MSBOTS: a Multiple Small Biological Organism Tracking System robust against non-Ideal detection and segmentation conditions (#55756)

First submission

Guidance from your Editor

Please submit by **20 Dec 2020** for the benefit of the authors (and your $200 publishing discount) .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files

Download and review all files from the materials page.

1 Latex file(s)

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Speculation is welcome, but should be identified as such.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

MSBOTS: a Multiple Small Biological Organism Tracking System robust against non-Ideal detection and segmentation conditions

Xiaoying Wang [Corresp., 1], **Eva Cheng** [2], **Ian S Burnett** [2]

[1] School of Engineering, RMIT University, Melbourne, Australia

[2] Faculty of Engineering & Information Technology, University of Technology Sydney, Sydney, New South Wales, Australia

Corresponding Author: Xiaoying Wang
Email address: xiaoying.wang@rmit.edu.au

Accurately tracking a group of small biological organisms using computers to obtain their movement trajectories is essential to many biomedical and pharmaceutical research. However, object miss-detection, segmentation errors and overlapped individual trajectories are especially common issues that restrict the development of automatic multiple small organism tracking research. Extended on previous work, this paper presents an accurate and generalised Multiple Small Biological Organism Tracking System (MSBOTS), robust against non-ideal object detection and segmentation based on the Kuhn-Munkres association algorithm and the proposed object position calculation using inter-frame knowledge. Evaluated on a set of zebrafish, Artemia and Daphnia videos, a wide variety of video conditions were tested. The proposed system exhibits decreased overall MOTP errors of up to 71.59%, 77.59% and 72.73% when tested on zebrafish, Artemia and Daphnia videos, respectively. The MOTA accuracy of MSBOTS increased by up to 63.01% compared with the state-of-the-art idTracker system. Moreover, MSBOTS obtains more reliable tracking trajectories with a smaller standard deviation of up to 47.68 pixels compared with idTracker. The analysis of detailed locomotive characteristics of individual organisms from tracking trajectories is also presented in this paper. The developed MSBOTS system, analysed data and the locomotive analysis code are made freely available online.

MSBOTS: a Multiple Small Biological Organism Tracking System robust against non-Ideal detection and segmentation conditions

Xiaoying Wang[1], Eva Cheng[2], and Ian S. Burnett[2]

[1]**RMIT University, School of Engineering, Melbourne, 3000, Australia**
[2]**University of Technology Sydney, Faculty of Engineering & Information Technology, Sydney, 2007, Australia**

Corresponding author:
Xiaoying Wang[1]

Email address: xiaoying.wang@rmit.edu.au

ABSTRACT

Accurately tracking a group of small biological organisms using computers to obtain their movement trajectories is essential to many biomedical and pharmaceutical research. However, object miss-detection, segmentation errors and overlapped individual trajectories are especially common issues that restrict the development of automatic multiple small organism tracking research. Extended on previous work, this paper presents an accurate and generalised Multiple Small Biological Organism Tracking System (MSBOTS), robust against non-ideal object detection and segmentation based on the Kuhn-Munkres association algorithm and the proposed object position calculation using inter-frame knowledge. Evaluated on a set of zebrafish, *Artemia* and *Daphnia* videos, a wide variety of video conditions were tested. The proposed system exhibits decreased overall MOTP errors of up to 71.59%, 77.59% and 72.73% when tested on zebrafish, *Artemia* and *Daphnia* videos, respectively. The MOTA accuracy of MSBOTS increased by up to 63.01% compared with the state-of-the-art idTracker system. Moreover, MSBOTS obtains more reliable tracking trajectories with a smaller standard deviation of up to 47.68 pixels compared with idTracker. The analysis of detailed locomotive characteristics of individual organisms from tracking trajectories is also presented in this paper. The developed MSBOTS system, analysed data and the locomotive analysis code are made freely available online.

INTRODUCTION

In recent years, small biological organisms such as zebrafish larvae (genetically and physiologically similar to humans), *Artemia franciscana*, and *Daphnia magna* have become powerful models and are widely used to study human disease James et al. (2019), toxicology, pharmacology Comeche et al. (2017) and ecotoxicology James et al. (2019); Comeche et al. (2017); Poynton et al. (2007). Accurate tracking techniques are vital for understanding the biology and ecology underlying their movement Martineau and Mourrain (2013); Nema et al. (2016); Colwill and Creton (2011); Alyuruk et al. (2013); Ekvall et al. (2013). Traditional manual observation approaches are very tedious, time consuming, and the results are difficult to reliably repeat and reproduce. Fluorescent labelling can improve the human visual distinction ability of specific targets, but fluorescent materials affect the behavioural response of organisms Ekvall et al. (2013).

However, while automatic object tracking techniques have assisted in developing approaches for the behaviour and interaction analysis of large organisms such as mammals, birds and adult fish, the tracking of small organisms (in millimetre scale, most are considerably smaller than 1 mm Marechal et al. (2004)) are hampered by the constraints of existing automatic tracking methods Ekvall et al. (2013); Dur et al. (2011). There are many challenges imposed by the small size factor especially of aquatic organisms. Firstly, radio frequency identification chips, also called u-chips or transponders, are widely

used in individual organism identification application, while the available smallest chip size is around 0.4 mm in current Usami (2004). Such sized devices will dramatically affect the natural dynamic behaviour of organisms in mm-scale Ekvall et al. (2013); Lard et al. (2010). In addtion, general object tracking is already a complicated problem due to object occlusion or overlapping, non-rigid object structure (object rotating and scale changing), and motion pattern changing and more Habibi et al. (2017). The difficulty level rises when the tracking targets are in small size because small organisms provide little information regarding the imaging noise.

Existing automatic tracking approaches for multiple organisms Zhou et al. (2014); Conklin et al. (2015); Liu et al. (2017) either use adult fish Pérez-Escudero et al. (2014) in a big container to limit the chance of object interactions or use a petri dish plate to separate individual objects, allowing only one target in each petri dish to avoid overlapped and swapped trajectories that commonly seen in multiple organisms housed in one container. A recent machine learning technique applied to biological organism tracking is CNNTracker Zhiping and Cheng (2017), which optimises the accuracy of the zebrafish head feature map classification with individual identities. However, CNNTracker only tested on adult zebrafish, which has different movement characteristics and higher intensity contrast than millimetre-scale small organisms. Thus this method may miss-classify organisms with multiple identities Zhiping and Cheng (2017). The generalized linear mixed model (GLMM) based approach Liu et al. (2017) attempts to analyse larval locomotive response, but can only detect whether movement exists by handling binary move/no-move classification, but object displacement cannot be computed. However, these methods do not allow for the study of small organism interaction and group behaviour.

Furthermore, the accurate estimation of an object's location from segmentation results provides the critical foundation for the performance of subsequent tracking processes. However, the performance of such Multiple Target Tracking (MTT) methods is affected by false positive segmentation (e.g., noise fragments remaining after segmentation), and false negative segmentation from missed object detections and occluded objects Mallick et al. (2013). Such segmentation challenges commonly occur in microscopic videos of multiple small organisms, especially when taken under realistic experimental conditions. In addition, due to the small organism size, miss-detection and object occlusion or overlap are inevitable problems for multiple small organism tracking, and it is near impossible to identify and segment all organisms when videos taken under realistic experimental conditions (not high quality) are used Noss et al. (2013).

IdTracker Pérez-Escudero et al. (2014) is a well-known biological organism tracking system with 'fingerprint' generation for organism identity differentiation, and the commercial LoliTrack Závorka et al. (2017) system can track multiple targets in a single container. However, both approaches require the video input to be under strictly constrained imaging conditions to obtain accurate organism detection and segmentation results. As reported in the paper of Zhou et al. (2014), even small impurities inside the water and lighting reflection will affect the object segmentation accuracy. Non-ideal tracking results then subsequently pose challenges for the movement behavioural analysis and maintaining individual organism tracking identities over time Martineau and Mourrain (2013).

3D systems with multiple cameras or super-resolution images built from multiple low-resolution images Ekvall et al. (2013); Günel et al. (2019); Noss et al. (2013) are presented to obtain more information for accurate tracking of small organisms. These systems increase computational complex, change detection and association tracking system structure and require further object location registration and association among cameras or images. These challenges constrain the application in real-time detection and tracking of multiple organisms in group or in small-scale larval organisms Günel et al. (2019). And the detection and tracking performance are found also affected by the environmental conditions of videos Noss et al. (2013).

Extended upon our previous work Wang et al. (2017b, 2018), this paper presents an automatic, accurate and effective Multiple Small Biological Organism Tracking System (MSBOTS) robust against non-ideal object detection and segmentation results obtained from microscopic time-lapse videos taken under practical laboratory experimental conditions. The proposed system applies Gaussian Mixture Model (GMM) based background subtraction in segmentation module Wang et al. (2017b) to detect and segment the small organisms from each video frame, and initially maps detected objects between successive video frames to generate individual tracking trajectories based on the (non-ideal) segmentation results using the Hungarian algorithm (also called Kuhn–Munkres algorithm) Bourgeois and Lassalle (1971); Munkres (1957), which guarantees one to one association. The positions of miss-detected and overlapped

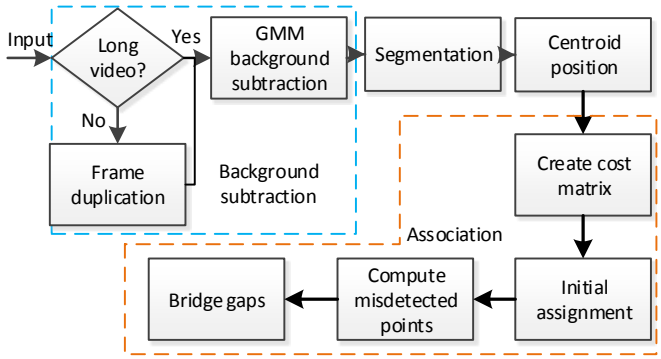


Figure 1. Flow chart of the overall multiple small organism tracking system.

objects are then calculated through knowledge of their neighbour's locations. And the theoretically computed locations are bridged in the individual tracking trajectories in the developed system. In addition, the tracking results enable locomotive analysis and behaviour study, where this paper further presents developed software to estimate the velocity, acceleration and movement direction. The performance and versatility of the proposed MSBOTS system is evaluated and demonstrated based on tracking accuracy, and compared with existing multiple organism tracking systems using three types of small organisms: zebrafish larvae, *Artemia franciscana*, and *Daphnia magna*.

METHODS

Figure 1 outlines the overall work-flow of the proposed MSBOTS System. The accurate differentiation of organisms from the video background and water impurities in each video frame is the critical foundation for multiple organism tracking systems. In the proposed MSBOTS platform, the video background is estimated by an adaptive Gaussian Mixture Model (GMM) Zivkovic and Heijden (2006). Organisms in every video frame are segmented after the background subtraction; details of this segmentation approach based on background subtraction has been summarised and reported in our previous work Wang et al. (2017b). The following source-target assignment procedure is based on the computed centroid point locations of segmented regions. This step plays an essential role in maintaining consistent identities for individual detected objects over time. To find the corresponding organismss between frames, the Kuhn-Munkres algorithm Bourgeois and Lassalle (1971) is applied in the association module.

The proposed mapping algorithm not only finds the targets in the next frame for the organisms detected in every frame, but also calculates the theoretical positions for miss-detected or occluded organisms. After obtaining the moving trajectories of all organisms associated with individual identities, the movement characteristics for each organism are then estimated for dynamic behaviour analysis.

Code of Ethics

Ethical approval is not applicable to this work, because no chemicals or medicine were tested with the organisms being filmed.

Organism Detection and Segmentation

Small biological organisms from time-lapse microscopic videos are difficult to detect and segment due to imbalanced movement characteristics, where organisms such as zebrafish larvae can exhibit a mean proportion of activities less than 0.075 over time Liu et al. (2017), with 'bursty' movements of sudden swimming locomotion interspersed with substantially stationary periods of little or no movement Liu et al. (2017). In addition, the relative small size between organisms and water impurities (e.g. the water bubbles and ripple surface injections) render the distinguishing between them a challenge. For example, the typical size used in dynamics and path tracking study is less than 0.4 mm Rashid et al. (2012).

For background subtraction in the proposed MSBOTS approach as reported in our previous work Wang et al. (2017b), an improved GMM Zivkovic and Heijden (2006) was chosen to estimate the stationary

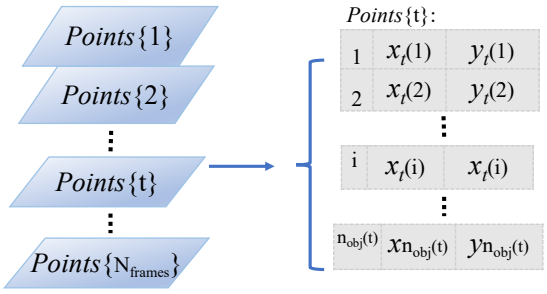


Figure 2. Storage structure of detected objects in a video sequence.

background due to its detection period extension ability. In the video background estimation, a moving object will be firstly represented by a new region cluster with a small weight, whose value will be gradually increasing if the region remains the motion state. The new region cluster will be classified as a foreground object only if its weight exceeds a threshold c_f (whose value is calculated as in Zivkovic and Heijden (2006)). Thus, with a decaying envelope factor α, the detection period of an object with no movement is extended for approximately $\log(1-c_f) \setminus \log(1-\alpha)$ frames Zivkovic and Heijden (2006).

Furthermore, to adapt to background changes, the adaptive GMM Zivkovic and Heijden (2006) adds flexibility to the number of Gaussian components that describe the background by adaptively and recursively selecting the number of Gaussian components. For the initialisation and selecting the number of Gaussian components, the approach from Zivkovic and Van Der Heijden (2004) and the Dirichlet criteria are applied, respectively. The component parameters are then updated for each video frame, in contrast to traditional GMM models that apply one or a fixed number of components to model each pixel. The foreground objects are then obtained by the differentiation of the video frames with the corresponding background estimated by the adaptive GMM Zivkovic and Heijden (2006). After the background subtraction is performed, a median filter and morphological grayscale erosion are applied to remove remaining image distortion and scattered small noise fragments Pnevmatikakis and Polymenakos (2006).

The MSBOTS background subtraction approach enables the removal of stationary backgrounds such as the organism container and labels drawn on the containers; hence, unlike existing techniques, the proposed system is able to process videos under practical experimental imaging conditions.

Representing Detected Organisms

To represent the positions of the detected organisms in each video frame, the centroid locations of segmented foreground regions in Cartesian coordinates are used and stored in a vertical cell array matrix, as shown by the parallelogram series (indicating video frames) and $Points\{t\}$ matrix in Fig. 2. In the $Points\{t\}$ cell array, the first column stores the temporary identity, numbered from 1 to the number of detected organisms in each frame, where $n_{obj}(t)$ indicates the number of detected organisms in frame t. The second and third column stores the horizontal and vertical positions of each detected organism in X and Y coordinates, respectively. This cell array and matrix representation allows for varying element length to indicate the number of detected foreground organisms, which can change frame-to-frame due to detection and segmentation errors.

Organism Assignment between Frames

The centroid positions of detected organisms are obtained in the segmentation processes and represented frame-by-frame using a cell array for a video sequence as described in the previous section; however, the organism identities in each frame are still unknown. That is, which organism in the current frame corresponds to which organism in the following frame has not been mapped. In addition, there are still some remaining miss-detected organisms that have been classified into background clusters or overlapped with other detected organisms. The assignment algorithm within the proposed MSBOTS approach constructs the individual tracking trajectories by linking the detected organisms in each video sequence to correspondences in the next successive frame, calculating the positions of miss-detections or occlusions, and reassigning the calculated organisms to their correct trajectories by adjusting initial assignment results.

Initial assignment

The initial assignment of detected organisms in each video frame is a partial assignment using an extension of the Kuhn-Munkres algorithm to frame-by-frame processing for rectangular arrays Bourgeois and Lassalle (1971). Extending the original Hungarian algorithm Harold (1955) from solving the assignment problem with an equal number of workers and tasks represented in a $n \times n$ matrix, the number of workers and tasks can be unequal and represented in a rectangular matrix. This approach can then be applied to organism tracking, where the number of detected organisms can change due to non-ideal segmentation resulting from organism miss-detection and occlusion.

In the initial assignment process in this extended Kuhn-Munkres approach Bourgeois and Lassalle (1971), all the detected organisms $Points\{t\}$ in frame t are taken as source points, and the segmented organisms $Points\{t+1\}$ in the following frame $t+1$ are seen as the target points between connected frames. The target points in $Points\{t+1\}$ are mapped to the source points in $Points\{t\}$ frame-by-frame across a video sequence. An $n \times m$ matrix $\mathbb{D}$ is created to annotate the cost of associating source organisms $O = \{O_1, O_2, \cdots, O_n\}$ in the frame t to the target organisms $T = \{T_1, T_2, \cdots, T_m\}$ in the frame $t+1$:

$$D(O,T) = \begin{pmatrix} d_{O_1,T_1} & d_{O_1,T_2} & \cdots & d_{O_1,T_m} \\ d_{O_2,T_1} & d_{O_2,T_2} & \cdots & d_{O_2,T_m} \\ \vdots & \vdots & \ddots & \vdots \\ d_{O_n,T_1} & d_{O_n,T_2} & \cdots & d_{O_n,T_m} \end{pmatrix} \tag{1}$$

where n is the detected number of organisms in the frame t, and m is the number of organisms segmented in the successive frame $t+1$. The element d_{O_i,T_j} in the matrix denotes the cost to connect the i-th organism in the frame t, to the j-th organism in the frame $t+1$, and is calculated by the Euclidean distance between the source point to the target point using Equation (2):

$$d_{O_i,T_j} = (x_j - x_i)^2 + (y_j - y_i)^2 \tag{2}$$

The frame-to-frame organism assignment based on the cost matrix is performed using the Muncres implementation of the Hungarian algorithm Pilgrim (2017), which searches for unique assignments to assign source organism i to only one target organism j in the successive frame. The final sum of the resultant complete assignment between $Points\{t+1\}$ and $Points\{t\}$ is a global optimal cost, which is the lowest summed distance amongst all of the possible assignments within two successive frames. The matched target points propagate the identities of their matched source points; thus, after obtaining the final assignment map for the whole video, connecting the points with the same identities over frames of a video gives the individual organism tracking trajectories.

To eliminate the false positive points from the segmentation results when building the individual trajectories, a distance constraint is set as a threshold in the source-target cost matrix when applying the initial frame-to-frame assignment. The threshold is calculated by $\delta * \text{median}(dis_{i,j})$ Zhiping and Cheng (2017). When the minimum value of the i-th row in the source-target cost matrix is larger than the threshold value of this video sequence, which indicates that the distance between the source point i to all of the points in the successive frame exceeds the threshold value, the source point i with its corresponding position information will be removed from the point matrix and considered as a segmentation noise fragment to not be further assigned to a target point.

When an organism disappears in a frame t due to miss-detection or occlusion, a source point in frame $t-1$ therefore cannot be assigned to a target. A gap will then occur in the tracking trajectory where the organism fails to be detected, and a new tracking trajectory will start from the frame when the organism is correctly detected again. This source point in frame $t-1$ without a mapped target is saved in an unmatched source matrix.

When the organism is re-detected in the frame $t+n$ after being missed for n frames, there is one more point in $Points\{t+n\}$ compared to $Points\{t+n-1\}$. To map the points $Points\{t+n-1\}$ to $Points\{t+n\}$, a point in the frame $t+n$ cannot be assigned to a source point in the previous frame; this point in frame $t+n$ is saved in the unmatched target matrix.

The methods to calculate the theoretical positions between the unmatched source points and unmatched target points and adding these points to their correct tracking trajectories are explained in the following two sections, respectively.

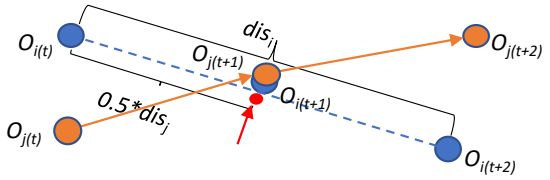


Figure 3. Point calculation for occluded organisms

Position estimation for miss-detected and occluded organisms

Fig. 3 illustrates the location computation of an overlapped organism. The two points O_i (shown by the blue dot) and O_j (shown by the yellow dot) overlap with each other in the frame $t+1$, and this overlapped point at time $t+1$ is assigned to the object O_j by the initial tracking process. Since the point O_i in the frame t cannot find a target in the frame $t+1$, and the point in the frame at time $t+2$ cannot be assigned to a source point in frame $t+1$, the point O_i is classified as an unmatched source point in the frame t, and unmatched target point in frame $t+2$ by the initial assignment approach described in the previous section, respectively.

The position of the missed point due to occlusion (as shown by the blue dot partially covered by the yellow dot in Fig. 3) or miss-detection (for example, when the blue dot is totally covered by the yellow dot in Fig. 3) is calculated by the unmatched source point $O_i(t)$ and unmatched target point $O_i(t+2)$. For example, as shown in Fig. 3, the location of the missed point O_i at frame $t+1$ is calculated by the median point between $O_i(t)$ and $O_i(t+2)$.

When there are multiple unmatched pairs, the mapping from unmatched target points to the unmatched source points is also based on the extended Hungarian assignment algorithm Bourgeois and Lassalle (1971). In searching for target points, the unmatched source points firstly search for possible correspondences in the following frame. If no assignment can be mapped, the search extends to the unmatched targets in the following 3rd frame. It was shown in Noss et al. (2013) that a trajectory fragment can be connected to its subsequent fragment track so long as the frame separation is less than 6 video frames. Thus, the default search range in the proposed MSBOTS system is from the second to the sixth frames following the frame when the miss-detections and occlusions originally occur. The positions of missed organisms (x_c, y_c) are calculated by Equations (4) and (5) using the matched point pair from an unmatched source point (x_s, y_s) and an unmatched target point (x_t, y_t).

$$x_c = x_s + \frac{1}{j} * (x_t - x_s) \tag{4}$$

$$y_c = y_s + \frac{1}{j} * (y_t - y_s) \tag{5}$$

where j indicates the following j-th frame from the unmatched source point.

Bridging trajectory gaps for miss-detected and occluded organisms

Individual tracking trajectories for each organism are obtained through connecting the matched points with the same identities after the initial assignment process frame-by-frame over a video sequence. However, the tracking trajectories obtained from the initial assignment process are usually trajectory fragments, separated when organisms are miss-detected or overlapped to result in segmentation errors. In the proposed MSBOTS approach, these trajectory gaps are bridged by adding the location points of these miss-detected or overlapped organisms as estimated and described in the previous section.

To connect trajectory fragments, the points stored in the unmatched source matrix are mapped to the points in the unmatched target matrix, and the positions of the missed points between the newly unmatched-source to unmatched-target are also calculated during this searching process. For example, as shown in Fig. 3, the unmatched source $O_i(t)$ as the end point of its trajectory fragment is connected to the unmatched target point $O_i(t+2)$, which is the start point of its trajectory fragment, and their middle point shown by the red dot is added between points $O_i(t)$ and $O_i(t+2)$ as the theoretical position of the overlapped point $O_i(t+1)$.

Locomotion characteristic analysis

After obtaining the individual tracking trajectories for each organism in video sequences, the movement characteristics of these organisms can be analysed based on the tracking results. The calculation of three movement parameters are presented in this work: movement velocity, acceleration and direction as represented by Equ. (6-8), respectively.

$$velocity = \frac{\sqrt{(x_{t+1}(i) - x_t(i))^2 + (y_{t+1}(i) - y_t(i))^2}}{dt} \tag{6}$$

$$acceleration = \frac{d(velocity)}{d^2t} \tag{7}$$

$$direction = \text{atan2}\,\frac{y_{t+1}(i) - y_t(i)}{x_{t+1}(i) - x_t(i)} \tag{8}$$

where $dt = 1 \backslash fs$ and fs is the video frame rate. $x_t(i)$ and $y_t(i)$ are the Cartesian coordination of organism i (i is the signed organism identity) in the frame t.

RESULTS AND DISCUSSION

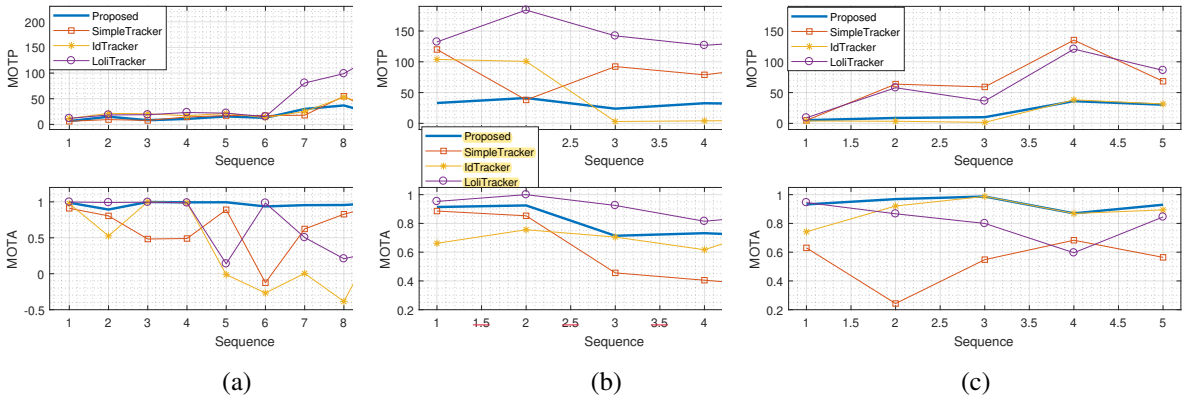


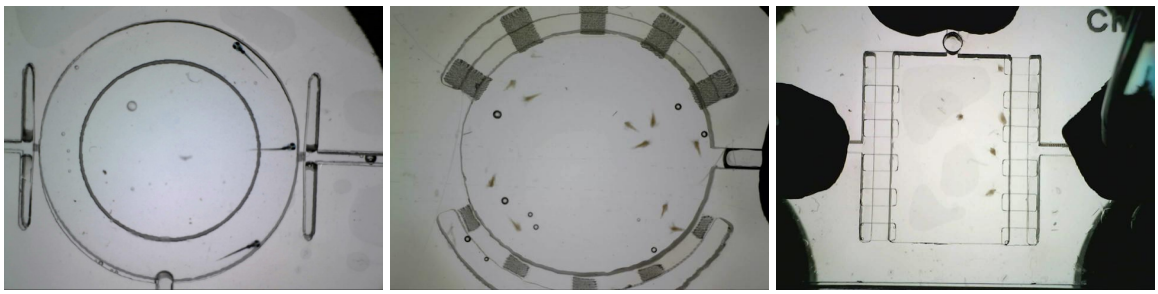
Figure 4. Tracking results evaluation among the comparison methods using videos of three types of small biological organisms.

To evaluate the proposed MSBOTS system, microscopic videos of three types of small biological organisms (zebrafish, *Artemia* and *Daphnia*) were tested, where the initial system test using zebrafish was presented in our previous work Wang et al. (2018). Evaluated on a set of single and multiple larvae and adult zebrafish, *Artemia* and *Daphnia* videos, a wide variety of (complex) video conditions were tested, including shadowing, labels, and background artefacts (such as water impurities, subject feces and water bubbles with different sizes). No chemical stimuli were tested on the studied organisms, so their behaviour response does not correspond to any specific chemical stimuli. In addition to the tracking accuracy evaluation, the natural locomotive characteristics as described by movement velocity, acceleration and direction are also analysed on the video dataset to test the dynamic behaviour analysis capability of the proposed system.

Small biological organism datasets

Microscopic time-lapse videos of three types of small organism models: zebrafish, daphnia and drasophila were applied to evaluate the proposed MSBOTS approach. Low frame rate videos were recorded with a Dino-Lite AD7013MT microscope at frame rates of 14 or 15 fps. High frame-rate videos were captured by an Imaging Development Systems (IDS) UI-3360CP-C-HQ microscope, with a high resolution 12.5mm focal lens.

Wild zebrafish embryos (Danio rerio) were incubated at 28°C in a petri dish filled with an E3 medium. Any dead and unfertilised embryos were manually removed three hours post-fertilization (hpf). Five days post-fertilization, the larvae were obtained from hatched zebrafish embryos. For data acquisition,



(a) (b) (c)

Figure 5. Microscopic video frame examples of (a) zebrafish larvae, (b) *Artemia franciscana*, and (c) *Daphnia magna*.

zebrafish larvae were transferred to poly (methyl methacrylate) (PMMA) housing wells. The zebrafish dataset consists of 10 video sequences with 3056 frames in total, of various durations and imaging conditions (a microscopic video frame example of zebrafish larvae is shown in Fig. 5a); details about the zebrafish dataset and its generation can be found in our previous work Wang et al. (2017a). The dataset was presented in the order that the first sequence has clearest background and the 10th (last) sequence has the most complex background.

Cysts of the marine crustacean *Artemia franciscana* and freshwater *Daphnia magna* were hatched and cultured according to the Artoxkit-M and Daphtoxkit-F (MicroBioTests Inc., Belgium) standard operating protocols. *Artemia franciscana* were hatched in a petri dish filled with sea water (pH 8.0 ± 0.5) at 24 ± 0.5°C under exposure to 3000-4000 lux light source for 30 hours. *Artemia* were placed into a group of 10 in a miniaturised Lab-on-a-Chip (LOC) chamber Solis et al. (2015) when shooting videos with microfluidics infused at a flow rate of 5.25 mL/h. Five *Daphnia magna* neonates were randomly selected and transferred into a petri dish monitored temperature at 20.0 ± 0.5°C. The *Artemia franciscana* and *Daphnia magna* dataset consists of 5 video sequences each with 4802 frames and 4804 frames in total, respectively.

- *Artemia franciscana* microscopic videos containing 5 organisms and artifacts (bubbles of different sizes, seqs. 1-5), with Fig. 5b a microscopic video frame examples of *Artemia franciscana*.
- *Daphnia magna* microscopic videos containing 10 organisms and artifacts (bubbles and impurities of different sizes, seqs. 1-5), with Fig. 5c a microscopic video frame examples of *Daphnia magna*.

Tracking evaluation metrics

To enable the objective and quantitative evaluation of the tracking performance from the proposed MSBOTS approach, this paper employs the widely utilised standard Multiple Object Tracking (MOT) metric: Classification of Events, Activities and Relationships (CLEAR MOT) Bernardin and Stiefelhagen (2008). CLEAR MOT consists of two metrics: Multiple Object Tracking Precision (MOTP), which estimates the location precision of all detected objects compared to that of the manually labelled object positions in each frame (known as ground truth); and, Multiple Object Tracking Accuracy (MOTA), which measures the accuracy in tracking object trajectories (producing exactly one trajectory per object), and the ability to consistently label objects over time. Mathematically, the MOTP and MOTA metrics are represented as:

$$MOTP = \frac{\sum_{i,t} |D_{i,t} - GT_{i,t}|}{\sum_t N_t} \tag{9}$$

$$MOTA = 1 - \frac{\sum_t (m_t + fp_t + mme_t)}{\sum_t g_t} \tag{10}$$

where $|D_{i,t} - GT_{i,t}|$ indicates the Euclidian distance between the pair-wise matched position of the i-th segmented object in frame t $D_{i,t}$ and the position of this object in the ground truth $(GT_{i,t})$, averaged by the total number of matches in the entire video sequence.

In the MOTA metric, m_t, fp_t, and mme_t for each frame t indicate the number of missed detections, false positive segmentation (i.e., image noise fragment detected as object), and the swapping of identities of individual objects, respectively. g_t represents the total number of objects present in frame t.

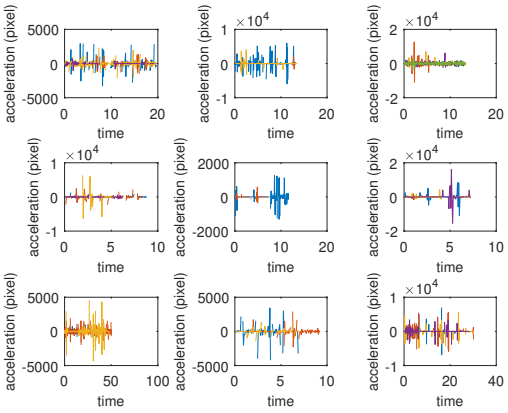
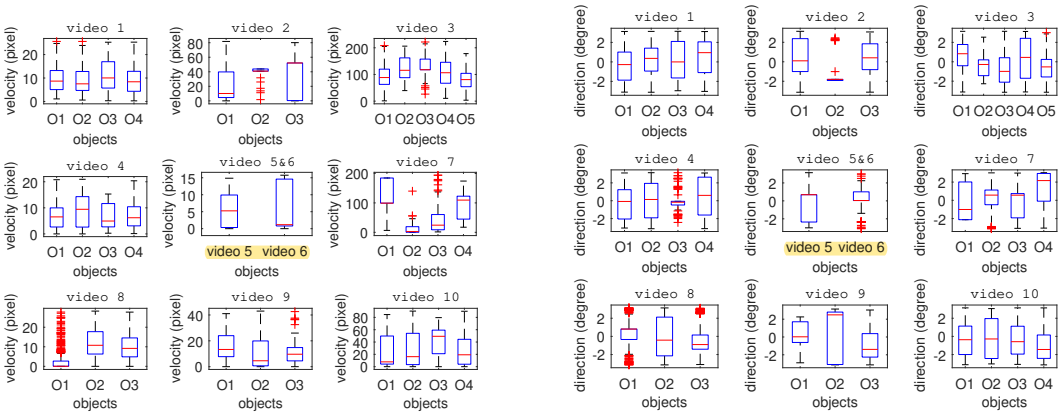


Figure 6. Organism movement speed acceleration over time.



(a) Movement speed analysis of individual organism(b) Movement direction analysis of individual organism.

Figure 7. Analysis of organism locomotion characteristics across video sequences.

Tracking accuracy evaluation

To evaluate the proposed tracking system, the overall tracking accuracy over a video sequence is compared with well-known multiple object tracking approach idTracker Pérez-Escudero et al. (2014), SimpleTracker Bourgeois and Lassalle (1971) (using the Kuhn-Munkres tracking algorithm for initial association and the nearest neighbourhood to directly connect trajectory fragments, without considering missed objects), and the widely used commercial LoliTrack system Závorka et al. (2017). Fig. 4 shows the tracking accuracy results measured by the MOTP and MOTA, with Fig. 4a showing results from zebrafish videos, Fig. 4b *Artemia* videos and Fig. 4c *Daphnia* videos.

It can be seen from Fig. 4a that the positions of detected organisms using adult and larvae zebrafish videos consistently exhibit the smallest distance differences with the manually labelled ground truth positions amongst all the methods tested. The proposed MSBOTS approach resulted in overall smaller MOTP values of 0.92, 25.59, and 44.48 pixels (71.59% increased) compared to SimpleTracker, idTracker, and LoliTrack, respectively (see Supplementary Table S1 for more detail). The general improved organism position detection results demonstrated the efficiency of the theoretical position estimation based on the organism detection and segmentation results using the adaptive GMM model and post-processing in the proposed MSBOTS system.

All of the methods performed well for position accuracy of the detected organisms when the videos had a clear background as shown by seqs. 1-6. However, the location errors measured by MOTP for the detected organisms in the zebrafish videos compared to the ground-truth did not decrease as dramatically

as LoliTrack or idTracker with increasingly complex video backgrounds, as shown by the MOTP values in seqs. 7-10 in Fig. 4a. Accordingly, the proposed MSBOTS method still out-performed the existing approaches when taking into account miss-detection, false positive segmentation and identity swapping, with 31.20%, 63.01%, and 24.61% higher MOTA values than SimpleTracker, idTracker, and LoliTrack, respectively. This was mainly achieved by the ability to estimate the positions of the missdetected or overlapped organisms using their neighbour position knowledge from the segmentation results in the proposed MSBOTS. In addition, the trajectory fragments bridging strategy of MSBOTS using the extended Hungarian assignment algorithm Bourgeois and Lassalle (1971) when there are multiple unmatched trajectory fragment pairs based on unmatched source-target points decreased the possibility of individual identity swapping occurrence, compared with SimpleTracker, which only used distance metric by nearest neighbour algorithm, and further in turn can generate identity swapping during the gap bridging process.

In addition, Fig. 4b and Fig. 4c shows the tracking accuracy evaluation using *Artemia franciscana* and freshwater *Daphnia magna* videos, respectively, to test the application of the proposed MSBOTS system on other small organisms with variant movement characteristics with zebrafish larvae (*Artemia* displays flexible movement and varies according to surrounding fluidics Tyson (1974); Williams (1994), and *Daphnia* with short, jerky hopping movement in water Rottmann et al. (1992)). The overall tracking accuracy performance of the proposed MSBOTS method is consistent for the tested videos on these two organism types, while the tracking accuracy measured by MOTP and MOTA values vary among these videos (see Supplementary Table S2 $ S3 for more detail).

As can be seen in MOTP values for *Artemia* videos in Fig. 4b, the proposed method exhibits 47.68 pixels smaller standard deviation than the idTraker system, which illustrated the usability of the proposed system on *Artemia* microscopic videos and the enhancement ability on improving the tracking accuracy. Though idTracker can produce smaller organism position estimation errors (MOTP) as shown by the seqs. 3-5 in Fig. 4b, the mean MOTA value is 7.07%, and 6.44% less than the proposed method and LoliTrack, respectively, which illustrates that the similar organism detection problem with zebrafish larvae detection (detected organism as background and impurities as organism due to their small size differentiation and similar movement character some time, especially when the organism stops moving or the water impurities stirred up by organism movement), which further cause identities confusion due to these detection errors).

Fig. 4c shows the overall tracking accuracy of the proposed MSBOTS system and idTracker consistently out-perform the comparing systems, measured by MOTP and MOTA. While the mean MOTA value of the proposed MSBOTS system is still 5.48% higher than idTracker. The tracking efficiency and accuracy of the proposed MSBOTS system illustrated application ability on *Daphnia* with short, jerky hopping movement characteristics compared with these exiting systems.

Analysis of organism movement characteristics

To explore the capability of the proposed MSBOTS approach to be applied to organism movement characteristic analysis due to the improved accuracy organism tracking trajectories, velocity, acceleration and movement direction were calculated for the zebrafish videos as an example.This paper also presents the basic zebrafish movement characteristics after obtaining individual tracking trajectories

Fig. 6 shows the movementacceleration analysis for each zebrafish video using the resultant tracking trajectories generated by the proposed MSBOTS method. It was found that in Hinz and de Polavieja (2017) the interaction and movement of zebrafish larvae was very close to zero by 7 dpf. As shown in Fig 6, in general the variation in the zebrafish acceleration values that is obviously visually perceptible to the human eye occurred only at approximately 10 seconds in the 10 adult and larvae zebrafish videos tested, due to the anxiety response from the zebrafish when turning on the imaging camera Peng et al. (2016)). The zebrafish speed are more subtle from this point on, which is consistent with the zebrafish movement characteristics found in Peng et al. (2016) Hinz and de Polavieja (2017).

As organism movement speed and changes in movement direction can provide insight into the interaction rules Hinz and de Polavieja (2017),Fig. 7 shows an example of the acceleration and movement direction results for each zebrafish video, respectively. The number of boxes show that each zebrafish larvae was successfully assigned with one identity (which also illustrated the one-to-one organism mapping guarantee of the association method in the proposed MSBOTS system), and the median velocity (labelled by the red line inside each box) of each larvae indicates the consistent movement characteristic within the same housing well. Mean, minimum and maximum velocity values and quartiles can be easily obtained

and analysed through the visual box plot representation shown in Fig. 7.

CONCLUSIONS

Accurate automatic tracking for multiple small biological organisms provides efficient approach to many biomedical and ecotoxicity applications. However, organism miss-detection and occluded organisms are inevitable problems when detecting and segmenting these small biological organisms from time-lapse microscopic videos (the detection and segmentation difficulty level rises when tracking small biological aquatic organisms compared with general objects), which in turn affects the subsequent organism tracking processes. To improve the tracking accuracy based on the non-ideal organism detection and segmentation results, extended upon and improved our previous work, this paper presents a Multiple Small Biological Organism Tracking System (MSBOTS), combining multiple object association algorithm for linking detected objects frame-by-frame and tracking trajectory adjustment techniques. To address segmentation errors due to miss-detected or occluded organisms, the proposed MSBOTS approach estimated positions of organisms in interim frames using corresponding points in neighbouring frames. Finally, the calculated points are applied to connect and adjust the tracking trajectory fragments from the initial association based on an extended Kuhn-Munkres algorithm. The proposed system was tested on three different types of small organism with variant movement characteristics using 20 videos in total. The resulted tracking accuracy of the proposed system outperforms three existing (state-of-the-art/commercial) tracking systems. Moreover, this system also provides locomotive characteristic analysis module using the generated individual tracking trajectories to facilitate small organism behaviour analysis research. Behavioural rules and new medicine or chemical effect on the dynamic behaviour of organisms can be further summarised using the provided movement analysis module to facilitate specified experiments.

DATA AVAILABILITY STATEMENTS

The datasets analysed during the current study were made fully publicly available in the online repository, https://github.com/Xiao-ying. The developed MSBOTS system and the software to analyse locomotive characteristics of individual organisms can also be downloaded from the same depository.

REFERENCES

Alyuruk, H., Demir, G. K., and Cavas, L. (2013). A video tracking based improvement of acute toxicity test on artemia salina. *Marine and Freshwater Behaviour and Physiology*, 46(4):251–266.

Bernardin, K. and Stiefelhagen, R. (2008). Evaluating multiple object tracking performance: the clear mot metrics. *EURASIP Journal on Image and Video Processing*, pages 1687–5281.

Bourgeois, F. and Lassalle, J. C. (1971). An extension of the munkres algorithm for the assignment problem to rectangular matrices. *Communications of the ACM*, 14(12):802–804.

Colwill, R. M. and Creton, R. (2011). Locomotor behaviors in zebrafish (danio rerio) larvae. *Behavioural Processes*, 86(2):222–229.

Comeche, A., Martín-Villamil, M., Picó, Y., and Varó, I. (2017). Effect of methylparaben in artemia franciscana. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology*, 199:98–105.

Conklin, E. E., Lee, K. L., Schlabach, S. A., and Woods, I. G. (2015). Videohacking: automated tracking and quantification of locomotor behavior with open source software and off-the-shelf video equipment. *Journal of Undergraduate Neuroscience Education*, 13(3):A120–A125.

Dur, G., Souissi, S., Schmitt, F., et al. (2011). Effects of animal density, volume, and the use of 2d/3d recording on behavioral studies of copepods. *Hydrobiologia*, 666(1):197–214.

Ekvall, M. T., Bianco, G., Linse, S., Linke, H., Backman, J., and Hansson, L. A. (2013). Three-dimensional tracking of small aquatic organisms using fluorescent nanoparticles. *PloS One*, 8(11):e78498.

Günel, S., Rhodin, H., Morales, D., et al. (2019). Deepfly3d, a deep learning-based approach for 3d limb and appendage tracking in tethered, adult drosophila. *eLife*, 8.

Habibi, Y., Sulistyaningrum, D. R., and Setiyono, B. (2017). A new algorithm for small object tracking based on super-resolution technique. In *AIP Conference Proceedings*, volume 1867, pages 20–24. AIP Publishing.

Harold, W. K. (1955). The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2:83–97.

Hinz, R. C. and de Polavieja, G. G. (2017). Ontogeny of collective behavior reveals a simple attraction rule. *Proceedings of the National Academy of Sciences*, 114(9):2295–2300.

James, D. M., Kozol, R. A., Kajiwara, Y. J., et al. (2019). Intestinal dysmotility in a zebrafish (danio rerio) shank3a; shank3b mutant model of autism. *Molecular Autism*, 10(1):3–18.

Lard, M., Backman, J., Yakovleva, M., Danielsson, B., and Hansson, L. A. (2010). Tracking the small with the smallest-using nanotechnology in tracking zooplankton. *PloS One*, 5(10):e13516.

Liu, Y. W., Ma, P., Cassidy, P. A., Carmer, R., Zhang, G. N., Venkatraman, P., Brown, S. A., Pang, C. P., Zhong, W. X., Zhang, M. Z., et al. (2017). Statistical analysis of zebrafish locomotor behaviour by generalized linear mixed models. *Scientific Reports*, 7.

Mallick, M., Vo, B., Kirubarajan, T., and Arulampalam, S. (2013). Introduction to the issue on multitarget tracking. *IEEE Journal of Selected Topics in Signal Processing*, 7(3):373–375.

Marechal, J. P., Hellio, C., Sebire, M., and Clare, A. (2004). Settlement behaviour of marine invertebrate larvae measured by ethovision 3.0. *Biofouling*, 20(4-5):211–217.

Martineau, P. R. and Mourrain, P. (2013). Tracking zebrafish larvae in group–status and perspectives. *Methods (San Diego, Calif.)*, 62(3):292—303.

Munkres, J. (1957). Algorithms for the assignment and transportation problems. *Journal of the society for industrial and applied mathematics*, 5(1):32–38.

Nema, S., Hasan, W., Bhargava, A., and Bhargava, Y. (2016). A novel method for automated tracking and quantification of adult zebrafish behaviour during anxiety. *Journal of Neuroscience Methods*, 271:65–75.

Noss, C., Lorke, A., and Niehaus, E. (2013). Three-dimensional tracking of multiple aquatic organisms with a two camera system. *Limnology and Oceanography: Methods*, 11(3):139–150.

Peng, X. L., Lin, J., Zhu, Y. D., Liu, X. Y., Zhang, Y. L., Ji, Y. X., et al. (2016). Anxiety-related behavioral responses of pentylenetetrazole-treated zebrafish larvae to light-dark transitions. *Pharmacology Biochemistry and Behavior*, 145:55–65.

Pérez-Escudero, A., Vicente-Page, J., Hinz, R. C., Arganda, S., and De Polavieja, G. G. (2014). idtracker: tracking individuals in a group by automatic identification of unmarked animals. *Nature Methods*, 11(7):743–748.

Pilgrim, R. A. (2017). Munkres' assignment algorithm, modified for rectangular matrices. *Course notes, Murray State University*.

Pnevmatikakis, A. and Polymenakos, L. (2006). Robust estimation of background for fixed cameras. In *Computing, CIC06. 15th International Conference on*, pages 37–42. IEEE.

Poynton, H. C., Varshavsky, J. R., Chang, B., et al. (2007). Daphnia magna ecotoxicogenomics provides mechanistic insights into metal toxicity. *Environmental Science & Technology*, 41(3):1044–1050.

Rashid, M. T., Frasca, M., Ali, A. A., Ali, R. S., et al. (2012). Artemia swarm dynamics and path tracking. *Nonlinear Dynamics*, 68(4):555–563.

Rottmann, R. W., Graves, J. S., Watson, C., and Yanong, R. E. (1992). *Culture Techniques of Moina: The ideal Daphnia for feeding freshwater fish fry*. Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida.

Solis, J., Huang, Y. S., Wlodkovic, D., and Reyes, C. (2015). Microfluidic environment and tracking analysis for the observation of artemia franciscana. *Proceedings of sthe 26th British Machine Vision Conference*.

Tyson, G. E. (1974). Ultrastructure of a spirochete found in tissues of the brine shrimp, artemia salina. *Archives of Microbiology*, 99(1):281–294.

Usami, M. (2004). An ultra-small rfid chip:/spl mu/-chip. In *Advanced System Integrated Circuits. Proceedings of 2004 IEEE Asia-Pacific Conference on*, pages 2–5. IEEE.

Wang, X. Y., Cheng, E., Burnett, I. S., et al. (2017a). Crowdsourced generation of annotated video datasets: a zebrafish larvae dataset for video segmentation and tracking ealuation. In *2017 IEEE Life Sciences Conference (LSC)*, pages 274–277. IEEE.

Wang, X. Y., Cheng, E., Burnett, I. S., Huang, Y. S., and Wlodkowic, D. (2017b). Automatic multiple zebrafish larvae tracking in unconstrained microscopic video conditions. *Scientific Reports*, 7(1):1–8.

Wang, X. Y., Cheng, E., Burnett, I. S., Wilkinson, R., and Lech, M. (2018). Automatic tracking of multiple zebrafish larvae with resilience against segmentation errors. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 1157–1160. IEEE.

Williams, T. A. (1994). A model of rowing propulsion and the ontogeny of locomotion in artemia larvae.

The Biological Bulletin, 187(2):164–173.

Zhiping, X. and Cheng, X. E. (2017). Zebrafish tracking using convolutional neural networks. *Scientific Reports*, 7.

Zhou, Y. Z., Cattley, R. T., Cario, C. L., Bai, Q., and Burton, E. A. (2014). Quantification of larval zebrafish motor function in multi-well plates using open-source matlab® applications. *Nature Protocols*, 9(7):1533– 1548.

Zivkovic, Z. and Heijden, V. D. F. (2006). Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recognition Letters*, 27(7):773–780.

Zivkovic, Z. and Van Der Heijden, F. (2004). Recursive unsupervised learning of finite mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(5):651–656.

Závorka, L., Koeck, B., Cucherousset, J., Brijs, J., Näslund, J., Aldvén, D., Höjesjö, J., Fleming, I. A., and Johnsson, J. I. (2017). Co-existence with non-native brook trout breaks down the integration of phenotypic traits in brown trout parr. *Functional Ecology*, 31(8):1582–1591.