ENVIRONMENTAL NOISE MONITORING - SOURCE IDENTIFICATION

I H Flindell and P Wright

Institute of Sound and Vibration Research, University of Southampton, Southampton, SO9 5NH

1. INTRODUCTION

1.1 The recent Department of the Environment and Welsh Office Planning Policy Guidance: 'Planning and Noise - Consultation Draft' [1] sets out a range of noise exposure categories for different sources. The categories reflect observed differences in community response at the same L_{Aeq}. It is important that noise exposure is determined accurately for each significant contributing source as assessment based on these categories can have significant consequences for planning and development.

1.2 Prediction and measurement
Noise prediction is essential for new sources, but will usually be subject to some degree of error under most circumstances, particularly where there is some uncertainty regarding the characteristics of the source. Existing noise sources can be measured but source identification is then a problem if there is more than one source present. Manual identification is usually only feasible for short term monitoring whereas automatic identification is currently only feasible for long term permanent noise monitoring systems where the microphones can be sited to avoid ambiguity, and where there is independent verification of particular events.

1.3 Automatic source identification
This paper describes progress made under a research project to establish the potential for automatic noise source identification for use in field portable noise measurement systems. Such measurement systems must be capable of breaking down the overall noise environment into the L_{Aeq} contributions made by each significant source to be of any practical use. It is important that the source identification system does not compromise the basic measurement accuracy achievable with a conventional sound level meter and microphone combination. On the other hand, the level of source identification required is limited to generic source types, e.g. aircraft, road traffic, etc. There is no need for identification down to the level of individual vehicles, as there is no practical use for this level of detail in this context. This project was limited to aircraft, road traffic, and railway noise as generic transportation noise sources, with discrimination against other sources of background noise having a high priority.

2 ALTERNATIVE SYSTEMS

2.1 The most convenient type of system for general field use would be based around a conventional sound level meter fitted with a standard measurement microphone and with appropriate signal processing enhancements to achieve source identification. The first stages of this research project [2] investigated the potential for pattern recognition techniques using the acoustic features available from a single microphone channel to achieve reliable identification. There are no measurement grade directional microphones because of the difficulty of constructing a directional microphone with uniform directivity throughout the working frequency range. This means that directional information is not available to a precision grade single microphone system, although it is certainly feasible to use directional microphones to discriminate against fixed noise sources for source identification as opposed to direct measurement purposes. Nevertheless, there is a great deal of temporal and spectral information available, from which it might be possible to extract suitable features for use in simple pattern recognition against a library of pre-determined templates.

2.2 Multitrack recordings

Aircraft, road traffic and railway noise are each built up from a sequence of individual vehicle events. There is usually considerable variability in terms of acoustic features from one event to the next, which might confound successive measurements of a particular source made at different distances or in different orientations with respect to the source. In order to investigate this a number of multitrack recordings using a wide spaced microphone array were made of aircraft, road traffic, and railway noise in carefully selected locations that had as little extraneous noise as possible. These recordings were then analysed in a number of different ways to look for particular acoustic features that could reliably differentiate between the different sources.

2.3 Spectral and temporal overlap

The general conclusion was that there are sufficient temporal and spectral cues available in most permanent noise monitoring situations to allow a single microphone system using template matching pattern recognition to work satisfactorily but only where discrete events are separated in time with a good signal to noise ratio, and where there is some flexibility with respect to microphone siting. On the other hand, the particular features which could be used for source identification varied across different measurement sites, such that a unique set of features would probably have to be adopted for each new site. This is because the overlap of most temporal and spectral features between different sources at different measurement sites is likely to exceed the differences between different sources at the same site. This means that template matching pattern recognition using a single microphone system is only likely to be feasible for permanent monitoring systems where a significant investment in setup time can be justified. It seems that the major noise sources on aircraft, road vehicles and trains have many similarities in terms of acoustic features, whilst different operating conditions and different measurement sites impose relatively large differences.

2.4 Spatial and Directional Information

The next stage was to consider methods of taking spatial or directional information into account. Human listeners with normal hearing can localise sources in three dimensional space, both passively in terms of the spectral and temporal differences of the acoustic signals received at the two ears, and adaptively, when the differences in the signals received at the two ears are compared at different head positions so as to 'home-in' on the source by moving the head where the source direction is otherwise ambiguous. In addition, a human listener generally has a considerable amount of non-acoustic information available to inform source identification decisions. It is unreasonable to expect a machine to perform as well as a human listener with very much less information available. It is also unnecessary for the machine system to perform better than a human listener, as such fine source distinctions would be of no interest for practical assessment purposes.

2.5 Close spaced microphone arrays

Bearing in mind that directional microphones are not suitable for precision measurements, the next stage was to consider the use of microphone arrays to obtain directional or spatial information. Close spaced arrays can provide directional sensitivity over a wide frequency range (depending on spacing) through appropriate signal processing of the combined signals. Adaptive beamforming techniques can be used to steer the directional sensitivity of the array to follow moving sources. There are two main disadvantages to this approach. First, a highly directional array requires a large number of microphones and sophisticated signal processing. Full bandwidth audio signals from each microphone channel must be appropriately processed in real time. This could be very expensive in terms of software development even though the price of digital signal processing hardware is falling rapidly. Second, whereas narrow directional sensitivity would be essential to give good discrimination between different sources, a steerable array would have problems with extended sources such as road traffic noise or railway noise, which are actually composed of many separate individual sources. This means that adaptive beamforming systems are more suitable for specialised applications than for general field use at the present time.

2.6 Wide spaced microphone arrays

Widely spaced arrays can give simple spatial information in terms of the relative levels across different frequency bands at each microphone as a result of differential attenuation from each source. A reference microphone is placed at the desired measurement position, and a number of remote microphones are distributed around the site to inform the identification system of the presence of contributions from each of the major contributing sources at that site. The identification system can then operate on frequency spectrum time histories from each microphone. This represents a considerable reduction in data throughput and brings the necessary signal processing and data storage requirements within the range of portable PCs. A second advantage of this system is that each microphone in turn can be considered as the reference where a blanket survey of a large site is required. For example, consider a site surrounded by main roads with a main line railway crossing underneath the busiest main road at right angles and with a regional airport nearby. The site is hypothetically under consideration for development for a noise sensitive use. Road traffic and railway noise levels will vary a great deal across the site, whereas the contribution from overflying aircraft will be relatively constant in comparison. There are considerable advantages to being able to deploy a multichannel data logging system at such sites, rather than moving around from one position to the next sequentially with conventional techniques.

2.7 Multichannel Data Logging Systems

Considering the above example, conventional technology using a short L_{Aeq} data logging sound level meter such as the Cirrus Research 2.36A can cope, but at a heavy manpower cost. The Cirrus system allows for manual source coding for later decomposition of the short L_{Aeq} time history at each measurement site into separate source contributions. Skilled manual interpolation of the combined data obtained at a number of positions within the site would be capable of providing all the necessary data for land use planning purposes, but cost limits on the amount of data available would limit the accuracy because of sampling variability. The multi-channel system proposed here could use a similar type of data logging sound level meter system at each microphone position (or possibly short range low bit rate radio links back to a central data logging station). The source identification system can then be applied to the combined data set to decompose the L_{Aeq} contributions from each significant source across the site as a whole.

2.8 Validating results

The fundamental difficulty in investigating any scheme for data processing is that of validating results. Whilst it is possible to demonstrate the practical validity of any technique for automatic source identification against real-life field data by simultaneous observation, it is not possible to demonstrate that the system is capable of correctly decomposing a complex noise environment into the L_{Aeq} contributions from each significant source as there is no independent method of determining these contributions in the field. Since this is the real objective of this type of identification system it is necessary to use simulation techniques to provide noise level time series data for experimental systems. A simulated time series must be generated to represent an actual time series that would be obtained at a real array of reference and remote microphones deployed around a site with multiple contributing noise sources, by combining actual time series data obtained with wide spaced arrays deployed at sites where there is only one major noise source. Thus the separate source contributions at the reference microphone can be precisely determined in advance.

3. STATIONARY SOURCES AND LINEAR REGRESSION

3.1 Linear regression techniques can resolve the contribution made by stationary sources at the reference microphone using data from remote microphones placed near to the stationary source. In this case, there are some advantages to using a directional microphone to discriminate against other noise sources at the remote microphone position, provided that the data at that position is not required for primary (reference) measurements. Linear regression of noise level time series data can then estimate the attenuation constant, over different frequency bands if necessary, for noise propagation from the stationary source between the remote and reference microphones. The regression calculation must be updated continuously by moving the regression time window along the data to take short term changes in propagation characteristics into account.

3.2 A number of trials were carried out using this technique, which showed that linear regression can precisely determine the actual contributions at the reference microphone due to remotely identified stationary sources, even under conditions with intermittent high background noise level and other disturbances.

5. MOVING SOURCES AND LINEAR REGRESSION

4.1 The next step was to investigate the applicability of the linear regression technique to moving sources. The problem here is that there is no longer a linear relationship between the noise level time series contributions due to the moving source at the remote and reference microphone positions. The only practical solution to this problem is to increase the sampling duration to encompass the entire noise source pass-by event, so as to effectively regress a time series of SELs (Single Event L_{Aeq} referenced to 1 second duration) due to separate events. It was envisaged that the variability of SEL attenuation from the remote to the reference microphone over individual events could then be treated as a random error which could then be taken into account by an appropriate sample size. A number of spreadsheet simulation exercises were carried out which showed that the assumed event duration was a crucial variable. Unfortunately, a short assumed event duration reduces the linear correlation between the level changes at each microphone, while a long assumed event duration causes the regression to effectively ignore the contribution from the events in question, producing a higher correlation but no useful attenuation constants. The only solution to this problem would be to identify the effective event duration at the reference microphone accurately for each event, but this is tautologous and therefore unworkable. This means, effectively, that there is no unique set of acoustic features which can be used to identify the different moving transportation noise sources against other background noise sources, even when spatial factors are taken into account.

5. SOURCE CODING

5.1 Manual source coding

This apparent impasse was only overcome by returning to a consideration of manual source coding, as in the Cirrus 2.36A system. This provides a workable technique for the decomposition of separate source contributions in the field, but it requires a skilled operator to be present to interpret all the acoustic features and non-acoustic cues present at any particular site, and there is no independent or objective verification of the operator's judgements. What is needed is a technique which can take the adaptability of manual source coding into account but then be left to continue for extended periods without the operator present. Automatic operation should then increase the reliability of long term measurements and allow for greatly increased measurement sample durations, without increasing labour costs. One way of achieving automatic operation is to deploy a neural network pattern recognition system which relies on an initial period of manual coding. Because the training data is obtained on site the particular site characteristics are taken into account.

5.2 Neural networks

Neural networks operate as pattern recognisers by calculating an appropriately weighted function of the input variables to determine an output quantity which represents the network classification of the input pattern. The key point is that the various weighting constants are adapted during the network training phase to provide that function of the input variables that achieves optimum pattern recognition performance against the training data. A neural network can represent non-linear relationships and can take correlated input variables into account. It is still necessary to select input variables which provide information which is useful for classification as input variables which do not contribute to the classification tend to reduce the overall performance. Two commercially written neural network packages for the PC were used. Whilst a commercial package greatly simplifies the application of neural network techniques by eliminating much of the learning curve it remains difficult to determine which input variables are actually contributing useful information, except by a process of trial and error. In this case, noise level time series data, broken down into octave and one third octave bands at both reference and remote microphone channels were used as inputs to the networks, together with overall A and L-weighted levels.

5.3 Simulations

The first investigations were carried out using data derived from separate train and road traffic noise recordings. These were fed into a spreadsheet and calculations were made to simulate the combined data at each microphone channel and to determine the overall and source contribution L_{Aeq}s at the reference microphone. The reference microphone data was coded at a preset 'coding threshold' designed to simulate manual coding by the operator. The coding threshold is defined as a railway to road traffic instantaneous (one second average) signal to noise ratio. This was set to a level above which railway noise was assumed dominant and below which road traffic noise was assumed dominant. The first half of the data was used to train a network and the second half presented to the network as test data. The signal to noise ratio was varied and the trained network was found to be capable of determining the railway noise contribution of the test data to within 1 dB even down to negative railway to road traffic L_{Aeq} signal to noise ratios. This promising result led to further investigation of the neural network approach.

5.4 Further Investigations

There are two main areas requiring further investigation of the neural network approach.

- The optimum operating characteristics in terms of input variables, signal to noise ratio, coding threshold, and number of training data samples must be determined under controlled test conditions using simulations of combined noise environments synthesised from separate source recordings.

- It is important to determine the levels of performance obtainable in the field under practical conditions where the training data is subject to the variability of manual source coding. Controlled trials of the variability of manual coding data obtained in the field are necessary, but in addition, it may also be desirable to carry out controlled manual coding trials using simulated test recordings in the laboratory to provide an absolute check on manual coding accuracy.

Further considerations of practicality for field use with less highly skilled operators can be left until the basic feasibility of the technique has been proved (or not).

5.5 Input variables

A wide range of input variables must be made available for the network to take advantage of those particular combinations of input variables which give the best source discrimination at a particular measurement site. On the other hand, it is important to avoid input variables which do not contribute at any measurement site, as these variables cannot aid overall performance but will increase the amount of computation required, and may in fact degrade performance. This problem can only be investigated on a trial and error basis for the present, but may require detailed investigation at some future stage. A continuous time series of one third octave band levels at one second intervals appears to be useful, but it is not known which overall frequency band limits should be set. Applying the network on a second by second basis to this frequency spectrum data does not take the recent time history of the data into account, whereas this probably also contributes useful discrimination. For example measures of rising or falling noise levels obtained by comparing a 't' second average level against an earlier or later 't' second average level could potentially assist the network to discriminate the beginning and end of particular noise events. It remains to be seen whether such time domain type measures can make a worthwhile contribution in the general case.

5.6 Signal to noise ratio

Signal to noise ratio in terms of the ratio of the separate source L_{Aeq}s is important, but this is intimately linked to the number of discriminable events per hour, and the maximum instantaneous levels of those events. In practice, the full identification system is only likely to be required within a relatively narrow range (perhaps plus or minus 10 to 12 dB) of separate source L_{Aeq} ratios as one source will be clearly dominant outside of this range. Nevertheless, source identification may still be desirable in the case of irregular and isolated events which are significant in themselves without significantly affecting the overall L_{Aeq}. Network performance is likely to be good in this situation, provided that the problem of sampling a sufficient number of events which have been manually coded for training data can be overcome.

5.7 Manual coding strategy

Typical manual coding thresholds depend on the adopted coding strategy. Most practical situations where the system would be deployed will have a more or less steady road traffic background noise, with regular railway or aircraft noise events, and occasional extraneous noises which must be discriminated against. In such cases, the operator could code railway or aircraft events whenever audible, or whenever judged to be making a significant contribution to the instantaneous overall level. The coded event L_{Aeq} contribution is then estimated by counting the coded samples as event plus road traffic noise, and then subtracting an assumed constant road traffic noise level, derived from the non-coded portions of the noise level time series, or by counting the coded samples as due to the event only, depending on the coding strategy. The first strategy (coding when audible) is likely to be more accurate, as audibility is probably easier to judge than 'contributing significantly' (whatever that means), but the data will probably be less clear cut for the purpose of training a neural network. The best compromise here can only be determined by a combination of practical field trials and simulations, bearing in mind that there is no independent measure of the 'correct' answer for real field data, as discussed above.

5.8 Measurement Errors

Potential measurement errors arise from inconsistency in coding and network classification errors. However there is also the fundamental error determined by the coding threshold generated in relation to the time series profile and L_{Aeq} signal to noise ratio. This error theoretically defines the upper limit of the performance of a network based system trained using manual coding. A spreadsheet simulation using real railway and road traffic level time series data was performed to estimate this fundamental error. L_{Aeq} signal to noise ratios were set in 5dB steps between +20dB and -20dB and were each tested over the range of coding thresholds between -1dB and -20dB. The results demonstrated the interaction between coding threshold and signal to noise ratio and were encouraging in that they predicted low errors. They showed that a coding threshold of -5dB would be optimal down to signal to noise ratios of -5dB for the particular noise level time series investigated. Further investigation is necessary to determine likely manual coding thresholds (and their variability) and to confirm that there are no unexpected problems with different types of noise event level time series profiles.

6. FURTHER WORK

6.1 Work is at present continuing in two main areas as follows;

- The main purpose of the simulations at present is to get a clearer idea of the types and amount of input information required for adequate network performance in different measurement conditions. Further simulations to evaluate subjective audibility (and hence provide realistic coding thresholds) are in preparation.
- A field measurement programme is underway to collect data at representative sites which are subject to two or more significant transportation noise sources and which could hypothetically be subject to consideration under the planning and noise policy guidance. This programme has two purposes, first, to assemble coded multimicrophone data that can be used to test the input information strategies suggested by simulation; and second, to build up a library of case studies for use in future reference materials setting out the technique for practical field use.

7. ACKNOWLEDGEMENT

This work was carried out at ISVR under contract to the Building Research Establishment as part of the DoE environmental noise research programme.

8. REFERENCES

[1] Department of the Environment and Welsh Office - Planning Policy Guidance: 'Planning and Noise - Consultation Draft', January 1992

[2] P Wright & I H Flindell, 'Automatic Noise Source Identification', Proceedings of the Institute of Acoustics, Vol. 13: Part 4, pp 25 - 32, September 1991.

Figure 1 Stationary sources and linear regression

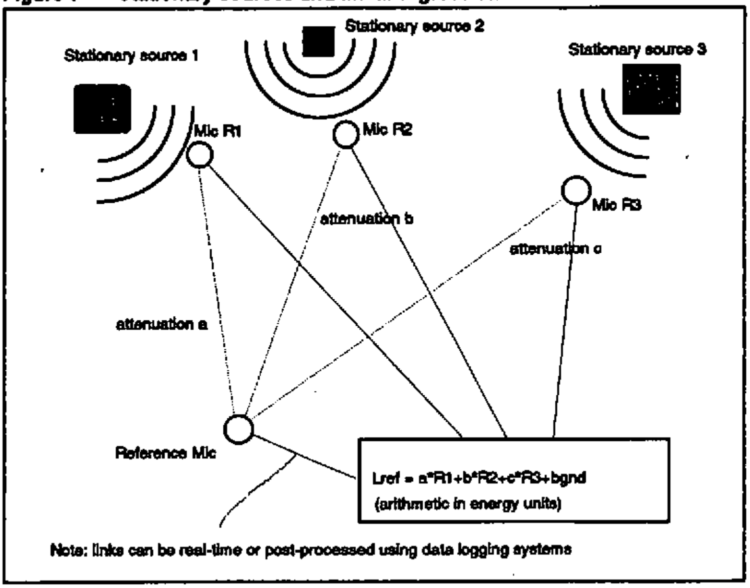

Figure 2 Neural network system