

TIME-FREQUENCY CLASSIFICATION USING A MULTIPLE HYPOTHESES TEST: AN APPLICATION TO THE CLASSIFICATION OF HUMPBACK WHALE SIGNALS

Geoff Roberts, Abdelhak M. Zoubir and Boualem Boashash

Signal Processing Research Centre
Q.U.T., GPO Box 2434, Brisbane Q.4001, Australia
email: ga.roberts@qut.edu.au

ABSTRACT

We present a non-stationary signal classification algorithm based on a time-frequency representation and a multiple hypothesis test. The time-frequency representation is used to construct a time-dependent quadratic discriminant function. At selected points in time we evaluate the discriminant function and form a set of statistics which are used to test the multiple hypotheses. The multiple hypotheses are treated simultaneously using the sequentially rejective Bonferroni test to control the probability of incorrect classification of one class. We show results for classifying three classes of humpback whale calls. The results demonstrate that this time-frequency method performs favourably when compared with a frequency domain method which assumes stationarity.

1. INTRODUCTION

In this paper we extend a frequency domain classifier for stationary signals [7] to a time-frequency classifier for non-stationary signals. The motivation for this extension is straightforward: the classical technique is only optimal (in the sense of minimising the probability of misclassifying an observation of one kind for a fixed misclassification rate of the other kind) if the signal is stationary. This led us to consider a technique that does not require the signal to be stationary. In particular, we introduce a time-varying quadratic discriminant function using the spectrogram. We apply the sequentially rejective Bonferroni test (SRBT) to the multiple hypotheses that can be constructed at different points in time from this discriminant function.

In the paper we will apply the method to the classification of humpback whales, *Megaptera novaeangliae*, recorded along the east coast of Australia. An important area of research is the study of the behaviour of the humpback whale so that their reaction to human

intervention can be monitored. Researchers are particularly interested in how the humpback whales song evolves over successive years - is there a pattern to the evolution of the song and ultimately, is the song indicative of the humpback's behaviour? The answers to these questions are beyond the scope of this paper, however, we present results for classifying humpback whale calls, which is the first step to a detailed analysis of the humpback's song.

In the next section we formalise the classification problem. In Section 3 we introduce a time-frequency based discriminant function for discriminating between two classes. In Section 4 we summarise the multiple hypotheses test which we use to make a classification. In Section 5 we show the results for classifying three humpback whale calls.

2. PROBLEM STATEMENT

In this section we formalise the problem of classifying two classes of signals. It is a straightforward extension for the case where there is more than two classes. Let $f_{\mathbf{X}}(\mathbf{x})$ be the density of $\mathbf{X} = [X_1, \dots, X_N]'$ a model for the observation of a discrete Gaussian random process and an element of the set $\{f_{\mathbf{X}}(\mathbf{x}|\boldsymbol{\vartheta}) : \boldsymbol{\vartheta} \in \Theta\}$ where the parameter vector $\boldsymbol{\vartheta} \in \Theta$ is unknown. Let Θ_H and Θ_K be subsets of Θ such that $\Theta_H \cup \Theta_K = \Theta$ and $\Theta_H \cap \Theta_K = \emptyset$. The global decision is viewed in the normal way as testing the hypotheses $H : \boldsymbol{\vartheta} \in \Theta_H$ against the alternative $K : \boldsymbol{\vartheta} \in \Theta_K$.

In the sequel we introduce a time dependent discriminant function which we use to construct a test at different times. Let $T_{n_i}(\mathbf{x}), i = 1, \dots, P$ denote the test statistics, where n_i is a time index and $P \leq N$. We would like to make a global decision based on the outcome of a collection of time localised tests given by $\{(H_1, K_1), \dots, (H_P, K_P)\}$ where (H_i, K_i) is the test corresponding to $T_{n_i}(\mathbf{x})$. The corresponding parameter subsets are $\{(\Theta_{H_1}, \Theta_{K_1}), \dots, (\Theta_{H_P}, \Theta_{K_P})\}$. Let I

be the set of indices for which $H_{i=1,\dots,P}$ are true. The global parameter set, under H , is given by the intersection of the local parameter sets, i.e., $\Theta_H = \bigcap_{i \in I} \Theta_{H_i}$. To control the error rate for each local decision we set a local level of significance, α_i . However, this does not control the global level of significance, α . In general, α will be much higher than the local levels, α_i [4]. In Section 4 we outline the SRBT, which makes a global decision based on multiple hypotheses, while maintaining a global level of significance.

In the next section we introduce a function based on a time-frequency representation for generating the test statistics, $T_{n_i}(\mathbf{x})$.

3. TIME-FREQUENCY DISCRIMINATION

In this section we introduce a time-frequency discriminant function which is an extension of the power spectrum discriminant function given in [7]. Let $S_{\mathbf{x}}(n_i, \frac{2\pi k}{N})$ be the discrete time, discrete frequency representation (TFR) of a random variable X_n , for $n = 1, \dots, N$. For the case of classifying a signal into one of two classes, we define the time-dependent discriminant:

$$T_{n_i}(\mathbf{x}) = \sum_{k=0}^{N-1} \hat{S}_{\mathbf{x}}(n_i, \frac{2\pi k}{N}) \left(\hat{S}_{\mathbf{x}|H_i}^{-1}(n_i, \frac{2\pi k}{N}) - \hat{S}_{\mathbf{x}|K_i}^{-1}(n_i, \frac{2\pi k}{N}) \right) \quad (1)$$

where: $\hat{S}_{\mathbf{x}}(n_i, \frac{2\pi k}{N})$ is an estimate of the TFR from $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$, a realisation of \mathbf{X} ; $\hat{S}_{\mathbf{x}|H_i}(n_i, \frac{2\pi k}{N})$, and $\hat{S}_{\mathbf{x}|K_i}(n_i, \frac{2\pi k}{N})$, are estimates of the TFRs representing the two different classes and are assumed to be non-zero; $n_i = 1, \dots, N-1$; and $k = 0, \dots, N-1$, are the discrete frequency samples.

The discriminant function given by Eq (1) returns a value at each time, n . Each value is used to construct a hypothesis, which are then combined and treated simultaneously. This approach differs from previous time-frequency based methods [2, 8, 1] where the solutions involve integration over time to form a single hypothesis which can lead, in practical situations, to misclassification.

Although any TFR could be used in Eq (1) we use the spectrogram because its statistics are well known. The spectrogram is defined as

$$I_{xx}(n_i, \frac{2\pi k}{M}) = M^{-1} \left| \sum_{m=0}^{M-1} x_m w_{m-n} e^{-j \frac{2\pi k}{M} m} \right|^2 \quad (2)$$

for $n_i = 0, \dots, N-1$, $k = 0, \dots, M-1$, $M \leq N$; and w_m is an appropriate window [3] of length M . Under the assumption that the spectrogram is flat over L adjacent frequencies, we can use the smoothed spectro-

gram as an estimate of the time-frequency representation, i.e.,

$$\hat{S}_{\mathbf{x}}(n_i, \frac{2\pi k}{M}) = L^{-1} \sum_{l=-(L-1)/2}^{(L-1)/2} I_{xx}(n_i, \frac{2\pi(k+l)}{M}) \quad (3)$$

Under H_i , $T_{n_i}(\mathbf{X})$ can be assumed to be normal. An estimate of the mean is given by [7]

$$E\{T_{n_i}(\mathbf{X})|H_i\} = \sum_{k=0}^{M-1} (\hat{S}_{\mathbf{x}|K_i}^{-1}(n_i, \frac{2\pi k}{M}) - \hat{S}_{\mathbf{x}|H_i}^{-1}(n_i, \frac{2\pi k}{M})) \hat{S}_{\mathbf{x}|H_i}(n_i, \frac{2\pi k}{M}) \quad (4)$$

and an estimate of the variance can be shown to be given by

$$\begin{aligned} \text{var}\{T_{n_i}(\mathbf{X})|H_i\} = & \sum_{k=0}^{M-1} (\hat{S}_{\mathbf{x}|K_i}^{-1}(n_i, \frac{2\pi k}{M}) - \hat{S}_{\mathbf{x}|H_i}^{-1}(n_i, \frac{2\pi k}{M}))^2 \\ & \times (E\{|I_{xx}(n_i, \frac{2\pi k}{M})|^2|H_i\} - \hat{S}_{\mathbf{x}|H_i}(n_i, \frac{2\pi k}{M})^2) \end{aligned} \quad (5)$$

Similarly, by replacing H_i with K_i the mean and variance can be estimated under the alternative.

In the next section we outline a multiple hypotheses test, which uses the results presented in this section, for treating the test statistics, $T_{n_i}(\mathbf{x})$ simultaneously.

4. A MULTIPLE HYPOTHESES TEST

As discussed in Section 2, we would like to test the global hypotheses H based on the outcomes of a set of local decisions, $\{(H_1, K_1), \dots, (H_P, K_P)\}$. In this section we describe the SRBT [5], which treats a collection of related hypotheses simultaneously while maintaining the global level of significance.

In addition to the test statistics, $T_{n_i}(\mathbf{x})$, the SRBT requires the probability values

$$\mathcal{P}_i = P(T_{n_i}(\mathbf{X}) \geq T_{n_i}(\mathbf{x})|H_i)$$

for $i = 1, \dots, P$. The \mathcal{P} -values are the probability that the test statistics, $T_{n_i}(\mathbf{X})$ exceeds the observed quantity, $T_{n_i}(\mathbf{x})$ under H_i . The \mathcal{P} -values are sorted in ascending order. Let $\mathcal{P}_{(1)}, \dots, \mathcal{P}_{(P)}$ be the sorted probability values and $H_{(1)}, \dots, H_{(P)}$ be the corresponding hypotheses. If $\mathcal{P}_{(1)} > \alpha/P$ then accept all hypotheses; otherwise reject $H_{(1)}$ and proceed to the next test $H_{(2)}$. This scheme is illustrated in Figure 1.

It is shown in [5] that this tests maintains the global level of significance.

To extend this method to classify more than one class we used Eq (1) to perform pairwise comparisons between each pair of classes.

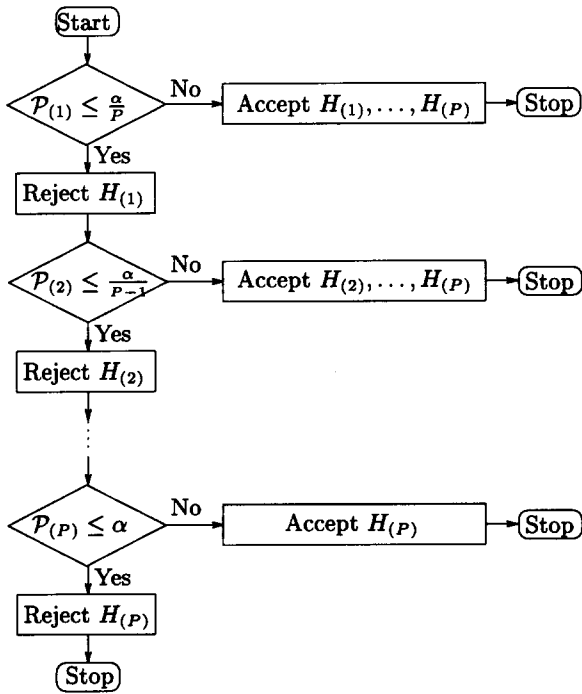


Figure 1: Sequentially rejective Bonferroni test.

5. RESULTS

In this section we show results for classifying three different humpback whale calls. The humpback whale signals were recorded along the east coast of Australia during their annual migration in 1995. The signals were digitised at 16kHz and decimated to 8kHz. Each call was segmented by analysing the signal aurally and with a spectrogram. For each class there were 15 members each 4096 samples long. Each signal was classified using the remaining signals to create the estimates of the class TFRs. The signals were filtered using a 70 tap FIR high pass filter with cutoff frequency at 120Hz. This was to remove low frequency noise which was present in some of the samples. The estimates of the class spectrograms for classes 1,2, and 3 are shown in Figures 2, 3, and 4 respectively. The discriminant function given in Eq (1) requires the inversion of a spectrogram, which for values close to zero will adversely affect the calculation. A small regularisation constant is added to each spectrogram to alleviate this problem. A window length of 256 was chosen which results in 16 local hypotheses. The results for the SRBT for two global level of significance, $\alpha = 10\%$ and $\alpha = 1\%$ are shown in Table 1 where each row, $C1 - C3$, represents the class of the input signal and each column, $C1 - C3$, represents the classification result for this input. The column, NC , gives the number of "no classification"

	$\alpha = 10\%$				$\alpha = 1\%$			
	$C1$	$C2$	$C3$	NC	$C1$	$C2$	$C3$	NC
$C1$	15	0	0	0	15	0	0	0
$C2$	0	15	0	0	0	15	0	0
$C3$	0	1	14	0	0	1	13	1

Table 1: Classification results using the SRBT.

	$\alpha = 10\%$				$\alpha = 1\%$			
	$C1$	$C2$	$C3$	NC	$C1$	$C2$	$C3$	NC
$C1$	15	0	0	0	15	0	0	0
$C2$	0	15	0	0	0	13	0	2
$C3$	0	0	10	5	0	2	7	8

Table 2: Classification results using Power Spectra.

results, i.e., when the algorithm could not discriminate between the three signal at the set level of significance. The spectrograms were not smoothed in either of these experiments.

If we take the spectrogram window length equal to the signal length then the resulting classifier is the power spectrum classifier proposed in [7]. The results for this method, which is optimal for a fixed significance level, for stationary signals are shown in Table 2. For $\alpha = 10\%$ a smoothing window of length 8 was used. However, for $\alpha = 1\%$ no smoothing was used. For lower values of α discrimination occurs more frequently in the tails of the distribution. Smoothing reduces the likelihood of the value of the test statistic being in the tails. Therefore, smoothing increases the number of no classification results and so was not always used.

From Tables 1 and 2 it can be seen that the SRBT based method performs at least as well as the power spectrum based method. For class 3 the SRBT consistently gives superior results due to a highly non-stationary feature at 0.3 second and 2kHz.

6. CONCLUSION

We have presented a new method for the classification of non-stationary signals by combining time-frequency analysis with multiple hypothesis testing. A time-frequency discriminant function was used to form a set of time localised tests. A global decision was made by treating the tests simultaneously using the sequentially rejective Bonferroni test. This multiple hypotheses test was used to control the global level of significance. Results are shown for the classification of three classes of humpback whale calls. The multiple hypotheses method showed superior performance when compared to a power spectrum based method which assumes stationarity.

7. ACKNOWLEDGEMENTS

The first author would like to thank Michael Noad, from the University of Sydney, and Trish and Wally Franklin from The Oceania Project for the humpback whale data.

8. REFERENCES

- [1] R.A. Altes. Detection, estimation, and classification with spectrograms. *J. Acoust. Soc. Am.*, 67:1232–1246, April 1980.
- [2] B. Boashash and P. O'Shea. A methodology for detection and classification of some underwater acoustics signals using time-frequency analysis techniques. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 38(11):1829–1841, nov 1990.
- [3] D.R. Brillinger. *Time Series Data Analysis and Theory*. McGraw-Hill, USA, 1981.
- [4] Y. Hochberg and A. Tamhane. *Multiple Comparison Procedures*, page 363. John Wiley & Sons, USA, 1987.
- [5] S. Holm. A simple sequentially rejective multiple test procedure. *Scand J Statist*, 6:65–70, Nov 1979.
- [6] G. Roberts, A.M. Zoubir, and B. Boashash. Time-frequency discriminant analysis for non-stationary Guassian signals. In *Proceedings of ISSPA '96*, volume 1, pages 33–36, Gold Coast, Australia, August 1996.
- [7] R.H. Shumway. Discriminant analysis for time series. In P.R. Krishnaiah and L.N. Kanal, editors, *Hand. of Stat.* 2, chapter 1. North-Holland, New York, USA, 1982.
- [8] I. Vincent, C. Doncarli, and E. Le Carpentier. Non stationary signal classification using time-frequency distributions. In *Proc. of the Inter. Symp. on Time-Frequency and Time-Scale Analysis*, pages 233–236, Philadelphia, USA, Oct 1994.

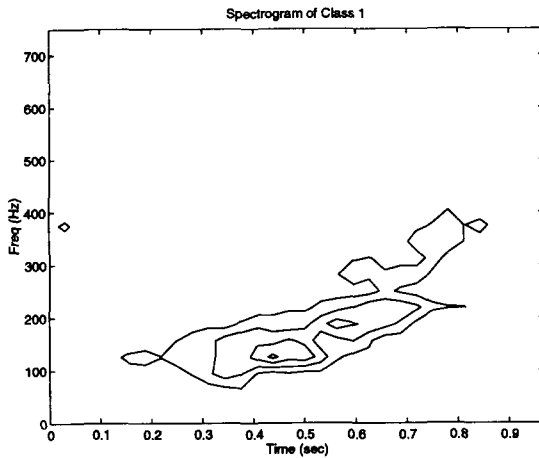


Figure 2: Estimate of class 1 spectrogram.

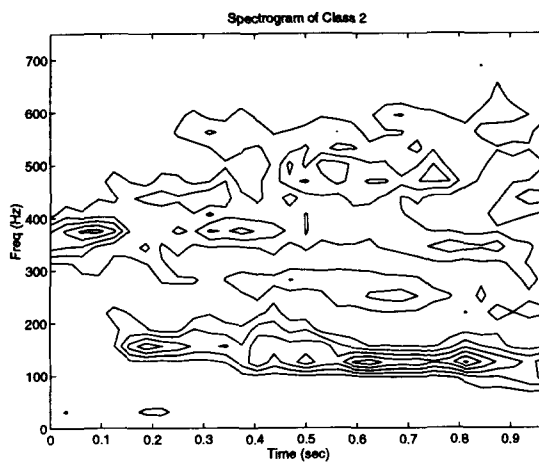


Figure 3: Estimate of class 2 spectrogram.

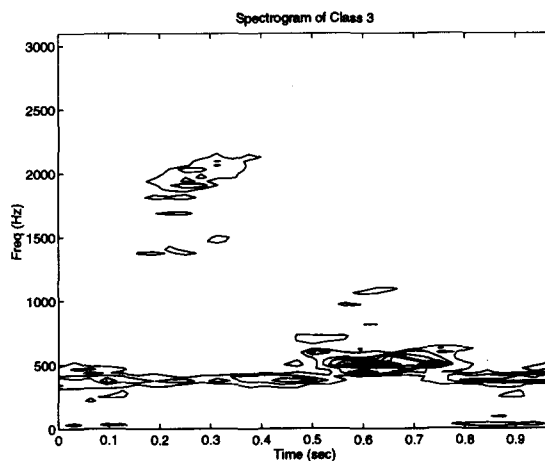


Figure 4: Estimate of class 3 spectrogram.