



**HAL**  
open science

## Vision of the Seas: Open Visual Perception Framework for Autonomous Sailing Vessels

André P D Araújo, Ganadev Prajapathy Chandrasekharan, Phillippe Preux,  
Esteban Walter Gonzalez Clua, Eduardo Charles Vasconcellos, Luiz Marcos

► **To cite this version:**

André P D Araújo, Ganadev Prajapathy Chandrasekharan, Phillippe Preux, Esteban Walter Gonzalez Clua, Eduardo Charles Vasconcellos, et al.. Vision of the Seas: Open Visual Perception Framework for Autonomous Sailing Vessels. The 30th International Conference on Systems, Signals and Image Processing, IWSSIP 2023, IEEE; Faculty of electrical engineering and information technologies - Skopje, Jun 2023, Ohrid, Republic of Macedonia, Macedonia. pp.5, 10.1109/IWSSIP58668.2023.10180295 . hal-04163513

**HAL Id: hal-04163513**

**<https://hal.science/hal-04163513>**

Submitted on 25 Jul 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Vision of the Seas: Open Visual Perception Framework for Autonomous Sailing Vessels

André P. D. Araújo<sup>1</sup>, Ganadev Prajapathy Chandrasekharan<sup>1</sup>, Esteban Walter Gonzalez Clua<sup>1</sup>,  
Phillipe Preux<sup>2</sup>, Eduardo Charles Vasconcellos<sup>2</sup>, Luiz Marcos Garcia Gonçalves<sup>3</sup>

<sup>1</sup> Institute of Computing – Universidade Federal Fluminense, 24210-346, Niteroi, Brazil

<sup>2</sup> Inria – Université de Lille, 59650, Villeneuve d’Ascq, France

<sup>3</sup> Elec. and Comp. Eng. Graduate Program – Univ. Fed. do Rio Grande do Norte, 59078-970, Natal, Brazil

andrepda@id.uff.br, ganadevprajapathy@id.uff.br, esteban@ic.uff.br, philippe.preux@univ-lille.fr

evasconcellos@id.uff.br, lmarcos@dca.ufrn.br

**Abstract**—We propose an Open Visual Perception framework for autonomous navigation and obstacle detection in marine sailing vessels. The proposed framework lays the groundwork for perceiving marine objects needed for vessel navigation and obstacle avoidance using only image processing as the primary method. Various methods are explored to extract features such as waterline, terrain, and other vessels. The proposed framework also incorporates image stabilization techniques without the need for an Inertial Measurement Unit to compensate for ocean wave oscillations. The work also introduces the system architecture of an autonomous marine vessel, besides the components of the open visual perception framework. Results demonstrate the capabilities of obstacle detection, horizon line identification, and water texture detection.

**Index Terms**—Horizon Line, Water Texture, Obstacle Detection, Stabilization

## I. INTRODUCTION

The development of autonomous robots rely on sensors such as lidars, radars, and camera images for perceiving the surrounding environment. The development of such perception systems for unmanned surface vessels (USV) of type sailboat is critical as this kind of vessel effectively uses the wind and the water current to navigate. Addressing and understanding such influence on the vessel from the weather and environment becomes imperative for developing its autonomous navigation and obstacle detection algorithms. In this experimental work, we explore the use of image processing techniques as a primary cornerstone for developing an open visual perception framework for self navigation of marine sailing vessels.

Extracting features from images is an essential tool to enable the vessel to distinguish the waterline, shoreline, rocks and mountains, and other vessels. Identifying the waterlines from images allows the boat to understand which area is safe to move. Estimating the distance of the obstacles around the vessel helps understanding which obstacles are a threat and at priority to avoid. By identifying the vessel’s angle of tilt coupled with the wind speed, it is possible to predict the type of sail maneuver that is safer to perform. In this

---

We thank CAPES-Brazil, grant number 88881.506890/2020-01, and CNPq-Brazil, grant number 311640/2018-4, for funding this research. Ph. Preux acknowledges the support of the Métropole Européenne de Lille (MEL), ANR, Inria, Université de Lille, through the AI chair Apprenf number R-PILOTE-19-004-APPRENF, as well as the Scool research group.

work we will introduce different image analysis methods [1] that we explored, implemented and incorporated in our open framework.

The data source for this project is collected by using the ZED stereo camera [2], [3], which is a commercially accessible stereo vision [4], [5] solution from Stereolabs [6], and which is positioned on the bow of our sailing vessel. Notice that the proposed open visual perception framework involves the identification of horizon line, vessel tilt/heel angle, terrain, coastline, and other marine vessels, even under the presence of reflections, dynamic illumination changes, and oscillations caused by ocean waves and wind.

The main goal is to lay the groundwork to allow the perception of all kinds of marine objects, which are necessary for the vessel navigation and obstacle avoidance. Nonetheless, we also need to stabilize the video in relation to the vessel motion, for the perception of objects and to make the framework adaptable to other marine vessels. We propose a method to image stabilization that compensates for the oscillation caused by ocean waves, by calculating the vessel tilt/heel angle, without the use of any Inertial Measurement Unit (IMU) devices.

The ZED stereo camera is provided with a Software Development Kit (SDK). Using ZED SDK, a 3D reconstruction of the environment is generated from the perspective of the surface vessel. Using this virtual 3D environment, we are able to estimate the distance of the objects detected around the surface vessel by applying the depth disparity from the ZED stereo cameras. With the depth map, the Open Visual Perception Framework can actively recognize marine objects within the range of 20 meters, allowing for future manipulations of optical flow and movement predictions.

## II. LITERATURE REVIEW

Cameras are undoubtedly the most popular sensors used for land-based autonomous robots projects. Single camera images can be used for horizon line detection, while stereo vision cameras can be used to enhance the estimation of objects distance. The combination of both enables autonomous navigation. From a marine perspective, a horizon line is the boundary between the sky and the waterline. The calculation of such a line allows the distance estimation of objects from the marine

vessel. Existing methods normally use a predetermined region of interest to identify the horizon line [7] or use convolution neural network to understand water surface [8]. However, in order to have a predetermined region of interest, the images from the camera are expected to be stabilized, something that is not easy to obtain on a vessel that is continuously in motion, due to the wind, waves and water movement [9].

The angle of heel is the inclination of the vessel from its horizontal equilibrium. A sailing vessel always utilizes the wind and waves to propel. Thus, calculating the angle of heel is paramount in deciding the safer sail maneuver to maintain equilibrium and efficiency. The estimation of the angle of heel normally depends on external sensors such as IMU. As said above, in this work we explore a method to calculate the angle of heel using the images from the marine vessel alone.

For depth estimation, stereo vision techniques [10] prove to be a valuable tool. From a stereo camera, identifying and correlating multiple points that reference the same object from the different perspective images allow the calculation of the camera parallax. Thereby allowing the determination of the geometry of the scene and the depth of the projected pixels. Faugeras [11] proposed the stereo pair calibration technique, which became popular. In this work, for the calibration of a stereoscopic pair, only one image from each camera is necessary and a set of five reference points from each image which point to the same object or scene.

Accurate calibration of stereo cameras are critical for stereoscopy to generate accurate distance estimation. This information allows mobile robotic systems to understand and navigate through their surrounding environment. Obstacle detection and trajectory re-planning are fundamental for the safety of the mobile robotic systems. Stereo camera image based methods offer lot more information about the surrounding environment as opposed to works that only use sensors such as ultrasound radars, LIDARS, etc. In this work, we explore distance estimation and understanding of surrounding environment with stereo cameras.

### III. SYSTEM ARCHITECTURE FOR AUTONOMOUS MARINE VESSELS

The conceptual system architecture for autonomous marine vessels is composed of the the following modules: Sensors, Data Processing, Visual Perception, Decision and Planning and Control System. Figure 1 illustrates the overall components including both hardware and software modules of the system architecture. Sensors module specifically include all the hardware equipment and accessories for collection data about the surroundings to be used later for environment perception and navigation. In this work, it primarily includes stereo cameras, and GPS sensors. More sensors such as sonars, Lidar, IMU, etc can be included but are however beyond the scope of this work.

Data Processing module includes all calibration and synchronization processes between the sensors needed to establish reliable data for further manipulations. Within this module, the stereo cameras are used to generate the depth map from

disparity. The data fusion of the GPS data and camera outputs are also synchronized so the real time location is available for the perception module.

Visual Perception module includes all processes and techniques involved with detecting and understanding the pose (position and orientation) of the autonomous vessel. The decision and Planning module assembles the comprehensive feature map from the Visual Perception module to understand the obstacles and the priority map for triggering trajectory re-planning.

Finally, the Vessel Control module includes all hardware control systems needed for the marine vessel motion. For our sailboat, the control module includes the control systems to change the direction of the sail surface and rudder for steering, breaking or accelerating.

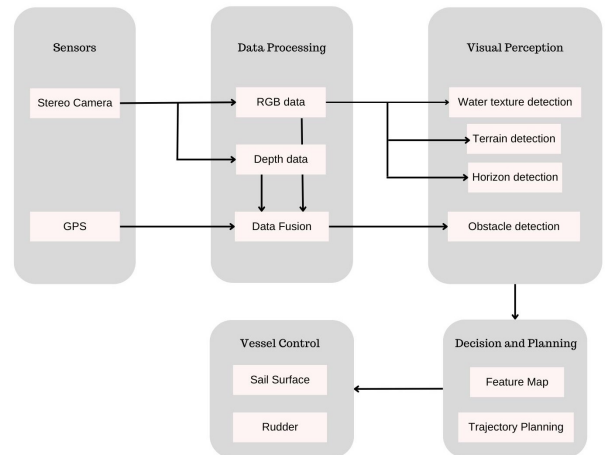


Fig. 1. System architecture for autonomous marine vessel.

### IV. COMPONENTS OF OPEN VISUAL PERCEPTION FRAMEWORK

One of the fundamental system requirements is the perception of the environment in which its navigating. The need to understand the location of the surrounding obstacles, their velocities and its future trajectory with respect to the autonomous vessel is paramount to ensure the safety of the autonomous vessel. In this work, we specifically tackle the challenge of understanding the environmental perception using the stereo cameras mounted on the bow of the autonomous vessel.

The data processing module collects the raw image data from the ZED camera, which is installed on the bow of the marine vessel to generate depth data map and Red-Green-Blue(RGB) data.

Perception module extracts several environmental information from the data processing module, which are used to model the environment feature map. It includes the identification of objects, the estimation of their distances from the vessel and provides the image stabilization using horizon detection. To extract information such as terrain, horizon line, water texture,



Fig. 2. Reference image captured by the camera onboard the marine vessel.

and angle of the heel, various image analysis methods are explored, as seen next.

#### A. Obstacle Detection

Enhanced in relation to its previous model [3], the ZED-2 camera has now a built-in artificial neural network (ANN) that aims to reproduce human vision [6]. This ANN is connected to a Digital Signal Processor and contributes, jointly, to create the depth map. From this depth map and RGB data, the ZED SDK groups the different depth clusters as objects and sets tracking id to individual objects using clustering techniques. The position of detected objects in the world map is estimated. With its coordinates  $(x, y, z)$ , the distance between the object and the vessel is calculated using the Euclidian distance. Based on the distances of the detected objects, the framework prioritizes and classify which objects are obstacles to be avoided, with the closest ones being avoided first. With the ZED SDK object tracking id, each object velocity is tracked independently. This information is then passed on to the decision and planning module.

#### B. Horizon Detection

Detecting the horizon line between the sky and the waterline can give an understanding of the orientation of the vessel with respect to the water surface. For extracting the horizon line, the image is first converted to grayscale, and then processed with the Canny Edge detector algorithm for border extraction. To get longer edges, its grayscale output are smoothed. The output from the edge detection also recognizes the edges of the boat to which the camera is attached to. Since it will interfere with the detection of the horizon line, we remove the boat from the image for further operations. To do so, we use a mask, related to the specific boat, which is subtracted from the original image. The edge detection generates a list of edges. By utilizing the Hough line transform, we extract a single horizon line that encompasses a set of edges that satisfy the threshold condition which is the minimum number of intersections needed to detect a line.

The Hough line transform algorithm is used to identify collinear edges from the list generated from the edge detection algorithm. For this, the thresholds for Hough line transform was a minimum of 5 edges to generate the horizon line estimation. The prominent Hough line that is identified from the algorithm is the horizon line. With this horizon line, its

possible to stabilize the images. To identify the angle in which the image has to be rotated, the slope of the horizon line is calculated using the linear equation  $(y = mx + b)$ . Based on this angle, the image is rotated, thus stabilizing the image horizontally.

#### C. Extracting Terrains

To extract terrain features such as rocks, boulders, hills, and coastal areas from the images, we explore image histogram filtering [1].

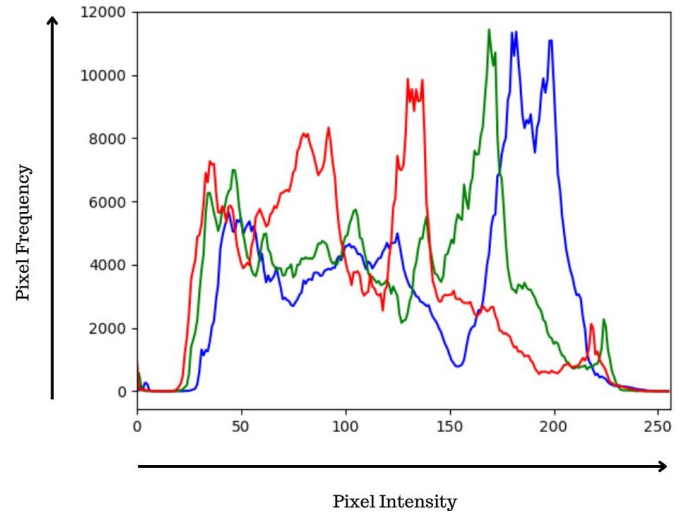


Fig. 3. Histogram data of the reference image Fig. 2.

Figure 3 shows the histogram data of the reference image (Figure 2) obtained with the marine vessel. From the histogram data, the mountainous area within the image is analyzed. To manipulate the pixels with value within a defined range, we use histogram band-pass filtering. This method can be used to do extract elements in the image that have distinctive RGB profiles. By identifying the pixel range of mountain regions from all Red, Green and Blue channels from the histogram of the reference image, it becomes possible to extract these elements using histogram band-pass filtering technique.

However, there can be potential false positives as the colors from the water body can change dynamically due to illumination changes and can be very similar to the pixel values for mountain region. To eliminate such false positives, it becomes paramount to understand and identify the water surface and eliminate the water surface region from being included in the histogram band-pass filtering operation for detecting mountains.

#### D. Water texture detection

To extract the water surface from the reference image, the grayscale image can be applied with the entropy function from Skimage [12]. For an image, local entropy is related to the complexity contained in a given neighborhood, typically defined by a circular structuring element of a radius that is



big enough to capture the local gray-level distribution. The entropy filter can detect subtle variations in the local gray level distribution. The wave patterns of the water surface can be differentiated using texture detection.

The entropy function gives a value that represents the level of complexity in a certain section of an image. To compute the local entropy, a circular structuring element disk is applied. For the reference image, the radius of the structuring disk is set to 6.

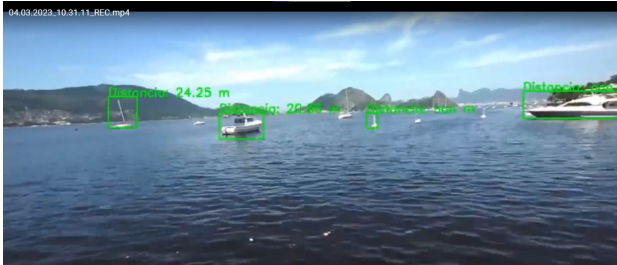


Fig. 4. Obstacles detected with the distances.

## V. RESULTS

The accurate perception of the surrounding enables reliable and safe operational efficiency of USV in marine applications. Analyzing moving obstacles requires a combination of robust algorithms, equipment and processing. To evaluate the modules of the proposed Open Visual Perception framework, the results from each said modules are analyzed.

Figure 4 shows the obstacle detection algorithm tracking the obstacles and its distance from the marine vessel. The depth data from the stereo camera could be used to accurately estimate the distance of the obstacles of up to 20 m from the USV.

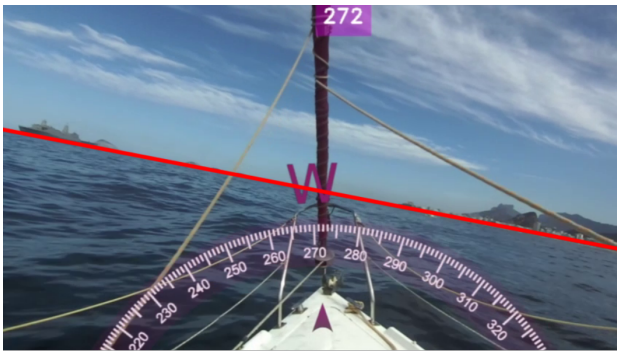


Fig. 5. Horizon line detected.

Figure 5 shows an accurate horizon line detected, between the sky and the waterline, even when there are other vessels and terrain surfaces in the horizon. These satisfactory detected horizon lines were enough for calculating of the heel angle.

Figure 6 shows the histogram band-pass filtering technique extracting terrain details from the image. Figure 7 shows the stabilized image using the angle of heel, horizon line and the

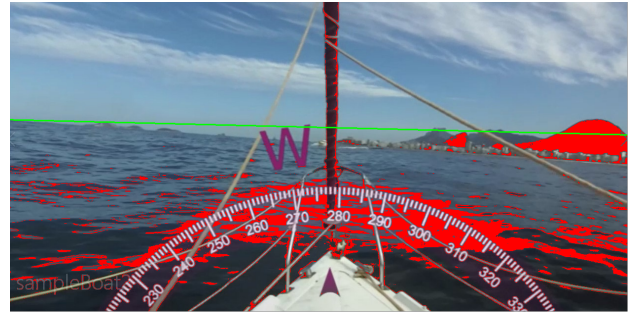


Fig. 6. Output from histogram band-pass filter.

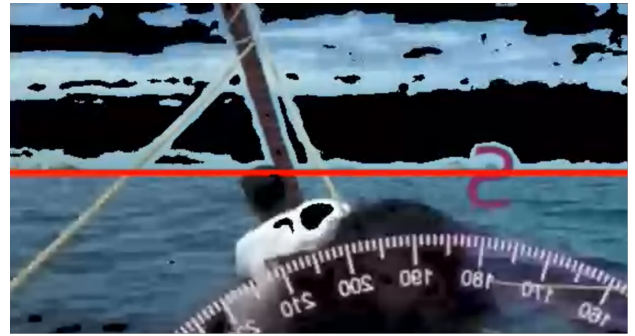


Fig. 7. Water texture detection after the stabilization using horizon line.

water texture highlight using entropy function. It clearly voids highlighting mountain surfaces in the image.

## VI. CONCLUSION

Perception plays a critical role in the development of autonomous vessels enabling them to gather information about their environment and make informed decisions. Even though cameras are used in both land based and marine systems, the perception needed for marine applications is by far different. This brings in a need for a standardization of perception algorithms that can be adapted from one marine vessel to another. Thus the proposed Vision of the Seas: Open Visual Perception framework enable the understanding of different environmental components from the onboard images from the marine vessel.

By leveraging these proposed perception algorithms within the framework, features such as horizon line, terrain, water surface texture, obstacles and other vessels are extracted from the data set collected from marine sailing boat. Thus allowing autonomous vessels to operate safely, adapt to changing environments, and navigate complex maritime scenarios.

The Open Visual Perception Framework is intended to be a starting point for different perception algorithms for marine applications. As such, some of the future works related to the perception techniques include utilizing machine learning techniques for recognizing the type of obstacles such as other vessels, boils, rocks etc. Such future works related to marine applications should be accommodated into the framework.

## REFERENCES

- [1] S. Milan, V. Hlavac, and R. Boyle, *Image processing, analysis, and Machine Vision*, 3rd ed. Cengage Learning, 2015.
- [2] L. Ortiz, E. Cabrera, and L. Gonçalves, "Depth data error modeling of the zed 3d vision sensor from stereolabs," *Electronic Letters on Computer Vision and Image Analysis*, vol. 17, 04 2018.
- [3] E. V. Cabrera, L. E. Ortiz, B. M. F. d. Silva, E. W. G. Clua, and L. M. G. Gonçalves, "A versatile method for depth data error estimation in rgb-d sensors," *Sensors*, vol. 18, no. 9, 2018. [Online]. Available: <https://www.mdpi.com/1424-8220/18/9/3122>
- [4] E. Trucco and A. Verri, *Introductory techniques for 3-D computer vision*. Prentice Hall, 1998.
- [5] R. Szeliski. (2011) *Computer vision algorithms and applications*. London; New York. [Online]. Available: <http://dx.doi.org/10.1007/978-1-84882-935-0>
- [6] Stereolabs, "Zed 2: Built for the spatial ai era. combine 3d sensing with ai to create next-generation spatial intelligence." [Online]. Available: [https://arxiv.org/abs/2203.10493\(05/23/2023\)](https://arxiv.org/abs/2203.10493(05/23/2023))
- [7] C. Y. Jeong, H. S. Yang, and K. Moon, "Fast horizon detection in maritime images using region-of-interest," *International Journal of Distributed Sensor Networks*, vol. 14, no. 7, p. 1550147718790753, 2018. [Online]. Available: <https://doi.org/10.1177/1550147718790753>
- [8] L. Steccanella, D. Bloisi, A. Castellini, and A. Farinelli, "Waterline and obstacle detection in images from low-cost autonomous boats for environmental monitoring," *Robotics and Autonomous Systems*, vol. 124, 11 2019.
- [9] U. Ganbold and T. Akashi, "The real-time reliable detection of the horizon line on high-resolution maritime images for unmanned surface-vehicle," in *2020 International Conference on Cyberworlds (CW)*, 2020, pp. 204–210.
- [10] D. Huynh, "Feature-based stereo vision on a mobile platform," Ph.D. dissertation, University of Western Australia, 10 1997.
- [11] S. J. Maybank and O. D. Faugeras, "A theory of self-calibration of a moving camera," *International Journal of Computer Vision*, 1992.
- [12] S. v. d. Walt<sup>1</sup>, J. L. Schönberger, J. Nunez-Iglesias<sup>3</sup>, F. Boulogne<sup>4</sup>, J. D. Warner<sup>5</sup>, N. Yager<sup>6</sup>, E. Gouillart<sup>7</sup>, T. Yu<sup>8</sup>, T. S.-I. Contributors, B. S, and et al., "Scikit-image: Image processing in python," Jun 2014. [Online]. Available: <https://doi.org/10.7717/peerj.453>