ACOUSTICAL FEATURE EXTRACTION FROM AIRCRAFT AND TRAFFIC NOISE

R. Mohajeri* and F. R. Fricke**
***Department of Mechanical and Mechatronic Engineering, University of Sydney, NSW 2006 Australia**
****Department of Architectural and Design Science, University of Sydney, NSW 2006 Australia**

This paper was awarded the 1995 PRESIDENT'S PRIZE
The President's prize, established in 1990 by the Australian Acoustical Society, is awarded to the best technical paper presented in the Annual Australian Acoustical Society Conference.

Abstract: For the purpose of developing a real-time transportation noise recognition system, a variety of acoustical features and statistical models of signals are reviewed. Finding an appropriate acoustical criteria to discriminate between several types of environmental noise is the goal of the present paper. The main emphasis is put on the discrimination of transportation noise, music and speech. The results could be used in transportation noise activated systems for the "intelligent" monitoring and control of aircraft and motor vehicle noise.

1. INTRODUCTION

The goal of this paper is to find some statistical features from transportation noise signals to be used in a noise activated monitoring and control system. A variety of different parameters and signal processing methods have been examined to distinguish between transportation noise and other environmental sound sources such as speech and music. Some parameters are required to distinguish between the low frequency, random nature of transportation noise and more complex spectra of speech, music and other sources of environmental noise. In this study, aircraft, heavy vehicle and mixture of traffic in dry and wet weather are considered as the transportation noise sources.

A real-time intelligent system requires real-time data acquisition, detection and pattern recognition. The pattern recognition problem can be divided into several stages of which feature extraction and source classification are two of the most important ones.

The feature extraction methods used in seismology, which deal with low frequency vibration recognition, are found applicable to transportation noise recognition. The methods for automatic discrimination between nuclear explosions and natural seismic activity are also worth examining. Speech and speaker recognition systems have employed a variety of signal processing and pattern recognition methods, some of which may be used to identify transportation noise.

The energy of a signal, the zero-crossing rate, the linear prediction coefficients and the autocorrelation function are the time domain functions which have been used for successful waveform recognition in other fields [1,2,3]. The frequency spectrum envelope, averaged-peak frequency, maximum-peak frequency in each band and the plot of the first and second most prominent frequencies are the other features that have already been used for transient events and aircraft type recognition [1,2,8]. The Euclidean distance and the nearest neighbourhood are applied to compare the pattern vector with the sample vector.

In the following sections the characteristics of transportation noise are discussed and a signal modeling procedure is developed. A pattern classification method is then presented together with a recognition algorithm.

2. CHARACTERISTICS OF TRANSPORTATION NOISE

Aircraft noise and traffic noise are non-stationary and time dependent. Transportation noise emitted by cars, heavy vehicles and aircraft is variable in both time and frequency domains. Most of the acoustic energy is concentrated in the lower frequencies. The amplitude and frequency of land transportation noise is strongly dependent on the acceleration, speed, operating mode and the conditions of the vehicle and road. Also, the distance between the source and the observer as well as the weather conditions are very important parameters. Many random parameters such as the noise of vehicle brakes and the sound of car horns and the vibration of trailers are mixed with transportation noise. In the case of aircraft noise, the type of aircraft and propulsion system, load, take off or the landing mode, the angle of flight and the weather conditions all determine the characteristics of the noise heard.

3. TIME VARIATION OF ENERGY AND SOUND PRESSURE LEVEL

In real-time signal processing, one of the most important parameters to be measured is the energy of the signal. This parameter has been used to differentiate between voiced and unvoiced sounds and silence in [3]. Fig. 1(a) is the result of monitoring the energy of the signal from different environmental noise sources. It shows that the energy of background noise is considerably less than the energy of the noise events of interest. When the sensing system is turned off the energy is equal to zero. In the present research an energy threshold level is used to activate the recognition system and discriminate between the "Off" position of sensing system and the "Background" noise. Also, the energy is an indication of the beginning and end points of noise events.The sound pressure level of an acoustical signal is another parameter which can be easily measured Fig. 1(b) shows monitoring of sound pressure level from the same events in Fig. 1(a) under the same conditions. As the logarithmic plot is much noisier than the linear plot, particularly for the values close to zero, the linear plot is more useful for activating the recognition system.

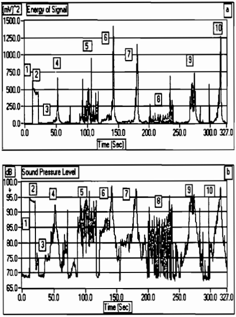

Figure 1a) Energy monitoring of environmental noise, b) Sound pressure level monitoring of the same noise events shown in (a). The sampling frequency was 5 KHz for 512 samples.

In Fig. 1(a) and (b) each number represents a separate class of noise events as follow: 1) the sensing system is off, 2) a B&K 4230 calibrator, pure tone 1 KHz, 3) Background noise, 4) Aircraft noise, 5) A piece of classical music, 6) Aircraft noise, 7) Aircraft noise, 8) Speech, 9) Rock music and 10) Aircraft noise.

4. LINEAR PREDICTION MODEL OF TIME SIGNAL

The "Linear Prediction Model" of time signals is one of the most powerful existing techniques available to discriminate between different waveforms. Different characteristics of this method have been used successfully in the field of speech and speaker recognition [3,4,5]. Also, there is a wide range of applications in the field of seismology to classify the nuclear explosions and earthquakes automatically [1,2,8]. The basic idea of the linear prediction model of a signal is to model the signal as a linear combination of its past and present values with a hypothetical input to the system. The input of such a model is white noise or an impulse and the output is the given signal. In a stationary and invertible process, if x_n is the time signal, the general form of such a statistical model can be represented as:

$$x_n = \sum_{k=1}^{p} a_k x_{n-k} + \sum_{l=1}^{q} b_l w_{n-l} + e_n \,, \quad 1 \le k \le p, \quad 1 \le l \le q \tag{1}$$

where, a_k, b_l are the model coefficients, w_n is the white noise sequence, p and q are the orders of model and e_n is the residual error. Thus, the output signal, x_n is a linear combination of past outputs and present and past inputs. The estimated coefficients, a_k and b_l, are useful parameters for pattern classification. The frequency domain representation of equation (1) is:

$$H(z) = G \frac{1 + \sum_{l=1}^{q} b_l z^{-l}}{1 + \sum_{k=1}^{p} a_k z^{-k}} \tag{2}$$

where G is the system gain. H(z) is the general pole-zero model which is called an autoregressive moving average model ARMA(p,q). Equation (1) can be simplified as an all-pole model or autoregressive model, AR(p), equation (3), when $b_l = 0$. It can be considered as a recursive filter with feedback as follows:

$$x_n = \sum_{k=1}^{p} a_k x_{n-k} + e_n \tag{3}$$

where, p is the order of the AR model. The most applicable frequency spectral match for the AR model is found by dividing G^2 by the magnitude squared of the FFT from the sequence of $1, a_1, a_2, \ldots, a_p$. Fig. 2(a) is the result of 20-pole fit (AR) to a power spectrum of a signal computed from the noise of an aircraft in landing mode.

5. DISCRIMINATION OF AR COEFFICIENTS

Transportation noise mainly contains low frequency energy and the randomness of the noise is greater than other acoustic sources such as speech and music. A low order AR model for the 1 KHz filtered data from different aircraft, heavy vehicles and mixture of traffic noise has been examined. There is a strong similarity between the plot of the first and second AR coefficients of aircraft and road traffic noise. The same similarity has been found for classical music, rock music and continuous speech. A second order AR model gives a considerable separation between speech, music and transportation noise. In Fig. 2(b) the results of 200 data from two pieces of classical music, a piece of rock music and an interview are compared with the same number of data from aircraft noise and a mixture of traffic noise. Fig. 2(b) shows that the second order AR coefficients are useful in recognising the transportation noise. They also have the potential of dealing with low frequency noise recognition.

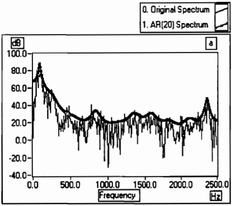

Figure 2. a) Frequency spectrum of the autoregressive model with 20 poles using 512 samples from aircraft noise in landing mode.

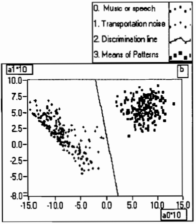

Figure 2.b) Plot of the first and the second coefficients of a 2-pole AR model to discriminate between speech/ music and transportation noise.

6. ACOUSTICAL PATTERN CLASSIFICATION OF ENVIRONMENTAL NOISE SOURCES

After studying different measurable and available features and their abilities to discriminate between different acoustical sources, an attempt was made to combine the extracted features to make a decision system. The energy of a signal is a good indication of the start and the end of noise events. In addition the energy of the signal is a discriminating factor between an "Off" signal and "Background" noise. In the present study the pattern recognition task is divided into two classes of recognition:

1. Short-time acoustical source recognition, which is quick enough not to miss the noise event. This is important as the ultimate goal of source recognition is to activate a system as the noise event starts. The data acquisition, data processing, feature extraction and patten recognition must not take more than a portion of a second. In such a short time extracting the detailed information from the acoustical source is impossible. (Basically a person's hearing system is also not able to recognise similar sound sources in a very short time either).
2. Long-time acoustical source recognition, which does not take more than 50 seconds duration. This case is useful to identify the particular vehicle or aircraft. To differentiate between different types of aircraft, a satisfactory number of data is required to compare with the data base already made from statistical manipulation of data recorded from aircraft noise.

7. SHORT-TIME ZERO-CROSSING OF TIME SIGNAL

The zero-crossing is the number of zero crossings level of signal per duration of sampling. In the case of a pure tone, zero-crossing is a good measurement for frequency, but for broadband signals eg. transportation noise, it can only be a rough indication of frequency content. Direct measurement of zero-crossing for a discrete signal is difficult because the value of the signal is rarely zero. The process of detecting zero-crossing at a given time is based on the sign change of the product of multiplication of one data point value before and one data point value after that time.

8. SHORT-TIME AUTOCORRELATION FUNCTION

The autocorrelation function has a lot of useful properties in real-time signal processing and detection. An important property of the autocorrelation function, for the present study, is the considerable difference in its form for transportation noise and the other acoustical sources. The main reasons for this difference are likely to be the low frequency nature of traffic noise and its randomness. The number of peaks and the number of zero-crossing in the case of speech and music, is considerably higher than that of aircraft and traffic noise. (The number of zero-crossings of the autocorrelation function is totally different with to meaning of the zero-crossing function of time signals). Fig. 3(a) is the result of 200 plots of zero-crossings of the autocorrelation function with respect to the zeroth element of the autocorrelation function.

9. THE LOCATION OF PEAKS

The acoustic energy of transportation noise is mainly concentrated in the low frequencies. It seems to be reasonable to draw the plot of the first and second peak frequencies which usually occur at frequencies lower than 1000 Hz. Fig. 3(b) is the plot of f1 and f2, the first and second frequencies of the most prominent energy peaks in the 200 samples of heavy vehicle noise and the speech and music, that were given in previous sections. The discrimination line shows a good separation between the two classes of acoustical sources.

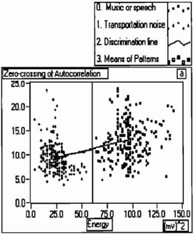

Figure 3. a) Feature extraction from the autocorrelation function for 200 samples of traffic noise and speech/music.

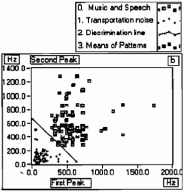

Figure 3.b) A plot of the first and the second frequencies of mixture of heavy vehicle noise compared with speech/ music. The sampling frequency is 5 KHz with 512 samples used.

10. EUCLIDEAN DISTANCE AND NEAREST NEIGHBOURHOOD

In the case of short-time data acquisition and pattern recognition the features (eg. AR coefficients or f1 vs f2 from sample data) have to be compared with the related features in pattern space. By considering the discrimination line, which has the same distance from the means of two classes of noise, the distance between the feature from the sampled data and the mean of the feature from each pattern can be used to discriminate between different types of sounds. The closest distance indicates the type of sample data statistically. One of the simplest distance measurements is the Euclidean distance. If the coordinates of the sample feature is *(x, y)* and the means of the feature from source-1 and source-2 are located at, (μ_{1x}, μ_{1y}) and (μ_{2x}, μ_{2y}), respectively, then the sample belongs to source-1 if the following condition satisfied:

$$[(x-\mu_{1x})^2+(y-\mu_{1y})^2]^{0.5} < [(x-\mu_{2x})^2+(y-\mu_{2y})^2]^{0.5} \qquad (4)$$

In equation (4) the Euclidean distances between sample and patterns are compared. This method is implemented to discriminate between speech and, music and transportation noise based on AR coefficients and zero-crossing rate in the autocorrelation function.

The pattern vectors made from the acoustical signature of the environmental sources can be compared with the sample vector made in the same manner. The statistical pattern recognition finds the nearest neighbourhood between the sample vector and pattern vector. If the reference pattern of m types of noise source is represented as a vector [6,7]:

$$P_i = (P_{1i}, P_{2i}, \ldots, P_{ni}) \qquad i=1,2,..m \qquad (5)$$

and the sample pattern of noise character is represented as:

$$S = (s_1, s_2, \ldots, s_n) \qquad (6)$$

then the sample vector and the reference pattern vectors can be compared. To evaluate the similarity between a sample and references, the following set of conditional distances has to be computed:

$$D(S,P_i) = [\sum_{j=1}^{n}(s_j - P_{ji})^2 . W_{ji}]^{0.5} \qquad i=1,2,..m,\ j=1,2,..n \qquad (7)$$

where W_{ji} is the weight of nth parameter for a noise source type and can be defined as the estimated variance of the nth parameter for the ith source. The sample is assigned to a type of noise source using the condition that the distance score is a minimum. The pattern made by the averaged spectrum from 5 types of aircraft have 256 elements per spectrum. The mean and variance of the data can be calculated and compared with the same parameters in sample vectors based on equation (7).

11. FREQUENCY SPECTRAL PATTERNS

A short-time data acquisition cannot provide enough information to discriminate between different types of aircraft and vehicles. The averaged frequency spectrum is a feature for long-time acquisition. An averaged spectrum has been used successfully to identify 5 different types of aircraft (after detection of transportation noise, 5 types of aircraft will be identified when the noise event is finished). The main feature used to identify the type of aircraft is the minimum Euclidean distance between the averaged frequency spectra of the sample and the patterns. The averaged spectrum of noise can be expressed as follows:

$$X_m = x_{m1}, x_{m2}, \ldots x_{mn} \qquad (8)$$

$$Y_j = \frac{\sum_{i=1}^{m} x_{ij}}{m} \qquad i=1,2,..m, \quad j=1,2,..n \qquad (9)$$

if the averaged frequency spectrum array is $Y_j = y_1, y_2, \ldots y_n$, n is the number of data in each spectrum array and m is the number of spectra. Fig. 4(a) illustrates the averaged frequency spectrum pattern made from the overflight duration of aircraft VH-TAD with 512 samples and a 2 KHz sampling frequency.

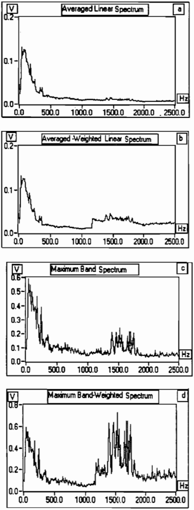

Figure 4. Comparison between different spectral patterns of an aircraft noise, a) A pattern of averaged frequency spectrum, b) A pattern of linear average-weighted frequency spectrum, above 1200 Hz, c) A pattern of linear maximum frequency spectrum, d) A pattern of linear maximum-weighted frequency, above 1200 Hz.

A further indicator of the source of noises, particularly aircraft, is the pattern of the maximum amplitude in each frequency band. This pattern is comparable to the averaged spectrum pattern but the maximum spectrum pattern captures the peaks of short-duration frequency components, which are not observable in the averaged spectrum.

$$Y_j = Max.X_{ij} \quad i = 1,2,...m, \quad j = 1,2,..n \tag{10}$$

The major problem with this feature is its large number of spikes and its instability. We can put emphasis on crucial parts of the averaged or maximum spectrum by weighting the frequencies of interest, eg. Fig. 4(c) shows the growth of frequencies higher than 1200 Hz and Fig. 4(d) shows the multiplications of those frequencies by 3.

12. DECISION MAKING AND AIRCRAFT RECOGNITION

The order of various feature extraction stages and classifiers is crucial in making a right and quick decision about the type of acoustical source. The first thing to know is the correlation between features and sources and realising the stage of decision making based on the extracted feature. The best method is to move from general criteria for a rough classification to specific criteria for a fine classification.

A decision making algorithm based on the discussed features, is presented in Fig. 5. The level of energy in short intervals is measured and compared with the threshold levels of background noise and noise events. When a noise event starts, the system will change the sampling rate and start to measure the zero-crossing rate of time signal, autocorrelation function and the FFT. Then the AR model of the signal is determined. All information from these features is compared with some selected threshold values. After evaluating the type of noise, if it belongs to the transportation noise class, the monitoring system will be activated and then the spectral features will be calculated. They are compared with the patterns in the data base to find whether they are matched to a recognisable aircraft. The acoustical sources to be classified in the present research are: a) Sensing system Off, zero voltage acoustical input, b) Background noise, c) B&K 4230 calibrator, 94 dB, 1 KHz, d) Traffic noise in rainy weather, e) Music and speech, 2 pieces of classic music, 1 piece of rock music and one piece of reading the news by male speaker, f) Transportation noise in dry weather including 5 types of aircraft, a mixture of traffic, a sample of heavy vehicles and a mixture of traffic, train and aircraft noise together.

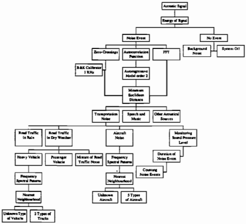

Figure 5. Decision tree for transportation noise recognition and monitoring.

Traffic noise in rainy weather contains a lot of high frequency signals that makes it easily identifiable from the other transportation noise. The high rate of zero-crossing in the time signal is the main characteristic of this signal. The 4230 B&K calibrator gives a constant zero-crossing rate and sound pressure level.

There are two major stages in the algorithm; the first is training and the second is operating. Both stages are adaptive and will be activated only by transportation noise. The main reason that they are made adaptive is that the duration of overflight is different for different aircraft noise events. Therefore, finding the exact start and end points of the event can avoid interferences caused in the subsequent stages due to incorrect data. An example of that is the process of averaging the frequency spectrum which has a crucial role in aircraft noise recognition.

13. INTELLIGENT TRANSPORTATION NOISE MONITORING

A sound pressure level monitoring system can be activated when the traffic noise or aircraft noise dominates the background noise. Such a recording and monitoring system can automatically record the desired noise. The described program is able to calculate the time duration of an event and the maximum and the mean of the sound pressure level during an event. In a quiet environment this program counts the number of vehicles and aircraft with a noise level more than a selected threshold. The results of operating an intelligent noise monitoring system for the noise given in Fig. 1, are presented in Fig. 6. The monitoring system is activated only by transportation noise. When the nature of sound changes or the level of energy and sound falls to less than some particular value, the system stops monitoring.

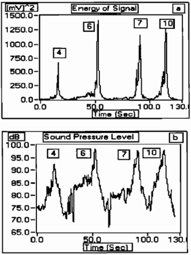

Figure 6. a) The energy monitoring of transportation noise by an intelligent noise monitoring system for the same noise events shown in Fig. 1, b) Sound pressure level monitoring of transportation noise by an intelligent noise monitoring system for the same noise events as in (a).

14. CONCLUSION

In this paper some useful features for discriminating between the transportation noise and other environmental noises, such as music and speech, have been introduced.

The features used to discriminate the transportation noise from the other sounds are the energy of the signal, sound pressure level, linear prediction coefficients, autocorrelation function, peak frequency and zero-crossing. The AR coefficients and the autocorrelation function provide reliable criteria to discriminate between transportation noise, and music and speech in short-time data acquisition. These features have been used successfully to activate an intelligent noise monitoring system. There is also some false recognition when the speech segments are not continuous.

Spectral features for long-time data acquisition are used to recognise the type of aircraft. The result of this research is likely to be applicable to other acoustical noise recognition, such as intelligent noise control [9], counting vehicles, acoustical diagnosis of defects in machines and even to medical diagnosis systems.

Further development of the model is planned. This will include the recognition of other environmental noises such as bird and animal noises, wind and thunderstorm sounds. Techniques will also be developed to recognise sounds when they occur concurrently with other noises in the environment.

ACKNOWLEDGMENT

The work reported here is sponsored by the New South Wales Road and Traffic Authority. Their financial assistance is gratefully acknowledged. We would like to thank Mr. Ali R. Mirhosseini and Mr. C. D. Field for their useful comments and Mr. Ken Stewart for his assistance.

REFERENCE

1. Chen. C. H. "Pattern Recognition Processing In Underwater Acoustics" *Pattern Recognition*, 16 (6), 627-640, (1983).
2. Chen. C. H. "Recognition of Underwater Transient Patterns" *Pattern Recognition*, 18 (6), 485-490, (1985).
3. Rabiner L. R. and Schafer R. W. *Digital Processing of Speech Signals* Prentice-Hall, Signal Processing Series, USA, (1978).
4. Makhoul. John. "Linear Prediction: A Tutorial Review" *Proceedings of IEEE*, 63 (4), 561-580, (1975).
5. Shuzo Saito and Kazuo Nakata *Fundamental of Speech Signal Processing* Academic Press, USA, (1985).
6. Therrien. W. Charles *Decision Estimation and Classification* Wiley & Sons, USA, (1989).
7. Yamada. I, and Yokota. A, Yamamoto. K. "Acoustic Recognition of Aircraft Types in Flight" *J. Acoust. Soc. Japan.* (E) 6 (3), 203-213 (1985).
8. Tjostheim. D. "Recognition of Waveforms Using Autoregressive Feature Extraction, *IEEE Transactions on Computers*, 268-270 (1977).
9. Mohajeri. R. and Fricke. F R., "An Intelligent Window for Minimising Noise Intrusion into Building" *Proceedings of 29th ANZAScA*, 80-87 (1995).