The Use of Time-Scale Analysis in Forensic Acoustical Investigations

LUIGI MAXMILIAN CALIGIURI, DOMENICA GIORDANO
Foundation of Physics Research Center (FoPRC)
ITALY
luigimaxmilian.caligiuri@foprc.org

Abstract: - The identification and characterization of noise events is one of the most important task in acoustical forensic analysis. In this field it is often fundamental to distinguish, within a complex acoustical framework, the different noise events, especially because, in many cases, the operator cannot be present at the measurements. It is fundamental to be able to distinguish the atypical or extraneous noise events from the specific ones under investigation and know what type of sources make up the noise climate. To this aim is essential to develop a time - frequency analysis technique able to overcame the known limitations of the "traditional" 1/3 of octave frequency analyses. In this paper a novel technique, based on multiresolution analysis, has been developed and applied to some forensic "typical" problems, showing that a suitable choice of the analysis parameters can be able to answer to the main questions of this field.

1 INTRODUCTION

The individuation and recognition of noise events play a fundamental role in forensic analysis. Within this framework is in fact required, in order to perform a given investigation, to be able to separate specific acoustic signals from atypical or sporadic events determining the so – called "noise background". The signals to be recognized can be very different in their time – frequency behaviour because they can include human voice as well as "natural" noise events and, virtually, any kind of "man made" noises.

The progress in computer technology has allow the development of digital analysis techniques and their implementation, in addition to the common personal computers, in commercial analyzers on the market.

In acoustical analysis (as in many other science areas), and especially in forensic ones, a fundamental goal of these techniques must be the "correct" characterization of the time – frequency content of a given signal.

The most part of digital techniques currently used in such analyses are based on different versions of windowed numeric Fast Fourier Transform (FFT) or 1/n of octave analysis. Although useful in the most common applications, these methodics show very critical limitations when applied to complex situations characterized by the presence of many noise sources with very different time – frequency and high "non – stationary" (both in frequency and time) behavior as almost always occurs in forensic acoustics.

Their most important limitation is, in fact, when applied to highly non- stationary signals , they does not tell us about the time location of the frequency components of the spectra and, consequently that they are not be able to give us the "signature" of the different noise sources within an acceptable precision (referred to the purpose of forensic analysis).

Multiresolution analysis can give us the solution to this problem getting the spectra associated to samples of signal that can be considered sufficiently stationary (both in time and frequency); one of the most important multiresolution analysis is that obtained by Wavelet Transform (WT) that leads to a multiresolution time – scale space of representation.

In this paper we'll discuss an application of a multiresolution technique, based on WT, to a typical forensic acoustics case showing how a suitable choice of analysis parameters can be able to properly characterize the noise signals in comparison with the failure of the traditional fixed resolution analyses.

2 FUNDAMENTALS OF MULTIRESOLUTION ANALYSIS

2.1 Multiresolution approximation

A multiresolution analysis gives a "zoomed" vision of a signal at different scales in such a way that we will dispose of different signal versions each of them characterised by a different detail level. If $L^2(\mathbb{R})$ designs the Hilbert space of square integrable signals $f(t)$, a multiresolution approximation (MRA) for f is a sequence of closed sub-spaces $(V_k)_{k \in \mathbb{Z}}$ such that

$$f(t) \in V_j \Leftrightarrow f(t - 2^j k) \in V_j \; j,k \in \mathbb{Z} \qquad (1)$$

$$\forall j \in \mathbb{Z}, V_{j+1} \subset V_j \qquad (2)$$

$$\forall j \in \mathbb{Z}, f(t) \in V_j \Leftrightarrow f(2t) \in V_{j-1} \qquad (3)$$

$$\bigcap_{j=-\infty}^{\infty} V_j = \{0\} \qquad (4)$$

$$\bigcup_{j=-\infty}^{\infty} V_j = L^2(\mathbb{R}) \qquad (5)$$

There exists furthermore a function ϑ such that the ensemble $\{\vartheta(t - n)\}_{n \in \mathbb{Z}}$ is a Riesz basis of V_0 that's to say the $\{\vartheta\}_n$ are linearly independent and there exist $A, B > 0$ such that, for any $f \in V_0$, one uniquely has

$$f(t) = \sum_{n=-\infty}^{\infty} a[n]\vartheta(t - n) \qquad (6)$$

satisfying the property

$$A\|f\|^2 \le \sum_{n=-\infty}^{\infty} |a[n]|^2 \le B\|f\|^2 \qquad (7)$$

According to this approach, the approximation of a signal f at scale 2^{-j} can be seen as the orthogonal projection of f on space V_j, and will be denoted by $P_{V_j} f$. Therefore if $\{b_j\}_n$ is an orthogonal basis of V_j we have

$$P_{V_j} f = \sum_{n=-\infty}^{\infty} b_j[n]\phi_{j,n} \qquad (8)$$

where $a_j[n] = \langle f, b_{j,n} \rangle$.

It's possible to prove [1] that an orthogonal basis of V_j can be constructed by dilating and translating a function ϕ called "scaling" function; if $(V_k)_{k \in \mathbb{Z}}$

is an MRA and ϕ is the scaling function whose Fourier transform is

$$\hat{\phi}(\omega) = \frac{\hat{\vartheta}(\omega)}{\left(\sum_{k=-\infty}^{\infty} |\hat{\vartheta}(\omega + 2k\pi)|^2 \right)^{1/2}} \qquad (9)$$

then the functions $\{\phi_{j,n}\}_{n \in \mathbb{Z}}$ where

$$\phi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \phi\left(\frac{t - n}{2^j} \right) \qquad (10)$$

represent an orthogonal basis of V_j. An *MRA* is then completely defined by means of the scaling function ϕ that generates an orthogonal basis for each space V_j. If we indicate with W_j the orthogonal complement of V_j in V_{j-1} we can write

$$V_{j-1} = V_j \oplus W_j \qquad (11)$$

and the approximation of f at scale 2^{j-1} $P_{V_{j-1}} f$ is given by

$$P_{V_{j-1}} f = P_{V_j} f + P_{W_j} f \qquad (12)$$

The projection $P_{W_j} f$ gives us the informations about the "details" of the signal at scale 2^{j-1} that we cannot "view" at less accurate scale 2^j. It has been shown [1] that an orthonormal basis of W_j can be constructed by means of the so called "wavelets" waveforms. The projection of a signal f in the "detail" space W_j is given by an expansion in terms of wavelet basis $\{\psi_j\}_n$

$$P_{W_j} f = \sum_{n=-\infty}^{\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n} \qquad (13)$$

A signal f can be then expressed as composition of the "details" at all scales

$$f = \sum_{j=-\infty}^{\infty} P_{W_j} = \sum_{j=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n} \qquad (14)$$

if we know an orthogonal wavelet basis.

2.2 Wavelet basis

A continuous wavelet is a function $\psi \in L^2(\mathbb{R})$ such that

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \qquad (15)$$

$$\int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty \qquad (16)$$

(where $\hat{\psi}$ is the Fourier transform of ψ) usually normalised $\|\psi\| = 1$ and centred in the neighbourhood of $t = 0$. A wavelet "family", whose elements are called "atoms", is obtained by translating ψ by u and by "scaling" ψ by a factor s:

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right) \qquad (17)$$

with $\|\psi_{u,s}\| = 1$. The continuous wavelet transform of a signal $f(t)$ is the image $WT^{\psi}[f](u,s)$ on $]0,\infty[\times \mathbb{R}$ given by

$$WT^{\psi}[f](u,s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{\infty} f(t)\psi^*\left(\frac{t-u}{s}\right)dt \qquad (18)$$

(where * indicates the conjugate complex) and quantifies the contribution of the signal f in the neighbourhood of $t = u$ at a scale s. The time translation of atoms permits to us to select the different portions of signal while by dilating and contracting it we are be able to analyse signal structures of very different size. When the atoms are dilated we analyse the signal components oscillating more slowly while when they are contracted we can analyse the components oscillating more quickly.

We can construct a wavelet function such that the family of translated and scaled functions

$$\left\{\psi_{j,n}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t-2^j n}{2^j}\right)\right\}_{j,n\in\mathbb{Z}} \qquad (19)$$

is an orthonormal basis of $L^2(\mathbb{R})$. The basis can be obtained from a scaling function ϕ

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2}}\hat{\phi}\left(\frac{\omega}{2}\right)\hat{g}\left(\frac{\omega}{2}\right) \qquad (20)$$

with

$$\hat{g}(\omega) = e^{-i\omega}\hat{h}^*(\omega + \pi) \qquad (21)$$

where h represents the conjugate mirror filter[1] corresponding to ϕ.

The multiresolution approach and its realization by means of wavelet basis leads, first of all, to a time – scale decomposition. Nevertheless, we can "convert" it into a time – frequency signal decomposition, by properly choosing the wavelet basis. We can establish such a relationship between

scale and frequency by assuming that the base function is placed, in the frequency domain, around a frequency ω_0, corresponding to the maximum value of the wavelet frequency spectrum. The simultaneous time-frequency resolution of analysis is limited by the Heisenberg uncertainty principle:

$$\Delta\omega\Delta t \geq \frac{1}{2} \qquad (22)$$

where the time resolution Δt is related to the wavelet time spread Δt_{ψ} (with $\Delta t = s\Delta t_{\psi}$) and the frequency resolution to the wavelet bandwidth $\Delta\omega_{\psi}$ ($\Delta\omega = \Delta\omega_{\psi}/s$). The product $\Delta\omega\Delta t$ and so the effectiveness of the analysis depends on the wavelet waveform shape and parameterization. To analyse the signals associated to electric and magnetic fields, some classes of wavelet atoms, with Fourier transforms obtained from equation (20), have been employed.

They are the Gauss, Morlet, Gabor, Franklin wavelets. The Gauss wavelet is the function whose Fourier transform is given by

$$\hat{\psi}_{Gauss}(\omega) = \exp\left[-\omega^2/2\right] \qquad (23)$$

the Gabor wavelet is a "phase displaced" version of a gaussian function

$$\hat{\psi}_{Gabor}(\omega) = \exp\left[-(\omega - \omega_0)^2/2\right] \qquad (24)$$

with $\omega_0 = \pi\sqrt{2/\log 2}$. In the same way we can define the Morlet and Franklin atoms:

$$\hat{\psi}_{Morlet}(\omega) = \exp\left[-(\omega - \omega_0)^2/2\right] + \\ -\exp\left[-(\omega + \omega_0)^2/2\right] \qquad (25)$$

$$\hat{\psi}_{Frank}(\omega) = \cos\left(\frac{\pi\omega}{16}\right)\sin\left(\frac{\pi\omega}{16}\right)^2 \times \\ \times\sqrt{\frac{2 - \cos(\pi\omega/16)}{[2 + \cos(\pi\omega/16)][2 + \cos(\pi\omega/8)]}}\left[\sin c\left(\frac{\pi\omega}{8}\right)\right]^2 \qquad (26)$$

It's important to note that the WT leads to a frequency analysis with constant percentage bandwidth (fraction of octaves) whose value (1/3, 1/6, 1/12 of octave for example) depends on the atom form and on the parameterisation chosen for the analysis. This choice must be made in function of the original signal features, so, in performing the WT, it is necessary to have a preliminary, at least qualitative, knowledge of signal to be analysed.

---

[1] A conjugate mirror filter can be defined as a discrete filter whose transfer function satisfies the condition $\forall \omega \in \mathbb{R}, |\hat{h}(\omega)|^2 + |\hat{h}(\omega + \pi)|^2 = 2$.

3 APPLICATION OF MULTIRESOUTION ANALYSIS TO A FORENSIC CASE

3.1 "Traditional" analysis

As an example, we'll considerer the signal showed in Fig. 1 (on the x axis is represented the time, on the y axis the wave intensity in arbitrary units normalized to 1), representing the acoustical wave, measured by a sound pressure level digital analyzer, associated with a complex noise scenario including typical sources considered in forensic acoustics.
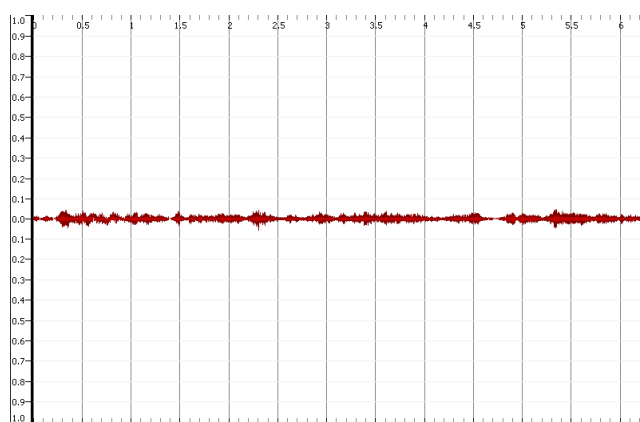

Figure 1: time history of the wave amplitude of digital signal considered.

In particular the noise climate is composed as follow:

  a) from $t = 0$ to the end ($t = 6.247\ s$) a male voice A superimposed to a music background;
  b) from $t = 0$ to the end a very low background composed by a stationary – continuous (almost monochromatic with few harmonics) noise;
  c) from $t = 2.278\ s$ to the end a female voice B characterized by overall pressure levels slightly greater than those of the voice A.

When we listen at the sound associated to the wave of fig. 1, we are be able to clearly recognize the voice A until the start of voice B when we clearly ear only the voice B (and not A) until the end of the sample. During the whole time interval, we not ear the stationary – continuous source at all. The voice B then masks the voice A almost completely and both of them mask the continuous source. The forensic problem can be here, for example, to show that the voice A is always present in the sample, together with the continuous source, even when we cannot hear them after $t = 2.278\ s$ and that this

voice in really the same as that heard until $t = 2.278\ s$ (the same person).

The analysis of the wave amplitude time –history gives us no useful information and the waveform amplitude doesn't reveal what is happening.

The fig. 2 shows the result of an 1/3 of octave not windowed frequency numerical analysis performed on the signal of fig. 1 (on the x - axis is reported the frequency band, on the y – axis the sound pressure level amplitude normalized do 0 dB for both the right and left recorded channels).
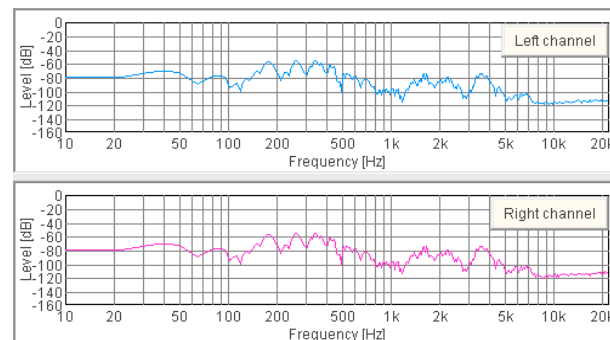

Figure 2: 1/3 of octave frequency analysis of the considered signal.

As obvious, due to the highly not – stationary time – frequency features of signal, from this kind of spectrum we cannot extract useful data in order to recognize and characterize the sources under investigations.

A better method can be the one based on a windowed version of FFT that can give a spectrogram like that showed in the Fig. 3 (on the x – axis is shown the time in sec, on the y – axis the frequency in Hz and the image points intensity is proportional to sound pressure level in dB),
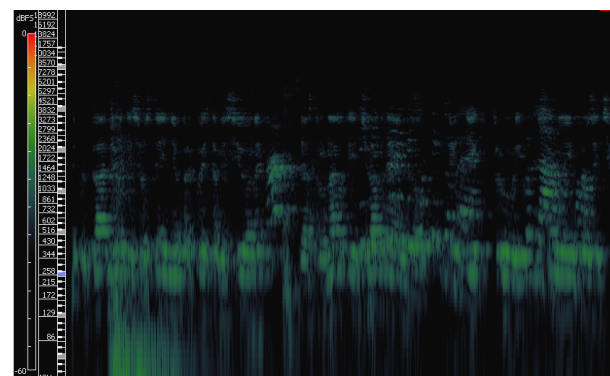

Figure 3: FFT spectrogram of the considered signal.

obtained considering a moving rectangular time window of 1024 points which achieve the best compromise between time – frequency resolution and image clarity (a different window width would not alter the meaning of the treatment). The

windowed FFT analysis is able to give us some interesting information about the time – frequency structure of sample. In particular it shows the presence of a middle – frequency (from about 215 to about 700 Hz) pattern during the whole sample whose time – frequency configuration meanly considerably changes after about $t = 2\,s$ (the beginning of the voice B !). The spectrogram also shows, at higher frequencies, a very different pattern for $t > t^* \simeq 2s$ associated with the more acute female voice, furthermore it is evident around $t = 2\,s$ an high frequency event associated to the start of female speaking. However, we don't see trace of the background stationary – continuous noise at low frequencies while a quasi - continuous noise appears at high frequencies (around about 3.4 kHz).

In summary the spectrogram shows a remarkable difference in the time –frequency (at low and middle – high frequencies) between the first and the second part of the sample (that we know respectively associated to the male and female speaking), doesn't reveal the other noise components (as, for example the continuous monochromatic and the musical background) and doesn't suggest to us the presence, for $t > t^*$, of the same male voice dominating the first part of the sample.

3.2 Multiresolution analysis of the sound sample

The fig. 4, 5 and 6 and 7 show the time-scale images corresponding to WT of the studied signal, using Morlet, Gabor, Gauss and Franklin wavelet base atoms respectively. All the base atoms and are characterized by a frequency resolution of 1/12 of octave. On the x axis is represented the time interval of measurement normalized between 0 and 1, on the y axis the quantity $h = \log_2\left(1/a\right)$ related to the scale factor a while the image points intensity is proportional to WT coefficients. We note that the high scale components (low frequencies) are always characterized by good frequency resolution (and correspondingly poor time resolution) while low-scale components (high frequencies) have a good time resolution (and correspondingly poor frequency resolution).
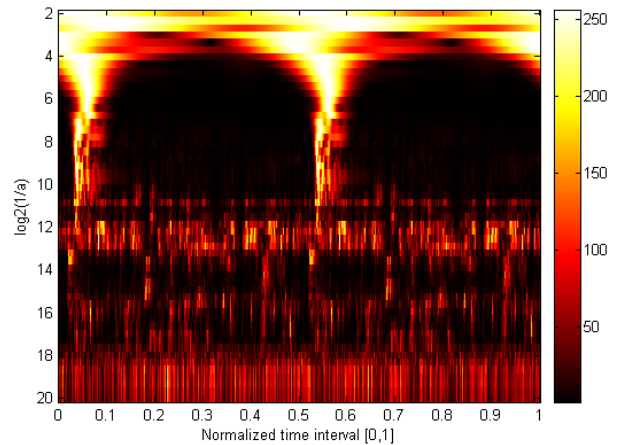

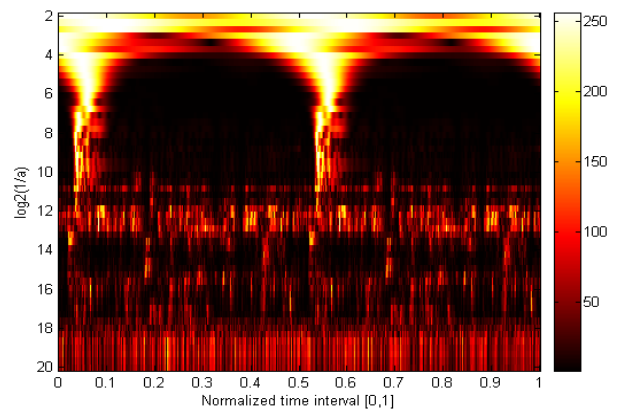Figure 4: WT multiresolution analysis of signal using Morlet base atoms.

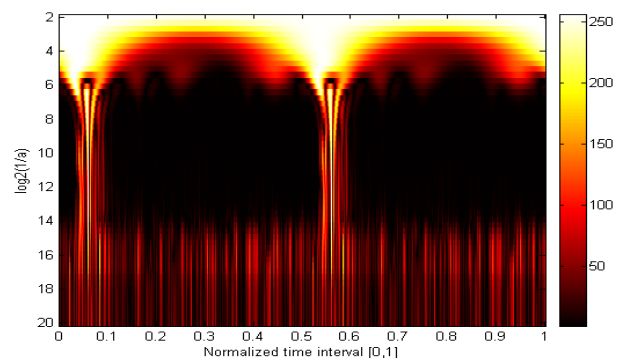Figure 5: WT multiresolution analysis of signal using Gabor base atoms.

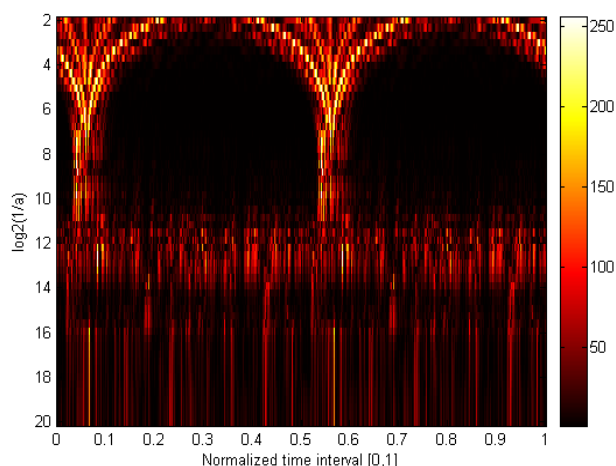Figure 6: WT multiresolution analysis of signal using Gauss base atoms.

Figure7: WT multiresolution analysis of signal using Franklin base atoms.

All the multiresolution analyses performed clearly show the subdivision of the time – scale pattern in two well distinct parts: the first one corresponding to the supposed activity of sources 1) + 2), the second one to the activity of the sources 2)+3). However, differently from "traditional" analyses, the WT shows the time – scale structure of the two part is substantially the same, suggesting an unexpected fact that is the noise sources active during the first part of the sample also work during the second one. Furthermore, a low frequency (high scale) and an high frequency (low scale) continuous patterns are now well evident in the graphics associated with the early monochromatic stationary – continuous noise source and to its high frequencies harmonics. Then, the multiresolution WT analysis has shown the presence, in the second part of the studied sample, of "hidden" noise sources, not revealed by "traditional" analyses, on the contrary simply showing a different time – frequency structure between these two parts of sample. In the multiresolution analysis the effect of the female voice is evidenced by a little increase in middle scale values in the second part of the WT image (for $t > \overline{t} \simeq 0.6$) that, however, doesn't practically alter the overall time – scale structure.

We can note, in conclusion, that some choices of wavelet base atoms are more suitable than others with respect the particular signal considered in this paper. In particular we note that Gauss and Franklin base atoms are less sensitive than the Morlet and Gabor ones to the time – scale signal variations in the specific case studied.

This shows, in general, that a suitable selection of the multiresolution base atoms is very important in order to ensure the representativeness of the analysis.

4 CONCLUSION

As we have seen from the above brief discussion a suitable choice of the parameters related to a multiresolution analysis based on WT allow us to properly analyse complex noise signals associated to the presence of many sources characterized by very different time –frequency (or equivalently time – scale) structures, making us able to distinguish the different noise events one from another, and deduce the relevant features of these noise sources. This ability can be very usefully exploited in forensic acoustic analysis where is often fundamental to identify the presence of different noise sources even when they are hidden and practically "invisible" to the traditional time – frequency and 1/n of octave analysis.

However more progresses must be still achieved about these multiresolution techniques in order to make them more specific and above all more versatile since they are very hardware consumptive and now practically usable only for very short signals.

References:
[1] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1999.
[2] Y. Meyer, *Wavelet Algorithms and Applications*, SIAM, Philadelphia, 1993.
[3] L. M. Caligiuri et al, *Time – scale and time - frequency analysis of complex noise scenarios by means of Wavelet and smoothed Wigner – Ville transforms*, Proceedings of INTER-NOISE 2003.
[4] L.M. Caligiuri et al., *Multiresolution analysis of noise generated by aircrafts during landing/take off operations*, Journal of the Acoustical Society of America, 114(4), 2341 (2003).
[5] L. M. Caligiuri et al., *Multiresolution analysis of environmental noise signals: an application to railways noise*, Proceedings of IMTC 2004, The 2004 Instrumentation and Measurement Technology Conference" IEEE, 18-20 May 2004 Como (Italy).