

Shape Normalizing and Tracking Dancing Worms

Carmine Sansone^{1,2}, Daniel Pucher¹(✉), Nicole M. Artner¹(✉),
Walter G. Kropatsch¹, Alessia Saggesi², and Mario Vento²

¹ Pattern Recognition and Image Processing Group, TU Wien, Vienna, Austria
{daniel.pucher,artner,krw}@prip.tuwien.ac.at

² Department of Information Engineering, Electrical Engineering and Applied Mathematics (DIEM), Faculty of Engineering, University of Salerno, Fisciano, Italy

Abstract. During spawning, the marine worms *Platynereis dumerilii* exhibit certain swimming behaviors, which are described as *nuptial dance*. To address the hypothesis that characteristic male and female spawning behaviors are required for successful spawning and fertilization, we propose a 2D tracking approach enabling the extraction of spatio-temporal data to quantify gender-specific behaviors. One of the main issues is the complex interaction between the worms leading to collisions, occlusions, and interruptions of their continuous trajectories. To maintain the individual identities under these challenging interactions a combined tracking and re-identification approach is proposed. The re-identification is based on a set of features, which take into account position, shape and appearance of the worms. These features include the *normalized shape* of a worm, which is computed using a novel approach based on its distance transform and skeleton.

Keywords: Object tracking · Appearance models · Shape normalization · Shape analysis

1 Introduction

Platynereis dumerilii [19] are marine worms, who reproduce through external fertilization when sexually mature. During spawning they exhibit particular swimming behaviors, which differ based on their gender and spawning phase [6]. The aim of the biologists is to verify that characteristic male and female spawning behaviors are required for successful spawning and fertilization. Therefore, they want to analyze the spawning behaviors in a quantitative manner and characterize and compare male- and female-specific behaviors. To achieve this, the aim is to develop methods that enable the 2D tracking of spawning worms and extract features describing their appearance (skeleton, curvature, normalized shape, etc.) and motion (trajectories of head and tail). These features are described and discussed in more detail in [13].

For the proposed spatio-temporal analysis, the nuptial dance [3] of the worms is recorded with the help of a special setup. In nature, spawning happens in the sea at night around new moon. Therefore, it is necessary to record the videos inside a light-tight box with an infrared camera. During the recordings, the worms are placed into a shallow, cylindrical bowl filled with sea water referred to as *arena*. On the water surface, the arena has a diameter of 7 cm. The average length of a worm is between 1 and 2 cm.

When the worms are not close to each other, analysis and tracking is straightforward. As the camera is fixed, the worms can be reliably segmented by subtracting a background image (showing the empty arena) and removing noise (morphological filters). The resulting binary regions of both worms are represented by their center line (extracted based on skeleton) limited by head and tail position. Tracking of the worms is done by employing a Kalman filter, which finds the correspondences of head and tail for both worms in each frame.

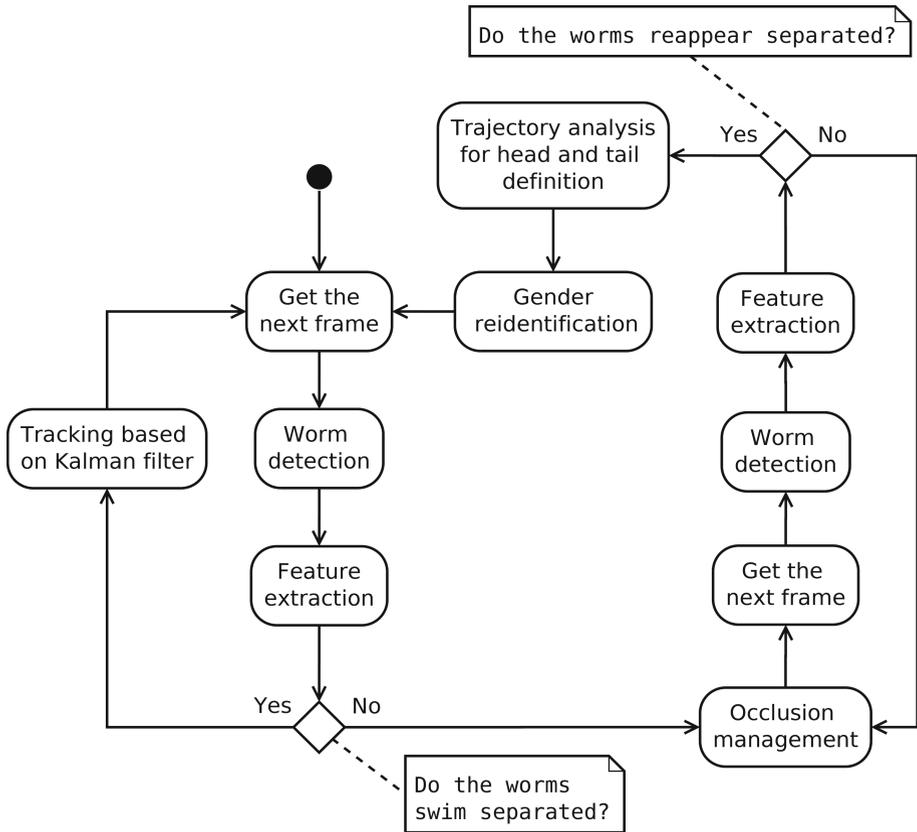


Fig. 1. Flow chart of the proposed method

The contribution and focus of this paper lies in the active phase of the spawning process, where the worms interact. During this active phase the worms collide and occlude each other in the 2D image, thus leading to ambiguities in segmentation and tracking. As the worms frequently change their appearance and motion, it is a challenging task to correctly identify them throughout the whole video sequence. The main aim of this paper is to re-identify the worms after their frequent interactions. This ensures that the collected spatio-temporal data is associated with the correct label (male or female) and is a reliable source of information for the following analysis of the biologists. Figure 1 gives an overview of the overall tracking pipeline.

The remaining part of this paper is organized as follows: Section 2 discusses existing tracking methods, Sect. 3 explains shape normalization in detail, Sect. 4 presents the tracking and re-identification method. Experimental results are shown in Sect. 5 and Sect. 6 concludes the paper.

2 State of the Art

There is a vast amount of research in tracking people [11] and vehicles [5], but only modest attention has been devoted to tracking worms. Traditional tracking methods cannot be directly used to track swimming worms as the worms change their direction and speed of motion in a fast and unpredictable way. Furthermore, the appearance of the worms changes rapidly, thus the history of the moving objects cannot be used for increasing the reliability of the tracking, as in [11].

As summarized in [9], a number of worm trackers focus on the tracking and feature extraction of one worm [2, 15, 17, 18] or a group of worms [14, 16]. The proposed methodologies use background subtraction and track the centroids of all foreground regions. When two or more blobs occlude each other, the information about these blobs is lost since the focus of these algorithms is on the behavior of the worm groups and not on the exact trajectory of a single worm.

Huang [8] describes an approach, where it is possible to keep track of the worms while they occlude each other. In contrast to Huang's approach, this paper does not try to segment the worms during occlusion, but aims at re-identifying them afterwards. They also propose a method to define the head and the tail points of each worm based on the fact that the *C. elegans* worms have an accumulation of fat in the head making the head area usually brighter than the tail area, which does not hold for *Platynereis dumerilii* worms. In the work of Hoshi and Shingai [7] another method to detect the worm's head and tail is proposed. This method cannot be applied to our problem because it is based on the assumption that the head swings more than the tail.

3 Shape Normalization

As the worms are highly deformable, it is necessary to come up with a description of their shape, which is independent of deformations. For this, we follow a recent strategy, which is known as co-registration, where shapes are first straightened or

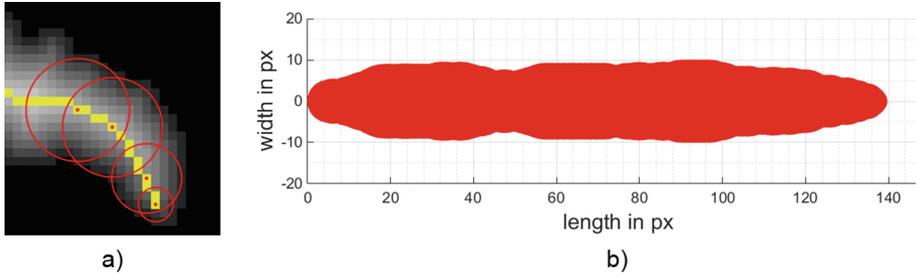


Fig. 2. (a) Four circles (red) representing the shape at four selected positions in the skeleton (red points). Yellow pixels represent the skeleton. Gray values visualize the distance to the boundary. (b) Normalized shape representation of a male worm. (Color figure online)

flattened to then register different views/deformations of the same normalized shape [1]. The shape normalization is based on the distance transform of the binary region and the skeleton of the shape. The skeleton is restricted to a line delimited by two end-points without any branches. For every point of the skeleton, the corresponding value of the distance transform holds the Euclidean distance to the nearest boundary pixel of the binary region. These distances hold information on the thickness of the shape and also serve as radii of circles used to draw the normalized shape representation. Figure 2a shows the circles of four points on the skeleton of a worm.

For every pixel of the skeleton a corresponding circle can be drawn. The union of these circles covers the original worm shape. By arranging the skeleton points in a straight line, the representation becomes independent of the deformations of the worm and gives the normalized shape (See Fig. 2b). To maintain the correspondence between deformed and normalized shape, the geodesic distance along the deformed skeleton is mapped to the Euclidean distance in the normalized shape. This one-to-one mapping also allows to return to the image space. In this paper, the normalized shape is used to compare worm shapes before and after an occlusion occurs.

4 Worm Tracking

4.1 Representation of Worms

Let $\omega(\tau, z, id)$ be a function that defines the position of a point along the worm's skeleton, where $\tau \in [0, 1]$ is the percentage of the worm length that defines the normalized distance along the skeleton between the head point $\omega(1, z, id)$ and the tail point $\omega(0, z, id)$. z represents the frame number of the video and id is the worm's identifier. In case of two worms $id \in [male, female]$. For a fixed frame \hat{z} of the video and for a fixed worm \hat{id} , the position of the tail point is $\omega(0, \hat{z}, \hat{id})$ and the position of the head point is $\omega(1, \hat{z}, \hat{id})$.

However, since the trajectory of the head and the tail of a worm is not smooth but jittery due to flicking and bending motions, it is not possible to correctly predict the position of these points using the Kalman filter [10]. Thus, the skeleton is partitioned into three parts, namely tail points $\tau \in [0, \frac{1}{3}]$, body points $\tau \in [\frac{1}{3}, \frac{2}{3}]$, and head points $\tau \in [\frac{2}{3}, 1]$, and the following two points having a smoother trajectory are introduced: $\omega(\frac{1}{3}, z, id)$ and $\omega(\frac{2}{3}, z, id)$.

4.2 Tracking Based on Kalman Filter

During a preliminary detection step based on a traditional background subtraction algorithm, the connected regions associated to the worms are detected and their skeletons are extracted [13]. Each skeleton is then represented by the two points p_1 and p_2 , having a geodesic distance of $\frac{1}{3}$ and $\frac{2}{3}$ from one of ending points of the skeleton, respectively. At the generic frame $\hat{z}+1$, the positions of the worms $w(\frac{1}{3}, \hat{z}, male)$, $w(\frac{2}{3}, \hat{z}, male)$, $w(\frac{1}{3}, \hat{z}, female)$ and $w(\frac{2}{3}, \hat{z}, female)$ are known from the previous frame \hat{z} . In the first frame of a video, the positions are defined manually. The tracking algorithm has to find the correct association among the above positions identified at the frame \hat{z} (the output of the tracking at the previous frame) and the points p_1 and p_2 defined in frame $\hat{z}+1$ (the output of the detection).

The Kalman filter provides predictions for the positions of the points of frame \hat{z} based on the trajectory of these points [10]. These predictions are compared with the points that describe the worms in frame $\hat{z}+1$. In the case of two worms, there are just eight possible associations. For every prediction, the corresponding error is calculated as the Euclidean distance between the position of the prediction and the position of the predicted point in frame $\hat{z}+1$. For every hypothesis four errors are taken in account: the errors of the head and tail predictions of the male worm and the errors of the head and tail predictions of the female worm. The hypothesis with the lowest root mean square error is chosen because the prediction errors follow a norm distribution [4]. Furthermore the evaluation showed that using the root mean square error brings better results.

4.3 Handling Occlusions by Re-Identification

An occlusion is defined as a set of frames in which the worms overlap each other and appear as a single connected region in the binary image. In the frames before an occlusion, a set of features is extracted to describe each worm, while during an occlusion it is not possible to reliably extract and associate these features. After an occlusion, there are again two connected regions: $region_1$ and $region_2$, each describing a worm, but their identities (male or female) are unknown. Furthermore, it is unknown which of the two end-points of each skeleton is the head and which is the tail. When the worms occlude each other, they often change the direction of their movement in an unpredictable way. Information about the movement before the occlusion can't be robustly used to predict the position of the worms after the occlusion.

As motion (speed and direction) is not a reliable indicator for the identity of the worms, this paper proposes an approach based on comparing the appearance immediately before and after the occlusion. For this comparison, the following features are taken into account:

- Normalized Shape, $f^s \in \mathbf{R}^N$
- Area in pixels, $f^a \in \mathbf{N}$
- Mean gray-scale value of skeleton, $f^m \in [0, 255]$
- Length of skeleton in pixels, $f^l \in \mathbf{R}$

A sliding window is used to analyze the worms over several frames to increase the reliability in the classification process. The features collected before the occlusion are used to build two models: $model_1$ for the male worm and $model_2$ for the female worm. To compare the worm models with the features extracted from the unidentified regions after the occlusion, a measure of similarity (see Eq. 1) and a method of comparison are necessary.

After the occlusion, a majority voting approach is used to combine the decisions taken in every frame of the temporal window to find the identity of the worms (male or female). The identity that receives the most votes is the final identity of each worm.

The size of the sliding window has to be chosen according to the rate in which the chosen features change over time. For *Platynereis dumerilii* worms, some of the features depend on how and where the worms are moving. When they swim in a straight line, the body area is bigger than when they are moving in circles. The luminosity also changes depending on where they swim. It is higher in the middle of the arena than at the border. Therefore a *temporal window* of five frames was chosen empirically in the current tracking setup.

To build the male and female model, the average value of the features in the last five frames before an occlusion is calculated. The models define how the worms should appear after the occlusion. For each of the five frames after the occlusion, the features of the two unidentified regions are compared with the male and female models. For the comparison, the following measures of similarity are computed: s_{ij}^s (normalized shape), s_{ij}^a (area), s_{ij}^m (mean gray scale value) and s_{ij}^l (length), where $i \in [1, 2]$ is an index referring to the two unidentified regions and $j \in [1, 2]$ is an index referring to the models. All these measures of similarity are normalized to one and combined according to Eq. 1.

$$s_{ij} = \sqrt{(s_{ij}^s)^2 + (s_{ij}^a)^2 + (s_{ij}^m)^2 + (s_{ij}^l)^2} \quad (1)$$

Let $x \in \{a, m, l\}$ be a variable that represents the value of one of the following features: area, mean gray scale value and length. d_{ij}^x is the difference between the value of the feature x of region i and the value of the same feature in model j (see Eq. 2). s_{ij}^x is computed according to Eq. 3 where the features are normalized with the maximum feature value d_{max}^x which is defined empirically.

$$d_{ij}^x = f_i^x - f_j^x \quad (2)$$

$$s_{ij}^x = 1 - \frac{|d_{ij}^x|}{d_{max}^x} \quad (3)$$

The normalized shape similarity s_{ij}^s is computed according to the Pearson correlation [12] (see Eq. 4), where N is an integer value computed as the minimum value between the geodesic length of the skeleton from $region_i$ and the geodesic length of the skeleton of $worm_j$. $r_i(n)$ and $r_j(n)$ are two functions that define the radius of the normalized shape for the $region_i$ and for the $worm_j$ of a fixed point n on the skeleton. Comparing normalized shapes 3D movements are managed.

$$s_{ij}^s = \frac{N \cdot \sum_{n=1}^N r_i(n) \cdot r_j(n) - \sum_{n=1}^N r_i(n) \cdot \sum_{n=1}^N r_j(n)}{\sqrt{N \cdot \sum_{n=1}^N r_i^2(n) - (\sum_{n=1}^N r_i(n))^2} \cdot \sqrt{N \cdot \sum_{n=1}^N r_j^2(n) - (\sum_{n=1}^N r_j(n))^2}} \quad (4)$$

There are two possible associations: $region_1$ is the male worm and $region_2$ is the female worm (hypothesis H_{p1}) or vice versa (hypothesis H_{p2}). A measure of confidence is assigned to both hypotheses: $s_{11} + s_{22}$ to H_{p1} and $s_{12} + s_{21}$ to H_{p2} . The association with the greater confidence value is chosen as result of the re-identification process.

From Eq. 1 it can be observed that all the features bring an equal contribution to the final decision. From the evaluation it has been observed that there is no need to weight the features.

4.4 Trajectory Analysis for Head and Tail Definition

After the re-identification, the regions are labeled as male and female, but the head and tail of each worm is not identified yet. To solve this open issue, a trajectory analysis is done. The basic idea is to analyze the direction of the motion of the worms, as they in general swim in a forward manner (tail point on skeleton is pulled from head point).

A new predefined temporal window after the occlusion is considered for the trajectory analysis. In each frame, the two end-points of the skeletons of both worms are tracked as shown in Fig. 3. In the synthetic example in Fig. 3, the Euclidean distance between the point $w(\frac{1}{3}, \hat{z} + 1, \hat{id})$ and the point $w(\frac{1}{2}, \hat{z}, \hat{id})$ is smaller than the Euclidean distance between the point $w(\frac{2}{3}, \hat{z} + 1, \hat{id})$ and the point $w(\frac{1}{2}, \hat{z}, \hat{id})$. Therefore, the point $w(0, \hat{z} + 1, \hat{id})$ votes for tail in this frame pair. This voting is repeated for every consecutive frame pair in the temporal window. Finally, the majority of votes decides on head and tail of each worm.

For the trajectory analysis it is assumed that the worms do not move backwards for more than half of the frames in the temporal window. The temporal window for the trajectory analysis has to be chosen considering the trade off between the probability of making the correct decision and the elaboration time. Therefore, a window of 30 frames was chosen empirically related to speed and longest backward movement of *Platynereis dumerilii* worms in the current tracking setup. Figure 4 shows an example of occlusion management.

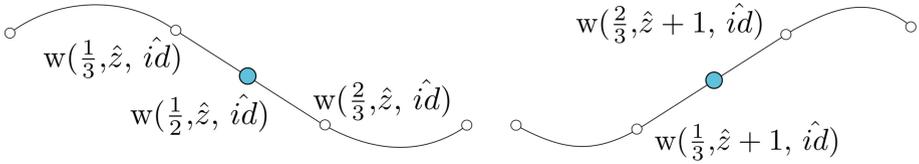


Fig. 3. Information taken in account to identify the head and the tail points in a frame pair (left and right).

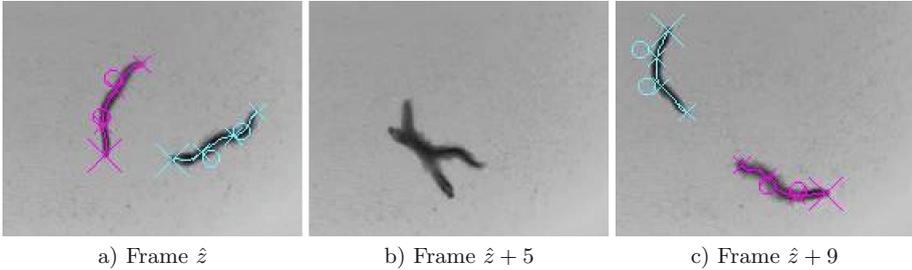


Fig. 4. Three frames taken out of a sequence of ten frames. (a) Male (blue) and female (pink) worm before occlusion. Small crosses identify the points $w(0, \hat{z}, \hat{id})$, $w(\frac{1}{3}, \hat{z}, \hat{id})$ and $w(\frac{2}{3}, \hat{z}, \hat{id})$. Circles identify the predictions of the points $w(\frac{1}{3}, \hat{z}, \hat{id})$ and $w(\frac{2}{3}, \hat{z}, \hat{id})$. The big cross on both worms identifies the head point $w(1, \hat{z}, \hat{id})$. (b) Interactions during occlusion. (c) Re-identification after occlusion. (Color figure online)

5 Evaluation

5.1 Normalized Shape

The normalized shape representation has been evaluated in [13] on a dataset containing 100 images selected from 7 different video sequences. This section gives an overview of the results. For the evaluation, the original binary worm region is compared with its normalized shape, which is projected into the image space to match the original worm region. Errors occur when pixels are missed from the original shape. Non-original pixels are not added when projecting the normalized shape, since the circles are always inside the original shape. Comparing the projected normalized shape to the original worm region, the minimum error is 1.02%, the maximum error is 8.89% and the mean-error for all images is 3.2%.

5.2 Tracking

To evaluate the performance of the proposed approach a MATLAB prototype has been realized. The prototype has been applied on a dataset of worm videos from the Max F. Perutz Laboratories in Vienna. All videos in the dataset are in gray scale with a size of 1280×960 pixels and a variable frame rate between 30

Table 1. Evaluation results on whole dataset.

Re-identifications		Head-tail decisions	
# of re-identifications	1.308	# of head-tail decisions	2.962
# of false re-identifications	3	# of false head-tail decisions	42
False re-identifications in %	0.229 %	False head tail decisions in %	1.418 %

and 60 frames per second (43 frames per second in average). To speed up the elaboration time the videos have been re-sized to 640×480 pixels. The whole dataset consists of 25 videos that have a total duration of more than 2h and are composed of 320.630 frames (elaborated videos can be found on our web site¹).

The ground truth of the worm identities has been generated manually by selecting head and tail position and the worm identifier (male or female) in every frame. Table 1 shows the result of the evaluation of the tracking on the whole dataset. The number of re-identifications is equal to the number of occlusions in which the worms overlap themselves. The number of head-tail decision is bigger than the number of re-identifications, because it is two times the number of re-identifications plus the times a single worm occludes itself (e.g. forms a circle). As can be seen in Table 1, the number of false gender decisions is different in comparison to the number of false head-tail decisions, because these decisions are made independent of each other. Furthermore, it is important to note that false head-tail decisions do not have repercussions on the following head-tail decisions, but false gender decisions can influence the following decisions on gender.

Regarding the gender association, it can be observed that in twenty-two videos there are zero association errors and in three videos, starting from a certain frame to the end of the video, the worms' genders are confused.

Regarding the head and tail association, it can be observed that the majority of false decisions are made when the worms lose their vitality and start to stay quiet or do not move at all. In these situations, the trajectories of the worms are ambiguous and depend on the movement of the water. Instead, when the worms swim in a *natural way*, it is unlikely that false decisions are made.

To illustrate the diversity of the worm features, Table 2 shows the values before and after the occlusion shown in Fig. 4 occurs. This example was chosen because the male worm appears bigger than the female worm before the occlusion, which changes after the occlusion. It shows why it is necessary to consider all features and use a sliding window.

Table 2. Worm features before and after the occlusion shown in Fig. 4 occurs.

Feature	Male worm	Female worm
Area (pixels)	477/384	441/485
Mean gray scale value	50/49	71 / 66
Geodesic length (pixels)	68,63/66,77	71,04/66,61

¹ <http://www.prip.tuwien.ac.at/research/worms.php>.

6 Conclusion

In this paper, a novel approach to track *Platynereis dumerilii* is proposed. It is able to handle occlusions and maintain the identity of the tracked worms with the help of a novel feature, the *normalized shape*. The normalized shape allows the comparison of the shape of worms independent of their deformation. It is used in conjunction with other features to correctly re-identify the worms after occlusions. Experimental evaluations on more than two hours of video material showed that the proposed approach is able to reliably analyze the nuptial dance of the worms. In 99.8% of the cases the gender of the worms was correctly re-identified after an occlusion. The head and the tail were correctly labeled in 98.6% of the cases.

The proposed method for the re-identification and the trajectory analysis to assign head and tail are not limited to the presented application and can be applied to other tracking problems. Especially the normalized shape is a suitable representation for all kinds of non-rigid objects having a main axis.

Acknowledgements. The authors thank Stephanie Bannister from the Max F. Perutz Laboratories GmbH for valuable discussions and providing videos of the spawning worms.

References

1. Aigerman, N., Poranne, R., Lipman, Y.: Lifted bijections for low distortion surface mappings. *ACM Trans. Graph* **33**(4), 69:1–69:12 (2014)
2. Baek, J.H., Cosman, P., Feng, Z., Silver, J., Baek, J., Schafer, W.: Using machine vision to analyze and classify *Caenorhabditis elegans* behavioral phenotypes quantitatively. *J. Neurosci. Methods* **118**, 9–21 (2002)
3. Bentley, M.G., Olive, P.J.W., Last, K.: Sexual satellites, moonlight and the nuptial dances of worms: the influence of the moon on the reproduction of marine animals. *Earth, Moon, Planets* **85**, 67–84 (1999)
4. Chai, T., Draxler, R.R.: Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding rmse in the literature. *Geosci. Model Dev.* **7**(3), 1247–1250 (2014)
5. Coifman, B., Beymer, D., McLauchlan, P., Malik, J.: A real-time computer vision system for vehicle tracking and traffic surveillance. *Transp. Res. Part C: Emerg. Technol.* **6**(4), 271–288 (1998)
6. Hardaker, L.A., Singer, E., Kerr, R., Schafer, W.R.: Serotonin modulates locomotory behavior and coordinates egg-laying and movement in *C. elegans*. *J. Neurobiol.* **49**, 303–313 (2001)
7. Hoshi, K., Shingai, R.: Computer-driven automatic identification of locomotion states in *Caenorhabditis elegans*. *J. Neurosci. Methods* **157**(2), 355–363 (2006)
8. Huang, K.M.: *Tracking and Analysis of Caenorhabditis Elegans Behavior Using Machine Vision*. ProQuest, Ann Arbor (2008)
9. Husson, S.J., Costa, W.S., Schmitt, C., Gottschalk, A.: Keeping track of worm trackers. In: *WormBook* (2012)

10. Kovvali, N., Banavar, M.K., Spanias, A.: An Introduction to Kalman Filtering with MATLAB Examples. Synthesis Lectures on Signal Processing. Morgan and Claypool Publishers, San Rafael (2013)
11. Lascio, R.D., Foggia, P., Percannella, G., Saggese, A., Vento, M.: A real time algorithm for people tracking using contextual reasoning. *CVIU* **117**(8), 892–908 (2013)
12. Osborne, J.W.: Best Practices in Quantitative Methods. Sage Publications, Thousand Oaks (2008). Ed. by Osborne, J.W.: Includes bibliographical references and index
13. Pucher, D.: 2D tracking of platynereis dumerilii worms during spawning. Technical report PRIP-TR-135, TU Wien, Austria, April 2016
14. Ramot, D., Johnson, B.E., Berry, T.L., Carnell, L., Goodman, M.B.: The parallel worm tracker: a platform for measuring average speed and drug-induced paralysis in nematodes. *PLoS ONE* **3**(5), e2208+ (2008)
15. Shingai, R.: Durations and frequencies of free locomotion in wild type and gabaergic mutants of *Caenorhabditis elegans*. *Neurosci. Res.* **38**(1), 71–84 (2000)
16. Swierczek, N.A., Giles, A.C., Rankin, C.H., Kerr, R.A.: High-throughput behavioral analysis in *C. elegans*. *Nat. Methods* **8**(7), 592–598 (2011)
17. Wang, S.J., Wang, Z.-W.: Track a worm, an open-source system for quantitative assessment of *C. elegans* locomotory, bending behavior. *PLoS ONE* **8**(7), 1–10 (2013)
18. Yemini, E.: High-Throughput, Single-worm Tracking and Analysis in *Caenorhabditis Elegans*. University of Cambridge, Cambridge (2013)
19. Zantke, J., Bannister, S., Rajan, V.B.V., Raible, F., Tessmar-Raible, K.: Genetic, genomic tools for the marine annelid *platynereis dumerilii*. *Genetics* **197**, 19–31 (2014). World Polychaeta Database (WPolyDb)