2007 ASME International Mechanical Engineering Congress and Exposition
November 11-15, 2007, Seattle, Washington, USA

IMECE2007-41700

BAYESIAN MILITARY IMPULSE NOISE CLASSIFER

Brian Bucci
University of Pittsburgh
560 Benedum
Department of Mechanical Engineering
University of Pittsburgh
Pittsburgh, PA 15261
brian_arthur_bucci@hotmail.com

Dr. Jeffrey Vipperman
University of Pittsburgh
531 Benedum Hall
Department of Mechanical Engineering
University of Pittsburgh
Pittsburgh, PA 15261
jsv@pitt.edu

Dr. Amro El-Jaroudi
University of Pittsburgh
348 Benedum Hall
Department of Electrical and Computer Engineering
University of Pittsburgh
Pittsburgh, PA 15261
amro@pitt.edu

ABSTRACT

Civilian noise complaints and damage claims have created the need for stations to monitor military impulse noise. However, the stations currently in service suffer from numerous false positive detections (due to wind noise) of impulse events and often miss many events of interest. To improve the accuracy of military impulse noise monitoring, an algorithm based upon a Bayesian classifier with inputs of conventional and custom acoustic metrics is proposed. To train and evaluate the noise classifier approximately 1,000 waveforms were field collected. The final Bayesian noise classifier used kurtosis and crest factor and, the frequency domain metrics, spectral slope and weighted square error as inputs. The EM algorithm is utilized to fit multi-Gaussian distributions to the different classes of data. In testing the classifier performed to accuracies of up to 99.6%.

INTRODUCTION

The production of high amplitude impulse noise and its effects on the community surrounding military installations have become of great concern in recent years[1]. In an effort to assist the investigation of noise complaints and damage claims[2], monitoring stations have been installed in the areas surrounding some military installations. However, these stations are not able to detect many impulse events, especially those with a peak level (L_{pk}) below 119dB, and are plagued by false positive detections, mainly due to wind noise. The goal of this effort and previous efforts is to produce a noise classifier with improved impulse detection below peak levels (L_{pk}) values of 119dB and greater resistance to false positive detections[3,4].

In previous work, a data set containing waveforms of military impulse noise and non-impulsive noise was classified using an artificial neural network. The metrics used in the classifier were kurtosis, crest factor (CF), spectral slope (m), and weighted square error (WSE)[3,4]. It was deemed of interest to classify the same data set with a Bayesian classifier[5-8] to determine if comparable performance to the artificial neural network classifiers[3,4] could be achieved, while offering greater simplicity and possible better physical understanding of the process. The proposed Bayesian classifier would utilize the same data as the previous study along with the same computed metrics.

NOMENCLATURE

L_{pk}	Peak Sound Level
CF	Crest Factor
PSD	Power Spectral Density
m	Spectral Slope
WSE	Weighted Square Error
EM	Expectation-Maximization

DATA COLLECTION

As described in the previous body of work[3,4], approximately 1,000 usable waveforms were field collected to train and evaluate classifiers. Within the data set, there were 330 waveforms of military impulse noise, 560 waveforms of wind, and 110 waveforms of aircraft noise. Of the 330 recordings of military impulse noise, 66 contained more than one impulse event within the 2.1 to 2.5 second recording. The military impulse noise recordings consisted of 155mm Howitzers, 81mm mortars, 60mm rockets, M67 hand grenades, and Bangalore Torpedoes (strings of 3, (27lbs HE)). The military impulse noise records had L_{pk} values ranging from 80 to 138dB. The non-impulse noise records were wind noise and aircraft noise (F-16, A-10, C-130). The military impulse noise recordings were made at ranges between 1.5 km and 6 km from the noise sources and military aircraft noise recordings were made at distances of approximately 0.5 to 8 km. Although most of the energy of the noise sources to be measured is within the 0 to 100 Hz bandwidth, the data were sampled at 10kHz to verify that no key features in the higher frequency range were being neglected. Data were also measured at a variety of different locations and under a variety of different conditions in attempt to witness the largest array of factors that may affect the data.

DATA PROCESSING

Prior work describes the investigation of several conventional acoustic metrics and the Power Spectral Density (PSD), and their corresponding utility in applications of identifying military impulse noise[3,4]. Of these conventional acoustic metrics, it was shown that kurtosis and crest factor were the most useful in applications of identifying military impulse noise. Two additional scalar metrics, spectral slope (*m*) and weighted square error (*WSE*), were also developed and their utility in producing a more accurate artificial neural network noise classifier was quantified[3,4].

BAYES CLASSIFIER

A Bayesian Classifier is a type of classifier based on Bayes' rule of probability[5-8]. Bayes' rule is stated as

$$P(C_i \mid \underline{x}) = \frac{p(\underline{x} \mid C_i)P(C_i)}{p(\underline{x})}, \tag{1}$$

where $\underline{x}$ is the vector of observations and C_i is the i[th] class of data from n classes. In equation (1) $P(C_i|\underline{x})$ is the probability that a data point is from class C_i given the observation vector $\underline{x}$, $p(\underline{x}|C_i)$ is the conditional probability density of class i, $P(C_i)$ is the prior probability of data class C_i, and $p(\underline{x})$ is the probability density of observation vector $\underline{x}$, which is defined in terms of the conditional probability densities as

$$p(\underline{x}) = \sum_{i=1}^{n} p(\underline{x} \mid C_i)P(C_i). \tag{2}$$

In the Bayesian Classifier, there is a cost associated with classifying a data point into a particular class. The classifier works to minimize the cost of a decision. Let the expected conditional cost be defined as

$$K(C_i \mid \underline{x}) = \sum_{j=1}^{n} k_{ij} P(C_j \mid \underline{x}), \tag{3}$$

where k_{ij} is the cost of deciding the data is from class i when the data is from class j. The classifier performs the function

$$\arg\min_{i=1,..,n} K(C_i \mid \underline{x}), \tag{4}$$

to minimize the cost of classification. To minimize the average probability of misclassification cost k_{ij} is set to

$$k_{ij} = \begin{cases} 0, & for\ i = j \\ 1, & for\ i \neq j \end{cases}. \tag{5}$$

Taking into account equations (1) and (2), the conditional cost is now defined as

$$K(C_i \mid \underline{x}) = 1 - P(C_i \mid \underline{x}). \tag{6}$$

The Bayes decision rule[7] is thus given by

$$\arg\max_{i=1,..n} P(C_i \mid \underline{x}) = \arg\max_{i=1,..n} p(\underline{x} \mid C_i)P(C_i). \tag{7}$$

In using a Bayesian classifier it is convenient to assume a Gaussian distribution of data, thus conditional probability density of class i would be defined as

$$p(\underline{x} \mid C_i) = \frac{1}{(2\pi)^{m/2}\left(\det \underline{Q}_i\right)^{1/2}} \exp\left(-\frac{1}{2}\left(\underline{x} - \underline{\mu}_i\right)^T \underline{Q}_i^{-1}\left(\underline{x} - \underline{\mu}_i\right)\right), \tag{8}$$

where m is the dimension of the observation vector, Q_i is the covariance matrix, and μ_i is the mean vector. However, in the case of this problem, the metrics are not truly Gaussian. Thus, fitting a single Gaussian distribution to each class of data will likely not produce a classifier with adequate accuracy. Consequently, a Gaussian mixture fit to the data may provide a classifier with greater accuracy. The EM (Expectation-Maximization) algorithm was used to construct these distributions, which were compared to single Gaussian fits.

The EM algorithm is an unsupervised learning method that can be used to fit a set number of probability distributions to a distribution of data. This technique can be used to find patterns or clusters of data when the composition of the data set is largely unknown[6,9]. For this particular case, the EM algorithm is used to fit Gaussian mixture distributions to the data points of a particular class of noise and the distributions to be fit to the data are Gaussian. For each Gaussian distribution generated by the EM algorithm, the algorithm returns the covariance matrix, mean vector, and a factor relating the particular Gaussian's contribution to the entire distribution. Thus, in the case of fitting r Gaussians to the distribution of a particular class of noise, the conditional probability distribution is now defined as

$$p(\underline{x} \mid C_i) = \sum_{j=1}^{r} w_j \frac{1}{(2\pi)^{m/2}\left(\det \underline{Q}_{ij}\right)^{1/2}} \exp\left(-\frac{1}{2}\left(\underline{x} - \underline{\mu}_{ij}\right)^T \underline{Q}_{ij}^{-1}\left(\underline{x} - \underline{\mu}_{ij}\right)\right), \tag{9}$$

where w_j is the contribution of Gaussian j to the entire probability distribution. It is important to note that

$$\sum_{j=1}^{r} w_j = 1, \tag{10}$$

so that the distribution is a true probability density function.

The next step is to compute the prior probabilities of a given noise source. Let the likelihood ratio be defined as

$$\Lambda_i(\underline{x}) = \frac{p(\underline{x} \mid C_i)}{\sum_{\substack{s=1 \\ s\neq i}}^{n} p(\underline{x} \mid C_s)} . \qquad (11)$$

Additionally, let a value known as threshold of test be defined as

$$\xi_i = \frac{\sum_{\substack{s=1 \\ s\neq i}}^{n} P(C_s)}{P(C_i)} . \qquad (12)$$

The classifier can now be defined as shown in Figure 1.

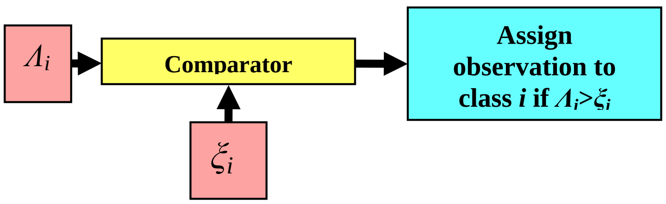

Figure 1: Bayesian Classifier

Now, by analyzing the distribution of the likelihood ratio it is possible to define the prior probability of a given class based on the desired characteristics of the classifier. If it is desired that the classifier be most sensitive to a particular class and false positive detections are not of great concern, then the probability of that particular class should be set higher. The inverse concept is also true. Additionally if $\log_{10}(\Lambda)$ is computed, the distributions of the likelihood ratios for particular classes will be approximately Gaussian[6]. This allows for the computation of the theoretical accuracy of a Bayesian classifier. Figure 2 provides a simple illustration of this concept. To compute the theoretical accuracy of the classifier the threshold of test is first chosen. The green curve represents the distribution of the likelihood ratio of the class of data we are attempting to identify and the blue curve is the distribution of the likelihood ratio of all other classes of data. The area under the green curve to the left of the threshold of test is the theoretical rate of false negatives and the area under the blue curve to the right of the threshold of test is the theoretical rate of false positives.

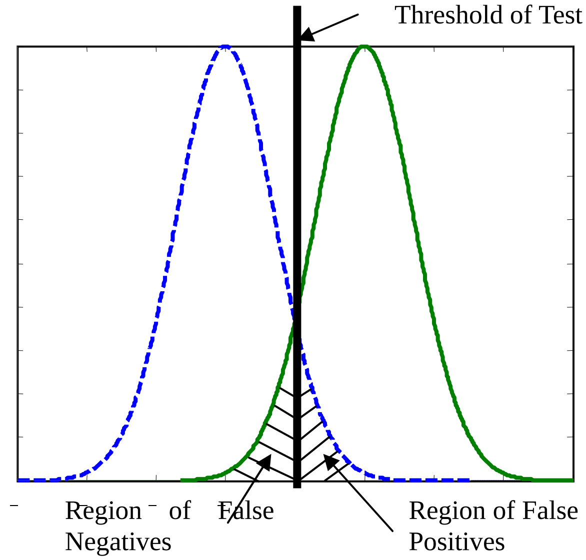

Figure 2: Simple Illustration of Log-Likelihood Distributions

Since both distributions are normal, the theoretical accuracy of the classifier is

$$A_{theoretical} = \frac{(1 - area_{false\_neg}) + (1 - area_{false\_pos})}{2} . \qquad (13)$$

For this example, selection the threshold where the two curves intersect as shown, would provide the highest accuracy.

RESULTS AND DISCUSSION

To compute the parameters of the classifier, half of the data was selected in a stratified random configuration (half of the aircraft noise data, half of the wind noise data, and half of the military impulse noise data) to serve as training data. The remaining half of the data was used to evaluate the performance of the classifier. In this classification problem there are three main classes of noise, aircraft noise, wind noise, and military impulse noise. However, it is only necessary to discern military impulse noise from non-impulse noise (aircraft noise and wind noise). Thus, aircraft noise and wind noise are grouped together as non-impulse noise.

The performance of the classifier is evaluated for a variety of configurations of number of Gaussians fit to each class of data. In the interest of presenting the results of this investigation in a clearer format, only the cases where equal numbers of Gaussians are used to fit each class of data are presented. The limiting factor in number of Gaussians that are fit to a distribution of data is that as the number of Gaussians increases, the covariance matrices can become poorly conditioned, thus making the EM algorithm unstable. The EM algorithm was relatively stable for fits of 8 Gaussians or less for each class of data. Figure 3 shows the results of algorithm for a typical case, (three Gaussians fit per class).

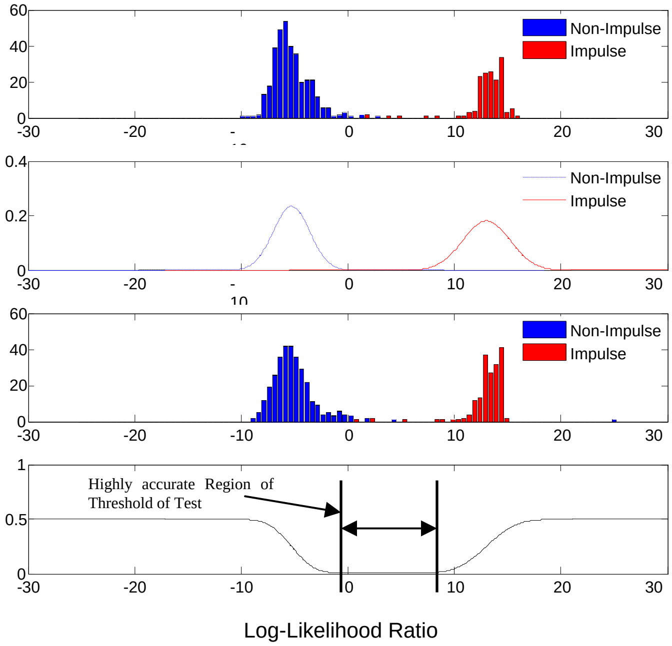

Figure 3: (top) Histogram of Log-Likelihood Ratio output for training data, (2nd) theoretical probability distribution for Log-Likelihood Ratio of classifier, (3rd) histogram of Log-Likelihood Ratio for testing data, (bottom) theoretical classifier error associated with value chosen for threshold of test.

From Fig. 3, it is noticeable that non-impulse and impulse distributions in the top and 3rd plots do resemble Gaussian distributions, as expected. The theoretical classifier, shown in the 2nd plot also appears to match the distributions presented in the top and 3rd plots. The accuracy of the classifier is taken to be the ratio of the total number of data points classified correctly to the total number of data points processed. In the bottom plot it is seen that there is quite a wide range of prior probability that should, theoretically produce classifiers with very similar accuracies. This is a positive quality because it enables the algorithm to be tailored to being more or less sensitive to military impulse noise, thus dealing with outliers, without sacrificing much accuracy.

A summary of the testing of the classifier is presented in Figure 4 and Tables I. and II. It is seen in Figure 4 that the accuracy of the classifier on the training data tended to increase as the number of Gaussian distributions used to model the data was increased. This is expected as modeling the distribution with more components should produce a more accurate approximation of the distribution of the data. The theoretical accuracy of the classifier also tended to increase as the number of Gaussians used was increased. Since the theoretical accuracy is based on the training data, this is also expected. In the case of the accuracy of the classifier on the testing data, accuracy first increases, peaks at 5 distributions, and begins to decrease slightly. The initial increase in accuracy is due to more Gaussian fits producing a better approximation of the actual data distribution. The ensuing decrease in accuracy is due to the increased number of Gaussians over-fitting the training data. This is a common problem with most classifier structures[7,8]. The classifier has become too specific to the training data and thus has lost some of its generalization capabilities.

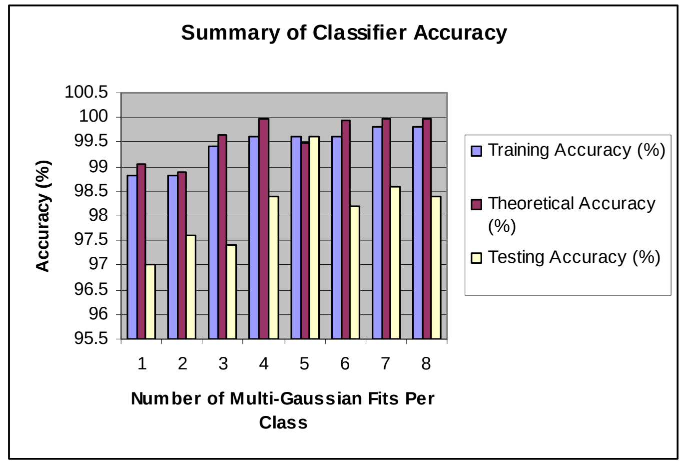

Figure 4: Summary of Training Accuracy, Theoretical Accuracy, and Testing Accuracy

Table I. summarizes the errors of the classifiers with respect to false positive and false negative detections. In all cases the ratio of false positives to false negatives appears to be of the same order of magnitude. This indicates that the choice of threshold of test is fairly neutral. It may be adjusted to favor mitigation of false positives or false negatives accordingly. A receiver operating curve is not included because the plot is not interesting since it is nearly rectangular, given the high level of *accuracy achieved*.

Table I. False Positive (FP) and False Negative (FN) Summary for Training, Theoretical, and Testing Cases

Number of Gaussians fit to each class of data	Training		Theoretical		Testing	
	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
1	0.80	0.40	0.45	0.48	1.50	1.50
2	0.80	0.40	0.31	0.79	1.40	1.00
3	0.40	0.20	0.12	0.25	1.60	1.00
4	0.40	0.00	0.01	0.02	1.20	0.40
5	0.20	0.20	0.15	0.39	0.40	0.00
6	0.20	0.20	0.03	0.03	1.40	0.40
7	0.00	0.20	0.01	0.00	1.20	0.20
8	0.00	0.20	0.01	0.00	1.40	0.20

Figure 5 shows the theoretical accuracy of the classifier for choices of the threshold of test parameter (ξ). It is notable that there is a wide range of prior probabilities that will produce a classifiers with comparable accuracies. Referring back to the bottom plot in Figure 3, this is also expected as there is a wide range of prior probabilities that may be selected to produce classifiers of comparable accuracy. In actual implementation, the choice of the priors would be based on the data observed and desired characteristics of the classifier. As previously stated, this gives the user of this algorithm the opportunity to mitigate either false positives or false negatives while retaining a classifier with a high degree of accuracy.

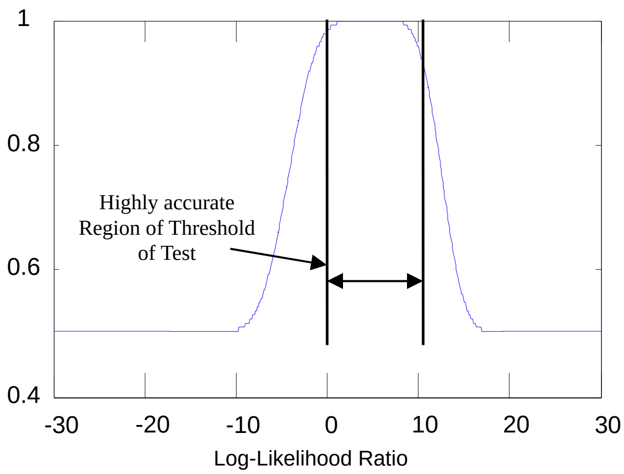

Figure 5: Theoretical classifier accuracy associated with value chosen for threshold of test.

CONCLUSION

This paper summarizes the extension of an effort to develop a simpler, yet accurate algorithm for identifying military impulse noise. In earlier work military impulse noise was identified using an artificial neural network with inputs of scalar metrics: kurtosis, crest factor, spectral slope, and weighted square error. Utilizing these same scalar metrics, a Bayesian classifier was developed. The Bayesian classifier was able to perform up to 99.8% accurate on training data and 99.6% accurate on testing data. This performance is very comparable to the ANN structures presented in earlier work. It was also shown that the Bayesian classifier could be tailored to reduce the amount of false positive or false negative detections, as per user request, while maintaining a classifier with a high degree of accuracy. Additionally, it is now possible to statistically characterize the observed values of the metrics, which may provide further insight into the behavior of the metrics for particular classes of noise.

.

ACKNOWLEDGMENTS

This research was supported wholly by the U.S. Department of Defense, through the Strategic Environmental Research and Development Program (SERDP). Special thanks to the program manager, Dr. John Hall, for his guidance and to the officials from Ft. Indiantown Gap, PA and MCB Camp Lejeune for their assistance in collecting the data.

REFERENCES

[1] Office of Economic Adjustment, Office of Assistant Secretary of Defense, and Economic Security. "Joint Land Use Study." November 1993.

[2] Luz, G. A.. "Suggested procedures for recording noise complaints at army installations." United States Army Environmental Hygiene Agency, Aberdeen Proving Ground, MD 21010-5422. HSE-08/WP Technical Guide, April 1, 1980.

[3] Bucci, B., Vipperman, J. S., "Artificial neural network military impulse noise classifier." IMECE2006-14065, Proceedings of IMECE 06: 2006 ASME International Mechanical Engineering Congress, November 5-10, 2006, Chicago, Illinois.

[4] Bucci, B. and Vipperman. "Performance of artificial neural network-based classifiers to identify military impulse noise." J. S. Journal of the Acoustic Society of America, (in press).

[5] D. Sivia and J. Skilling. *Data Analysis: A Bayesian Tutorial.* Oxford University Press, New York, New York, 2006.

[6] H. Stark and J. Woods. *Probability and Random Processes with Application to Signal Processing* Prentice Hall, Upper Saddle River, New Jersey, 2002.

[7] S. Abe. Pattern Classification: Neuro-fuzzy Methods and Their Comparisons, Springer-Verlag, London, England, 2001.

[8] Simon Haykin. *Neural Networks: A Comprehensive Foundation*, Prentice Hall, Upper Saddle River, New Jersey, 1999.

[9] Edwin Chong and Stanislaw Zak, *An Introduction to Optimization*. John Wiley and Sons, Inc. New York, NY, 2nd ed. 2001.