WIRELESS SENSOR NETWORKS WITH NOISE SOURCE RECOGNITION CAPABILITIES

PACS.: 43.50.Rq

Valero González, Xavier; F. Alías Pujol.
GTM - Grup de Tecnologies Mèdia, La Salle – Universitat Ramon Llull.
Quatre Camins, 2, 08022.
Barcelona. España
Tel: +34 93 290 24 40
E.mail: xvalero@salleurl.edu; falias@salleurl.edu.

ABSTRACT

Current noise sensor networks are generally limited to collect basic sound signal information, typically the noise levels (in dB(A)). As a step forward, in this work, we aim at coming up with noise sensors that are able to automatically recognise the origin of the measured noise signal (i.e., road traffic, aircraft, railway noise, etc.) by making use of advanced audio signal processing techniques and artificial intelligence algorithms. Preliminary results employing data collected from different nodes of a sensor network are presented and discussed. In addition, a novel approach is proposed to tackle the complex problem of distinguishing among road noise sources.

RESUMEN

Las redes de sensores de ruido actuales generalmente se limitan a recoger información básica sobre la señal sonora, típicamente los niveles de ruido (en dB (A)). En este trabajo, nuestro objetivo es dar un paso adelante en el desarrollo de sensores de ruido que sean capaces de reconocer automáticamente el origen de la señal de ruido medido (por ejemplo, el tráfico por carretera, de aeronaves, ferroviario, etc.), haciendo uso de técnicas avanzadas de procesamiento de la señal y algoritmos de inteligencia artificial. Se presentan los resultados de los experimentos preliminares realizados usando datos de diferentes nodos de una red de sensores. Además, se propone un nuevo enfoque para abordar el complejo problema de la distinción entre los diferentes fuentes de ruido de tráfico rodado por carretera.

INTRODUCTION

It is a matter of fact that environmental noise affects human quality of life. Besides, it has been already proved that it provokes harmful effects to human health [1], [2]. Facing this situation, the EU Commission promoted the Environmental Noise Directive (END) to inform the citizens about their exposure to noise and draw up appropriate action plans to mitigate the harmful noise effects [3]. One of the main tools described in the END are the noise maps, which represent the acoustic reality of cities, transport infrastructures and industrial sites. Such noise maps are computed averaging the noise levels during one year, being reviewed once every five years.

Rather than just having a static acoustic picture periodically reviewed, disposing of dynamic pictures would allow more detailed assessments, as well as checking the effectiveness of the conducted noise actions plans in a very short term. In order to achieve this goal, acoustic sensor networks that monitor the noise levels in urban areas have been recently proposed [4], [5]. However, it would be highly interesting that acoustic sensor networks could relate the measured levels with the noise source origin (e.g., road traffic, aircraft, industry, etc.) at every instant of time. In this way, they would be able to provide the exposure to each environmental noise source independently. In addition to environmental noise assessment, this new sensor network functionality could have additional uses, such as in soundscape characterization [6] or security applications [7].

In order to incorporate such functionality to acoustic sensor networks, the implementation of automatic noise recognition systems is required. In this paper, we report recent experiments carried out to study the feasibility of implementing these recognition systems in low-cost noise sensor networks. In addition, we present the results of a research study devoted to improve the discrimination among different types of road traffic vehicles.

AUTOMATIC NOISE RECOGNITION SYSTEM

The system devoted to automatically recognise the environmental noise sources is based on monaural signal processing, thus, it can be deployed in any monitoring station disposing of a single microphone. It follows a pattern recognition approach that is divided into two main processes: the signal feature extraction block and the recognition block (see Figure 1).
The signal feature extraction block aims at parameterizing the noise signal by means of a set of coefficients. This processing step is especially relevant, since it models the spectro-temporal characteristics of the noise signals to be recognised. To this aim, the continuous audio stream recorded by the microphone is segmented into shorter frames using a Hanning window. Subsequently, a feature set is extracted from each signal frame, yielding representative patterns from each noise source. In this work, we select the biologically-inspired Gammatone Cesptral Coefficients (GTCC), which have recently shown an improved performance in environmental sound recognition tasks [8].

The second block takes the parameterised noise signals and labels them according to their acoustic content, i.e., it determines which of the environmental noise sources each signal belongs to. Two different approaches can be taken to conduct this task: supervised or unsupervised machine learning techniques. The first one consists in training the system with a set of known data (i.e., noise samples with its corresponding label), thus learning about the noise signal characteristics. Based on the acquired knowledge, the system is able to recognise new incoming non-labelled noise samples. This two-step process is depicted in Figure 1. On the other hand, the second approach takes the unlabelled parameterised noise signals and groups them into clusters according to their similarity, also providing a visual representation of the typical sound events occurring in a given location [9].

The selection of one of these two approaches depends on the prior information we dispose from the given location where the noise source recognition will be placed in. If there is no prior knowledge about the occurring sound events (or they are highly unpredictable), we should follow an unsupervised learning approach to separate out the sounds from the acoustic environment [9]. On the contrary, if relevant information about the typical noise sources is available a priori, we will rather make use of supervised machine learning techniques. In this work we will follow this last assumption, since we consider a monitoring station in a fixed location that can be trained beforehand (i.e., prior to entering into the operational mode). Among the multiple supervised learning techniques available in the related literature, in this work we consider the simple but effective K-Nearest Neighbor algorithm [8].

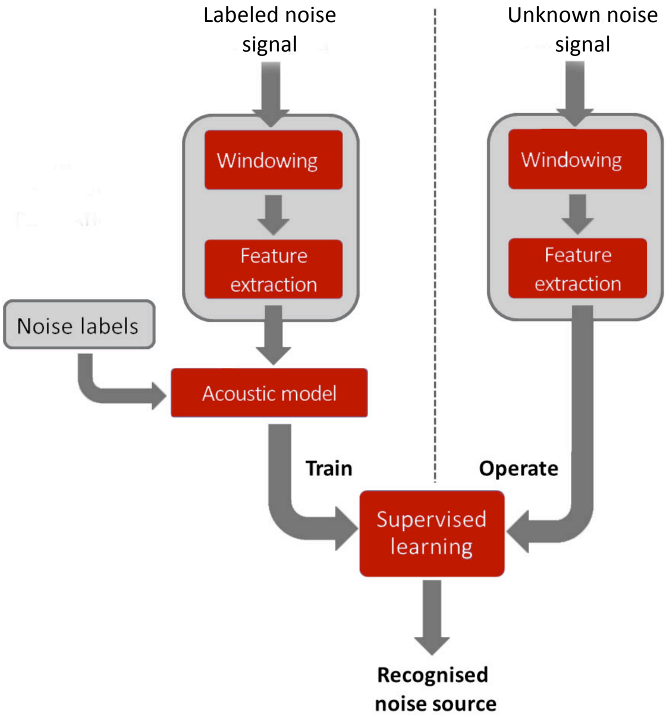

Figure. 1: Basic block diagram of the algorithm employed to automatically recognise environmental noise events.

NOISE SOURCE RECOGNITION IN WIRELESS NOISE SENSOR NETWORKS

Employing the recognition algorithm described in the former section, we aim at experimentally testing the feasibility of having noise source recognition capabilities in the wireless noise sensor networks, despite their particular characteristics (cheap microphones with poor frequency responses, compressed digital audio signals, etc.).

Experimental Work

The objective of the experimental work is to test the performance of the automatic noise recognition system with a set of recordings collected from the low-cost noise monitoring stations. More specifically, data was collected from the SENSEable network [5], located in the Italian city of Pisa. To this aim, two locations presenting different characteristics were selected.

The first noise monitoring station is located in a residential area of the city. Specifically, the microphone is placed 1.5 m in height and 8 m from the central line of the street. The acoustic environment is composed of road traffic noise (especially cars, scooters and buses) and residential noise. Series of recordings were carried out at different time periods of the day. Two of them were used to train the system and the remaining two to test it.
The second monitoring station is located in a residential area below one of the main flight paths from Pisa. The microphone is placed 7 m in height and 40 m from the central line of the street. In this location, aircraft and road traffic are the two primary noise sources. As in the first location, different series of recordings were performed to train and test the system.

The system output will be compared to the ground truth (known labels of the samples) and will be evaluated using precision, recall and F1 metrics [11] to obtain a robust result analysis.

Results

Road traffic noise was the main source in the first location. Thus, we aimed at distinguishing among the different types of road vehicles, i.e., scooters, cars, and buses. There were also fragments when people talking in the street were clearly audible. The quietest periods were labelled as *Background noise*, which consisted of the distant murmur from close streets traffic, birds chirping, wind softly blowing or some distant voices, just to name a few of them.

As shown in Figure 2, the three bus events are correctly recognised, as well as the excerpts with people talking. The few cases where the system fails are mainly concentrated in the scooter events. While the recall measure for this type of noises is high, the precision measure is quite lower, as shown in Table 1. Sometimes the system confuses scooters with cars, as already noticed in [10].

Besides road traffic, in the second location another important noise source was present: aircrafts flying over. As depicted in Figure 3, the aircraft noise events are almost perfectly recognised, yielding a F1 score about 90% (see Table 2). In only one case the system confuses an aircraft with a bus. Regarding the road traffic noise, similar conclusions as the ones obtained in the first location are extracted: good recognition of bus events and the occasional confusion between motorbikes and car events. It should also be noted that the recording distance to the center of the road is larger than in the first location, thus some quiet car events are confused with background noise, yielding a lower F1 score (66.3%).

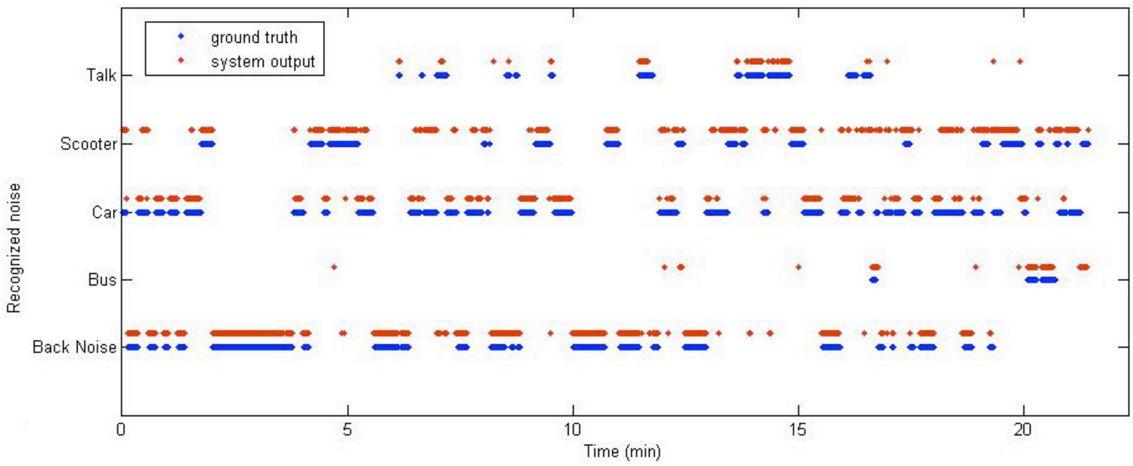

Figure. 2: Recognition system output and ground truth at each time stamp from the first location.

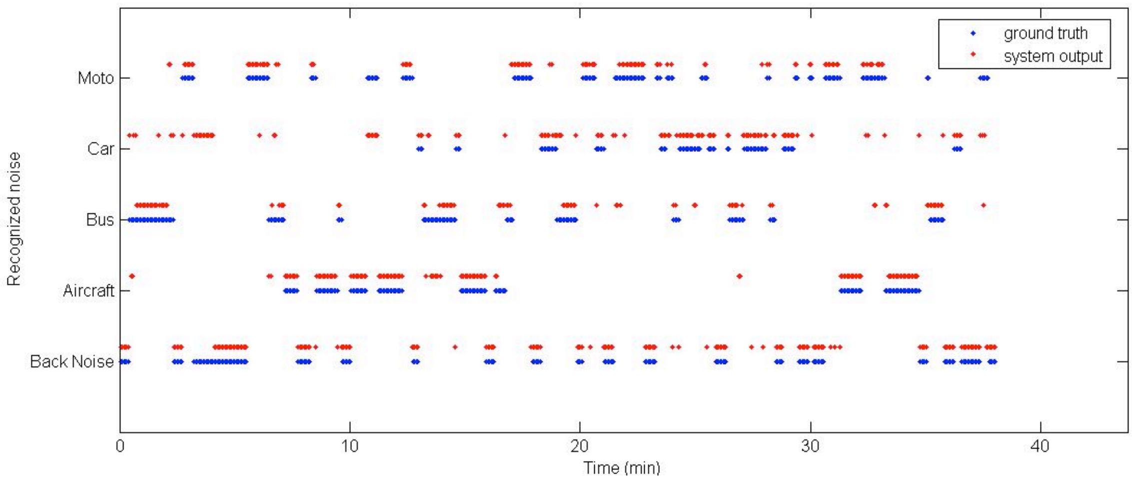

Figure. 3: Recognition system output and ground truth at each time stamp from the second location.

Noise	Precision (%)	Recall (%)	F1 score (%)
Backgr. noise	91.5	93.0	92.3
Bus	50.0	82.9	62.4
Car	87.3	63.7	73.7
Scooter	51.7	76.5	61.7
Talking	84.9	57.1	68.3

Table 1: Precision, recall and F1 measures of the system for the experiment in the first location

Noise	Precision (%)	Recall (%)	F1 score (%)
Backgr. noise	88.8	90.4	89.6
Aircraft	88.4	91.5	89.9
Bus	80.0	62.6	70.2
Car	53.1	88.5	66.3
Moto	91.6	69.5	79.0

Table 2: Precision, recall and F1 measures of the system for the experiment in the second location

IMPROVING ROAD TRAFFIC VEHICLE RECOGNITION

As observed in the previous experiments, the most frequent confusions are produced between different types of road traffic vehicles. Hence, specific research is conducted to improve the recognition of such noise sources. As a first step, we study the case of clear vehicle pass-bys (non-mixed sources).

Proposed Solution

The proposed solution consists in dividing the vehicle pass-by into three phases: when the vehicle is approaching (approaching), when the vehicle is perpendicular to the microphone (passing), and when the vehicle is moving away (receding). The sound signal recorded in those phases presents differences that can be perceived by simply listening to the recordings.
The physical explanation of the perceived responds to the multi-source nature of the road vehicles (engine, tyre, exhaust and aerodynamic noise) as well as to the Doppler effect. A more extensive explanation is provided in [10]. Several examples of vehicle pass-by division are depicted in Figure 4.

In order to tackle this goal, we propose a recognition algorithm adapted to the pass-by characteristics of this type of environmental noise signals (see Figure 5). The algorithm divides the time pattern of the parameterized noise sample into three parts, corresponding to the approach, passing and receding phases of the vehicle pass-by. An independent recognition decision is taken for each of these three phases and, finally, a second decision layer applies a simple voting scheme to come up with a unique solution (i.e., the recognized noise source). In case of tie (i.e., all three noise sources recognized are different), the second decision layer takes the identified noise source on the passing phase as final decision. We decided to follow this criterion since passing is the phase when the source is closer to the receiver (i.e., the microphone), thus it might be less contaminated by background noise).

Experiments

Rather than performing a continuous monitoring of the noise sources, as in the previous experiment, here we employed a corpus of isolated noise samples from road traffic vehicles, which should be recognised by the automatic recognition system. The database was composed of 90 noise samples from three types of vehicles: *motorcycles* (i.e. scooters and motorbikes), *light vehicles* (i.e. cars) and *heavy vehicles* (i.e., trucks and buses).

The recognition scheme uses Gaussian Mixture Models (GMM) as machine learning technique. For further details we address the interested reader to [10]. The experiment consisted in comparing the performance of the pass-by adapted recognition scheme (Figure 5) with a non-adapted one (Figure 1).

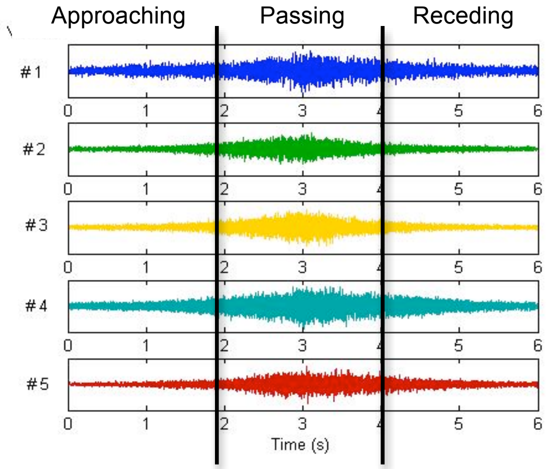

Figure 4. Time evolution of five road traffic vehicle pass-bys and their division into approaching, passing and receding phases.

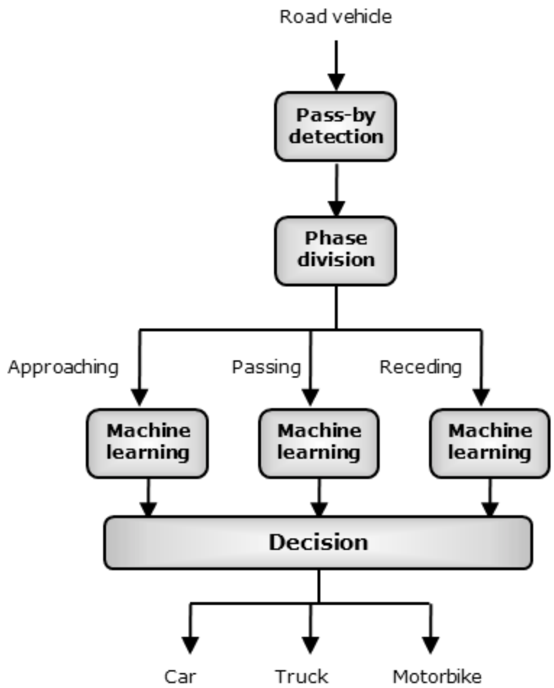

Figure 5. Recognition scheme adapted to road vehicles pass-by

Results

An averaged recognition accuracy of 84.6% is obtained with a classical non-adapted recognition scheme. However, if we adapt the recognition scheme to the vehicle pass-by characteristics (employing one GMM per pass-by phase), the averaged recognition rate increases up to 88.2% in average. As shown in Table 3, the two types of vehicles that showed the lower accuracy (light vehicles and heavy vehicles) attain a significantly improved recognition with the proposed scheme (7% and 4%, respectively).

Vehicle	Non-adapted system	Pass-by adapted system
Motorbike	89.2%	89.1%
Light vehicles	78.6%	82.4%
Heavy vehicles	86.1%	93.0%
Mean	**84.6%**	**88.2%**

Table 3. Averaged recognition rates yielded by both the adapted and non-adapted system for each type of road traffic vehicle.

CONCLUSIONS

In this work, we have proposed a supervised learning system to automatically recognise the environmental noise events recorded by low-cost wireless noise sensor networks. Experimental work has been carried out with noise recordings from two different noise sensor nodes in order to test the performance of the proposed system. Results have shown that the system is able to provide a notable transcription of the noise events occurring in the environment. Experiments have also revealed that the most frequent confusions occur among road traffic vehicles. In this sense, this paper has also reported some ongoing research studies on improving the road vehicle identification by dividing the acoustic pass-by signature into three phases when conducting the recognition process.

Future work will include the integration of those last advances on road traffic vehicle discrimination as well as the automatic computation of the equivalent noise levels from each noise source recognised by the system.

ACKNOWLEDGEMENTS

We would like to acknowledge the collaboration of Dustlab [12] for putting at our disposal the SENSEABLE wireless noise monitoring network from which the test noise samples where extracted. We would also like to thank Luca Nencini for his support in this project.

REFERENCES

[1] W. Babisch, "Transportation noise and cardiovascular risk: Updated review and synthesis of epidemiological studies", in Noise&Health, vol. 8, n.30, pp. 1-29, 2006.
[2] F. Rasche, Arousal and aircraft noise – Environmental disorders of sleep and health in terms of sleep medicine, in Noise & Health, vol. 6, n.22, pp. 15-26, 2004.
[3] EU Directive: Directive 2002/49/EC of the European parliament and the Council of 25 June 2002 relating to the assessment and management of environmental noise. Official Journal of the European Communities, L 189/12, July 2002.
[4] P. Wessels, T. Basten, F. van der Eerden, "Model based monitoring of urban traffic noise – A wireless sensor network design", in Proc. Internoise'12, New York, 2012.
[5] L. Nencini, P. De Rosa, E. Ascari, B. Vinci, N. Alexeeva, "SENSEable Pisa – a wireless sensor network for real-time noise mapping", in Proc. Euronoise'12, Prague, 2012.
[6] X. Valero, F. Alías, "Narrow-band Autocorrelation function features for the automatic recognition of acoustic environments", in Journal of the Acoustical Society of America, vol. 134, n.1, pp. 880-890, July 2013.
[7] S. Ntalampiras, I. Potamitis, N. Fakotakis, "An adaptative framework for acoustic monitoring of potential hazards", in EURASIP Journal on Audio, Speech and Music Processing, vol. 2009, January 2009.
[8] X. Valero, F. Alías, "Gammatone Cepstral Coefficients- Biologically Inspired Features for Non-Speech Audio Classification", in IEEE Transactions on Multimedia, vol. 14, n.6, pp. 1684-1689, December 2012.
[9] X. Valero, D. Oldoni, F. Alías, D. Botteldooren, "Support Vector Machines and Self-Organizing Maps for the recognition of sound events in urban soundscapes", in Proc. Internoise'12, New York, 2012.
[10] X. Valero, F. Alías, "Hierarchical classification of environmental noise sources considering the acoustic signature of vehicle pass-bys", Archives of Acoustics, vol. 37, n.4, pp. 423-434, December 2012.
[11] F. Sebastiani, "Machine learning in automated text categorization", in ACM computing surveys, vol.34, n.1, pp. 1-47, 2002.
[12] Dustlab (Distributed Ubiquitus Sensor Tecnology): http://www.dustlab.org/