

Automatic Classification of Cetacean Vocalizations Using an Aural Classifier

Paul C. Hines and Carolyn M. Binder
Defence R&D Canada–Atlantic
PO Box 1012 Dartmouth
NS, Canada, B2Y3Z7
phone: (902) 426-3100 fax: (902) 426-9654
email: paul.hines@drdc-rddc.gc.ca, phines50@gmail.com

Award Number: N000141210139
<http://www.drdc-rddc.gc.ca/>

LONG-TERM GOALS

To develop a robust automatic classifier with a high probability of detection and a low false alarm rate that can classify vocalizations from a variety of cetacean species.

OBJECTIVES

In this research, we wish to apply a unique automatic classifier developed by the PI that uses perceptual signal features – features similar to those employed by the human auditory system – to classify cetacean species vocalizations and reject anthropogenic false alarms. This *aural* classifier has been successfully used to distinguish active-sonar echoes from man-made (i.e. metallic) structures and naturally occurring clutter sources [1, 2] and performs as well or better than expert sonar operators [3]. Many of the features were inspired by research directed at discriminating the timbre of different musical instruments – a passive classification problem – which suggests the method should be able to classify marine mammal vocalizations since these calls possess many of the acoustic attributes of music.

APPROACH

The research is part of a PhD program undertaken by Ms. Carolyn Binder under the supervision of Dr. Paul C. Hines. The postgraduate program is being conducted in the Oceanography department at Dalhousie University where Dr. Hines is an adjunct professor and at Defence R&D Canada–Atlantic where Dr. Hines is Principal Scientist/Underwater Sensing and Ms. Binder is a Research Assistant. In this project we examine anthropogenic transients and vocalizations primarily from four¹ cetacean species – the sperm whale, northern right whale, the bowhead whale and the humpback whale. These species were chosen for the following reasons:

¹ Vocalization data from other cetacean species will be tested with the classifier as time permits. For example, Minke whale vocalizations, available on the Mobysound website, were the focal topic for the 5th International Workshop on Detection, Classification, Localization, and Density Estimation of Marine Mammals using Passive Acoustics [4]. Data such as these provide comparative a performance measures against other classifiers and tests the versatility of the classifier.

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE 30 SEP 2013	2. REPORT TYPE	3. DATES COVERED 00-00-2013 to 00-00-2013			
4. TITLE AND SUBTITLE Automatic Classification of Cetacean Vocalizations Using an Aural Classifier		5a. CONTRACT NUMBER			
		5b. GRANT NUMBER			
		5c. PROGRAM ELEMENT NUMBER			
6. AUTHOR(S)		5d. PROJECT NUMBER			
		5e. TASK NUMBER			
		5f. WORK UNIT NUMBER			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Defence R&D Canada -Atlantic,PO Box 1012,Dartmouth, NS, Canada B27 3Z7,		8. PERFORMING ORGANIZATION REPORT NUMBER			
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)			
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)			
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified	Same as Report (SAR)	10	

- all are present in US and Canadian waters;
- sperm whale clicks are often confused with false alarms from impulsive anthropogenic;
- transients and hydrophone self-noise (RF crackle, sensor knocks and bumps);
- the northern right whale is critically endangered (estimates of a few hundred remaining);
- the bowhead and the humpback have proven particularly difficult to discriminate automatically because the duration and bandwidth of vocalizations from the two species are similar.

The marine mammal vocalizations being used throughout the project have been obtained from DRDC's data archive [5] and the Mobysound website [6].

In addition to classifying archived calls, the robustness of the aural classifier will be quantified. That is, we shall address the question: "Will it work on vocalization data from these species collected under different environmental conditions?" To examine this, discriminant analysis (DA) [7] was used to rank the aural features in terms of their ability to separate the vocalizations between species [8]. Then, the more highly ranked features will be tested for robustness. The testing will be done by using the calls from [5] and [6] along with synthetically generated calls as source signals in propagation experiments conducted on board CFAV QUEST. (Note that the experiments are facilitated through *in kind* contribution from DRDC.) The measurements will be complemented by modeling the experiment using OASP—the pulse propagation component of the OASES propagation model [9].

WORK COMPLETED (FY2013)

The focus of the effort during FY 2013 has been two-fold:

1. Identify other cetacean vocalizations with which to extend the test cases of the classifier; this includes data from other species to examine the versatility of the aural classifier and additional data from the current selection of species to allow for further testing.
2. Preparing for and conducting the propagation experiments to test the robustness of the aural classification feature set.

Extending the test cases of the aural classifier: The 6th International Workshop on Detection, Classification, Localization, and Density Estimation of Marine Mammals using Passive Acoustics (DCLDE workshop) provided additional North Atlantic right whale data with which to train and test the classifier. These data included logged North Atlantic right whale upswEEP and gunshot calls, as well as a noise only dataset in which there was a high degree of certainty that no right whales were present. Only the gunshot and noise data have been analyzed. The results of this analysis conducted during FY 2013 are contained in the Results section.

Aural feature robustness: Results thus far have shown that the aural classifier can successfully discriminate between several species of cetacean vocalizations. This part of the research aims to examine the robustness of the classifier under various environmental conditions. The significance of this work goes beyond validating the aural classifier algorithm and extends to automatic recognition research in general; properties of the ocean environment such as sound speed profile, bathymetry, and boundary properties determine how acoustic signals will change as they propagate through the ocean.

Clearly, this affects all automatic recognition techniques since any successful system must be robust enough to operate effectively in diverse environmental conditions.

Surprisingly, there is scant research published in the field of acoustic propagation applied to passive acoustic monitoring (PAM) of marine mammals. Helble *et al.* [10] recently demonstrated the impact of the ocean’s environmental properties on PAM for the *detection* problem. They showed that the probability of detecting a humpback whale is a function of environmental properties and location. In the current project, we examine the impact of the environment on the *classification* problem. In an effort to determine the environmental impacts on the aural classifier, two experiments were designed and conducted to evaluate the robustness of the perceptual features to propagation effects.

Two field trials conducted on board CFAV QUEST provided an opportunity to collect data for testing the robustness of the aural features with respect to underwater sound propagation. The trials occurred in the spring of 2012 on the Scotian shelf and in the spring of 2013 in the Gulf of Mexico. To investigate the impacts of propagation on aural classification, a set of bowhead and humpback whale vocalizations were transmitted from QUEST to a set of moored hydrophones. To help model the effect of propagation, synthetic bowhead and humpback vocalizations were also transmitted. The synthetic signals were designed to have similar mean and variance values to the cetacean calls for three of the aural features found to be important to bowhead/humpback discrimination. The experiments are presented in the Results section.

RESULTS

Extending the test cases of the aural classifier: The first step in processing the workshop data for the 6th International DCLDE was to apply an automatic band-limited energy detector to isolate detections. There were 972 contacts from the right whale workshop data that overlap in time with the calls in the logs and so were considered truth detections. The 3636 contacts identified by the detector that didn’t overlap were assumed to be false alarms. All 465 contacts from the workshop noise data were assumed to be false alarms. The workshop data were augmented with the 86 right whale gunshot calls from the DRDC data archive referred to earlier. The results are listed in Table 1.

First, the classifier was trained using the 972 true gunshot detections from the workshop dataset and the 465 false alarms from the noise dataset. The results of training the classifier are shown in the left panel of Figure 1. The line dividing the blue from the red background in the figure represents the decision threshold; for example, any vocalizations that are mapped onto the discriminant function to the left of the line will be classified as false detections and any that are

Table 1 North Atlantic right whale gunshot-call detections and false alarms used for aural classification.

Dataset	True detections	Logged calls	False detections
<i>Gunshots (workshop)</i>	972	1042	3636
<i>Noise data (workshop)</i>	0	0	465
<i>Gunshots (Fundy)</i>	86	N/A	0

mapped to the right side of the line will be classified as gunshot calls. The classification accuracy for the training set is 95%, and the area under the rock curve (*not shown*) is 0.99 indicating a successfully

trained classifier. Next, the classifier was tested using the 86 Bay of Fundy gunshots calls and correctly identified all 86 calls (right panel of Figure 1). This is considered a reasonably challenging case since the Fundy data were collected in a different year and location, under different environmental conditions, using entirely different monitoring equipment.

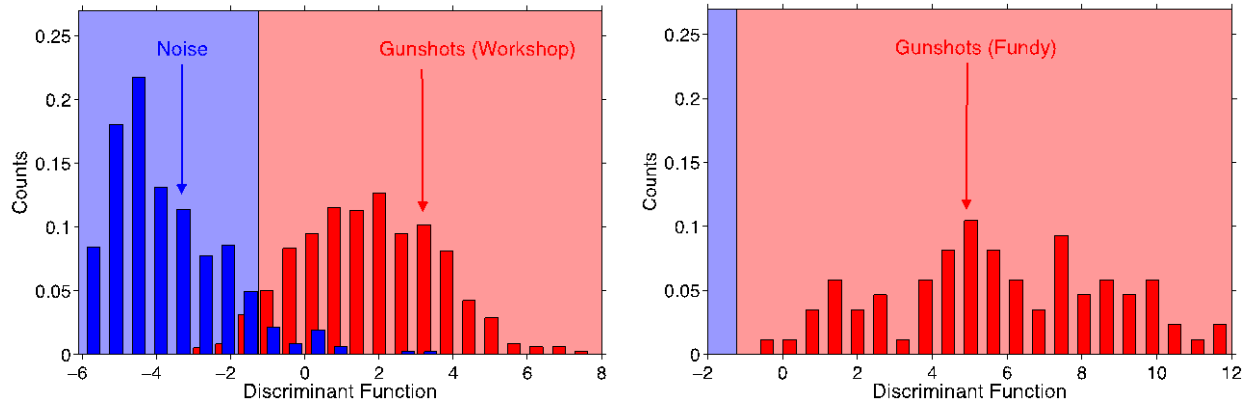


Figure 1 Left panel shows the results of training the classifier with gunshots from the workshop dataset and detections from the noise dataset. Right panel shows the results of testing the classifier model shown in the left panel on gunshots from the Bay of Fundy dataset. All vocalizations were correctly identified.

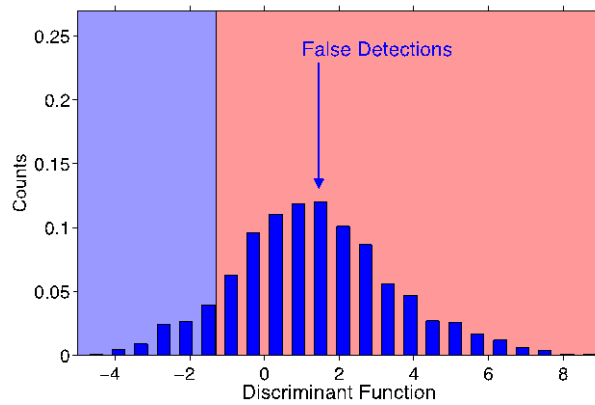


Figure 2 Testing the aural classifier with the 3636 false alarms encountered in the DCLDE workshop data set.

Then the trained classifier was tested using the 3636 false alarms that DRDC’s automatic band-limited energy detector identified in the right whale workshop data. Unfortunately approximately 70% of the false alarms were classified as right whale detections (Figure 2). This was not anticipated given the previous success of the aural classifier. (It should be noted that it was generally accepted at the DCLDE workshop that the data set represented an extremely challenging classification problem.) Two hypotheses are offered for the classifier’s poor performance: The first is that many of the false alarms are not “noise” in the conventional sense that the classifier was trained for. That is to say, the false alarms in the right whale dataset include many calls from other cetacean species such as

humpbacks for example. These sounds share greater similarity with right whale vocalizations than with noise and so are being misclassified as right whales; the second is that the noise detections are so aurally-diverse in nature that they cannot be accurately represented as a single class. To address either of these hypotheses one would need to sub-divide the noise class into “other marine mammal calls” and “noise”, re-train the classifier and perform a three-class classification as has been done to discriminate multiple classes of mysticetes in the past with the aural classifier. At present, this work is not scheduled for the final year of the grant but could be addressed in future should there be sufficient interest from the program manager.

There are currently two further data sets identified that could be tested with the aural classifier. The first is a collection of Pacific humpback vocalizations [11] from the endangered Oceania subpopulation. The PI has contacted the two lead authors of Ref. 11 and both expressed interest in providing the data for testing with the aural classifier. The second data set is a collection of South Pacific blue whale and fin whale vocalizations collected by the Australian Navy. Following a marine mammal workshop sponsored by The Technical Cooperation Program (TTCP) Panel 9 on ASW (Canberra, Nov. 2013), the Maritime Environmental Compliance branch of the RAN expressed considerable interest in providing these data. These data would extend the versatility of the multiclass classification capability of the algorithm. To date neither dataset has been obtained due to insufficient personnel resources; however, they remain available should time and resources permit.

Aural feature robustness: Two field trials were conducted on board CFAV QUEST to collect data for testing the robustness of the aural features with respect to underwater sound propagation. The trials occurred in the spring of 2012 on the Scotian shelf and in the spring of 2013 in the Gulf of Mexico. A set of pre-recorded bowhead and humpback whale vocalizations and a set of synthetic bowhead and humpback vocalizations were transmitted from QUEST to a set of moored hydrophones. The synthetic signals were designed to have similar mean and variance values to the cetacean calls for three of the aural features found to be important to bowhead/humpback discrimination. Environmental measurements including sound speed profiles, bottom properties, and wind speed were measured at the sites to support the modeling effort. A sample geometry is shown in Figure 3.

The signals (155 of each type) were transmitted from a projector deployed from the quarterdeck of QUEST, as the ship drifted, and received on moored recorders 0.5-20 km away from the ship. For each field trial the experiment was repeated several times, on different days and at different locations, so as to capture various propagation conditions. High SNR vocalizations were selected with the assumption that high SNR indicates the vocalizing whale was relatively close to the recording equipment so that these vocalizations were less affected by propagation. From the spectrograms of the selected vocalizations, it was determined that the frequency ranges of the vocalizations are 50–800 Hz for bowheads and 100–2000 Hz for humpbacks. Since no projector was available to transmit the signals over the approximately 5-octave band of the vocalizations, the signals were filtered and scaled to take advantage of the two-octave passband (1–4 kHz) of an ITC-2010 projector. The RMS averaged power spectra of the signals after a 200–800 bandpass filter was applied are compared to the full bandwidth signals in Figure 4. The reduced frequency band contained 74% of the energy in the bowhead vocalizations and 72% of the energy for the humpback vocalizations.

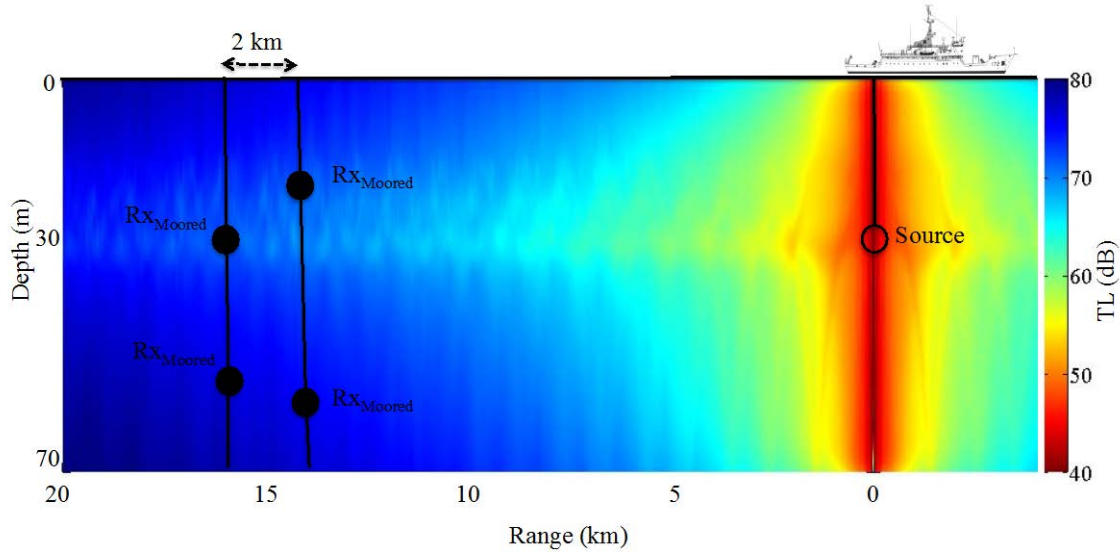


Figure 3. Experimental setup for PAM propagation experiment. Rx_{Moored} refers to the moored hydrophones. Measurements were made at ranges from 2 to 20 km by steaming to a set range, deploying a projector from QUEST and transmitting the vocalizations. Colour background represents transmission loss in dB using Bellhop propagation model.

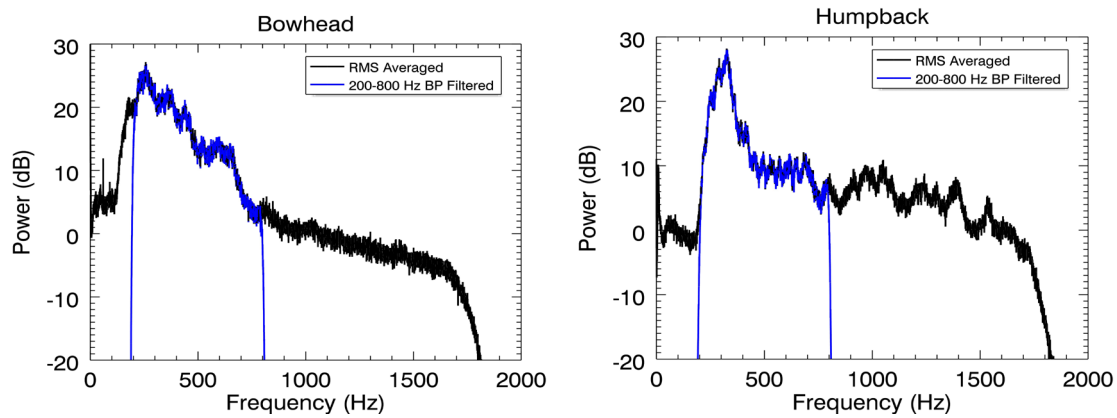


Figure 4 RMS averaged power spectra for the selected bowhead (left panel) and humpback (right panel) vocalizations. The black line represents the full-band spectra of the vocalizations, and the blue line represents a 2-octave band that contains a significant proportion of the energy in the vocalizations.

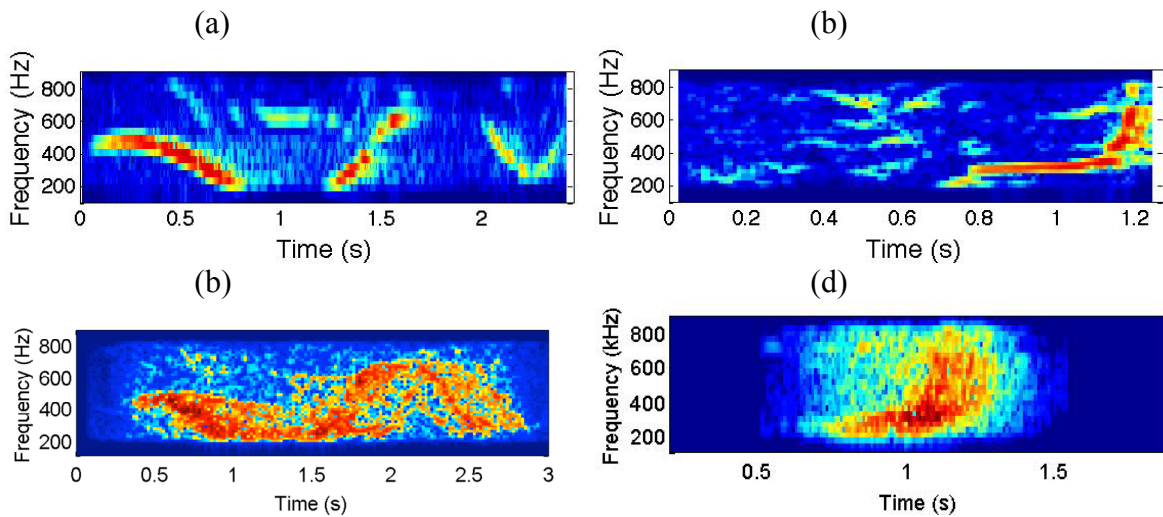


Figure 5 Spectrograms of example band-pass filtered signals used for propagation experiments. These are (a) real bowhead, (b) real humpback, (c) synthetic bowhead, and (d) synthetic humpback vocalizations.

To ensure that the reduced frequency band contained sufficient information to discriminate between the species, aural classification was performed on both the full-bandwidth and reduced-bandwidth signals. Classification accuracy reduced slightly from 94% for the full-bandwidth signals to 92% for the reduced-bandwidth signals; the *area under the ROC curve (AUC)* also reduced from 0.99 to 0.98. Many of the same perceptual features were highly ranked discriminators for both signal types; this is important since it suggests that applying the 200–800 Hz bandpass filter does not remove the information required to calculate the important perceptual features. Altogether, this evidence suggests that sufficient information is contained in the reduced-bandwidth signals to accurately represent both species’ vocalizations. The final step in processing the signals for transmission was to increase the playback speed of each filtered signal by a factor of five to shift the signals into the passband of the ITC-2010 source.

All the bowhead and humpback vocalizations available include some propagation effects. To minimize the impact of propagation effects already embedded in the signals we chose high SNR vocalizations for the experiment; that is to say, louder vocalizations were presumed to come from animals that were nearest in proximity. However, to gain additional insight into environmental impacts, synthetic vocalizations were also used. In contrast to real vocalizations, synthetic vocalizations provided a known signal with no embedded propagation effects. The synthetic signals were designed using empirical orthogonal functions (EOFs) to have similar mean and variance values for the perceptual features that were considered important in discriminating bowhead and humpback vocalizations. Figure 5 shows example sonograms of both the real and synthetic bowhead and humpback vocalizations.

Preliminary analysis of the data from the two field trials indicate that sufficient data were obtained to examine the effects of propagation on the perceptual features used by the aural classifier. This is currently being undertaken and includes examining changes to the general aural classification results, as well as examining changes to individual perceptual features to identify those features that may be

robust to propagation effects. This will be the primary focus of this year's research and will form the basis of Ms. Binder's PhD dissertation.

Finally, it should be mentioned that during the first year of the project [2012], the primary objective was to quantify the ability of the aural classifier to discriminate four cetacean species from one another and from anthropogenic transients. This was reported on in last year's annual report but has culminated with a paper summarizing those results being submitted to a peer reviewed journal [8] during this reporting period.

IMPACT/APPLICATIONS

Detection and classification of cetaceans has become critically important to the US Navy due to an ever increasing requirement for environmental stewardship. Passive acoustics continues to be the best method to carry out this task but current techniques provide only a partial solution; most detectors are either too specialized (i.e., species-specific) leading to many missed detections, or are too general, leading to unacceptably high false alarm rates. Furthermore, future military platforms will have to support smaller complements and deal with ever-increasing data throughput, so that automation of on-board systems is essential. In addition, the technique is well suited to autonomous systems since a much smaller bandwidth is needed to transmit a classification result than to transmit raw acoustic data. The success of the machine classifier in discriminating cetacean vocalizations suggests that it could be applied to other passive acoustic classification problems which currently employ human audition. This would be particularly useful if expert listeners aren't available –such as diagnosing heart murmurs in remote communities that lack a cardiologist, or as part of the triage process in a hospital emergency department. Alternatively, the machine classifier is ideally suited when the sheer volume of data makes human audition untenable – such as classifying ocean acoustic data for species population monitoring. Finally, testing the classifier on passive marine mammal vocalizations is also a first step to testing the algorithm on passive transients generated by submarines to examine its potential for passive detection and classification of submarines.

RELATED PROJECTS

This research will benefit from DRDC Atlantic's SUBTRACTION Applied Research Project in which DRDC's aural classification algorithms (including the marine mammal classification algorithm) is being integrated into DRDC's System Test Bed (STB). The STB is used to evaluate sonar algorithms in a military context. Some of the insights to be gained will be: whether the aural classifier can reduce false alarms from marine mammals; does the classifier reduce operator workload required by environmental considerations (the so-called green navy) to enable greater concentration on potential targets; is the aural classifier easily integrated into a navy platform?

This research also benefits substantially from a recently completed project at DRDC [5] during which anthropogenic transients and cetacean vocalization data were compiled, extracted into *.wav* files, and manually classified with assistance from expert listeners.

REFERENCES

- [1] Victor W. Young and Paul C. Hines, "Perception-based automatic classification of impulsive-source active sonar echoes," *J. Acoust. Soc. Am.*, 122 (3) pp. 1502-1517, (2007).

- [2] Paul C. Hines, Victor W. Young, and Jeff Scrutton, "Aural classification of coherent-source active sonar echoes," *Proceedings: International Symposium on Underwater Reverberation and Clutter*, Lerici, Italy, eds. P.L. Nielsen, C.H. Harrison, and J.-C. LeGac, pp. 293-301, (2008).
- [3] Nancy Allen, Paul C. Hines, and Victor W. Young "Performances of human listeners and an automatic aural classifier in discriminating between sonar target echoes and clutter," *J. Acoust. Soc. Am.*, 130 (3), pp. 1287-1298, 2011.
- [4] Carolyn M. Binder, Paul C. Hines, and Joey Hood, "Passive Classification of Marine Mammal Vocalizations Using an Automatic Aural Classifier," 5th International Conference on Detection, Classification, Localization, and Density Estimation of Marine Mammals Using Passive Acoustics, Mt. Hood, Oregon, August, 2011.
- [5] Joe Hood and Derek Burnett, "Compilation of Marine Mammal Passive Transients for Aural Classification," DRDC Atlantic CR 2008-287, 93 pages, (2009).
- [6] S. Heimlich, D. Mellinger, H. Klinck, "MobySound.org," <http://www.mobysound.org/index.html> (2007).
- [7] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification*, (2nd ed, Wiley-Interscience, 2001).
- [8] Carolyn M. Binder and Paul C. Hines, (2013), Automated aural classification used for inter-species discrimination of cetaceans, *J. Acoust. Soc. Am.* Submitted 4 Sep 2013, 13 pages. Submission number: 13-13589.
- [9] H. Schmidt, OASES 2.2, Massachusetts Inst. Technol., Cambridge, MA, 1999 [Online]. Available: <http://acoustics.mit.edu/faculty/henrik/oases/oases.html>
- [10] T.A. Helble, G.L. D'Spain, J.A. Hildebrand, G.S. Campbell, R.L. Campbell, and K.D. Heaney, "Site specific probability of passive acoustic detection of humpback whale calls from single fixed hydrophones," *J. Acoust. Soc. Am.*, 134 (3), 2556–2570 (2013).
- [11] L.M. Munger, M.O. Lammers, P. Fisher-Pool, and K. Wong, Humpback whale (*Megaptera novaeangliae*) song occurrence at American Samoa in long-term passive acoustic recordings, 2008–2009, *J. Acoust. Soc. Am.*, 132 (4), pp. 2265-2272, 2012.

PUBLICATIONS

- [1] Binder, C.M. and Hines, P.C. (2013), Automated aural classification used for inter-species discrimination of cetaceans, *Jour. Acoust. Soc. Am.* Submitted 4 Sep 2013, 13 pages. Submission number: 13-13589 [refereed].
- [2] Binder, C.M., Hines, P., Pecknold, S.P., and Scrutton, J. (2013) Robustness of features used for automatic aural classification to propagation effects. *Proceedings of Meetings on Acoustics: International Congress on Acoustics*, Montreal, QC.
- [3] Binder, C.M. and Hines, P.C. (2013) Passive acoustic monitoring of North Atlantic right whales using computer-based aural classification. 6th International Workshop on Detection, Classification, Localization and Density Estimation of Marine Mammals using Passive Acoustics, St. Andrews, Scotland.

HONORS/AWARDS/PRIZES

Carolyn Binder (Defence R&D Canada – Atlantic / Dalhousie University), Fessenden Prize in Underwater Acoustics, sponsored by the Canadian Acoustics Association.

Carolyn Binder (Defence R&D Canada – Atlantic / Dalhousie University), Kathy Ellis Memorial Book Prize, sponsored by Dalhousie Dept. of Oceanography.