

Infrared Depth Camera System for Real-time European Lobster Behavior Analysis

Sheng Yan and Jo Arve Alfredsen

Department of Engineering Cybernetics, Norwegian University of Science and Technology, Norway

Keywords: Infrared Depth Camera, Behavior Analysis, European Lobster, Animal Tracking.

Abstract: European lobster is a highly treasured seafood, but aquaculture production based on traditional communal rearing practices has proved challenging for this species due to its inherent agonistic behavior. This paper presents a novel computer vision system that is designed for analysis of lobster behavior and can serve as a tool to assist selection of breeding stock in a prospective selective breeding program for European lobster. The automated tracking system provides large quantities of behavioral data for boldness and aggressiveness analysis, and the infrared light source causes less disturbance to the nocturnal animal under observance. In addition, because the object is recognized based on depth information instead of color or grayscale pattern recognition, there are no restrictions on the selection of color or material for the substrate in the experimental setup. This paper also contributes towards diminishing tracking error caused by water surface reflection and robust body orientation estimation in case of inaccurate body segmentation. We tested ten European lobsters sized between 25-30 cm to demonstrate the performance and effectiveness of our proposed algorithm.

1 INTRODUCTION

The European lobster was once an important economic special for the fishery but the stock collapsed between 1960 and 1980, to less than 10% of its pre-1960 level (Agnalt et al., 2007). Nowadays, the situation has not improved much and there are still many attempts to recover the natural stock. In order to meet the huge market demand for the European lobster, a long-lasting interest for controlled intensive production of lobster has raised, where the aquaculture industry mainly focuses on two different strategies; in-house production from egg to plate-sized lobster sold directly to the consumer market, or sea ranching where lobster juveniles are released into the sea in selected areas for growth and recapture after some years when they reach the minimum landing size. However, both approaches require that the lobster is kept in a relatively high density environment, and inherent behavioral traits such as cannibalism and aggressiveness make it expensive and complicated to run high-volume production systems.

One apparent solution to this problem could be to identify potential non-aggressive individuals through behavior analysis, which subsequently could be used as breeding material for a cultivation-friendly docile strain of lobster where the lobsters can be kept in sim-

ple communal cultures similar to those used successfully for fish.

Accordingly, the motivation of this paper is to create a real-time automatic tracking system which can robustly gather large quantities of high resolution behavioral data during carefully designed challenge tests and hence give the possibility for more accurate boldness and aggressiveness analysis of European lobster. In addition, the infrared light source of our system allows gathering of data in complete darkness giving less disturbance to the nocturnal animals under observance. Because the European lobster is recognized in our system based on depth information instead pattern recognition based on color or grayscale images, there are no restrictions on the selection of substrate color or material in the experimental setup.

In previous studies of lobster behavior, it has been common to manually determine the type of the behavior such as fighting and approaching, and only record the time span of various behaviors (Gherardi et al., 2010; Aspaas S, 2016). The statistical analysis is based merely on the measurement of time span, while other useful data, such as the speed of body movement and body orientation along the trajectory are omitted.

Even though there are tremendous amounts of publications on automatic human gesture and behavior

analysis, only a few focus on specific animals, especially for lobster behavior analysis. (Kato et al., 2004) developed a computer image processing system for quantifying zebrafish behavior based on two color cameras. Later (Qian and Chen, 2017) extend the system for tracking of multiple fishes from multi-view images. (Straw et al., 2010) used a multi-camera system for tracking of flying animals. The flying animals were modeled as small blobs and their positions were calculated by triangulation with known camera positions. (Yan and Alfredsen, 2017b) tried to extract the gesture from a single lobster in view based on a skeleton and distance transform. However the algorithm requires that the background is restricted to have a color that is very different from the lobster and its performance relies heavily on color based segmentation. Furthermore, attempts of aggressive behavior analysis for stage IV European Lobster juveniles was presented in (Yan and Alfredsen, 2017a).

Almost all previous studies on animal behavior are using RGB cameras and the objects are extracted from the background based on noticeable differences in color or grayscale image pattern between objects and the background. However, in real applications, shadows and noise in images are inevitable and the object extraction algorithm normally puts strong restrictions on the backgrounds that can be used, which subsequently complicates the setup procedure of the animal behavior experiments.

The main contribution of our algorithm is that it is the first attempt using an infrared depth camera for lobster behavior research which introducing minimum disturbance to the nocturnal animal under observation. The paper also addresses the water reflection problem when tracking and determining the orientation of the animal under water using infrared depth camera.

The structure of this paper is as follows: we describe our proposed algorithm in detail in Section 2. This section contains three subsections describing each module contained in the automated real-time tracking system. The experiments where our algorithm is tested on ten wild European lobsters sized between 25-30 cm are described in Section 3. Discussion and further work are given in Section 4.

2 ALGORITHM DESCRIPTION

The system configuration is shown in Figure 1(a). The infrared camera is mounted in front of a water filled arena holding an European lobster with the optical axis being approximately perpendicular to the bottom surface of the arena. A typical depth map is shown in

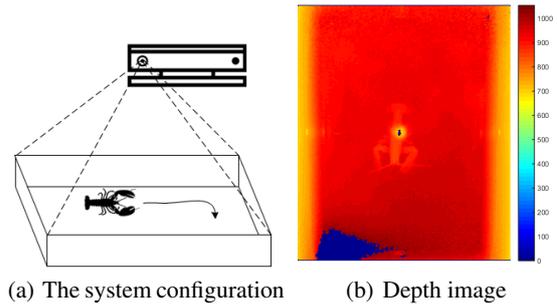


Figure 1: Infrared Depth Camera System for European Lobster Tracking.

Figure 1(b). The value on the color map bar is the distance of the point to the camera plane with unit of millimeter. It is noticeable that there are erroneous depth areas in the center and the left bottom of the image due to the water surface reflection and deflection. Moreover, because the absorption coefficient of water for the infrared light source (825 - 850 nm) used in the depth camera is much higher than that of air (Pegau et al., 1997), and the infrared signal being attenuated exponentially with respect to the distance travelled in water, it will appear more noisy at the bottom of the most distant parts of the arena. In the following subsections, we will deal with difficulties caused by these problems in order to track the lobster and obtain the orientation of lobster as accurate as possible.

2.1 Lobster Segmentation

Because of the relative position between the camera and the background is fixed in experiment's setup, it is effective to use the background subtraction method.

$$B_t(x,y) = \begin{cases} 1 & d_a(x,y) - d_t(x,y) > T_d \\ 0 & otherwise \end{cases} \quad (1)$$

where B_t is the primitive segmented foreground containing the lobster at time t . $d_t(x,y)$ is the depth map value at pixel (x,y) and $d_a(x,y)$ represents the depth to the bottom of arena obtained by calculating the average of the first N depth maps prior to introducing the lobster into the arena. Because the lobster is always located above the bottom of the arena, we can segment out the lobster simply by calculating the depth difference.

However, this method is not able to segment out the lobster in the area where depth camera renders a wrong depth map due to the reflection or deflection of water. Normally, the areas with erroneous measurements are marked with distance zero and we have to interpolate depths in this region by using the value from the depth region that is measured correctly by the camera. Because the arena bottom is flat, we

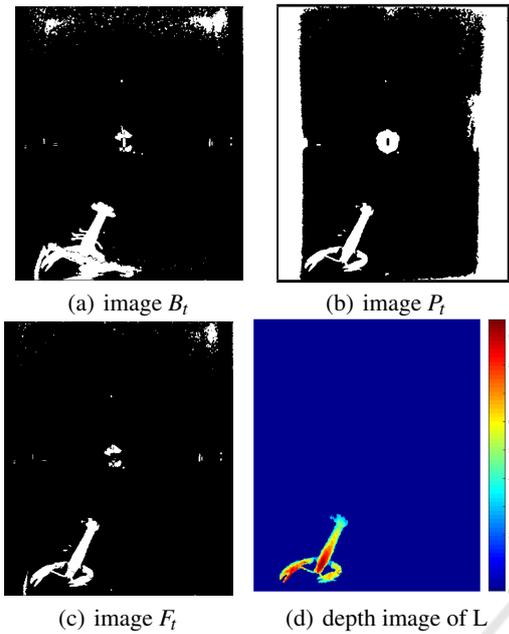


Figure 2: An example of the lobster segmentation procedure in which the lobster crosses the area where the depth map contains erroneous disparities. The example illustrates the robustness of the segmentation procedure.

run the MSAC (M-estimator SAmple and Consensus) (Torr and Zisserman, 2000) algorithm to obtain the plane model $a_px + b_py - z + d_p = 0$. We define the cost function that the MSAC tries to minimize as

$$C = \sum_{(x,y)} \rho\left(\frac{[a_px + b_py - d_a(x,y) + d_p]^2}{a^2 + b^2 + 1}\right) \quad (2)$$

Where C is a redescending M-estimator with threshold T that classifies the points as outliers when distances to the plane are larger than T .

$$\rho(e^2) = \begin{cases} e^2 & e^2 < T^2 \\ T^2 & e^2 \geq T^2 \end{cases} \quad (3)$$

With the estimated plane parameter $\theta = (a_p, b_p, d_p)$ we can obtain the flat bottom depth map denotes as $d_f(x, y) = a_px + b_py + d_p$ and obtain the second segmentation binary image based on the fact that the lobster should be above the estimated bottom plane but also below T_{max} that is determined by the height of the lobster's body.

$$P_t(x, y) = \begin{cases} 1 & T_{max} > d_f(x, y) - d_t(x, y) > T_{min} \\ 0 & otherwise \end{cases} \quad (4)$$

Combining the two segmented binary images by applying the logical AND operation, we obtain a more robust lobster body segmentation.

$$F_t(x, y) \leftarrow B_t(x, y) \wedge P_t(x, y) \quad (5)$$

Because all the previous procedures are done pixel-wise, there are still many unavoidable noise spots or small regions caused by the reflection of water surface. Therefore, we run the two-pass, connected components algorithm (Horn, 1986) to divide the binary image F_t into blobs and choose the largest blob as lobster and we denote the blob point set as $L = \{(x_i, y_i) | i = 1 \dots n\}$. The procedure is illustrated in Figure 2.

2.2 Body Orientation

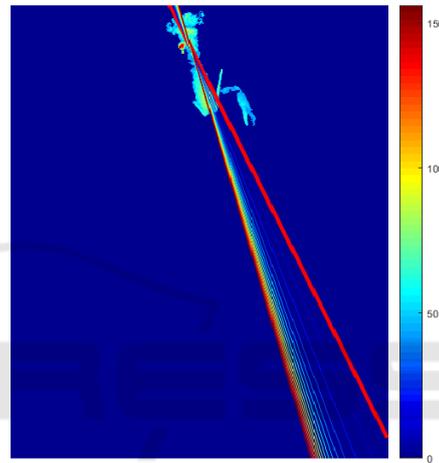


Figure 3: Example showing the performance of the body line fitting algorithm. In this example, one claw is not included in the connected component blob and the initial line of orientation is indicated with a thick red line. The asymmetry of the lobster body causes the line fail to represent the true body orientation. The thin lines with colors from blue to red show the steps taken by the refinement procedure for finding the true main body orientation.

The body orientation is a very important quantity in behavior analysis. However, the lobster has two large flexible claws with a wide range of possible positions making the orientation somewhat difficult to calculate. In addition, because the connecting joints between the lobster's body and claws are sometimes positioned above the bottom of arena with small height, it is possible that one or both of the claws are missing in L due to blob disconnection. By observing that the main body stem is always highest, we could fit a line to L in terms of weighted perpendicular offsets. Suppose the line function is $y = ax + b$ and we can obtain the model by minimize the line cost function

$$R = \sum_{i=1}^n \frac{[y_i - (a + bx_i)]^2 w_i}{1 + b^2} \quad (6)$$

Where weight $w_i = \lambda(d_f(x_i, y_i) - d_t(x_i, y_i))$ is proportional to the height from lobster's back to the interpolated plane. The equation (6) can be shown to have a closed-form solution as follows:

$$B = \frac{1}{2} \frac{[\sum_{i=1}^n y_i^2 w_i - (\sum_{i=1}^n y_i w_i)^2 / \bar{n}]}{(\sum_{i=1}^n x_i w_i \sum_{i=1}^n y_i w_i) / \bar{n} - \sum_{i=1}^n x_i y_i w_i} - \frac{1}{2} \frac{[\sum_{i=1}^n x_i^2 w_i - (\sum_{i=1}^n x_i w_i)^2 / \bar{n}]}{(\sum_{i=1}^n x_i w_i \sum_{i=1}^n y_i w_i) / \bar{n} - \sum_{i=1}^n x_i y_i w_i} \quad (7)$$

$$b = -B \pm \sqrt{B^2 + 1} \quad (8)$$

$$a = \frac{\sum_{i=1}^n y_i w_i - b \sum_{i=1}^n x_i w_i}{\bar{n}} \quad (9)$$

Where $\bar{n} = \sum_{i=1}^n w_i$ and we choose the sign in equation (8) that minimize the line cost function.

After obtaining the initial estimate of the body line, a subsequent iterative refinement procedure is implemented to obtain a more accurate line fitting the main body stem based on the subset of points in L that are within some vertical distance to the line. This iterative procedure is run in steps until convergence or maximum number of iterations is reached. Figure 3 shows a challenging case with a large initial asymmetry, and illustrates how the body line refinement algorithm iterates through multiple steps to eventually find the true body line.

Algorithm 1: Refine the body line.

Input: Lobster blob L , weight w_i , distance threshold T_l

Output: refined body line $y = ax + b$
initialize (a, b) according to equation (7)-(9);

repeat

 obtain point set

$$L_s = \{(x, y) \in L \mid \frac{|y - ax - b|}{\sqrt{a^2 + b^2}} < T_l\};$$

$$(a, b) = \operatorname{argmin}_{a, b} \sum_{(x_i, y_i) \in L_s} \frac{[y_i - (ax_i + b)]^2 w_i}{1 + b^2};$$

until reach maximum iteration OR

 Convergence;

return $y = ax + b$ and body orientation

$$\alpha = \operatorname{actan}(a)$$

For fast convergence, we can also take the calculated line model at $t - 1$ as the initialization value for (a, b) at time t .

To resolve the ambiguity of the head direction, points in set L are projected onto the line $y = ax + b$ and we count the number of projected points in the two halves of the line segment whose end points are defined by the two projected points from L on $y = ax + b$ having the largest distance. The direction

is pointing from the half line segment with less points to the half with more points based on the fact that the carapace of a lobster is has larger height and covers a larger area than the tail. Because it is impossible for a lobster to make a turn between two adjacent video frames, it is reasonable to assume that the orientation α_t at frame t and the orientation α_{t-1} at frame $t - 1$ forms an acute angle. We can then keep the direction at $t - 1$ and utilize it to eliminate the direction ambiguity by adding $\pm\pi$ to α_t if α_{t-1} and α_t forms an obtuse angle. Thus, the head orientation detection does only have to be calculated once in the first frame.

2.3 Position and Tracking

One easy solution to obtain the animal position is to calculate the centroid point of the lobster blob. However, because the claws are moving or claws can be missing from the blob, the centroid point is not a stable representation of the position. To reduce the instability and at the same time alleviate the computational burden, we project the centroid point to the lobster's body line. We then make the assumption that the distribution of measurement error is Gaussian and the lobster's dynamics of motion is linear. A Kalman filter can then be applied to obtain a more stable tracking trajectory. We define the state vector $\mathbf{s} = (x, y, \dot{x}, \dot{y})$.

The measurement model is \mathbf{z}_t at time t is defined as

$$\mathbf{z}_t = I\mathbf{s}_{t-1} + \mathbf{v} \quad (10)$$

The process update model can be written as

$$\mathbf{s}_t = A\mathbf{s}_{t-1} + \mathbf{w} \quad (11)$$

where the random variable \mathbf{v} and \mathbf{w} are normal distribution with zero mean, representing measurement and process update noise respectively. I is 4×4 unit matrix. The state transition matrix A relates the state at the previous time step $t - 1$ to the state at current time step t , and is defined as

$$A = \begin{pmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (12)$$

Because the movement of the lobster in a real situation is quite sophisticated and the linear dynamical model represents a relatively coarse approximation, we account for the uncertainty by setting a larger variance in \mathbf{w} compared to the noise in the measurement model \mathbf{v} .

3 EXPERIMENT

The experiment was carried out in an indoor environment and we chose to use a Kinect v2 as the infrared depth camera. The depth image had a resolution of 512 by 424 pixels and was obtained at a frame rate of 30 frames per second. The European lobster under observance were put in an shallow arena of size 115 cm by 95 cm filled with water of height to a height around 19 cm. The Kinect v2 was fixed around 80 cm above the bottom of the arena.

We used ten European lobsters of wild origin caught off west coast of Norway. The profiles of the lobster are listed in the Table (1).

We tested our system in two different scenarios; one where the water arena had no substrate, and another with sand as substrate, giving 20 tests in total. When a test for one lobster was finished, we emptied the arena and cleaned it in order to reduce potential odors released from lobster, preventing possible behavioral influence between tests.

The implementation of our algorithm was done in MATLAB R2016 on a PC with an Intel i5-3470 CPU and the data acquisition module was implemented in C#. The first 9000 frames from each test was used for the analysis. We calculated the background d_a using the first 90 frames before the lobster was put into the arena and waited for another 36 seconds to allow for the transfer of the lobster into the arena as well as some time to let the animal settle in the tank. Then the algorithm ran for the remaining 7830 frames. The average processing speed for the system was about 30 frames per second.

The system successfully tracked the lobsters movement in real-time, and recorded the position and orientation of the lobster in each frame allowing for analyses at a resolution and detail that otherwise would virtually impossible be done manually. For illustrative purposes, we overlay the lobster trajectory on the average of the first and last IR images in the tracking process as shown in Figure 4. Even though the substrate is different, the specimen shows some similarity in the trajectory pattern. For instance, lobster No.2 is unwilling to explore the four walls of the arena in the two tests. Lobster No.4 tends to be more active in exploring its surroundings than the other two lobster in the two tests. The orientation versus time plot is even more interesting, as shown in Figure 5. The consistency of the different lobsters orientation pattern can easily be seen. For instance, lobster No.2 tends to change the direction of motion more frequently than the other two. Such patterns are not easily discovered if records of the span of time of certain behaviors were the only quantities measured

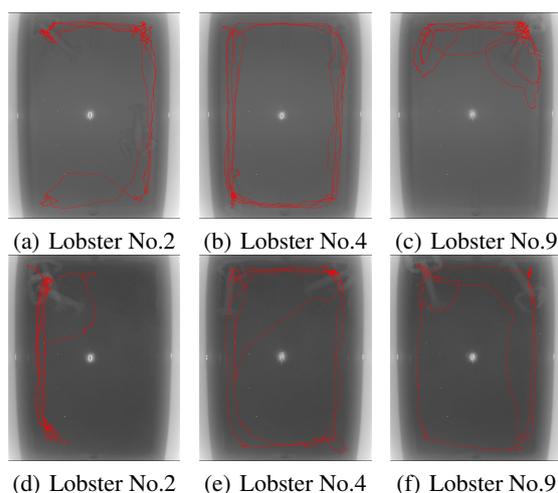


Figure 4: Selected results of European lobster tracking system. The three columns are for Lobster No.2, No.4 and No.9 respectively. First row is the trajectory of lobster with no substrate and the second row is with sand substrate.

during the behavior test.

We also calculated four additional metrics for the ten lobster in the two scenarios. The results are shown in table (2). The column 'normal' in the table represents the scenario with no substrate, while 'sand' represents the scenario with sand as substrate. Trajectory length gives the distance travelled by the lobster during a test and can be used as an indicator of its activity level when exploring a new environment. Explored area gives the percentage of the tank area the animal has explored during a test and can be used as a metric related to the lobsters willingness to explore new areas in an unknown environment. When calculating this metric, we represented the lobster by a circular disk covering an area with the same size as the lobster itself to compensate for the influence of variations in the segmentation results caused by the water reflection and noise. Mean angular speed could serve as a statistical indirect measurement with respect to the characteristics of individual to change its direction of exploration. From the mean speed measurement, we can also observe some tendencies that can not be seen directly from Figure 4, for this instance that all the lobster has higher mean speed in the normal scenario than in one with sand.

4 DISCUSSION AND FURTHER WORK

This is the first attempt to use an infrared depth camera for real-time European lobster tracking and behavior analysis. It represents an ideal solution for

Table 1: Profiles of the ten lobsters used in experiment.

Lobster No.	No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8	No.9	No.10
Body length (cm)	26.9	27.0	24.9	25.5	25.0	25.9	29.4	25.0	26.2	27.0
Gender	M	M	F	F	M	M	F	M	F	M

Table 2: Real-time lobster tracking metrics.

Lobster No.	trajectory length (pixel)		explored area		mean speed (pixel/s)		mean angular speed (radian/s)	
	normal	sand	normal	sand	normal	sand	normal	sand
No.1	6730	4922	63.9%	55.3%	25.8	18.9	0.204	0.147
No.2	5707	5072	46.0%	34.2%	21.9	19.4	0.133	0.139
No.3	7030	6343	63.9%	66.1%	26.9	24.3	0.168	0.158
No.4	7360	5907	46.0%	61.0%	28.2	22.6	0.144	0.126
No.5	5646	5023	72.4%	74.1%	21.6	19.2	0.115	0.097
No.6	7137	6849	72.9%	72.1%	27.3	26.2	0.186	0.157
No.7	6303	3803	68.9%	26.3%	24.1	14.6	0.147	0.085
No.8	7697	3985	79.1%	56.1%	29.5	15.3	0.165	0.096
No.9	5107	3982	34.8%	67.0%	19.6	15.3	0.149	0.091
No.10	8340	6225	61.8%	73.4%	32.0	23.8	0.179	0.143

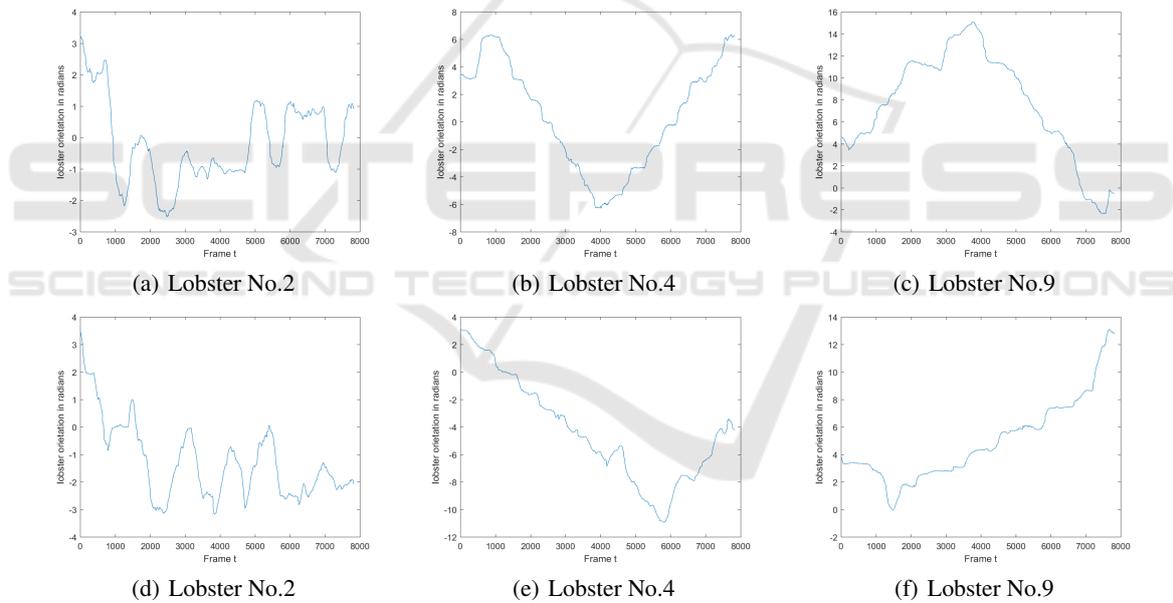


Figure 5: Lobster body orientation versus time plots. The three columns represent Lobster No.2, No.4 and No.9 respectively. First row is from the test scenario with no substrate and second row with sand substrate.

nocturnal animals such as lobster which require darkness to behave naturally. We also proved the applicability of the algorithm in terms of processing speed and the consistent results obtained from the observation of ten wild caught European lobster. The infrared depth camera together with the proposed algorithms enable quantitative high resolution analyses of behavioral traits like boldness and aggressiveness that were previously unattainable. The great amount of data that can be extracted through the system in

real-time allows for efficient scanning and behavioral characterization of populations of considerable size in searching for specimen suitable for breeding. More advanced machine learning algorithms can be applied to such data to reveal possible hidden patterns and hence to predict the lobster aggressiveness level from camera recordings.

We did not use the object association algorithm for tracking mainly because of the conditions of the application. The lobster is very aggressive animal and

we seldom put more than two lobsters together and let them fight. The lobsters can be seriously injured because of fight. However, with a simple extension that keeps two largest blob from F_t , this algorithm can be applied directly to the situation in which two lobsters are kept in the same arena with objects that prevent the two approaching or fighting with claws. In this way, the aggressiveness level between the two lobsters can be automatically inferred by the metrics from the measurements.

Further work may be done to improve the robustness of the lobster segmentation. Also, it is likely that a multiple camera system may obtain more robust and occlusion free observations.

national Conference of Pattern Recognition Systems (ICPRS 2017), pages 41 (6.)–41 (6.)(1).

Yan, S. and Alfredsen, J. A. (2017b). Real time lobster posture estimation for behavior research. *Proc. SPIE, Eighth International Conference on Graphic and Image Processing (ICGIP 2016)*, 10225:102250F–102250F–5.

REFERENCES

- Agnalt, A.-L., Kristiansen, T. S., and Jrstad, K. E. (2007). Growth, reproductive cycle, and movement of berried european lobsters (*homarus gammarus*) in a local stock off southwestern norway. *ICES Journal of Marine Science*, 64(2):288.
- Aspaas S, Grefsrud ES, F. A. J. K. T. H. A. A.-L. (2016). An enriched environment promotes shelter-seeking behaviour and survival of hatchery-produced juvenile european lobster (*homarus gammarus*). *PLoS ONE*, 11(8):e0159807.
- Gherardi, F., Cenni, F., Parisi, G., and Aquiloni, L. (2010). Visual recognition of conspecifics in the american lobster, *homarus americanus*. *Animal Behaviour*, 80(4):713 – 719.
- Horn, B. K. P. (1986). *Robot Vision*, pages 69–71. MIT Press.
- Kato, S., Nakagawa, T., Ohkawa, M., Muramoto, K., Oyama, O., Watanabe, A., Nakashima, H., Nemoto, T., and Sugitani, K. (2004). A computer image processing system for quantification of zebrafish behavior. *Journal of neuroscience methods*, 134(1):1–7.
- Pegau, W. S., Gray, D., and Zaneveld, J. R. V. (1997). Absorption and attenuation of visible and near-infrared light in water: dependence on temperature and salinity. *Appl. Opt.*, 36(24):6035–6046.
- Qian, Z.-M. and Chen, Y. Q. (2017). Feature point based 3d tracking of multiple fish from multi-view images. *PloS one*, 12(6):e0180254.
- Straw, A. D., Branson, K., Neumann, T. R., and Dickinson, M. H. (2010). Multi-camera real-time three-dimensional tracking of multiple flying animals. *Journal of The Royal Society Interface*.
- Torr, P. H. and Zisserman, A. (2000). Mlesac: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78(1):138–156.
- Yan, S. and Alfredsen, J. A. (2017a). Automatic video analysis of stage iv european lobster juveniles for their aggressive behavior assessment. *Proc. IET, 8th Inter-*