




## In situ three-dimensional video tracking of tagged individuals within site-attached social groups of coral-reef fish

Anael Engel <sup>1</sup>\* Yaela Reuben,<sup>1,2</sup> Irina Kolesnikov,<sup>2</sup> Dmitri Churilov,<sup>2</sup> Ran Nathan <sup>1</sup>, Amatzia Genin <sup>1,2</sup>

<sup>1</sup>Ecology, Evolution and Behavior, Edmond Safra Campus, Hebrew University of Jerusalem, Jerusalem, Israel

<sup>2</sup>The Interuniversity Institute for Marine Sciences in Eilat and Department of Ecology, Evolution and Behavior, The Hebrew University of Jerusalem, Eilat, Israel

### Abstract

Tracking the movement of all individual group members in their natural environment remains a challenging task. Using advances in computer vision and Deep Learning, we developed and tested a semi-automated in situ tracking system to reconstruct simultaneous three-dimensional trajectories of marked individuals in social groups of a coral-reef fish. Our system has a temporal resolution of 10s of milliseconds, allowing for multiple 30-min tracking sessions that have been repeated over weeks to months. We present the technique and illustrate its application for *Dascyllus marginatus*, a planktivorous damselfish that lives in social groups associated with branching corals. Our technique identified all individuals 85–100% of the time, with a mean spatial error of ~ 1.3 cm. It provides a cost-effective semi-automated tool for in situ research on movements and foraging of individuals within small site-attached groups of animals in their natural environment.

Tracking individual animals provides the means to link their movement to features of their biotic and abiotic environment, fostering better understanding of their ecology and behavior (Nathan et al. 2008; Hussey et al. 2015). Existing methods to track animals vary greatly according to the research questions, the biology of the focal species, the properties of the environment and the available wildlife tracking technologies. Information on simultaneous movements of individuals within social groups of animals is fundamental to our understanding of the mechanisms underlying the structure, dynamics, and ecological functioning of social groups (Parrish and Hamner 1997; Couzin et al. 2002; Herbert-Read 2016). Tools to acquire such information have been developed mostly for animal groups in controlled experimental settings, ranging from small arenas and aquariums (Viscido and Parrish 2004) to large farms and pastures (Šárová et al. 2010). Such experimental systems are typically confined to cover an area that is much smaller than the home range of the animals in the wild, and are designed to control for the effects of various factors that likely play an important role in determining the movement and behavior of animals in their natural environments. Studies of animal groups under natural settings in the wild remain scarce, with only a few examples of GPS tracking of whole groups of domestic pigeons (Nagy et al. 2013) and most group members in primates (Strandburg-Peshkin et al. 2015). Nevertheless, application of GPS tracking

for high-resolution tracking of individual group members in aquatic ecosystems (seas, lakes, rivers) is infeasible due to the rapid attenuation of the GPS signals in the water.

Acoustic telemetry has been extensively used to track multiple individuals at relatively high temporal resolution and duration in the sea (Hussey et al. 2015; Crossin et al. 2017; Lowerre-Barbieri et al. 2019; Thomas et al. 2019). Few systems have reached sub-meter resolution in two and three dimensions (Cote et al. 1998; Cooke et al. 2005; Rillahan et al. 2009; Deng et al. 2011; Bohaboy et al. 2020), but have not yet reached the spatial precision of ~ 1 cm. Such a precision is needed to characterize the movements and behavior of animals that live in the same location for long periods. With the rapid advent of acoustic technologies and a dense array of multiple receivers this goal might be achieved (Deng et al. 2011). Nevertheless, data obtained with acoustics lacks important information that can be obtained by cameras. These include information on tracked individuals (e.g., feeding bites, antagonistic attacks), their dynamic environment, including disturbances and possible interactions with non-tagged conspecifics and predators. Clearly, at present optical techniques provide superior information on behavior and fine-scale movements of animals (Dell et al. 2014).

Multiple cameras have long been used to reconstruct the three-dimensional (3D) positions of both flying animals over land (Ballerini et al. 2008; Wu et al. 2009; Liu et al. 2016) and swimming animals in aquatic environments. Earlier studies of aquatic animals were by and large limited to aquaria (Cullen et al. 1965; Pitcher 1973; Partridge et al. 1980) and cages

\*Correspondence: anael.engel@mail.huji.ac.il

(Pitcher et al. 1985), and data were processed manually. Recent studies applied programming tools to automatically obtain 3D trajectories in aquaria and tanks (Viscido and Parrish 2004; Cachat et al. 2011; Wang et al. 2017; Ruberto et al. 2019). However, only a few studies successfully applied automated detection and tracking methods for fish in their natural environment (Jäger et al. 2017; Labao and Naval Jr 2019; Salman et al. 2019). To our knowledge, only a single study managed to automatically reconstruct fish 3D trajectories in situ (Francisco et al. 2020).

Problems hindering automated tracking using underwater video include variations in lightning conditions, unexpected visual perturbations, and repeated changes in the background (e.g., surface waves). Various approaches were applied to overcome those challenges, including the use of Gaussian mixture modeling (Salman et al. 2019) and convolutional neural networks (Jäger et al. 2017; Labao and Naval Jr 2019; Francisco et al. 2020). The latter technique is based on computational training that allows the detection of the target objects (e.g., a fish) in changing settings.

An important advantage of video tracking techniques in natural settings lies in its non-invasive approach alleviating the need to trap, tag or mark individuals which is time consuming and expensive, and can alter the behavior of tagged individuals and of other individuals that interact with them (Dell et al. 2014; Francisco et al. 2020). Notwithstanding these considerations, in the common cases where individuals are morphologically very similar one to another, an important prerequisite of repetitive in situ tracking of grouping individuals is artificial tagging. Long-term tracking of the identified individuals allows for assessing interactions among particular group members, their response to changes in their biotic and abiotic environment, and differences in behavioral consistency (behavioral types, or “personality”), social hierarchy and other key characteristics of each group member.

In this study, we developed an in situ technique for repetitive, high-resolution tracking of individual fish in social groups of site-attached coral-reef fish over periods of weeks to months. The limited foraging space of such fish enabled their uninterrupted presence in the cameras’ field of view, and tiny, seemingly non disruptive tags allowed the identification of the same individuals over long time.

Below, we describe our tracking system, from tagging through the data acquisition and its processing. We provide detailed assessment of the performance of the system, focusing on the following six criteria: tracking duration, data acquisition rate, spatial precision, automation, concurrency (simultaneous tracking of multiple individuals), and cost-effectiveness, as well as brief guidelines for best-practice application and an overview of the technique’s key limitations. The results we present are based on the raw data, without filtering, smoothing, or interpolating.

## Material and procedures

### Study fish

Our study focused on the site-attached group-forming damselfish *Dascyllus marginatus* (Pomacentridae (Rüppell 1829)). *D. marginatus* is a small (~ 6 cm in TL) pomacentrid, common throughout the Red Sea and Gulf of Oman (Fishelson et al. 1974; Randall 1986). It forms social groups ranging in size from 2 to 25 individuals, most groups have five individuals or less (Kent et al. 2006). It resides in live corals only, mostly in the branching corals *Stylophora* and *Acropora* (Kent et al. 2006).

### Study site

Our study was carried out at the fringing reef along the northwestern coast of the desert-enclosed Gulf of Aqaba (Eilat), Red Sea. This fringing reef is dominated by stony corals, that live on a steep slope extending from the subtidal zone to more than 50 m depth (Rickel and Genin 2005). The *D. marginatus* groups we studied were found in the shallow (9–14 m) reef off the Interuniversity Institute for Marine Sciences in Eilat, Israel.

### Fish tagging

To gently capture fish for tagging with minimal disturbance, the fish were first partly anesthetized by squirting clove oil (~ 0.03% nominal concentration) on the home coral where the fish hide upon arrival of divers. When a fish became dizzy, it was gently trapped with a small aquarium net. After trapping, each fish was placed for a few minutes in a transparent “zip-lock” bag (with small holes to allow water flow) and injected with black visible implant elastomer (VIE) dye (Northwest Marine Technology). The dye was implanted beneath the fish’s translucent scales, remaining visible for months. Each fish was tagged at a different location on its body, enabling individual identification. Immediately after tagging, the fish were returned to their home coral, providing a 1 d sheltered recovery under a large cage. The entire process, lasting approximately 60–90 min for groups of 3–5 individuals, took place in the vicinity of the home coral, without ever taking the fish to the laboratory. No changes in the fish behavior were apparent after tagging, as based on our long-term observations on the behavior of *D. marginatus* (e.g., Kiflawi and Genin 1997; Kent et al. 2006). The study was done under the permits 2014/40483 and 2017/41742 from the Israel Nature and Parks Authority according to the ethics procedures of animal treatment.

### Video recording

The 3D reconstruction of the positions of a fish required synchronized video recordings from at least 2 points of view. In the coral reef, as on land, target animals are sometime hidden from a stationary camera behind obstacles such as corals, stones or plants. Therefore, 3D tracking that requires

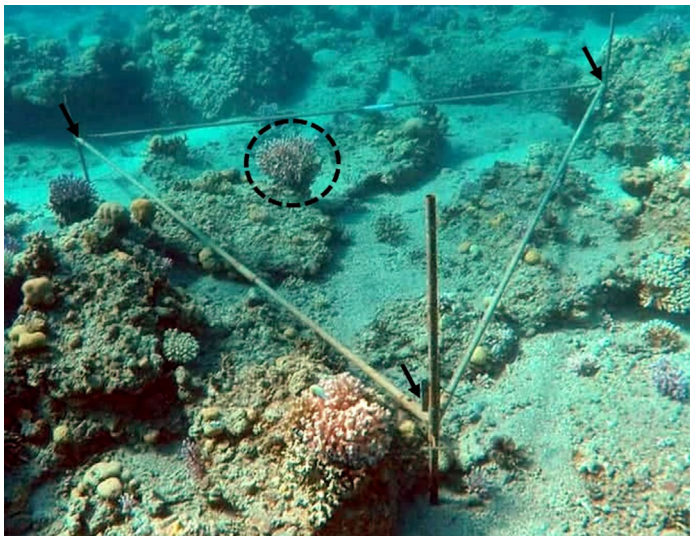
simultaneous views from at least 2 points required the use of  $\geq 3$  cameras (Francisco et al. 2020).

In this study, we used three cameras attached at three corners of a 3 m equilateral triangle frame positioned around the fish's home coral, hereafter "camera triangle" (Fig. 1). Thereby, each fish was seen by at least two cameras when it was outside its home coral. The cameras we used were GoPro (Hero 3+ Black,  $2704 \times 1524$  resolution, and Hero 5 Black  $2704 \times 1520$  resolution) set to acquire images at a rate of 29.97 frames per second (fps). The three cameras were synchronized every  $\sim 10$  min to within one frame using a sharp acoustic cue, generated by a diver hammering a metal cylinder. The need for this repetitive synchronization was an occasional loss of the cameras' synchronization over longer period. Synchronized frame numbers in each video record (i.e., the frame at which the acoustic cue started) were determined using VirtualDub (VirtualDub 1.10.4, Avery Lee, <http://virtualdub.org>).

### Camera calibration

The use of the three cameras for computations of fish 3D positions required extensive calibration, as follows:

1. The intrinsic parameters of each camera were determined using a checkerboard with  $8 \times 8$  squares 37 mm each, recorded in situ from different angles, and processed using Camera Calibration Toolbox for MATLAB<sup>®</sup> (Bouguet 2010).
2. The extrinsic parameters were determined for the camera triangle at the end of each trial. Here we used the open program easyWand5, which implements the sparse bundle adjustment (SBA) calibration algorithm (Theriault et al. 2014). We used manually digitized wand points, each digitized in the three



**Fig 1.** The underwater experimental setup. Three-dimensional underwater camera system consisting of three high-resolution GoPro cameras that were tightly attached at the three corners of the 3 m equilateral triangle frame (black arrows), positioned around the home coral of the fish (dashed black circle), in the coral reef of Eilat.

cameras using DLTdv5 (Hedrick 2008). For that, we marked two 1 cm conspicuous points 10 cm apart on a transparent wand attached to a metal stick that was repeatedly moved by a diver across the volume recorded by the cameras (Fig. 2). The distortions of the camera lenses were corrected using easyWand5 based on the distortion model extracted from the checkerboard calibration.

3. To determine absolute compass directions in the records, the calibrated space was aligned to compass metal axes placed at the end of the trial inside the "camera triangle" at a point that was visible by the three cameras.

### Detection of fish in the videos

To detect each fish in each camera we used convolutional neural networks. This procedure required training of the network to detect *D. marginatus* in the images. Our training used Google's TensorFlow Deep Learning platform (Abadi et al. 2016) and the faster-RCNN (Ren et al. 2017) inception V2 model (Huang et al. 2017), pre-trained on the COCO (Common Objects in Context) dataset (Lin et al. 2014) that contains more than 200,000 images in 80 object categories.

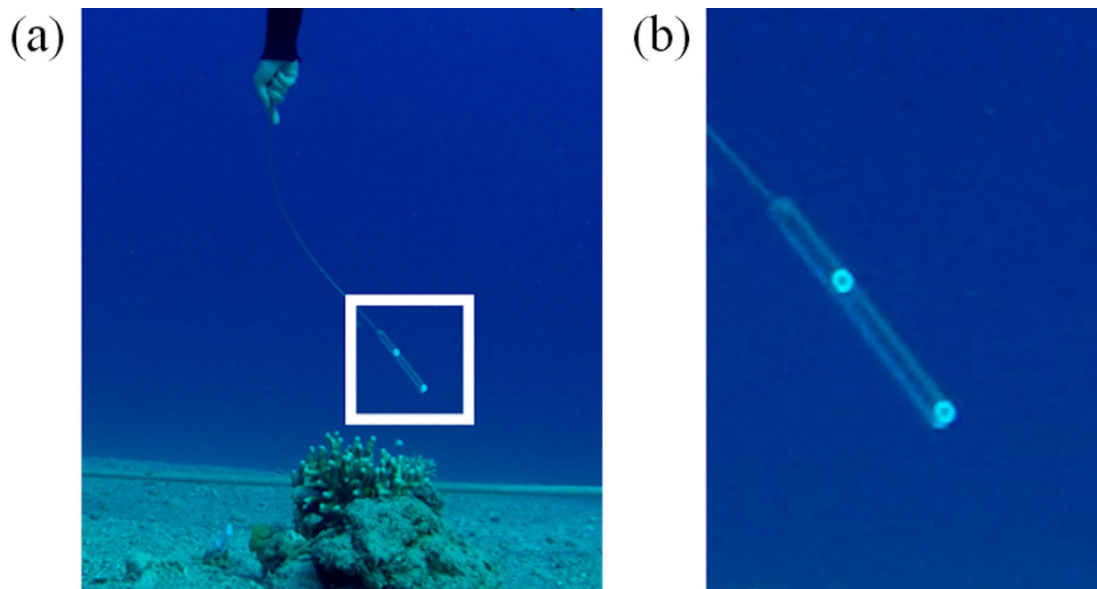
We used transfer learning to train the model on  $> 3500$  images with  $> 11,000$  manually identified *D. marginatus*, obtained during seven different days at three different locations in the coral reef. The images covered different conditions of lighting, backgrounds, and fish orientations, including cases when the fish were partially masking one another. After training, the model was used to detect *D. marginatus* in all the frames of our video records (Fig. 3). Model training and data processing were performed using Python 3.7.1, Tensorflow 1.10.0 on a computer equipped with NVIDIA GeForce GTX 1080 Ti graphical processing units (GPUs).

### Two-dimensional tracking

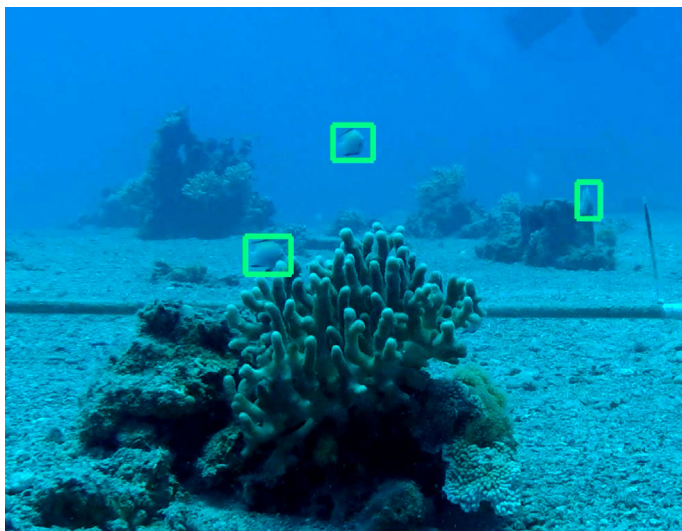
After detecting a fish in each camera and each frame, two-dimensional (2D) tracks were reconstructed by automatically linking close detections of the same fish in consecutive frames. Ambiguities such as occlusions or fish entering and exiting the branching coral were resolved manually using a custom code written in MATLAB<sup>®</sup> that enabled the user to add or remove detections from the trajectories and to fix detection errors, based on visual identification of each marked fish. This process was performed for each video separately and manually-resolved segments accounted for less than 2% of the 2D trajectories.

### Calculating 3D trajectories

To reconstruct the 3D tracks, corresponding tracks (of the same individual) had to be identified in at least two cameras. Manual matching of tracks of the same fish recorded by different cameras was possible in our study since the groups were relatively small and each individual was differentially tagged. The calculations of the 3D trajectories were based on the



**Fig 2.** The calibration object. (a) An image of the calibration wand (indicated with the white square) during the calibration. (b) A close-up view of the calibration wand, showing the two marks 10 cm apart (white circles with an inner black dot).



**Fig 3.** An example of a video frame, with the detections of three fish (marked with green squares). This image is cropped to improve visualization.

direct linear transformation (DLT) technique as implemented in DLTdv5 (Hedrick 2008).

### Assessment

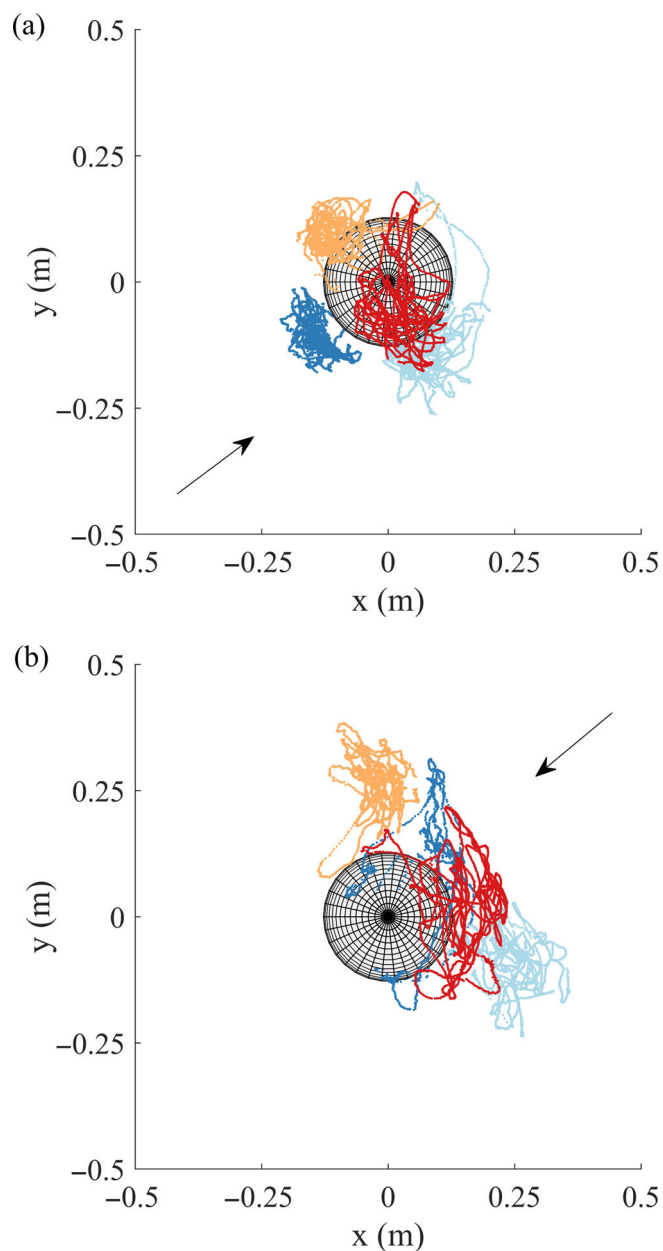
We used the system to track four groups of 3–5 fish during the years 2015–2017, recording a total of ~20 h of data. We illustrate application of the method by two examples: 3D tracks of four *D. marginatus* in a group occupying the branching coral *Stylophora pistillata* in two opposing current

directions (Fig. 4), and inter-individual variation in trajectories toward a prey (Fig. 5). Each recording dive lasted about 1 h, including mounting of the cameras on the tripod, recording of the fish, calibrating, and dismounting the cameras. Net tracking time per dive (hereafter “session”) lasted ~30 min each.

Assuming that the fish are visible by at least two video cameras when out of their coral shelter, the data acquisition rate depended on the frame rate of the video cameras and the system’s actual 3D detection rate. We define the system’s 3D detection rate as the percentage of 3D reconstructed points of the total expected points recorded by the GoPro cameras at 29.97 fps. The 3D detection rate was calculated by processing three randomly chosen datasets of 3.5-min long segments of video records from 3 different days of two groups of *D. marginatus* (with three and four fish) during which all fish were visible and foraged outside the coral. A 3D detection rate of 85–100% was found, with a mean of 97% (Table 1). Considering the cameras’ frame rate, our system’s average 3D data acquisition rate was 29 Hz.

The calibration error was calculated for each recording day based on  $\geq 100$  records of the wand that were not included in those used for the calibration. The mean error was always  $< 2.25\%$ . To assess the spatial error of the reconstructed 3D trajectories, we compared the 3D position of specific points along the automatically calculated tracks to their manually determined 3D position. Since the manual procedure was time-consuming and labor-intensive, we applied this test to small subsets of the full trajectories. Specifically, 100 points, 50 frames apart, were manually determined for each fish, serving as our ground truth. The mean Euclidian distance between

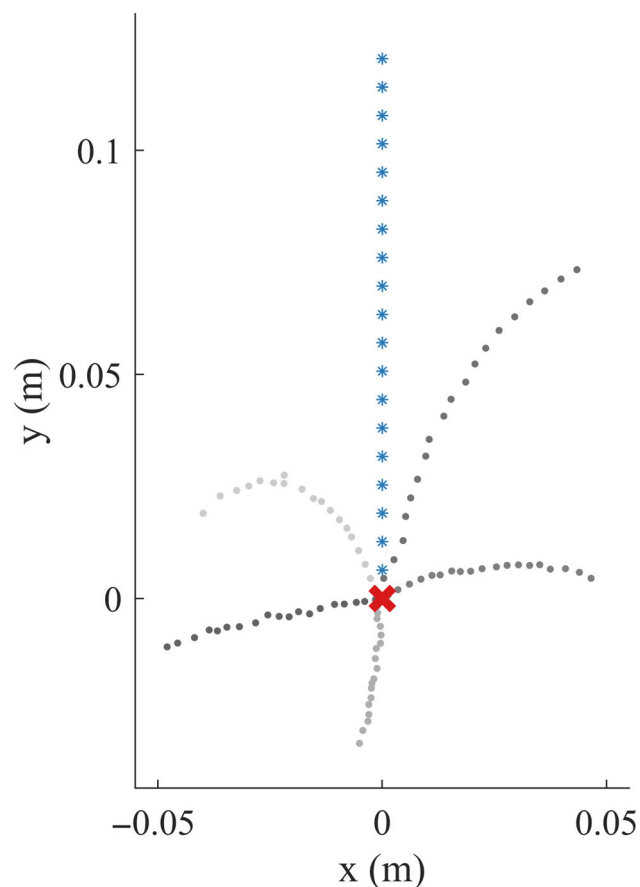




**Fig 4.** Top views of 3D reconstruction of 2.5 min of trajectories of four *Dascyllus marginatus* (different colors) recorded at 11 m depth in the coral reef of Eilat. **(a)** Tracks recorded on the 9th of March 2015 during NE currents (black arrow). **(b)** Tracks recorded on 29th of April 2015 during SW currents. Gray spheres indicate the home coral. Note the spatial separation between individual fish within the group.

the automatically-calculated and ground-truth positions ranged 0.6–2.0 cm across fish and datasets (Table 1), with a grand-mean of 1.3 cm and maximum of 4.7 cm, less than the fish's body length (~ 6 cm).

The method described in this paper was relatively inexpensive. Aside from the cost of personnel and field operations (e.g., scuba diving), the system required funds to cover the cost of three cameras (a few hundred US\$), the VIE dye (~ 100



**Fig 5.** Top view of five feeding events (dots with different shades of gray) by the same fish during a single session. Dots indicate the consecutive positions of the fish during the strikes. Blue asterisks indicate the expected positions of the planktonic prey assuming it passively drifted with the flow, which was concurrently measured at the tracking site using a current meter ( $19 \text{ cm s}^{-1}$ ;  $219^\circ$  heading). Positions in each track were based on 20 consecutive video frames recorded at 29.97 fps. The red cross indicates the position at which the fish extended its jaw to capture (suck) the prey. The figure's vertical axis was rotated to align with the current direction and all positions were shifted so that the capture point will be positioned at the origin.

US\$ for tagging ~ 100 fish), and the cost of using a computer for training a neural network.

## Discussion

We present a method to reconstruct high-resolution 3D trajectories of individual coral-reef fish in social groups. In terms of tracking duration, our system provided uninterrupted data on the movement of the same individuals during ~ 30-min that could be repeated over weeks to months, perhaps even years. During this project (2015–2017), we focused on four *D. marginatus* groups that were repeatedly recorded during 2–3 months, with 7–16 sessions per group. Thereby, our recording sessions covered a wide range of environmental conditions, including different

**Table 1.** Performance of our system. The 3D detection rate is the percentage of 3D reconstructed points in a segment of 3.5-min (6295 frames). The values of the distance to the ground truth are computed on 100 points. Values for each fish are given.

Dataset	# fish	3D detection rate (%)	Distance in 3D between ground truth and automatically calculated data		
			Mean (cm)	SD (cm)	Max (cm)
1	3	99, 100, 97	2.0, 1.8, 1.6	0.6, 0.3, 0.4	3.7, 2.6, 2.7
2	3	98, 96, 100	1.8, 1.7, 1.8	0.5, 0.3, 0.3	4.7, 2.8, 2.4
3	4	99, 100, 100, 85	0.6, 0.6, 0.7, 0.6	0.2, 0.2, 0.2, 0.2	1.2, 1.1, 1.3, 1.6
Mean values		97	1.3	0.3	2.4

current directions and speeds. It revealed consistent spatial partitioning of the foraging spaces among group members (Engel et al. unpubl.).

In terms of acquisition rate, our video records were acquired using 29.97 fps, with an average 3D detection rate of 97%, thereby matching the present state-of-the-art techniques that are used in the laboratory (Sridhar et al. 2019). To make our method comparable, system performance was limited to times when all fish were swimming outside of the coral.

Regarding spatial precision, the mean positioning error we had was 1.3 cm. For a small fish such as the one we used (~ 6 cm in length) and for the main objective of our study—examine foraging movements and an occurrence of space partitioning among the group members, such a precision sufficed. Studies where better precision is needed should use cameras with higher resolution, with smaller calibration points, and based the automated tracking on a certain point on the fish's body (e.g., eye, snout), rather than the position of the approximate center of the fish's body.

An important feature of our technique was the automation. The 2D track reconstruction was nearly (~ 95%) automated, requiring minor manual corrections accounting for occlusions among neighboring individuals and small errors. Track matching before calculating the 3D trajectories was done manually. Several automated solutions were suggested for solving occlusions (Perez-Escudero et al. 2014; Fukunaga et al. 2015; Xu and Cheng 2017) and also for track matching (Attanasi et al. 2014; Qian and Chen 2017). The recent version of Hedrick's DLTdv8 (Hedrick 2008) might also resolve this problem and other new tools are expected in this fast-growing field. Nevertheless, as the groups of fish we studied were small (3–5 individuals), both cases of occlusions and track matching could be easily resolved manually. Our calibration of the cameras using the wand was manual. This calibration can be made automated if the right setup is programmed. Our automated detection of the fish in 2D used a trained neural network that required the building of a dataset of manually tagged images showing the fish from different points of view and under different postures. Such a dataset should be prepared at the outset of the study and can be used for repetitive sessions where similar spatial settings are used. Here we used the same trained network

to track different groups of *D. marginatus* inhabiting branching corals in different locations within the coral reef of Eilat.

Our tracking method was used with groups of 3–5 fish, which was ecologically relevant as most of the *D. marginatus* groups at our study site had ≤ 5 individuals (Kent et al. 2006). Larger groups and more aggregated distribution within groups are likely to increase the number of occlusions, rendering a need for more manual work. Clearly, such cases should greatly benefit from automated resolution of occlusions. Another limiting factor is the position of the tagging marks. We used four different tagging locations: head, upper dorsal spot, above the tail and below it. Fish were symmetrically marked on both sides. The use of colored dye proved unsuccessful as colors were hard to identify in the video records. Therefore, we used black dye exclusively throughout this study. Multiple marks should increase the number of individually tagged fish, but incur higher risk of harming the fish. Therefore, in our study we chose to mark the fish in no more than two positions per side (e.g., head + above tail). All the fish we tagged survived, none showing any sign of stress throughout the extended period of our study.

Finally, in terms of cost-effectiveness, our system was inexpensive. The GoPro cameras as well as the dye to mark the fish were off-the-shelf products, and the metal frame required simple welding.

The method presented here was tailored for tracking site-attached fishes, where group members remain at all times within a limited distance (~ 1 m) from their home coral. The fish we studied, *D. marginatus*, was similar to many other species belonging to the ubiquitous guild of planktivorous fish that forage for prey in the proximity of stationary shelters such as branching corals and structurally-complex rocks (Wilson et al. 2008; Coker et al. 2014). For mobile groups and those occupying larger spaces, a system consisting of series of stationary cameras or a mobile system with stereo cameras that follows the group as it moves can be used (e.g., Francisco et al. 2020).

The system is currently used to investigate patterns of spatial segregation and mechanisms of zooplankton hunting in *D. marginatus* groups at the Red Sea (Engel et al. unpubl.). The system provides valuable data to explore a variety of questions

on the ecology of coral-reef fish that remain mostly unknown since the works of Forrester (1991) and Webster and Hixon (2000). These include the relationships between the position of a fish and its social rank in the group, and an assessment of the quality of different positions (e.g., whether holding an upstream position results in higher feeding rate). Furthermore, data from this system can complement laboratory work in studies such as those investigating individual movement differences among group members (Herbert-Read et al. 2012; Jolles et al. 2017) and escape responses (Shashar et al. 2005). Our method could also be used on in situ studies of the spatial organization among groups consisting multiple species (Shpigel 1982).

### Comments and recommendations

By applying emerging tools such as convolutional neural networks for detecting individuals, our semi-automated video tracking system portrays effective implementation of advance computational techniques for ecology and behavioral science. The ability to simultaneously track individual fish in their natural environment using a fine ( $\sim 1$  cm) spatial scale and high temporal (milliseconds) resolution that can be repeated over days and months, opens a new horizon for ecological and behavioral research of foraging, territoriality, social interactions, and behavioral repertoires of the individual (“personality”).

Further improvements in system performance can be obtained by following several technical lessons we gained during this project. First, to put all the reconstructed data on the same exact coordinate system, it is useful to reconstruct several fixed points on the coral or other environmental structures and to align the coordinate system based on these points. Second, the inner clock of the GoPro’s cameras proved unreliable, resetting at unexpected times. We therefore recommend to independently document by one of the cameras the exact time using an underwater digital clock at the beginning or end of each session. This practice could be particularly useful for synchronizing data obtained with auxiliary sensors (e.g., a current meter). Finally, in most recording sessions, the fish occupied only a certain area in the video frame (only one side of the coral, for example). To improve the detection process, we first cropped the video frames to different sections, and later converted the detected fish coordinates back to the coordinates of non-cropped original video frame.

On a broader perspective, we note that computer vision and deep learning tools are advancing fast, hence envision wider use and further improvements of this and similar systems. The simple inexpensive cameras we used can be replaced by available higher-performance cameras for studying smaller fish or to address questions requiring higher-resolution data to track, for example, specific body parts (Mathis et al. 2018). Furthermore, while in our current application session length is limited to  $\sim 30$  min (determined by the camera’s battery and air consumption of the scuba divers), full representation of

environmental conditions and temporal dynamics (hours to seasons and even years) can be accomplished by using cabled cameras connected to an on-shore power supply and a data logger. This, however, will generate enormous amounts of data that will render the manual parts of data processing and data analysis overly laborious, motivating the development of automated detection and big-data analysis tools, which necessitates tight collaborative work between biological and computer sciences.

### References

- Abadi, M., and others. 2016. Tensorflow: A system for large-scale machine learning, p. 265–283. In 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16). USENIX Association. Available from: <https://dl.acm.org/doi/10.5555/3026877.3026899>
- Attanasi, A., and others. 2014. Information transfer and behavioural inertia in starling flocks. *Nat. Phys.* **10**: 691–696. doi:[10.1038/nphys3035](https://doi.org/10.1038/nphys3035)
- Ballerini, M., and others. 2008. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proc. Natl. Acad. Sci. USA* **105**: 1232–1237. doi:[10.1073/pnas.0711437105](https://doi.org/10.1073/pnas.0711437105)
- Bobaboy, E. C., T. L. Guttridge, N. Hammerschlag, M. P. M. Van Zinnicq Bergmann, and W. F. Patterson III. 2020. Application of three-dimensional acoustic telemetry to assess the effects of rapid recompression on reef fish discard mortality. *ICES J. Mar. Sci.* **77**: 83–96. doi:[10.1093/icesjms/fsz202](https://doi.org/10.1093/icesjms/fsz202)
- Bouguet, J.-Y. 2010. Camera calibration toolbox for MATLAB. Available from: [http://www.vision.caltech.edu/bouguetj/calib\\_doc/index.html](http://www.vision.caltech.edu/bouguetj/calib_doc/index.html)
- Cachat, J., and others. 2011. Three-dimensional neurophenotyping of adult zebrafish behavior. *PLoS One* **6**: e17597. doi:[10.1371/journal.pone.0017597](https://doi.org/10.1371/journal.pone.0017597)
- Coker, D. J., S. K. Wilson, and M. S. Pratchett. 2014. Importance of live coral habitat for reef fishes. *Rev. Fish Biol. Fish.* **24**: 89–126. doi:[10.1007/s11160-013-9319-5](https://doi.org/10.1007/s11160-013-9319-5)
- Cooke, S. J., G. H. Niezgod, K. C. Hanson, C. D. Suski, F. J. S. Phelan, R. Tinline, and D. P. Philipp. 2005. Use of CDMA acoustic telemetry to document 3-D positions of fish: Relevance to the design and monitoring of aquatic protected areas. *Mar. Technol. Soc. J.* **39**: 31–41. doi:[10.4031/002533205787521659](https://doi.org/10.4031/002533205787521659)
- Cote, D., and others. 1998. A coded acoustic telemetry system for high precision monitoring of fish location and movement: Application to the study of nearshore nursery habitat of juvenile Atlantic cod (*Gadus morhua*). *Mar. Technol. Soc. Mar. Technol. Soc. J.* **32**: 54. Available from: <https://www.proquest.com/scholarly-journals/coded-acoustic-telemetry-system-high-precision/docview/211720226/se-2?accountid=14546>

- Couzin, I. D., J. Krause, R. James, G. D. Ruxton, and N. R. Franks. 2002. Collective memory and spatial sorting in animal groups. *J. Theor. Biol.* **218**: 1–11. doi:[10.1006/yjtbi.2002.3065](https://doi.org/10.1006/yjtbi.2002.3065)
- Crossin, G. T., M. R. Heupel, C. M. Holbrook, N. E. Hussey, S. K. Lowerre-Barbieri, V. M. Nguyen, G. D. Raby, and S. J. Cooke. 2017. Acoustic telemetry and fisheries management. *Ecol. Appl.* **27**: 1031–1049. doi:[10.1002/eap.1533](https://doi.org/10.1002/eap.1533)
- Cullen, J. M., E. Shaw, and H. A. Baldwin. 1965. Methods for measuring the three-dimensional structure of fish schools. *Anim. Behav.* **13**: 534–543. doi:[10.1016/0003-3472\(65\)90117-X](https://doi.org/10.1016/0003-3472(65)90117-X)
- Dell, A. I., and others. 2014. Automated image-based tracking and its application in ecology. *Trends Ecol. Evol.* **29**: 417–428. doi:[10.1016/j.tree.2014.05.004](https://doi.org/10.1016/j.tree.2014.05.004)
- Deng, Z. D., M. A. Weiland, T. Fu, T. A. Seim, B. L. LaMarche, E. Y. Choi, T. J. Carlson, and M. B. Eppard. 2011. A cabled acoustic telemetry system for detecting and tracking juvenile salmon: Part 2. Three-dimensional tracking and passage outcomes. *Sensors* **11**: 5661–5676. doi:[10.1002/eap.1533](https://doi.org/10.1002/eap.1533)
- Fishelson, L., D. Popper, and A. Avidor. 1974. Biosociology and ecology of pomacentrid fishes around the Sinai peninsula (northern Red Sea). *J. Fish Biol.* **6**: 119–133. doi:[10.1111/j.1095-8649.1974.tb04532.x](https://doi.org/10.1111/j.1095-8649.1974.tb04532.x)
- Forrester, G. E. 1991. Social rank, individual size and group composition as determinants of food consumption by humbug damselfish, *Dascyllus aruanus*. *Anim. Behav.* **42**: 701–711. doi:[10.1016/S0003-3472\(05\)80116-2](https://doi.org/10.1016/S0003-3472(05)80116-2)
- Francisco, F. A., P. Nührenberg, and A. Jordan. 2020. High-resolution, non-invasive animal tracking and reconstruction of local environment in aquatic ecosystems. *Mov. Ecol.* **8**: 27. doi:[10.1186/s40462-020-00214-w](https://doi.org/10.1186/s40462-020-00214-w)
- Fukunaga, T., S. Kubota, S. Oda, and W. Iwasaki. 2015. GroupTracker: Video tracking system for multiple animals under severe occlusion. *Comput. Biol. Chem.* **57**: 39–45. doi:[10.1016/j.compbiolchem.2015.02.006](https://doi.org/10.1016/j.compbiolchem.2015.02.006)
- Hedrick, T. L. 2008. Software techniques for two- and three-dimensional kinematic measurements of biological and biomimetic systems. *Bioinspir. Biomim.* **3**: 34001. doi:[10.1088/1748-3182/3/3/034001](https://doi.org/10.1088/1748-3182/3/3/034001) Available from: <https://biomech.web.unc.edu/dltdv/> [DLTdv5 assessed 30 April 2015, DLTdv8 assessed 4 February 2021]
- Herbert-Read, J. E. 2016. Understanding how animal groups achieve coordinated movement. *J. Exp. Biol.* **219**: 2971–2983. doi:[10.1242/jeb.129411](https://doi.org/10.1242/jeb.129411)
- Herbert-Read, J. E., S. Krause, L. J. Morrell, T. M. Schaefer, J. Krause, and A. J. W. Ward. 2012. The role of individuality in collective group movement. *Proc. R. Soc. B Biol. Sci.* **280**. doi:[10.1098/rspb.2012.2564](https://doi.org/10.1098/rspb.2012.2564)
- Huang, J., and others. 2017. Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors, p. 3296–3297. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE. doi:[10.1109/CVPR.2017.351](https://doi.org/10.1109/CVPR.2017.351)
- Hussey, N. E., and others. 2015. Aquatic animal telemetry: A panoramic window into the underwater world. *Science* **348**: 1255642. doi:[10.1126/science.1255642](https://doi.org/10.1126/science.1255642)
- Jäger, J., V. Wolff, K. Fricke-Neudert, O. Mothes, and J. Denzler. 2017. Visual fish tracking: Combining a two-stage graph approach with CNN-features, p. 1–6. In *Ocean. 2017-Aberdeen*. IEEE. doi:[10.1109/OCEANSE.2017.8084691](https://doi.org/10.1109/OCEANSE.2017.8084691)
- Jolles, J. W., N. J. Boogert, V. H. Sridhar, I. D. Couzin, and A. Manica. 2017. Consistent individual differences drive collective behavior and group functioning of schooling fish. *Curr. Biol.* **27**: 2862–2868.e7. doi:<https://doi.org/10.1016/j.cub.2017.08.004>
- Kent, R., R. Holzman, and A. Genin. 2006. Preliminary evidence on group-size dependent feeding success in the damselfish *Dascyllus marginatus*. *Mar. Ecol. Prog. Ser.* **323**: 299–303. doi:[10.3354/meps323299](https://doi.org/10.3354/meps323299)
- Kiflawi, M., and A. Genin. 1997. Prey flux manipulation and the feeding rates of reef-dwelling planktivorous fish. *Ecology* **78**: 1062–1077. doi:[10.2307/2265858](https://doi.org/10.2307/2265858)
- Labao, A. B., and P. C. Naval Jr. 2019. Cascaded deep network systems with linked ensemble components for underwater fish detection in the wild. *Ecol. Inform.* **52**: 103–121. doi:[10.1016/j.ecoinf.2019.05.004](https://doi.org/10.1016/j.ecoinf.2019.05.004)
- Lin, T. Y., M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. 2014. Microsoft COCO: Common objects in context, p. 740–755. In *European Conference on Computer Vision*. Cham: Springer. doi:[10.1007/978-3-319-10602-1\\_48](https://doi.org/10.1007/978-3-319-10602-1_48)
- Liu, Y., S. Wang, and Y. Q. Chen. 2016. Automatic 3D tracking system for large swarm of moving objects. *Pattern Recognit.* **52**: 384–396. doi:<https://doi.org/10.1016/j.patcog.2015.11.014>
- Lowerre-Barbieri, S. K., R. Kays, J. T. Thorson, and M. Wikelski. 2019. The ocean's movescape: Fisheries management in the bio-logging decade (2018–2028). *ICES J. Mar. Sci.* **76**: 477–488. doi:[10.1093/icesjms/fsy211](https://doi.org/10.1093/icesjms/fsy211)
- Mathis, A., P. Mamidanna, K. M. Cury, T. Abe, V. N. Murthy, M. W. Mathis, and M. Bethge. 2018. DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**: 1281–1289. doi:[10.1038/s41593-018-0209-y](https://doi.org/10.1038/s41593-018-0209-y)
- Nagy, M., G. Vásárhelyi, B. Pettit, I. Roberts-Mariani, T. Vicsek, and D. Biro. 2013. Context-dependent hierarchies in pigeons. *Proc. Natl. Acad. Sci.* **110**: 13049–13054. doi:[10.1073/pnas.1305552110](https://doi.org/10.1073/pnas.1305552110)
- Nathan, R., W. M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, and P. E. Smouse. 2008. A movement ecology paradigm for unifying organismal movement research. *Proc. Natl. Acad. Sci. USA* **105**: 19052–19059. doi:[10.1073/pnas.0800375105](https://doi.org/10.1073/pnas.0800375105)



- Parrish, J. K., and W. M. Hamner. 1997. Animal groups in three dimensions: How species aggregate. Cambridge University Press
- Partridge, B. L., T. Pitcher, M. Cullen, and J. Wilson. 1980. The three-dimensional structure of fish schools. *Behav. Ecol. Sociobiol.* **6**: 277–288. doi:[10.1007/BF00292770](https://doi.org/10.1007/BF00292770)
- Perez-Escudero, A., J. Vicente-Page, R. C. Hinz, S. Arganda, and G. G. de Polavieja. 2014. idTracker: Tracking individuals in a group by automatic identification of unmarked animals. *Nat. Methods* **11**: 743–748. doi:[10.1038/nmeth.2994](https://doi.org/10.1038/nmeth.2994)
- Pitcher, T. J. 1973. The three-dimensional structure of schools in the minnow, *Phoxinus phoxinus* (L.). *Anim. Behav.* **21**: 673–686. doi:[10.1016/S0003-3472\(73\)80091-0](https://doi.org/10.1016/S0003-3472(73)80091-0)
- Pitcher, T. J., A. E. Magurran, and J. I. Edwards. 1985. Schooling mackerel and herring choose neighbours of similar size. *Mar. Biol.* **86**: 319–322. doi:[10.1007/BF00397518](https://doi.org/10.1007/BF00397518)
- Qian, Z.-M., and Y. Q. Chen. 2017. Feature point based 3D tracking of multiple fish from multi-view images. *PLoS One* **12**: e0180254. doi:[10.1371/journal.pone.0180254](https://doi.org/10.1371/journal.pone.0180254)
- Randall, J. E. 1986. Red Sea reef fishes. Immel
- Ren, S., K. He, R. Girshick, and J. Sun. 2017. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**: 1137–1149. doi:[10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031)
- Rickel S., and A. Genin. 2005. Twilight transitions in coral reef fish: the input of light-induced changes in foraging behaviour. *Anim. Behav.* **70**: 133–144. doi:[10.1016/j.anbehav.2004.10.014](https://doi.org/10.1016/j.anbehav.2004.10.014)
- Rillahan, C., M. Chambers, W. H. Howell, and W. H. Watson. 2009. A self-contained system for observing and quantifying the behavior of Atlantic cod, *Gadus morhua*, in an offshore aquaculture cage. *Aquaculture* **293**: 49–56. doi:[10.1016/j.aquaculture.2009.04.003](https://doi.org/10.1016/j.aquaculture.2009.04.003)
- Ruberto, T., M. Violet, S. Sukhewanpreet, N. Daniele, and P. Maurizio. 2019. Zebrafish response to a robotic replica in three dimensions. *R. Soc. Open Sci.* **3**: 160505. doi:[10.1098/rsos.160505](https://doi.org/10.1098/rsos.160505)
- Rüppell, E. 1829. Reisen in Nubien, Kordofan und dem peträischen Arabien: vorzüglich in geographische-statistische Hinsicht. Wilmans
- Salman, A., S. Maqbool, A. H. Khan, A. Jalal, and F. Shafait. 2019. Real-time fish detection in complex backgrounds using probabilistic background modelling. *Ecol. Inform.* **51**: 44–51. doi:[10.1016/j.ecoinf.2019.02.011](https://doi.org/10.1016/j.ecoinf.2019.02.011)
- Šárová, R., M. Špinková, J. L. A. Panamá, and P. Šimeček. 2010. Graded leadership by dominant animals in a herd of female beef cattle on pasture. *Anim. Behav.* **79**: 1037–1045. doi:[10.1016/j.anbehav.2010.01.019](https://doi.org/10.1016/j.anbehav.2010.01.019)
- Shashar, N., G. G. Rosenthal, T. Caras, S. Manor, and G. Katzir. 2005. Species recognition in the blackbordered damselfish *Dascyllus marginatus* (Rüppell): An evaluation of computer-animated playback techniques. *J. Exp. Mar. Biol. Ecol.* **318**: 111–118. doi:<https://doi.org/10.1016/j.jembe.2004.12.037>
- Shpigel, M. 1982. Niche overlap among two species of coral dwelling fishes of the genus *Dascyllus* (Pomacentridae). *Environ. Biol. Fishes* **7**: 65–68. doi:[10.1007/BF00011825](https://doi.org/10.1007/BF00011825)
- Sridhar, V. H., D. G. Roche, and S. Gingins. 2019. Tracktor: Image-based automated tracking of animal movement and behaviour. *Methods Ecol. Evol.* **10**: 815–820. doi:[10.1111/2041-210X.13166](https://doi.org/10.1111/2041-210X.13166)
- Strandburg-Peshkin, A., D. R. Farine, I. D. Couzin, and M. C. Crofoot. 2015. Shared decision-making drives collective movement in wild baboons. *Science* **348**: 1358–1361. doi:[10.1126/science.aaa5099](https://doi.org/10.1126/science.aaa5099)
- Therault, D. H., N. W. Fuller, B. E. Jackson, E. Bluhm, D. Evangelista, Z. Wu, M. Betke, and T. L. Hedrick. 2014. A protocol and calibration method for accurate multi-camera field videography. *J. Exp. Biol.* **217**: 1843–1848. doi:[10.1242/jeb.100529](https://doi.org/10.1242/jeb.100529)
- Thomas, M. J., M. L. Peterson, E. D. Chapman, N. A. Fanguie, and A. P. Klimley. 2019. Individual habitat use and behavior of acoustically-tagged juvenile green sturgeon in the Sacramento-San Joaquin Delta. *Environ. Biol. Fishes* **102**: 1025–1037. doi:[10.1007/s10641-019-00888-1](https://doi.org/10.1007/s10641-019-00888-1)
- Viscido, S. V., and J. K. Parrish. 2004. Individual behavior and emergent properties of fish schools: A comparison of observation and theory. *Mar. Ecol. Prog. Ser.* **273**: 239–249. doi:[10.3354/meps273239](https://doi.org/10.3354/meps273239)
- Wang, S. H., J. Zhao, X. Liu, Z.-M. Qian, Y. Liu, and Y. Q. Chen. 2017. 3D tracking swimming fish school with learned kinematic model using LSTM network, p. 1068–1072. *In* 2017 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE. doi:[10.1109/ICASSP.2017.7952320](https://doi.org/10.1109/ICASSP.2017.7952320)
- Webster, M. S., and M. A. Hixon. 2000. Mechanisms and individual consequences of intraspecific competition in a coral-reef fish. *Mar. Ecol. Prog. Ser.* **196**: 187–194. doi:[10.3354/meps196187](https://doi.org/10.3354/meps196187)
- Wilson, S. K., S. C. Burgess, A. J. Cheal, M. Emslie, R. Fisher, I. Miller, N. V. C. Polunin, and H. P. A. Sweatman. 2008. Habitat utilization by coral reef fish: Implications for specialists vs. generalists in a changing environment. *J. Anim. Ecol.* **77**: 220–228. doi:[10.1111/j.1365-2656.2007.01341.x](https://doi.org/10.1111/j.1365-2656.2007.01341.x)
- Wu, Z., N. I. Hristov, T. L. Hedrick, T. H. Kunz, and M. Betke. 2009. Tracking a large number of objects from multiple views, p. 1546–1553. *In* 2009 IEEE 12th International Conference on Computer Vision. IEEE. doi:[10.1109/ICCV.2009.5459274](https://doi.org/10.1109/ICCV.2009.5459274)
- Xu, Z., and X. E. Cheng. 2017. Zebrafish tracking using convolutional neural networks. *Sci. Rep.* **7**: 42815. doi:[10.1038/srep42815](https://doi.org/10.1038/srep42815)

#### Acknowledgments

We thank the Interuniversity Institute for Marine Science of Eilat (IUI)

and its local staff for their helpful assistance with the field and especially Moty Ohevia for skillfully constructing the 3D camera system. We thank the marine ecology group at the IUI and the movement ecology group at the Hebrew University for insightful discussions, and especially Shir Bar for many shared dives and useful comments. We are grateful to Ty Hedrick for the free distribution of DLTdv and EasyWand and for his many useful advice to resolve several issues with the camera calibration. The study was supported by Israel Science Foundation grant ISF-1211/14 to A.G. and grant ISF-964/13 to R.N. A.E. was supported by the Minerva Center for Movement Ecology, and by fellowships from the Advanced School of Environmental Studies at the Hebrew University of Jerusalem and the IUI.

#### **Conflict of Interest**

None declared.

*Submitted 14 February 2021*

*Revised 30 May 2021*

*Accepted 28 June 2021*

*Associate editor: David Suggett*