the odd-numbered lines, the second set consisting of the even-numbered lines. Thus, there is a brief interval (16.7 ms for video in North America) between adjacent scan lines. Movement within this interval causes motion interlace blur. This artifact was eliminated and the effective frame rate was doubled to 59.94 frames/s by deinterlacing with the missing scan lines interpolated using bicubic interpolation. Each image was then contrast enhanced by subtracting a 2-D median filtering of the image from the original and adjusting intensity values. Finally, the images were resized from 720×486 (nonsquare TV pixels) to 720×540 pixels. For our studies, video of 120 prey capture events of 1–2 s in duration were digitized, resulting in 6 gigabytes of image data.

Using standard video resolution test patterns we determined the resolving power of the final images to be ≈ 1 1ine/mm along both dimensions, representing the minimum resolvable width of alternating black and white lines. The prey used in our study, 2–3 mm in length, are just resolvable under these conditions. For additional technical information on video resolution (Jack, 1993; Young et al., 1995; Poynton, 1996). How to estimate system resolution from camera and recording format specifications and other video information is provided in http://soma.npa.uiuc.edu/labs/nelson/model_based_tracking.html.

3.3. Implicit image correction and 3-D reconstruction

Recovering accurate 3-D position information from 2-D camera images is an active topic of research in the field of machine vision. Current methods often employ a model of the camera based on physical parameters such as focal length and principal point (Tsai, 1987). Such methods based on physical parameters of the camera are termed explicit methods. An alternative approach, termed implicit image correction, utilizes a set of non-physical parameters without reference to a camera model. The implicit method is motivated by the observation that the physical parameters of the camera are of little interest when only the relationship between 3-D reference coordinates and 2-D image coordinates is required. Implicit image correction (Heikkilä and Silvén, 1996, 1997) can achieve very high accuracy without the complexity and computational overhead of a rich camera model. We chose to use a simple implicit method, which in our application has the additional advantage that distortions due to water refraction are automatically taken into account.

The geometry of the reconstruction problem for our two-camera system is schematized in Fig. 4. The portion of the scene in view for each camera is termed the camera's window, while the portion that is finally displayed on the monitor after the video passes through the video splitter is called the splitter window. The 720×540 pixel space of the digitized image is the image coordinate system (ICS). We identify two portions of the ICS, the side viewport (i, j) and top viewport (k, l), corresponding to the area imaged by the side and top camera splitter windows. The $383 \times 293 \times 186$ mm space of the tank is the tank coordinate system (TCS). The cameras are positioned so that their sight lines are approximately orthogonal to the face of the tank closest to the camera.

In our implicit image correction method, which does not take into account radial or tangential distortion, the $[(i, j), (k, l)]$ image coordinates are related to the (x, y, z) tank coordinates by the following transformation matrix:

$$\begin{bmatrix} i \\ j \\ k \\ l \end{bmatrix} = \begin{bmatrix} 0 & 0 & -S_i(y) \\ S_j(y) & 0 & 0 \\ 0 & -S_k(z) & 0 \\ S_l(z) & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} b_i(y) \\ b_j(y) \\ b_k(z) \\ b_l(z) \end{bmatrix} \quad (1)$$

where $S_{i,j}(y)$, $S_{k,l}(z)$ are scale factors (pixel/mm), $b_{i,j}(y)$, $b_{k,l}(z)$ are offsets (pixel); (i, j); (k, l) are side viewport and top viewport image coordinates (pixel), and (x, y, z) are tank coordinates (mm).

Because of the camera perspective, the scale factors and offsets depend on the distance of the image plane from the camera. Thus the side view parameters (S_i, S_j, b_i, b_j) depend on the tank coordinate y of the imaged

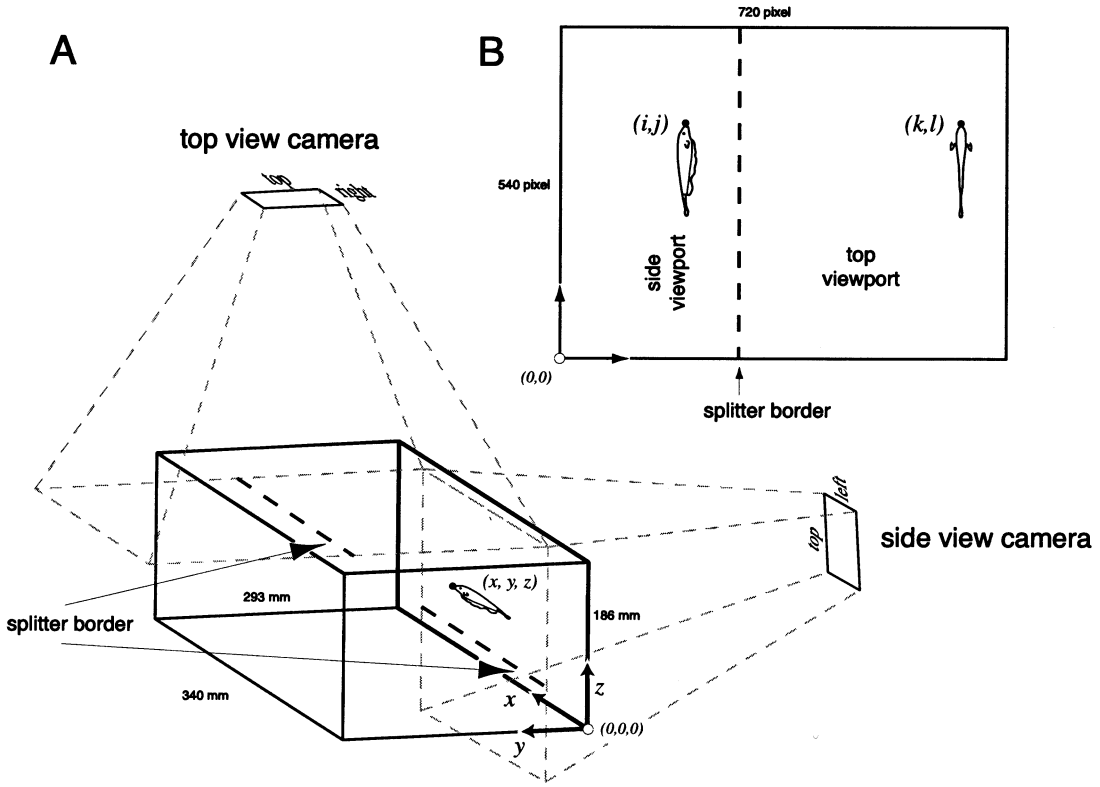

Fig. 4. The geometry of the reconstruction problem. (A) The behavioral tank and configuration of the two cameras. (B) Each video image displays two projections of tank objects, as illustrated by the knifefish outline: one in the side viewport and one in the top viewport. The tank (x, y, z) and image coordinate $[(i, j), (k, l)]$ labels for the snout of the fish are shown.

point, while the top view parameters (S_k, S_i, b_k, b_l) depend on z.

To measure these distance-dependent scale and offset parameters we populate the proximal and distal planes in both the side and top camera splitter windows with control points using an accurate 1-cm planar grid. The pixel coordinates of the intersection of all grid lines in view within the ICS were measured and recorded semi-automatically using a custom MATLAB script. A constrained optimization function was then called for each of the four sets of digitized and measured calibration points to fit 2-D scaling and offset factors that minimize the aggregate Euclidean distance between the measured and predicted ICS points, where the prediction is obtained by transforming the known TCS coordinates according to Eq. (1).

Having determined scaling factors at the proximal and distal tank walls in each view, we linearly interpolate to arrive at the appropriate scale and offset factors for intermediate positions. For example, the scale factor $S_i(y)$ for image coordinate i is computed as

$$S_i(y) = S_i^{\text{prox}} + (S_i^{\text{dist}} - S_i^{\text{prox}})\frac{y}{y_{\text{tank}}} \quad (2)$$

where S_i^{prox} (S_i^{dist}) is the scaling factor for the calibration grid that is proximal (distal) to the camera, y is the coordinate of the point in the TCS, and y_{tank} is the total extent of the tank along the y-dimension. Similar equations apply for other scaling and offset parameters. The model therefore has a total of 16 free parameters: two scale factors and two offsets for each of four viewplanes (top proximal and distal, side proximal and distal).

3.4. 3-D reconstruction validation

To validate the calibration and 3-D reconstruction procedures, we digitized 2 s (60 frames) of video of a 150 mm rod being randomly moved through the tank. The (i, j, k, l) image coordinates of a point at each end of the rod were measured. We then inverted Eq. (1) to numerically solve for the (x, y, z) tank coordinates of each end of the rod. Fig. 5 shows a plot of the computed length of the moving rod versus time. The RMS error was 0.48 mm, with a maximum error of 1.0 mm.

The accuracy of the reconstruction is also continually validated during model-based tracking (Section 3.5). Correspondence between the two projected polygonal fish models and the position of the real fish provides a cross-check that helps alert the user to unintended changes in the optical pathway, such as slight shifts in camera position, which necessitate recalibration of the system.

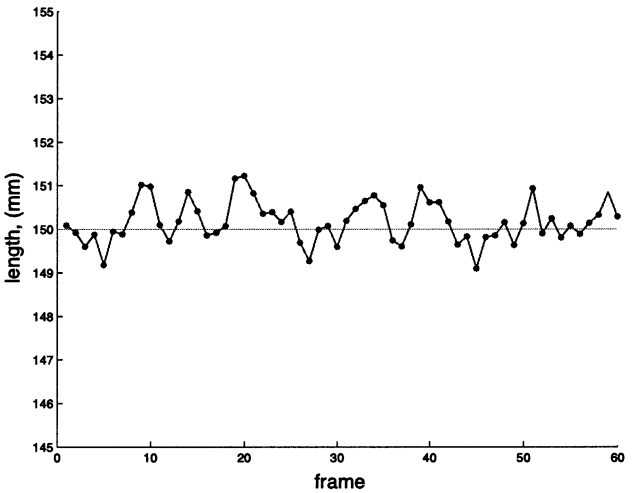

Fig. 5. Reconstructed rod length over 60 frames of video. Actual rod length was 150 mm, shown by the horizontal solid line. RMS error: 0.48 mm.

3.5. Creation of a parametric fish model

Having established a method for accurately transforming points between tank and image coordinate systems, we can use the video images to determine the position and shape of the fish during behavioral sequences. Because we are interested in the conformation of a non-rigid object (the fish body) we do not simply digitize key points on the image (e.g. head, tail, fins, etc.). Rather we represent the entire surface of the fish using the polygonal model described earlier. The model is parameterized with suitable degrees of freedom (DOF) to allow it to translate, rotate, and change shape.

We implement six rigid-body DOF: three for position (x, y, z), and three for rotation (roll, pitch, yaw), using standard geometric methods (Mortenson, 1985). Determining the non-rigid DOF adequate to describe the range of body conformations of interest is an iterative process. In knifefish, locomotion is achieved by generation of traveling waves along the ventral ribbon fin, while the trunk of the fish remains relatively straight or follows a shallow spline-like curve (Blake, 1982; Lighthill and Blake, 1990; Sfakiotakis et al., 1999). We modeled the lateral bend of the trunk with one non-rigid DOF that specified the deviation of the tip of the tail from the midline. The lateral displacement of the body was described by a cubic spline curve, which is widely used in geometric modeling of natural objects (Terzopoulos et al., 1987). The cubic spline was computed with a MATLAB function (de Boor, 1978; Hearn and Baker, 1997). The input points to the spline function were midline points for the non-flexing anterior of the fish body and one point at the tip of the tail. The nodes of the polygonal model were then displaced such that the midline followed the spline curve.

Initial tracking studies using this model revealed that the fish also flexes its spine in a dorsal–ventral plane, a subtle movement we had not previously noticed and which is difficult to see without the visual aid of the overlaid polygonal model. This observation required the use of one additional non-rigid DOF to parameterize the dorsal–ventral bend, modeled and computed in the same manner as the lateral bend. Thus the final model that we used had eight DOF (six rigid, two non-rigid). Using this model we were able to describe the vast majority of the fish's conformations, with rare exceptions such as when the fish briefly enters into an 'S' shape. Representing the fish body in this parametric way is compact: rather than save the (x, y, z) coordinates of each node of the polygonal model for each video field, we save only the eight fitted parameters needed to reconstruct the polygonal model's position and shape. When digitizing prey capture sequences, the 2–3 mm diameter prey (*D. magna*) was modeled using three DOF representing the (x, y, z) coordinates of its center.

3.6. Fitting the model to images

Model-based tracking involves determining the position and shape of fish in the tank by manipulating the fish model until its projections are congruent with the imaged fish. First, a representative polygonal model is scaled to the size of the particular individual being studied. Second, a digitized image from the behavioral sequence is displayed on the computer screen within the animal tracking program interface (see Fig. 6). The user interface contains eight controls corresponding to the eight DOF of the parametric model. Using the implicit image correction transformation (Eq. (1)), a wireframe visualization of the polygonal model is projected from tank coordinates into image coordinates, resulting in two projected wireframes, one in the side viewport and one in the top viewport. Third, by adjusting the values of the eight DOF, the user moves the polygonal model in the tank coordinate system until the two wireframe projections are congruent with the projections of the real fish on the digitized image. In order to attenuate manual model placement jitter, each adjustment is zero-phase filtered through a digital 5th-order Butterworth filter (6 Hz passband corner frequency). Some aspects of this method are similar to those used by Assad (1997).

In general, it is not necessary to fit the polygon model to every field of the video sequence. In our application, intervals of approximately eight video fields (133.5 ms) are used initially with intermediate positions estimated by cubic spline interpolation of each DOF. A second pass through the sequence is made to verify the accuracy of the interpolation and set additional fields as necessary.

4. Discussion

4.1. Linking behavior to neurophysiology

Nocturnal black ghost knifefish (*A. albifrons*) are able to locate objects without visual cues by sensing perturbations in a weak self-generated electric field (Bastian, 1986, 1995; Bullock and Heiligenberg, 1986; Turner et al., 1999). Perturbations in the field caused by objects that differ in impedance from the surrounding water cause changes in the voltage across the skin. These transdermal potential changes are transduced into trains of action potentials by $\approx 10^4$ electroreceptors that cover most of the body surface. By dynamically controlling the positioning of their surface sensory array these fish actively influence the strength and spatiotemporal pattern of the incoming electrosensory signals (Nelson and MacIver, 1999).

In our application, quantitative behavioral analyses of black ghost knifefish and *Daphnia* trajectories allows us to infer properties of the sensory signals reaching the brain through the primary electrosensory afferents during prey capture behavior (Nelson and MacIver, 1999). To characterize the incoming electrosensory signals, fish and prey trajectories were reconstructed at time steps of 16.7 ms. At each time step, we compute the spatial distribution of transdermal voltage changes on the skin based on the physics of electric image formation (Rasnow, 1996). Fig. 7A shows the resulting pattern of transdermal potential change for a representative prey capture sequence. Note that the electric image is weak and diffuse at the beginning of the sequence and becomes both more intense and more tightly focused as the *Daphnia* comes closer to the electroreceptor array. Based on the estimated transdermal potential change, we then compute the corresponding change in afferent firing rate based on a model of electrosensory afferent response dynamics (Nelson et al., 1997). Fig. 7B shows the estimated change in afferent firing rate corresponding to the change in transdermal potential shown in Fig. 7A.

These electrosensory image reconstructions have given us a better understanding of the interactions between sensory and motor aspects of active sensory acquisition (Nelson and MacIver, 1999). We have also determined that in the brief period (≈ 600 ms) between the fish's initial reaction to the presence of the prey and the capture of the prey, the fish are able to dynamically change their post-detection posture to compensate for displacement of the prey during the prey strike (MacIver and Nelson, 1999). We can conclude that the animal is able to use feedback control of its position

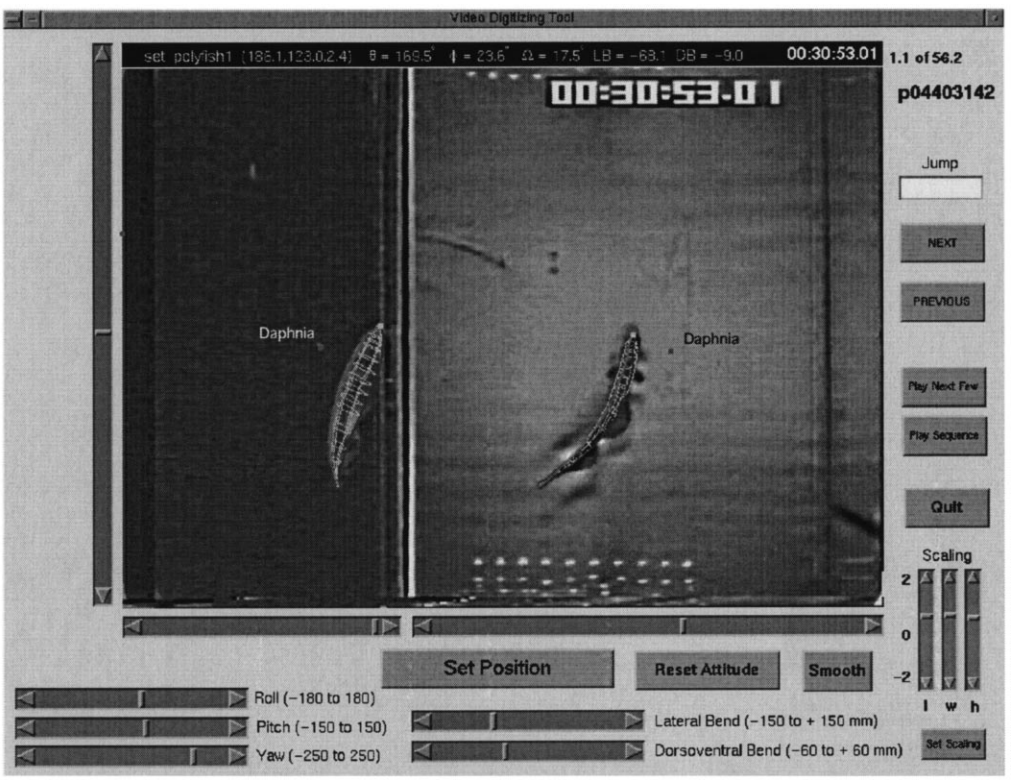

Fig. 6. Snapshot of the animal-tracking interface. The sliders adjacent to the image control the (x, y, z) position of the snout; the sliders below control yaw, pitch, roll, lateral bend, and dorsal-ventral bend. The scaling sliders control the subject-specific scaling of the polygonal model. The user navigates through the fields of the behavioral sequence and manipulates the polygonal model so that its top and side viewport projections are congruent with those of the real fish.

during the strike, rather than using a ballistic strike (Gilbert, 1997). Observing and quantifying these behaviors without the use of model-based tracking would have been exceedingly difficult.

4.2. Other applications

The techniques that we have presented for body modeling and model-based tracking worked particularly well for our application. In part, this can be attributed to three factors: the simple body plan of the knifefish and the corresponding small number of degrees of freedom, the occlusion-free viewing environment of the aquarium, and the relatively short duration (typically <2 s) of the behavioral sequences that needed to be reconstructed. While our specific application therefore represents a relatively simple case, the techniques that we have described could be extended to handle more complicated problems including (1) body modeling of animals with more complex body plans and more degrees of freedom; (2) fitting of the model to partially occluded video images, and (3) reconstructing longer behavioral sequences.

For more complex body plans, it may be useful to build up a body model from multiple components (e.g. head, neck, trunk, limbs, etc.). Using methods similar to those outlined above, a surface representation could be obtained either for the entire animal, or for each body component separately. The rigid and non-rigid degrees of freedom for each component could then be modeled. Finally, the components would need to be linked together, using a technique such as the hierarchical method described in Jung and Wohn (1997). Also, depending on the requirements of the study, it may be possible to simplify the problem by modeling only a small subset of the body components and degrees of freedom, or by tracking movements in two dimensions rather than three.

In video tracking studies, occlusions can occur either due to objects in the environment or due to self-occlu-

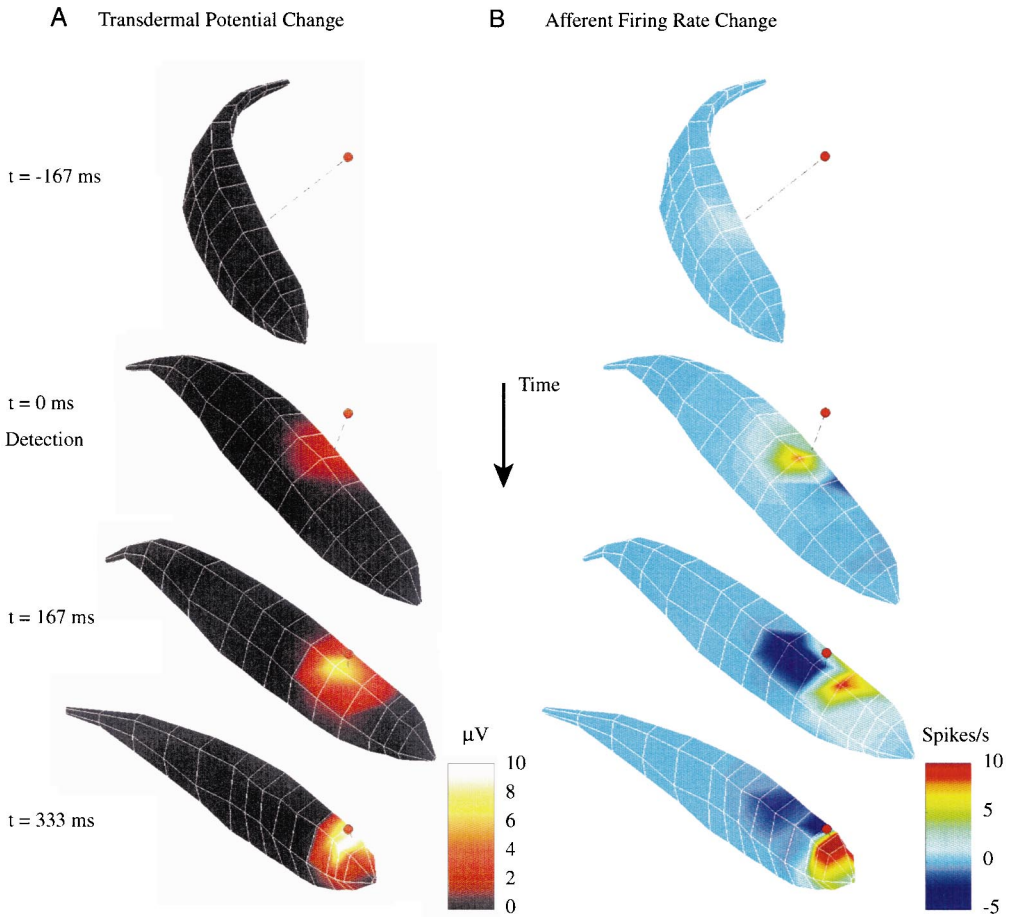

Fig. 7. False color maps of reconstructed electrosensory images generated from model-based tracking of a single prey capture sequence. The weakly electric fish (*Apteronotus albifrons*) is able to detect prey in the dark by sensing small perturbations in a self-generated electric field. Each column shows 'snapshots' of the polygonal fish model at four different times in the prey capture sequence. The left-hand column (A) shows the voltage change across the skin (transdermal potential) induced by the prey. The right-hand column (B) shows the corresponding change in electrosensory afferent firing rate due to the voltage perturbations shown in (A). The prey *(Daphnia magna)* is shown as a red dot; the dashed line represents the shortest distance between the fish and the prey. Modified from Nelson and MacIver (1999).

sion when one part of the body overlaps with another part in the projected image. Predictive tracking methods can help in these cases (Jung and Wohn, 1997), as can the addition of more camera views. These methods typically employ physics-based deformable models (Terzopoulos et al., 1987; Essa et al., 1993; Metaxas, 1996) or active contours (Blake and Isard, 1998) to constrain the placement of the model in images of a scene from multiple perspectives. This is done using computer-automated fitting techniques, rather than manual model placement as described in this study (Mochimaru and Yamazaki, 1994; Gavrila and Davis, 1996; Jung and Wohn, 1997; Tillett et al., 1997).

Manual model fitting is time consuming and establishes a practical limit for the length of video sequences that can be reconstructed. In our application, reconstruction of a 1–2 s sequence required ≈ 45 min per sequence. Automatic model fitting techniques reduce the amount of user involvement required and thus have the advantage of enabling longer-term behavioral observations. We are currently exploring the use of automatic model fitting methods in our studies.

4.3. Future directions

The data that we obtain from model-based tracking is rich. One of the challenges we face is the visual display of 3-D data for selection and quantitative analysis of behavioral patterns. In 2-D projections of prey capture sequences, the absence of depth cues makes interpretation of the movements difficult. Thus, we have recently utilized a virtual reality system developed at the University of Illinois (CAVE, Beckman Institute, Urbana, IL, USA) to visualize prey capture reconstructions in 3-D. In this facility, a stereo image of the fish and prey is projected onto three walls of a room and the floor using four projectors. Liquid crystal stereo glasses provide the illusion that the fish and prey are floating in space within the CAVE. A computer tracks the user's position and gaze direction and dynamically changes the visual display accordingly. The playback speed, direction, position, and zoom level of the prey capture sequences are controlled by a hand-held joystick. We have used this system to identify subtle aspects of the behavior that were not previously observed while viewing the prey capture sequences on 2-D workstation monitors.

Finally, we are in the process of designing a biomimetic robot based on the weakly electric fish in order to test active sensing hypotheses in the electrosensory system. We have utilized a very high resolution (258 609 polygon) version of the fish model discussed above to generate a physical realization of the model using a stereolithography apparatus (SLA-50, 3D Systems, Valencia, CA, USA). This apparatus consists of a tank of photosensitive resin and a computer controlled laser. The laser scans the tank of resin to build up a rigid model in layers that are 4.2 μm thick.

The sensory acquisition mechanisms of interest in our research are the adaptive control of body posture and the descending control of sensory filtering in the brain. By combining precision behavioral quantification, neurophysiology, neural simulations, and biomimetic robotics we hope to elucidate these mechanisms subserving the remarkable sensory abilities of weakly electric fish. In general, these methods may provide a bridge between analytical methods of studying adaptive behavior and synthetic approaches (Ekeberg et al., 1995; Terzopoulos et al., 1995, 1997; Chiel and Beer, 1997; Beer et al., 1998).

Acknowledgements

We thank Ben Grosser and the staff of the Visualization, Media, and Imaging Laboratory of the UIUC Beckman Institute for their assistance with video digitizing and the MicroScribe. This research was supported by a grant from the National Institute of Mental Health (ROlMH49242).

References

Assad C. Electric field maps and boundary element simulations of electrolocation in weakly electric fish. California Institute of Technology, 1997. Unpublished doctoral dissertation.

Bastian J. Electrolocation: behavior, anatomy, and physiology. In: Bullock T, Heiligenberg W, editors. Electroreception. New York: Wiley, 1986:577–612.

Bastian J. Electrolocation. In: Arbib M, editor. The Handbook of Brain Theory and Neural Networks. Cambridge, MA: MIT Press, 1995:352–6.

Beer RD, Chiel HJ, Quinn RD, Ritzmann RE. Biorobotic approaches to the study of motor systems. Curr Opin Neurobiol 1998;8:777–82.

Blake A, Isard M. Active Contours. New York, NY: Springer, 1998.

Blake R. Swimming in the electric eels and knifefishes. Can J Zool 1982;61:1432–41.

Boardman ET. Techniques of life casting of small vertebrates. Museum News Washington 1950;28(11):7–8.

Bromage TG. Systematic inquiry in tests of negative/positive replica combinations for SEM. J Microsc 1985;137(2):209–16.

Bullock TH, Heiligenberg W, editors. Electroreception. New York: Wiley, 1986.

Chiel HJ, Beer RD. The brain has a body: adaptive behavior emerges from interactions of nervous system, body, and environment. Trends Neurosci 1997;20:553–7.

de Boor C. A Practical Guide to Splines. New York: Springer, 1978.

Ekeberg Ö, Lansner A, Grillner S. The neural control of fish swimming studied through numerical simulations. Adapt Behav 1995;3(4):363–84.

Essa I, Sclaroff S, Pentland A. Physically based modelling for graphics and vision. In: Martin E, editor. Directions in Geometric Computing. UK: Information Geometers, 1993.

Fernald RD. Aquatic adaptations in fish eyes. In: Atema J, Fay R, Popper A, Tarolga W, editors. Sensory Biology of Aquatic Animals. New York: Springer, 1988:433–66.

Gardner GS. Casting lifelike models from living animals. Curator 1974;17(1):10–5.

Gavrila D, Davis L. 3-D model-based tracking of humans in action: a multi-view approach. In: 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition: June 18–20, 1996, San Francisco, California. Los Alamitos, CA: IEEE Computer Society Press, 1996:73–80.

Gilbert C. Visual control of cursorial prey pursuit by tiger beetles (Cicindelidae). J Comp Physiol A Sens Neural Behav Physiol 1997;181(3):217–30.

Hearn D, Baker MP. Computer graphics, C version, 2nd edition. Upper Saddle River, NJ: Prentice Hall, 1997.

Heikkilä J, Silvén O. Calibration procedure for short focal length off-the-shelf CCD cameras. In: Kropatsch W, editor. Proceedings of the Thirteenth International Conference on Pattern Recognition. Los Alamitos, CA: IEEE Computer Society Press, 1996:166–70.

Heikkilä J, Silvén O. A four-step calibration procedure with implicit image correction. In: Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Los Alamitos, CA: IEEE Computer Society Press, 1997:1106–12.

Hughes N, Kelly L. New techniques for 3-D video tracking of fish swimming movements in still or flowing water. Can J Fish Aquat Sci 1996;53(11):2473–83.

Jack K. Video Demystified. A Handbook for the Digital Engineer. Solana Beach, CA: HighText, 1993.

James T. The Prop Builders Molding and Casting Handbook. Cincinnati, OH: Betterway, 1989.

Jung K, Wohn S. Tracking and motion estimation of the articulated object: a hierarchical Kalman filter approach. Real-Time Imag 1997;3(6):415–32.

Kruk M. Measuring behaviour into the twenty-first century. Trends Neurosci 1997;20(5):187–9.

Lighthill J, Blake R. Bigfluiddynamics of balistiform and gymnotiform locomotion. 1. Biological background, and analysis by elongated-body theory. J Fluid Mech 1990;212:183–207.

MacIver M, Nelson M. Evidence for closed-loop control of prey capture in weakly electric fish. In: Abstracts, Eighth Annual Computational Neuroscience Meeting, Pittsburgh, PA, July, 1999:18–22.

McHenry MJ, Pell CA, Long JH Jr. Mechanical control of swimming speed: stiffness and axial wave form in an undulatory fish model. J Exp Biol 1995;198:293–2305.

Metaxas DN. Physics-Based Deformable Models. Boston, MA: Kluwer, 1996.

Mochimaru M, Yamazaki N. Three-dimensional measurement of unconstrained motion using a model-matching method. Ergonomics 1994;37(3):493–510.

Mortenson ME. Geometric Modeling. New York: Wiley, 1985.

Nelson M, MacIver M. Prey capture in the weakly electric fish *Apteronotus albifrons*: sensory acquisition strategies and electrosensory consequences. J Exp Biol 1999;202(10):1195–203.

Nelson ME, Xu Z, Payne J. Characterization and modeling of P-type electrosensory afferent response dynamics in the weakly electric fish *Apteronotus leptorhynchus*. J Comp Physiol A 1997;181:532–44.

Parsons KC. Precision casting: a new method in museum technology. Am J Phys Anthropol 1973;38(3):789–802.

Piegl L, Tiller W. The NURBS Book. New York: Springer, 1995.

Poynton CA. A Technical Introduction to Digital Video. New York: Wiley, 1996.

Rasnow B, Assad C, Hartmann MJ, Bower JM. Applications of multimedia computers and video mixing to neuroethology. J Neurosci Methods 1997;76(1):83–91.

Rasnow B. The effects of simple objects on the electric field of *Apteronotus*. J Comp Physiol A 1996;178(3):397–411.

Sfakiotakis M, Lane D, Davies JBC. Review of fish swimming modes for aquatic locomotion. IEEE J Oceanic Eng 1999;24(2):237–52.

Spruijt B, Thorwald H, Rousseau J. Approach, avoidance and contact behavior of individually recognized animals automatically quantified with an imaging technique. Physiol Behav 1992;51:747–52.

Terzopoulos D, Platt J, Barr A, Fleischer K. Elastically deformable models. Comput Graph 1987;21(4):205–14.

Terzopoulos D, Tu X, Grzeszczuk R. Artificial fishes: autonomous locomotion, perception, behavior, and learning in a simulated physical world. Artif Life 1995;1:327–51.

Terzopoulos D, Rabie T, Grzeszczuk R. Perception and learning in artificial animals. In: Artificial life V: Proceedings of the Fifth International Workshop on the Synthesis and Simulation of Living Systems. Nara, Japan, May, 1996. Cambridge, MA: MIT Press, 1997:313–20.

Tillett RD, Onyango CM, Marchant JA. Using model-based image processing to track animal movements. Comput Electron Agric 1997;17(2):249–61.

Tsai RY. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. IEEE J Robotics Automation 1987;3(4):323–44.

Turner R, Maler L, Burrows M. Electroreception and electrocommunication. Cambridge, UK. J Exp Biol 1999;202:10.

Vatine J, Ratner A, Dvorkin M, Seltzer Z. A novel computerized system for analyzing motor and social behavior in groups of animals. J Neurosci Methods 1998;85(1):1–11.

Waters PH. A review of the moulding and casting materials and techniques in use in the Palaeontology Laboratory British Museum (Natural History). Conservator 1983;7:37–43.

Watt A, Watt M. Advanced Animation and Rendering Techniques. New York: ACM Press, 1992.

Winberg S, Nilsson GE, M SB, Hoglund U. Spontaneous locomotor activity in arctic charr measured by a computerized imaging technique: role of brain serotonergic activity. J Exp Biol 1993;179:213–32.

Young L, Poynton C, Schubin M, Watkinson J, Olson T, editors. Pixels, Pictures, and Perception: The Differences and Similarities Between Computer Imagery, Film, and Video. White Plains, NY: Society of Motion Picture and Television Engineers, 1995.