Towards three-dimensional point cloud reconstruction of fish swimming

Mert Karakaya[a], Chen Feng[a,b], and Maurizio Porfiri[a,c]

[a]Department of Mechanical and Aerospace Engineering, New York University Tandon School of Engineering, Six MetroTech Center, Brooklyn, NY 11201, USA
[b]Department of Civil and Urban Engineering, New York University Tandon School of Engineering, Six MetroTech Center, Brooklyn, NY 11201, USA
[c]Department of Biomedical Engineering, New York University Tandon School of Engineering, Six MetroTech Center, Brooklyn, NY 11201, USA

ABSTRACT

Zebrafish is extensively used in behavioral, pharmacological, and neurological studies due to a number of methodological and practical advantages, including genetic and neurobiological homologies with humans and a fully sequenced genome. Critical to a biologically-based understanding of zebrafish behavior is the ability to reconstruct their complex behavioral repertoire in three-dimensions. Toward this aim, several efforts have been made to score their ethogram in three-dimensions, but most of these studies are constrained by a single-view imaging. A promising line of approach to extract refined information about the mechanosensory and perceptual systems of zebrafish is point cloud reconstruction. Here, we provide an initial review of the state of knowledge in zebrafish tracking and we propose a potential methodology that can capture the dynamic three-dimensional geometry of fish swimming. We utilize a stereo vision camera, calibrated with a pinhole camera model with refraction correction to allow for multi-medium imaging. The corrected pinhole camera model accounts for refraction through multiple mediums and allows for more accurate point cloud reconstruction from two cameras. From the point cloud data, we could recreate the three-dimensional geometric model of the fish and analyze its swimming behavior in three dimensions. The extracted dynamic fish geometry should allow for an improved understanding of mechanosensation and perception, which are critical to elucidate how zebrafish process visual cues and perceive flow structures.

Keywords: camera calibration, mechanosensation, point cloud reconstruction, stereo-vision, zebrafish

1. INTRODUCTION

Zebrafish is a popular model organism for behavioral, pharmacological, and neurological research.[1,2,3,4] Several factors have fueled the widespread adoption of zebrafish as an animal model in biomedical research, including homologies with humans at the genetic and neurobiological levels[5] as well as a fully sequenced genome.[6] A critical step in zebrafish experimentation is the tracking of its motion, which has historically been performed manually by scoring the behavior from recorded videos and, more recently, through automated tracking softwares.[7]

Despite clear advantages of automated tracking software with respect to high throughput testing and accuracy, they largely rely on the mere analysis of the motion of the zebrafish centroid in two dimensions. The added value of three-dimensional analysis of zebrafish behavior has been recently examined by Macrì *et al.*,[8,9] where it has been shown that tracking the centroid of the fish in three dimensions could lead to different predictions than standard practice, based on single views from the top or the side of the tank. Across a range of conditions, Macri et al.[8,9] have demonstrated that two-dimensional analysis could beget inaccurate and sometimes erroneous predictions.

Whether or not extending the three-dimensional analysis to full body tracking could beget further improvements is yet to be experimentally tested. There is evidence, however, that supports the need for technical

Further author information: (Send correspondence to M.P.)
M.P.: E-mail: mporfiri@nyu.edu, Telephone: 1 646 997 3681

progress on full-body reconstruction. For example, fear response and mating strategies in zebrafish are likely to be better identified from the shape of the tail as well as the body undulations.[10] Likewise, a three-dimensional model of a fish could help resolve gills and eyes, which would aid in the study of mechanosensing in zebrafish.[11]

Our ability to score the rich ethogram of zebrafish is limited to data extracted from point tracking, based on the animal centroid.[7] The objective of this paper is twofold, first we review existing methods for zebrafish tracking and, second, we propose a tenable alternative that could assist in the reconstruction of the full-body of a zebrafish swimming. Through the review, we examine two- and three-dimensional tracking techniques that are currently available and widely used, and we summarize some of the recent efforts on three-dimensional full-body reconstruction method – yet to be tested on small fishes, like zebrafish.

More specifically, we first review readily available software capable of two-dimensional tracking. Then, we introduce open-source software for three-dimensional trajectory reconstruction. Finally, we discuss the possible methods that could be used for full-body reconstruction, highlighting pros and cons of each method. Grounded in the state-of-art review, we propose a novel experimental framework to extract point cloud of a fish swimming using a stereo camera. Cameras are calibrated using a pinhole camera model with refraction correction that accurately reconstructs the three-dimensional geometry of the fish and minimizes the effects of light refraction. The approach allows for robust representation of zebrafish swimming with only one stereo camera setup, by leveraging in-plane constraints of stereo triangulation and estimates depth through geometry.

2. LITERATURE SURVEY

In experimental assays, zebrafish behavior is typically analyzed from the reconstructed trajectories of the centroid. There is a variety of open-source software available to extract trajectories from experimental videos in two- and three-dimensions.[7] Besides trajectories, other methods have been introduced to build full-body movement of a fish with state-of-art computer vision technologies.

2.1 Two-dimensional tracking

There is a variety of freely available and high-throughput software to extract zebrafish trajectory from videos.[7] Some of these software are based on the use of a threshold, along with a predictor and a cost assignment algorithm.[8,12] Others rely on machine learning methods to extract trajectories while maintaining identities of the individual animals.[13,14,15] The former class of software has high levels of success, but it requires manual corrections, whereas machine learning-based software demand minimal manual input while maintaining high accuracy, but are computationally expensive.

Regardless of their specific attributes most of them don't account for light refraction at the water surface that might confound experimental output, and are limited to two-dimensional scoring of animal behavior. These software are generally successful with experiments conducted in shallow water, where the position in the water column of a zebrafish plays a secondary role. However, these settings could be quite unnatural for live animals, which display a rich behavioral repertoire along the water column, such as geotactic response due to an imminent or a potential threat.[10] Macrì *et al.*[8,9] showed that two-dimensional analysis could result in false-positives as well as false-negatives in comparison with three-dimensional analysis, thereby questioning the use of two-dimensional approaches.

2.2 Three-dimensional tracking

Three-dimensional tracking has been implemented with success with a variety of algorithms.[8,14,16,17,18,19] Most commonly, experiments are recorded using two cameras, one from the front view and other from the top view.[8,16,18,20] In some cases, both a stereo-camera setup and a third camera is used to film the experiments.[17] Trajectories are extrapolated by triangulating the centroid with respect to known positions of the cameras.

As a result, it is possible to integrate multiple views toward a three-dimensional representation of the centroid motion, from which one could better elucidate zebrafish ethogram. While the use of two different views is sufficient for reconstructing the centroid trajectory, it is not possible to identify feature points that are needed to populate a point cloud for potential full-body reconstruction. Recently, Al-Jubouri *et al.*[17] have proposed a setup that might be capable of extracting point clouds, however their results indicate limited accuracy through stereo vision, without accounting for light refraction at the water surface.

2.3 Full-body reconstruction

Full-body reconstruction of a zebrafish is part of the state-of-the-art in neurological research, where a zebrafish larva is held stationary and its images are taken from multiple angles.[21] With this method, it is not possible to capture the dynamic behavior of an adult zebrafish, thereby prompting the exploration of other techniques. For example, Lin *et al.*[22] proposed to use a binocular stereo vision setup to extract a point cloud from images of a fish swimming, where their method accounts for refraction through an improved disparity image calculation algorithm. This method has not been tested with fish of similar size to zebrafish, therefore its success in obtaining the point cloud for small subjects is unknown. Ichimaru *et al.*[23] recommends the use of a structured light and a stereo camera to reconstruct the full-body of a fish, and accounts for refraction using Machine Learning techniques. However, a structured light might cause an unwanted stimulus that could alter zebrafish behavior.[24]

Current methods present viable options for full-body reconstruction, however they have not been experimentally tested on adult zebrafish swimming in typical experimental arenas. The three-dimensional imaging of a zebrafish larvae is routinely performed on a stationary subject in a test tube, which is a very different setting from behavioral phenotyping of adult animals. In the studies by Lin *et al.*[22] and Ichimaru *et al.*,[23] state-of-art algorithms for disparity mapping are introduced, which aid in the full-body reconstruction of a fish. In both efforts, the fish is at least three times the body-length of a zebrafish, and fish behavior is not analyzed. Lin *et al.*[22] demostrated the accuracy of their approach in the body-length estimation of the fish, but accuracy in the depth estimation is presently unknown. The latter is of particular significance in the analysis of geotactic response in fear-evoking experiments.[10] Ichimaru *et al.*[23] introduced an external light source through a projector to create a dynamic pattern, which, however, might confound zebrafish response.

3. APPROACH

Here, we explore the feasibility of an alternative method of full-body reconstruction of zebrafish that does not rely on external light sources other than ambient lighting and can beget accurate estimates of depth. Our approach leverages in-plane constraints of stereo vision triangulation,[25] and the use Snell's law to estimate the depth of a point with high computational efficiency. Through pilot observations, we confirm the applicability of the method for full-body reconstruction of a freely swimming zebrafish.

3.1 Experimental setup

Two overhead cameras (Logitech C930, Logitech, Lausanne, Switzerland) spaced 9.5 cm apart and fixed on an aluminum frame, were placed 9.5 cm above a circular tank (Fig. 1). The cameras were calibrated using a checkerboard pattern and Matlab R2020a stereo camera calibration toolbox with a pinhole camera model.[26] We used 75 pairs of checkerboard images and selected 50 image pairs that provided a mean projection error less than 0.39 pixels. From the calibration output, we took the focal point of the first camera as the origin so that the second camera's focal point was located at (9.42, 0.48, -0.89 cm) with respect to our world Cartesian reference system (X, Y, Z).

We drew 25 black dots on the bottom of the tank to add detectable features over the homogeneous background provided by the tank. We performed experiments at different water depths, to demonstrate the reliability of the proposed approach. Specifically, we filled the tank with 0, 5, 7, and 10 cm of water, and took photos of the tank with the different water levels. First, photos were undistorted using the intrinsic camera parameters. Then, we highlighted the feature points by applying a threshold on the intensity values of the pixels. In the binary mask created from the thresholded images, we detected the centroids of the feature points by looking at the connected pixel regions. We ordered the detected centroids in a radial manner, going from the outer points toward the center of the tank. Using the rotation and translation matrices between the two cameras extracted from the calibration, we triangulated the world coordinates of the feature points.

3.2 Refraction correction

Following Suresh *et al.*,[25] we assumed that refraction had no effect on the in-plane ($X - Y$ plane) coordinates of the triangulated points, but on the depth estimation in the water. This claim is made evident in Figure 2, where we show the triangulated points for different water levels. For the same feature point, changing the water level

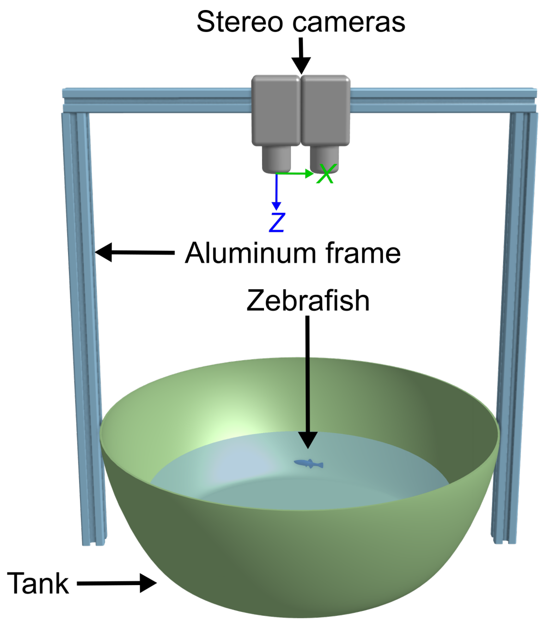


Figure 1: Computer assisted design of the experimental setup including the tank, zebrafish, a stereo camera, and the aluminum frame.

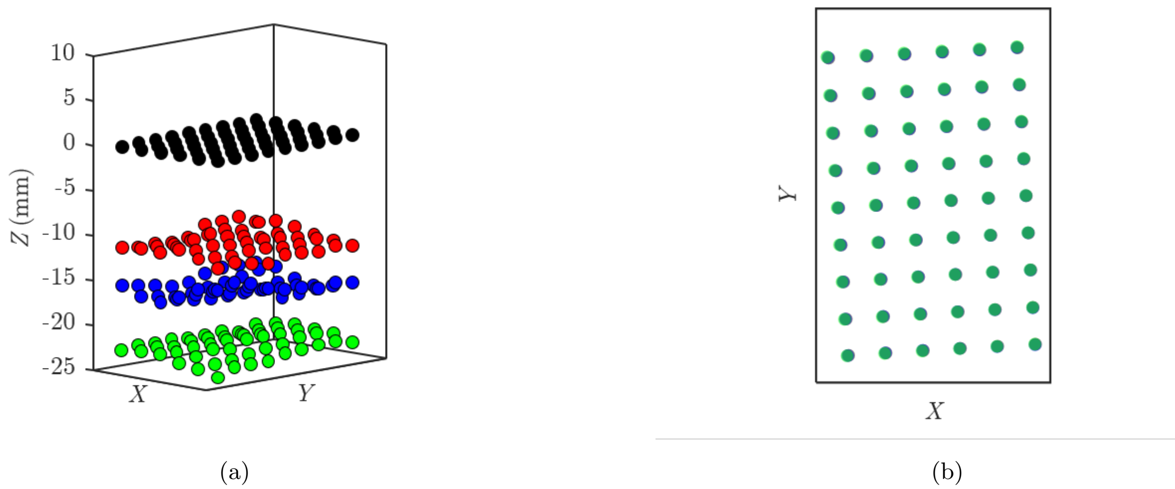


Figure 2: World coordinates of the triangulated feature points. (a) Feature points in three-dimensions; black points correspond to data extracted in an empty tank (constituting a baseline for all calculations) and red, blue and green points represent triangulated feature points at 5, 7, and 10 cm of water level, respectively. Variations in the value of Z for the same feature points across different water depths demonstrate the refractive effect on triangulation. (b) Feature points triangulated with varying water levels overlaid on each other.

has a dramatic effect on the value of the Z-coordinate (Figure 2a), while no change is recorded with respect to data on the $X-Y$ plane (Figure 2b).

Figure 3 illustrates our approach to incorporate the effect of refraction into the triangulation. Rays from the cameras to a point P (r_1 and r_2) were calculated by triangulating the pixel coordinates of the features in the images through the stereo camera extrinsics discarding light refraction. The camera height from the water surface (h) was measured for each water level, and the incidence angles to the water surface was calculated from h using Snell's law,[27] that is,

$$n_1 \sin(\theta) = n_2 \sin(\theta'), \tag{1}$$

where n_1 is the refractive index of air and n_2 is the refractive index of water (equal to 1.00 and 1.33, respectively).

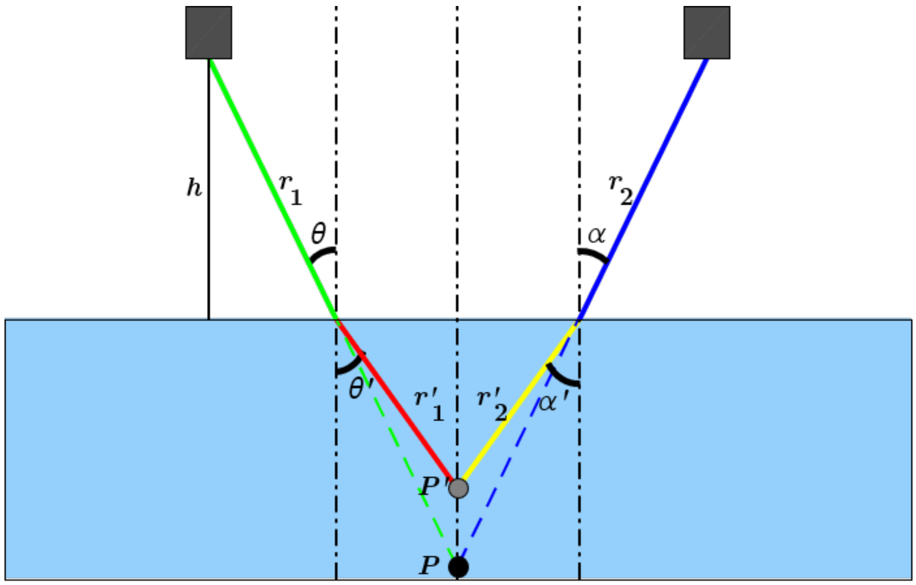
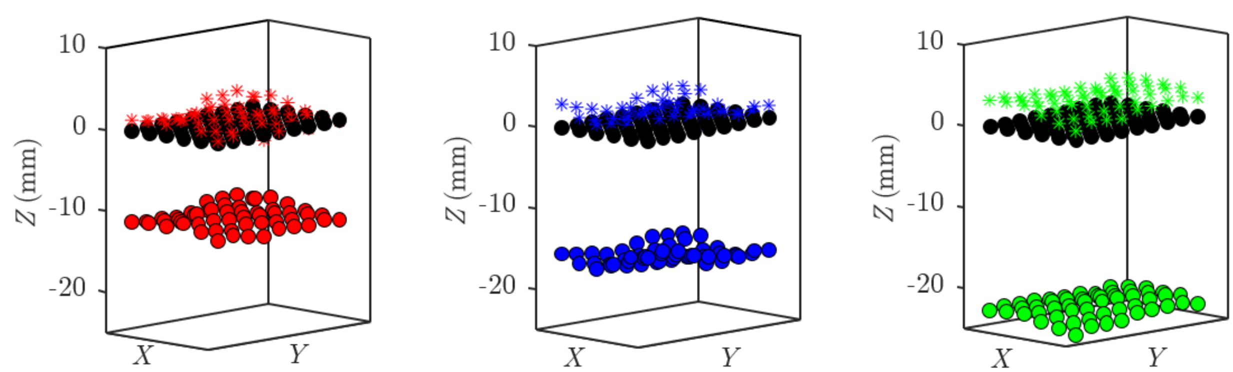


Figure 3: Three-dimensional triangulation algorithm with refraction effects. Dark grey boxes represent cameras, the black point P is the triangulated point, and the grey point P' is the actual point. h is the distance from the cameras and the water surface, r_1 and r_2 are the rays from cameras to point P before refraction at the water surface, r'_1 and r'_2 are the refracted rays to the actual point P', θ and α are the incidence angles, and θ' and α' are the refraction angles.

We simply calculated the depth of the point using the refraction angle and the in-plane distances between P and the intersection points of the rays with the water surface. We then averaged the two depths obtained from both cameras to estimate the depth of a feature point. Figure 4 shows the updated Z-axis position of the feature points upon applying the transformation. While the original points were scattered away from the baseline in the range $21.98 - 33.45$ mm, correcting for light refraction leads to a small scatter within $0.47 - 9.54$ mm, which is less than the typical size of an adult zebrafish that measures about 3 cm.



(a) (b) (c)

Figure 4: Feature point coordinates adjusted with the refraction algorithm, at water levels of (a) 5 cm, (b) 7 cm, and (c) 10 cm. Black points are the baseline coordinates. Colored points are triangulated coordinates without refraction taken into account, and colored asterisks are the updated coordinates with the refraction algorithm.

Water Level (mm)	Mean error			
	X-axis (mm)	Y-axis (mm)	Z-axis (mm)	Corrected depth (mm)
50	0.34	0.14	11.48	0.85
70	0.32	0.12	16.45	1.28
100	0.57	0.26	22.92	2.33

Table 1: Triangulation error in millimeters with relation to the baseline coordinates. One pixel corresponds to 1.5 mm.

3.3 Preliminary point cloud reconstruction of a zebrafish

To assess the proposed algorithms, we filmed a wild-type zebrafish. First, we filled the experimental tank with 7.5 cm of water at room temperature (26°C). The fish was gently hand netted from its housing tank to the experimental tank, and let habituate for one minute. Upon the end of a one-minute habituation period, we recorded videos of fish swimming using the stereo camera for two minutes. Then, the fish was hand netted back into its housing tank.

Towards obtaining the three-dimensional reconstruction of a fish, we populated a point cloud using a stereo camera. First, we obtained a background image, which was then subtracted from the frames including the fish, to reduce complexity in calculation. In standard practice,[28] point clouds are constructed using the disparity map between the two cameras, however refraction would cause inaccuracies in the disparity map calculation. Therefore, we used the algorithm described earlier to estimate the three-dimensional coordinates of the pixels associated with a fish. Once we populated the point cloud, we fitted an alpha shape over the points to visualize the fish.[29] We applied this technique to all the frames in the video to capture the fish body undulation while swimming.

4. EXPERIMENTAL RESULTS

Triangulation error across three-dimensions were calculated from the difference between the coordinates and the baseline (Table 1). Error along all three axes increased as the water level increased. Mean errors for the X-axis and Y-axis at different water levels were negligible, as they stayed under a millimeter (Table 1). This accuracy validates our initial assumption that the triangulated in-plane coordinates of feature points at different water levels would be the same. The mean errors in Z-axis were magnitude larger than the other axes. With the standard pinhole camera model, error values ranged between $10.24 - 12.62$ mm, $13.94 - 17.72$ mm, and $21.74 - 24.57$ mm for the water levels of 5, 7, and 10 cm respectively, and with the proposed algorithm their ranges were reduced to $0.00 - 2.34$ mm, $0.03 - 4.65$ mm, and $0.48 - 3.84$ mm. Calculating the depth with the refraction correction algorithm showed a considerable decrease in error. Mean error for depth calculation were reduced by 93%, 92%, and 90% for water levels of 5, 7, and 10 cm, respectively.

By identifying feature points, such as, eyes, fins, tail, and head, on a zebrafish, a sparse point cloud was successfully reconstructed using the refraction correction algorithm described. The sparse point cloud was fitted with an alpha shape to identify the body shape of a zebrafish (Fig. 5). Furthermore, the body length of the fish was identified as 25.83 mm, which is within range of body-lengths of a zebrafish.[10]

5. CONCLUSIONS

We elaborated on some of the limitations of existing tracking methods to assist in the complete analysis of the rich behavioral repertoire of zebrafish. In fact, most of the available methods to study zebrafish behavior are limited to point tracking in two- and three-dimensions. Few studies explored full-body reconstruction, but their success in the analysis of zebrafish behavior in authentic behavioral assays remain untested. Borrowing algorithms from robotics, we presented a potential alternative to estimate the three-dimensional position of features in zebrafish experiments.

The approach takes into consideration the presence of light refraction at the water surface, which challenges the use of a traditional pinhole camera model. Our results confirm the intuition that errors of the pinhole camera

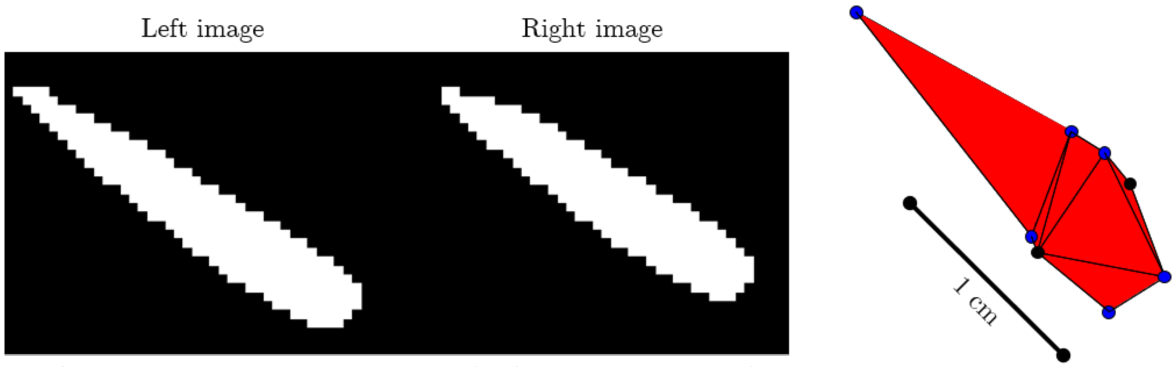


Figure 5: Sparse point cloud reconstruction of a fish, based on the refraction corrected triangulation. From left and right images feature points were selected for triangulation. These points are represented with blue dots on the three-dimensional model, and eyes are indicated with black dots. The red surface shows the alpha shape overlaid on top.

model increase with the water depth.[25,27] Correcting for the depth estimation offers a considerable improvement in the estimation at a limited computational cost. It is indeed tenable to assume that in-plane coordinates of a feature point would not change with increasing water depths, thereby limiting the correction for the presence of the water surface to only the vertical coordinate. Preliminary results on live zebrafish offer initial evidence in favor of the use of the proposed approach in the study of zebrafish behavior.

The accuracy in the estimation might be further improved by selecting feature points automatically, which would eliminate possible user error. With a better initial estimation, the refraction correction would perform with a reduced error in calculation. Also, the use of higher resolution cameras encased in a 3D-printed housing could provide densely populated point clouds, and allow for a better shape fit around the body of a fish. Pursuing further improvements on the three-dimensional body of the fish extracted from videos could allow for iris and gill tracking, thereby paving the way to elucidate the driving mechanisms in fish schooling.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation under Grant No. CMMI-1901697, by the National Institutes of Health, National Institute on Drug Abuse under grant number 1R21DA042558-01A1, and by the Office of Behavioral and Social Sciences Research that co-founded the National Institute on Drug Abuse grant. The authors are thankful to Vrishin Rajiv Soman for his help in preparing some of the experiments and conducting preliminary analysis.

REFERENCES

[1] Grunwald, D. J. and Eisen, J. S., "Headwaters of the zebrafish—emergence of a new model vertebrate," *Nature Reviews Genetics* **3**(9), 717–724 (2002).

[2] Rubinstein, A. L., "Zebrafish: from disease modeling to drug discovery," *Current Opinion in Drug Discovery and Development* **6**(2), 218–223 (2003).

[3] Vascotto, S. G., Beckham, Y., and Kelly, G. M., "The zebrafish's swim to fame as an experimental model in biology," *Biochemistry and Cell Biology* **75**(5), 479–485 (1997).

[4] Spence, R., Gerlach, G., Lawrence, C., and Smith, C., "The behaviour and ecology of the zebrafish, danio rerio," *Biological Reviews* **83**(1), 13–34 (2008).

[5] Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., Muffato, M., Collins, J. E., Humphray, S., McLaren, K., Matthews, L., et al., "The zebrafish reference genome sequence and its relationship to the human genome," *Nature* **496**(7446), 498–503 (2013).

[6] Broughton, R. E., Milam, J. E., and Roe, B. A., "The complete sequence of the zebrafish (danio rerio) mitochondrial genome and evolutionary patterns in vertebrate mitochondrial dna," *Genome R esearch* **11**(11), 1958–1967 (2001).

[7] Franco-Restrepo, J. E., Forero, D. A., and Vargas, R. A., "A review of freely available, open-source software for the automated analysis of the behavior of adult zebrafish," *Zebrafish* **16**(3), 223–232 (2019).

[8] Macrì, S., Clément, R. J., Spinello, C., and Porfiri, M., "Comparison between two-and three-dimensional scoring of zebrafish response to psychoactive drugs: identifying when three-dimensional analysis is needed," *PeerJ* **7**, e7893 (2019).

[9] Macrì, S., Neri, D., Ruberto, T., Mwaffo, V., Butail, S., and Porfiri, M., "Three-dimensional scoring of zebrafish behavior unveils biological phenomena hidden by two-dimensional analyses," *Scientific Reports* **7**(1), 1–10 (2017).

[10] Kalueff, A. V., Gebhardt, M., Stewart, A. M., Cachat, J. M., Brimmer, M., Chawla, J. S., Craddock, C., Kyzar, E. J., Roth, A., Landsman, S., et al., "Towards a comprehensive catalog of zebrafish behavior 1.0 and beyond," *Zebrafish* **10**(1), 70–86 (2013).

[11] Bollmann, J. H., "The zebrafish visual system: From circuits to behavior," (2019).

[12] Hedrick, T. L., "Software techniques for two-and three-dimensional kinematic measurements of biological and biomimetic systems," *Bioinspiration & Biomimetics* **3**(3), 034001 (2008).

[13] Pérez-Escudero, A., Vicente-Page, J., Hinz, R. C., Arganda, S., and De Polavieja, G. G., "idtracker: tracking individuals in a group by automatic identification of unmarked animals," *Nature Methods* **11**(7), 743 (2014).

[14] Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., and Bethge, M., "Deeplabcut: markerless pose estimation of user-defined body parts with deep learning," *Nature Neuroscience* **21**(9), 1281 (2018).

[15] Branson, K., Robie, A. A., Bender, J., Perona, P., and Dickinson, M. H., "High-throughput ethomics in large groups of drosophila," *Nature Methods* **6**(6), 451 (2009).

[16] Kim, C., Ruberto, T., Phamduy, P., and Porfiri, M., "Closed-loop control of zebrafish behaviour in three dimensions using a robotic stimulus," *Scientific Reports* **8**(1), 1–15 (2018).

[17] Al-Jubouri, Q., Al-Nuaimy, W., Al-Taee, M. A., and Young, I., "Computer stereovision system for 3d tracking of free-swimming zebrafish," in [*2017 10th International Conference on Developments in eSystems Engineering (DeSE)*], 188–192, IEEE (2017).

[18] Cachat, J., Stewart, A., Utterback, E., Hart, P., Gaikwad, S., Wong, K., Kyzar, E., Wu, N., and Kalueff, A. V., "Three-dimensional neurophenotyping of adult zebrafish behavior," *PloS One* **6**(3) (2011).

[19] Ladu, F., Butail, S., Macrí, S., and Porfiri, M., "Sociality modulates the effects of ethanol in zebra fish," *Alcoholism: Clinical and Experimental Research* **38**(7), 2096–2104 (2014).

[20] Macrì, S., Karakaya, M., Spinello, C., and Porfiri, M., "Zebrafish exhibit associative learning and memory of an aversive robotic stimulus.," *Lab Animal* ("Under consideration").

[21] Guo, Y., Veneman, W. J., Spaink, H. P., and Verbeek, F. J., "Three-dimensional reconstruction and measurements of zebrafish larvae from high-throughput axial-view in vivo imaging," *Biomedical Optics Express* **8**(5), 2611–2634 (2017).

[22] Lin, C.-f., Xu, L.-h., and Liu, Z.-c., "Digitization of free-swimming fish based on binocular stereo vision," in [*2015 8th International Symposium on Computational Intelligence and Design (ISCID)*], **2**, 363–368, IEEE (2015).

[23] Ichimaru, K., Furukawa, R., and Kawasaki, H., "Cnn based dense underwater 3d scene reconstruction by transfer learning using bubble database," in [*2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*], 1543–1552, IEEE (2019).

[24] Emran, F., Rihel, J., and Dowling, J. E., "A behavioral assay to measure responsiveness of zebrafish to changes in light intensities," *JoVE (Journal of Visualized Experiments)* (20), e923 (2008).

[25] Suresh, S., Westman, E., and Kaess, M., "Through-water stereo slam with refraction correction for auv localization," *IEEE Robotics and Automation Letters* **4**(2), 692–699 (2019).

[26] Zhang, Z., "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(11), 1330–1334 (2000).

[27] Tewinkel, G., "Water depths from aerial photographs," *Photogrammetric Engineering* **29**(6), 1037–1042 (1963).

[28] Hiep, V. H., Keriven, R., Labatut, P., and Pons, J.-P., "Towards high-resolution large-scale multi-view stereo," in [*2009 IEEE Conference on Computer Vision and Pattern Recognition*], 1430–1437, IEEE (2009).

[29] Linsen, L., "Point cloud representation," technical report, Faculty of Computer Science, University of Karlsruhe (2001).