

Research Article

Real-Time Perceptual Simulation of Moving Sources: Application to the Leslie Cabinet and 3D Sound Immersion

R. Kronland-Martinet and T. Voinier

Laboratoire de Mécanique et d'Acoustique, CNRS, 31 Chemin Joseph Aiguier, 13402 Marseille Cedex 20, France

Correspondence should be addressed to R. Kronland-Martinet, kronland@lma.cnrs-mrs.fr

Received 31 October 2007; Accepted 29 May 2008

Recommended by Sen M. Kuo

Perception of moving sound sources obeys different brain processes from those mediating the localization of static sound events. In view of these specificities, a preprocessing model was designed, based on the main perceptual cues involved in the auditory perception of moving sound sources, such as the intensity, timbre, reverberation, and frequency shift processes. This model is the first step toward a more general moving sound source system, including a system of spatialization. Two applications of this model are presented: the simulation of a system involving rotating sources, the *Leslie Cabinet* and a 3D sound immersion installation based on the sonification of cosmic particles, the *Cosmophone*.

Copyright © 2008 R. Kronland-Martinet and T. Voinier. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

The simulation of moving sources is of great importance in many audio sound applications, including musical applications, where moving sources can be used to generate special effects inducing novel auditory experiences. Motion of instruments while they are being played can also subtly affect the sound, and hence the expressiveness of the performance. Wanderley et al. [1] have described, for example, that the motion of the clarinet follows specific trajectories depending on the type of music played, independently of the player. Although the effect of this motion on sound has not yet been clearly established, it probably contributes to the rendering and should be taken into account in attempts to synthesize musical sounds. Virtual reality is another field, where moving sources play an important role. To simulate motion, the speed and trajectories are crucial to creating realistic acoustical environments, and developing signal processing methods for reconstructing these contexts is a great challenge.

Many authors have previously addressed these problems. Two main approaches have been used so far for this purpose: the physical approach, where sound fields resembling real ones as closely as possible are simulated, and the perceptual

approach, where the resulting perceptual effects are taken into account.

The physical approaches used so far in this context have involved modelling sound fields using physical models based on propagation equations. In this case, the distribution of the acoustical energy in the 3D space requires a set of fixed loudspeakers precisely- and accurately-controlled. Several techniques such as ambisonics [2], surround sound [3] and, more recently, wave field synthesis [4], and VBAP [5] have been developed and used in studies on these lines. Specific systems designed for headphone listening have also been developed [6], which involve filtering signals recorded under anechoic conditions with head-related transfer functions (HRTFs). However, the specificity of individual HRTF gives rise to robustness issues, which have not yet been solved. In addition, it is not clear how a system of spatialization may be suitable for simulating rapidly moving sound sources, since they do not take the dynamics of the source into account. Lastly, Warren et al. [7] have established that different brain processes are responsible for mediating static and dynamic moving sounds, since the perceptual cues involved were found to differ between these two categories of sounds.

The perceptual approaches to these issues have tended to focus on the attributes that convey the impression

that sounds are in motion. Chowning [8], who conducted empirical studies on these lines, established the importance of specific perceptual cues for the synthesis of realistic moving sounds.

In the first part of this paper, the physical and perceptual approaches are combined to develop a real-time model for a moving source that can be applied to any sound file. This model, which was based on Chowning's studies, was calibrated using physical knowledge about sound propagation, including air absorption, reverberation processes, and the Doppler effect. The second part of this paper deals with two audio applications of this system. The first application presented is the *Leslie cabinet*, a rotating source system enclosed in a wooden box, which was modelled by combining several moving sound elements to simulate complex acoustic phenomena. In this application, we take the case of a listener placed far from the sound sources, which means that the acoustic environment greatly alters the original sound. The second application focuses on a virtual reality installation combined with cosmic particle detectors: the *Cosmophone*. Here, the listener is immersed in a 3D space simulating the sonified trajectories of the particles.

2. WHAT IS PERCEPTUALLY RELEVANT?

Based on previous studies (see, e.g., [9] and the references therein, [8, 10–16]), four important perceptual cues can be used to draw up a generic model for a moving sound source. Most of these cues do not depend on the spatialization process involved, but they are nevertheless greatly influencing the perception of sounds, including those emitted by fixed sources.

Sound pressure

From the physical point of view, the sound pressure relates to the sound intensity, and in a more complex way, the loudness. Sound pressure varies inversely with the distance between the source and the listener. This rule is of great importance from the perceptual point of view [15], and it is possibly decisive in the case of slowly moving sources. It is worth noting that only the relative changes in the sound pressure should be taken into account, since the absolute pressure has little effect on the resulting perceptual effect.

Timbre

Timbre is a perceptual attribute which makes it possible to discriminate between different sounds having the same pitch, loudness, and duration [17]. From a signal processing point of view, timbre variations are reflected in changes in both the time evolution and the spectral distribution of the sound energy. Subtle changes of timbre can also make it possible to distinguish between various sounds belonging to the same class. For example, in the class consisting of impact sounds on geometrically identical bars, it was established in a previous study that it is possible to differentiate perceptually between various wood species [18].

Changes in the timbre of moving sound sources, which are physically predictable, play an important perceptual role. Composers such as Maurice Ravel used cues of this kind in addition to intensity variations to make a realistic sensation of an an-coming band in his *Bolero*: the orchestra starts in a low-frequency register to simulate the band playing at a distance, and the brightness gradually increases to make the musicians seem to be coming closer. Schaeffer [10] also used changes of timbre in a radiophonic context to simulate auditory scenes, where the speakers occupied different positions in the virtual space.

The changes of timbre due to distance can be accounted for physically in terms of air absorption. The main perceptual effect of air absorption on sounds is due to a low-pass filtering process, the result of which depends on the distance between source and listener. Note that, under usual conditions, the 0–3 kHz frequency band, in which most human communications occur varies very little, even at large source-to-listener distances. To simulate moving sound sources which cover large distances, effects due to air absorption must be taken into account.

The doppler effect: a frequency shift

From the physical point of view, moving sound sources induce a frequency shift known as the Doppler effect. Actually, depending on the relative speed of the source with respect to the listener, the frequency ω_l measured at the listeners position is [19]

$$\omega_l = \omega_s \left(\frac{1 + v_{ls}/c}{1 - v_{sl}/c} \right), \quad (1)$$

where ω_s is the frequency emitted by the source, v_{ls} and v_{sl} denote the relative speed of the listener in the direction of the source and the relative speed of the source in the direction of the listener, respectively, and c is the sound velocity. During a given sound source trajectory, the perceived frequency is time-dependent and its specific pattern seems to be a highly relevant cue enabling the listener to construct a mental representation of the trajectory [15]. Chowning [8] used such a pattern to design efficient signal processing algorithms accounting for the perception of moving sources. It is worth noting here that the Doppler effect integrates changes in intensity as well as the frequency shifts. The perceptual result is, therefore, a complex combination of these two parameters, since an increase in the intensity tends to be perceived as a pitch variation due to the close relationship between intensity and frequency [13]. The Doppler effect is a dynamic process, which cannot be defined by taking motion to be a series of static source positions, and this effect is robust whatever the system of spatialization uses, including fixed mono speaker diffusion processes.

Environment: the effects of reverberation

In everyday life, quality of sound depends on the environment. Scientists and engineers working on room acoustics (see, e.g., [11]) have studied this crucial issue intensively. The influence of the environment is a complex problem,

and modelling sounds taking architectural specificities into account are not the scope of this study. In particular, the effects of reverberation can be explained by the physical laws of sound propagation, which impose that distant sound sources lead to more highly reverberated signals than nearby sound sources because with distant sound sources, both the direct and reflected sound paths are of similar orders of magnitude, whereas with nearby sources, the direct sound is of greater magnitude than the reflected sounds. Moving sound sources, therefore, involve a time-dependent direct-reverberated ratio, the value of which depends on the distance between source and listener.

2.1. A real-time moving source model

In line with the above considerations, a generic model was drawn up simulating the motion of an acoustic source by processing a sound file corresponding to the acoustic radiation emitted by a fixed source. This model consists of a combination of the four main components described above (Figure 1). The relative speed and distance between the listener and the moving source control the parameters of the model. Efficient interfaces can, therefore, be added to simplify the modelling of the trajectories. The resulting sound is intended for monophonic listening, but it could be linked to a system of spatialization, enhancing the realism of the motion.

2.2. Implementation

We describe how each elementary process can be modelled algorithmically. The global implementation scheme is shown in Figure 3. The whole model was implemented in real time under Max/MSP [20] development environment. The implementation, which can be downloaded on the web (see Section 6), allowed to check the perceptual accuracy of the model.

2.2.1. Intensity variations

Intensity variations are controlled directly by the level of the sound. Assuming the sound propagation to involve spherical waves, the sound level will vary with respect to $1/x$, where x is the source-to-listener distance. From the practical point of view, care must be taken to avoid divergence problems at $x = 0$.

2.2.2. Timbre variations

As mentioned above, timbre variations due to the air absorption mainly affect the high-frequency components. Since this factor is probably of lesser perceptual importance than other motion cues, it is possible to simplify its treatment in the implementation process. Huopaniemi et al. [12] have established that the magnitude response of the low-pass filter accounting for air absorption can be modeled using low-pass IIR filters. The frequency response of these filters must vary with respect to the listener-to-source distance. However, no information seems to be available in the literature giving

cues as to how accurately these filters have to be designed to ensure the realism of the distance perception. We, therefore, designed a model based on a compromise between perceptual accuracy and real-time performance. This constraint actually requires the number of control parameters (the so-called “mapping”) as well as the algorithmic complexity to be minimized. A classical high-shelving second-order IIR filter was used as described in [21] to model the timbre variations due to the air absorption. This kind of filter, which was originally designed for parametric equalizers, makes it possible to either boost or cut off the high-frequency part of the audio spectrum. To simulate air absorption, the control parameters (cutoff frequency and gain) have to be linked to the listener-to-source distance. At a given listener-to-source distance x , one “air transfer function” $A(f)$ can be computed using formulae given in [22]. An optimization procedure, based on a least square minimization method, then gives the gain and cutoff frequency minimizing $\| |A(f)| - |H(f)| \|^2$, where $H(f)$ is the transfer function of the high-shelving filter. Since the cutoff frequency was found to depend weakly on the distance, it was set to 10 kHz. This led to a single control parameter: the gain G . Furthermore, this gain in dB can be related to the distance x in meters via the simple relation:

$$G(\text{dB}) \simeq -0.5x \text{ (m)}. \quad (2)$$

The computed air transfer functions and the simulated filter magnitude responses are compared in Figure 2 at distances up to 50 meters, with the parameters given above. Although the simulation differs from reality (especially in the high-frequency range), it yielded to perceptually satisfactory results. In addition, the factor -0.5 , applied between the filter gain and the source-to-listener distance, can be changed, so that the effects of timbre variations can be easily adjusted (increased or decreased).

2.2.3. Doppler frequency shift

The Doppler frequency shift is due to changes in the path length between source and listener, and hence to changes in the propagation time, $\tau(t)$. The Doppler frequency shift (1) can then be controlled by a variable delay line. In the case of a sound source emitting a monochromatic signal and moving with respect to a fixed listener, Smith et al. [23] obtained the following expression:

$$\frac{d\tau(t)}{dt} = -\frac{v_{sl}}{c}. \quad (3)$$

For a given trajectory, (e.g., in the case of a source moving along a straight line and passing in front of the observer), the source velocity projected onto the source-to-listener line can be precalculated at each time sample. The delay value can then be computed as a function of time. However, when the source trajectory is unpredictable, derivative of the delay can be used as in (3). Strauss [24] suggested approximating complex trajectories as linear piecewise curves in order to obtain an analytical solution of $\tau(t)$.

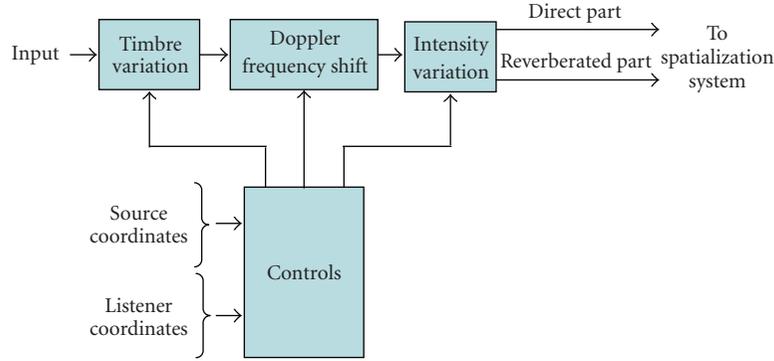


FIGURE 1: Scheme of the moving source model.

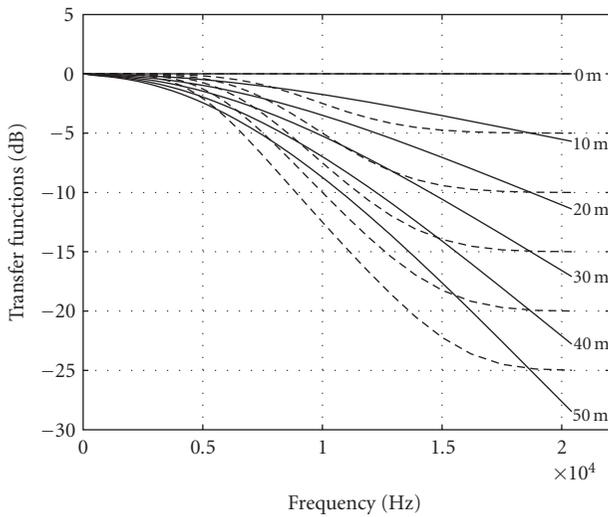


FIGURE 2: Air transfer functions (solid lines) and simulated filter transfer functions modules (dotted lines) obtained by optimization for various source-to-listener distances. Air transfer functions were computed with a temperature of 20°C, an atmospheric pressure of 1013 hPa, and 45% hygrometry. The cutoff frequency of the simulated filter was set at 10 kHz, and the filter gain was computed using (2).

Here, we adopted the approach proposed by Tsingos [25] who gave the following expression for $\tau(t)$:

$$\tau(t) = \frac{1}{c} \|L(t) - S(t - \tau(t))\|, \quad (4)$$

where $L(t)$ and $S(t)$ are the respective positions of the listener and the source at time t , and $\|\cdot\|$ denotes the Euclidian distance. This expression was simplified in our implementation, since similar perceptual effects were still obtained, even at source speeds of 100 km/h,

$$\tau(t) = \frac{1}{c} \|L(t) - S(t)\| = \frac{x(t)}{c}. \quad (5)$$

Note that the delay line must deal with fractional values of τ . This problem has been previously addressed (see, e.g., [26]).

2.2.4. Reverberation effect

Reverberation depends on the local environment and its treatment is usually left to the user. However, a few reverberation archetypes can be defined. In line with Chowning [8], we split the reverberation into its *global* and *local* components. The global reverberation originates from the whole space, whereas the local reverberation originates from the direction of the source. Actually, as Chowning stated, this corresponds to a fair approximation of a real acoustical situation, where the increase of the distance between the listener and the sound source leads to a decrease of the distance between the source and the reflecting surfaces, giving the reverberation some direction emphasis. The global reverberation level can be defined as $1/(x\sqrt{x})$, and the local reverberation level is given by $(1/\sqrt{x})(1 - (1/x))$. This ensures the following:

- (i) the sum of global and local reverberation levels varies as $1/\sqrt{x}$;
- (ii) the ratio between the global reverberation level and the direct sound level varies as $1/\sqrt{x}$.

The modelling of the effects of reverberation can be enhanced with specific systems of spatialization. Actually, in the case of multiple speaker arrays, the global reverberation should be equally distributed to all the speakers, while the local reverberation follows the moving source. This method has been found to greatly improve the realism of the perceptual effects simulated.

3. A LESLIE CABINET SIMULATOR

3.1. The Leslie cabinet

The Leslie cabinet is an interesting application of the moving sound source model. Originally designed to add choral effect to Hammond organs, Leslie cabinets have been successfully used as an effect processor for many other musical instruments [27]. A Leslie cabinet is a wooden box, containing a rotating horn radiating high frequencies and a rotating speaker port adapted to a woofer radiating low frequencies. Each rotating source is driven by its own motor and mechanical assembly, and the rotating speeds

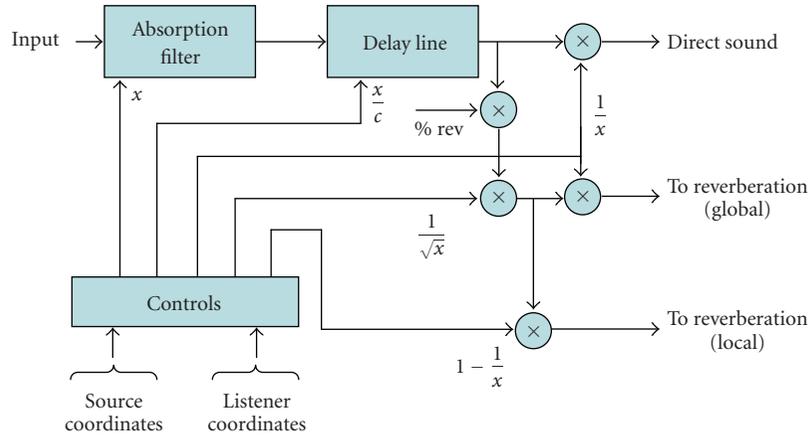


FIGURE 3: Implementation of the moving source model.

of the sources are, therefore, all different. The crossover frequency of this two-way speaker system is about 800 Hz. A diffuser is mounted at the end of the horn to approximate an omnidirectional pattern of radiation. The box is almost completely closed and contains only the vents from which the sound radiates. The rotating speed of the horn is fast enough to obtain pitch and amplitude modulations due to the Doppler effect. In the woofer port, the frequency modulation is assumed not to be perceptible [27], the main perceptual effect is the amplitude modulation. In addition to these effects, the rotation of both low- and high-frequency sources results in time-dependent coupling with the room, creating a particular spatial modulation effect.

Smith et al. [23] investigated the Leslie effect, focusing mainly on the simulation of the sound radiated by the rotating horn. In this study, the authors concluded that under free field conditions, without the box, far from the rotating source, both the Doppler frequency shift and the amplitude modulation are likely to be almost sinusoidal. They also stated that the reflections occurring inside the wooden cabinet should be taken into account when simulating Leslie effects.

3.2. Measurements

To assess the perceptual effects of these factors, measurements were performed on a model 122A Leslie cabinet (Figure 4). The cabinet was placed in an anechoic room and driven by a sinusoidal generator. The acoustic pressure was measured using a microphone placed 1.2 m from the cabinet, at the same height from the floor as the rotating plane of the horns.

From the signal recorded, $s(t)$, the analytic signal [28], given by $Z(t) = s(t) + iH[s](t) = A(t)e^{i\phi(t)}$, (where H denotes the Hilbert transform operator) was calculated in order to deduce both amplitude $A(t)$ and instantaneous frequency $d\phi/dt$ modulation laws.

The middle panel in Figure 5 shows the amplitude modulation law of the signal obtained with a 800 Hz input signal. The bottom panel shows the frequency modulation

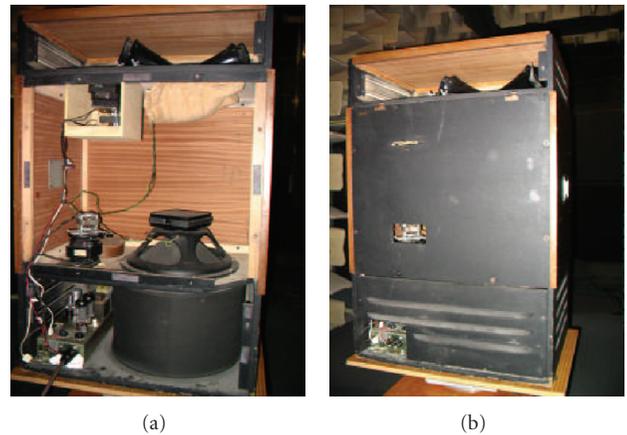


FIGURE 4: View of the 122A Leslie cabinet (open and closed) used for our measurements.

law of this signal. The instantaneous frequency showed a typical pattern, where the high-positive and negative peaks occur synchronously with a quasizero time amplitude signal. Patterns of this kind have been observed in situations where, for example, the vibrato of a singing voice is perturbed due to the room acoustics [29]. To determine the origin of these components, additional measurements were performed using sinusoidal input signals driving the horn alone. In this case, the interference was still observed, which means that radiation interference due to the woofer and the horn alone did not account for the complexity of the modulations. Other sound sources due to the enclosure, therefore, have to be taken into account in Leslie cabinet modeling procedures.

3.3. Implementation

The moving sound source model makes it easy to use the well-known image method [30] to account for the box wall reflections in the simulation procedure. The coordinates of the image sources can easily be deduced from the geometry of the cabinet, that is, the coordinates of the directly

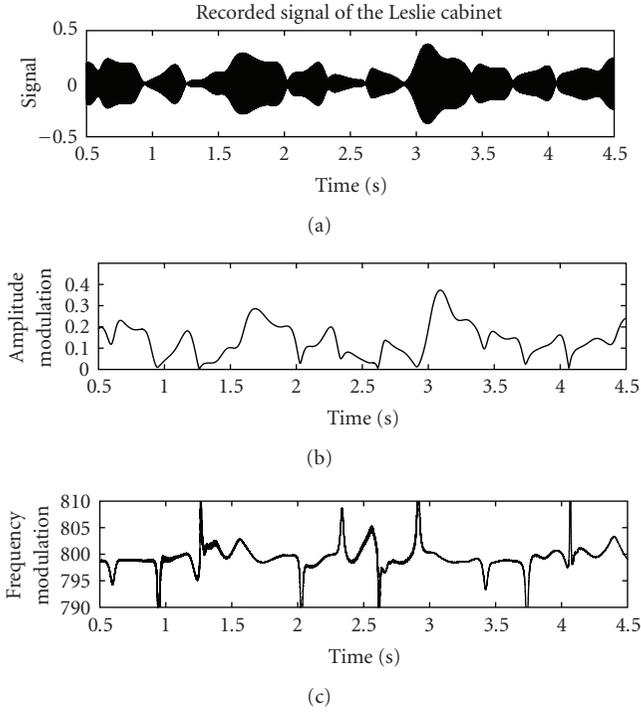


FIGURE 5: Analysis of the acoustical output signal from the Leslie cabinet driven with a 800 Hz sinusoidal input signal. Both the woofer and the horn have been activated. (a) microphone signal, (b) amplitude modulation, (c) frequency modulation.

radiating source and those of the reflecting planes. Since the computational complexity of the image method increases exponentially with the number of reflections taken into account, perceptual assessments were performed to estimate the minimum number of source images required. It was concluded that one image source for each reflecting plane (first order) sufficed to obtain satisfactory perceptual results.

The implementation of the Leslie horn simulator is shown in Figure 6. The sound produced by the horn is composed of the sum of the direct sound source and the five image sources (the back wall of the horn part of our cabinet was removed). Each source was processed using the moving source model. In addition, the signals injected into the moving image source models were filtered to account for the frequency-dependent sound absorption by the wood material. The wood absorption filter was an FIR filter and its impulse response was based on wood absorption data available in the literature [31]. The same procedure was used for the woofer simulator. As in the real Leslie cabinet, crossover filtering of the input signal gives the input to both the woofer and the horn simulators. It is worth noting that to obtain a more realistic simulation of the Leslie cabinet, the distortion due to the nonlinear response of the Leslie tube amplifier has to be taken into account.

3.4. Results

To assess the perceptual quality of the model, listening tests have to be run. In addition, these tests should be

entrusted to musicians experienced with the use of the Leslie cabinet manipulation. Nevertheless, to check the accuracy of the model, the main characteristics of the simulated signal obtained can be compared with the recorded one. For this purpose, we fed the model with a sinusoidal input signal with a frequency of 800 Hz (the crossover frequency) in order to include the effects of both the horn and the woofer. When the images source part was not active, the output signal showed periodic amplitude and frequency modulations, the extent of which was comparable to the data given by [23]. This can be seen in Figure 7, which gives both the signal and its amplitude and frequency modulation laws. In this case, the resulting audible effect (which can also be obtained as described in [32]) is a combination of the so-called vibrato and tremolo effects that does not correspond at all to the typical Leslie effect. When the source images were active, the signal characteristics were much more complex, as shown in Figure 8, where the aperiodic behavior of the modulation laws, which we believe to be responsible for the particular “Leslie effect,” can be clearly seen. Actually, these features can also be seen in Figure 5, which shows the output signal recorded from a real Leslie cabinet driven by an 800 Hz monochromatic signal. Using musical signals, the sounds obtained with the Leslie cabinet and the simulator output have been described by professional musicians as being of a similar quality. A Max-MSP implementation of the Leslie cabinet simulator can be downloaded on the web (see Section 6).

3.5. Spatialization

Another important feature of the Leslie cabinet effect is the spatial modulation resulting from the time-dependent coupling between the cabinet and the listening room. To simulate this effect, a time-dependent directivity system was used. The directivity of this system should ideally be the same as that of the Leslie cabinet. A generic approach to this directivity simulation such as that described in [33] can be used here, which involves measuring the simulating system and the target directivity. From these measurements, a set of filters is obtained by optimization methods. In the case of the Leslie cabinet simulation, rotation of the sources increases the complexity of the problem. In the first step, we designed a simplified, easy to control system of spatialization preserving the concept of rotating source. Our system of spatialization consisted of four loudspeakers placed back to back (Figure 9) to cover the whole 360-degree range. The set of loudspeakers can be defined as two orthogonal dipoles (x^+ , x^- and y^+ , y^-) which are able to generate a variable pattern of directivity. The input signal fed to each speaker satisfies the following expressions:

$$\begin{aligned}
 x^+ &= s(t)(\beta + (1 - \beta)\cos(\omega_M t)), \\
 x^- &= s(t)(\beta - (1 - \beta)\cos(\omega_M t)), \\
 y^+ &= s(t)(\beta + (1 - \beta)\sin(\omega_M t)), \\
 y^- &= s(t)(\beta - (1 - \beta)\sin(\omega_M t)).
 \end{aligned} \tag{6}$$

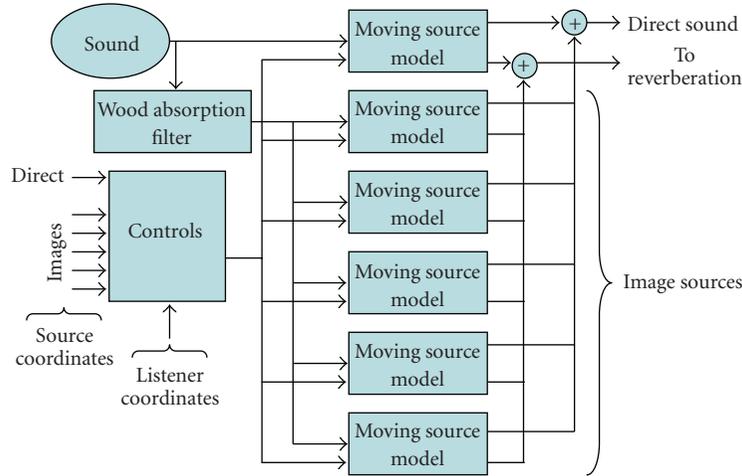


FIGURE 6: Overview of the Leslie horn simulator with 5-image sources.

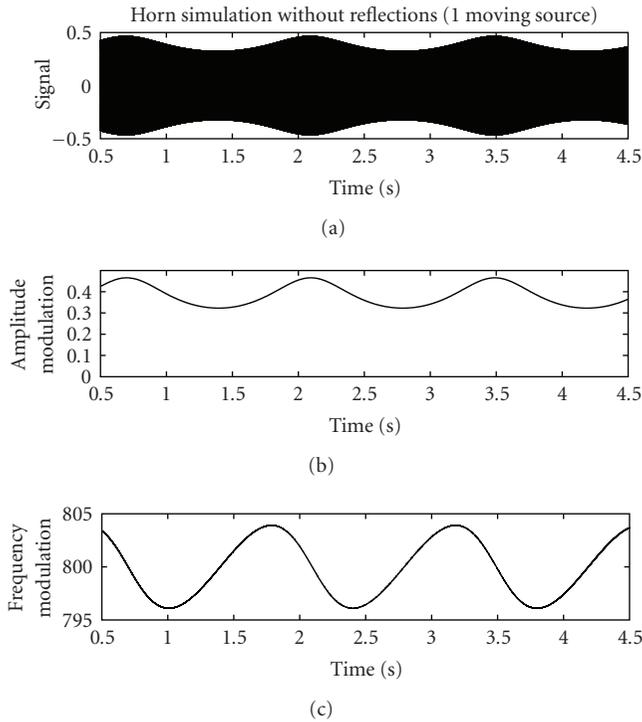


FIGURE 7: Analysis of the output signal from the horn simulator driven with a 800 Hz sinusoidal input signal. The part simulating the image sources has been disconnected. (a) microphone signal, (b) amplitude modulation, (c) frequency modulation.

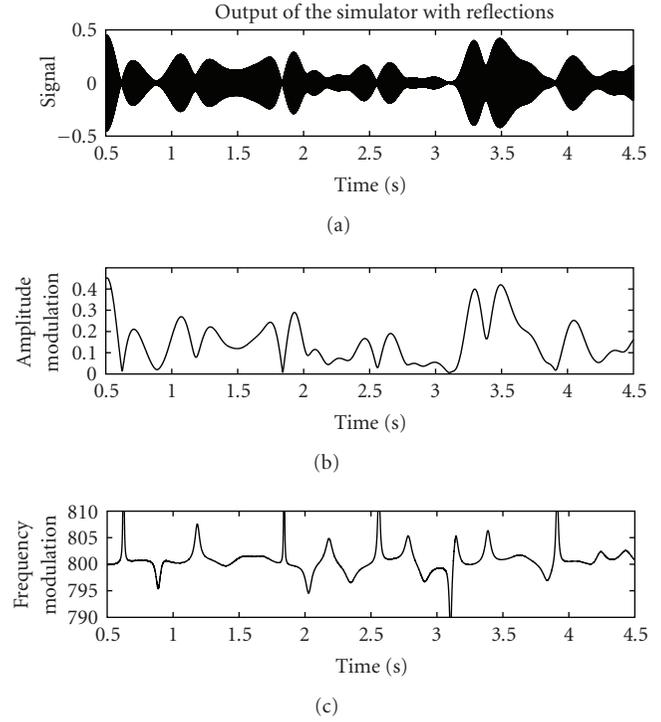


FIGURE 8: Analysis of the output signal from the complete Leslie simulator driven with a 800 Hz sinusoidal input signal. (a) microphone signal, (b) amplitude modulation, (c) frequency modulation.

The β parameter can be set at any value ranging between 0 and 1, so that the pattern of directivity can be adjusted from the omnidirectional to the bidirectional pattern. When $\beta = 1$, each speaker receives the same signal, and the system is, therefore, omnidirectional. When $\beta = 0$, the speakers corresponding to each dipole receive signals with opposite phases. Each dipole then distributes the energy with a “figure of eight” pattern of directivity. Since the two dipoles are

in phase quadrature, the resulting directivity of the whole system corresponds approximately to that produced by a rotating dipole at an angular speed of ω_M . When $\beta = 1/2$, which corresponds theoretically to a rotating cardioid pattern, satisfactory perceptual results were obtained.

In the real Leslie cabinet, the woofer port and the horns rotate at different angular frequencies. Two identical system of spatializations can thus be used to control the simulation

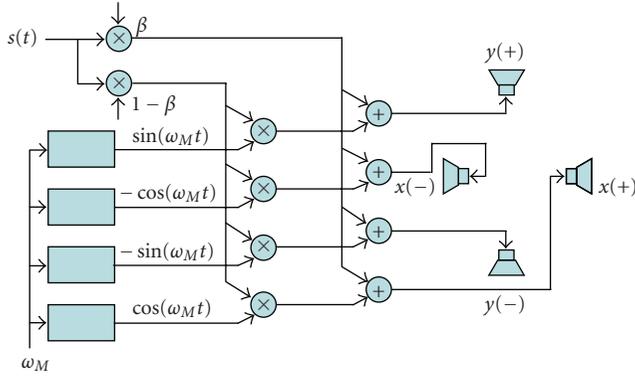


FIGURE 9: Scheme of the system of spatialization used for Leslie cabinet simulations.

process separately for the woofer and horn, each system being controlled by different angular rotation speed values.

4. COSMOPHONE

Sound is an interesting way of making invisible events perceptible. Actually, sounds produced by invisible or hidden sources can provide information about both the motion and the location of the sources. The cosmophone is a 3D sound immersion installation designed to sonify invisible cosmic particles, using synthetic sounds eliciting physically relevant sensations. The design of the cosmophone as a sound and music interface has been described in [34, 35]. We will describe below how the moving sound model was used in this framework to generate sounds evoking the trajectories of cosmic particles.

4.1. The cosmic rays

Interstellar space contains a permanent flux of high-energy elementary particles called “cosmic rays.” These particles were created by violent events, such as those occurring when a huge and aged star explodes and becomes a supernova. The particles then remain confined in the galaxy for millions of years because of the galactic magnetic fields before reaching our planet. When colliding with the Earth’s atmosphere, cosmic rays create showers of secondary particles. Although they are partly absorbed by the atmosphere, these showers have many measurable effects, including a flux of muons. Muons, which resemble heavy electrons but are usually absent from matter because of its short lifetime, are present in high levels in cosmic showers. Thanks to their outstanding penetrating properties, they are able to reach the ground. At sea level, they arrive at a rate of about a hundred muons per second per square meter. High-energy cosmic rays produce bunches of muons or multimuons, having the same direction and falling a few meters apart from each other.

4.2. The cosmophone installation

Human beings are unaware of the particles passing through their body. The cosmophone is a device designed to make the

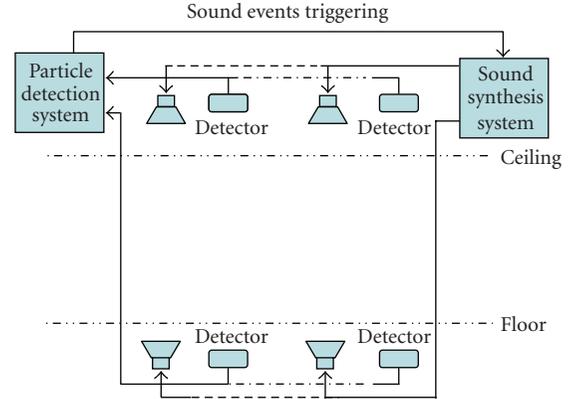


FIGURE 10: Scheme of the cosmophone device.

flux and properties of cosmic rays directly perceptible within a three-dimensional space. This is done by coupling a set of elementary particle detectors with an array of loudspeakers via a real-time data acquisition system and a real-time sound synthesis system (Figure 10). In this device, the information received from the detectors triggers the onset of sounds. Depending on the parameters of the particles detected, various types of sounds are generated. These parameters and the rate of occurrence of the various cosmic phenomena give rise to a large variety of sound effects. Many strategies for generating sounds from random events of this kind are currently being explored.

The system of synthesis has to generate sounds in response to signals emitted by the particle detection system. To simulate a rain of particle, in which listeners are immersed, the loudspeakers were placed in two arrays: one above the listeners (above a ceiling) and the other one below them (under a specially built floor). The arrays of loudspeakers were arranged so that the ears of the listeners (who were assumed to be standing up and moving about inside the installation) were approximately equidistant from the two groups. Both ceiling and floor were acoustically transparent, but the speakers were invisible to the listeners. A particle detector was placed near each loudspeaker. When a particle first passed through a detector in the top group, then through a detector in the bottom group, a sound event was triggered. This sound event consisted of a sound moving from the ceiling to the floor, thus “materializing” the trajectory of the particle.

4.3. Sound generation and spatialization

The sound generator system was based on the moving sound source model described above. It also includes a synthesis engine allowing for the design of various sounds and a sampler triggering the use of natural sounds. Because of the morphology of human ears, one can accurately localize sources moving in a horizontal plane, but far less accurately those moving in the vertical plane [36]. Accordingly, initial experiments have shown that the use of a panpot to distribute the signal energy between two loudspeakers do not suffice to create the illusion of a vertically moving sound source.



FIGURE 11: A picture of the cosmophone installed in the Cité des Sciences et de l'Industrie (Paris).

In particular, listeners were unable to exactly distinguish the starting and final positions of the moving source in 3D space. To improve the localization of the extreme points on the particle trajectory, we, therefore, added two short cues (called localization indices) to the sound event. The first cue is emitted by the upper loudspeaker at the beginning of the sound event and the second by the lower loudspeaker, at the end of the event. Since these two cues were chosen so as to be very exactly localizable, they have greatly improved the subjects perception of the vertical trajectory by giving the impression of a sound crossing the ceiling before hitting the floor.

A 24-channel cosmophone device was built for the *Cité des Sciences et de l'Industrie* in Paris, as part of a particle physics exhibition stand: the *Théâtre des Muons* (Figure 11). It was recently updated for the exhibition called *Le Grand Récit de l'Univers*. In this installation, two arrays of twelve speakers and detectors were placed in two concentric circles: the inner one comprises four speakers and detectors and the outer one, eight others. The outer circle was about five meters in diameter, which is wide enough to allow several listeners to stand in the installation.

In practice, three different events could be distinguished: a single muon reaching a pair of detectors (by successively hitting a detector placed above the ceiling, then one located under the floor), a “small bunch,” where more than one, but less than four pairs of detectors are hit simultaneously, and a “large bunch,” when at least four pairs are hit. The three cases corresponded to different sound sequences (sound examples can be found at: <http://cosmophone.in2p3.fr/>).

5. CONCLUSION

To make virtual moving sound events realistic, some important features of the physical processes of real moving sources can be modeled. When dealing with synthesis processes or sounds recorded from fixed sources, a preprocessing step is required to induce in listeners a coherent mental representation of the motion. The real-time preprocessing model designed for this purpose accounts accurately for four main perceptual cues, namely, the intensity, timbre, and reverberation, as well as the Doppler effect. This model renders moving sound sources accurately, even in the case

of monophonic diffusion systems, which shows the relative independence existing between sound motion and sound localization. The model parameters can be based on physical considerations. By simplifying the process, while keeping the most fundamental aspects of the situation, an accurate method of implementing and controlling the model in real time was developed.

The moving sound model could now be used as the basis of more complex systems involving the influence of room acoustics, for example. The Leslie Cabinet is a good example of systems of this kind, since the perceptual effects produced by the cabinet results from the effects of both the rotating source and the sound enclosure. We have also described here how a combination of several elementary moving sound source models can be used to accurately simulate this special choral effect and how the realism can be enhanced by connecting these models to a system of multiple speakers. Likewise, the moving source model has been used to construct a 3D sound immersion system for detection of cosmic particles. The cosmophone, which is based on a combination of moving source effects and spatialization techniques, is a good example of applications, where only a few features, such as localization indices improving our ability to localize vertically moving events, have been successfully added to our generic model.

The simulation of moving sound sources is an exciting field of research, always opening new domains of applications. Various techniques can be combined to generate novel audio effects such as those obtained by incorporating the Leslie cabinet simulator to the cosmophone installation. As far as the musical applications of this approach are concerned, we are currently developing an interface including a motion sensor for controlling a clarinet synthesis model in which the motion of the instrument is accounted for. Simulating the motion of sound sources is undoubtedly one of the keys to realistic sound modelling.

6. METHODS

Cosmophone: <http://cosmophone.in2p3.fr/>.

Java atmospheric sound absorption calculators: <http://www.csgnetwork.com/atmosphndabsorbcalc.html><http://www.me.metu.edu.tr/me432/soft15.html>.

Moving Sound Max/MSP patches downloadable from: <http://www.lma.cnrs-mrs.fr/~kronland/MovingSources>.

ACKNOWLEDGMENTS

Part of this work has been supported by the French National Research Agency (A.N.R.) in the framework of the “senSons” project (JC05-41996), headed by S. Ystad (see <http://www.sensons.cnrs-mrs.fr>). The cosmophone was developed by D. Calvet, R. Kronland-Martinet, C. Vallée, and T. Voinier, based on an original idea by C. Vallée. The authors thank T. Guimezanes for his participation in the Leslie cabinet measurements.

REFERENCES

- [1] M. M. Wanderley, B. W. Vines, N. Middleton, C. McKay, and W. Hatch, "The musical significance of clarinetists' ancillary gestures: an exploration of the field," *Journal of New Music Research*, vol. 34, no. 1, pp. 97–113, 2005.
- [2] M. A. Gerzon, "Periphony: with-height sound reproduction," *Journal of the Audio Engineering Society*, vol. 21, no. 1, pp. 2–10, 1973.
- [3] ITU-Recommendation BS.775-1, "Multichannel stereophonic sound system with and without accompanying picture," 1994.
- [4] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *The Journal of the Acoustical Society of America*, vol. 93, no. 5, pp. 2764–2778, 1993.
- [5] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, 1997.
- [6] J. Schroeter, C. Poessel, H. Opitz, P. L. Divenyi, and J. Blauert, "Generation of binaural signals for research and home entertainment," in *Proceedings of the 12th International Congress on Acoustics (ICA '86)*, vol. B1–6, Toronto, Canada, July 1986.
- [7] J. D. Warren, B. A. Zielinski, G. G. R. Green, J. P. Rauschecker, and T. D. Griffiths, "Perception of sound-source motion by the human brain," *Neuron*, vol. 34, no. 1, pp. 139–148, 2002.
- [8] J. M. Chowning, "The simulation of moving sound sources," *Journal of the Audio Engineering Society*, vol. 19, no. 1, pp. 2–6, 1971.
- [9] A. Välijamäe, P. Larsson, D. Västfjäll, and M. Kleiner, "Traveling without moving: auditory scene cues for translational self-motion," in *Proceedings of the 11th International Conference on Auditory Display (ICAD '05)*, Limerick, Ireland, July 2005.
- [10] P. Schaeffer, *Traité des Objets Musicaux*, Seuil, Paris, France, 1966.
- [11] J.-M. Jot and O. Warusfel, "A real-time spatial sound processor for music and virtual reality applications," in *Proceedings of the International Computer Music Conference (ICMC '95)*, pp. 294–295, Banff, Canada, September 1995.
- [12] J. Huopaniemi, L. Savioja, and M. Karjalainen, "Modeling of reflections and air absorption in acoustical spaces: a digital filter design approach," in *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA '97)*, p. 4, New Paltz, NY, USA, October 1997.
- [13] S. S. Stevens, "The relation of pitch to intensity," *The Journal of the Acoustical Society of America*, vol. 6, no. 3, pp. 150–154, 1935.
- [14] J. G. Neuhoff and M. K. McBeath, "The Doppler illusion: the influence of dynamic intensity change on perceived pitch," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 22, no. 4, pp. 970–985, 1996.
- [15] L. D. Rosenblum, C. Carello, and R. E. Pastore, "Relative effectiveness of three stimulus variables for locating a moving sound source," *Perception*, vol. 16, no. 2, pp. 175–186, 1987.
- [16] A. Merer, S. Ystad, R. Kronland-Martinet, M. Aramaki, M. Besson, and J.-L. Velay, "Perceptual categorization of moving sounds for synthesis applications," in *Proceedings of the International Computer Music Conference (ICMC '07)*, pp. 69–72, Copenhagen, Denmark, August 2007.
- [17] S. McAdams and E. Bigand, *Thinking in Sound: The Cognitive Psychology of Human Audition*, Oxford University Press, Oxford, UK, 1993.
- [18] M. Aramaki, H. Baillères, L. Brancheriau, R. Kronland-Martinet, and S. Ystad, "Sound quality assessment of wood for xylophone bars," *The Journal of the Acoustical Society of America*, vol. 121, no. 4, pp. 2407–2420, 2007.
- [19] P. M. Morse and K. U. Ingard, *Theoretical Acoustics*, MacGraw-Hill, New York, NY, USA, 1968.
- [20] D. Zicarelli, "An extensible real-time signal processing environment for max," in *Proceedings of the International Computer Music Conference (ICMC '98)*, pp. 463–466, International Computer Music Association, Ann Arbor, Mich, USA, October 1998.
- [21] U. Zölzer, *Digital Audio Signal Processing*, John Wiley & Sons, New York, NY, USA, 1997.
- [22] ANSI-S1.26, "Method for calculation of the absorption of sound by the atmosphere," American National Standards Institute, New York, NY, USA, 1995.
- [23] J. Smith, S. Serafin, J. Abel, and D. Berners, "Doppler simulation and the leslie," in *Proceeding of the 5th International Conference on Digital Audio Effects (DAFx '02)*, Hamburg, Germany, September 2002.
- [24] H. Strauss, "Implementing Doppler shifts for virtual auditory environments," in *Proceedings of the 104th Audio Engineering Society Convention (AES '98)*, Audio Engineering Society, Amsterdam, The Netherlands, May 1998, paper no. 4687.
- [25] N. Tsingos, *Simulation de champs sonores de haute qualité pour des applications graphiques interactives*, Ph.D. thesis, Université de Grenoble 1, Saint-Martin-d'Hères, France, 1998.
- [26] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay: tools for fractional delay filter design," *IEEE Signal Processing Magazine*, vol. 13, no. 1, pp. 30–60, 1996.
- [27] C. A. Henricksen, "Unearthing the mysteries of the leslie cabinet," *Recording Engineer/Producer Magazine*, pp. 130–134, April 1981.
- [28] J. Ville, "Théorie et applications de la notion de signal analytique," *Cables et Transmission*, vol. 2, no. 1, pp. 61–74, 1948.
- [29] I. Arroabarren, X. Rodet, and A. Carlosena, "On the measurement of the instantaneous frequency and amplitude of partials in vocal vibrato," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 14, no. 4, pp. 1413–1421, 2006.
- [30] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.
- [31] G. Ballou, *Handbook for Sound Engineers*, Focal Press, Woburn, Mass, USA, 1991.
- [32] S. Dish and U. Zölzer, "Modulation and delay line based digital audio effects," in *Proceeding of the 2nd COST-G6 Workshop on Digital Audio Effects (DAFx '99)*, pp. 5–8, Trondheim, Norway, December 1999.
- [33] O. Warusfel and N. Misdariis, "Directivity synthesis with a 3D array of loudspeakers-application for stage performance," in *Proceedings of the COST-G6 Conference on Digital Audio Effects (DAFx '01)*, Limerick, Ireland, December 2001.
- [34] P. Gobin, R. Kronland-Martinet, G.-A. Lagesse, T. Voinier, and S. Ystad, "Designing musical interfaces with composition in mind," in *Computer Music Modeling and Retrieval*, vol. 2771 of *Lecture Notes in Computer Science*, pp. 225–246, Springer, Berlin, Germany, 2003.
- [35] C. Vallée, "The cosmophone: towards a sensuous insight into hidden reality," *Leonardo*, vol. 35, no. 2, p. 129, 2002.
- [36] J. Blauert, *Spatial Hearing*, The MIT Press, Cambridge, Mass, USA, 1983.