

Autonomous Tracking of Sea Turtles based on Multibeam Imaging Sonar: Toward Robotic Observation of Marine Life

Toshihiro Maki* Hiroumi Horimoto* Takashi Ishihara**
Kazuya Kofuji***

* *The University of Tokyo, Tokyo, Japan*
(*e-mail: maki@iis.u-tokyo.ac.jp*).

** *Suma Aqualife Park, Hyogo, Japan (e-mail: t-ishihara@sumasui.jp)*

*** *Ibaraki Prefectural Oarai Aquarium, Ibaraki, Japan*
(*e-mail: kofuji@aquaworld-oarai.com*)

Abstract: This paper proposes a method for autonomous underwater vehicles to track sea turtles without attaching any tag to them, toward efficient and long-term observation of marine life. The method uses a multibeam imaging sonar as the main sensor to detect sea turtles. The method utilizes convolutional neural network (CNN) for detecting a sea turtle in sonar imagery. Surge and yaw movements of the vehicle are controlled to maintain the relative distance and direction to the detected target. The proposed method was implemented in the AUV HATTORI. The AUV succeeded in tracking a sea turtle in natural condition for 270 seconds in shallow sea.

© 2019, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.

Keywords: Autonomous underwater vehicles, Field robotics, Imaging sonar, CNN, Sea Turtle

1. INTRODUCTION

Tracking and logging marine life are important for understanding their habitat, behavior, and ecology. Recently, the method called bio-logging is widely used to monitor wild life. The method employs miniature animal-attached electronic tags to collect scientific data on the movement, behavior and physiology of free-ranging animals. The method is especially important in underwater environment, where other sensing methods are difficult to be applied. There have been a number of new findings on marine life such as whales, dolphins, penguins, sea turtles, fishes, birds, and so on, brought by the method (Hays et al. (2016)). On the other hand, it is often difficult to attach tags to marine life, especially when the target is savage, rarely surfaces, or has slippery skins. Recovery of the tags is also a big challenge. There is also a concern that the tag itself might affect the target.

This paper proposes a method for autonomous underwater vehicles (AUVs) to track sea turtles without attaching any tag to them, toward efficient and long-term observation of marine life. The method uses a multibeam imaging sonar to detect sea turtles. There are some related works. For example, an AUV actively tracked acoustic tags attached to sharks (Skomal et al. (2015) Lin et al. (2017)) and sea turtles (Dodge et al. (2018)). Vision-based tracking methods have been also proposed for animals such as jellyfish (Rife and Rock (2003) Yann et al. (2005)) and fishes (Jun and Clark (2006)). However, optical methods have difficulties at poor lighting or highly turbid conditions.

The advantage of the proposed method is that there is no need to attach tags to the target. As the method relies only

on sensors mounted on the AUV, the method mitigates troubles relating to attaching/recovering tags. The method is also applicable to animals those are difficult to pre-capture for tagging such as deep sea sharks. Furthermore, the method has a large potential on measurable data type, quality, and duration, as any sensor can be used as long as the AUV can house it. This paper focuses on the tracking algorithm and results of sea experiments using our test-bed AUV HATTORI. See Horimoto et al. (2018) for details of the detection algorithm.

The structure of the paper is as follows. Section 2 introduces the overview of the proposed method. Section 3 explains about the convolutional neural network (CNN) based algorithm to detect sea turtles. Section 4 reports the results of sea experiments using the AUV HATTORI. In the experiments, the AUV succeeded in tracking a sea turtle in natural condition for 270 seconds.

2. TRACKING METHOD

There are two types of sensors, optics and acoustics, those can be used to detect sea turtles. In general, optical sensors such as camera can detect more detailed features than sonars, as light has much higher frequency than sound. However, the range is limited to typically several meters. It is often less than one meter in coastal water. On the other hand, sonars have a longer range at the cost of resolution. The proposed method employs sonar, so that the AUV can track a sea turtle at a distance of several meters in turbid waters.

Figure 1 shows the concept of the method. A forward looking multibeam imaging sonar is mounted at the nose of the AUV. This sonar can scan wide area at once by hundreds of fan-shaped acoustic beams. The AUV is

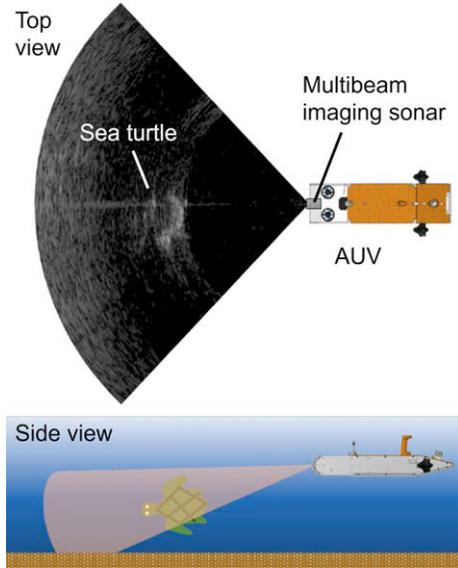


Fig. 1. Concept of the proposed method

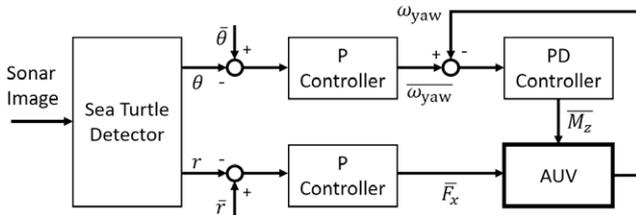


Fig. 2. Block diagram of the tracking method

assumed to be able to measure its depth and attitude (roll, pitch) and yaw angular velocity. The AUV is assumed to be able to control at least two degree of freedom, surge and yaw. It is also assumed that the AUV is at the same depth with the sea turtle.

The block diagram of the method is shown in Fig. 2. The sea turtle detector estimates the relative position of the target from the acoustic image by the CNN based algorithm, and then outputs the distance r and direction θ to the target. Definition of r and θ are shown in Fig. 3. The references of surge force and yaw moment, \bar{F}_x and \bar{M}_z , are generated by the independent PID controllers, so that the AUV maintains the distance and direction to the target at the reference value of \bar{r} and $\bar{\theta}$, respectively. Typically, $\bar{\theta} = 0$. In case a surge velocity sensor such as flow meter is available, velocity reference \bar{v}_x can be used. Then the reference force \bar{F}_x is generated by another PID controller with velocity feedback.

3. SEA TURTLE DETECTOR

Interpretation of acoustic images is challenging due to acoustic shadow, low S/N ratio and ambiguity with acoustic beam aperture. Changing shape and various texture of sea turtles also make the reflection complicated. Recently, CNN (Convolutional Neural Network) based image recognition and detection methods are successfully applied to optical images. These methods have been also applied to acoustic images to detect specific target, such as ROV (Kim and Yu (2016)) and debris (Valdenegro-Toro (2016)).

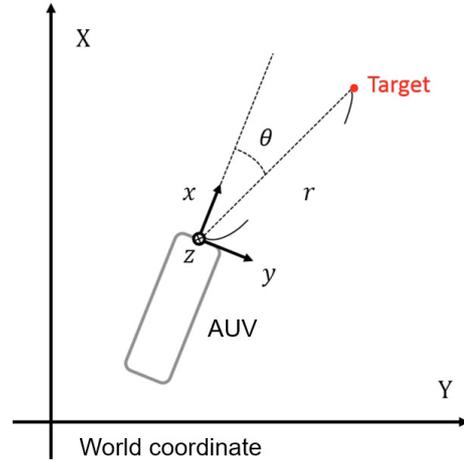


Fig. 3. Definition of the coordinates

The proposed method employs YOLOv2 (Redmon and Farhadi (2017)), convolutional neural network based object detection method, as the method has capability of real-time detection with high detection precision. The method is also robust against changes in scale and shape of the target. The method can handle multiple detection. In case more than one candidates were found in a single sonar image, the one with the highest probability is selected as the target.

Comparing to optical images, collecting acoustic images takes more time and effort. For training the model with fewer datasets, the fine-tuning method is applied, where the network was initially trained by other datasets, and retrained by the actual acoustic images. YOLOv2 was originally developed for optical images those have 3 color channels. However, sonar images have only 1 channel. The authors trained and tested the network, considering the sonar image as 3 channels image where all the channels have the same value.

After trained with the large dataset of optical images (Deng et al. (2009)), the network was retrained with a dataset of acoustic images. The dataset was collected by an ROV with a multibeam imaging sonar at a tank and shallow sea. Specifications of the sonar is shown in Table 1. The two species, loggerhead turtle (*Caretta caretta*) and green turtle (*Chelonia mydas*), were observed. The turtles were manually labelled on the acoustic images. Optical images were also collected at the tank as a reference. The dataset of 1190 still images were divided into training dataset (415 images obtained at the tank), validation dataset (126 images obtained at the tank), and the test dataset (649 images obtained at the shallow sea). The final score was 88 % precision and 64 % recall. See Horimoto et al. (2018) for details.

4. EXPERIMENTS

Sea experiments were conducted in November 2018, at the shallow and closed environment, the same place where the dataset used to test the trained network was collected. The dataset was used only to test, or for measuring the performance of the trained network. So, the network had not been trained using the information of this site.

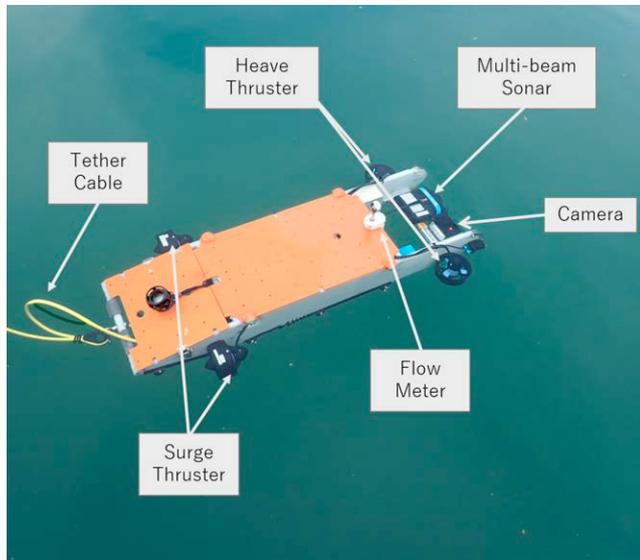


Fig. 4. AUV HATTORI

The proposed method was implemented in our test bed AUV HATTORI (Maki et al. (2018)). The outlook and specifications of the vehicle are shown in Fig. 4 and Table 2, respectively. The sonar is the same with the one used to collect the dataset. Although a tether cable was attached for easy recovery and communication, all the computations to detect and track the sea turtle were conducted in the vehicle in real-time. The frequency of sonar image acquisition and target detection was around 8 Hz. A forward-looking camera was installed for verification in post processing.

A loggerhead turtle with the straight carapace length of 0.85 m was brought several meters in front of the vehicle by a diver. The vehicle started the tracking mission as soon as the diver moved away. The reference distance and direction were set to be $\bar{r} = 4.0$ m, $\bar{\theta} = 0.0$ deg, respectively.

The vehicle succeeded in tracking the free-swimming sea turtle for around 270 sec. Time series of sonar and camera images obtained by the vehicle, and images taken by an outside camera are shown in Fig. 5. Detected sea turtle is indicated by red rectangles in the sonar images. The target is dimly seen in the camera images at 0 and 140 sec. The

Table 1. Specifications of multibeam imaging sonar

Model	Blueprint Oculus M750d
Frequency	750 kHz
Sampling rate	10 Hz
Range	5 - 10 m
Horizontal aperture	130 deg
Vertical aperture	20 deg
Angular resolution	1 deg
Mounting angle	10 deg down

Table 2. Specifications of AUV HATTORI

Weight (in air)	20 kg
Size (L, W, H)	1.0, 0.5, 0.3 m
Thrusters	Blue Robotics T200 × 4
Computer	Jetson TX2
IMU	MPU9250
Camera	GoPro HERO4 (stand alone)

target is clearly seen and detected in the sonar image at 270 sec, although it is hardly seen in the camera image taken at the same time. Relative distance and direction to the target, estimated by the vehicle in real-time, are shown in Figs. 6 and 7, respectively. The tracking error in distance and direction were within 1 m and 10 deg, respectively. Figure 8 indicates the number of sea turtle detections per second. There were constantly 8 detections throughout the mission. Although the target was lost for two times at around 60 and 110 sec, it was soon recovered within several seconds.

Trajectory of the sea turtle estimated from the information collected by the AUV is shown in Fig. 9. The trajectory is based on the AUV's trajectory, which was estimated by the control reference to the thrusters and IMU measurements. Origin of the coordinates is the initial position of the AUV. Segments with the estimated velocity of over 5 m/s were filtered out to eliminate noises. Moving average filter with the window size of 5 sec was also applied.

The external camera images indicates that the sea turtle submerged at around 270 sec. As the target disappeared from the sonar image simultaneously, the sea turtle was considered to have gone out of the field of view by descending. Extension to 3D environment is necessary to increase robustness against vertical movement of the target. As the vertical field of view of an multibeam imaging sonar is limited, stochastic search algorithm will be necessary.

5. CONCLUSION

This paper proposed a method for autonomous underwater vehicles to track sea turtles without attaching any tag to them, toward efficient and long-term observation of marine life. The method uses a multibeam imaging sonar to detect sea turtles. The method utilizes YOLOv2, a CNN based detector, to detect sea turtles in sonar imagery. Surge and yaw movements of the vehicle are controlled to maintain the relative distance and direction to the detected target. The proposed method was implemented in the AUV HATTORI. In the experiment held at shallow sea, the AUV succeeded in tracking a free-swimming sea turtle in natural condition for 270 sec.

ACKNOWLEDGEMENTS

The authors would like to thank to the staff of Suma Aqualife Park and Ibaraki Prefectural Oarai Aquarium for their kind support during data collection and experiments. The authors are also grateful to Prof. K. Sato, The University of Tokyo, for helpful suggestions that made this work happen.

REFERENCES

- Deng, J., Dong, W., Socher, R., Li, L., Kai, L., and Li, F.F. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255. doi: 10.1109/CVPR.2009.5206848.
- Dodge, K.L., Kukulya, A.L., Burke, E., and Baumgartner, M.F. (2018). Turtlecam: A smart autonomous underwater vehicle for investigating behaviors and habitats

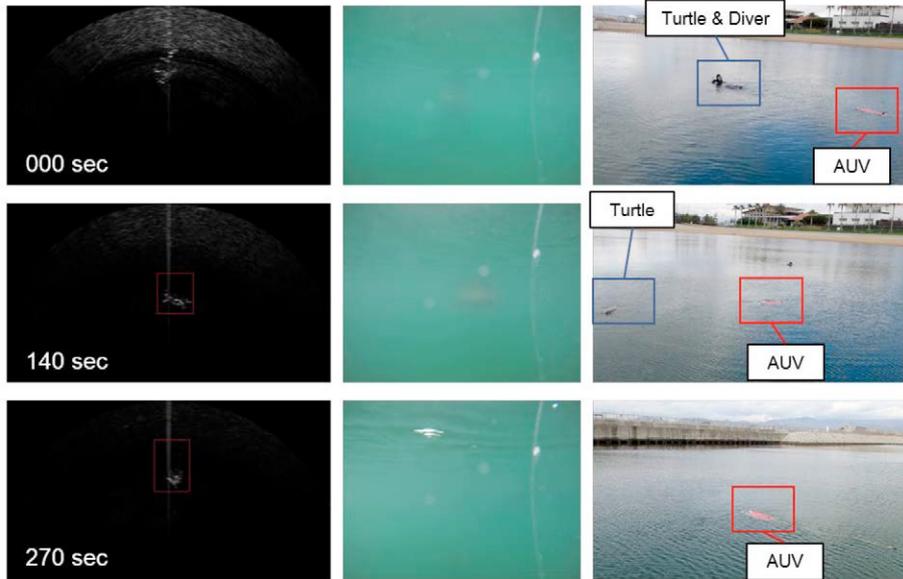


Fig. 5. Images of the sonar (left), the forward-looking camera (center), and the external camera (right) during the tracking mission. Elapsed time from starting the mission is indicated on the left.

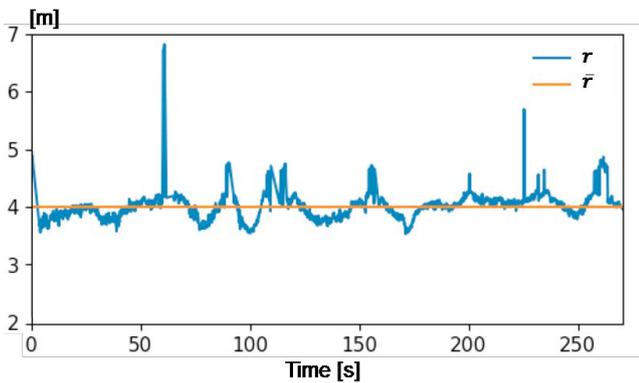


Fig. 6. Distance to the sea turtle from the AUV

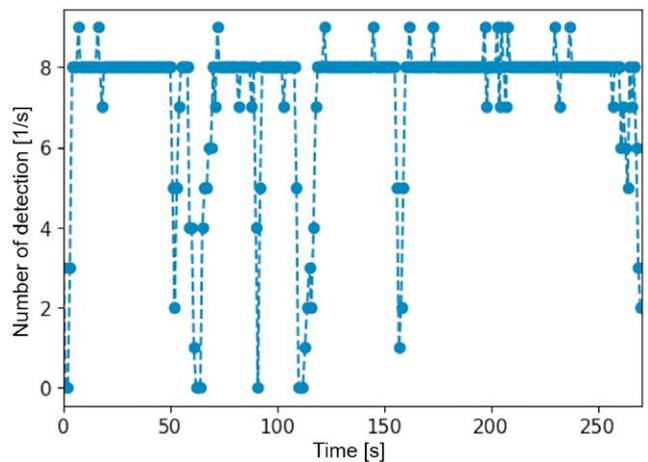


Fig. 8. Number of sea turtle detection per second

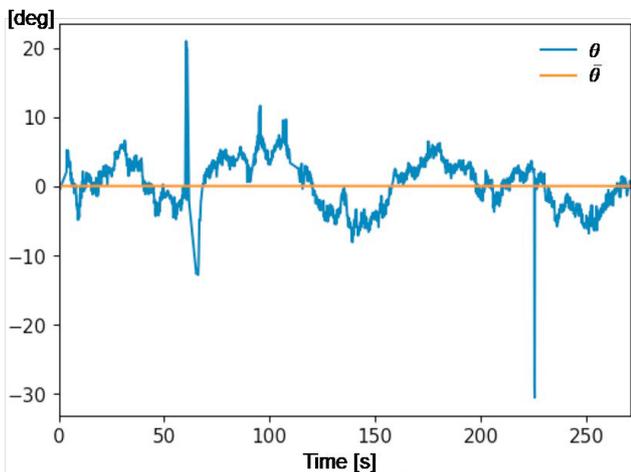


Fig. 7. Direction to the sea turtle from the AUV

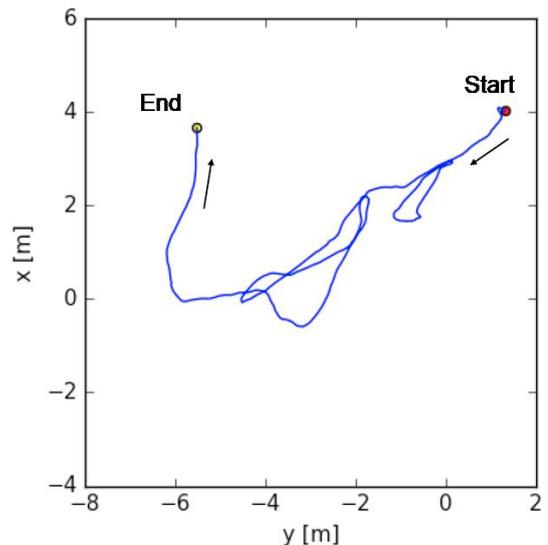


Fig. 9. Estimated trajectory of the sea turtle

of sea turtles. *Frontiers in Marine Science*, 5(90). doi: 10.3389/fmars.2018.00090.
 Hays, G.C., Ferreira, L.C., Sequeira, A.M.M., Meekan, M.G., Duarte, C.M., Bailey, H., Bailleul, F., Bowen,

- W.D., Caley, M.J., and Costa, Daniel P., e.a. (2016). Key questions in marine megafauna movement ecology. *Trends in Ecology & Evolution*, 31(6), 463–475. doi: <https://doi.org/10.1016/j.tree.2016.02.015>.
- Horimoto, H., Maki, T., Kofuji, K., and Ishihara, T. (2018). Autonomous sea turtle detection using multi-beam imaging sonar: Toward autonomous tracking. In *Autonomous Underwater Vehicles (AUV), 2018 IEEE/OES*.
- Jun, Z. and Clark, C.M. (2006). Autonomous fish tracking by roV using monocular camera. In *The 3rd Canadian Conference on Computer and Robot Vision (CRV'06)*, 68–68. doi:10.1109/CRV.2006.16.
- Kim, J. and Yu, S. (2016). Convolutional neural network-based real-time roV detection using forward-looking sonar image. In *2016 IEEE/OES Autonomous Underwater Vehicles (AUV)*, 396–400. doi: 10.1109/AUV.2016.7778702.
- Lin, Y., Hsiung, J., Piersall, R., White, C., Lowe, C.G., and Clark, C.M. (2017). A multi-autonomous underwater vehicle system for autonomous tracking of marine life. *Journal of Field Robotics*, 34(4), 757–774. doi: 10.1002/rob.21668.
- Maki, T., Noguchi, Y., Kuranaga, Y., Masuda, K., Sakamaki, T., Humblet, M., and Furushima, Y. (2018). Low-altitude and high-speed terrain tracking method for lightweight auvs. *Journal of Robotics and Mechatronics*, 30(6), 971–979. doi:DOI 10.20965/jrm.2018.p0971.
- Redmon, J. and Farhadi, A. (2017). Yolo9000: Better, faster, stronger. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6517–6525. doi:10.1109/CVPR.2017.690.
- Rife, J. and Rock, S.M. (2003). Segmentation methods for visual tracking of deep-ocean jellyfish using a conventional camera. *IEEE Journal of Oceanic Engineering*, 28(4), 595–608. doi:10.1109/JOE.2003.819315.
- Skomal, G.B., Hoyos-Padilla, E.M., Kukulya, A., and Stokey, R. (2015). Subsurface observations of white shark carcharodon carcharias predatory behaviour using an autonomous underwater vehicle. *Journal of Fish Biology*, 87(6), 1293–1312. doi:10.1111/jfb.12828.
- Valdenegro-Toro, M. (2016). End-to-end object detection and recognition in forward-looking sonar images with convolutional neural networks. In *2016 IEEE/OES Autonomous Underwater Vehicles (AUV)*, 144–150. doi: 10.1109/AUV.2016.7778662.
- Yann, E., Nose, Y., and Ura, T. (2005). Autonomous underwater sampling using a manipulator and stereovisual servoing. In *Europe Oceans 2005*, volume 2, 731–736 Vol. 2. doi:10.1109/OCEANSE.2005.1513146.