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 algorithms to obtain their movement trajectories is essential to biomedical and pharmaceutical research. However, object mis-detection, segmentation errors and overlapped individual trajectories are particularly common issues that restrict the development of automatic multiple small organism tracking research. Extending on previous work, this paper presents an accurate and generalised Multiple Small Biological Organism Tracking System (MSBOTS), whose general feasibility is tested on three types of organisms. Evaluated on zebrafish, Artemia and Daphnia video datasets with a wide variety of imaging conditions, the proposed system exhibited decreased overall Multiple Object Tracking Precision (MOTP) errors of up to 77.59%. Moreover, MSBOTS obtained more reliable tracking trajectories with a decreased standard deviation of up to 47.68 pixels compared with the state-of-the-art idTracker system. This paper also presents a behaviour analysis module to study the locomotive characteristics of individual organisms from the obtained tracking trajectories. The developed MSBOTS with the locomotive analysis module and the tested video datasets are made freely available online for public research use.

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 algorithms to obtain their movement trajectories is essential to biomedical and pharmaceutical research. However, object mis-detection, segmentation errors and overlapped individual trajectories are particularly common issues that restrict the development of automatic multiple small organism tracking research. Extending on previous work, this paper presents an accurate and generalised Multiple Small Biological Organism Tracking System (MSBOTS), whose general feasibility is tested on three types of organisms. Evaluated on zebrafish, *Artemia* and *Daphnia* video datasets with a wide variety of imaging conditions, the proposed system exhibited decreased overall Multiple Object Tracking Precision (MOTP) errors of up to 77.59%. Moreover, MSBOTS obtained more reliable tracking trajectories with a decreased standard deviation of up to 47.68 pixels compared with the state-of-the-art idTracker system. This paper also presents a behaviour analysis module to study the locomotive characteristics of individual organisms from the obtained tracking trajectories. The developed MSBOTS with the locomotive analysis module and the tested video datasets are made freely available online for public research use.

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), 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). The traditional method relying on human visual observations is very tedious and time consuming. Also, the related experiments are difficult to reliably repeat. Though fluorescent labelling can improve the visual distinction of specific targets, fluorescent materials affect the behavioural response of these organisms (Ekvall et al., 2013).

Automatic object tracking techniques have assisted in developing approaches to the behaviour analysis of large organisms such as mammals, birds and adult fish. However, the tracking of small organisms are hampered by the constraints of existing automatic tracking systems (Ekvall et al., 2013; Dur et al., 2011) as most organisms are considerably smaller than 1 mm (Marechal et al., 2004). There are many challenges imposed by the small sized organisms. Firstly, the widely used transponder (also called u-chips) in individual organism identification dramatically affect the natural behaviour of organisms within millimetre-scales (Ekvall et al., 2013; Lard et al., 2010) when affixing to their bodies, because the currently available smallest chip size is around 0.4 mm (Usami, 2004; Rashid et al., 2012). In addition, general object tracking is already a complex problem due to object occlusion, non-rigid structure

(caused by object rotation and scale changes), and motion pattern changes (Habibi et al., 2017). The challenge increases when the tracking targets are small in size, because small organisms provide very little information compared with imaging noise.

Existing automatic tracking systems for multiple organisms (Zhou et al., 2014; Conklin et al., 2015) either use adult subjects (Pérez-Escudero et al., 2014) in a large container to limit object interaction, or use a petri dish plate to separate individual objects, allowing only one object in each dish well to avoid overlapped and swapped trajectories. The machine learning based CNNTracker (Zhiping and Cheng, 2017) attempts to optimise individual identification accuracy by creating zebrafish head feature maps. However, this system only tested on adult zebrafish, which have very different movement characteristics and higher target-background imaging contrast than small sized larval organisms. Thus, this method may mis-classify larval organisms with multiple identities (Zhiping and Cheng, 2017). The Generalized Linear Mixed Model (GLMM) (Liu et al., 2017) studies the analysis of larval locomotive activities, but can only detect whether movement exists by reporting binary move/no-move classes, but does not enable object displacement estimation.

Accurate object detection and segmentation provides a critical foundation for the performance of the subsequent tracking process. However, the results of existing Multiple Target Tracking (MTT) algorithms degenerate caused by false positive segmentation (i.e., noise fragments remaining after segmentation) and false negative segmentation (i.e. mis-detected objects and occluded objects) (Mallick et al., 2013). Such segmentation challenges commonly occur in microscopic small organism videos when taken under realistic experimental conditions (without deliberate imaging control). It is nearly impossible to identify and segment all small organisms without picking up noise using these microscopic videos (Noss et al., 2013). Non-ideal tracking results further pose challenges for behavioural analysis and maintaining individual identities over time (Martineau and Mourrain, 2013).

IdTracker (Pérez-Escudero et al., 2014) is a well-known biological organism tracking system with 'fingerprint' generation for subject identity differentiation, and the commercial LoliTrack (Závorka et al., 2017) system can also track multiple targets in a single container. However, both systems require the input videos to be taken under strict imaging conditions. As reported by Zhou et al. (2014) and Noss et al. (2013), even small impurities inside water (e.g., water bubbles) or lighting reflections (e.g., surface ripples) affect the object segmentation accuracy.

3D systems with multiple cameras or super-resolution images built from multiple low-resolution images are presented to obtain more information for accurate tracking of small organisms (Ekvall et al., 2013; Günel et al., 2019; Noss et al., 2013). However, these systems increase computational complexity, change data association structure, and require further location registration among cameras or images. Thus, these challenges constrain the real-time application of these techniques in tracking multiple small-scaled organisms (Günel et al., 2019).

Extending on our previous work (Wang et al., 2017a, 2018), this paper presents an automatic and accurate Multiple Small Biological Organism Tracking System (MSBOTS). The approach is robust against non-ideal object detection and segmentation results that are obtained from microscopic time-lapse videos taken under practical laboratory experimental conditions. The system applies Gaussian Mixture Model (GMM) based background subtraction in the segmentation module (Wang et al., 2017a) to detect and segment small organisms from each video frame, and initially maps detected objects between successive video frames based on the (non-ideal) segmentation results using the Hungarian algorithm (also called Kuhn–Munkres algorithm) (Bourgeois and Lassalle, 1971; Munkres, 1957). The positions of mis-detected and overlapped objects are then calculated through their neighbour's locations. Then the theoretically computed locations are bridged in the individual tracking trajectories (Wang et al., 2018). This paper extends on this existing work to present a novel behaviour analysis module to implement locomotive behaviour analysis such as estimating object velocities, accelerations and movement directions. The performance and versatility of MSBOTS is then evaluated and demonstrated by tracking accuracy, and compared with existing multiple organism tracking systems using a zebrafish larvae video dataset (Wang et al., 2017a) as used in our previous work. In this paper, to evaluate the generalisation of the proposed system to track other small-sized organisms, we also apply the system to video datasets of another two types of small organisms *Artemia franciscana*, and *Daphnia magna*).

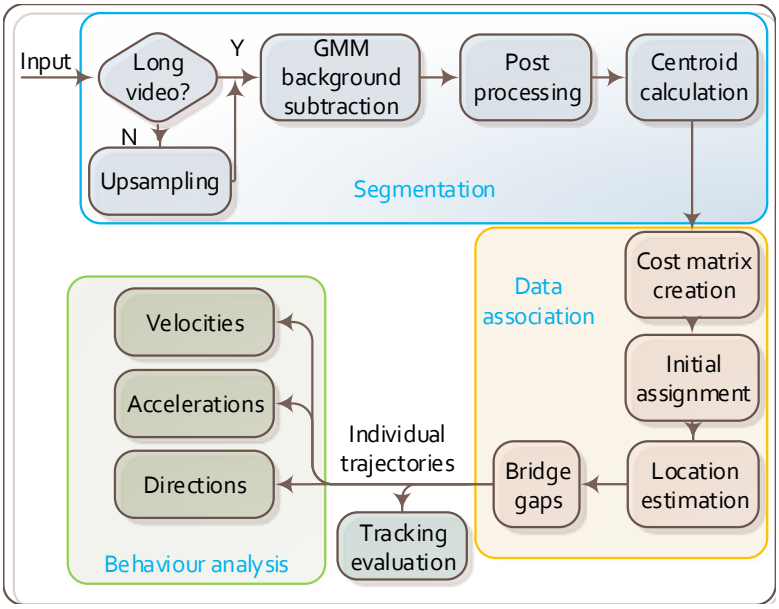

Figure 1. Flow chart of the proposed MSBOTS platform. It takes microscopic time-lapse frames as input, and consists of three main modules: object segmentation, data association based on the centroid locations of detected objects, and the behaviour analysis using the obtained individual trajectories.

METHODS

Fig. 1 outlines the overall workflow of the proposed MSBOTS platform, which comprises an object segmentation module(Wang et al., 2017a), data association module (Wang et al., 2018) and the behaviour analysis module novel to this paper. The accurate differentiation of organisms from the image background and foreign matter (such as water impurities and objects faeces) in each video frame is the critical foundation for this system. Here, the image background is estimated by an adaptive Gaussian Mixture Model (GMM) (Zivkovic and Heijden, 2006) in the segmentation module. Organisms in every video frame are segmented after the background subtraction; details of this segmentation approach can be found in our previous work (Wang et al., 2017a). The following data association module assigns detected objects (as sources, represented by the computed centroids of segmented regions) to their corresponding targets in the successive frames. This part plays an essential role in maintaining consistent individual identities for the detected organisms over time; details of this approach are in Wang et al. (2018). The data association (also called mapping) algorithm not only finds the most likely targets in the following frames for detected organisms, but also calculates the theoretical positions for mis-detected or occluded organisms. After obtaining the individual trajectories of all organisms from the output of the data association module, the movement characteristics of each organism are then estimated in the novel behaviour analysis module proposed in this paper as shown in Fig. 1.

This platform as a whole (as shown in Fig. 1) is tested on a zebrafish video dataset as used in our previous work (Wang et al., 2017a, 2018), and another two new video datasets of two different types of small organisms to evaluate the generalisation of our approach.

Code of ethics

Ethics approval is not required when filming videos of larval organisms. Further, no chemicals were tested with the larval organisms being filmed. The adult zebrafish video analysed is from a publicly available online repository (Pérez-Escudero et al., 2014).

Organism detection and segmentation

In the background subtraction step of the proposed MSBOTS platform, an improved GMM (Zivkovic and Heijden, 2006) was chosen to estimate the stationary background due to its ability to extend the detection period (Wang et al., 2017a). This is particularly effective for addressing the imbalanced movement problem of small biological organisms that occur in time-lapse microscopic video frames. As reported by

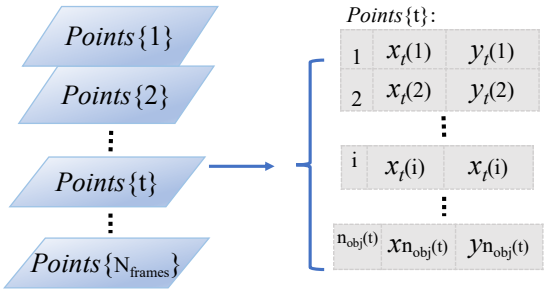

Figure 2. Storage structure of detected objects in a video sequence. The light blue parallelogram series represent the detected organisms within each time-lapse frame. And the cell array $Points\{t\}$ on the right side shows the detailed data storage structure (that is how the individual identities and object Euclidean locations are arranged) of the detected objects in the example frame t.

Liu et al. (2017), small organisms such as zebrafish larvae exhibit a mean proportion of activities less than 7.5% over time with a 'bursty' movement pattern of sudden swimming motion interspersed with mainly stationary periods.

After video background estimation, the moving objects can be initially classified as the foreground organisms. Due to the 'bursty' movement pattern of small organisms, a moving organism will be firstly represented by a new region cluster with a small weight, whose value will be gradually increasing if the region remains stationary over time. This new region will not be classified as the background until its weight value exceeds a threshold c_f , which is calculated as in Zivkovic and Heijden (2006). Thus, the detection period of a stationary organism can be extended for approximately $\log(1-c_f) \setminus \log(1-\alpha)$ frames, where α denotes the constant factor shaping an exponential decay envelope introduced in the GMM work by Zivkovic and Heijden (2006). This enables the detection and tracking of organisms that stop at a position for a certain period of time before restarting their motion.

Furthermore, to adapt to background changes in microscopic videos over time, the adaptive GMM (Zivkovic and Heijden, 2006) adds flexibility when selecting the number of Gaussian components and their parameters. The approach by Zivkovic and Van Der Heijden (2004) and the Dirichlet criteria are applied, respectively, in model initialisation and selection of the number of Gaussian components. The Gaussian component parameters are then adaptively updated for each video frame to accurately represent background pixels, in contrast to traditional GMM models that apply one or a fixed number of components. The foreground objects are then obtained by the differentiation of the video frames from the corresponding estimated background. Furthermore, post-processes such as median filtering, morphological grayscale erosion, and size-based noise removal are applied to eliminate image distortion and scattered noise fragments from the foreground. Within these, the morphological grayscale erosion deploys a flat diamond-shaped structuring element to erode the obtained organism foreground image. Thus, the boundaries among organisms in very close proximity are further widened; thus, the estimated centroid positions of the detected regions are more accurate.

The detection and segmentation of MSBOTS enables the removal of stationary backgrounds (such as the organism container and labels/markers drawn on its surface) and is robust against unbalanced organism movement patterns (e.g., 'bursty' movement) and water impurities. Hence, unlike existing systems, MSBOTS 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 video 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 matrix representation allows for varying element length (indicating the number of detected organisms in frame t), which can change frame-to-frame

due to detection and segmentation errors.

Organism assignment between frames

The centroid locations of detected organisms are obtained in the segmentation module and represented frame-by-frame using a cell array for each video sequence as described in the previous section. However, the organism identities over video frames 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 mis-detected organisms that have been classified into background regions or overlapped with other detected organisms. The data association module of MSBOTS thus builds tracking trajectories of individual organisms by mapping the detected organisms between successive pairs of frames, calculating the positions of mis-detected or occluded organisms, and bridging the theoretical locations back to their correct trajectories by adjusting the initial assignment results.

Initial assignment

The initial frame-by-frame data association of detected organisms is a partial assignment using an extension of the Kuhn-Munkres algorithm (Bourgeois and Lassalle, 1971) to process the rectangular cell array $Points\{t\}$. Extending on the original Hungarian algorithm (Harold, 1955) that solves the assignment problem with an equal number of workers and tasks represented by a $n \times n$ matrix, the number of workers and tasks can be unequal and represented in a rectangular matrix (Bourgeois and Lassalle, 1971). This extended algorithm can then be applied to multiple organism tracking, where the number of detected organisms can change due to non-ideal segmentation resulting from organism mis-detection and occlusion.

In the initial assignment process, all the detected organisms annotated by $Points\{t\}$ 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. The target points $Points\{t+1\}$ need to be mapped to the source points $Points\{t\}$ frame-by-frame across a video sequence. The Euclidean distance of a source point S_i to a target point T_j calculated by Equation (1) is the cost to connect this point pair. A matrix $\mathbb{D}$, as shown in Equation (2), is created to represent the cost of assigning source organisms $S = \{S_1, S_2, S_3, \cdots, S_n\}$ in the frame t to the target organisms $T = \{T_1, T_2, T_3, \cdots, T_m\}$ in the frame $t+1$.

$$d_{S_i,T_j} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \tag{1}$$

$$\mathbb{D}(S,T) = \begin{pmatrix} d_{S_1,T_1} & d_{S_1,T_2} & d_{S_1,T_3} & \cdots & d_{S_1,T_m} \\ d_{S_2,T_1} & d_{S_2,T_2} & d_{S_2,T_3} & \cdots & d_{S_2,T_m} \\ \vdots & \vdots & \ddots & \vdots & \\ d_{S_n,T_1} & d_{S_n,T_2} & d_{S_n,T_3} & \cdots & d_{S_n,T_m} \end{pmatrix} \tag{2}$$

where n and m are the detected number of organisms in the frame t and the successive frame $t+1$, respectively.

The frame-to-frame objects mapping searches for unique assignments in the cost matrix $\mathbb{D}(S,T)$ by connecting the source organism S_i in the frame t to only one target organism T_j in the successive frame $t+1$. The sum of the resultant complete assignments between $Points\{t+1\}$ and $Points\{t\}$ is the global optimum with the lowest overall cost amongst all the possible assignments within two successive frames. Thus, the assignment seeks for the most likely correspondences for all detected source points in the successive frame (combinatorial optimisation). The matched target points propagate the identities of the source points. Thus, connecting the points with the same identities across video frames gives the individual organism tracking trajectories after calculating the assignment maps for all the successive video frames.

It is possible to avoid the propagation of false positive detections within the segmentation results when constructing the individual trajectories. In the initial frame-to-frame assignment step, a distance threshold is set as a constraint in the source-target cost matrix $\mathbb{D}(S,T)$. The threshold value is calculated by $\delta * \text{median}(d_{S_i,T_j})$, as in Zhiping and Cheng (2017). When the minimum value of the i-th row in the source-target cost matrix $\mathbb{D}(S,T)$ is larger than the estimated threshold value, this indicates that the distances between the source point S_i to all of the points in the successive frame exceed the threshold value.

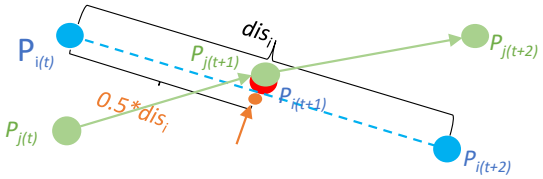

Figure 3. A point calculation example for occluded organisms. The object P_i in the frame $t+1$ as shown by the partial red dot disappears due to occlusion by the object P_j. However, an approximate location of it in the frame $t+1$ annotated by the orange point can be estimated by the relative locations of its correspondences in the neighbour frames t and $t+2$.

Thus, the source point S_i is considered as a segmentation noise fragment to not be further assigned to a target point and its corresponding position information will be removed from the point matrix $Points\{t\}$. This enables the estimation of reliable tracking results, resilient to 'bursty' movement (tested by zebrafish larvae videos as seen in Section 'Tracking accuracy evaluation'), flexible and jerky hopping movement patterns (tested by *Daphnia* videos as seen in Section 'Tracking accuracy evaluation'). Only in one occasion where an organism suddenly moves extremely fast in the same direction (resulting in a large distance change within a short period of time over the distance threshold), this process will incorrectly eliminate the organism as a noisy fragment.

When an organism disappears in frame t due to mis-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 for this video sequence.

Similarly, when an 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 with its previous fame $Points\{t+n-1\}$. To map the points $Points\{t+n-1\}$ to $Points\{t+n\}$, a point in the frame $t+n$ cannot find a source point in the previous frame. This leftover point in frame $t+n$ is saved in the unmatched target matrix for the same video sequence.

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.

Position estimation for mis-detected and occluded organisms

Fig. 3 illustrates the location computation for the overlapped organism P_i in the frame $t+1$. The two points P_i (shown by blue dots) and P_j (shown by green dots) overlap with each other in the frame $t+1$, and this overlapped point at time $t+1$ is assigned to the object P_j in the initial mapping process. Since the point P_i in the frame t cannot find a target in the frame $t+1$, and this point in the frame $t+2$ cannot be assigned to a source point in frame $t+1$, this point is classified as an unmatched source point in frame t, and an unmatched target point in frame $t+2$, respectively. The position of the missing point due to occlusion (as shown by the red dot occluded by the green dot in Fig. 3) or mis-detection (for example, when the red dot is completely occluded by the green dot in Fig. 3) is calculated using the locations of the unmatched source point $P_i(t)$ and unmatched target point $P_i(t+2)$. In this example, as shown in Fig. 3, the location of the missing point P_i in frame $t+1$ is calculated by the median point between $P_i(t)$ and $P_i(t+2)$.

When there are multiple unmatched pairs, the mapping from unmatched source points to the unmatched target points is also according to the extended Hungarian assignment algorithm (Bourgeois and Lassalle, 1971). In this step, the unmatched source points firstly search for possible correspondences in the following $2nd$ 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 bridged to its subsequent tracking fragment so long as the frame separation is less than 6 video frames. Thus, the default search range in MSBOTS is from the $2nd$ to the $6th$ frame following the frame when the mis-detection and occlusions originally occur. The positions of missed organisms (x_{miss}, y_{miss}) are calculated by Equations (4) and (5) using the matched point pair from the the locations of unmatched source point (x_s, y_s) and an unmatched target point (x_t, y_t).

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

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

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

Bridging trajectory gaps for mis-detected and occluded organisms

Individual tracking trajectories for each organism are obtained through connecting the matched points with the same identities 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 mis-detected or overlapped, caused by segmentation errors. In the proposed MSBOTS, these trajectory gaps are bridged by adding the estimated points of these mis-detected or overlapped organisms as 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 matched unmatched-source to unmatched-target pairs are also calculated during this mapping process. For example, as shown in Fig. 3, the unmatched source $P_i(t)$ as the end point of its trajectory fragment is connected to the unmatched target point $P_i(t+2)$, which is the start point of its trajectory fragment, and the middle point shown by the orange dot is added between points $P_i(t)$ and $P_i(t+2)$ as the theoretical position of the overlapped point $P_i(t+1)$.

Locomotion characteristic analysis

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

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

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

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

where $dt = 1/fs$ and fs is the video frame rate, and $x_t(i)$ and $y_t(i)$ are the Cartesian coordinates of organism i (i is the assigned organism identity) in frame t.

RESULTS AND DISCUSSION

To evaluate the performance of MSBOTS, microscopic videos of three types of small biological organisms (zebrafish, *Artemia* and *Daphnia*) were tested, where preliminary evaluation using zebrafish was presented in our previous work (Wang et al., 2018; Wang, 2018). Evaluated on a set of single and multiple larvae and adult zebrafish, *Artemia* and *Daphnia* videos here, a wide variety of (complex) imaging conditions were tested, including shadowing, labels (manually drawn on the petri dish), and background artefacts (such as water impurities, object faeces and water bubbles of varying sizes). No chemical stimuli were tested on the studied organisms in this work, so their behaviour analysed corresponds to natural response. 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 datasets 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, *Artemia* and *Daphnia* were applied to evaluate the proposed MSBOTS platform. Both low frame rate videos (of 14 or 15 fps recorded by a AD7013MT Dino-Lite microscope) and high frame rate videos (captured by an UI-3360CP-C-HQ microscope) are tested in this work.

Wild zebrafish (*Danio rerio*) embryos were incubated at 28°C in a petri dish filled with an E3 medium. Zebrafish larvae were obtained from the hatched embryos five days post-fertilization. Zebrafish larvae were transferred to round poly (methyl methacrylate) (PMMA) housing wells for shooting microscopic time-lapse videos. The zebrafish dataset consists of 10 video sequences with 3056 frames in total: nine zebrafish larvae videos captured as prescribed here and one adult zebrafish video provided by Pérez-Escudero et al. (2014) from their publicly available online repository ('Example video of 5 zebrafish' at http://www.idtracker.es/download). The size of the zebrafish larvae video frames is 960 pixels x 1280 pixels, and the average zebrafish larvae size is around 1900 pixels. More details about the larvae zebrafish dataset and its generation can be found in our previous work (Wang et al., 2017b).

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 as per 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 with the temperature maintained 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, video sequences 1-5), with Fig. 4b a microscopic video frame example of *Artemia franciscana* in 480 pixels x 640 pixels. The average size of an *Artemia franciscana* object is approximately 500 pixels. *Daphnia magna* microscopic videos containing 10 organisms and artifacts (bubbles and impurities of different sizes, video sequences 1-5), with Fig. 4c a microscopic video frame example of *Daphnia magna* in 480 pixels x 640 pixels. The average size of *Daphnia magna* tested is approximately 400 pixels. Both the *Artemia* and *Daphnia* videos were ordered randomly to test the flexibility and feasibility of the proposed system.

Tracking evaluation metrics

To objectively and quantitatively evaluate the tracking performance of the proposed MSBOTS platform, the widely utilised standard metric, Classification of Events, Activities and Relationships (CLEAR MOT) (Bernardin and Stiefelhagen, 2008), for Multiple Object Tracking (MOT) is employed in this paper. CLEAR MOT comprises two metrics: Multiple Object Tracking Precision (MOTP) as presented by Equation (9) and Multiple Object Tracking Accuracy (MOTA) as shown by Equation (10). MOTP measures the position precision of all segmented organisms compared to that of the manually labelled ground truth in every video frame, whereas MOTA estimates the individual trajectory accuracy (the ability to produce exactly one trajectory per organism with a consistent label over time).

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

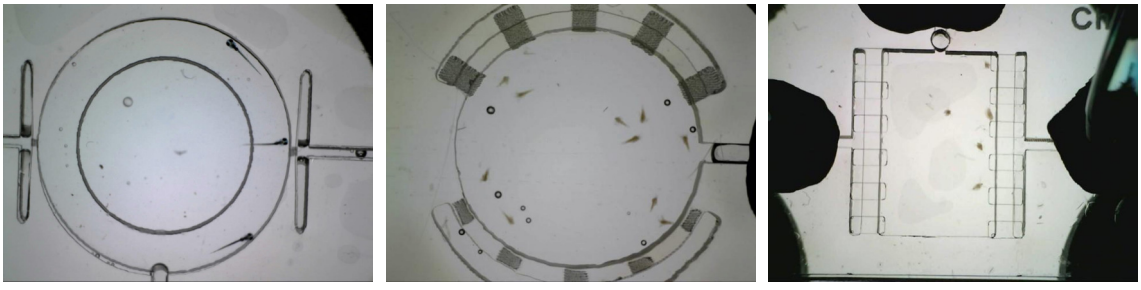

Figure 4. Microscopic video frame examples of (a) zebrafish larvae (960 pixels x 1280 pixels), (b) *Artemia franciscana* (480 pixels x 640 pixels), and (c) *Daphnia magna* (480 pixels x 640 pixels). The average sizes of zebrafish larvae, *Artemia* and *Daphnia* tested are approximately 1900, 500 and 400 pixels, respectively.

$$MOTA = 1 - \frac{\sum_t (M_t + FP_t + SII_t)}{\sum_t g_t} \tag{10}$$

where $|P_{i,t} - GT_{i,t}|$ is the Euclidian distance between the estimated centorid position of the i-th detected organism $P_i(t)$ in the frame t. Its position in the manually labeled ground truth is denoated as $(GT_{i,t})$ and N_t indicates the total number of segmented organisms in the ground truth in frame t. The smaller the MOTP value, the more precise the segmentation result.

MOTP measures the organism segmentation accuracy compared with the ground truth, whereas the MOTA metric emphasises the evaluation of the individual tracking trajectories, where M_t, FP_t, and SII_t indicate the number of missed detections, false positive detection (i.e., image noise fragments segmented as organisms), and the Swapping of Individual Identities (SII), respectively, in the frame t. And g_t means the total number of organisms detected in frame t. The ideal value of MOTA is 1, and the value decreases with the occurrence of detection errors and identity swapping. Comparing with the total number of detected organisms, the MOTA value will generate negative numbers (as shown in Fig. 5a) when the combination of detection errors and identity swapping is high, which implies low system reliability and the resultant tracking trajectories should be considered as unreliable.

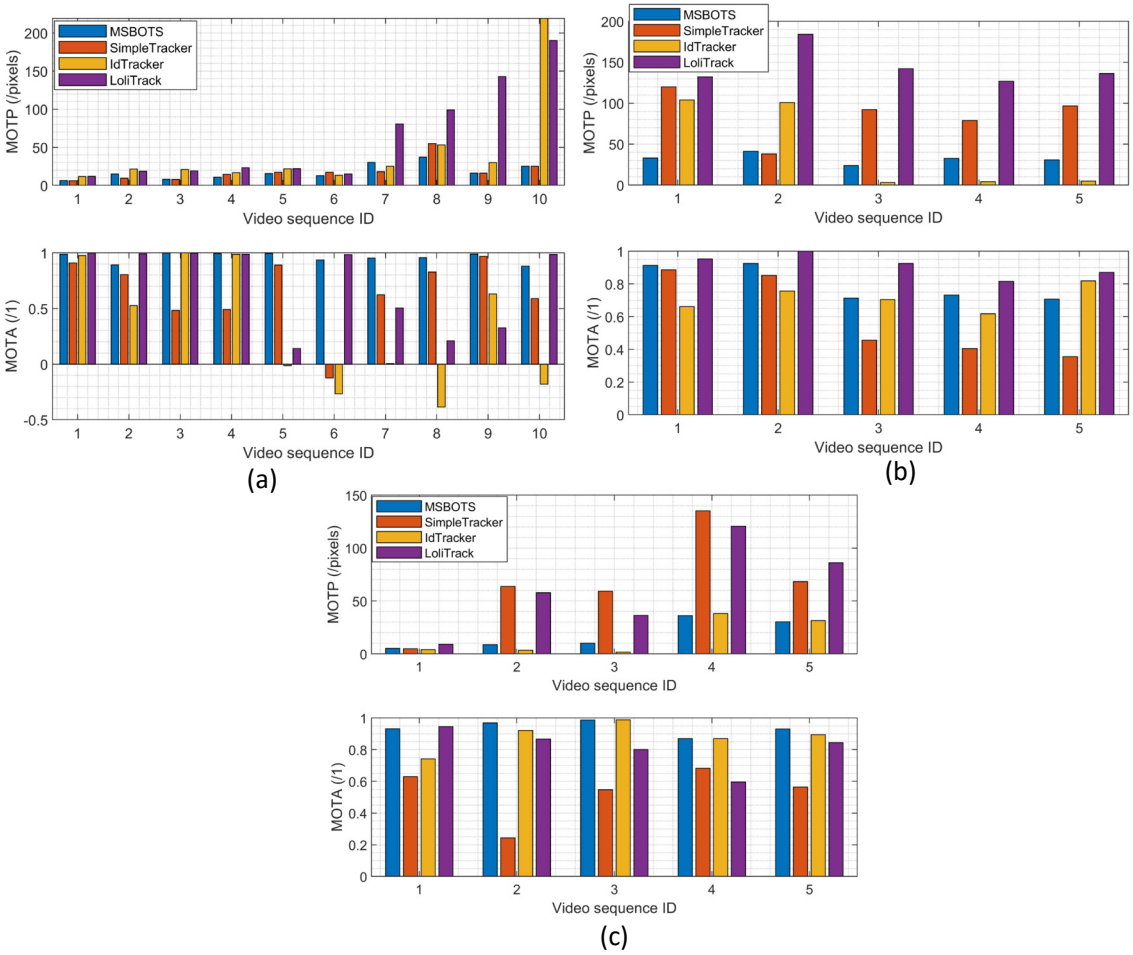

Figure 5. Evaluation of tracking results comparing the proposed system MSBOTS to existing systems - SimpleTracker (Bourgeois and Lassalle, 1971), idTracker (Pérez-Escudero et al., 2014) and LoliTrack (Závorka et al., 2017) - using videos of three types of small biological organism. (a), (b) and (c) show the estimated MOTP and MOTA values for zebrafish, *Artemia* and *Daphnia* video datasets, respectively. MOTP values measure the segmentation performance, and the smaller the value the higher the segmentation accuracy. MOTA compares the accuracy of the resultant individual tracking trajectories, with a higher value denoting a more accurate result.

Tracking accuracy evaluation

To evaluate the proposed tracking system, its tracking performance over the microscopic video dataset is compared with the well-known multiple object tracking platform 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 organisms), and the off-the-shelf commercial LoliTrack system (Závorka et al., 2017). Fig. 5 shows the estimated tracking accuracy measured by the MOTP and MOTA metrics, with Fig. 5a showing results from zebrafish videos, Fig. 5b showing *Artemia* videos, and Fig. 5c showing *Daphnia* videos.

It can be seen from Fig. 5a (the upper MOTP figure) 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% increase) compared to SimpleTracker, idTracker, and LoliTrack, respectively. The improved organism position detection results demonstrate the accuracy 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. Supplementary Table S1 summarises the estimated tracking accuracy measured by MOTP and MOTA values across all the zebrafish video sequences.

All of the methods performed well for the position accuracy of detected organisms when the videos had a clear background, as shown by sequences 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 sequences 7-10 in Fig. 5a. Accordingly, the proposed MSBOTS method still out-performed the existing approaches when taking into account mis-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 mis-detected or overlapped organisms using their neighbour position knowledge from the segmentation results in the proposed MSBOTS platform. In addition, the bridging of trajectory fragments in MSBOTS based on the extended Hungarian assignment algorithm (Bourgeois and Lassalle, 1971) when there are multiple unmatched trajectory fragments decreased the possibility of individual identity swapping, compared with SimpleTracker, which only used a distance metric by nearest neighbour algorithm (which in turn can generate identity swapping during the gap bridging process).

In addition, Fig. 5b and Fig. 5c show the tracking accuracy evaluation using *Artemia franciscana* and freshwater *Daphnia magna* videos, respectively, to test the generalisation of the proposed MSBOTS system on other small organisms with movement characteristics different to zebrafish larvae. *Artemia* display flexible movement and vary according to surrounding fluidics (Tyson, 1974; Williams, 1994), and *Daphnia* exhibits 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. The detailed data on tracking accuracy measured by MOTP and MOTA values can be seen in Supplementary Table S2 & S3.

As can be seen from the MOTP values in Fig. 5b, the proposed method exhibits 47.68 pixels smaller standard deviation than the idTracker system, which illustrated the usability of the proposed system on *Artemia* microscopic videos and the enhanced ability on *Artemia* detection accuracy. Though idTracker can produce smaller organism position estimation errors (MOTP) as shown by sequences 3-5 in Fig. 5b, the mean MOTA value is 7.07%, and 6.44% less than the proposed MSBOTS and LoliTrack, respectively, which illustrates the existence of a similar detection problem when testing zebrafish larvae. That is, the organism is detected as background, and impurities as an organism, due to their small size differentiation and similar movement characteristics when the organism stops moving or the water impurities are stirred up by organism movement; this further causes identity confusion.

Fig. 5c shows the overall tracking accuracy of the proposed MSBOTS system and idTracker consistently outperformed the other two systems under comparison, measured by MOTP and MOTA. However, the mean MOTA value of the proposed MSBOTS system is still 5.48% higher than idTracker. The generalised tracking accuracy of the proposed MSBOTS system is further illustrated by applying it to *Daphnia* video sequences with short, jerky hopping movement characteristics, compared with these existing systems.

Analysis of organism movement characteristics

To explore the capability of the proposed MSBOTS approach in analysing organism movement characteristics, this paper presents the estimation of movement velocity, acceleration and direction calculated from the individual tracking trajectories.

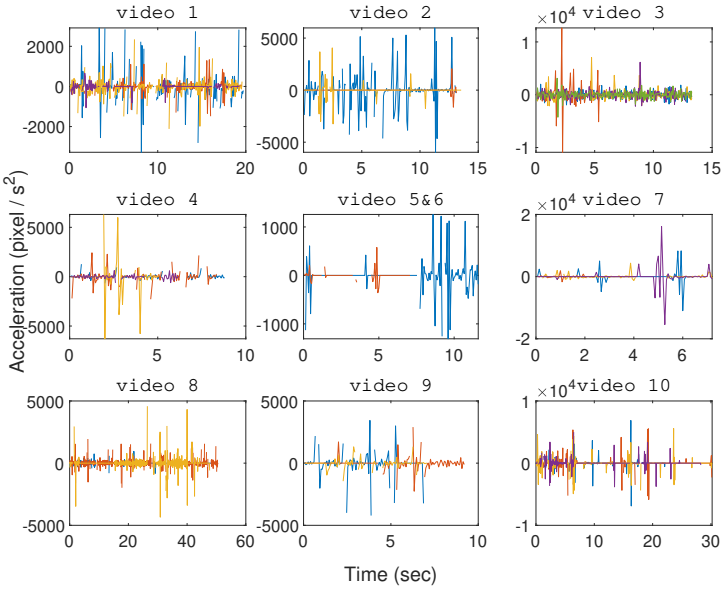

Figure 6. Organism movement acceleration over time for every tested zebrafish video (one video per subplot), with the centre subplot combining the results of videos 5 and 6 housing a single zebrafish larvae each. The line colors represent different zebrafish individuals in every video.

Fig. 6 shows the movement acceleration analysis for each zebrafish video using the resultant tracking trajectories generated by the proposed MSBOTS. 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 of movement speed (shown by the computed acceleration values) that is obviously visually perceptible to the human eye occurred only within approximately 10 seconds, as tested in the 10 adult and larvae zebrafish videos. This is due to the natural (anxious) response of zebrafish when turning on the imaging camera (Peng et al., 2016)). The zebrafish movement speeds 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 direction changes can provide insight into the interaction rules (Hinz and de Polavieja, 2017), Fig. 7 shows an example of the calculated velocity and movement direction results for each zebrafish video, respectively. The equal number of boxes with the number of zebrafish housing in each video show that all individuals were successfully assigned with one identity (which also illustrated the one-to-one organism mapping accuracy 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 can also be easily obtained from the visual box plot shown in Fig. 7.

There is no specific speed requirement for the organisms that can be tracked by the MSBOTS platform. Both organisms travelling at high speed or average low speed as illustrated by Fig. 6 and Fig. 7a, respectively, can be accurately detected and tracked over time by the system. Thus, this system can be used for automatic tracking, comparison and analysis of small organisms in natural response or under the exposure of testing chemicals. The recommended lowest frame rate of time-lapse videos is 14 f/s as in the tested dataset from trial-and-error comparison.

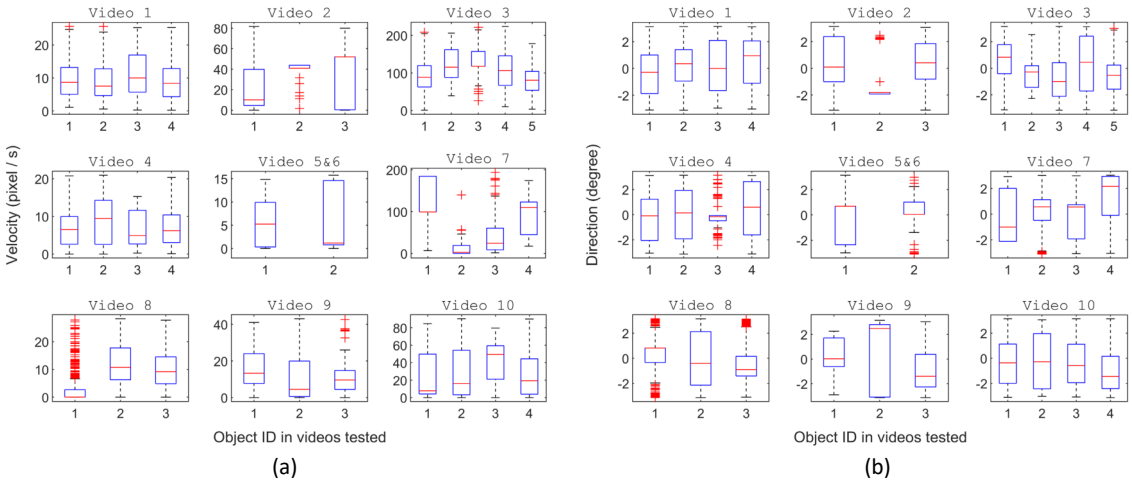

Figure 7. Analysis of zebrafish locomotion characteristics across video sequences, where (a) and (b) show movement speed and direction analysis of individual organisms, respectively, in very video frame (the centre subplots combines the results of single zebrafish larvae in videos 5 & 6).

CONCLUSION

Accurate automatic tracking for multiple small biological organisms provides an efficient approach for many biomedical and ecotoxicity applications. However, organism mis-detection and occlusion are inevitable problems when detecting and segmenting these small biological organisms from time-lapse microscopic videos. The detection and segmentation becomes more challenging 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, extending on and improving our previous work, this paper presents a Multiple Small Biological Organism Tracking System (MSBOTS), combining a multiple object association algorithm for linking detected objects frame-by-frame and tracking trajectory adjustment techniques. To address segmentation errors due to mis-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 organisms with variant movement characteristics, using 20 videos in total for evaluation. The resulted tracking accuracy of the proposed system outperformed three existing (state-of-the-art or commercial) tracking systems. Moreover, this system also provides locomotive characteristic analysis using the generated individual tracking trajectories to facilitate small organism behaviour analysis research. Behavioural rules and new medicine or chemical effects on the dynamic behaviour of organisms can thus be investigated using the proposed behaviour analysis module, enabling an automatic and quantitative movement analysis to the related 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 and https://github.com/Xiao-ying/moving-zebrafish-larvae-segmentation-and-tracking-dataset-/tree/master/Data. The analysed adult zebrafish video was downloaded from the publicly available online repository http://www.idtracker.es/download (the 'Example video of 5 zebrafish'). The developed MSBOTS system and the software to analyse locomotive characteristics of individual organisms can also be downloaded from the same repository.

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., G., M. F., S., M. M., and S., H. J. (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., J., C., P., P. R., and P., F. (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., L., W. A., C., S. E., D., B. J., M., K., B., B. M., and E., D. J. (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., and F., L. Y. (2017). Statistical analysis of zebrafish locomotor behaviour by generalized linear mixed models. *Scientific Reports*, 7.

Mallick, M., Vo, B. N., 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., Y., Z., N., G., and Q., L. (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.

Poynton, H. C., Varshavsky, J. R., Chang, B., G., C., S., C., S., H. P., V., L. A., J., B. D., K., K., C., T. E., E.J., P., O., H., and D., V. C. (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., L., F., and G., X. M. (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. (2018). *computer vision techniques applied to automatic biological organism segmentation and tracking*. PhD thesis, School of Engineering.

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

Wang, X. Y., Cheng, E., Burnett, I. S., S., H. Y., and D., W. (2017b). 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., 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.