



---

# Audio Engineering Society

# Convention Paper 6656

Presented at the 120th Convention  
2006 May 20–23 Paris, France

*This convention paper has been reproduced from the author's advance manuscript, without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. Additional papers may be obtained by sending request and remittance to Audio Engineering Society, 60 East 42<sup>nd</sup> Street, New York, New York 10165-2520, USA; also see [www.aes.org](http://www.aes.org). All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.*

---

## Visualization of perceptual parameters in interactive user interfaces: Application to the control of sound spatialization

Olivier Delerue<sup>1</sup>

<sup>1</sup> IRCAM, 1 Place Igor Stravinsky 75004 Paris, France.  
[Olivier.Delerue@ircam.fr](mailto:Olivier.Delerue@ircam.fr)

### ABSTRACT

This work addresses the general problem of designing graphical user interfaces for non expert users. The key idea is to help the user anticipating his/her actions by displaying, in the interaction area, the expected evolution of a quality criterion according to the degrees of freedom which are being monitored. This concept is first applied to the control of sound spatialization: various perceptually based criteria such as “spatial homogeneity” or “spatial masking” are represented as a grey shaded map superimposed to the background of a bird’s eye view interface. After selecting a given sound source, the user is thus informed how these criteria will behave if the source is being moved to any other location of the virtual sound scene.

### 1. INTRODUCTION

Beyond spatial sound synthesis and rendering techniques for simulating sound source localization and their associated room effect, spatialization tools require the creation of convenient graphical user interfaces for manipulating the numerous parameters of the virtual sound scene. However, the wide varieties of applicative contexts as well as the high complexity of the underlying signal processing algorithms suggest a high level design of these control interfaces so that non-expert end users can manipulate the virtual sound scenes without risking perceptual inconsistency in the auditory result.

The goal of the work presented in this paper is to provide an original means for visualizing arbitrary perceptual parameters over a 3D sound scene: this work was carried out in the context of the SemanticHIFI European project ([11]) which objective is precisely to provide new features to domestic audio equipment by integrating results from the computer music research area.

This system provides a grey-shaded map representation integrated as the background of a bird’s eye view of the sound scene. The goal of this system is to indicate to the user how a given perceptual parameter is going to behave if a sound source is moved to a given location in

the space. Typically the chosen perceptual parameter underlies a quality criterion, using a bright shade of grey when the auditory result is satisfying according to this criterion, and a darker shade for locations that lead to a poor auditory result. Thus, when manipulating the sound scene, the user can predict what the result of his actions is going to be with respect to the quality criterion.

This paper first focuses on early experiments with a simple and didactic perceptual parameter named “Spatial Homogeneity”. This criterion stipulates simply that in order to create a good mixing, the loudness perceived at each ear of the listener should be relatively balanced. This parameter is highly dependant on the restitution system (stereo loudspeakers, headphones, 5.1 setup...) as well as the spatialization or panning method used (pair wise panning, X-Y microphone simulation, ambisonic, binaural...). We describe the influence of such characteristics of the rendering engine by comparing the resulting maps for different panning methods in the case of a stereo loudspeaker setup.

Another example of important parameter to consider when performing spatialization is the spatial masking effect that occurs when sound sources match in the same region of space and might result in complete masking or loss of intelligibility. In such case, for a given sound source, the space is represented with a bright shade of grey for locations where the source will be fully audible and will not interfere with other sources, and, on the opposite with a dark shade for locations where the source will be either inaudible or might disrupt the listening of another source.

Beyond individual parameters, we consider visualizing weighted combinations of criteria in order to lead to a system that provides a global indication on the auditory result even if a single dimensional shade of grey does not allow explaining why a given location is either satisfactory or not.

The system has been implemented within ListenSpace ([2],[3]), a prototyping control and authoring environment, developed at IRCAM and that communicates with the Spatialisateur ([5]) for sound localization and room effect rendering. This implementation takes the advantage of metadata such as individual loudness of the sound sources, or spectral information, either pre-calculated or, when possible, evaluated on the fly.

Finally, we believe that there exist a large number of applications for such visualization system, ranging from non-expert end users interaction to critical situations such as live performances where acting over the positions of sources – as well as other types of parameters – can not be performed unheard: we conclude on a possible generalization of our concept to reach a wide range of computer music applications.

## 2. RELATED WORK

Control interfaces for sound spatialization have been the purpose of a number of researches. The Ircam Spatialisateur ([5]) for instance was the opportunity to focus on the perceptual aspects of the sound scene and provide the “SpatOper”, a control interface made of classical interaction components (typically sliders), but presenting the sound scene as a set of perceptual acoustical properties instead of low level digital signal processing parameters. Other works such as Move in Space ([9]) or Holophon ([8]) for instance focused more on the temporal aspects of sound spatialization by making use of the notion of trajectory in real time. Such idea has been also studied in [6] in the OpenMusic environment from a more symbolical approach for spatialization scripting.

The research presented in this paper is based on existing representations of sound scenes and interaction means, but focuses on computer assisted aspects in order to help the users when performing sound spatialization. In a previous work ([7]) we described how the use of constraint programming could restrain the interaction space to a set of values that always provide a satisfactory auditory result. In this work our goal is not to restrict the interaction possibilities but to take advantage of the use of information visualization to provide real time information on the auditory result with respects to a given perceptual criterion.

The world of information visualization is now a mature research area and there exist already a number of toolkits such as [4] for instance to create visualization environments using a predefined set of advanced and reusable interaction components. However, to our knowledge, such systems are not oriented toward real-time applications and have never been applied to the particular context of real time control of sound spatialization.

Finally, each perceptual criterion implemented will be based on results originating from perceptual studies. Concerning our case study, the ‘‘Spatial Homogeneity’’ and ‘‘Spatial Masking’’ criteria, basic information can be found in [1] for instance, as well as in [10] where spatial masking information is exploited to simplify the spatialization process of complex sound scenes.

### 3. GENERIC APPROACH

In this approach we consider an interactive system as a set of variables  $V_1, \dots, V_n$  that take values in  $\mathfrak{R}^N$ . These variables are represented in a subspace  $\mathfrak{R}^k$  by means of a projection  $P$  (Typically the representation is a two-dimensional representation with  $k=2$  for a screen-based interactive application). The user can then interact with the system through a set of components (sliders, push buttons, toggles,...) in a user graphical interface and adjust consequently the values of the variables in the representational subspace. We call  $Vp_1, \dots, Vp_n$  the values of the variables in the representational subspace.

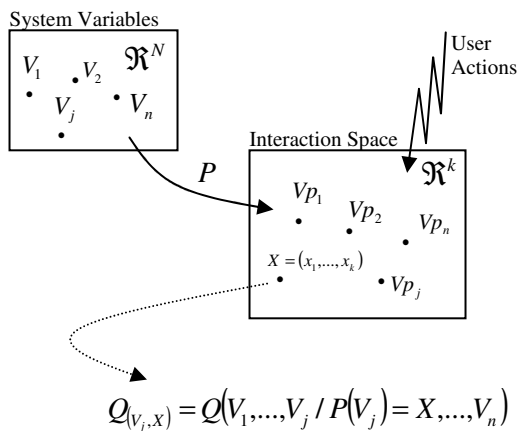


Figure 1: illustration of the quality criterion for a given ‘‘target variable’’ ( $V_j$ ) evaluated at the target value  $X=(x_1, \dots, x_k)$  in the interaction space.

We introduce the idea of a ‘‘quality criterion’’ as a function  $Q: (\mathfrak{R}^N)^n \rightarrow [0,1]$ , that associates  $(V_1, \dots, V_n) \mapsto Q(V_1, \dots, V_n)$  a coefficient taking the value

1 when the given criterion is satisfied and 0 when it is not.

The system we propose consist for a given set of values to the variables  $(V_1, \dots, V_n)$ , and a given ‘‘target variable’’  $V_j$  representing in each point  $(x_1, \dots, x_k)$  of the interaction subspace the value of  $f(V_1, \dots, V_j', \dots, V_n)$  where  $V_j'$  represents the interaction modification to  $V_j$  so that  $P(V_j') = (x_1, \dots, x_k)$  by using a background shade of grey for instance (a value close to 0 will result in a dark shade and a value close to 1 will give a bright shade).

Thus, the shade of grey represented at point  $(x_1, \dots, x_k)$  reflect directly what would be the value of the quality criterion if the user had changed the target variable  $V_j$  so that its projection in the representation space was  $(x_1, \dots, x_k)$ .

This concept could be applied typically to any type of interactive system controlled by a set of continuous variables. We present in the next section our implementation, applied to the control of sound spatialization.

### 4. APPLICATION TO THE CONTROL OF SOUND SPATIALISATION

In our case, we consider the control of source localization in a sound spatialization application. We used the ListenSpace environment to implement this visualization system. The variables of the system are the position of each sound source (relatively to a reference listening point) in a two dimensional scene.

The user uses the mouse (drag) to change the position of the sound source and can point out which source with be considered as the ‘‘target variable’’: this source is then represented with a red circle (such as ‘‘Source2’’, on Figure 2).

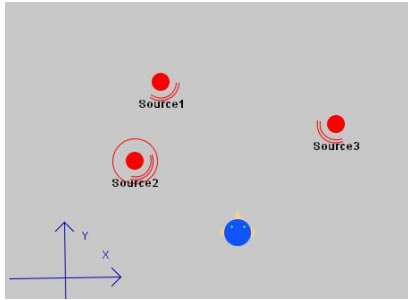


Figure 2: 2D representation & interaction space for a sound scene

In this context, the representation and interaction space has the same dimension than the original space so the projection is straightforward (at the exception of a Polar to Cartesian conversion and a truncation from floating point numbers to integers). It corresponds to a graphical “screen” representation as shown in Figure 2. The value of the chosen criterion is represented in the manner of a background shade of grey map (see Figure 3).

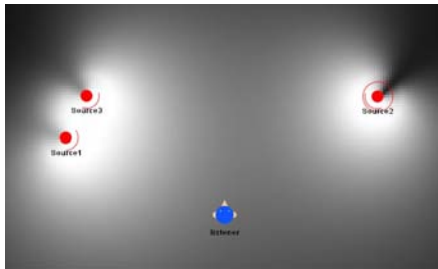


Figure 3: example of the use of a background shade of grey map to represent the directivity pattern of sound sources.

We discuss in the next section several examples of perceptual aspect of the sound scene that could be converted into suitable criteria for such visualization system.

## 5. CRITERION EXAMPLES

One can imagine a wide range of criteria that could be used in our system: perceptual characteristics of a sound scene can be defined over many different aspects. However, when focusing on the localization of sound sources the most immediate criteria that come to mind are for instance the spatial homogeneity criterion of the scene. Its use would lead to a even distribution of en

energy in the auditory space, according to the rendering setup and the underlying spatialization algorithm.

Similarly, the “Spatial masking” criterion would help the user positioning sources so that they do not interfere and prevent from a possible masking effect or loss of intelligibility in the auditory result.

The “crosstalk criterion” can as well provide useful information: this idea reflects the case of the spatial reconstitution of a live recording for instance in which the sound sources originating from acoustical recordings might be correlated. This criterion would then help deciding to which extent these sound sources can be put apart in the rendering space without losing precision in the perception of sound localization. This criterion applies as well when the audio material originates from source separation algorithm. In such case, the risk is to let separation artifacts become hearable because of the spatial distance between the extracted signals.

We focus in the next sections of this document on two particular examples, the « Spatial Homogeneity » and the « Spatial masking » criteria for which we proposed an implementation in our system, ListenSpace.

### 5.1. Spatial Homogeneity

Intuitively, one can consider that a “good mix” has to take up the auditory space evenly. We introduce the “spatial homogeneity” criterion as a clue on how evenly the energy is distributed within the rendering space.

Thus, for a given target source, we represent in each location of the interaction space the quality of the resulting mix in terms of “spatial unity” as if this sound source was moved to this location.

The notion of spatial unity is dependant on the restitution system. For a stereo setup for instance it can be defined as the difference in the resulting loudness coming from each loudspeaker. We present in the following section, an implementation of the Spatial Homogeneity criterion, in the case of a stereo rendering setup, considering arbitrary that the sound scene starts to be significantly unbalanced when the resulting loudness at each loudspeaker shows a difference of 10 decibels.

### 5.1.1. Stereo panning case:

For a number of sound sources  $\{S_i\}$  with locations corresponding to azimuth  $az_i$  and distance  $d_i$ , and a basic panning method resulting in a panning attenuation for left and right channels ( $Pan_{right}(az), Pan_{left}(az)$ ) as well as an attenuation factor depending on the distance of the source  $Att(d)$ , we define the resulting loudness on each loudspeaker as follow:

$$Loud_{left} = 10 * \log_{10} \left( \sum_i \frac{10^{\frac{Loud_i}{10}} * Pan_{left}^2(az_i)}{d_i^2} \right)$$

and

$$Loud_{right} = 10 * \log_{10} \left( \sum_i \frac{10^{\frac{Loud_i}{10}} * Pan_{right}^2(az_i)}{d_i^2} \right)$$

If the user had moved the source  $S_j$  to a location  $p$  defined by azimuth  $az_p$  and distance  $d_p$  the resulting loudness would then be:

$$Loud_{left}(S_j, p) = 10 * \log_{10} \left( \left( \sum_{i \neq j} \frac{10^{\frac{Loud_i}{10}} * Pan_{left}^2(az_i)}{d_i^2} \right) + \frac{10^{\frac{Loud_j}{10}} * Pan_{left}^2(az_p)}{d_p^2} \right)$$

and

$$Loud_{right}(S_j, p) = 10 * \log_{10} \left( \left( \sum_{i \neq j} \frac{10^{\frac{Loud_i}{10}} * Pan_{right}^2(az_i)}{d_i^2} \right) + \frac{10^{\frac{Loud_j}{10}} * Pan_{right}^2(az_p)}{d_p^2} \right)$$

We then define the “spatial uniformity” factor for source  $S_j$  at location  $p$  factor as:

$SU(S_i, p) = f_{SU}(loud_{left}(S_i, p) - loud_{right}(S_i, p))$  where  $f_{SU}$  is a Gaussian function (typically  $f_{SU}(x) = e^{-\frac{x^2}{100}}$ ) as represented in Figure 4. This function gives intuitively a relatively dark shade of grey for a loudness difference of 10 decibels which seems to be a perceptively significant value.

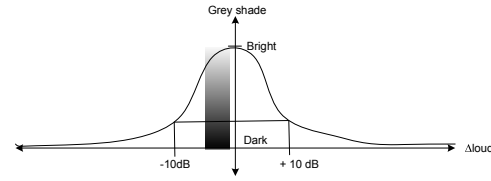


Figure 4: conversion function from loudness difference to spatial uniformity criterion

### 5.1.2. Practical approach:

In practice the IRCAM Spatialisateur provides a value “Es” that describes the energy attenuation that results in the direct sound and early reflections for a sound source located at a given reference distance  $d_{ref}$ . We use this couple  $(Es, d_{ref})$  as a reference energy value and deduce a distance attenuation if the source was moved to distance  $d'$  as  $-20 \log_{10} \left( \frac{d'}{d_{ref}} \right)$

In the “PanC2” algorithm, the Spatialisateur simulates a XY couple of microphones (coincident microphones with a 90 degrees angle). The signal coefficients corresponding to azimuth are defined as:

$$Pan_{left}(az) = (\sqrt{2} - 1) \left( 1 + \cos\left(\frac{\pi}{4} - az\right) \right)$$

and

$$Pan_{right}(az) = (\sqrt{2} - 1) \left( 1 + \cos\left(\frac{\pi}{4} + az\right) \right)$$

Therefore, for a sound source  $S_i$  with the given azimuth  $az_i$ , distance  $d_i$  and attenuation  $Es_i$ , its contribution to each speaker is:

$$loud_{left}(S_i) = 10 \log_{10} \left( \frac{10^{\frac{loud_i + Es_i}{10}} * Pan_{left}^2(az_i)}{\left(\frac{d_i}{d_{ref_i}}\right)^2} \right)$$

and

$$loud_{right}(S_i) = 10 \log_{10} \left( \frac{10^{\frac{loud_i + Es_i}{10}} * Pan_{right}^2(az_i)}{\left(\frac{d_i}{d_{ref_i}}\right)^2} \right)$$

If source  $S_j$  was moved to distance  $d'$  and azimuth  $az'$  its contribution would be:

$$loud_{left}(S_j, az', d') = 10 \log_{10} \left( \frac{10^{\frac{loud_j + Es_j}{10}} * Pan_{left}^2(az')}{\left(\frac{d'}{d_{ref_j}}\right)^2} \right)$$

and :

$$loud_{right}(S_j, az', d') = 10 \log_{10} \left( \frac{10^{\frac{loud_j + Es_j}{10}} * Pan_{right}^2(az')}{\left(\frac{d'}{d_{ref_j}}\right)^2} \right)$$

We finally express the resulting transmitted loudness (for a corresponding move of source  $S_j$  to distance  $d'$  and azimuth  $az'$ ) at each speaker as:

$$Loud_{left}(S_j, p(az', d')) = 10 * \log_{10} \left( \left( \sum_{i \neq j} \left( \frac{10^{\frac{loud_i + Es_i}{10}} * Pan_{left}^2(az_i)}{\left(\frac{d_i}{d_{ref_i}}\right)^2} \right) \right) + \frac{10^{\frac{loud_j + Es_j}{10}} * Pan_{left}^2(az')}{\left(\frac{d'}{d_{ref_j}}\right)^2} \right)$$

and

$$Loud_{right}(S_j, p(az', d')) = 10 * \log_{10} \left( \left( \sum_{i \neq j} \left( \frac{10^{\frac{loud_i + Es_i}{10}} * Pan_{right}^2(az_i)}{\left(\frac{d_i}{d_{ref_i}}\right)^2} \right) \right) + \frac{10^{\frac{loud_j + Es_j}{10}} * Pan_{right}^2(az')}{\left(\frac{d'}{d_{ref_j}}\right)^2} \right)$$

In the sound scene example represented in Figure 5, we use 3 sound sources (Source1, Source2, and Source3), and display in the background the Spatial uniformity criterion when Source2 is the target source. As described previously a light background describes areas where the auditory result should be evenly balanced between the loudspeakers and a dark shade of grey (typically close to the listener on its sides) would correspond to a high difference in energy level between the speakers.

Since the sound scene is already balanced with Source1 and Source 3 that have the same loudness energy, the system advises locating Source2 ideally in the central axis of the scene, defined by the listener's orientation.

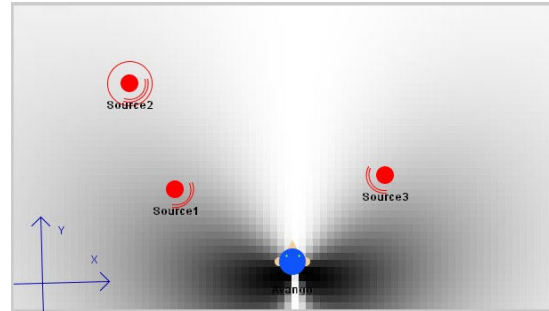


Figure 5: spatial uniformity criterion #1. Moving Source2 in a scene which is already balanced

If we now lower the energy delivered by Source3 (by adjusting its “presence” parameter in the Spatialisateur), we see that the sound scene becomes unbalanced and that Source2 should ideally be located in a circular-like area centered on the right side of the listening reference position (see Figure 6).

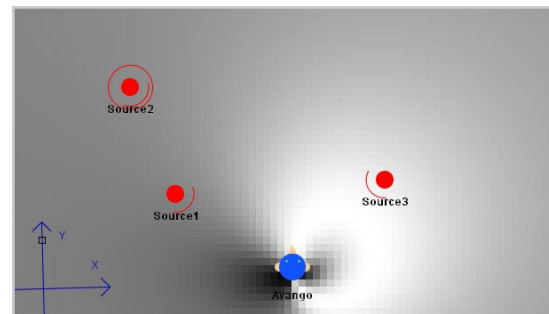


Figure 6: Spatial Uniformity Criterion #2. Moving Source2 in a scene with a high loudness difference between Source1 and Source3

Finally, in the last example Source2 has been moved to a bright background area: this signifies that the Spatial Uniformity Criterion is optimized.

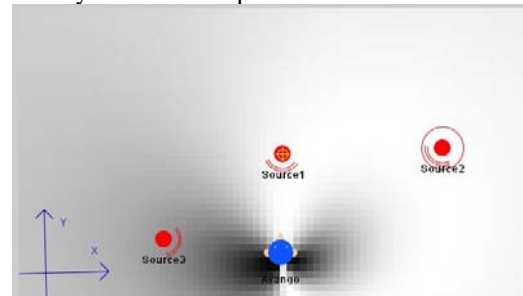


Figure 7: representation of a scene where the quality criterion is optimized

### 5.1.3. Importance of the panning techniques and rendering setup

To compare the effect of the panning law used (which depends on the chosen spatialization technique and the rendering setup) we implemented a different panning law for the same Spatial Homogeneity Criterion: we propose a simple cosinus panning law defined as:

$$Pan_{left}(az) = \cos\left(\frac{\pi - 2 * trim(az)}{2}\right)$$

and  $Pan_{right}(az) = \sin\left(\frac{\pi - 2 * trim(az)}{2}\right)$

where:

$$trim(az) = az \text{ if } \frac{-\pi}{2} \leq az \leq \frac{\pi}{2}, \pi - az \text{ if } az > \frac{\pi}{2}, \text{ and } -\pi - az \text{ if } az < \frac{-\pi}{2}$$

Figure 8 proposes a comparison of the representation of the spatial homogeneity criterion for a given sound scene for two different panning laws: it shows that the resulting area where the criterion is optimized is more symmetrical in the front / back positions and more lateralized in the cosines law method than in the XY couple simulation.

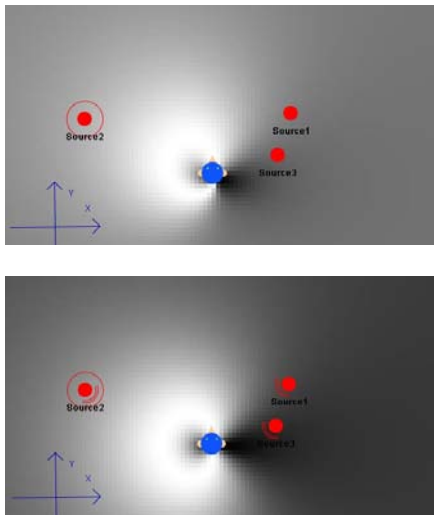


Figure 8: comparison of the effect of the panning technique over the spatial homogeneity criterion (XY panning law on the top, and cosinus panning law on the bottom representations).

Naturally, these results would differ even more when choosing a different rendering setup such as 5.1

rendering techniques using either ambisonics or pair wise panning methods.

### 5.2. Masking effect

Another interesting criterion we wish to illustrate in this paper is the “masking effect” criterion. We know that the spatial coincidence of sound sources may induce a masking effect and / or a loss of intelligibility in the auditory result.

We then introduce the “risk of being masked” perceptual criterion as the difference, for a given target point between the energy delivered by a target source and the energy delivered by the contribution of the other sources of the sound scene.

We estimate the contribution in loudness of each sound source in the direction corresponding to the target point as described in Figure 9.

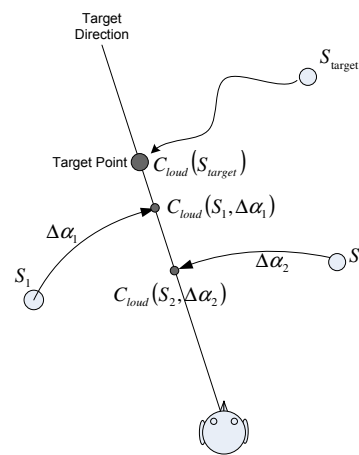


Figure 9: estimation of the loudness contribution of the sound sources in the target direction

The contribution of the target source at the target point is given as its original loudness combined with an attenuation function corresponding to the distance:

$$C_{Loud}(S_{target}, az_{target}, d_{target}) = (Loud_{target} - 10 * \log_{10}(d_{target}^2))$$

Whereas contribution of other sound sources is:

$$C_{Loud}(S_i, az_{target}, d_{target}) = 10 * \log_{10} \left( \frac{10^{\frac{Loud_i}{10}} * F_{att}^2(az_i - az_{target})}{d_i^2} \right)$$

with  $F_{att}(\Delta az)$  an attenuation function representing the fact that the masking influence of a sound source decreases as much as we consider a direction with an increasing angular variation from the position of the source. In our experiment, this function was modeled using a Gaussian function.

As a quality criterion for the risk of being masked, we compare the contribution of the target source at the target point with the sum of the contributions of all other sources in the target direction. If the target source is for instance  $S_j$  we express the resulting loudness difference as:

$$\Delta Loud = (Loud_j - 10 * \log_{10}(d_{target}^2)) - 10 * \log_{10} \left( \sum_{i \neq j} \left( \frac{10^{\frac{Loud_i}{10}} * F_{att}^2(az_i - az_{target})}{d_i^2} \right) \right)$$

We consider arbitrary that for a loudness difference smaller than -10db the target source would be considered as “masked”. This value is converted into a quality criterion using the function:

$$Q_{masking}(\Delta Loud) = e^{\frac{-(\Delta Loud)^2}{100}} \text{ if } \Delta Loud < 0, \text{ and } 1 \text{ if } \Delta Loud \geq 0$$

as represented in Figure 10

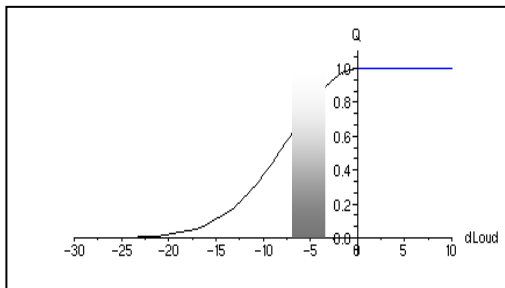


Figure 10: conversion function from loudness difference to masking quality criterion

Finally, the Figure 11 represents the “risk of being masked” criterion according to the loudness value of each sound source. In these 4 examples, the target source is always “Source2” (on the right of the sound scene).

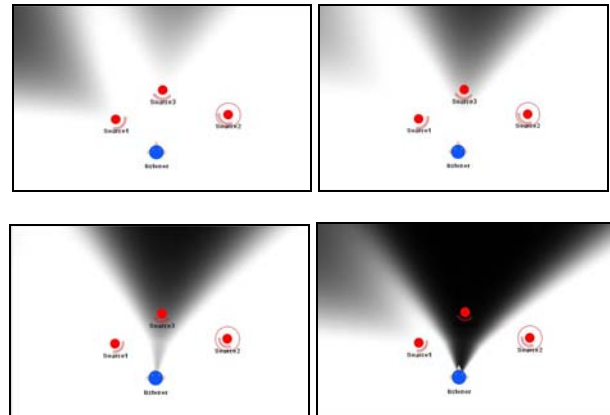


Figure 11: representation of the masking zones according to the loudness of the sound sources  
 (a) (upper left)  $E_{S1} = -10\text{dB}$ ,  $E_{S2} = -10\text{dB}$ ,  $E_{S3} = -10\text{dB}$ ,  
 (b) (upper right)  $E_{S1} = -15\text{dB}$ ,  $E_{S2} = -5\text{dB}$ ,  $E_{S3} = 0\text{dB}$   
 (c) (lower left)  $E_{S1} = -20\text{dB}$ ,  $E_{S2} = -10\text{dB}$ ,  $E_{S3} = 0\text{dB}$ ,  
 (d) (lower right)  $E_{S1} = -20\text{dB}$ ,  $E_{S2} = -20\text{dB}$ ,  $E_{S3} = 0\text{dB}$

Beside the “risk of being masked” for a given target sound source, it is also necessary in order to consider completely the spatial masking criterion to represent simultaneously the “risk of masking” for that source. This leads naturally to the idea of combining criteria together as described in the following section.

### 5.3. Combining Criteria

The system we presented allows predicting the effect of a given user action with respect to a unique perceptual criterion. We believe that a global view over the auditory scene would be more adapted to the control by a user and introduce for this purpose the idea of a “meta-criterion” expressed as a weighted combination of elementary criteria.

This idea is already necessary in order to represent entirely the “spatial masking” criterion as it should result in the combination of the “risk of masking” simultaneously with the “risk of being masked” criteria. We extend this idea to any weighted combination of criteria in order to represent a general quality criterion over the sound scene: the background shade of grey would then provide a hint on how the combination of quality criteria is satisfied, although for a dark grey value, it would not explain why a given action from the user would not be suitable from the perceptual point of



view (e.g. which of the criteria would not be satisfied from such action).

## 6. CONCLUSION & FUTURE WORK

We have presented a new visualization method for interactive systems, applied to the specific case of the control of sound spatialization. We showed how this system could help the user predict what is going to be the effect of one of his actions in the auditory result according to a given perceptual criterion.

A number of such perceptual criteria were mentioned and for two of them, the “spatial homogeneity” and the “spatial masking” criteria, we gave details of an early implementation stage in the ListenSpace environment.

Further development will consist first in a more detailed study of such perceptual criteria, and especially trying to relate our models to precise results in the field of auditory perception. More, the basic masking result for instance will be refined taking into account spectral aspects of the audio content.

Second, we believe that other fields of the Computer Music area can take advantage of our visualization system. Another direction of future work is then the implementation of these concepts for different types of applications by integrating it to common interaction components such as sliders, knobs, push button,... This concepts should find application to any situations where the context (such as the concert diffusion for instance) makes valuable the possibility to predict the result of an action before it has an effect in the auditory result.

## 7. REFERENCES

- [1] Blauert, Jens, “Spatial Hearing: The Psychophysics of Human Sound Localization. MIT Press, Cambridge, MA, 1983.
- [2] Delerue, Olivier, Warusfel, Olivier. “Authoring virtual sound scenes in the context of the LISTEN project”. Presented at the AES22, 22nd Audio Engineering Society Conference, Espoo, Finland, June 2002.
- [3] Delerue, Olivier, “A Mixed Physical and Perceptual approach to control spatialization in audio augmented realities”. Proceedings of the 6th International Conference on Digital Audio Effects (DAFX03), London, 2003
- [4] Fekete, J.D., “The InfoVis Toolkit”, in Proceedings of the 10th IEEE Symposium on Information Visualization (InfoVis'04), IEEE Press, 2004, pp. 167-174.
- [5] Jot, J.M. “Efficient Models for Distance and Reverberation Rendering in Computer Music and Virtual Audio Reality” Proceedings of the International Computer Music Conference ICMC97. San Francisco, USA, 1997
- [6] Nouno, Gilbert, Agon, Carlos, « Contrôle de la spatialisation comme paramètre musical », presented at the JIM2002, Journées d’informatique musicale, Marseille, France, 2002.
- [7] Pachet, François, Delerue, Olivier, “MidiSpace: a Temporal Constraint-Based Music Spatializer”, presented at the ACM98, sixth ACM international conference on Multimedia, pp 351 – 359, Bristol, England, 1998
- [8] Pottier, Laurent, Holophon : « projet de spatialization multi-sources pour une diffusion multi-haut-parleurs », Presented at the JIM2000, Journées d’Informatique Musicale, Bordeaux, 2000, pp 96 – 102.
- [9] Todoroff, Todor, Traube, Caroline, Ledent, Jean-Marc, “NeXTSTEP Graphical Interfaces to Control Sound Processing and Spatialization instruments”. in Proceedings ICMC97, International Computer Music Conference, Thessaloniki, Greece, 1997, pp.325 – 328, 1997.
- [10] Tsingos, Nicolas, Gallo, Emmanuel, and Drettakis, George, “Perceptual Audio Rendering of Complex Virtual Environments”, in ACM Transactions on Graphics (SIGGRAPH Conference Proceedings) number 3 volume 23 July 2004.
- [11] Vinet, Hugues, “The Semantic Hifi Project” presented at the ICMC05, International Computer Music Conference, Barcelona, 2005.