Automatic Noise Source Recognition

Didier Dufournet*, Philippe Jouenne* and Adam Rozwadowski**

* 01dB, 69100 Villeurbanne, France and **01dB inc., Skaneateles, NY, 13152, USA

Abstract: Madras is an intelligent noise monitoring system able to classify the acoustic signatures of different kinds of noise sources. The purpose of this paper is to present the use and the adaptation of several methods of signal processing and artificial intelligence to acoustic signal characterization and specially the “meta-class” recognition step. The MADRAS system works with Symphonie, a PC-based dual-channel real time analyzer developed by 01dB.

INTRODUCTION

The MADRAS European project (1) deals about the development of a new generation of acoustics instruments able to automatically identify in real time various acoustic sources and to assess their impact effect. The automatic recognition process uses different methods of signal and shape recognition processing, organized in an treelike classifier in time and spectral domains.

In this paper we present the global architecture of the MADRAS process and its real-time implantation on the 01dB PC-based two-channels analyzer Symphonie. We focus on a important step of the classifying task which consists to choose the general class of the noise, before applying a specific treatment.

A FIVE STEPS ANALYSIS

The MADRAS functional architecture has been adopted to consider various constraints linked to the realization of a light and autonomous system in a moving environment: decreasing calculation charge by pertinent audio recording, adaptation capacities in-situ, integration of moving measurements conditions and separation of decisions in time and frequency/scale domains. The result in a five steps process, integrating detection and post-treatment phases (fig.1), as described below:

Step 1 : Detection
The purpose is to produce some pertinent audio files, according to some statistical criteria of energy. A detector – 01dB software dBTrig- records audio events which the Leq A 125ms (a good nuisance indicator) exceeds a statistical and adaptive threshold based on L90+20dB (it allows robustness and adaptation to background noise). The audio files are transmitted to MADRAS software to be recognized and archived.

Step 2 : Segmentation
This step allows the MADRAS process to identify with accuracy the temporal limits of interesting segments within the audio recording. Methods from mathematical morphology allows to access to local audio patterns and global tendency (1)(2). The result is a list of audio events to treat in parallel way by steps 3 to 5.

Step 4: Specialists
After Broad Classification (see next section), each signal selected as interesting will be analyzed by the appropriate specialist among a specific strategy.
Pass-by of isolated vehicle: a neural network (classical multi-layer perceptron) learned more than thousand patterns of third-octave spectra to distinguish cars, motorcycles, motorbikes, trucks and heavyweight.

Stationary noise: an in situ-learning protocol based on a third-octave L90 signature is achieved. The user will select interesting bands, averaging different measurements and define a tolerance range to construct a template for each stationary source. The final templates are written in a dictionary to be reusable.

Impulsive noise: as for stationary noises, an in-situ protocol is defined, using a wavelet decomposition (3) to construct the pattern. In both cases, less than ten learning-examples are sufficient.

Step 5: Post-processing
The measurement file with encoded sources is post-treated by software dB Trait which computes individual contribution of each source and applies specific regulation.

BROAD CLASSIFICATION

This third step will classify the different segments found in segmentation process into time-patterns allowing to choose the correct specialist to apply in frequency/scale domain. According to experiments and expert knowledge, a shape classification is realized. Five geometrical criteria are defined on Leq 50ms acoustic signature: duration (sec.), dynamic range (dB), rising slope (dB/sec.), surface ratio (between Leq signal and a constant max valued same length signal ) and activity (number of crosses by the average value). Each signature is defined in a 5 five-dimensions space, that we reduce using a neural non-linear version of Principal Component Analysis

The learning process consists to present the normalized geometrical values of several patterns both on the input and output layers of the clustering memory (fig.2). The network will reduce the quadratic error between output and input compressing data on the hidden layer on only two neurons. It will define the decision representing space. At the end of the learning process, the different weights \( v1 \) and \( v2 \) reveal the contribution of each variable to the new axis (fig.3a). The outputs of the hidden layer \( X_1 \) and \( X_2 \) for each learning pattern reveal its disposition in the new space (fig.3b).

![FIGURE 2. Clustering memory](image)

The natural separation among two axis (a power axis -slope versus activity- and a shape axis –dynamic versus surface ratio-) of different pattern families is satisfying. Thanks to a quadratic boundary separation, we define five areas in the plan: impulsive noises, stationary signals, pass-by of vehicles, pass-by of heavy carriers (planes and trains) and out of these area, energy blasts (e.g. shouts). The real-time recognition process consists now simply to measure the geometrical parameters of the candidate and to leave the network choosing the good ellipse.

The results are encouraging (success ratio around 80% for specialists and broad classification). But we must improve the detection step using spectral shapes to assure robustness to changes in measurements conditions.

REFERENCES