19th IMEKO TC 4 Symposium and 17th IWADC Workshop
Advances in Instrumentation and Sensors Interoperability
July 18-19, 2013, Barcelona, Spain

Taking advantage of machine learning and pattern recognition in acoustic measurements

C. Asensio[1], M. Ruíz[1], M. Recuero[1]

[1] Universidad Politécnica de Madrid (I2A2 Research Group), 28031, EUITT – Ctra. de Valencia km 7, + 34 91 336 7807, casensio@i2a2.upm.es

Abstract- Traditionally, acoustics measurement researchers and developers have focused on: a) the development of novel instruments and sensors, trying to improve their accuracy or trying to fix specific problems deriving from specific applications; b) the definition of uncertainty estimation methodologies, either in instruments and testing methods, c) the definition and customization of new measurement indexes.

In very recent years, the research lines have been changing, in an attempt to take advantage of new signal processing techniques and solve classical handicaps in acoustic measurements. This is the case, for example, of blind source separation or beamforming.

Following this trend, this paper focuses on new acoustic measurement approaches deriving from the application of machine learning and pattern recognition techniques for the development of smart instruments based on the measurement of sound pressure level.

I. Introduction

Speech recognition has been one of the leading research fields throughout the last two decades, achieving great results that have boosted its implementation in real world applications (voice user interfaces, call routing, speech-to-text converters, etc).

Despite their closeness, speech recognition techniques have not been naturally imported to other adjacent fields of acoustics, where sound pressure level continues to be the most widely extended index for describing the strength of a sound or a noise. This index is the basis for calculating many other indexes for describing environmental noise, acoustic emissions of appliances, sound insulation of walls, acoustic performance of rooms… Acoustic analysers, commonly used in these fields, perform essentially as sound level meters for standardized frequency bandwidths (typically octave or third-octave bandwidths). They are dumb instruments requiring user intervention, as only time-level threshold techniques can be applied for the triggering or the detection of the sound events to be measured. Nevertheless, pattern recognition, or, in general, machine learning techniques, can enable new opportunities in the measurement and exploitation of acoustic quantities and audio signals.

In a traditional approach, pattern recognition can be used in standalone classification and control systems based on sound and vibration signals feature extraction, as, for example: i) the classification of the boiling water stages to be implemented in a cooking appliance control system [1]; ii) the acoustic-based classification of the state of the road surface [2-5].

Beyond these approaches, in which acoustic measurements are the means to get a classification, it is possible to use machine learning as a means of controlling the acoustic measurements, converting the sound analyser into a smart instrument in which unattended acoustic measurements will improve its reliance, precision and accuracy.

The objective of the research presented in this paper, was to improve the performance of a noise monitoring unit by adding pattern recognition capabilities to detect sound events by analysing the sound time-frequency characteristic. With this approach, any sounds caused by a specific and previously specified source can be detected among other sounds, triggering measurements or recordings.

The methodology has been satisfactorily applied in aircraft noise monitoring [6], reducing both, the detection, and identification uncertainties of noise assessment. It has also shown a good performance for the measurement of construction site noise in an environment where railway noise polluted the measurement. It would also be very useful for the measurement of residual (background) noise, or for façade sound insulation tests, in which the insulation is estimated by the difference between the interior and outdoor noise at aircraft flyover sound events.

II. Methodology

The sound produced by aircraft is one of those that the human ear can recognize almost everywhere, no matter how long the duration of the sound, in interior or exterior locations, just because the audio signal contains knowledge wide enough to distinguish this kind of sound from others. Under this hypothesis, the method proposed in this paper is based on the similarity between the input sound and the noise produced by aircraft. This similarity has been estimated by applying statistical pattern recognition techniques, in the way described in Figure 1, in order to obtain the Aircraft noise likeness (ANL) estimator.

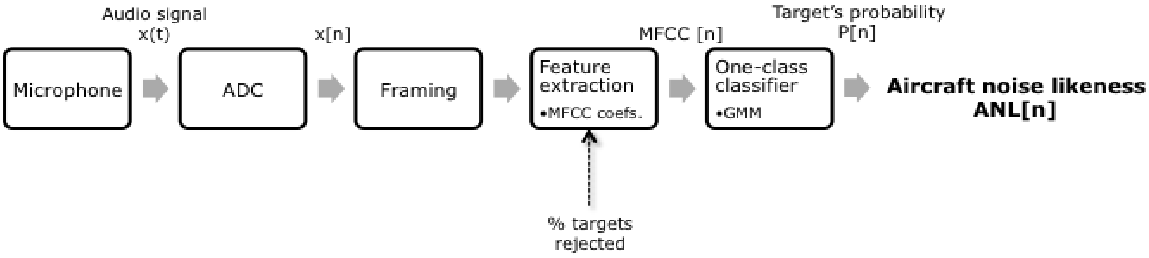

Figure 1. ANL estimation scheme

The analogue signal acquired at the microphone is converted to digital, and windowed in 100 ms duration frames. For every frame, thirteen Mel frequency cepstral coefficients (MFCC) are obtained, and normalized (see Figure 2). These data are used as features in the pattern recognition process. A one-class mixture of Gaussian classifier estimates the probability that an aircraft produced each audio frame. This probability is interpreted as the ANL index that tracks the similarity between the input sound and generic aircraft sounds (where ANL equals 1 means pure aircraft noise, while ANL equals 0 means no similarity to aircraft sound).

Figure 2. MFCC coefficients calculation using an audio frame.

The MFCC coefficients have previously shown a good performance in the recognition of sound sources. In this case, in order to generalize the recognition process with respect to the sound intensity (dependent on the distance from microphone to aircraft), the coefficients have been normalized to obtain nMFCC, which have shown a good separation among the classes (as shown in Figure 3).

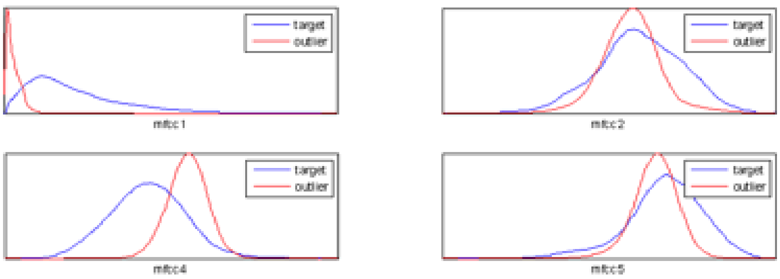

Figure 3. Probability density functions for some of the features (target means "aircraft"and outlier "non aircraft")

A. The one-class approach

The classifier is in charge of estimating the probability that an aircraft produced the input sounds. It is created by defining a statistical model for each class to be recognized, so that, for a given input (features vector) the model estimates the *a posteriori* class probability. In this application, two classes have been defined: aircraft and non-aircraft. The aircraft class can be reasonably sampled for a specific location by just trying to make recordings for most of the aircraft models, flight paths and meteorological conditions, which are the main conditions affecting aircraft sound emissions. On the other hand, it is very difficult to describe the non-aircraft class in a random location, as it consists of different noise sources like road traffic, railways, industries, people… and it is always possible to find new sounds for which the classifier was not trained. In order to fix this problem, a one-class approach has been used in this research. In this approach, it is necessary to define a target class, which is the aircraft class. Then, the statistical model fits the target class, and can be tightened by pre-establishing the targets rate that the will be misclassified. If we have samples available from the outlier class (non-aircraft), they can also be used to tighten the model, and optimize performance.

Using the Matlab PRTools toolkit, a mixture of Gaussian one-class classifier was used to model the target class:

$$\begin{aligned} f(x) = \sum_{i=1}^{Kt} P_i \cdot \exp\left(-(x-\mu_i)^T \cdot \Sigma_i^{\ -1} \cdot (x-\mu_i)\right) - \\ - P_* \cdot \exp\left(-(x-\mu)^T \cdot \Sigma_*^{\ -1} \cdot (x-\mu)\right) - \\ - \sum_{j=1}^{Ko} P_j \cdot \exp\left(-(x-\mu_j)^T \cdot \Sigma_j^{\ -1} \cdot (x-\mu_j)\right) \end{aligned} \quad (1)$$

where the first line stands for the target class distribution, the third stands for the outliers distribution, and the second line stands for a 'background outlier'.

When we model the target class using 20 Gaussian functions, and we set a 10% target rejection rate we achieve the best performance. In a practical case, this means that in an aircraft noise event having a duration of 10 s, one of the ten noise measurements performed by a noise monitoring unit will be rejected: the one that is most polluted by background noise, having the lowest intensity and being absolutely insignificant in the measurement results.

B. The aircraft sound events extraction

To improve the performance, the sequence of classifier outputs is smoothed to get a smoothed version of ANL (defined as soft ANL). This process adds a kind of structural analysis to the classification process removing sporadic classification data, and taking advantage of the structure of the physical process (the aircraft flyover takes some time, and the sound remains for several seconds). The soft ANL index can be used for automatically controlling the measurement of aircraft sound events, marking events and residual noise periods, triggering recordings, etc. The detection process is described in Figure 4, while Figure 5 shows an example of the output obtained using ANL and soft ANL.

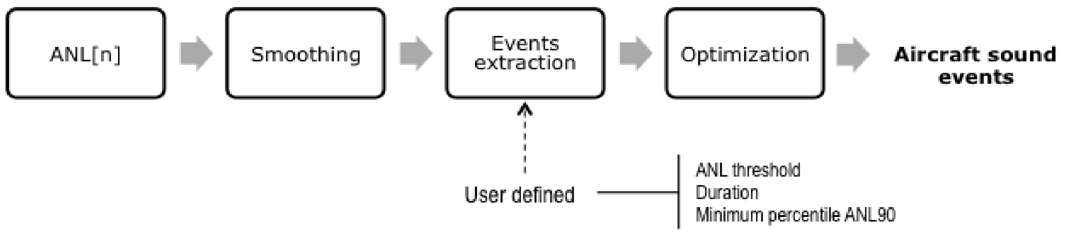

Figure 4. Aircraft sound events extraction

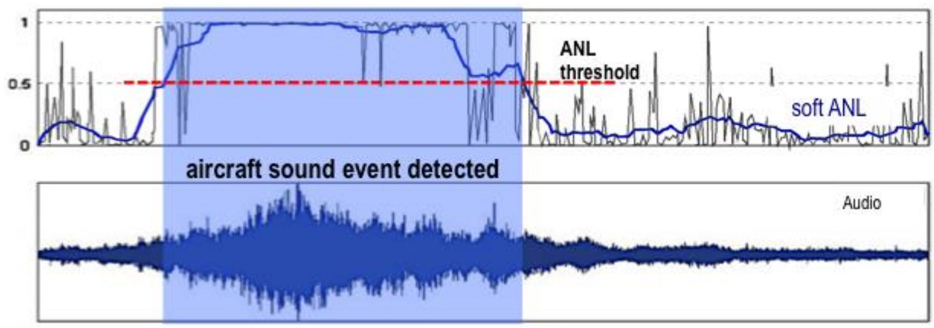

Figure 5. Classification outputs for an aircraft sound event

II. Results

C. Identification results

More than 1000 aircraft and 1700 non-aircraft sound events were manually labelled from the source recordings, taken at 30 different outdoor locations, for the training and testing stages. The system has shown a great performance in the classification and identification of previously detected events. In the case of 'pure aircraft' sound, the error was 2.3% for the selected setup. Table 1 shows the error rates for different sound sources, using two different durations (d=5 s an d=10 s)) for a fixed ANLS threshold (ANL threshold=0.6)

Table 1. Error rates for different sound sources (ANL threshold=0.6)

	Error			**Error**	
Sound source	**d=5 s**	**d=10 s**	**Sound source**	**d=5 s**	**d=10 s**
Pure aircraft	2.3%	5.9%	Emergency vehicles	0.0%	0.0%
Polluted aircraft	23.4%	33.9%	People	2.9%	0.0%
Car	0.0%	0.0%	Animals	1.8%	0.0%
Bus	0.0%	0.0%	Industry	0.0%	0.0%
Truck	1.0%	0.5%	Machinery	0.0%	0.0%
Motorcycle	1.8%	1.8%	Other	0.0%	0.0%
Train	2.5%	0.6%			

D. Detection results

The ANL index has shown a very good performance as a detection tool. The system has proved to recognize over 90% of aircraft sound fragments in which the signal-to-noise ratio is over 10 dB (Figure 6).

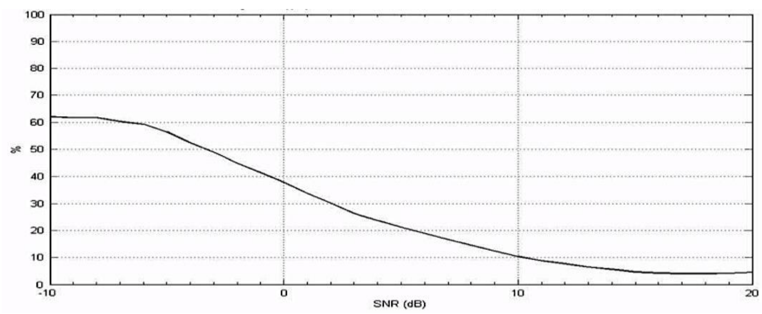

Figure 6. Target measurements error rates as a function of SNR

The good performance in the short-term measurements results in a full-event detection rate of over 90% when the signal-to-noise ratio of the event is over 8% (linked to a range of the event of about 6 dB). Figure 7 shows the classification rates as a function of the signal-to-noise ratio (SNR).

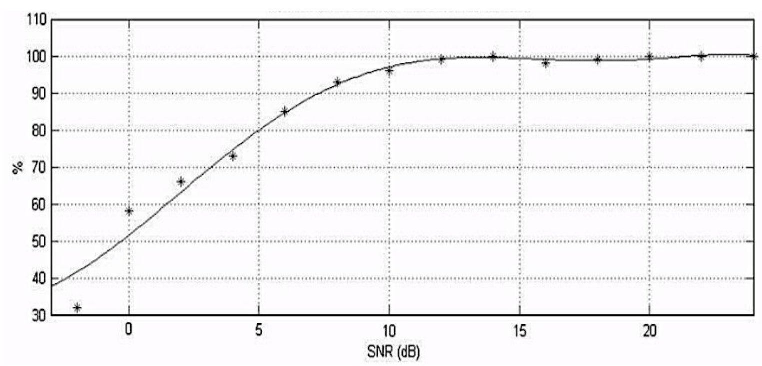

Figure 7. Aircraft events detected as a function of events' SNR

II. Conclusions

A close link can be established between pattern recognition and acoustic measurements. On the one hand, pattern recognition needs features to be measured for the classification purpose. On the other, it is possible to take advantage of machine learning capabilities to create smart instruments. For instance, to control unattended aircraft or background noise monitors, to perform sound insulation measurements using aircraft sound events (ISO 150-5), and to detect thrust reverse activation on aircraft landings [7].
The aircraft noise likeness detector has shown a good performance for the classification and identification of previous detected sound events, and also for the detection itself. The ANL detector does not use sound level measurements for the detection (like traditional monitoring units), but the audio signal itself. This way, the detection is carried out by similarity analysis. Non-aircraft sound events will not be detected, even in the event that their duration and level are high, or in the event that a “quiet aircraft” is nearby. On the other hand, aircraft events can be quite efficiently detected even in highly polluted acoustic environments. The replacement of sound level detector algorithms by those based on ANL would seriously reduce both the identification and detection uncertainty contributions in a noise monitoring unit, especially in polluted environments.
This approach and the designed system can be easily adapted to other usages. Just a new training of the model will customize the tool for specific purposes, like recognition at specific locations or recognition of new target sources. Furthermore, the detector can work either for the detection of targets, or non-targets, which provides a good performance of the background noise measurement.

References

[1] Tabacchi, M.; Asensio, C.; Pavón, I.; Recuero, M., Mir, J.; Artal, M.C.; A statistic pattern recognition approach for the classification of cooking stages. The boiling water case. Applied Acoustics, 2013.
[2] Kongrattanaprasert W, Nomura H, Kamakura T, Ueda K. Detection of road surface conditions using tire noise from vehicles. IEEJ Transactions on Industry Applications 2009;129(7):761-7.
[3] Kongrattanaprasert W, Nomura H, Kamakura T, Ueda K. Automatic detection of road surface conditions using tire noise from vehicles. The Institute of Electronics, Information and Communication Engineers 2009;108(411):55-60.
[4] Kongrattanaprasert W, Nomura H, Kamakura T, Ueda K. Detection of road surface states from tire noise using neural network analysis. IEEJ Transactions on Industry Applications 2010;130(7):920-5.
[5] Alonso, J.; López, J.M.; Recuero, M.; Pavón, I.; Asensio, C.; Arcas, G. On-board Road Surface climate condition evaluation using tyre/road noise and Support Vector Machines. Applied Acoustics, 2013 (under review).
[6] Asensio, C.; Ruíz, M.; Recuero, M.; Real time aircraft noise likeness detector. Applied Acoustics, 2010
[7] Asensio, C.; Moschioni, G.; Ruiz, M.; Tarabini, M.; Recuero, M. Methodology and implementation of a thrust reverse noise detection system for airports. Transportation research Part D, 2013.