A simulation and restitution technique for the perceptive evaluation of road traffic noise

Julien Maillard, Jacques Martin
Centre Scientifique et Technique du Bâtiment (CSTB),
24 rue Joseph Fourier, 38400 Saint Martin d'Hères, France, e-mail: **julien.maillard@cstb.fr**

Noise pollution of future road infra-structures is generally presented to the public using sound pressure levels expressed in decibels. For the non-specialist, it is difficult to translate these numbers into actual sound levels and therefore, to evaluate the real impact of the road in terms of comfort. In addition, masking effects and other frequency dependent variations can be hard to reveal through global quantities such as averaged sound pressure levels. As a result, a tool providing a perceptive rather than quantitative evaluation of noise pollution is much needed. The technique presented here allows an accurate 3D audio restitution of traffic noise from road infra-structures. The restitution is performed in real time and therefore enables interactive audio navigation through the site as well as on-line changes of road source characteristics. This is usually coupled with a visual representation of the synthesized environment. The approach combines real traffic noise signals and numerical results from CSTB software Mithra which estimates sound levels at a number of receiver points through ray tracing algorithms. The paper discusses the various signal processing techniques involved in the audio restitution in order to reproduce accurately the estimated sound pressure levels within a realistic auralized environment.

1 Introduction

For the past few years, noise regulation policies introduced in the European Union [1] have pushed the development of new monitoring and prediction tools. In the field of ground transportation noise, research has provided methods for assessing the impact of existing road or train infrastructures as well as predicting the impact of future infrastructures. Noise-mapping softwares such as Mithra© give sound pressure levels at any given points within a site using ray tracing techniques [2, 3, 4, 5]. Other tools have also been developed to predict noise within micro-scale urban areas such as city streets [6, 7, 8]. In all these tools, sound pressure levels are expressed in decibels either in frequency bands or as a global level, e.g. A-weighting decibels, allowing for example to check compliance with existing noise regulations. However, the results can be difficult to interpret in terms of noise pollution, especially for the non specialist. As a result, the need for a tool providing a perceptive evaluation of traffic noise pollution has emerged.

With the progress in signal processing , audio 3D recording and restitution techniques, as well as ever increasing processor speed, it is now possible to generate realistic soundscape environments including user interaction such as walk-through motion [9, 10, 11, 12]. As an example, recent video games use such techniques. The main goal in this case is realism and interactivity. On another level, research in virtual acoustics aims at providing physically-based sound environments in which the listener should be exposed to the same stimuli as in the real environment. In this case, the restitution must use accurate sound emission and propagation models in addition to proper 3D audio restitution systems and signal processing techniques.

While much research has been done in the field of physically-based virtual soundscape inside rooms [13, 14], less effort is focusing on outdoor noise environments possibly due to the presence of more complex sound sources such as road traffic. Kang *et al.* [15] recently proposed an acoustic animation tool based on simple empirical formulae for the restitution of traffic noise and other common sources in micro-scale urban areas. In their work, the sources are considered fixed and the user evaluates the sound field at given locations.

The present paper proposes a new technique entirely based on simulation results pre-computed by CSTB software Mithra© and processing of specific traffic recordings. The developed software module works as an addition to Mithra© and allows realistic audio restitution of road noise traffic within the modelled site. The technique ensures that the restituted levels correspond to the predicted levels within each octave frequency band. The restitution is also performed in real time and therefore enables interactive audio navigation through the site as well as on-line changes of road source characteristics.

The paper first gives an overview of the proposed technique followed by a short description of the required sound recording data and Mithra© pre-computations. The techniques and algorithms implemented in the real time module are then presented along with the user interface.

2 General architecture

The general architecture of the audio restitution module is presented in Figure 1. The module uses two main sets of data files: audio files containing recordings of road traffic and fixed sources, and acoustic data files containing results from pre-computed simulations. As explained below, the audio files are obtained from specific on-site recordings of generic road traffic and additional fixed sources representing familiar sound events commonly found in urban environments. Each acoustic data file is the output of a Mithra© simulation performed on the project under study. This file contains the sound pressure levels per octave bands and sources over a grid of receivers. In addition, a text file defines the association between each acoustic source (road or fixed type) in the Mithra© project and a given set of audio files. Each sound source has a restitution driver unit which computes its restitution parameters based on the source geometry and associated sound pressure levels at the current listener position. The source restitution parameters update the signal processor unit which processes the input signals obtained from the associated audio files (through an audio file streamer) to construct the output signal. The following sections now give more details about each one of the above components.

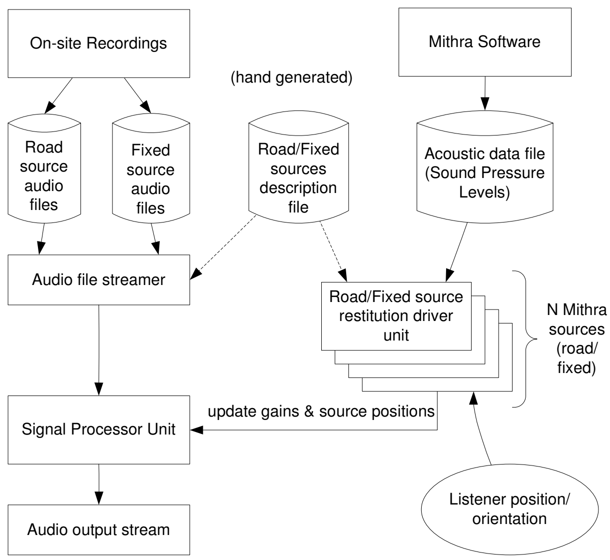

Figure 1: General architecture

A Mithra© project is constructed either manually or directly from generic format files defining the geometry of the site including the ground levels, buildings, and road geometry. Additional road or fixed sources may be added and configured by setting parameters such as road cover and traffic data. Horizontal maps are then computed using an automatically generated grid of receivers for various configurations, e.g., different road protection types. More details about the theory and algorithms implemented in Mithra© can be found in [2, 3, 4, 5]. The result data files contain the sound pressure levels per octave bands between 63 and 8000 Hz associated with each source over all receivers. These levels will be used by the source restitution driver unit to calibrate associated source audio signals. The grid of receivers uses a mesh of triangular cells, each receiver being located at the center of the cell. A table of adjacent cell indexes allows fast retrieval of the current receiver index as the listener moves from one cell to another.

The road traffic restitution technique is based on the construction of a realistic road traffic signal at any given point from a set of pre-recorded generic traffic signals. On-site traffic recordings use a set of stereo microphone pairs at several distances along a normal to the road. The recording site should offer free field conditions and very little environmental noise besides the road traffic noise. All signals are recorded simultaneously on a multi-channel digital interface. After proper calibration, the recorded signals of two adjacent recording distances will be mixed with proper gains to construct a signal which is a good representation of the actual traffic at the current road/listener distance. Note that to reduce the phase interferences due to the difference in propagation time between two recording points in the mixing process, the signals are time shifted in order to cancel the time delay when vehicles cross the measurement line, i.e., when the recorded level is maximum. Also, using stereo recordings preserve the perception of moving sources along a given direction in addition to better represent the spatial extent of the road. Ambisonic recordings could be used in a similar way but at the expense of an increased signal processing load. In addition to closely match the pre-computed sound pressure levels, the perception of the output signals must also correspond to the road type and traffic density under study. Therefore, several recording sessions must be performed on different sites and for various traffic conditions in order to best cover the modelled site. Unlike road sources, fixed sources require a single mono recording which should also be performed in free field conditions.

3 Real time module

The real time module includes all the software components represented below the system input data files in Figure 1. The algorithm implemented in the signal processor unit is now detailed (see Figure 2). The audio file streamer output contains the recorded signals associated with the sources as defined in the sources description file. Each road source includes N pairs of stereo signals and each fixed source is represented by a single mono signal. The frame containing all signals is processed by three

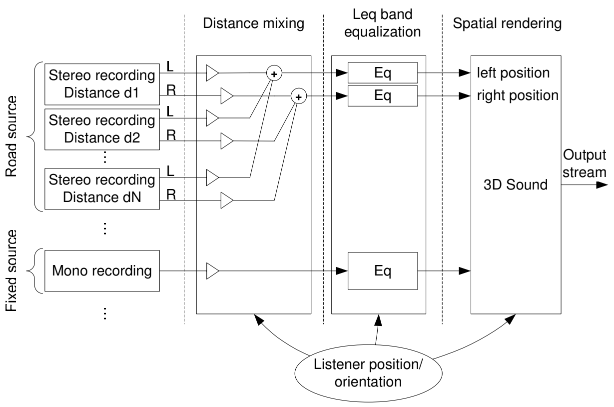

Figure 2: Real time module block diagram

successive algorithms: distance mixing, level equalization, and spatial rendering. The parameters controlling each of these three steps are updated in real time based on the listener position and orientation. The following sections describe in more details each step.

3.1 Distance mixing

For road sources, the left and right channels of the associated signals are multiplied by a linear gain depending on the current listener/road distance, d. The source restitution driver unit first retrieves d as the shortest distance from the listener location to the road segments. It then sets the mixing gains α_i according to:

$$\alpha_i = 1 - \beta,\ \alpha_{i+1} = \beta,\ \beta = \frac{d - d_i}{d_{i+1} - d_i} \qquad (1)$$

if $d_i \leq d \leq d_{i+1}$, $\alpha_1 = 1$ if $d \leq d_i$, and $\alpha_N = 1$ if $d_{i+1} \leq d$, all other gains being set to zero. As explained earlier, the mixer output signal is a weighted sum of the signals recorded on each side of the current distance which tends toward the actual recorded signal as the distance approaches one of the recording distances or moves past the minimum or maximum value.

For fixed sources, the single gain may be used to compensate for the difference in geometric attenuation between the current receiver-source and listener-source distance. If d_r and d_l denote the receiver-source and listener-source distance, respectively, the linear gain is set to $\alpha = d_r/d_l$.

3.2 Level equalization

The equalization module introduces a gain in each octave frequency band such that the signal levels of a source match the pre-computed levels. Each equalization unit in Figure 2 includes a bank of filters to separate the signal in octave frequency components followed by a set of linear gains. The modified components are then summed together to form the equalized signal. The bank filters are all IIR filters designed such that the global response is unity across the frequency bandwidth when all gains are set to one, i.e., the filter bank conserves the energy and do not introduce frequency coloration in the signals.

The source restitution driver unit updates the associated equalizer gains according to the following procedure. For road sources, the two equalizer units processing the left and right channels have identical gains. The gain in a given frequency band expressed in dB, α_{dB}, is the difference between the averaged sound pressure level computed by Mithra©, Leq_m, and the averaged level of the recorded audio signals after distance mixing, Leq_s. The source driver unit first retrieves the Mithra© levels at the current receiver. It then estimates the levels of the recorded signals after distance mixing based on the current distance mixing gains. In the general case where two stereo signals $x_1^L(n)$, $x_1^R(n)$, and $x_2^L(n)$, $x_2^R(n)$, respectively, are mixed together, i.e., the listener is within the recording distances, the resulting sound pressure signal, $y(n)$, at the listener location is

$$\begin{aligned} y(n) \;=\; & \alpha_1 \left(x_1^L(n) + x_1^R(n)\right) \\ & + \alpha_2 \left(x_2^L(n) + x_2^R(n)\right) \end{aligned} \qquad (2)$$

where α_1 and α_2 denote the mixing gains. The associated power level, Ψ_y^2, is obtained as

$$\begin{aligned} \Psi_y^2 \alpha_1^2 & \left(\Psi^2_{x_1^L} + \Psi^2_{x_1^R} + 2\Psi_{x_1^L x_1^R}\right) \\ + \alpha_2^2 & \left(\Psi^2_{x_2^L} + \Psi^2_{x_2^R} + 2\Psi_{x_2^L x_2^R}\right) \\ + 2\alpha_1\alpha_2 & \left(\Psi_{x_1^L x_2^L} + \Psi_{x_1^R x_2^R} + \Psi_{x_1^L x_2^R} + \Psi_{x_2^L x_1^R}\right) \end{aligned} \qquad (3)$$

where the auto- and cross-correlation terms are defined by $\Psi_y^2 = \mathrm{E}[y^2(n)]$ and $\Psi_{y_1 y_2} = \mathrm{E}[y_1(n) y_2(n)]$, respectively. The averaged sound pressure level in dB, re 20 μPa is given by $Leq_s = 10\log\left(\Psi_y^2/4 \times 10^{-10}\right)$. In practice, all auto- and cross-correlation terms of two adjacent recording signals are pre-estimated at startup. The gain update therefore only involves a few additions and multiplications (see Equation (3)). For fixed sources, the recorded signal level reduces to a single term, $\Psi_y^2 = \alpha^2 \Psi_x^2$.

Note that the recorded signal levels may be adjusted when the traffic density over the estimation period does not match the traffic density set in Mithra©. If T is the estimation period including N vehicles, the estimated levels then become

$$Leq_{T,m} = Leq_T + 10\log\left(T/T_m\right) \qquad (4)$$

where T_m represents the period over which N passing vehicles will yield the desired traffic density.

Most of the signal processing power required by the above level equalization is included in the bank filtering separating the signals in octave frequency components. When running on slower systems, an alternative equalization method can be selected, which pre-filter all audio signals prior to real time operation. In this case, the total number of channels returned by the audio file streamer is multiplied by 8 which increases the memory requirements (all signals can be pre-loaded in memory to avoid additional load from running disk access threads) but the resulting CPU load is decreased as the equalization now reduces to simple gain multiplications.

In order to validate the above equalization procedure, the averaged sound pressure levels were estimated from the output signal of the restitution module. The spatial rendering mode (see Section 3.3) was set to mono restitution in which case the left and right channels of each road source are simply summed together. For a given listener position, the estimated levels were compared to the desired levels pre-computed by Mithra© at the associated receiver. As an example, Figure 3 shows the comparison in the case of two road sources. The top two graphs correspond to each individual source levels and the bottom graph to the overall levels. At the chosen location, the distance mixing involved two recorded signals with respective gains -22.3 dB and -0.7 dB for the first road (top graph) and -5 dB and -7.2 dB for the second road (middle graph). Excellent agreement between restituted signal and Mithra© levels is obtained in all frequency bands except for the first and last band, as well as for the overall weighted dB(A) level. The differences at both end bands is explained by the finite roll-off of the band filters which, combined with the large gain difference between these two bands and their neighbors (almost 30 dB), yields an overall filter response which does not quite follow the desired response at these frequencies. Precisely, the recorded signals contain energy in the low and high frequency bands which does not match the power spectrum of the traffic sources defined by Mithra©. As a result, the equalization gain in this two bands is high compared to the value in the neighboring bands. Nonetheless, the good overall agreement validates the band filter design as well as the gain adjustment procedure based on both auto- and cross-correlation terms. It should be noted that neglecting the cross-correlation between left and right channels but also between signals at two adjacent recording distances results in significant errors in the output signal levels.

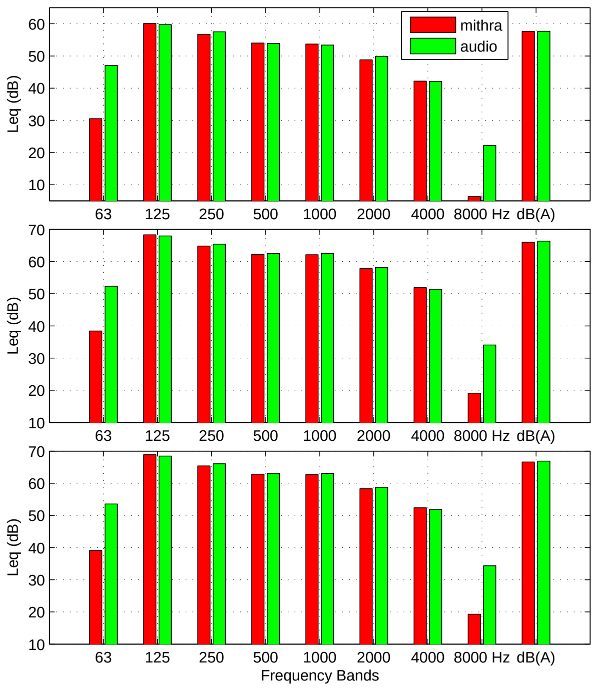

Figure 3: Predicted and restituted Leq levels

3.3 Spatial rendering

The last stage in the signal processor unit shown in Figure 2 is the 3D restitution module. The input signals to the 3D module include all equalized channels of the active sources, i.e., 2 channels per road source and 1 channel per fixed source. The module can be configured at run time for different restitution systems: in addition to the standard mono and stereo restitution, binaural [16, 17] and ambisonic [18] restitutions are also available. Binaural output implements an efficient filtering technique based on the decomposition of the Head Related Transfer Functions (HRTF) using Independent Component Analysis.

In addition to computing distance mixing and equalizer gains, each source restitution driver unit calculates the source position relative to the current listener location based on its geometry.

For road sources, the direction from the listener to the closest point on the road, i.e., along the normal to the closest road segment, is calculated. If θ_r represents the angle between a reference axis and this direction, the source left and right signals are positioned symmetrically on each side in directions $\theta_r - \theta_h + \Delta(d)$ and $\theta_r - \theta_h - \Delta(d)$, respectively. In the expression above, θ_h is the head orientation and $2\Delta(d)$ is the angle between the left and right sources. This angle is a function of the distance to the road: it decreases as d increases thereby rendering the road source apparent width.

For fixed sources, the source signal is positioned along the direction from the listener to the source location. Including the head relative orientation, the actual source angle is $\theta_f - \theta_h$ where θ_f is the angle between the listener-source direction and the reference axis.

For both road and fixed sources, the overall gain of the 3D processing unit is one. Indeed, all geometric attenuation is included in the gain equalization stage prior to 3D positioning according to Mithra© pre-computed levels.

Also, note that all gains and source positions are updated using incremental steps over a number of frames, e.g., 10 ms, in order to avoid audible artifacts related to large gain or delay changes between 2 consecutive samples.

4 User interface

The module is designed such that the user can evaluate different scenes under different traffic conditions and at different locations. Real time processing enables the user to switch from one configuration to another at will and makes comparative evaluations more reliable. The user interface includes controls to load a given scene, (Mithra© project file including the site geometry, the roads geometry, roads protections, and weather data), define the traffic related data of a given road and move within the site (translation and rotation).

At the time of this writing, loading a scene is not immediate due to the time required to evaluate source signals energy (see Section 3.2). In future versions, the user will be able to pre-load several scenes at once and switch from one to another immediately.

After selecting a road, associated traffic data may be modified during restitution through a specific interface window including all traffic parameters available in Mithra© (traffic density, road cover, percentage of heavy vehicles, averaged speed, or direct power spectrum per source unit length). This updates the source power spectrum and in turn the sound pressure levels per octave bands over all receivers and finally, the equalizer gains at the current listener location.

Two types of navigation interface are available: simple 2D view from the top using a mesh like representation of the site and fully lighted 3D model of the site.

The first type (see Figure 4) shows the building contours, the road segments, as well as the receiver points and associated triangular cells. Fixed sources are also shown in addition to the current listener position and orientation. The user can zoom in and out for precise positioning with the mouse device. This type of interface uses very little CPU cycles. However, it lacks a realistic visual output.

The second navigation interface (see Figure 5) is based on simulations performed by CSTB software *Phanie*. This tool generates a fully lighted 3D model of the site using photo-realist image synthesis techniques including precise definition and rendering of complex light sources. A specialized interactive renderer, coupled to the audio restitution module, is then used to navigate within the model. In this interface, the user can navigate either in “subject mode”, i.e., the camera is the listener, or in “third person mode”, i.e., the camera is a third person as shown in Figure 5. The visual output is performed on a stan-

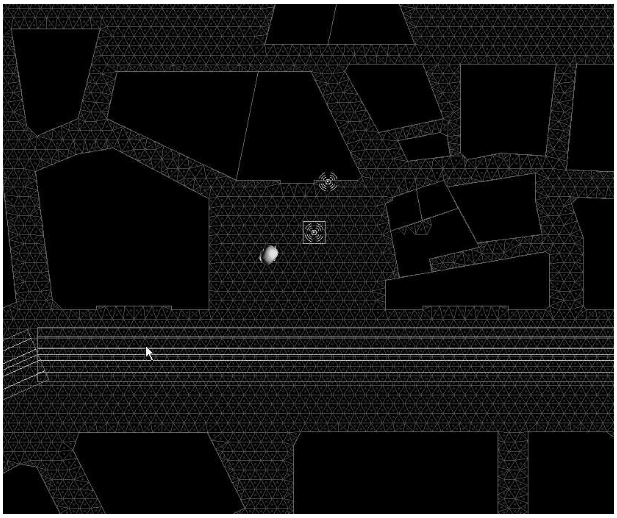

Figure 4: Prototype interface showing sources, buildings, and receivers

Figure 5: Picture of fully-lighted navigation interface

dard monitor or using stereo projection for 3D visualization. This type of output is implemented in CSTB Virtual Reality Center, “Le Corbusier”, in Sophia-Antipolis, France. The room is equipped with a conical projection screen (4.6 m by 2.8 m), Ambisonic audio system, as well as, CSTB transaural restitution seats allowing interactive audio and visual rendering of 3D scenes.

5 Conclusions

The audio restitution tool presented in this paper implements a novel approach for the perceptive evaluation of road traffic noise under different conditions. The tool works as an addition to CSTB simulation software Mithra© which provides pre-computed maps of sound pressure levels. Specific recording and processing techniques allow 3D audio restitution of traffic noise with accurate sound levels following predicted levels in each octave band. In addition, real time processing enables the user to navigate within the modelled site as well as modify sources characteristics in order to compare various traffic conditions.

The current limitations of the proposed tool lie in the large recording data base required to render realistically different types of traffic noise with, e.g., varying density or road cover type. The technique ensures that restituted sound levels always match predicted levels. However, it will fail to render other traffic noise signal characteristics specifically related to precise conditions, unless such conditions have been recorded using the proposed technique. Current work includes carrying on further recordings of traffic conditions as well as developing efficient classification and access methods to automatically select the proper signals based on the chosen traffic parameters.

6 Acknowledgements

The authors wish to thank Pascal Ducruet and Pierre Bruder who contributed to the traffic noise recordings.

References

[1] Proposal for a directive of the european parliament and of the council relating to the assesment and management of environmental noise. Technical report, The European Commission, 2000.

[2] *Mithra - Manuel technique, CSTB - 01dB*, 1999.

[3] J. Defrance, M. Bérengier, and J.F. Rondeau. Validation and evolution of the road traffic noise prediction method NMPB-96. Part 2: Improvements based on theoretical models. In *Inter-Noise'2001, Den Haag, The Netherlands*, 2001.

[4] Yannick Gabillet. Une méthode inverse de recherche de rayons: le logiciel MITHRA. In *Cahiers du CSTB, Cahier 2444*. CSTB, 1990.

[5] J. Defrance and Y. Gabillet. A new analytical method for the calculation of outdoor noise propagation. *Applied acoustics*, 57:109–127, 1999.

[6] D. J. Oldham and M. M. Radwan. Sound propagation in city streets. *Building Acoustics*, 1:65–87, 1994.

[7] J. Picaut, L. Simon, and J. Hardy. Sound field modelling in streets with a diffusion equation. *Journal of the Acoustical Society of America*, 106:2638–2645, 1999.

[8] J. Kang. Numerical modelling of the sound field in urban streets with diffusely reflecting boundaries. *Journal of Sound and Vibration*, 258(5):793–819, 2002.

[9] J. Blauert. *Spatial hearing, the psychophysics of human sound localisation*. MIT Press, 1997.

[10] M. Kahrs. *Applications of Digital Signal Processing to Audio and Acoustics*. Kluwer Academic Publishers, Boston, 1998.

[11] M. Kleiner, B. I. Dalenbäck, and P. Svensson. Auralization - An overview. *Journal of the Audio Engineering Society*, 41:861–875, 1993.

[12] J. M. Jot. Real-time spatial processing of sounds for music, multimedia and interactive human-computer interfaces. *Multimedia Systems*, 7(1):55–69, 1999.

[13] Andreas Silzle, Pedro Novo, and Holger Strauss. IKA-SIM: A system to generate auditory virtual environments. In *Audio Engineering Society-116th Convention May 8-11 Berlin, Germany*, 2004.

[14] Thomas Funkhouser, Nicolas Tsingos, Ingrid Carlbom, Gary Elko, Mohan Sondhi, James E. West, Gopal Pingali, Patrick Min, and Addy Ngan. A beam tracing method for interactive architectural acoustics. *The Journal of the Acoustical Society of America*, 115(2):739–756, 2004.

[15] J. Kang, Y. Meng, and G. J. Brown. Sound propagation in micro-scale urban areas: simulation and animation. In *EuroNoise, Naples, Italy*, 2003.

[16] M. Emerit, E. Dudouet, and J. Martin. Head related transfer function and high order statistics. In *15th International Congress on Acoustics*, pages 437–440, Trondheim, Norway, 1995.

[17] Eric Dudouet and Jacques Martin. A new HRTF decomposition and algorithm for sound spatialization. In *ASA/DAGA Joint Meeting, Berlin, 14-19 March*, 1999.

[18] Jérôme Daniel. *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*. PhD thesis, Thèse de doctorat de l'Université Paris 6, 2000.