Clean Data Training Approach to Active Sonar Classification

Alexei Kouzoubov (1), Binh Nguyen (1), Shane Wood (1) and Chris Gillard (1)

(1) Defence Science and Technology Organisation, Edinburgh, SA, Australia

ABSTRACT

An approach to forming a training data set for active sonar classification from the free-field noiseless echoes of underwater objects is considered. The approach has been tested on simulated data based on laboratory measurements. Both the test and the training data are generated from the clean data by applying two environment parameters: noise and a channel maltipath. It was investigated how well the parameters of the channel used for forming the training data should match those used for generating the test data. The dependence of classification accuracy on the pulse frequency bandwidth was also investigated and shown to increase with increasing bandwidth.. Some knowledge of the environment is required for the application of this approach to real sonar data.

INTRODUCTION

An ability to classify active sonar echoes is important in Anti-Submarine Warfare (ASW). The supervised classification techniques, which are used in this paper, require training data for each object class (Theodoridis & Koutroumbas, 2006). In the most common approach to active sonar classification, all sonar echoes are divided into two classes: those originating from the target of interest, such as a submarine, and all other, non-target, echoes. The training data for both classes should be representatively sampled from all possible echoes to provide good classification accuracy. In the underwater environment this task is complicated by the fact that the echoes from the objects are distorted by the propagation effects, such as multipath, reverberation etc. As a result of this, training classifiers on the echoes obtained from targets in the free-field, noiseless conditions does not necessarily lead to accurate classification of the echoes from the same targets because they are modified by the propagation in the underwater environment (Ainslie, 2010). Obtaining the free-field data for the targets of interest is not an easy task. It is much more difficult, however, to obtain echoes from these targets in all possible environments to form a universal training data set. It is therefore of interest to develop a classification approach based on the noiseless free-field target echoes, or in other words clean training data, to classify target echoes distorted by the underwater propagation.

In mathematical terms the problem can be formulated as follows. Assuming that we have clean target echoes $x_i(t)$, classify the test echo $x(t)$ distorted by the propagation, so the recorded echo available to the classifier is

$$z(t) = h(t) \otimes x(t) + w(t),$$

where $h(t)$ is the propagation transfer function and $w(t)$ is the noise. There are several solution strategies for this problem. They include:

- Use of channel invariant features, i.e. the features not affected by the propagation (Okopal et al., 2008).
- Use of the so-called blind deconvolution to reconstruct the clean target signal (Kil & Shin, 1996). This method required some knowledge of the environment parameters.
- Use of forward modelling to modify the clean data by the propagation and use it as the training data. This method also requires the knowledge of the environment (Liu et al., 2004).

In the present research we apply forward modelling for a range of environmental parameters to form the training data set and investigate how well we need to know the environment for building the training data set.

DATA DESCRIPTION

The clean acoustic data for this research were obtained in the water tank located in the Underwater Acoustic Scattering Laboratory (UASL) at Maritime Operations Division of DSTO. Two objects were used to represent different classes: a model of a generic submarine and a concrete cylinder of similar size. The generic submarine model is made of brass with a hull length of about 700 mm and a diameter of 80 mm. The concrete cylinder length is 560 mm and the diameter is 72 mm.

A schematic diagram of the experiment setup for the acoustic scattering measurements is shown in Figure 1. To collect acoustic scattering data, the object was suspended in the tank on four thin wires at a depth of 2 m. The tank is of 9 m length, 6 m width and 4 m depth. The acoustic unit transducer, which combines both transmitter and receiver in one housing, was positioned about 4 m away from the object at the same depth. The object was insonified with short pulses of 200 μs in duration. The echoes of the objects were recorded by the data acquisition system. The aspect angle of the object with respect to the transmitter was changed in the range from -5 to 362 degrees with one degree increment resulting in 368 echoes for each object. In case of the submarine model, zero degree aspect angle corresponds to the position of the model with its bow directed towards the transmitter. Similar for the cylinder, zero aspect angle means that its axis is aligned with the direction to the transmitter.

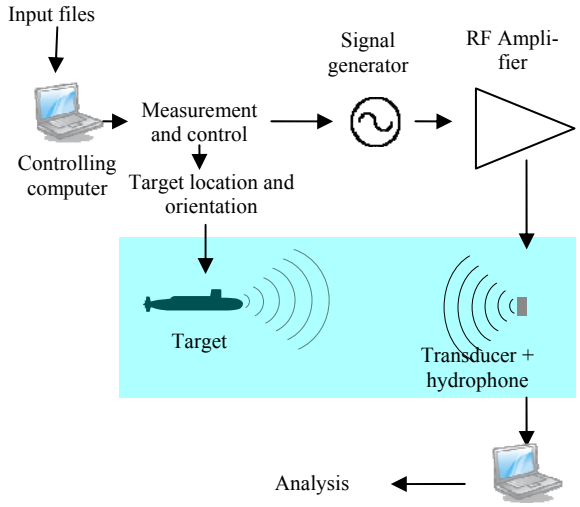

Figure 1. Setup for scattering measurements in UASL tank.

At each aspect angle four incident pulses with different frequency bandwidth were used. A technique of modifying the input pulse to account for the transfer function of the transmitter in order to generate incident pulses with sufficiently flat spectrum within the bandwidth was applied (Swincer et al., 2012). The following four frequency bands of the incident pulses were used:

- band 1: 80 – 220 kHz;
- band 2: 100 – 200 kHz;
- band 3: 120 – 180 kHz;
- band 4: 140 – 160 kHz.

The power spectra of the four incident pulses are shown in Figure 2.

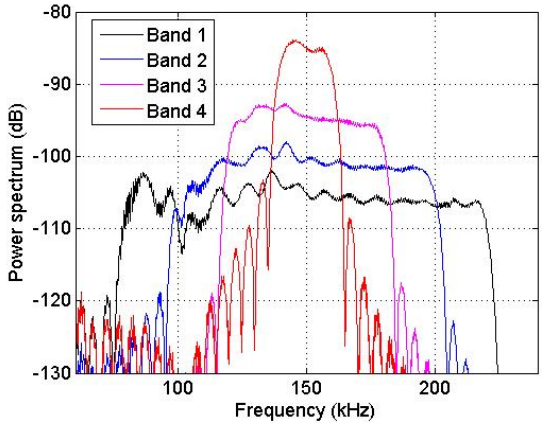

Figure 2. Power spectra of the incident pulses.

The echoes of the objects are strongly aspect dependent. This is clearly visible in Figure 3, where the so-called hour-glass plots of echoes of the generic submarine and the concrete cylinder are shown for the incident pulse with the frequency bandwidth of 80-220 kHz (band 1). These plots show the absolute value of the return signal in two dimensions: time and aspect angle. Thus, at the aspect angle of zero degrees, the earliest highlight is from the bow of the submarine model, the middle is from the fin, and the last is from the tail structure. At the aspect angle of 180 degrees the sequence of highlights is opposite. Of course the shortest and strongest echoes are at the aspect angles of 90 and 270 degrees, when the object is positioned with its side facing the transmitter. One can see from the figure that the clean data from the two different objects have clearly different structure. The submarine model data have fewer well structured highlights reflecting the simple design of the generic model. The data in the bottom plot have many randomly positioned highlights due to the rough surface of the cylinder with numerous cavities randomly located on the surface.

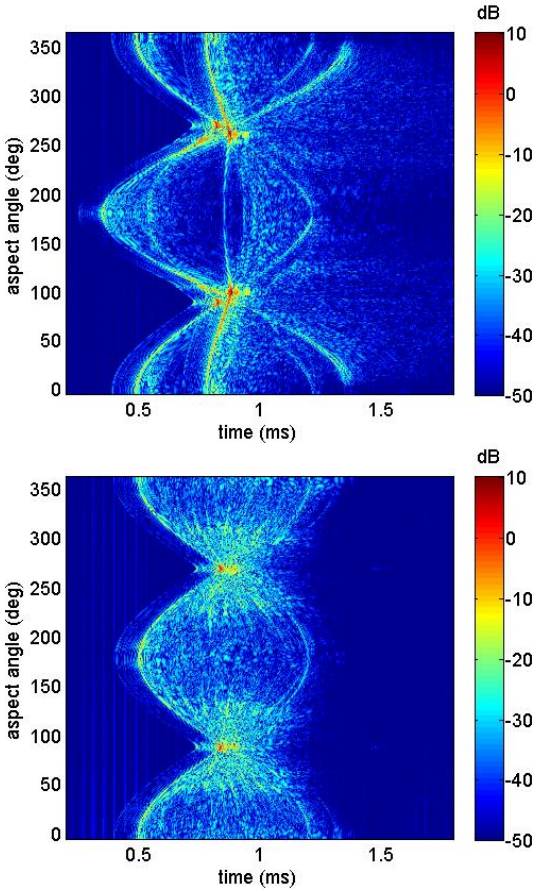

Figure 3. Hour-glass plot of echoes from the generic submarine (top) and concrete cylinder (bottom) for the incident pulse with the frequency bandwidth of 80 - 220 kHz (band 1).

The power spectra of the two objects are summarised in Figure 4 as two-dimensional frequency – aspect plots. The data in this plot are shown again for the bandwidth of band 1, 80 – 220 kHz.

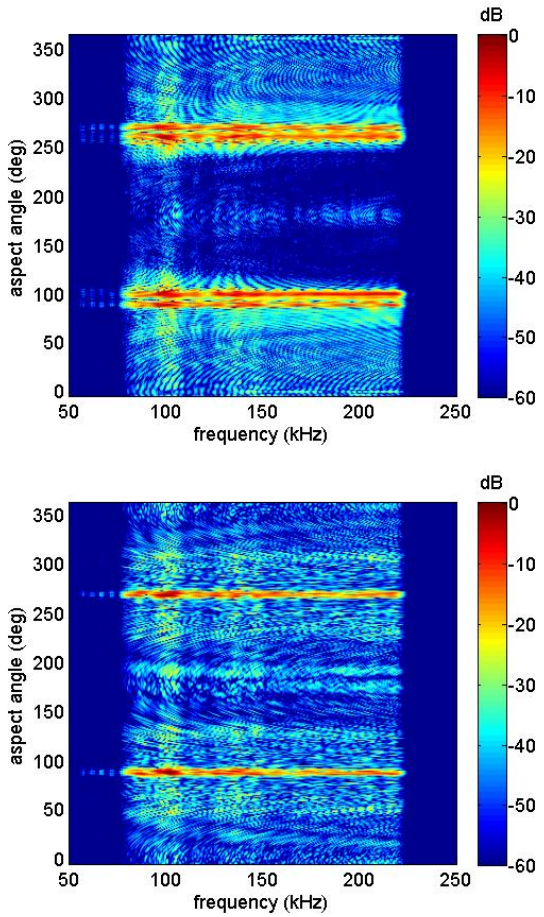

Figure 4. Frequency – aspect power spectrum plots for the generic submarine model (top) and the concrete cylinder (bottom) for the incident pulse with the frequency bandwidth of 80 – 220 kHz (band 1).

CLASSIFICATION FEATURES AND ALGORITHM

In this research we use the kernel ridge regression classifier as described in Saunders et al. (1998). The algorithm may be summarized as minimizing a quadratic loss function:

$$L = \lambda\|\mathbf{w}\|^2 + \sum_{i=1}^{l}(y_i - g(\mathbf{x}_i))^2,$$

where $\mathbf{x}_i, 1 \le i \le l$ is the data vectors of the training samples, y_i is their class labels, and $\lambda > 0$ regulates the norm. Given the training data, the classification algorithm finds the real valued linear function $g(\mathbf{x}) = \langle \mathbf{w}, \phi(\mathbf{x}) \rangle$ that minimizes the above quadratic loss function. The function $\phi(\mathbf{x})$ maps data vector $\mathbf{x}$ into the feature space. It can be shown that the label g($\mathbf{x}$) of a test vector $\mathbf{x}$ is calculated as

$$g(\mathbf{x}) = \mathbf{y}'(\mathbf{K}_{\text{train}} + \lambda \mathbf{I}_\lambda)^{-1}\mathbf{K}_{\text{test}}$$

where $\mathbf{y}$ is the vector of class labels of the training data. For example we may assign y = -1 for elements of class 1 and y = 1 for elements of class 2. $\mathbf{I}_\lambda$ is the identity matrix of dimension l corresponding to the number of training examples. The training and test kernel matrices, $\mathbf{K}_{\text{train}}$ and $\mathbf{K}_{\text{test}}$ are formed from training and test data vectors according to a selected kernel function. In this article we use the polynomial kernel, $\mathbf{K(x,z)} = (< \mathbf{x,z} > + c)^p$.

There are many approaches to calculating features from the active sonar echoes. In this report we limit ourselves to the set of twenty five features described in Tucker&Brown (2005). The features are based on the Short-Time Fourier Transform of the object echoes.

The results of classification can be presented in a different form such as the Receiver Operating Characteristic (ROC) curve, confusion matrix etc. Here we use two criteria to estimate the performance of the classification: the area under the ROC curve (AUC) and the average accuracy of classification. Those criteria are single valued parameters allowing presentation of the results of classification in a simple form. The average accuracy of classification is the average of the diagonal elements of the confusion matrix. In other words, it is the average of true positive and true negative classifications of the test data echoes.

The ROC curve is built by changing the decision threshold and calculating corresponding True Positive Rate (TPR) versus False Positive Rate (FPR). The True Positive Rate is calculated as the ratio of True Positives to the total number of Positives, and the False Positive Rate is calculated as the ratio of False Positives to the total number of Negatives. In our implementation the elements of class one are considered as Positives and the elements of class two are as Negatives. The decision threshold is specific to a classifier.

To build the ROC curve for the Ridge Regression Classifier we threshold the decision function $g(\mathbf{x})$. In its traditional form the decision is made according to the following rule:

$$\text{if } g(\mathbf{x}) < 0,\ \mathbf{x} \in P$$
$$\text{if } g(\mathbf{x}) > 0,\ \mathbf{x} \in N$$

where P stands for Positives, or class one, and N denotes negatives, or class two. Introducing threshold in the decision rule will result in the following rule:

$$\text{if } g(\mathbf{x}) < T,\ \mathbf{x} \in P$$
$$\text{if } g(\mathbf{x}) > T,\ \mathbf{x} \in N$$

The ROC curve is built by sweeping the threshold in the interval $-\max(|g(x)|) \le T \le \max(|g(x)|)$ and counting True Positives and False Positives for each value of the threshold.

The AUC can be calculated by numerical integration of the ROC curve or directly according to the following equation:

$$AUC = \frac{\sum_{i=1}^{n^+}\sum_{j=1}^{n^-} 1_{g(x_i^+) > g(x_j^-)}}{n^+ n^-},$$

where x^+ and x^- denote the positive and negative samples, respectively, n^+ and n^- are the number of positive and negative samples, respectively. 1_π is defined to be 1 if the predicate π holds and 0 otherwise.

SYNTHETIC DATA GENERATION

To test the classification approach we use synthetic active sonar data generated from the laboratory acoustic scattering data described above. The synthetic data are generated by

applying the channel filter and noise to the clean data echoes obtained in the laboratory in a free-field noiseless environment. Both test and training data are constructed in a similar way but using different values of parameters for the channel filter and noise. The algorithm for building the synthetic data is as follows. The clean data echo, $x_k(i)$, where i is the time sample index and k is the aspect angle index, is convolved with the specially constructed channel filter, $h(i;S_f)$. The parameter S_f characterises echo elongation due to multipath. The randomly generated noise, $w(i)$ is then added to the convolved signal:

$$z_k(i) = h(i;S_f) \otimes x_k(i) + w(i)$$

$$h(i;S_f) = h_0([nS_f]),\ 0 \le i \le [nS_f]$$

$$h_0[n] = \sum_{m=1}^{M} \gamma_m \delta[n - d_m],$$

$$M = 15,\ 0 \le n \le 99,\ d_m = rand(0,99)$$

$$\gamma_m = \pm e^{-\beta d_m},\ \beta = 0.024$$

The construction of the base channel filter, h_0, is taken from Anderson&Gupta (2008). We have introduced the parameter S_f to describe additional elongation of the echoes.

The noise is generated as follows:

$$w(i) = \alpha r_i$$

where r_i is the random Gaussian number with zero mean and unit standard deviation. The coefficient α is calculated from the required SNR:

$$\alpha = \frac{1}{\max(|r_i|)} 10^{(-SNR/10)}$$

Figure 5 and Figure 6 show the effect of applying the noise and channel filter, respectively, to the clean data at a certain aspect angle.

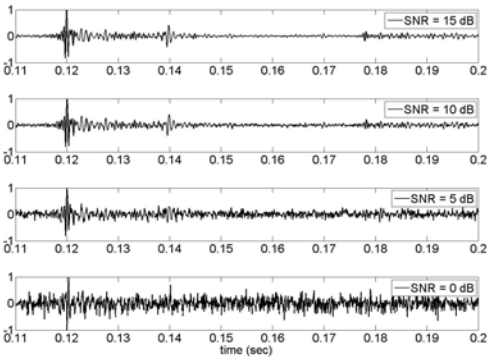

Figure 5. Effect of adding random noise to the clean data.

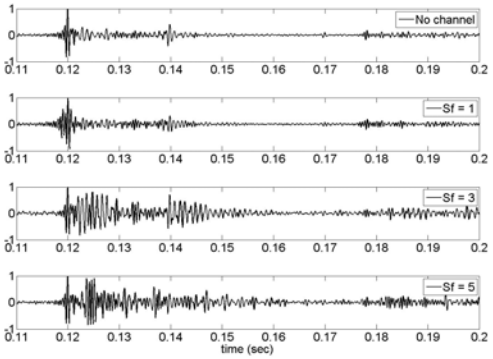

Figure 6. Effect of applying channel filter to the clean data.

RESULTS AND DISCUSSION

In this research we generate the test data using a fixed single value of SNR and the channel parameter, S_f, for all four frequency bands. The training data are generated for SNR and channel parameter values selected randomly from an interval:

$$\mathrm{SNR}^{(tr)} \in [\mathrm{SNR}_{\min}, \mathrm{SNR}_{\max}] \quad (1)$$

$$S_f^{(tr)} \in [S_{f\min}, S_{f\max}] \quad (2)$$

The above intervals include the corresponding values used for generating the test data. For each clean data echo N_{tr} training data echoes are generated with N_{tr} pairs of SNR and channel parameter each time randomly selected from the above intervals. In this research $N_{tr} = 3$ is used. Here we investigate how the width of the above intervals affects the classification accuracy.

First, we consider effect of the SNR interval. In this case we do not apply the channel filter to either the training or to the test data. The test data set is formed with SNR = 10 dB. Figure 7 shows the plots of the accuracy of classification as a function of the training data SNR range, or the length of the interval (1), $\mathrm{SNR}_{\max} - \mathrm{SNR}_{\min}$. One can see from the figure that increasing the SNR range of the training data certainly leads to the decrease of the classification accuracy. However, it is still sufficiently high for the first two frequency bands. This plot also clearly shows the importance of the wide frequency bandwidth for accurate classification.

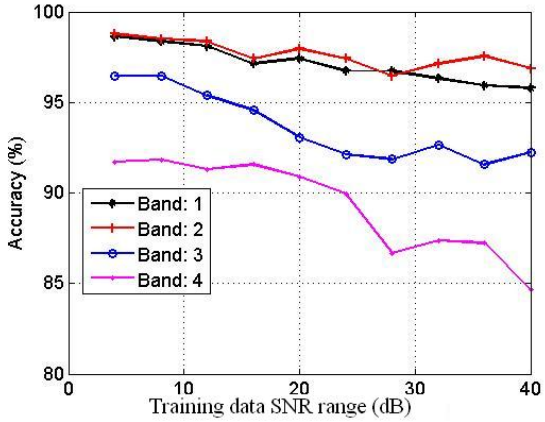

Figure 7. Classification accuracy versus width of the training data SNR interval for four different frequency bands. No channel filter applied.

Next, we apply a channel filter to both the test and the training data. First, we use the same channel filter parameter for both data sets, $S_f = 2$. In other words, $S_{f\max} = S_{f\min} = 2$ in the equation (2). We should note here that even if we use the same channel filter parameter, S_f, the filter still has an element of randomness in it due to the way it is generated. Therefore every echo in both test and training sets is generated using a slightly different channel filter. This leads to a change in the dependence of the classification accuracy on the training data SNR range (Figure 8).

The variation of the channel filter from echo to echo leads to some decrease in the classification accuracy for the first two bands. However, the classification for two narrower bands, bands 3 and 4, actually improves. It could be explained by the following. Pulses of the narrow frequency bands are more elongated comparing to the pulses of the wide frequency bands. As a result the hour-glass plots look more blurry for narrow frequency bands (Swincer et al., 2012). Obviously, multipath environment has the same effect of elongating, or blurrying, the object echoes. As a result, application of the channel filter blurrs the originally sharp wide band echoes and thus reduces the difference between the echoes of different bands.

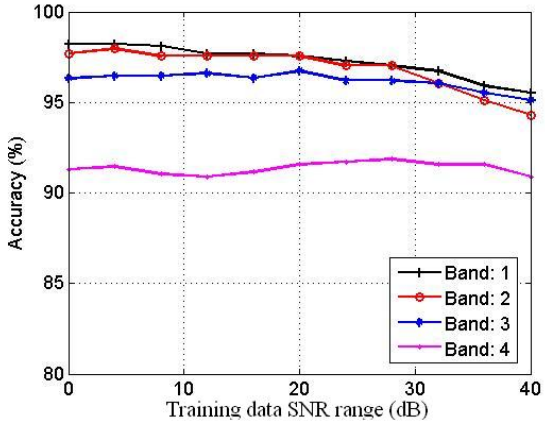

Figure 8. Classification accuracy versus width of the training data SNR interval for four different frequency bands. Channel filter parameter S_f =2 applied to both test and training data.

Now we will use the same SNR of 10 dB for both the test and the training sets and change the channel filter parameter range for the training set. As before this means that in equation (1) $\mathrm{SNR}_{\min} = \mathrm{SNR}_{\max} = 10\mathrm{dB}$. The channel filter parameter for generating the test data is kept constant, $S_f = 3$, at the centre of the parameter S_f interval of the training data. The results of this test are presented in Figure 9 for the classification accuracy and in Figure 10 for the area under the ROC curve. In both figures the horizontal axis shows the length of the channel filter parameter interval, i.e. $S_{f\max} - S_{f\min}$.

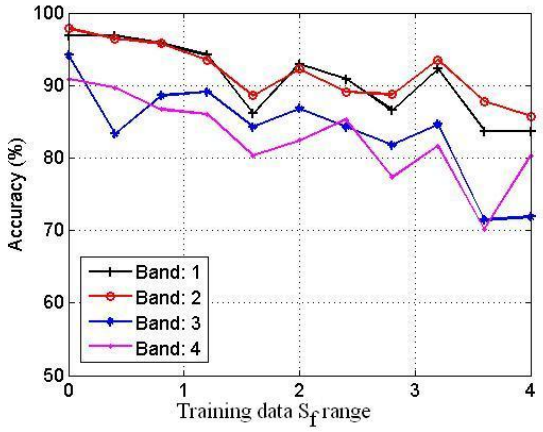

Figure 9. Classification accuracy versus width of the training data S_f interval for four different frequency bands. SNR=10 dB for both test and training data.

One can see from these figures that trying to include a wide range of multipath environments into the training data set may lead to a significant reduction in classifier performance. However, a reasonable range of multipath environments in the training data keeps the classifier performance at a sufficient level. The question of course is what is the reasonable range of multipath environments. From Figure 6 we can see that change from no multipath to the channel filter with the parameter $\boldsymbol{S_f = 5}$ gives a significant stretch to the echo. The question how this corresponds to the real environment can be answered using a higher fidelity acoustic environment model, which is left for future work.

We also see from these figures that again, the performance of the classifier is better in the two widest frequency bands.

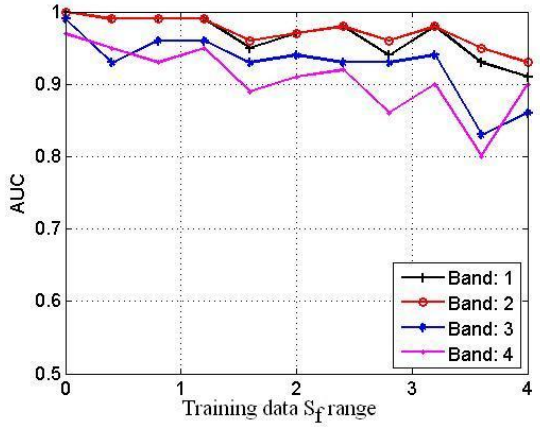

Figure 10. Area under the ROC curve versus width of the training data S_f interval for four different frequency bands. *SNR*=10 dB for both test and training data.

Finally, we generate the training data by randomly selecting both SNR and the channel filter parameter S_f in gradually increasing intervals. Again, the procedure of selecting both parameters is repeated three times, so for each combination of the intervals of SNR and the channel filter parameter the number of echoes in the training data set is three times the number of clean data echoes. As before, the test data set is generated from the clean data using SNR=10 dB and $S_f = 3$.

The results for the accuracy of classification are presented as surface plots in Figure 11 to Figure 14. As before, the performance of the classifier decreases with increasing SNR and channel filter parameter intervals for the training data. Apparently in these plots the effect of the channel filter parameter is much stronger than that of SNR.

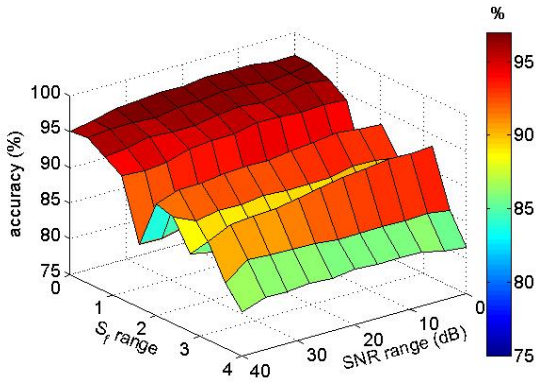

Figure 11. Classification accuracy versus training data intervals of SNR and the channel filter parameter. Test data parameters: SNR=10 dB, S_f =3. Band 1.

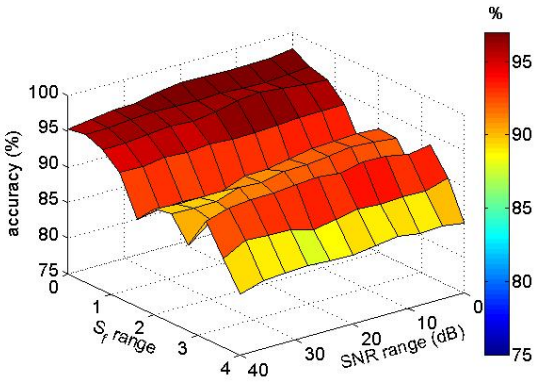

Figure 12. Classification accuracy versus training data intervals of SNR and the channel filter parameter. Test data parameters: *SNR*=10 dB, S_f =3. Band 2.

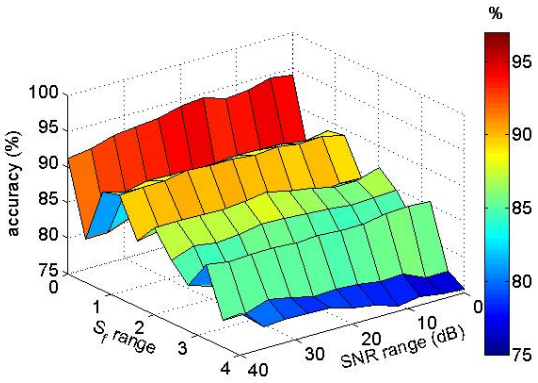

Figure 13. Classification accuracy versus training data intervals of SNR and the channel filter parameter. Test data parameters: SNR=10 dB, S_f =3. Band 3.

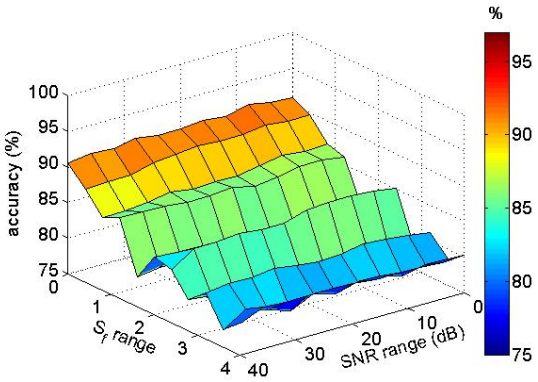

Figure 14. Classification accuracy versus training data intervals of SNR and the channel filter parameter. Test data parameters: SNR=10 dB, S_f =3. Band 4.

CONCLUSION

We considered an important issue in the active sonar classification of correct selection of the training data. Obtaining training data *in situ* for all targets and environments is not possible. That is why various techniques based on the use of the so-called clean data obtained from either mathematical modelling or controlled free-field experiments are gaining popularity. Here we considered one of such approaches based on the forward modelling to generate the training data. The question we investigated here is how well the environment parameters used for generating training data should match to

those of the test data. We based our research on the clean free-field echoes obtained at the DSTO Underwater Acoustic Scattering Laboratory test tank for two objects: scale model of a generic submarine and a concrete cylinder of similar size. The measurements were conducted for four frequency bands to additionally investigate the influence of the frequency bandwidth on the classifier performance. Both the test and the training data were generated from the clean data by applying two environment parameters: noise and a channel filter to model echo distortion due to multipath. While the test data echoes were generated using single values of SNR and the channel filter parameter, the parameters for generating the training data were randomly picked in gradually increasing intervals centred on the values of the test data.

As expected, the increase of the parameter intervals for generating the training data leads to a decrease of the classifier performance. However, the performance remains relatively high for reasonable intervals of SNR especially at the low values of the channel filter parameter. An increase in the channel filter parameter leads to a decrease in classification accuracy even in the cases when the same single value channel filter parameter was used for generating the training data. This is explained by the fact that the channel filter has additional randomness even at the same value of the channel filter parameter. This randomness results in difference between the echoes, which increases with increasing channel filter parameter. In the current research the channel multipath was modelled artificially to purely investigate the influence of the echo distortion on the classifier performance. In future work more realistic approach to accounting for the environment will be conducted using, for eaxmple, high-fidelity simulation of echoes from objects in acoustic waveguide (Kouzoubov, 2005).

All the results also show that the performance of the classifier is much higher when the pulses with wider frequency bandwidth are used.

REFERENCES

Ainslie, MA 2010, *Principles of Sonar Performance Modeling,* Springer.

Anderson, HS & Gupta, MR 2008, 'Joint deconvolution and classification with applications to passive acoustic underwater multipath', *J. Acoust. Soc. Am.*, vol. 124, pp. 2973-2983.

Kil, D & Shin, F 1996, *Pattern Recognition and Prediction with Applications to Signal Characterization*, AIP Press.

Kouzoubov, A 2005, *Modelling of Acoustic Scattering from an Object in a Waveguide*, DSTO-TR-1801, Defence Science and Technology Organisation.

Liu, H, Runkle, P, & Carin, L 2004, 'Classification of distant targets situated near channel bottoms', *J. Acoust. Soc. Am.*, vol. 115, pp. 1185-1197.

Okopal, G, Loughlin, P, & Cohen, L 2008, 'Dispersion-invariant features for classification, *J. Acoust. Soc. Am*., vol. 123, pp. 832-841.

Saunders, C, Gammerman, A, & Vovk, V 1998, 'Ridge regression learning algorithm in dual variables', *Proc. Of the 15th International Conference on Machine Learning (ICML98)*, pp. 515–521. Madison-Wisconsin.

Swincer, P, Nguyen, B, & Wood, S 2012, 'Method for the Generation of Broadband Acoustic Signals', *Australian Acoustical Society, Proceedings of Acoustics 2012 – Fremantle*.

Theodoridis, S & Koutroumbas, K 2006, *Pattern Recognition 3rd edition*. Academic Press.

Tucker, S & Brown, G 2005, 'Classification of Transient Sonar Sounds Using Perceptually Motivated Features', *IEEE Journal of Oceanic Engineering*, vol. 30, pp. 588-600.