Environmental Sound Recognition: A Statistical Approach

Christophe Couvreur[1]

Introduction

Noise pollution has become an important problem in our society. Noise assessment regulations require the measurement and evaluation of noise. The basic equipment for this measurement is the noise monitoring systems (or NMS). The current generation of NMS consists of "dumb" systems that record noise level and, possibly, noise spectra, over programmed periods of time. The interpretation of the recordings is left to a human expert acting "off-line." The goal of our research is to develop the tools necessary to build an "intelligent" NMS able to identify automatically the nature of the sources of the noise events in addition to recording their acoustic characteristics (Figure 1).

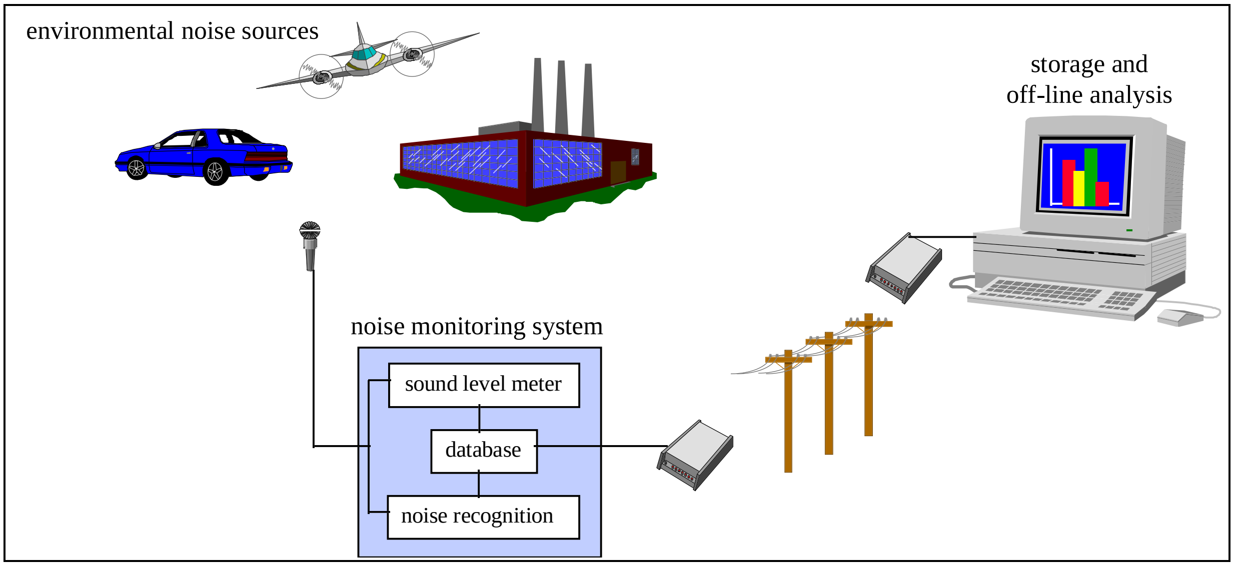

Figure 1: "Intelligent" Noise Monitoring System.

In our approach, the desired noise recognition capabilities are obtained by combining statistical pattern recognition methods with statistical signal processing techniques. The proposed statistical framework can then be used to address the problem of environmental sound recognition at various levels of complexity and to develop a series of solutions presenting various trade-offs between performance and computational requirements. This communication presents a general overview of our research. The reader interested in technical details is referred to the bibliography where some of the publications containing our findings are listed.

Noise Event Recognition by Statistical Classifiers

We started by considering the following notional noise recognition problem.Given the spectrum of a noise obtained by a third-octave spectral analyzer, is it possible to identify its source among a known set of possible sources? We chose to look for a solution in the

[1] Aspirant F.N.R.S., Faculté Polytechnique de Mons (F.P.Ms) , Service de Physique Générale et Service de Théorie des Circuits et de Traitement du Signal, Rue de Houdain 9, 7000 Mons, Tel. +32/(0)65/37.40.42 ; Fax : +32/(0)65/37.40.45 ; E-mail : couvreur@thor.fpms.ac.be

statistical paradigm of pattern recognition [1]. A database of samples of environmental sound recordings including car, truck and airplane noises has been built. This database has then been used to design various types of statistical classifiers including linear and quadratic discriminant classifiers, Gaussian Bayesian classifiers, and nearest-neighbors classifiers. The classifiers were trained on sets of samples from the database manually labeled. Their performance was evaluated on other sets of manually labeled samples from the database. The results obtained showed a correct recognition rate of as much as 94% for Gaussian classifiers using properly conditioned spectra as their inputs. Furthermore, since Gaussian classifiers have modest computational and storage requirements and therefore could be readily implemented inside the current generation of NMS with the available embedded microprocessor, they appear a potentially good candidate for a noise recognition sub-system in a NMS.

Adaptable and Adaptive Classifiers

One of the major drawbacks of classifies designed using the approach described above is their sensibility to variations in utilization conditions. For example, if the NMS is moved, the distance between its microphone and the noise sources changes and the spectra presented to the classifier undergo some frequency dependent attenuation. Of course, retraining the classifier for the new situation is possible, but it is costly since it requires the consistitution of a new database of labeled training noise samples. As an alternative solution, we developed a version of the Gaussian classifier for noise sources that adapts to changes in spectra. The modified classifier can either be manually adapted by a human expert using *a priori* knowledge on the modification occurring to the noise sources (i.e., it is an *adaptable* classifier) or it can adapt itself automatically without human supervision (i.e., it is also an *adaptive* classifier) [2]. The adaptive capabilities are obtained by combining the EM algorithm for mixtures of densities of statistics with a physical model of acoustic variations [3]. Experimental results show that the performance of the adaptable and adaptive Gaussian classifier is preserved under variations of utilization conditions without the need for retraining. Moreover, even if its computational requirements are higher than that of the "plain" Gaussian classifier, the adaptive Gaussian classifier remains reasonably easy to implement.

Hidden Markov Model Based Classifiers

While simple classifiers based on the average spectra of the noise sources (adaptive or not) can already offer good recognition accuracy they do not take advantage of the fact that, for most environmental noise sources, the spectra are evolving in time in a structured fashion, i.e., the noises are not stationary. Taking into account the time-frequency structure of the noise events can increase the discriminative power of noise classifiers at the cost of an increase in complexity. The dynamic evolution aspect of the spectra of noise sources is introduced in our statistical framework by using hidden Markov models (HMMs). Hidden Markov modelling has been the most successful recognition technique in speech recognition for nearly two decades. It is thus reasonable to expect that it will also offer good results for the noise recognition problem. An experimental study of the performance of HMMs for noise recognition is currently under way. It must also be noted that the adaptation techniques that have been developed for Gaussian classifiers can also be applied to HMM-based classifiers.

Decomposition and Classification of Mixtures of Signals

All the classifiers that have been presented so far are meant to classify environmental sounds occuring separately. However, in practice, multiple sound events are likely to happen simultaneously. The concept of mixture of ARMA processes [4] has been introduced in an attempt to develop a classifier for multiple simultaneous signals.

A mixture of ARMA processes is simply a linear combination of stationary ARMA processes taken from a dictionary of known ARMA models. The mixture decomposition and classification problem can then be formulated as follows. Given a sample of the mixture signal, find the ARMA components from the dictionary that are present in the signal and their proportions. The number of components and their proportions are *a priori* unknown. We proposed an optimal solution to the mixture decomposition problem based on the minimum description length (MDL) principle of information theory. An algorithm for the efficient computation of this solution has also been presented. The algorithm is based on a combination of the EM algorithm with Wiener or Kalman filters and original results on "greedy" algorithms for subset selection. It has been tested on mixtures of stationary noise sources and shown to yield indeed the desired results.

In order to remove the stationarity assumption necessary for the utilization of the mixture of ARMA models, we have also extended the theory of hidden Markov models to accomodate mixtures of signals [5]. The computational requirement of the current versions of the algorithms does not allow them to be tested on realistic data though.

Estimation of the Speed of Moving Sources From Doppler Shift

In addition to noise sources classification, it would also be desirable for the NMS to be able to estimate some characteristics of the noise sources. For example, the speed of the cars or trucks passing by the microphone could be estimated from their acoustic signatures. Such moving noise sources will be affected by the Doppler effect. An original stochastic process model for Doppler-shifted noise sources has thus been developed which can be used to estimate the speed of a moving noise source from measurements made at a single fixed microphone [6][7].

References

[1] C. Couvreur & Y. Bresler, "A Statistical Framework for Automatic Classification of Environmental Noise Sources," Submitted to Int. J. of Noise Control Eng., 1996.

[2] C. Couvreur & P. Chapelle, "Utilization of Expert Knowledge in Automatic Classifiers of Noise Sources," Proc. INTERNOISE’96, Liverpool, England, Jul. 30-Aug. 2, 1996, pp. 2855-2858.

[3] C. Couvreur & P. Chapelle, "Un classificateur adaptatif pour sources de bruit environnementales,", Actes CFA'97, Marseille, France, Apr. 14-18, 1997. To appear.

[4] C. Couvreur & Y. Bresler, "Optimal Decomposition and Classification of Linear Mixtures of ARMA Processes," Submitted to IEEE Trans. on Signal Processing.

[5] C. Couvreur, "Hidden Markov Models and Their Mixtures," DEA Report, Dep. of Mathematics, Université catholique de Louvain, aug. 1996.

[6] C. Couvreur & Y. Bresler, "Doppler-Based Movement Estimation for Wide Band Sources From Single Sensor Measurement, " Proc. ICASSP’97, Munich, Germany, April 21-24, 1997. To appear.

[7] C. Couvreur & Y. Bresler, "Modeling and Estimation of Doppler-shifted Gaussian Random Processes, " Proc. IEEE Workshop SSAP'96, Corfu, Greece, June 24-26,1996, pp. 28-31.