



Smart vector-inspired optical vision guiding method for autonomous underwater vehicle docking and formation

YUE ZHANG,^{1,2} XINWEI WANG,^{1,2,3,*} PINGSHUN LEI,¹ SHAO MENG WANG,^{1,2} YUQING YANG,^{1,2} LIANG SUN,^{1,2} AND YAN ZHOU^{1,2,3}

¹ Optoelectronic System Laboratory, Institute of Semiconductors, CAS, Beijing 100083, China

² College of Materials Science and Opto-Electronics Technology, University of Chinese Academy of Sciences, Beijing 100049, China

³ School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

*Corresponding author: wangxinwei@semi.ac.cn

Received 21 February 2022; revised 5 May 2022; accepted 13 May 2022; posted 16 May 2022; published 31 May 2022

A smart vector-inspired optical vision guiding (VIOVG) method for autonomous underwater vehicle (AUV) docking and formation is proposed. Unlike traditional optical guiding methods based on LED arrays, the method is inspired by a vector, and uses four laser diodes to form a wing-light pattern to realize smart optical guiding for AUVs. Due to the light scattering effect from water, the four laser diodes can emit four bright and slim laser beams in water, which can be captured by underwater cameras. The intersections of the laser beams form wing-lights as markers for AUV distance and pose estimation. The wing-light pattern is easily tuned and extended by changing the propagation direction and the power of the laser beams beyond the limitation of the physical size of the AUV or docking station. The simulations and experiments show that the proposed method can achieve high precision positioning. A relative distance error of 3.35% is achieved in a positioning experiment at ~10 m. It has great potential for AUV docking and formation, especially for small AUVs. © 2022 Optica Publishing Group

<https://doi.org/10.1364/OL.456544>

In the past decades, autonomous underwater vehicles (AUVs) have drawn significant interest in oceanographic surveys, pipeline inspection, natural resources exploration, and environment monitoring [1,2]. The limitations of battery volume and detection range make it difficult for a single AUV to work in a large venue. Instead of performing tasks with one AUV, multiple AUVs collaborating in a specific formation provide great improvement in working efficiency, and can easily perform more complicated tasks [3–6]. However, due to the difficulty of underwater positioning and communication, coordinating a group of AUVs has been a big challenge. Since electromagnetic waves are strongly attenuated by water, global positioning system signals cannot be received by AUVs for underwater navigation. Therefore, various guiding methods have been proposed to solve the problem of underwater cooperative positioning for AUV docking and formation [7–9]. Among these methods, acoustic and optical ways account for the vast majority. For acoustic

positioning methods, the accuracy is limited by the long wavelength of sound waves. Compared with acoustic methods, optical vision guiding (OVG) methods have higher accuracy, and can provide a six degrees of freedom (6-DoF) pose of AUVs, which is crucial for AUV docking and formation. In OVG methods, AUVs or docking stations are usually equipped with cameras and multiple lamps. Typically, more than four LEDs are fixed on the AUV housing to form a specific pattern as markers [10–13]. When one AUV and its neighbor AUVs or docking station are in different relative positions, the LED array shows different patterns captured by AUV cameras. Then the relative position of the AUVs can be calculated by pose estimation algorithms [14–16]. However, traditional optical methods can only work at a short distance [17]. Especially in turbid water, light spots of the LED array are severely diffused by a strong scattering effect, and thus the lights cannot be segmented and distinguished from each other, which leads to the failure of pose estimation [10,11]. What is worse, for a small docking station or small-AUV formation, under the limitation of their mechanical dimensions, LEDs have to be mounted compactly, as illustrated in Fig. 1(a), which makes it more difficult to distinguish and recognize each light spot in the images.

In this paper, a smart vector-inspired optical vision guiding (VIOVG) method for AUV docking and formation is proposed. For traditional OVG methods, the marker lights can be regarded as scalars, since the light only has intensity but no or weak direction characteristics. However, in the VIOVG method, the marker lights can be regarded as vectors, since lasers with slim beams are used, which have intensity and strong direction characteristics. Therefore, one can use the vector lights to form different patterns. For example, in Figs. 1(b) and 1(c), four lasers are mounted along the circumference of a docking station or AUVs. Each laser emits a collimated beam and forms a slim and bright ray under the scattering effect of water. In Fig. 1(b), laser rays L1 and L2 intersecting each other form a cross-point P1, and P2 can also be formed by laser rays L3 and L4. The opposite extensions of L1 and L4 generate a cross-point P3, and similarly P4 can be generated by L2 and L3. The four points P1, P2, P3, and P4 form a wing-light pattern and serve as feature points for

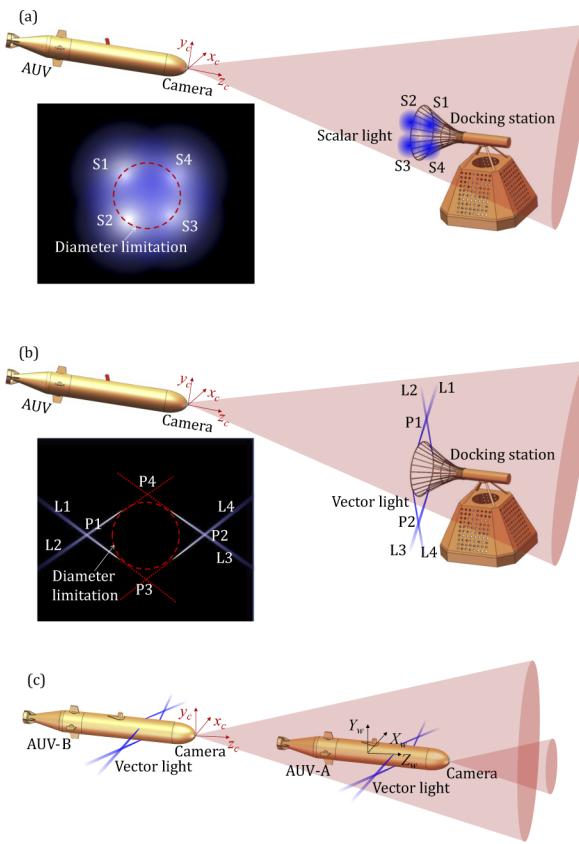


Fig. 1. (a) Scalar lights in AUV docking diagrammatic sketch of traditional OVG method. (b) Vector lights in AUV docking diagrammatic sketch of VIOVG method. (c) Multi-AUV formation diagrammatic sketch under the VIOVG method.

pose estimation. P1 and P2 are called as real wing-lights, and P3 and P4 are called as virtual wing-lights. Different from the traditional scalar marker lights in Fig. 1(a), the wing-lights can break through the limitation of the mechanical dimensions, as shown in Figs. 1(b) and 1(c), which contributes to the precision of pose estimation. For the traditional OVG method using LEDs, image processing is required to detect light spots as feature points. However, in the VIOVG method, there are no real light spots. Therefore, an algorithm for detecting laser beams instead of light spots needs to be developed. Since the laser beam is bright and slim, it can be precisely detected. Then, the sub-pixel coordinates of wing-light points in the VIOVG images are calculated by the intersections of laser rays. As illustrated in Fig. 1(c), when performing multi-AUV formation task by the VIOVG method, each AUV is equipped with vector lights, as well as optical cameras. The image of vector lights on AUV-A is captured by the camera of AUV-B. Similarly, for an AUV docking task under the VIOVG method, the vector lights are mounted on a docking station, as shown in Fig. 1(b), and their pattern can be captured by the camera of a pre-docking AUV. Let the world coordinate system locate on AUV-A or the docking station while the camera coordinate system locates on the camera of AUV-B or the pre-docking AUV. Since the size dimensions of wing-light pattern are known, the relative rotation matrix \mathbf{R} and translation vector \mathbf{t} from world to camera coordinate systems can be calculated. Thus, the relative 6-DoF pose of the AUV can be recovered. The pose estimation algorithm used in the VIOVG method is the

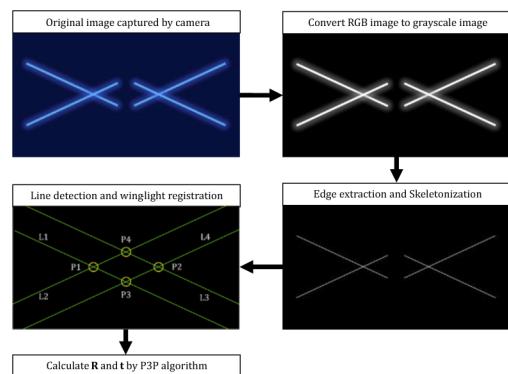


Fig. 2. Process of pose acquisition algorithm for the VIOVG method.

perspective-three-point (P3P) algorithm [18]. Shown in Eq. (1), point $\mathbf{X}_w \in \mathbb{R}^{3 \times 1}$ in world coordinate system is transformed to $\mathbf{X}_c \in \mathbb{R}^{3 \times 1}$ in camera coordinate system through 3D rotation and translation. The rotation can be described by rotation matrix $\mathbf{R} \in \mathbb{R}^{3 \times 3}$, while the translation can be noted as a vector $\mathbf{t} \in \mathbb{R}^{3 \times 1}$. Then, as shown in Eq. (2), using homogeneous coordinates, the point is projected to point $\mathbf{x} = [u, v, 1]^T$ on the camera image plane via a pinhole camera model, where (u, v) is the pixel position in the image. The transformation and projection relation can be expressed as

$$\mathbf{X}_c = \mathbf{R}\mathbf{X}_w + \mathbf{t}, \quad (1)$$

$$\mathbf{x} = \mathbf{K}\mathbf{X}_c, \quad (2)$$

where \mathbf{K} is the intrinsic camera matrix, expressed as

$$\mathbf{K} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}, \quad (3)$$

where f_x, f_y are the camera focal lengths along the x and y axes, c_x, c_y are offsets of camera principal point in pixel coordinates, and \mathbf{K} can be obtained from underwater camera calibration. Since the degree of freedom of \mathbf{R} and \mathbf{t} is 6, three $\mathbf{x}-\mathbf{X}_w$ point pairs are needed for the solution. Another point pair is used for the selection of the most reasonable solution from four candidate solutions, which has the smallest reprojection error. There are four wing-light points in the VIOVG method, which is enough for 6-DoF pose calculation. The process of pose acquisition algorithm is shown in Fig. 2. The camera mounted on the AUV keeps capturing images continuously during the process of formation or docking. Once the image is acquired, the RGB image is converted to gray scale image according to

$$I_{gray} = \alpha I_r + \beta I_g + \gamma I_b, \quad (4)$$

where I_{gray} is the converted gray scale image, I_r, I_g, I_b are the red, green, and blue channels of an RGB image, and α, β, γ are the weights of the different channels. The weights are determined by the wavelength of the laser and the water quality. For multi-AUV formation, each AUV can be equipped with vector lights of different wavelengths, which can be distinguished by adjusting the weight. In environments with poor water quality, weight adjustment can be used to filter scattered light and make the laser lines clear.

Edge extraction is performed on the gray scale image. The gray scale image is first processed by Sobel filtering as

$$D = |\nabla I_{gray}| = |S_x * I_{gray} + S_y * I_{gray}| \quad (5)$$

where D is the filtered image, ∇ is the gradient operator, and S_x and S_y are the Sobel filters along the x and y axes. The filtered image is then binarized by Otsu's method. After that, the binary image is skeletonized by morphological close operation to refine the laser rays and remove image noise. Then, the progressive probabilistic Hough transform (PPHT) [19] is used to detect lines from the skeletonized image. Five parameters are required to perform PPHT line detection, including distance resolution, angle resolution, threshold, minimum line length, and maximum allowed gap between points on the same line to link them. The detected lines are then clustered so that very close lines are considered to be the average of lines in the same cluster. Then four lines representing the laser rays are extracted. Thus, the four wing-light feature points in the image can be obtained as the intersections of the four lines. However, since the four lines intersect each other at up to six points, the fake point omitting algorithm must be employed to select the four wing-light points. Then, wing-light registration i.e., the one-to-one correspondence between the four wing-light points detected in the image and the four wing-light points in the world coordinate system is made. Finally, the relative 6-DoF pose of AUV is solved by the P3P pose estimation algorithm.

The most critical step in the pose estimation algorithm is to detect wing-lights in VIOVG images. Unlike traditional OVG methods, line detection replaces point detection. In our algorithm, four laser rays are detected first and then the wing-light points are calculated as the cross-points of four lines. Owing to the scattering effect of water, the laser ray lines are usually bright and slim, which makes it easy for line detection.

To prove the feasibility of the proposed VIOVG method, numerical simulations on the precision of pose estimation using the VIOVG method are performed. Compared with AUV docking, multi-AUV formation is challenging, especially for small AUVs, since their diameters are too small for traditional scalar LED lights. Therefore, the simulation for AUV formation is chosen, and Fig. 3 simulates the translation and rotation errors under (1) different angle θ between intersected laser beams, (2) different distance R between the vector light and the camera, (3) different yaw angle β of the vector light, separately. In the simulation, the camera has an image resolution of 1920×1080 pixels and a focal length of 5600 pixels. The principal point lies in the image center. As shown in Fig. 3(a), four laser outlets are mounted along the AUV housing, forming a square with a side length of 110 mm. For each simulation, the projected image of vector-light beams is generated using anti-aliased-line algorithm with a linewidth of 3 pixels. For every generated image, 64 PPHT line detections are performed, with the distance resolution of 1 pixel, angle resolution of 0.5°, detection threshold of 50, 83, 117, 150 pixels, minimum line length of 100, 133, 167, 200 pixels, and maximum allowed gap of 10, 40, 70, 100 pixels. When simulating the translation and rotation errors under different θ , R is set to be 10,000 mm and $\beta = 0$. As shown in Figs. 3(c) and 3(d), the translation and rotation errors decrease as θ grows. When θ is larger than 20°, the translation error remains smaller than 2.5% and the rotation error remains smaller than 10°. When simulating the translation and rotation errors under different R , θ is set to be 30° and $\beta = 0$. In Figs. 3(e) and 3(f), the translation and rotation error grows as R increases. The translation error is

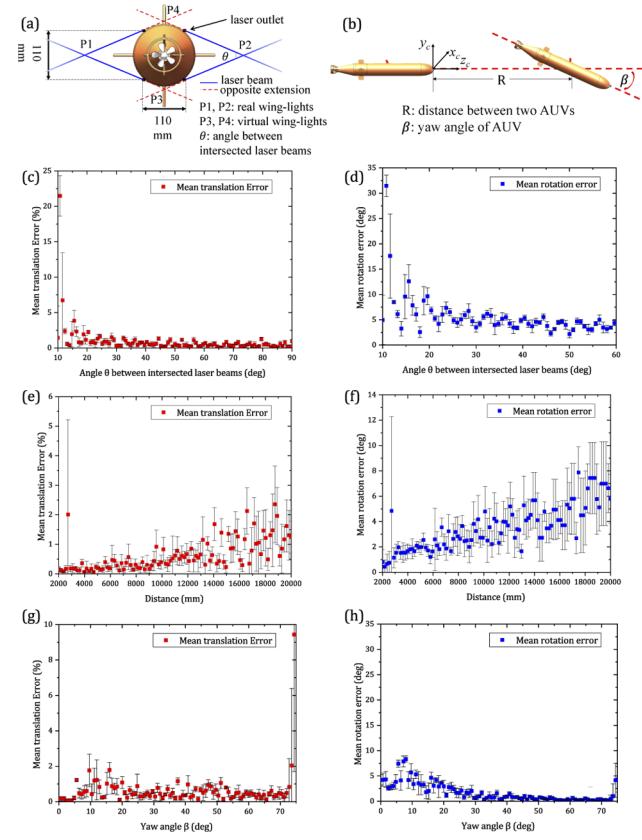


Fig. 3. Simulation configuration and results on the precision of pose estimation of AUV equipped with the VIOVG system. Translation and rotation errors are simulated under different θ , R , and β . (a) Back view of vector lights on the AUV. (b) Relative pose of two AUVs. (c), (d) Translation error and rotation error under different angle θ between intersected laser beams. (e), (f) Translation error and rotation error under different distance R . (g), (h) Translation error and rotation error under different yaw angle β .

always smaller than 4% and the rotation error is always smaller than 11°, with R varying from 2,000 mm to 20,000 mm. When simulating the translation and rotation errors under different β , R is set to be 10,000 mm and $\theta = 30^\circ$. As shown in Figs. 3(g) and 3(h), the translation error is under 3% and the rotation error is under 10°, with β varying from 0 to 70°. The simulation results show that the proposed method can provide sufficient accuracy for pose estimation in AUV formation. To test the VIOVG method in experiment, the experimental prototype system is established with a laser wavelength of 450 nm and angle between laser beams $\theta = 30^\circ$. In practical applications, when the water quality is poor, red lights are recommended since they are less affected by water scattering and the laser beam will be sharper in the images. In an environment with good water quality, blue-green light is recommended since it is less attenuated, and can work over long distances. As shown in Fig. 4, the underwater camera and vector light are placed on two tripods in a water pool to adjust their pose. Distance calibration wires are used to determine the distance between camera and vector light, providing the ground truth distance value for the positioning experiment. We performed position estimation at the distance of 9.804 m. The calculated distance $R = 9.476$ m, with a relative distance error of 3.35%. The experiment results show that the

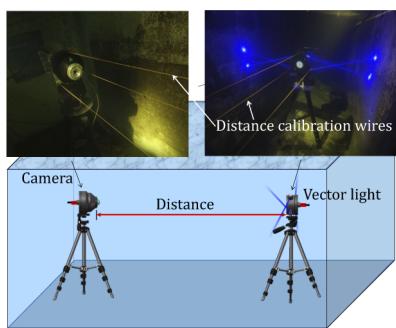


Fig. 4. Experiment setup of the VIOVG system. The underwater camera and vector light are placed on tripods in a water pool. Distance calibration wires are used to determine the distance between the camera and the vector light.

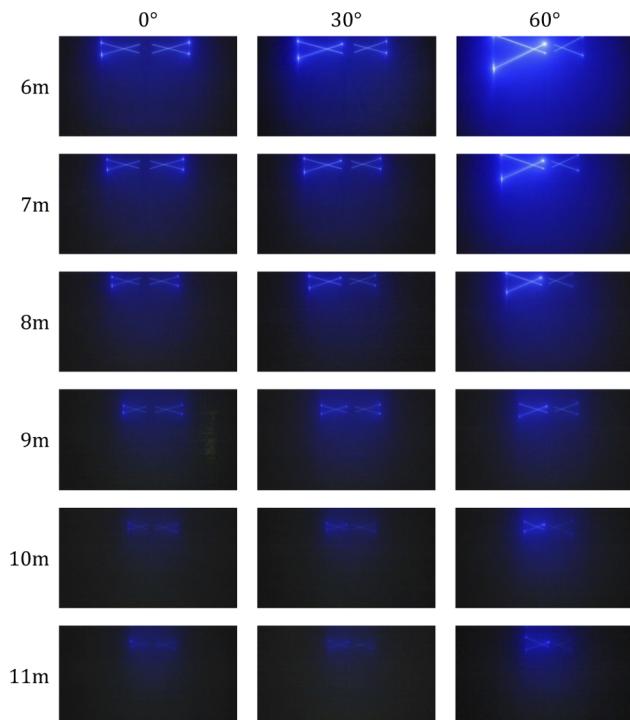


Fig. 5. Wing-light patterns at different distances and yaw angles for the smart VIOVG system.

VIOVG system has good accuracy in the positioning task. We performed experiments at different distances from 6 m to 11 m and yaw angles of 0°, 30°, and 60°. The images of vector light are shown in Fig. 5, from which one can see the vector light pattern is clear at a distance within 10 m and at different yaw angles. The attenuation coefficient of the experimental water is 0.15 m^{-1} measured by WetLabs AC-9 absorption and attenuation meter. Compared with the light spot detection used in traditional OVG methods, line detection can provide the sub-pixel location of wing-light points in VIOVG images, which can improve the precision of pose solution.

In a brief summary, the proposed VIOVG method provides a novel way of positioning for AUV formation and docking. The proposed guiding system uses lasers instead of LEDs as the markers for pose acquisition. Taking advantage of the scattering

effect of water, a laser can form a bright and slim line, which can be easily detected from the camera image. The cross-points of laser rays form wing-lights, and the sub-pixel coordinates can be obtained by solving the intersections of the laser lines. The wing-light breaks the limitation of mechanical dimensions, making the guiding system suitable to be mounted on small AUVs or docking stations. Simulation and experiment results show that the proposed method has good accuracy in pose estimation, which reveals that the VIOVG method has great potential for AUV formation and docking. In future work, we will continue to carry out the study of optimizing wavelength selection, designing wing-light patterns, and investigating new pose estimation algorithms.

Funding. Youth Innovation Promotion Association of the Chinese Academy of Sciences (Y2021044); Strategic Priority Program of the Chinese Academy of Sciences (XDC03060103); National Natural Science Foundation of China (61875189).

Acknowledgment. The authors acknowledge the financial funding of this work.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

REFERENCES

- M. Carreras, J. D. Hernandez, E. Vidal, N. Palomeras, D. Ribas, and P. Ridao, *IEEE J. Oceanic Eng.* **43**, 344 (2018).
- M. P. Brito, R. S. Lewis, N. Bose, and G. Griffiths, *IEEE Trans. Eng. Manage.* **66**, 98 (2019).
- A. Martins and Silva, *Oceans 2003. Celebrating the Past: Teaming Toward the Future (IEEE Cat. No.03CH37492)*, Vol. 1, San Diego, CA, USA, September 22–26, 2003, pp. 347–352.
- G.-X. You, Y. Pang, and D. Jiang, *J Mar. Sc. Appl.* **4**, 7 (2005).
- M. F. Fallon, G. Papadopoulos, and J. J. Leonard, *2010 IEEE International Conference on Robotics and Automation*, Anchorage, AK, USA, May 3–10, 2010, pp. 4256–4263.
- J. Manley, *2004 IEEE/OES Autonomous Underwater Vehicles (IEEE Cat. No.04CH37578)*, Sebasco, ME, USA, June 17–18, 2004, pp. 20–25.
- P. Baccou, B. Jouvencel, V. Creuze, and C. Rabaud, in *MTS/IEEE Oceans 2001. An Ocean Odyssey. Conference Proceedings (IEEE Cat. No.01CH37295)*, Vol. 3, Honolulu, HI, USA, November 5–8, 2001, pp. 1816–1821.
- G. Rui and M. Chitre, *Oceans'10 IEEE Sydney*, Sydney, Australia, May 24–27, 2010, pp. 1–6.
- K. G. Kebkal and A. I. Mashoshin, *Gyroscopy Navig.* **8**, 80 (2017).
- J.-Y. Park, B.-H. Jun, P.-M. Lee, and J. Oh, *Ocean Eng.* **36**, 48 (2009).
- Y. Li, Y. Jiang, J. Cao, B. Wang, and Y. Li, *Ocean Eng.* **110**, 163 (2015).
- S. Liu, H. Xu, Y. Lin, and L. Gao, *Sensors* **19**, 1889 (2019).
- J. Bosch, N. Gracias, P. Ridao, K. Istenič, and D. Ribas, *Sensors* **16**, 429 (2016).
- A. B. Figueiredo, B. M. Ferreira, and A. C. Matos, *OCEANS 2016 - Shanghai*, Shanghai, China, April 10–13, 2016, pp. 1–7.
- K. N. Lwin, N. Mukada, M. Myint, D. Yamada, M. Minami, T. Matsuno, K. Saitou, and W. Godou, *Artificial Life Robotics* **23**, 409 (2018).
- Y. Hamamatsu, T. Matsuda, T. Sakamaki, and T. Maki, *OCEANS 2019 - Marseille*, Marseille, France, June 17–20, 2019, pp. 1–5.
- R. Pérez-Alcocer, L. A. Torres-Méndez, E. Olgún-Díaz, and A. A. Maldonado-Ramírez, *J. Sens.* **2016**, 8594096 (2016).
- X.-S. Gao, X.-R. Hou, J. Tang, and H.-F. Cheng, *IEEE Trans. Pattern Anal. Machine Intell.* **25**, 930 (2003).
- J. Matas, C. Galambos, and J. Kittler, *Comput. Vis. Image Underst.* **78**, 119 (2000).