AUTOMATIC CLASSIFICATION OF IMPULSIVE-SOURCE ACTIVE SONAR ECHOES USING PERCEPTUAL SIGNAL FEATURES FROM MUSICAL ACOUSTICS

Victor W. Young, Paul C. Hines, and Sean Pecknold
Defence R&D Canada – Atlantic, 9 Grove St., Box 1012, Dartmouth, NS, B2Y 3Z7 victor.young@drdc-rddc.gc.ca

1. INTRODUCTION

False alarm returns from naturally occurring objects in the environment often plague active sonar systems. These types of echoes are known collectively as clutter.

Typically a human operator is responsible for discriminating between echoes from true targets and echoes from clutter objects using visual displays, such as time-frequency sonograms. There have also been attempts to create automatic classifiers based upon purely statistical signal features. Both of these conventional techniques for active sonar classification ignore a potentially valuable tool: the human auditory system.

There is mounting experimental evidence to suggest that human listeners can aurally discriminate between target and clutter echoes[1]. This paper investigates the possibility of using cues known to be perceptually relevant in the human auditory system as signal features in an automatic classifier.

Drawing an analogy between active sonar echoes and percussive musical timbre, this paper examines signal features that have been identified as underlying the perception of timbre. These perceptual signal features are used to automatically classify impulsive-source active sonar echoes recorded on a towed-array.

The purpose of this paper is to demonstrate that active sonar echoes can be successfully classified using perceptual signal features, and not to suggest that this technique is superior to human-operator classification using visual displays or to automatic classification with statistical signal features. Indeed, the optimum classification technique probably involves a combination of all three approaches.

2. EXPERIMENTAL DATA

The experimental data consists of 98 target echoes from two different objects and 100 clutter echoes from 28 different objects. The two target objects are an oil rig and the tanker ship that attends it. The 28 clutter objects are naturally occurring seafloor structures.

The data were collected during a sea trial on the Malta Plateau using signals underwater sound (SUS) charges and a towed-array. A total of nine SUS charges were deployed during the experiment. Each SUS charge contained 0.82 kg of TNT and was set to detonate at a depth of 87.0 m. The towed-array consisted of 96 omni-directional elements sampled at a rate of 4096 Hz, and it was towed at a speed of 10 knots and a depth of 40 m. Average water depth in the area was about 100 m.

The towed-array data were beamformed to obtain a total of 81 horizontal beams. Each beam was spectrally whitened using a Butterworth filter and then normalized to eliminate reverberation using a two-pass mean technique[2]. Then an energy threshold was applied and samples that exceeded the threshold were taken to be detections. To eliminate signal-to-noise ratio (SNR) as a possible target-clutter discrimination cue, care was taken to balance the target and clutter SNR distributions so that the distribution of SNR values within the target class was very nearly identical to the distribution of SNR values within the clutter class.

3. PERCEPTUAL SIGNAL FEATURES

Musical timbre is defined as "that attribute of auditory sensation in terms of which a subject can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar"[3]. There have been many musical acoustics studies that investigate the signal features that underlie the perception of timbre[4]. This paper will apply several perceptual signal features identified in those studies to the problem of active sonar classification.

Perceptual signal features considered in this paper include: duration, sub-band attack and decay time, sub-band synchronicity, spectral character of the pre-attack noise, and the peak value, centroid, and roughness of the perceived loudness spectrum. Because these features are inherently perceptual, they must be measured using a model of the human auditory system. The auditory model employed in this paper consists of two main components: an auditory filter bank[5], which breaks the sonar echoes down into sub-bands, and a loudness model[6], which converts the output of the filter bank into a perceived loudness spectrum.

Time-frequency features – like sub-band attack and decay time and sub-band synchronicity – are measured at the output of the filter bank (*i.e.*, prior to applying the loudness model). Purely spectral features – like the peak value, centroid, and roughness of the loudness spectrum – are measured at the output of the loudness model. To eliminate total loudness (*i.e.*, the integral across the perceived loudness spectrum) as a target-clutter discrimination cue, each echo is scaled to have the same total loudness.

Many of the features considered in this paper are multi-valued in that they consist of multiple values for each echo. For example, consider the feature sub-band attack time: for a single echo there are sub-band attack time values for each channel of the filter bank, but the analysis which follows requires single-valued features. Conversion of a single multi-valued feature into multiple single-valued features is achieved using summary statistics. Sub-band attack time is thus converted into three single-valued features: minimum, mean, and maximum sub-band attack time. In this way a total of 58 single-valued features are constructed.

Each of the 58 perceptual features is normalized so that, over all 210 returns, the mean feature value is 0 and the standard deviation is 1. This normalization process ensures that each feature is weighted equally in the analysis that follows.

4. PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) is a statistical technique for projecting a multi-dimensional feature-space defined by M possibly correlated signal features down onto a multi-dimensional feature-space defined by N uncorrelated signal features, where $N \leq M$. PCA is applied by first obtaining the eigenvectors of the correlation matrix defined by the M original features. The eigenvectors are sorted according to the magnitudes of their corresponding eigenvalues and the first N eigenvectors are then used to define a transformation matrix, which projects points (*i.e.*, echoes) in the old M-dimensional feature-space down onto the new N-dimensional feature-space.

The following section presents N=2 PCA results for several different M values. The M features included in the PCA are those which, when considered in isolation, have the least overlap between the target and clutter classes.

5. AUTOMATIC CLASSIFICATION

The full data set (consisting of 210 echoes described by N=2 PCA features) is split into two sub-sets, which are used to train and test a Gaussian-based automatic classifier. During the training phase, separate target and clutter Gaussian probability density functions (PDF) are defined using sample mean and covariance matrices estimated from the echoes in the training sub-set. During the testing phase, each echo in the testing sub-set is classified as either target or clutter based upon its relative position on these two PDFs. The error rate (*i.e.*, the fraction of misclassified echoes in the testing sub-set) is then calculated and used as a metric to quantify the success of the classification process.

Automatic classification using this technique is carried out twice for the M=5 / N=2 PCA results: the first time training on echoes from the tanker ship and 14 of the 28 clutter objects then testing on echoes from the oil rig and the other 14 clutter objects (scheme A), and the second time reversing the training-testing echoes (scheme B). Results for training-testing scheme A are presented in the Figure 1. Automatic classifier results for other M values are presented in Table 1.

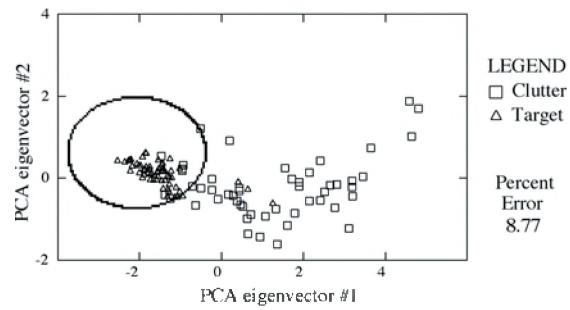

Fig. 1. M=5/N=2 automatic classifier results for scheme A; the ellipse represents the decision surface

Table 1. N=2 automatic classifier error rates

M	2	5	58
Scheme A	12.28%	8.77%	7.89%
Scheme B	9.52%	13.10%	14.29%

6. DISCUSSION

Results presented in this paper demonstrate that perceptual features can be used to successfully classify impulsive-source active sonar echoes: using perceptual features with a Gaussian classifier, error rates less than 10% can be achieved. However, care must be taken when selecting features for inclusion in the analysis since in many situations fewer features (*i.e.*, smaller M) yield better results. Moreover, A-B comparison suggests that the "best" choice of features for any classification task depends on the data used for training. Future plans include applying the perceptual feature automatic classifier to echoes from real submarines and to coherent-source active sonar data.

REFERENCES

1. N. Allen et al., *151st Meeting Acoust. Soc. Am.* (Providence, RI, 2006), pp. 3395.
2. W. A. Struzinski and E. D. Lowe, *J. Acoust. Soc. Am.* **76** (6), 1738 (1984).
3. W. A. Yost, *Fundamentals of Hearing*, 4th ed. (Elsevier, Toronto, 2000).
4. J. M. Grey and J. W. Gordon, *J. Acoust. Soc. Am.* **63** (5), 1493 (1978).
5. M. Slaney, *An Efficient Implementation of the Patterson-Holdsworth Auditory Filter Bank* (Apple Comp. Inc. TR35, 1993)
6. J. E. Appel, *Loudness Models for Rehabilitative Audiology* (Ph.D. dissertation, U. of Oldenburg, Germany, 2002).