Vision-Based Autonomous Surface Ship

Eric Nieves, Travis Moscicki, Jared Wampler, Armando Sinistera, Andrew Tarantul Soto
College of Engineering and
Computer Science
Florida Atlantic University
Boca Raton, Florida 333431

***Abstract*—The Florida Atlantic University Marine Robotics RoboBoat Team has designed the 7th generation of the Vision-Based Autonomous Surface Vessel (V-BASS) for the 2018 AUVSI RoboBoat Competition. Major improvements to the vehicle occur mainly in the vision, navigation, control, and electronics systems. Some small adjustments were made in the acoustics module. Once agglomerative hierarchical clustering (AHC) detected objects of interest, The RGB values from a camera were projected onto their 3D LiDAR point cloud representation. The results were fed into a classifier and the information was stored in the object persistence checker (OPC), which acted as a look up table for unique objects encountered by V-BASS. This vision pipeline allowed efficient recognition while preserving important data. In the Navigation pipeline, simultaneous localization and mapping (SLAM) was used. A global planner acted as a long term mission planner whereas the a local planner provided updates for obstacle avoidance.**

Fig. 1. Mechanical Structure of V-BASS

I. Competition Strategy

The Marine Robotics Club at FAU strategic vision included making the most of what systems were already in place and innovating aspects that aligned with the strengths of the team members. Specific areas of knowledge that the team displayed including signal processing, data analysis, machine learning, computer vision, autonomous navigation for ocean vehicles. Previous iterations of V-BASS included some detailed electrical and mechanical components that were leveraged to cover weaknesses of the team. Therefore, innovations persisted mostly in the vision (detection, classification, etc.) and navigation pipelines. Since the acoustics pipeline was reliable in previous competitions, it was a lower priority and less crucial updates were made.

The following subsections will talk about the priorities and focus for the vehicle design as well as each individual challenge.

A. Vehicle Design

The structure of the vehicle was largely re-used and can be seen in Fig.1 [1].

The V-BASS platform is a catamaran type vessel with an length overall of 58” and beam length of 36”. The propulsion relies on two brushless DC motors, one on each hull just forward of the transom, to apply differential thrust. A Nickel Cadmium bank is stored in each hull to provide the onboard electrical power. Mechanical remodeling was avoided because the return of time investment would be low. Since the beam length of the vehicle is 3 feet, it fulfilled the tighter space possibilities in the automated docking and find the path tasks (4 feet) in this year’s competition. However, the electronics box received some updates to fit the competition strategy and is explored later in the paper.

B. Introductory Tasks

The introductory course strategy relies purely on detection and color classification, since there is no need for shape recognition. However, as an elective priority, the team sought to use shape classification to allow V-BASS to autonomous detect the challenges and perform them. This would remove as much human interaction as possible.

The circumnavigation, maintaining heading, and slalom maneuvers all depend on our new vision system. For circumnavigation, since the specific path can change from day to day, the user will simply choose which option is expected and the vision system will properly execute the challenge. There is no need for this during the maintaining heading challenge, however, as the vision system can determine a clear line of sight to follow no matter the buoy’s configuration and spacing. For the slalom course, V-BASS was programmed to traverse to the left, traverse to the right, or circumnavigate depending on the color of the buoy.

C. Mission Tasks

Although the start gates are mandatory to began the advanced course, the vision pipeline was capable of handling the task reliably and focus shifted to the more complicated tasks in the course.

1) Speed Challenge: Detection and color classification were chosen to position the vehicle and start the challenge. The vision system was used in conjunction with the controls system to determine the highest velocity capable that would allow the V-BASS to effectively circumnavigate the blue buoy. On the return trip, V-BASS was programmed to put thrust on full to minimize completion time.

2) Automated Docking: Due to the complexity of adding an unmanned aerial vehicle (UAV), it was decided to focus on the acoustic localization portion of the challenge and use the vision system to navigate to the proper dock. Once this is accomplished, V-BASS was programmed to leave the dock and choose one of the other docks to show consistency in the algorithm.

3) Find the Path: Whereas most challenges, when entered, only need color classification, this challenge must rely on the shape classification. Therefore that was chosen to help V-BASS correctly find the can buoy and complete circumnavigation. To reduce complexity, the path chosen to leave was the same path the boat entered through.

4) Follow the Leader: Once stationed at the challenge, color classification was decided to correctly identify the red flag. Then the autonomous vehicle attempt to keep a circular trajectory just to the outside of the red flag.

II. Design Strategy

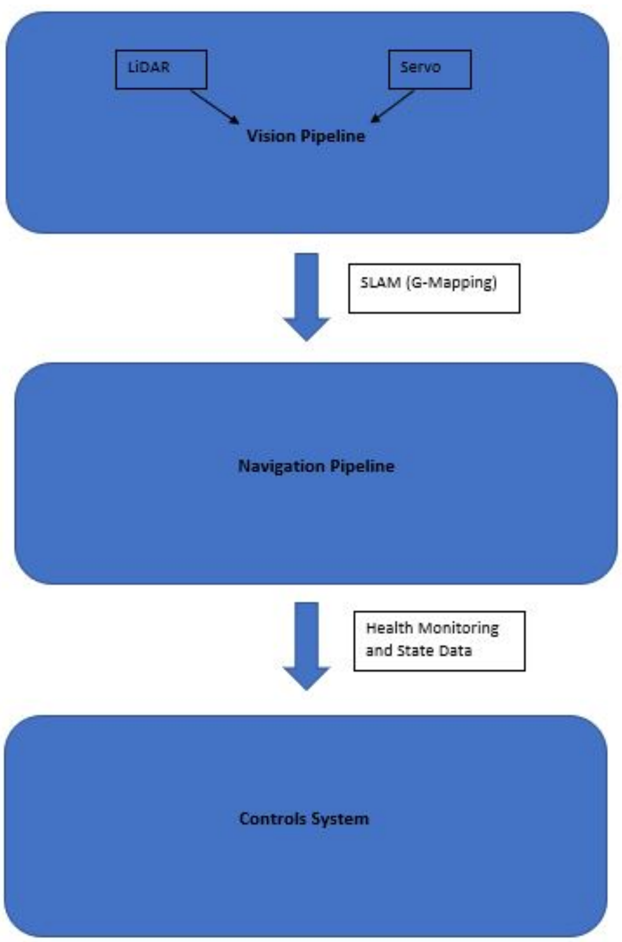

Fig. 2. Overview of the System Pipelines

A. Vision System

A flow chart of the vision system can be seen in Fig. 3.

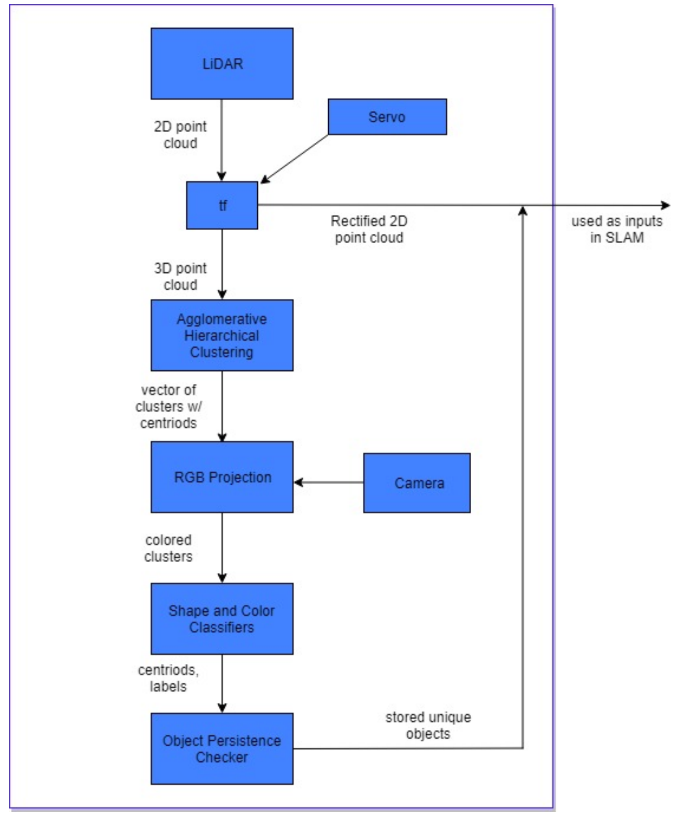

Fig. 3. The Vision Pipeline

A Hokuyo LiDAR sensor is the base vision tool, which relies on a transform function with servo inputs to turn 2D scans into a 3D point cloud. The 3D image then enters the agglomerative hierarchical clustering algorithm (AHC) to provide detection of important objects such as buoys associated with the challenge. These buoys are isolated from the larger point cloud. For the isolated objects, a parallel camera system projects the RGB values onto the object point cloud, creating a sparse representation of the object.

These colored clusters enter a shape detector, which has a library of bounding boxes based off the typical shapes used throughout the challenge. The means of the detected objects' bounding box dimensions are compared with the library and used to determine correct shape classification. To provide color classification, histogram matching is leveraged. Once the classification is done, the labels and centroid of the objects are output.

The labels are then fed into the Object Persistence Checker (OPC). This tool is used as a look up table for unique objects. If the detected object is one already seen before within some tolerance threshold (to account for buoy drift), then the position of that particular object is updated. Otherwise the OPC creates a new entry for it.

B. Navigation System

Simultaneous Mapping and Localization (SLAM) is used to build global and local planners. In the move-base node of the navigation stack of Fig. 4 [2], SLAM is used in place of the optional provided nodes for a map server and adaptive monte carlo localization. This is because buoys and other obstacles in the water environment may change position, and thus relying on a pre-mission map may not be a good strategy.

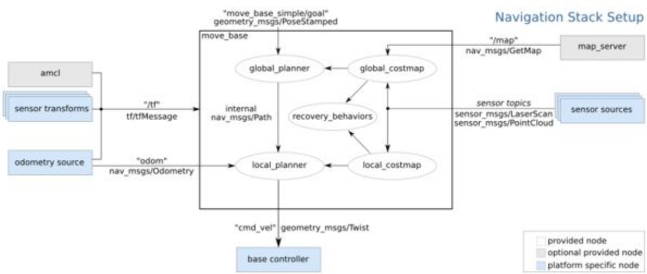

Fig. 4. Simple Move-Base Module

For the vehicle, the sensor sources in the platform specific node includes the LiDAR and camera (in the vision pipeline). The vehicle relies on a IMU-GPS coupled device with an internal extended Kalman filter to feed in the odometry data.

The rest works just as pictured in Fig. 4. Two cost maps are updated. The global map and sensor sources are used to update the global cost map, whereas only the sensor sources update the local cost map. These cost maps are fed into their corresponding planners.

A global planner starts empty and experiences object population as V-BASS navigates the course, allowing a long term mission plan to reach a particular way point. To avoid running into objects, the SLAM algorithm takes a rectified 2D point cloud from the vision system and inputs it into a local planner. This planner is used to detect obstacles that might have moved within the vehicle's path, providing a powerful update technique. The combination of planners are used to feed state data in the base controller.

Since SLAM builds a field view that does not uniquely label objects, the OPC entries can be projected onto the global map to make sure each object has a persistent set of coordinates.

Fig. 5 illustrates the frames and their corresponding transforms that the navigation system must consider. Both the map and the odometry frame will share the same origin and orientation, but they are used for different purposes. The map frame origin is positioned conveniently around the navigation area. The odometry data is represented in the odometry frame and is provided by the IMU-GPS Xsens sensor, which defines the odometry to base link (frame at the robot's center of rotation) transformation. Finally, the base link to laser link transformation is defined as a static transform which locates the LiDAR (and corresponding laser data) relative to the base link.

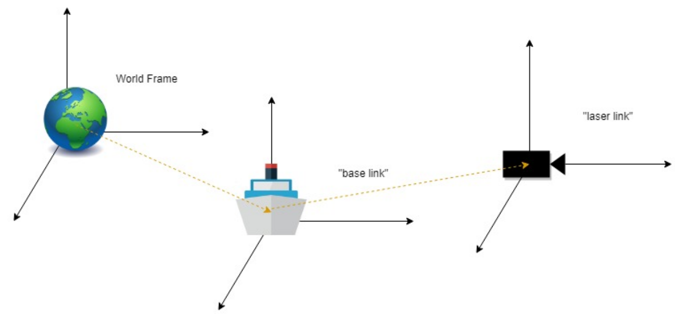

Fig. 5. Reference Frame Transformations

C. Controls System

The control strategy chosen for the vehicle was a multiple input multiple output (MIMO) transiting backstepping controller as developed in [3]. Using an appropriately tuned Lyapunov exponent gain matrix, this controller was shown to overcome environmental disturbances that could not be modeled. This controller implementation also remedies issues of trying to apply gain-scheduling techniques to a linear model during simultaneously control of surge, sway, and yaw.

D. Electronics Box

The inspiration for the electronics box was largely derived from [4]. Mission critical components in the electronics box include the Jetson TK1 Compute boards, a STM32 microcontroller, power distribution circuits, telemetry systems, and the motor control interface. Emphasis was placed on making sure each component can act individually and concurrently without overloading the power distribution systems.

The Jetson TK1 Compute boards were chosen to act as the CPUs for the system. The design ensured that these boards were able to communicate with the other systems during operation of all condition. In order to sync this communication with accurate timing, a Real Time Clock (RTC) was chosen to be battery operated so that during communication failure, the Jetsons would still have the correct time.

The STM32 Microcontroller acts as the Bottom Level Manager (BLM), where the low level real time devices are managed. This board has a primary function of measuring capacity in the battery supplies and indicating their state to the system. This was important to ensure the power distribution system runs properly.

The telemetry system has two channels, a primary link with 80 Mb/sec used for link the vehicle to the ground station as well as a back-up link with 115Kb/sec for mission critical data.

The motor control (PWM) interface was designed to provide hardware control over the motors from shore with a standard RC remote. A multiplexer control line originates from a PIC microcontroller with a singular job of detecting which way a switch is set on the RC remote, exceeding the level of control a user could have got from software solutions. A rendered view of the motherboard can be seen in Fig. 6

Fig. 6. Rendering of Motherboard

E. Acoustics Module

The acoustic signal received by a hydrophone can be described by:

$$s_i(t) = s_0 sin(2\pi f_p(t - t_i)) \tag{1}$$

where i denotes the specific hydrophone. Whereas time difference of arrival methods can be derived [5], the method used relied on the an analogous phase difference to determine the bearing of the acoustic signal. If there is a pair of hydrophones, then the fast fourier transform (FFT) of each received signal can be calculated (F_1 and F_2). The ratio of the FFTs can be expressed as:

$$\frac{F_1}{F_2} = |\frac{F_1}{F_2}|e^{j(\phi_2 - \phi_1)} \tag{2}$$

The phase differential can therefore be calculated by solving:

$$\phi_2 - \phi_1 = arctan[\frac{Im(\frac{F_2}{F_1})}{Re(\frac{F_2}{F_1})}] \tag{3}$$

The bearing of the acoustic signal is typically expressed for time difference of arrival methods as:

$$\theta_{12} = arcsin[\frac{c_w(t_2 - t_1)}{L}] \tag{4}$$

where c_w is the speed of sound in water and L is the distance between hydrophones, or the baseline length. L should be the half the minimum wavelength.

The phase difference is related to the time delay via:

$$\phi_2 - \phi_1 = 2\pi f_p(t_2 - t_1) \tag{5}$$

Therefore, substituting Eq. 5 into Eq. 4 creates the following equation:

$$\theta_{12} = arcsin[\frac{c_w(\phi_2 - \phi_1)}{2\pi f_p L}] \tag{6}$$

Using two pairs of hydrophones and performing least squares analysis allows the system to resolve a 2D bearing to the acoustic signal. Fig. 7 shows the hydrophone array used on V-BASS.

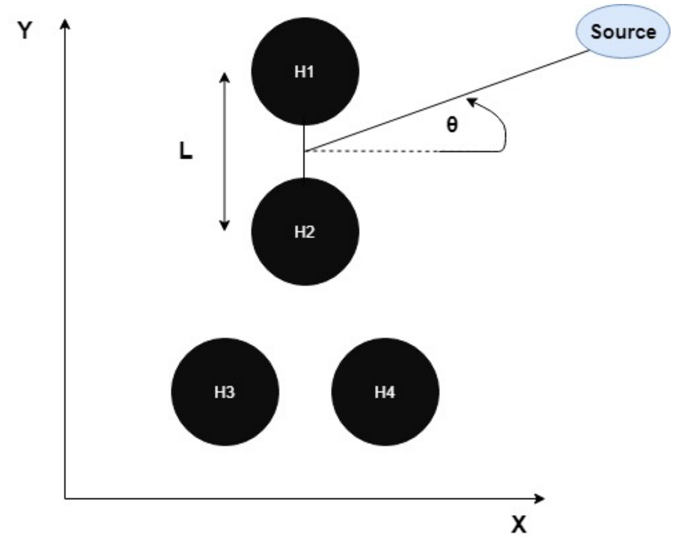

Fig. 7. A diagram of the hydrophone set-up (not drawn to scale). The distance between hydrophones should be set to half the minimum wavelength seen by the system to ensure it stays within - π to + π wrapping

In previous competitions, the motors have produced frequency disturbances within the 25 - 40 kHz range. While in competition, the motors would be turned off and the vehicle was allowed to freely drift during the acoustic bearing calculation. In order to avoid this, hardware filtering solutions were explored.

III. Experimental Results

A. Acoustics Experiments

Fig. 8 and Fig. 9 showcase some results of the acoustic positioning system onboard V-BASS. A source was used to create a 25 kHz signal and two different data collections were initiated. Both figures show the frequency and bearing of the source being resolved. Using these results, insights on how to improve the system were noted.

B. Vision Experiments

V-BASS was controlled to pass by the two buoys in Fig. 10 while the LiDAR was operating. Not only did the LiDAR detect the buoys, but the AHC algorithm was able to properly cluster them, as seen in the tree of Fig. 11. In Fig. 12, you can clearly see the two buoys detected as separate clusters with their points designated a unique number (1 and 2).

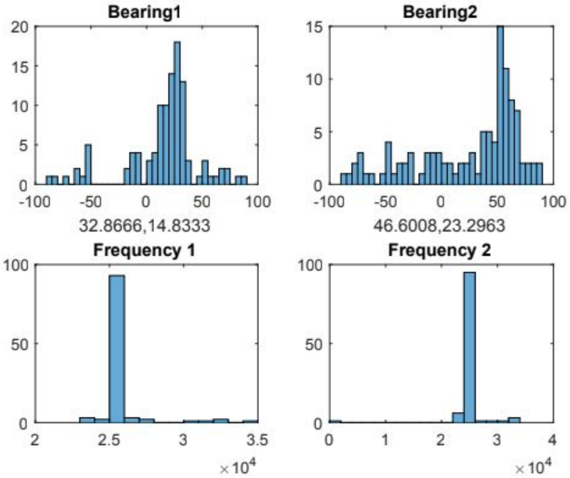

Fig. 8. The acoustics position module locating a 25 kHz emitting source

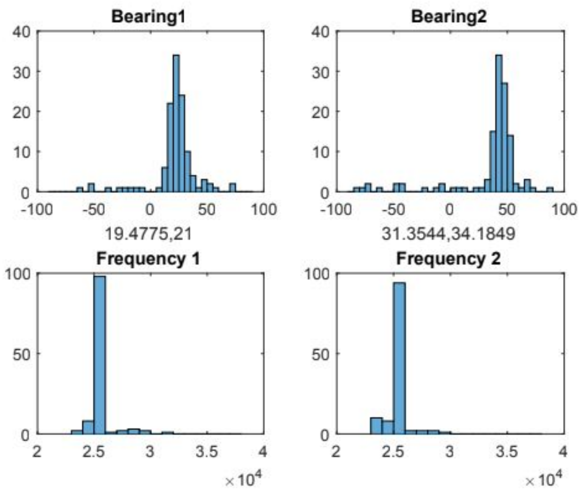

Fig. 9. Another example of the acoustic positioning system localizing a 25 kHz signal

Fig. 10. The set-up view of V-BASS as it passes over the buoys with a LiDAR

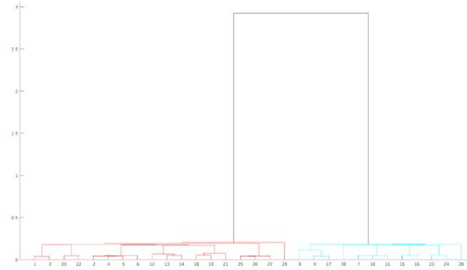

Fig. 11. Dendrogram of the AHC output, which shows how points are grouped and split into unique objects

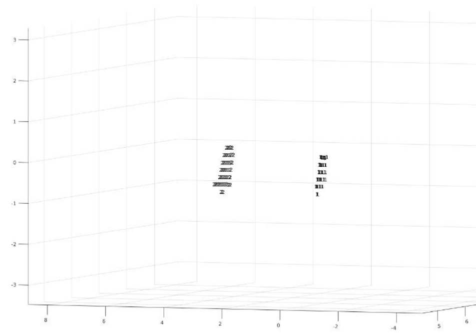

Fig. 12. The clusters for the two buoys plotted together with their unique markers

ACKNOWLEDGMENT

The authors would like to thank Manhar Dhanak and Karl Von Ellenrieder for their advisory roles and experience in project development. The team is also grateful for Ed Henderson and John Kielbasa for their invaluable assistance in electrical design.

APPENDIX

Component	**Vendor**	**Model**
ASV Hull form	Developed in house	Bailey Long Hull
Waterproof Connectors	Switchcraft	Con-x
Propulsion	SeaBotix	BTD150
Power systems	Battery-Space	LiFePo4
Motor Controls	RoboTeq	SDC2130
CPU	Nvidia	Jetson Tk1
Compass	OceanServer	os5000
IMU	Xsense	mti-g 700
Cameras	Logitec	c210
LiDAR	Hokuyo	utm-30lx
Team size	3 Students	1 Master
	1 PhDc	1 PostDoc
Testing Time: Simulation	40 hours	
Testing Time: In Water	30 hours	
Programming Language(s)	c++	ros
	matlab	python

REFERENCES

[1] Moscicki, Travis, "Vision-based autonomous surface ship 6th generation," Boca Raton, FL, 2016.

[2] O. S. R. Foundation, *MoveBase*, 2016.

[3] E. I. Sarda, H. Qu, I. R. Bertaska, and K. D. von Ellenrieder, "Station-keeping control of an unmanned surface vehicle exposed to current and wind disturbances," *Ocean Engineering*, vol. 127, pp. 305 – 324, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0029801816304206

[4] Bianca Mesa, Travis Moscicki, et. al, "Team worx: Journal paper for the 2016 robotx competition," Boca Raton, FL, 2016.

[5] A. K. Tellakula, "Acoustic source localization using time delay estimation," Master's thesis, Indian Institute of Science, Bangalore, India, 2014.