

Invited Paper



Tools for studying noise effects based on spectral and temporal content

Ernesto Accolti, Federico Miyara

Laboratory of Acoustics and Electroacoustics, National University of Rosario, Riobamba 220 (bis),
2000 Rosario, Argentina
e-mail: {eaccolti, fmiyara}@fceia.unr.edu.ar

Abstract

A research paradigm for the perception, effects in the human being and assessment of environmental noise is described. It is based on the analysis of several kinds of sounds that compose such environmental noise. The proposed paradigm studies instantaneous psychoacoustic parameters of each sound that composes environmental noise and its relationship to the corresponding time-averaged parameters and in relation to the multi-layer-timbre composition. The use of algorithms published by the authors' research group in several scientific meetings is described in an approach that emphasizes the applications to the study of the effects of the noise in the human being. The tools developed within this framework are: a combination tool that generates a realistic environmental sound with controlled parameters (such as the spectrum), a system of virtual auralization that uses an algorithm for calculating the impulse response of rooms, a calibration protocol for audiometric earphones and interactive interfaces that use several anchor sounds for the evaluation of test sounds.

Keywords: Effects of noise, environmental noise, psychoacoustics, soundscape.

1 Introduction

Sound Quality evaluation of tools, machines and vehicles is a current research topic ([1] and [2]). These evaluations are based on the psychoacoustic paradigm that uses variables based on the spectral envelopes of sound through the specific loudness, such as Roughness, Fluctuation Strength, Sharpness and Loudness ([3] and [4]).

In contrast, most of the results referred to human assessment of environmental noise are based on A-weighted equivalent sound pressure level (and sometimes with corrections or other weightings due to specific situations). Many of these results give exposure-response relationships ([5] and [6]) intended for noises with almost homogeneous spectral

characteristics, such as road, rail and air traffic. These noises differ from each other in their spectral and temporal envelope content and, as predictable, in human response ([7]). These results summarize the synthesis and homogenization of surveys with simultaneous measurements made in many countries ([8]). Although these results involve several kinds of noises, they are unsuitable for environmental noises composed of diverse sound sources. Less work has been done seeking for environmental descriptors related to the psychoacoustic parameters ([9]) and trying to expand these parameters to get a better description of some specific behaviour ([10], [11] and [12]).

Specific noises in relation to the spectral (low frequency noise) and the temporal content (impulsive noise) have been treated specially and there is much work on these topics ([13], [14], [15] and [16]).

Furthermore, attempts to isolate the physical qualities of signals from their semantic content have been made by analyzing the temporal envelopes of the critical-bands ([17] and [18]).

The well known Soundscape approach paradigm differs from the classical paradigms in many aspects and particularly in the idea of preserving acceptable sounds. Both are similar as regards the concept of controlling unwanted sounds, although in the classical approach there was no attempt to separate desirable sounds from unwanted ones. This new paradigm is still under development. There are many papers that add positive categories ([19] and [20]) to the classical verbal categories (e.g. annoyance categories). Although the relationship between the spectral or temporal content and these positive categories is currently a research topic, it is well known that source segregation is related to concept of timbre which in turn is related to the spectral and temporal content. Other research groups involved in soundscape approach are working in blind source separation which may be used to classify environmental noise sources and, afterwards, suitable descriptors can be developed ([22]).

2 The research paradigm

The main goal is to study the perception and the human assessment of noise in connection with its spectral and temporal envelope content. We intend to study the reactions to interference caused by noise in concentration and tasks, to develop a mathematical model based in objective parameters capable of explaining those reactions, to achieve an objective model to describe the subjective concept of *semantic content* and to investigate its importance as a factor in the object of study.

The research paradigm is close to the soundscape approach combining some results from classical papers in this field. This paradigm also aims to describe impulsive noise and low frequency effects as special cases of envelope evolution and spectral content effects respectively. One aspect from soundscape approach incorporated in the present paradigm is the study of environmental noise as a combination of sounds from different kind of sources.

Even though psychoacoustic parameters have been validated for one sound source with stationary envelope, they are valid for instantaneous evaluations. However, some authors conclude that the overall scene cannot be assessed just as the average of the parameter values for individual instants or the superposition of individual sources ([21] and [11]).

Some auxiliary tools are necessary to advance in this paradigm and they may be reused in related paradigms. These tools are an aural simulation system and an algorithm for the controlled combination of sounds.

3 Tools

Recent advance in this paradigm include: An algorithm for the combination of sounds to get a simulated environmental noise with user-specified characteristics; computational tools for spectral and amplitude analysis of individual or combined environmental sounds; some computational blocks for the implementation of a virtual auralization system developed for audiometric headphones and an interactive platform for psychoacoustic experiments has been designed and implemented.

3.1 Spectral analysis

This tool was developed to get the spectral data of all the individual sounds that compose the database for the controlled combination algorithm. The tool is also useful for the analysis of environmental sounds either real or simulated by the artificial combination of sounds.

3.1.1 Spectrum

Spectral analysis is achieved by the energy average of the Fast Fourier Transform (FFT) of successive frames of $N = 4096$ samples. Each frame is previously multiplied by a Blackman window to reduce the spectral leakage. This window was selected to reduce mainlobe width, increase sidelobe attenuation and reduce computational cost.

The m -th FFT frequency f_m ($m = 0, 1, \dots, N-1$) of a signal with sample rate F_s is $f_m = mF_s/N$. Hence, the spectral resolution is F_s/N .

Most of the sounds in the database are sampled at $F_s = 44100$ Hz reaching a resolution of about 10 Hz. In order to get better resolution below 680 Hz without time resolution loss at higher frequencies, signals were decimated by a factor of 32 and a new FFT was applied. This version of the FFT has a 0.3 Hz resolution for sounds originally sampled at $F_s = 44100$ Hz. Finally all the frames are averaged in energy. Figure 1 shows the two aforementioned spectra for a motorcycle passby sound (in green for the original and in purple for the decimated version).

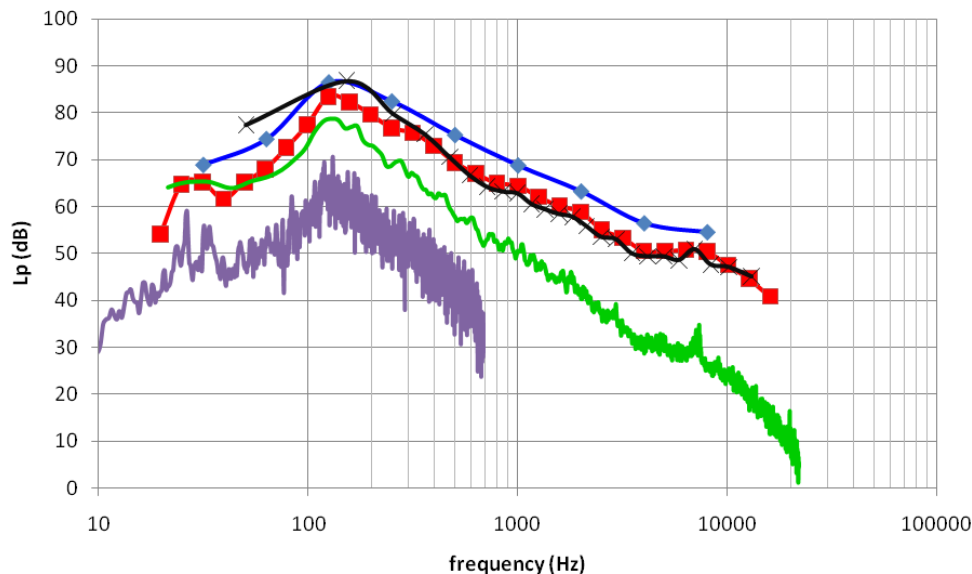


Figure 1 – Spectrum of a motorcycle passby. Green: original sample rate ($F_s = 44100$ Hz). Purple: decimated version ($F_s \approx 1400$ Hz). Red squares: 1/3 octave bands. Blue diamond: 1/1 octave bands. Black X: critical bands.

3.1.2 Band spectrum

Once both FFTs are obtained, the energy in each band is estimated by adding the energy of each FFT sample that falls inside the band, where the frequency of each FFT sample is f_m ([23] and [24]) (see blue diamonds and red squares in Figure 1). For the critical band levels the extremes of each band f_{ik} and f_{sk} in Hz are calculated with the inverse of the function reported in [3] for the critical band calculation. Finally the energies of the spectral lines that fall between f_{ik} and f_{sk} are added (See Figure 1)

3.2 Spectral amplitude envelopes

The spectral amplitude envelopes of the critical bands of a signal are important in this paradigm because they are strongly related to the psychoacoustic measurement of specific loudness. They are obtained by an algorithm that implements the block diagram in Figure 2.

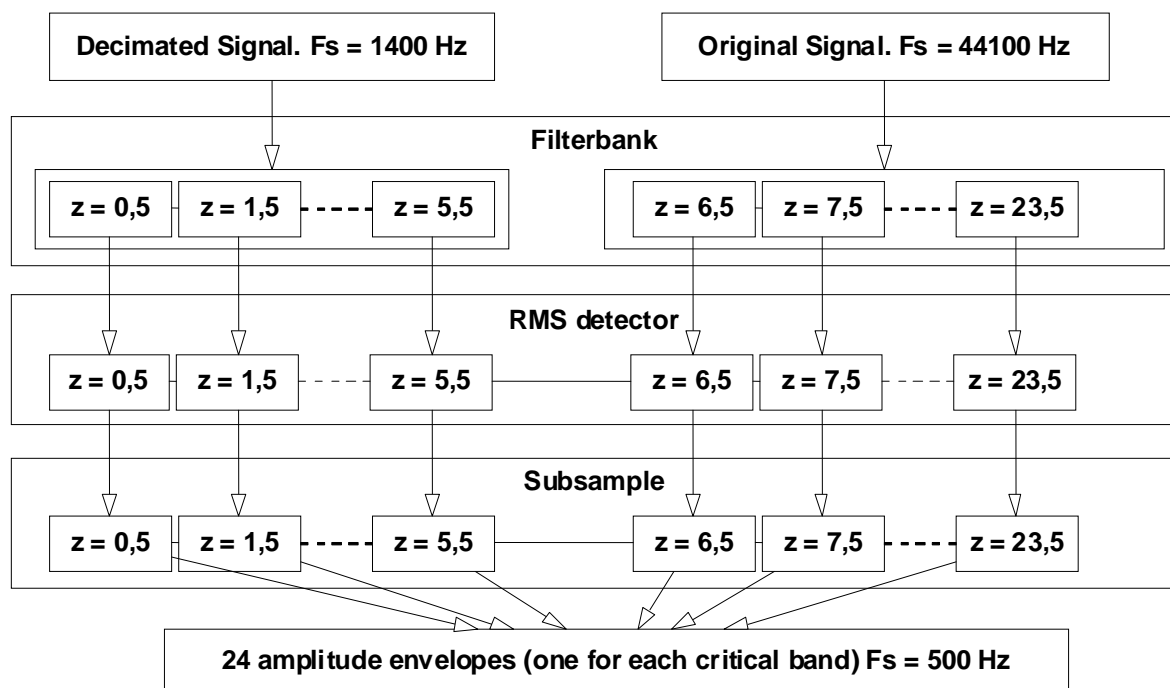


Figure 2 – Block diagram of the algorithm for the spectral envelopes extraction.

The decimated version of the original signal is used for frequencies under 680 Hz and both, the original and the decimated version, pass through a filterbank with 24 finite impulse response (FIR) filters designed by the method of frequency sampling. The first six critical bands use the decimated version for better resolution since the lower bands have narrower bandwidth than the higher ones. The critical band centered at 5.5 barks is the higher one that satisfies the Nyquist theorem for the decimated signal. The filtering block convolves each FIR with the signal by the overlap-add method. The implementation and validation of this algorithm can be seen in [23] and [24].

The RMS detector block is implemented as a first order bass-pass filter with time constant $\tau = 2$ ms applied to the squared signal.

Finally all the envelopes are decimated to a uniform sample rate $F_s = 500$ Hz and saved. The original information is thus reduced by a factor $44100 / 500 / 24 = 3.68$ and ready to use in psychoacoustic algorithms.

3.3 Controlled combination of sounds

The application of algorithms for the combination of sounds of different kinds by controlling the spectral and temporal characteristics of the combined sound has many benefits. One of these is the possibility of assessing the temporal and spectral cues responsible of the psychoacoustic parameters of an environmental noise. The control of these cues is very important for the experimental design when assessing the human responses to environmental noise.

It is known that the same global spectral and temporal results could be achieved from synthesis based on random noise but this technique underestimates the effect of the semantic content of real sound sources. The combination algorithm allows incorporating the effects of semantic content of the sources and the acoustic space related to the kind of source, location, meaning and acoustic field.

In [25] an algorithm has been developed that selects and combines sounds from a sound bank to get a simulated environmental sound with specified average spectrum. Currently it is being updated to include specified statistical envelope behaviour and constraints related to the acoustic space. Input sounds from the sound bank are first analyzed as described in 3.2, and the resulting spectra and critical band envelopes stored in an especially designed database.

Then an algorithm is applied to solve the undetermined system

$$\begin{aligned}
 A_{11} x_1 + A_{12} x_2 + \dots + A_{1M} x_M &= d_1 \\
 A_{21} x_1 + A_{22} x_2 + \dots + A_{2M} x_M &= d_2 \\
 &\vdots \\
 A_{N1} x_1 + A_{N2} x_2 + \dots + A_{NM} x_M &= d_N
 \end{aligned}
 \quad , x_j \geq 0 \tag{1}$$

where row i corresponds to the i -th spectral band, $A_{ij} = p_{ij}^2 t_i$ is the energy of the j -th sound in band i , x_j is the the square of the gain to be applied to sound j so that the linear combination has a target energy d_i in band i (expressed as a time-square-pressure product). N is the number of spectral bands, M is the number of sounds available in the bank.

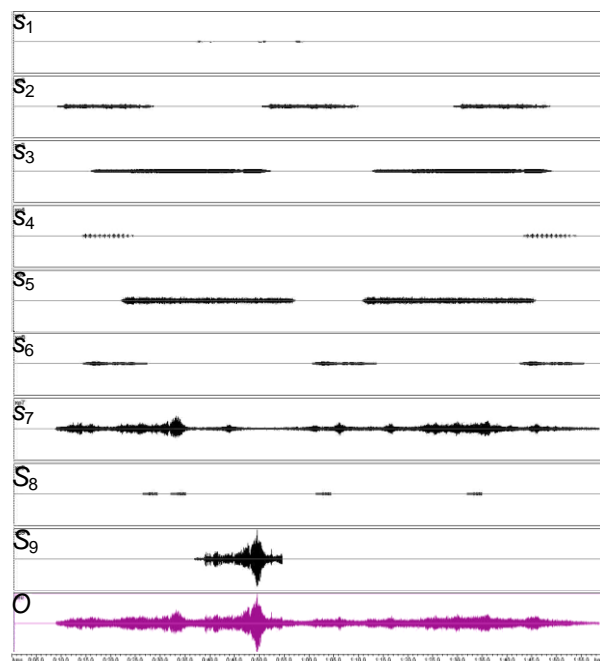


Figure 3 –Sounds s_i used to achieve sound O . Note the repetition pattern on a random basis.

The algorithm implements a least-squares solution with nonnegativity constraints ($x_j \geq 0$) that ensures the minimum quantity of nonzero values x_j .

Using a randomization technique some sounds are repeated introducing a suitable change in the required gain. Furthermore, the algorithm is adapted to allow the combination of several solutions obtained by successive removal of already used sounds ([25]).

Figure 3 shows sounds s_i required to obtain the output sound O for an arbitrary given spectrum. See [25] for a more detailed description of this algorithm.

Figure 4 shows two examples of octave band spectra. In the first one (right) there is a large content in the 2 kHz band. In the second one (left) the peak is reached at the 250 Hz band. Values specified by the user, predicted by the algorithm and calculated applying the spectral analysis tools to the combined sound are shown. Note that errors are at most about 2 dB.

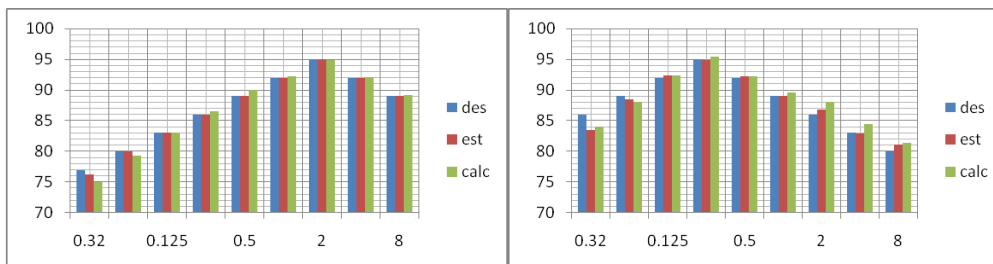


Figure 4 – Octave band spectrum of two solutions implemented by the controlled combination algorithm. Blue: User-specified values. Red: Predicted values. Green: Values calculated applying the spectral analysis tools to the output sounds.

3.3.1 Aural simulation of road traffic noise

Road traffic is a predominant noise source in the soundscapes of most cities around the world ([19]) and in this work it will be included in a more specialized manner allowing to control parameters such as the instantaneous velocity patterns of each individual and the percentage of each kind of vehicles.

This work is also under development. Some advances in this line are the simulation of a car moving at constant speed in a freeway and along a “U” profile street formed by building façades (See [26], [27] and [28]). This tool will allow studying traffic noise effects on the human being as regards speed, flow or traffic composition.

3.4 Virtual auralization system

The experiment tasks will use audiometric earphones. For this reason an auralization system has been developed. To date, we have already developed calibration protocol for audiometric earphones ([29]) and a fast algorithm to compute the impulse response (RIR) of box-shaped rooms ([30]).

The calibration protocol allows the design of a digital compensation filter for achieving an auralization system capable to present a calibrated free-field to a subject from an input sound file recorded or generated as if it had been recorded in a free field condition. The calibration applies the frequency-sampling filter design method to the measured response of the earphones in order to replace the classical analogue circuits used for this task. This calibration allows compensating all the blocks in the electroacoustic chain (Figure 5). Diffuse field calibration is used when calculating the late-reverberation part of the RIR.

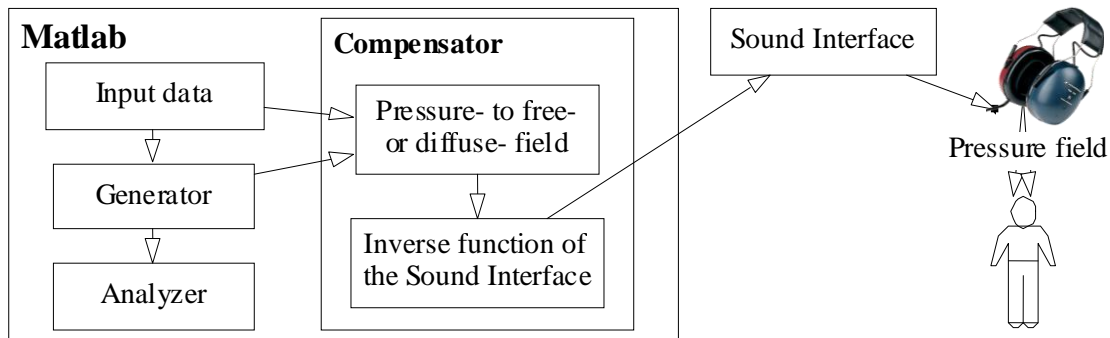


Figure 5 – Block diagram for the calibration protocol

The RIR algorithm that has been implemented is based on the classical image source method (ISM) for the early reverberation (ER) and amplitude modulated random phase noise for the late part of reverberation (LR). Two new techniques have been introduced in [30] for faster calculation. The first one replaces the frequency-sampling method commonly used for the ER calculation with a linear combination of impulse responses of 1/1 or 1/3 octave band filters. These responses are computed once and reused each time they are required. For each spectral band, the weightings depend on the cumulative absorption and air attenuation along the multiple reflection path of each ray.

The second technique corresponds to a new model for the transition response between ER and LR. After the end of ER, discrete time arrivals are still computed by the ISM, but instead of computationally heavy impulse responses, they are assigned isolated values taken from a bimodal random noise. A special procedure described in [30] is used to accomplish a smooth energy balance between the three parts of the response.

The head related transfer functions (HRTF) of a specific or generic listener are included in the algorithm to account for directional effects of the source and rays.

This algorithm is useful to include the effects of different kinds of rooms, for example their influence on the psychoacoustic parameters. The post-masking effects are very important for the psychoacoustic parameters. Post-masking is defined as the time dependent sound level required for a very short test signal following a masker signal to be just audible. A consequence of post-masking is that sounds ending abruptly in the free-field (i.e., without reverberation) are perceived as if they would have a decay time of the order of hundreds of milliseconds but in a room, this effect is reinforced by reverberation.

3.5 Interactive interfaces

3.5.1 Classical magnitude estimation method

In the psychoacoustic experimentation context, this method consists in the evaluation of the test sound in relation to a reference sound (usually called *anchor*) with a previously known value of the variable under test. The test sound is usually compared with one or two anchors in a paired comparison procedure (one at a time) and the subjects are expected to estimate the value of the subjective variable of the test sound in relation to the anchor(s) (See [3]).

Using more than one anchor is useful, particularly if the ratio of the test sound to the anchor is far from unity (e.g. 0,05 or 20). However the number of tests increases proportionally to the number of anchors in a paired comparison procedure.

Magnitude estimation is often a difficult task for subjects and in many cases require previous training, especially when the expected value is very different from the anchor's value.

3.5.2 Magnitude estimation by free anchor selection

Preliminary experiences ([11]) confirm the observation in [3] about the difficulty for subjects of magnitude estimation tasks. Some experiments have been conducted by equality estimations in the same experiences with less difficulty for subjects but the time required for this kind of studies depends on the number of tests. The magnitude estimation method is less difficult for subjects if the anchor is closer to the test sound value so the use of 10 anchors was introduced. When a paired comparison method is used, the number of required tests is proportional to the number of anchors, so a free selection of these anchors was introduced. An interactive interface for this method was developed and implemented in [11].

A group of ten buttons play sounds with known and equidistant values of the magnitude under study. The difference between the consecutive anchors (e.g. button 5 and button 6) is larger than the JND for this parameter. A “Play Test” button is used to play the test sound. The subjects are expected to listen to the test sound and to the anchors and choose the best fit. They are allowed to rehearse as many times as needed before making their choice. Once the estimation is done and introduced in a textbox, a “Next” button allows going ahead to the subsequent test.

Note that the JND and the relation between the objective parameters and the subjective sensation must be known when designing the set of anchors.

This interface has been successfully used in training procedures by replacing the test sounds with sounds whose magnitude values are known. The training procedure is the same as for the experiments, except that the magnitude that corresponds to the training sounds is informed to the subjects after they introduce their estimation. A set of $i = 10$ anchors was designed for training purposes. They consisted in amplitude-modulated tones with fixed carrier frequency $f = 1$ kHz, fixed modulation frequency $f_{\text{mod}} = 4$ Hz and different values of modulation index m_i . Zwicker’s model (See [3]) for the fluctuation strength of an amplitude modulated tone was used to get a set of m_i that provides evenly spaced values of fluctuation strength so that consecutive values exceed the JND. Amplitude modulated broadband noises were used as training sounds, estimating their fluctuation strength with Zwicker’s model ([3]). This interface was used in preliminary experiments on the fluctuation strength of mixed sources ([11]). Relative errors for each subject get smaller for a longer training period.

4 Conclusions

A set of tools have been developed for research tasks within the paradigm for the study of perception, assessment and effects of noise in the human being in connection with the spectral content and temporal pattern of noise. This paradigm complements the classical paradigms with new factors taking advantage of the advances in computer technology. Furthermore, the methods and techniques of this paradigm are consistent with ongoing paradigms such as Soundscape.

An algorithm for the controlled combination of sounds has been introduced, along with some possible applications. It is a powerful tool for studying the human response to environmental noise in connection with overall psychoacoustic magnitudes and their relation to those of each isolated sound. One benefit of this approach, in contrast with the use of synthesized noise, is that the semantic content of real sounds is present. Other application, useful for soundscape research, is the possibility of studying the influence of a particular sound within an environmental noise composed by sounds of diverse kinds.

The algorithm is complemented by a virtual auralization system that incorporates a calibration protocol for free-field and a RIR algorithm for diffuse-field. The calibration protocol is the digital version of the analogue electronics headphones equalizers. The RIR algorithm introduces some new techniques, making it faster yet realistic.

The interactive interface introduces a new method for psychoacoustic experimentation based the magnitude estimation and paired comparison methods. This method can be described as a magnitude estimation procedure by free selection of anchors. A training procedure has also been implemented using the same interface.

Acknowledgments

This work is part of research project ANPCyT - PICT N° 38109, funded by the Agencia Nacional de Promoción Científica y Tecnológica (National Agency of Scientific and Technological Promotion) from Argentina. The attendance of the first author to INTERNOISE 2010 has been made possible through the Young Scientist Grants from I-INCE.

References

- [1] Sato, S.; You, J.; Yong J. Sound quality characteristics of refrigerator noise in real living environments with relation to psychoacoustical and autocorrelation function parameters. *J. Acoust. Soc. Am.*, Vol 122(1), 2007, pp. 314-325.
- [2] Wang, Y. Sound quality estimation for non-stationary vehicle noises based on discrete wavelet transform. *Journal of Sound and Vibration*, Vol 324, 2009, pp. 1124-1140.
- [3] Zwicker, E.; Fastl, H. *Psychoacoustics: Facts and Models*. Springer, Berlin, 2007.
- [4] Chalupper, J.; Fastl, H. Dynamic loudness model (DLM) for normal and hearing-impaired listeners. *Acustica united with Acta Acustica*, Vol 88, 2002, pp. 378-386.
- [5] Miedema H.; Vos H. Noise annoyance from stationary sources: Relationships with exposure metric day-evening-night level (DENL) and their confidence intervals. *J. Acoust. Soc. Am.*, Vol 116(1), 2004, pp. 334-343.
- [6] Brown, A.; Van Kamp, I. Response to a change in transport noise exposure: A review of evidence of a change effect" *J. Acoust. Soc. Am.* Vol. 125(5), 2009, pp. 3018.
- [7] Miedema, H.; Vos, H. Exposure-response relationships for transportation noise. *J. Acoust. Soc. Am.*, Vol 104(6), 1998, pp. 3432-3445.
- [8] Fields, James M. An Updated Catalog of 521 Social Surveys of Residents' Reactions to Environmental Noise (1943-2000). Wyle Laboratories, Hampton, Va. National Aeronautics and Space Administration, Langley Research Center, 2001.
- [9] Schomer, P.; Suzuki, Y.; Saito F. Evaluation of loudness-level weightings for assessing the annoyance of environmental noise. *J. Acoust. Soc. Am.*, Vol 110 (5), 2001, pp. 2390- 2397.
- [10] Genuit, Klaus; Fiebig, André. Psychoacoustics and its Benefit for the Soundscape Approach. *Acta Acustica united with Acustica*, Vol 92, 2006, pp. 1-7.
- [11] Accolti, Ernesto; Miyara, Federico. Fluctuation Strength of Mixed Fluctuating sound Sources. *Mecánica Computacional (Computational Mechanics)* Vol. XXVIII (2), 2009, pp. 9-22.
- [12] Accolti, E.; Miyara, F. Herramientas para el estudio de la percepción, valoración y efectos en el ser humano en relación al contenido espectral y envolvente temporal del ruido. *1as Jornadas Regionales de Acústica AdAA Rosario(Argentina)*, 2009, code A026.
- [13] Berglund, B.; Hassmén, P.; Soames R. F. Sources and effects of low-frequency noise. *J. Acoust. Soc. Am.*, Vol 99 (5), 1996, pp. 2985-3002.

- [14] Schomer, P. Evaluation of a re-analysis of the relationship between the results obtained in laboratory and field studies on the annoyance caused by high-energy impulsive sounds. *Noise Control Engineering Journal*, Vol. 45 (6), 1997, pp. 251-255.
- [15] Beaman, P. Auditory distraction from low-intensity noise: a review of the consequences for learning and workplace environments. *Applied Cognitive Psychology*, Vol. 19, 2005, pp. 1041-1064.
- [16] Moorhouse, A.; Waddington, D.; Adams, M. Proposed criteria for the assessment of low frequency noise disturbance. *Report by University of Salford*, 2007, URN 07/1235
- [17] Blauert, Jens (Editor). *Communication Acoustics*. Springer, Verlag Berlin 2005.
- [18] Cabanellas S.; Pasch V.; Yanitelli M.; Miyara F. Estudio exploratorio de algunos parámetros perceptivos del ruido vinculados al contenido semántico. *V Congreso Iberoamericano de Acústica FIA*, Santiago(Chile), 2006, code A011 pp. 1-10.
- [19] Raimbault, M.; Dubois, D.; Urban soundscapes: Experiences and knowledge. *Cities*, Elsevier, Vol. 22(5), 2005, pp. 339-350.
- [20] Maffei, Luigi. Urban and quiet areas soundscape preservation. Conferencia Plenaria. *VI Congreso Iberoamericano de Acústica, FIA*, Bs As, 2008, code CP1.
- [21] Kuwano, S.; Namba, S.; Kato, T.; Hellbrück, J. Memory of the loudness of sounds in relation to overall impression. *Acoust. Sci. & Tech*, Vol 24(4) 2003, pp. 194-196.
- [22] Bunting, O.; Stammers, J.; Chesmore, D.; Bouzid, O.; Tian, G.; Karatsovis, C.; Dyne, S. Instrument for soundscape recognition, identification and evaluation (ISRIE): technology and practical uses. *EURONOISE*, Edinburgh(Scotland), 2009
- [23] Miyara, F.; Cabanellas, M.; Pasch, V.; Yanitelli, S.; Accolti, E.; Miechi, P.; Contrastación de algoritmos de análisis de espectro sonoro con un instrumento normalizado. *1as Jornadas Regionales de Acústica AdAA*, Rosario (Argentina), 2009, code A032.
- [24] Miyara, F.; Cabanellas, M.; Pasch, V.; Yanitelli, S.; Accolti, E.; Miechi, P. Contrastación de algoritmos de análisis de espectro sonoro con un instrumento normalizado. *Mecánica Computacional (Computational Mechanics)*, Vol XXVIII(2), 2009b, pp., 199.
- [25] Accolti, Ernesto; Miyara, Federico. Combinación digital controlada de ruidos diversos. *VI Congreso Iberoamericano de Acústica, FIA*, Bs As, code A090
- [26] Miyara, F.; Cabanellas, S.; Mosconi, P.; Pasch, V.; Yanitelli, M.; Vazquez, J.; Aural simulation of the effect of façades on vehicle noise. *SePaViAm*, Bs As, 2003.
- [27] Marengo R.; F., Miyara, F.; Design and implementation of a doppler removal method for parameters extraction and aural simulation of traffic noise. *Mecánica Computacional (Computational Mechanics)*, Vol XXVIII, 2009, pg. 89-100.
- [28] Marengo R., F.; Miyara, F. Parameters extraction and aural simulation of traffic noise by means of a novel Doppler removal method. *XIII Reunión de Trabajo en Procesamiento de la Información y Control, RPIC*, Rosario (Argentina), 2009. CD-ROM.
- [29] Accolti, Ernesto; Miyara, Federico; Mignini, Ezequiel. Protocolo de calibración de auriculares audiométricos para su uso en investigación psicoacústica *1as Jornadas Regionales de Acústica AdAA*, Rosario (Argentina), 2009, code A025.
- [30] Accolti, Ernesto; Miyara, Federico. Fast and controllable box-shaped room impulse response algorithm. *XIII Reunión de Trabajo en Procesamiento de la Información y Control, RPIC*, Rosario (Argentina), 2009. CD-ROM.