# Using non-local background modeling to quantify the schooling behaviour of sticklebacks

Reza Ardekani[1], Anna K. Greenwood[2], Catherine L. Peichel[2], Simon Tavaré[2,3]
[1] *Program in Molecular and Computational Biology, University of Southern California*
[2] *Division of Human Biology, Fred Hutchinson Cancer Research Center*
[3] *DAMTP, University of Cambridge, Centre for Mathematical Sciences*
*dehestan@usc.edu, {agreenwo,cpeichel}@fhcrc.org, st321@cam.ac.uk*

## Abstract

*Sticklebacks have been used as model organisms in behavioural biology for a long time. We plan to use quantitative trait locus mapping to identify the genetic basis for differences in schooling behaviour between marine and benthic sticklebacks. To do this, we need to quantify the schooling behaviour of thousands of fish. We have developed a robust high-throughput video analysis system that allows us to screen a few thousand individuals automatically. We propose a non-local mean background modeling approach that allows us to detect and track sticklebacks and obtain the schooling parameters efficiently.*

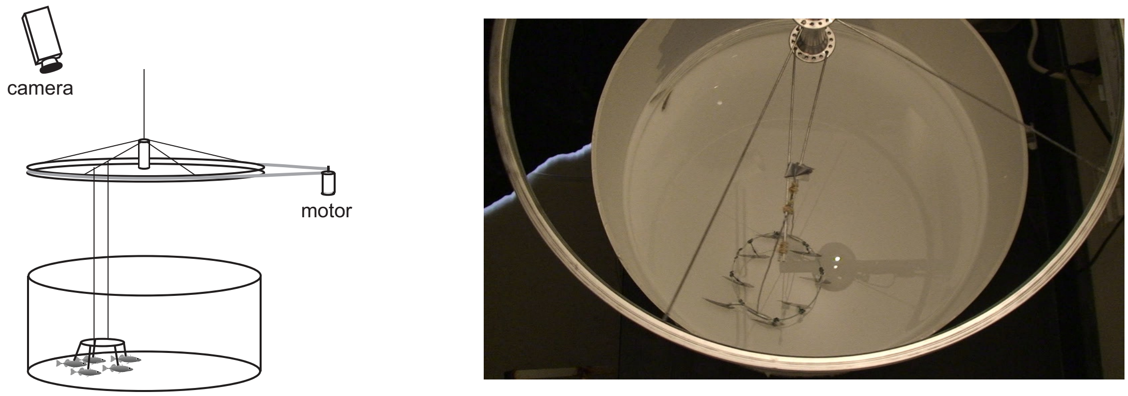
## 1. Introduction

Sticklebacks have been a model organism in behavioural biology since the pioneering work of Niko Tinbergen over half a century ago [9]. Much is understood about stickleback behaviour in both the field and the laboratory. More recently, sticklebacks have become a model system for understanding the genetic basis for divergence in phenotypic traits, including behaviour [5]. We have characterized differences in schooling behaviour between two populations of sticklebacks that inhabit dissimilar environments. Marine sticklebacks live in open water and school very strongly whereas freshwater bottom-dwelling lake populations (benthics) exhibit reduced schooling [11]. We developed an assay using an array of artificial stickleback models to elicit and quantify schooling behaviour [11]. Using this assay, we showed that marine sticklebacks spend significantly more time schooling.

Our goal is to dissect the genetic basis for the divergent schooling behaviour between marine and benthic sticklebacks. Quantitative trait locus (QTL) mapping has successfully identified the genetic basis for many variant traits in sticklebacks [5]. We plan to use QTL mapping in benthic-marine hybrids to identify genetic loci that contribute to differences in schooling behaviour. To assay the hundreds of fish necessary for this technique, a robust high-throughput video analysis system is essential. Here we present a custom approach for analysis of videos from our assay. We propose a method for background modeling for videos that are (semi-)periodic, those in which some or all of the background in each frame is repeated in at least a few other frames in the video. We show the result of this simple yet effective method for processing videos from our experiments.

## 2. Target detection for video tracking

For any video tracking system, target detection is an essential ingredient. One approach is to detect an object of interest based on appearance features such as geometric shape, texture and color [12]. In this approach, the visual features should chosen so that the target can be distinguished easily from other objects in the scene. Another approach to detect moving objects in the scene is background subtraction [7]. This approach is especially useful for surveillance systems, such as for parking lots, offices, and controlled experimental environments, in which cameras are fixed and directed to the area of interest. The main property of these systems is that background is to some extent static, and a model of background can be calculated for each frame [10]. Different methods have been developed to maintain robustly the background model in scenes with possible changes in background such as gradual change in lighting and sudden changes in illumination due to light switches [7, 10]. Moreover, there




(a) (b)

**Figure 1. (a) Physical setup (b) Sample frame from the video**

are studies that address background modeling in dynamic scenes with significant stochastic motion, such as water or waving trees [8, 2]. Unfortunately, the aforementioned approaches are not applicable for our experiments due to our special setup (see section 3.1).

## 3. Experimental setup

The model school is composed of eight plastic model sticklebacks that are arranged to mimic the formation of an actual school of sticklebacks [11]. The models are attached to an apparatus and driven by a motor in a circular path within a circular tank. Trials are videotaped using a video camera mounted above the tank, as shown in Fig. 1 (a). The features we quantify in each video are the time taken for the fish to move within one body length of the model, the time of schooling with the model (i.e. swimming in the same direction as the model, within one body length), and the number of schooling bouts (i.e., if the fish stops then starts schooling again, how many times does it do that). These data can be obtained from the position and direction of the fish and the model in each frame.

### 3.1. Challenges

There are two properties that make the task of tracking sticklebacks in our setup challenging. First of all, model school fishes, as intended, look very similar to the real fish (see Fig. 1 (b)). Therefore, there is no visual feature that could distinguish between real fish and model fish. So, even though it is possible to detect the real fish in the frames in which this fish is not close to the models using visual clues such as shape and intensity of fish contour, in the frames that the real fish is schooling with model fish it is almost impossible to distinguish them. Problematically, these are the frames in which we are most interested, because they show the schooling behaviour. Moreover, since the model school is rotating, there are frequent occlusions of the fish by the poles and wires around the model school. Basic methods for background modeling such as running Gaussian average [7] will not work here since not only does the real fish move between consecutive frames, but also the whole model school and the wires and poles are moving.

## 4. Proposed Method

### 4.1. Model school detection

As can be seen in Figure 1 (b) there is a circle on top of the model. An obvious choice for circle detection is the generalized Hough transform [4] and since the radius of the circle (aside from the negligible variation due to perspective effect) is constant, the model fish are effectively located. The process of model detection can be expedited by using the previous frame information for each frame and searching for a circle in the neighbourhood of the region of interest (close to the last frame detection) instead of searching the whole image. By finding the centre of the circle at each frame, the movement direction of model fishes is extractable; this is needed to calculate the statistics we need from each experiment.

### 4.2. Real fish detection

We want to build a background model for each frame such that the only "foreground" would be the real fish. This means we want to have the model school, poles and wires as background. Our proposed approach for background modeling for videos has some similarities with the NL-means algorithm described in [1]. In [1], for denoising a pixel, instead of just using the neighbours of the pixel or local pixels, all other pixels in the entire image that are "similar" to the current pixel are used. The measure of similarity is based on the intensity value of a square neighbourhood of fixed size.

The interesting property of the videos from our system is that for each frame, since the model school is turning around almost periodically, there are some other 'similar' frames in the video in which the position of the model school, as well as poles and wires and even shadows are almost the same. We exploit this specific feature of these videos to build a background model for each frame using the similar frames that exist in the whole video. So instead of using the neighboring frames (neighbor in terms of time), we look at the whole video, and find the frames that are similar to the current frame. Unfortunately, the period of turning varies slightly between and within experiments due to imperfections in the motor. Therefore, although periodicity helps us limit the search for finding similar frames, we still need to look in a window of frames in each period to pick the most similar frames to the current one. Our similarity measure is based on the Euclidean distance between frames. More precisely, $S_{f_1,f_2}$, the similarity score between frame $f_1$ and $f_2$ is defined as

$$S_{f_1,f_2} = 1 - C \times \sum_{i=0}^{w} \sum_{i=0}^{h} |I_{f_1}(i,j) - I_{f_2}(i,j)|$$

in which $h$ and $w$ are the height and width of the region of interest, respectively, $C$ is a normalization factor, and $I_f(i,j)$ is the intensity value of the pixel $(i,j)$ at frame $f$. In our videos, $w$ and $h$ are 700 and 540 respectively.

Since the area of the real fish is only about 0.1% of the whole image, the position of the fish does not make that much contribution to the value of the similarity score. This means that frames that are similar to each other have the same or very similar "background". To speed up the process of calculating the similarity score between frames, each frame is summarized as a vector of Haar-like features [6] that can be computed very efficiently using integral image [3]. In this case, the similarity distance

$$\hat{S}_{f_1,f_2} = 1 - \hat{C} \times \sum_{k=0}^{L} |V_{f_1}(k) - V_{f_2}(k)|$$

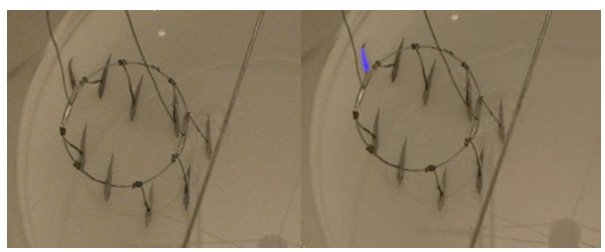
in which $V_f$ is a vector containing $L$ rectangular Haar-like features and $\hat{C}$ is a normalizing constant. We used $L = 100$.

| Segment No. | # MD | # FD, | % CD |
|---|---|---|---|
| 1 | 32 | 0 | 96.8 |
| 2 | 54 | 4 | 94.3 |
| 3 | 131 | 1 | 86.8 |
| 4 | 33 | 0 | 96.7 |
| 5 | 25 | 0 | 97.5 |
| Average | 55 | 1 | 94.4 |

**Table 1. Detection performance in 5 segments: Missed detection (MD), False Detection (FD) and Correct Detection (CD)**

For each frame, after ranking the similarity scores we pick the $N$ frames that have the highest scores; we used $N = 3$. The background for the current frame is then calculated using these frames. The latter step is needed to remove the water waves and other non-periodic changes in the image. Finally, we filter the components in the change mask based on their size and remove those components that are much smaller or larger than the real fish.

## 5. Results



(a) (b)

**Figure 2. Detection: (a) Original frame, (b) Processed frame; detected fish is indicated in blue**

Fig. 2 shows the result of real fish detection. The detected area is indicated in blue in Fig. 2(b). This shows our method is able to find the "foreground" or real fish, even in hard situations. To quantify the performance of our algorithm, we manually checked 5 segments of video of length 1000 frames. Table 1 shows the performance of the proposed method in terms of the number of missed/false detections. On average we correctly detect fish in 94.4% of the frames. This shows our detection algorithm works efficiently.

We present the result of processing three sample videos with the proposed method. For behavioural trials, fish are removed from their home tank and placed

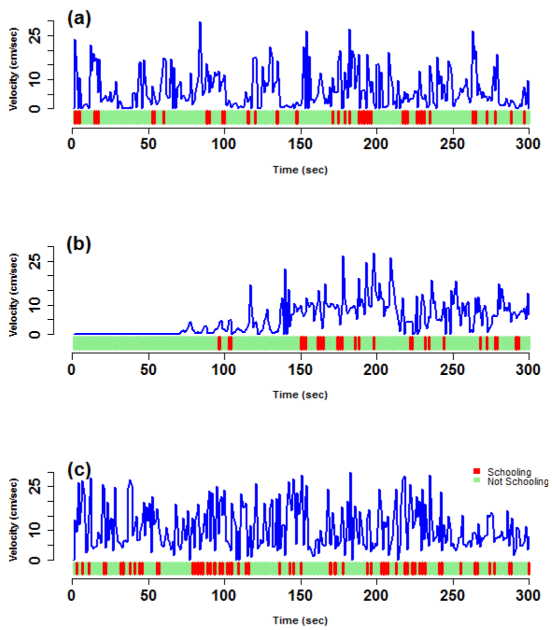



**Figure 3. Speed of movement and schooling behaviour for three sample videos. Red bars indicate inferred periods of schooling.**

into individual isolation chambers for at least 1.5 hours before the trial. The motor controlling the artificial school is then turned on remotely and the fish are given 5 minutes to interact with the models. For each frame, the distance between the model and the fish, and the speed of the fish, are obtained. If the distance between the fish and the model is less than a predefined threshold (5 cm) and the speed of the fish is more than a threshold (2 cm/sec), we identify that frame as schooling. Fig. 3 shows the result of quantifying speed and schooling behaviour. As can be seen, the patterns of schooling and activity differ between individuals. From these data, many parameters of schooling behaviour such as the latency to first approach to the model and average time of schooling can be obtained.

## 6. Conclusion

We have proposed a method to automate the quantitative analysis of stickleback schooling behaviour. We exploit the semi-periodic nature of the videos to build an accurate background model for each frame. Since we are processing recorded videos our background modeling algorithm does not need to be causal, however it can be extended for causal systems, e.g. real-time applications. The proposed method enables us to detect the fish in difficult situations, for example when the fish is very close to the model and/or is partially occluded. Using our approach, we can find the important parameters of schooling behaviour. This enables us to screen many individuals with different genotypes efficiently and conduct association studies between genotype and schooling behaviour.

## 7. Acknowledgments

Research reported in this publication was supported by the National Human Genome Research Institute of the National Institutes of Health under awards number P50HG002790 (RA, ST) and P50HG002568 (AKG, CLP). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## References

[1] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *CVPR*, volume 2, pages 60–65, 2005.

[2] A. B. Chan, V. Mahadevan, and N. Vasconcelos. Generalized Stauffer-Grimson background subtraction for dynamic scenes. *Mach. Vision Appl.*, 22:751–766, 2011.

[3] F. C. Crow. Summed-area tables for texture mapping. In *SIGGRAPH*, pages 207–212. ACM, 1984.

[4] R. O. Duda and P. E. Hart. Use of the Hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15:11–15, 1972.

[5] D. M. Kingsley and C. L. Peichel. *The Molecular Genetics of Evolutionary Change in Sticklebacks*. CRC Press, Boca Raton, 2007.

[6] C. P. Papageorgiou, M. Oren, and T. Poggio. A general framework for object detection. In *ICCV*, pages 555–562, 1998.

[7] M. Piccardi. Background subtraction techniques: a review. In *IEEE International Conference on Systems, Man and Cybernetics*, volume 4, pages 3099–3104, 2004.

[8] Y. Sheikh and M. Shah. Bayesian modeling of dynamic scenes for object detection. *PAMI*, 27:1778–1792, 2005.

[9] N. Tinbergen. The curious behavior of the stickleback. *Scientific American*, 187:22–26, 1952.

[10] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers. Wallflower: principles and practice of background maintenance. *ICCV*, 1:255 –261, 1999.

[11] A. R. Wark, A. K. Greenwood, E. M. Taylor, K. Yoshida, and C. L. Peichel. Heritable differences in schooling behavior among threespine stickleback populations revealed by a novel assay. *PLoS ONE*, 6, 2011.

[12] A. Yilmaz, O. Javed, and M. Shah. Object tracking: A survey. *ACM Comput. Surv.*, 38:13–57, 2006.