Evaluation of an acoustic detection algorithm for reactive collision avoidance in underwater applications

by

Oscar Alberto Viquez Rojas

Submitted to the Department of Mechanical Engineering in partial fulfillment of the requirements for the degree of Bachelor of Science in Mechanical Engineering at the MASSACHUSETTS INSTITUTE OF TECHNOLOGY June 2013

© Oscar Alberto Viquez Rojas, MMXIII. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Author
Department of Mechanical Engineering
May 10, 2013

Certified by......
Henrik Schmidt
Professor of Mechanical and Ocean Engineering
Thesis Supervisor

Accepted by .................
Anette Hosoi
Professor of Mechanical Engineering
Undergraduate Officer
Evaluation of an acoustic detection algorithm for reactive collision avoidance in underwater applications

by

Oscar Alberto Viquez Rojas

Submitted to the Department of Mechanical Engineering
on May 10, 2013, in partial fulfillment of the requirements for the degree of Bachelor of Science in Mechanical Engineering

Abstract

This thesis sought to evaluate a vehicle detection algorithm based on a passive acoustic sensor, intended for autonomous collision avoidance in Unmanned Underwater Vehicles. By placing a hydrophone at a safe distance from a dock, it was possible to record the acoustic signature generated by a small motor boat as it navigated towards, and then away from the sensor. The time-varying sound intensity was estimated by Root Mean Square of the sound amplitude in discrete samples. The time-derivative of the sound intensity was then used to estimate the time to arrival, or collision, of the acoustic source. The algorithm was found to provide a good estimate of the time to collision, with a small standard deviation for the projected collision time, when the acoustic source was moving at approximately constant speed, providing validation of the model at the proof-of-concept level.

Thesis Supervisor: Henrik Schmidt
Title: Professor of Mechanical and Ocean Engineering
Acknowledgments

I would like to thank my thesis advisor, Prof. Henrik Schmidt, for his guidance, support and encouragement throughout this work. His suggestions and comments, especially during the analysis phase, helped make this project a reality.

My gratitude also goes out to Dr. Michael Benjamin, who first suggested the concept of this project as an area of interest for me.

Furthermore, I would like to thank Sheida Danesh, PhD Candidate at MIT’s Laboratory for Autonomous Marine Sensing Systems, for her patience and support, especially during the experimental phase of this thesis.

I would also like to express my deepest gratitude to Alon Yaari, Research Software Engineer at MIT’s Laboratory for Autonomous Marine Sensing Systems, for his invaluable contributions to this project. Without his creativity, skill and support, this thesis would not have been possible.

Finally, I would like to thank my parents and siblings, who have always given me their best wishes, unwavering support and words of encouragement.
List of Figures

4-1 Basic experimental setup for passive acoustic detection system. . . . . 21

5-1 Boat position versus time, with switch sensor triggers marked. . . . . 23
5-2 Boat position through pass 7, April 26. . . . . . . . . . . . . . . . . 24
5-3 Boat distance to hydrophone from GPS data through pass 7, April 26. 24
5-4 Time to collision estimate, sound intensity, boat speed and standard
deviation of collision time through pass 7, April 26. . . . . . . . . . . . 25
5-5 Time to collision estimate, sound intensity, boat speed and standard
deviation of collision time through pass 5, April 26. . . . . . . . . . . . 26
5-6 Time to collision estimate, sound intensity, boat speed and standard
deviation of collision time through pass 3, April 26. . . . . . . . . . . . 27
5-7 Boat position through pass 2, April 26. . . . . . . . . . . . . . . . . 28
5-8 Time to collision estimate, sound intensity, boat speed and standard
deviation of collision time through pass 2, April 26. . . . . . . . . . . . 29
1 Introduction

The last few decades have seen considerable amounts of research in autonomous vehicles and their applications. Specially in recent years, projects such as the Unmanned Aerial Vehicle (UAV) Predator or the Unmanned Ground Vehicle (UGV) BigDog have become of great interest in military applications [5], while Google’s self-driving car has attracted the attention of the general public. As with unmanned aerial and ground vehicles, their marine counterparts have also seen a renovated interest from the scientific community. Unmanned Surface Vehicles (USVs) and Unmanned Underwater Vehicles (UUVs) have caught the attention of scientists, the off-shore industry and the military, for their applications in areas such as sea bottom exploration, hull inspection and repairing, mine-hunting and economic zone policing [10].

"The need for monitoring and securing harbor environments has grown in recent years, as a result of increased attention to pollution from runoff or other sources, natural processes such as sediment transport, water properties, and algal blossoms, as well as security against threats" [1]. This need has in turn increased the use of USVs and UUVs in high-traffic areas such as harbors and the surrounding littoral, for security and scientific applications alike. However, the highly dynamic nature of these environments requires that autonomous vehicles be capable of reactive collision avoidance [5].

While USV autonomous systems are now at a sufficiently advanced level of maturity for implementation in harbor observation missions, there is still much work to be done in obstacle avoidance for underwater vehicles [6]. This thesis continues development on the subject by evaluating a passive acoustic detection algorithm presented by Prof. Henrik Schmidt and Dr. Michael Benjamin under project name ALPACA1 [9]. Results from field tests are presented in order to validate the algorithm.

---

1The ALPACA technology is owned by MIT. A patent has been filed under United States of America Serial No. 13/536037, "System And Method For Collision Avoidance In Underwater Vehicles" by Michael Richard Benjamin and Henrik Schmidt. Filed June 28, 2012.
2 Background

2.1 Motion Control, Sensing and Communication

In the case of USVs, many technical similarities may be drawn with UGVs with regards to the number of degrees of freedom and operation in the presence of ambient traffic. The problem of motion control, however, becomes more challenging in USVs due to the harsh environmental disturbances [5]. This same comparison holds true for UUV dynamics, with the only exception that depth is added to the system's controllable degrees of freedom.

Despite the similarities between the kinematic models of USVs and UUVs, major differences exist between these vehicle types in the fields of sensing and communication. Whereas electromagnetic-domain devices such as laser scanners and cameras are useful for obstacle detection and mapping in surface vehicles [1], they are generally impractical for obstacle detection in underwater applications due to their limited range. Similarly, USVs rely on satellite systems for position tracking and communication, allowing for occasional operator guidance and remote control, as is the case for the Israeli Protector project [5]. UUVs may only use satellite systems when surfacing, but are otherwise limited to self-contained navigation systems for position tracking, and to low-bandwidth, low-frequency acoustic communication.

Taking a more general perspective to sensing in collision avoidance systems, the equipment that is currently used can be classified as either passive or active. The main difference between these two types is whether the sensor relies on an external source to serve its purpose, known as passive, or it provides its own source. Laser scanners, SOund Navigation And Ranging (SONAR) and Doppler radar systems are then examples of active systems, while cameras, microphones and other input-only devices are considered passive [11].
2.2 Autonomy and MOOS

The purpose for developing unmanned platforms is to have them perform complex tasks in situations where a human would be unable to perform well due to physical limitations, or due to the elevated risk of the scenario. Part of the challenge in this domain comes from understanding the system’s dynamics to achieve a particular desired outcome. Another element is developing an understanding of the environment through sensing. However, when communication is limited and the vehicle is unable to request guidance from the operator, as is the case in underwater applications, having the ability to make decisions about what the aforementioned desired outcome is becomes fundamental to the success of a mission.

Because autonomous vehicles have to handle a wide range of problems simultaneously in order to complete their missions, development in this domain often becomes hindered by the complexity of the software and controls systems as a particular project grows, it becomes increasingly difficult to adapt the associated platform. In an attempt to overcome this challenge, Paul Newman of Oxford University began working in 2001 on an innovative software package for mobile robot systems, named MOOS for “Mission Oriented Operating Suite” [8].

In order to simplify the development and contribution process, MOOS uses a centralized topology. At the core of every robotic system running MOOS lies a variable database, MOOSDB. Every other application in the system, called MOOSApp, may then subscribe to a particular variable in the database to receive reports whenever its value changes. In this way, some applications may connect to sensors in the system and update the values in the database with the most up-to-date information, while others link to the actuators to actually move the vehicle based on the desired heading and speed in the database. Other applications can simply monitor one variable and post another in response.
2.2.1 Interval Programming and the IvPHelm

Behavior-based controls systems have been used for many years to satisfy the decision-making requirements mentioned above. The origin of such systems is often attributed to Rodney Brooks [4], and one of their more important attributes is the ease of development of independent modules [2].

Given their modularity, MOOS implements behavior-based controls to choose the best action for the system. However, the decision making process itself has evolved since the “subsumption architecture” originally presented by Brooks. The Interval Programming (IvP) architecture presented by Dr. Michael Benjamin uses multiobjective optimization by having each behavior generate an objective function instead of a single desired outcome\(^1\) [3].

As an example, two behaviors `BHVAvoidCollision`, a collision avoidance protocol based on the Coast Guard Collision Regulations, and `BHVWaypoint`, a vehicle displacement behavior based on a given list of waypoints, can produce an interest value for each speed and heading pair. The `pHelmIvP` application, which embodies the architecture described by Dr. Benjamin, then considers the weighed sum of both objective functions to determine the most beneficial course of action for the overall mission. These results are then posted to the variable database such that the motor control process may drive the vehicle as instructed.

\(^1\)Additional information on the MOOS project and Interval Programming is available at www.themoos.org and www.moos-ivp.org
3 Mathematical Model

The model implemented in this thesis is derived from the concept envisioned for project ALPACA (Autonomous Littoral Passive Acoustic Collision Alarm), hereby explained. This concept is based on the cylindrical spreading loss associated with sound propagation in shallow water [9]. For the depth range of normal harbor or littoral operation, ignoring dissipation, the decay of sound intensity in decibel (dB) can be expressed as

\[ I_{dB} = I_0 - 10 \log_{10}(r) = I_0 - \frac{10 \log(r)}{\log(10)}, \quad (3.1) \]

where \( I_0 \) is a constant dependent on the properties and location of the acoustic source [7]. For a moving source, then, the rate of change of acoustic intensity may be expressed as:

\[ \Delta = \frac{dI_{dB}}{dt} = \frac{\partial I_{dB}}{\partial r} \frac{dr}{dt} = -v \cdot \frac{\partial I_{dB}}{\partial r} = \frac{10v}{\log(10)r}. \quad (3.2) \]

Rearranging Eq. 3.2 then gives an estimate of the time to collision \( dT \), in the form:

\[ dT = \frac{r}{v} = \frac{10}{\log(10)\Delta}. \quad (3.3) \]

In real applications, the time-local sound intensity \( \Delta \) may be estimated from a sound recording or acoustic input by computing the Root Mean Square (RMS) of the signal amplitude over a predefined amount of time \( \Delta T \).

The estimated time to collision will be sensitive to fluctuations in ambient noise as well as in the signal from the approaching acoustic source. While the estimation of intensity through RMS will reduce the apparent noise in the system, these variations will still be directly reflected in the estimated time to collision. Consequently, it is necessary to perform a statistical regression analysis to determine, with an adequate level of confidence, whether evasive action is necessary. In order to simplify the
analysis, the deviation may be observed in terms of the predicted collision time $T_n$ by adding the current time $T$ to the estimated time to collision $dT_n$ at some measurement $n$,

$$T_n = T + dT_n.$$ \hfill (3.4)

By creating a running average $\hat{T}_m$ of the predicted collision time over the last $N$ measurements, it is then possible to obtain a robust estimate of the collision time at measurement $m$. Should the standard deviation among said $N$ measurements be small enough, the system would then be allowed to make a decision regarding evasive action based on the time to collision.

$$\sigma_{T_m} = \sqrt{\frac{1}{N-1} \sum_{n=m-N+1}^{m} (T_n - \hat{T}_m)^2}.$$ \hfill (3.5)

In addition, the model is expected to provide a conservative estimate of the time to collision, since sound dissipation would effectively increase the range derivative of sound intensity [9].
4 Experimental Configuration

The threat detection algorithm studied in this thesis is aimed at Unmanned Underwater Vehicles. By integrating a hydrophone to a UUV and processing its input appropriately, the vehicle's autonomy platform, such as MOOS (Sec. 2.2), would be capable of triggering some collision avoidance behavior. For practical reasons, however, the UUV was replaced by a single hydrophone in the experimental setup, and the data was processed at shore instead of aboard the UUV.

Where the hydrophone line imposed a range limitation, and where the Charles basin serves as an adequate testing environment for shallow water tests, all experiments were conducted from the MIT Sailing Pavilion. As such, the hydrophone was installed approximately 35m away from the dock, at a depth of approximately 2m. Its input was recorded at a frequency of 44100Hz.

Another fundamental element in this experiment is the acoustic source. While a number of larger vessels do transit the Charles Basin on a regular basis during the late Spring and Summer, a smaller motor boat was considered safest given the proximity of the hydrophone to the dock. One such vehicle was borrowed from the MIT Sailing Pavilion to serve this purpose.

The components above would be considered sufficient in a practical implementation of the algorithm, with the exception of processing data aboard the unmanned vehicle. However, some additional elements were necessary in the experimental setup to provide the data required for a detailed evaluation of the algorithm's performance; given the nature and purpose of this detection model, the most important metric is the comparison of the estimated values with ground truth.

4.1 Ground truth of acoustic source

While a number of sensors and methods were given consideration in the initial development of a system to acquire the boat's true speed on the approach, such as laser
guns or indirect measurements of the speed by using a chronometer and a predefined path of known length, the resolution of these was ultimately regarded as insufficient for appropriate validation of the algorithm.

In place of the solutions presented above, a GPS sensor was selected to provide the actual boat location throughout the tests. In particular, the GARMIN GPS 18x 5Hz was deemed capable of meeting the needs of this project, and was made available through MIT's Laboratory for Autonomous Marine Sensing Systems.

4.1.1 GPS data logging and time matching

Given that the test data provided in this setup is divided in two main blocks – the acoustic data at shore and the boat true position aboard the vehicle itself – real-time processing was not an option during the experiments. Instead, it was necessary to store the GPS data for later analysis. For this purpose, a virtual vehicle was created under the MOOS architecture, which already allows for generation of detailed logs of unmanned vehicle state variables, such as operating state and position, within its pLogger process.

An added benefit of using a virtual MOOS vehicle to store the GPS data logs was that, by fitting the on-board computer with a long-range wireless network antenna, the true position of the boat could be monitored from shore during the tests. This made it possible to provide initial characterization of the results by concurrently monitoring the position of the boat and the audio recording.

While these initial observations helped understand how the two datasets should fit together for processing, an additional component was necessary to ensure that they were correctly synchronized. In order to achieve this, the device would have to create a signature on both datasets. For the specifics of this setup, that meant the device would have to communicate directly with the virtual vehicle to add an entry to the log, and produce some recognizable sound to be received through the hydrophone.

A simple Arduino-based switch sensor was chosen for this purpose, given the flexibility of the microcontroller for serial communication applications. By adding a pull-down resistor to the chosen input pin, and using two pieces of metal connected to
the input pin and to 5V, the microcontroller was then programmed to send a message through the serial port whenever the pieces of metal were struck together, drawing the input pin to 5V and also creating a particular sound. A process was added to the virtual vehicle to create an entry in the log whenever the microcontroller sent the desired message through the serial port.

Fig. 4-1 illustrates the basic experimental setup used for this project, including all elements discussed above.

4.2 Data Processing

The final element of the experimental setup was a platform to facilitate the data analysis. Through preliminary measurements with the hydrophone and an initial round of experiments to become familiarized with the system, it was possible to develop a series of scripts to extract the data from the vehicle logs and then present the results in a readable format. In its current form, the platform requires that the user identifies the start and end times of each segment of interest, namely the instances where the boat was performing an approach and departure maneuver.

The working principle of the audio processing script is drawn from the difference
in sampling frequency between the datasets and represents the intended approach to on-board processing. Where the audio signal is recorded at 44100Hz, the GPS data is stored at about 5Hz. In a real implementation of this algorithm, the application that records the input from the hydrophone would subscribe to the GPS reports, which would then give the application a working frame to estimate sound intensity by computing the Root Mean Square (RMS) of all points in the audio track between the previous position report and the latest one. Besides providing a smoother estimate of the sound intensity, the RMS would enable a real vehicle to track only one or two such segments of audio at a high sampling rate, resulting in reduced memory usage.

The script, then, computes the sound intensity data by identifying the segment of the hydrophone recording that belongs to each step in the vehicle position log. These values are then handed to the algorithm to generate an estimate of the time to collision and appropriate confidence metric as explained in Sec. 3.
5 Results and Discussion

After all data was recorded and the offset was identified, the vehicle log and audio recording were processed together. In order to identify the subsets of interest to this study, the vehicle position was plotted in a 3-dimensional graph with time on the vertical axis, as illustrated in Fig. 5-1.

![3D graph of boat position versus time with switch sensor triggers marked.](image)

Figure 5-1: Boat position versus time, with switch sensor triggers marked.

This representation allowed to break the data into segments, corresponding to the different approach-and-departure pairs of the boat from the perspective of the hydrophone. Each path segment was then plotted individually, to provide a visual aid for the analysis of its corresponding data (Fig. 5-2). Similarly, the distance to the hydrophone, or range, was plotted against the corresponding subset of the audio track and its RMS intensity estimation (Fig. 5-3). This representation shows that
the peak in sound intensity occurs after the boat has already passed the hydrophone. However, the spread of clear growth in sound intensity over approximately 30s before the range minimum already suggests this acoustic detection system may be practical in real applications.

Figure 5-2: Boat position through pass 7, April 26.

Figure 5-3: Boat distance to hydrophone from GPS data through pass 7, April 26.

In order to appropriately validate the model, however, it was necessary to confirm whether the algorithm is able to approximate the time to collision provided by the
on-board GPS system. After performing the statistical regression analysis described in Sec. 3, the output values were multiplied by a calibration factor. The resulting estimated time to collision and sound intensity were plotted against the confidence metric given by Eq. 3.5, as shown in Fig. 5-4. It may be observed that the time to collision predicted by the acoustic system in this particular case is indeed a generally conservative estimate of the value given by the GPS tracker.

![Graphs showing time to collision, sound intensity, boat speed, and standard deviation](image)

Figure 5-4: Time to collision estimate, sound intensity, boat speed and standard deviation of collision time through pass 7, April 26.

Fig. 5-5 show an instance where the predicted time to collision displays more variation and a less conservative estimate. However, the estimates presented in this case are still considered valuable for a vehicle’s behavior decision protocol, in particular for high risk-averse conditions. Here, the standard deviation drops below 5 seconds.
in multiple occasions during the first 60 seconds of the approach, giving a hypothetical UUV sufficient time to take evasive action such as sinking to the bottom and anchoring, or initiating travel to a known safe location.

![Time to collision (ETA), Audio segment (RMS), Boat speed from GPS, Standard deviation for projected collision time](image)

Figure 5-5: Time to collision estimate, sound intensity, boat speed and standard deviation of collision time through pass 5, April 26.

### 5.1 Confidence metric for estimation

As is explained in the ALPACA project white paper, the vehicle would only initiate an alarm or reactive behavior if the confidence level is high enough (small standard deviation) and the time to collision is smaller than a predefined value. Similarly, the alarm or behavior would only be cleared once the time estimate is sufficiently large in
the negative domain and the standard deviation is also large enough. In the case of Fig. 5-6, the estimated time to collision from $t = 814s$ to $t = 822s$ does not match the value provided by the GPS data. It must also be noted that the particular conditions for clearing a reactive behavior could be met by this case, depending on the preset values of trigger time $\Delta T$ and standard deviation $\hat{\sigma}$.

Figure 5-6: Time to collision estimate, sound intensity, boat speed and standard deviation of collision time through pass 3, April 26.

As such, it is strongly recommended that the system tracks not only the active values $\hat{T}_m$ and $\sigma_{T_m}$, but also the last instance where trigger conditions were met and its corresponding timestamp. Should this be the case, the scenario presented in Fig. 5-6 would be able to use this detection algorithm by holding the estimate given around $t = 808s$, since the standard deviation in the timeframe $t = 814s$ to $t = 822s$ is large.
and a sufficiently recent, high-confidence prediction exists.

5.2 Effect of accelerating acoustic source

As with Fig 5-2, passes 3, 5 and 7 of the April 26 experiment, discussed above, shared a common path feature: they started near the northern side of the Harvard Bridge and ended before reaching the Charles River Yacht Club, heading east. This feature was also present in other data sets that support the validity of the detection algorithm.

In order to acquire additional test data for validation, the reverse segments were also evaluated. The consequence of this path choice, however, was that the approach leg was significantly shorter than in the previous cases. As an example, Fig. 5-8 shows the boat as it turns around from its previous pass and into the next one. The corresponding estimate of time to collision for this case meets the required conditions to enable evasive action less than 10 seconds before the boat passes over the range minimum.

Two different arguments arise from these results. On the one hand, the short amount of time available after detection makes any evasive action impractical in this scenario. On the other hand, the self-correcting nature of the statistical regression analysis allowed the system to produce a good estimate of the time to collision, with sufficient confidence to meet the established trigger conditions.

In addition to the arguments presented above, it is possible that other variations
in the system affected the predictions provided by the detection algorithm. One such variation would be the boat speed as it moved towards the hydrophone. It is possible that due to the short length of the westward approach legs, the initial estimate of time to collision in these tests was affected by the boat picking up speed and perhaps even transitioning into planing. For this reason, a second look is taken to the time estimations discussed above, with respect to the speed logged via the GPS system.

At first glance, the vehicle speed in the first two cases discussed (Fig. 5-4 and 5-5) appears to be approximately constant. By comparison, Fig. 5-8 displays pronounced variations in speed around the same times where the confidence metric (the standard deviation of the projected collision time) increases significantly. In comparing these
observations with the seemingly unexpected behavior of the algorithm in pass 3 of the April 26 tests (Fig.5-6), it becomes apparent that the section where the model failed to approximate the GPS data coincides with a notorious increase in speed.

Accounting for the effects of an accelerating source in the perception of sound intensity related to the detection system is a non-trivial problem. However, the scenario envisioned for this algorithm is one where acoustic sources, such as barges or other motor boats, are traveling at approximately constant speed in a waterway, near a harbor or in littoral areas.

5.3 Estimation for departing acoustic source

A final observation is made, regarding the departure leg of all measurements presented in this document. From the graphs above, it may be drawn that the estimation of collision time as the boat moves away does not fit well the data obtained from the GPS. It may be possible that the faster drop in sound intensity is related with the motion of water in the wake. Despite this inaccuracy of the estimate, the algorithm appears to hold a negative value for the time to collision, with high confidence in the estimation. As such, it would be possible to modify the behavior or alarm release conditions to observe the amount of time this estimate has held a negative value before resetting.
6 Conclusions

Where the envisioned usage of the detection algorithm hereby evaluated involves vehicles traveling at approximately constant speed, the results provide proof-of-concept validation of the model. It has been demonstrated that detection of other vehicles through acoustics is possible, with sufficient confidence to command a collision avoidance behavior.

It is strongly recommended, however, that additional testing be performed to appropriately characterize the departure leg of a boat’s path, as well as the transition from approach to departure.

Given the possible effects of the boat’s wake on acoustic perception, it is also recommended that tests be performed in parallel lines, moving incrementally away from the detection system, be it a hydrophone setup or an actual UUV.

Additionally, as reaction time may be a concern in different applications, it is recommended that tests be performed in locations with greater water depth, and with a range of boat types as acoustic sources, to better understand the effects of these parameters in the algorithm and its calibration.
Bibliography


