SOUND EFFECT METAPHORS FOR NEAR FIELD DISTANCE SONIFICATION

Gaëtan Parseihian & Brian FG Katz

LIMSI-CNRS
BP133, 91403, Orsay, France
Gaetan.Parseihian@limsi.fr
Brian.Katz@limsi.fr

Simon Conan

LMA-CNRS
31 chemin Joseph-Aiguier
13402 Marseille, France
conan@lma.cnrs-mrs.fr

ABSTRACT

This article presents a concept of distance sound source sonification for virtual auditory displays in the context of the creation of an assistive device for the visually impaired. In order to respond to user needs, three sonification metaphors of distance based on sound effects were designed. These metaphors can be applied to any type of sound and thereby satisfy all aesthetic desires of users. The paper describes the motivation to use this new type of sonification based on sound effects, and proposes guidelines for the creation of these three metaphors. It then presents a user evaluation of these metaphors by 16 subjects through a near field sound localization experiment. The experiment included a simple binaural rendering condition in order to compare and quantify the contribution of each metaphor on the distance perception.

1. INTRODUCTION

Thanks to the development of research in auditory display, the use of sound as a means to convey information has considerably grown over the past few decades. One of the most obvious applications is the sensory substitution of visual information when it is not available. Visually impaired people have a variety of needs for non-visual information. Accessing computer information, avoiding obstacles, finding a route or a desired inanimate object are examples of tasks that can be challenging for them. Some of these problems could be resolved by the use of auditory displays.

This study takes place within the context of the development of an electronic device based on rapid object localization and auditory augmented reality for helping people with visual impairments in near field guidance (hand reaching movement for grasping objects) [1]. This device combines a bio-inspired vision system able to quickly recognize and locate objects [2] and a 3D sound rendering system [3] which will map a spatialized sound to the location of the targeted object. Sound guidance will be provided through binaural rendering, allowing a full exploitation of the human perceptual and cognitive capacity for spatial hearing.

Even though the basic mechanisms of directional sound localization are well documented and can be easily reproduced in virtual auditory display through binaural rendering [4], those allowing listeners to determine the distance of a sound source are less understood. Literature on distance perception of sound sources [5, 6] reports that humans significantly underestimate the distance of far sources and overestimate the distance of near sources. They report at least four auditory cues involved in the mechanisms of distance auditory estimation:

- In open space, *intensity* plays a major role with familiar sounds, it ideally decreases by 6 dB with doubling of distance between the source and listener. For unfamiliar sound sources, this cue is insufficient as it is confounded with the level of the sound itself [7].
- *Direct-to-reverberant energy ratio* is also an important cue in reflective and indoor environments. Mershon and King [8] have shown that distance perception is greater in reverberant environments compared to anechoic environment. Contrary to intensity, reverberation can allow the listener to make an absolute judgment of distance.
- If the listener has enough familiarity with the sound, the *spectrum* may convey distance cues as well. The spectral filtering, especially effective for far distances (particularly in the upper part of the auditory range) is induced by the absorption properties of the air and the eventual multiple reflections over non-ideal surfaces, which help one to estimate the distance of a sound source[9].
- For nearby sources, Brungart [10] has highlighted the importance of *binaural differences* in both intensity and time that are no longer independent of radial distance, as they are for far field planar waves. A study by Shinn-Cunningham *et al.* [11], provides a detailed analysis of binaural cue variations for nearby sound source location.

Despite the multiplicity of distance perception cues, the synthesis of range information in auditory display still remain a major issue and leads to poor quality results, especially for near field sound sources.

In an attempt to provide a linear relationship between perceived and physical distance, Devallez *et al.* [12] modeled a virtual listening environment consisting of a trapezoidal membrane with specific absorptive properties at the boundaries. This approach has been more recently extended by Fontana and Rocchesso [?] who studied the effect of exaggerating the acoustic cue of the reverberation by placing a real sound source in a pipe. They also demonstrated the possibility of creating flexible and virtual models for distance rendering with a simple physical system such as the acoustic pipe [13].

In the context of near field guidance (for distances inferior to 1.5 meters), distance perception is quite limited compared to the required precision. Instead of linearizing or exaggerating distance acoustics cues, this study aims to explore the influence of adding new acoustic cues for distance perception. It consists of representing distance cues instead of simulating them exactly. This can be realized through the use of sonification techniques.

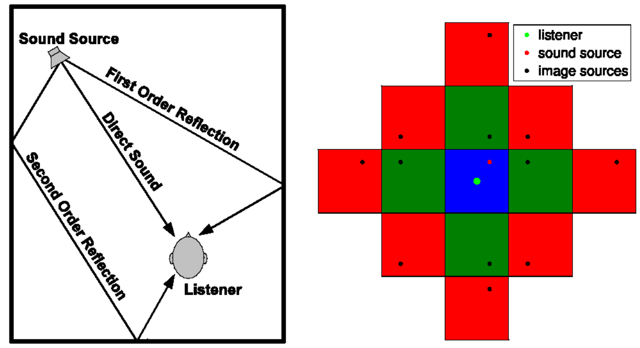
In [14], Kramer defined sonification as "the use of non-speech audio to convey information or perceptual data". Many studies have investigated methods of conversion from data to sound. The Sonification Handbook [15] provides a good introduction to various methods. In this study, a parameter mapping sonification approach was used. This method consists in representing changes in data dimension through an auditory variation [14, 16]. Most existing parameter mapping sonification applications use pitch, time, loudness, or timbre as the principal mapping parameters applied on sound synthesis. While the transfer function between sonified data and sound synthesis parameters is very easy, one problem is that the sounds produced can be unpleasant and irritating for daily use.

In the past few years, despite the development of many sound interfaces, aesthetic and user acceptance issues have been absent from the scope of most research. Very few studies have investigated the customization of sound information by the user and its impact on the effectiveness and efficiency of the system. In [17], the authors worked on the aesthetics of sonification and found that musical sounds were more pleasant and appropriate than natural sounds. In [18], Brungart and Simpson describe the design of an audio display that modified the acoustic properties of an arbitrary audio input signal (e.g. pilot-selected music) to provide the pilot with information about the altitude of the aircraft.

In this article, the concept of parameter mapping sonification is extended to the use of any type of audio signal by mapping the parameters to audio effects (these are then applied to the sound). In this concept, the data no longer relies on the sound parameters but on the audio effect parameters. This allows for the application of the sonification metaphor to any type of sound while maintaining coherency with the data displayed. Applying this concept, three sound effect metaphors were created and initially evaluated with a near-field localization test designed with laboratory sounds.

2. SOUND EFFECT METAPHORS

In the context of a commercial project, several constraints are imposed on the development of the prototype and therefore on the distance sonification design. First of all, the use of binaural sound display imposes the use of large spectrum sound samples (to increase HRTF cues perception) with sharp attacks (to improve ITD perception). Then, the design of an accessible, aesthetically pleasing, and ergonomic device takes into account the end user's needs in terms of output user interface. These were evaluated using several questionnaires as well as a creativity session held with six visually impaired participants (see [1] for further details). In general, the visually impaired panel did not favor the use of sound as a method of guidance. In addition to the sound environment-masking problem due to the use of headphones, they reported a severe fatigue from the kind of sounds generally used (such as beeps, noise, and tones) in interfaces, and to the excessive length of messages in the case of text-to-speech based systems. As sound information may interfere with natural auditory cues in the real environment and cause supplementary cognitive load, the amount of information provided should be minimal, presenting only what is necessary and sufficient to aid the user. Presented messages should be highly efficient and minimally intrusive. The level of detail and display frequency of messages must be adjustable by the user. The sounds must be short and different from urban environmental sounds. One of the most important results of these investigations on user needs was the differing desires of system sounds amongst potential users. Some users asked for electronic sounds (such as video game sounds) in order to easily differentiate



(a) (b)

Figure 1: (a) Sound path in a room. (b) 2D schematization of the image-sources method. The simulated room is in blue, first order reflections are located in green areas and second order reflections are located in red areas. The listener is a green •, the source a red •.

them from the natural ambient sounds, while others preferred decontextualized natural sounds (animal, sea, cave, or forest sounds) or instrumental sounds. Regarding these results, it was not possible to find a general agreement on the types of sounds to use for the design of a navigation aid. Instead, a decision was made to design the sonification device using a customizable sound strategy.

2.1. Effect based sonification

To answer all of these constraints, distance sonification was designed as a digital audio effect applicable to the sound. With this concept, the distance is mapped to one or several parameters of the audio effect and the resulting sound pattern is thus distance dependent. This method allows for the design of several distance metaphors while leaving the user the possibility to customize the actual sounds of the interface. Furthermore, it has the advantage that once the metaphors are understood and learned, the user is able to change the sounds without relearning the sonification mapping.

On the basis of this idea, three distance metaphors were developed. The first one consists of reproducing a natural perceptual phenomena (sound reflection from walls), based on a simple room acoustic simulation. The other two metaphors are symbolic. There is no ecological link between the effect and the parameter represented. These metaphors are defined in the next section with the chosen mapping corresponding to the experimental setup, detailed in Sec. 3.

2.2. Early Reflection (ER)

As explained in Sec. 1, several studies highlighted the improvement of distance perception using reverberation cues [8, 20]. In [21], Begault showed the benefit of an artificial reverberation in a virtual auditory display. The addition of room reverberation led to better externalization and distance perception of the sound source, but slightly decreased azimuth localization performance. From literature on distance perception of nearby sources, a hypothesis was made that distance perception of sound sources in peripersonal space is improved by early reflections [22].

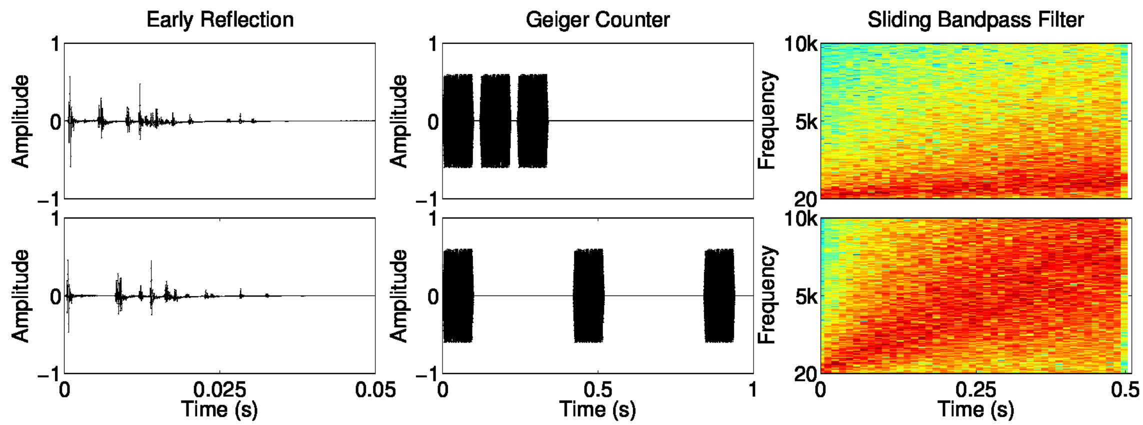


Figure 2: Representation of the sound resulting from the application of the three effect metaphors for two distances: $top = 0.6\ m$, $bottom = 1.5\ m$; (Left) Impulse response of the Early Reflection effect metaphor. (Center) Geiger Counter effect metaphor applied to a 10 ms burst. (Right) Spectrogram of the sounds resulting from the Sliding Bandpass Filter effect metaphor applied to a 0.5 sec burst.

The concept of this metaphor is therefore to create an effect based on the simulation of spatialized early reflection of second order (ie, reflecting off of one or two walls, considering an omnidirectional sound source, see Fig. 1) for a given room. In order to improve distance perception through the increase of natural audio cues with the simulation of room reverberation without decreasing the horizontal localization performances, a decision was made to simulate only early reflections. The image-source simulation method was used to simulate the early reflections [23]. Each reflection (called image-source) is a copy of the primary sound source coming from a different location. It is attenuated as a function of distance and filtered according to the absorption characteristics of the walls it encounters. These reflections allow for spatial information multiplication through the binaural spatialization of each reflection in addition to the direct sound source.

For the experiment, early reflections are based on the acoustic response of a $5 \times 5 \times 3m^3$ room. The head of the listener is placed at the center of this virtual room at a height of 1m40. 24 image-sources (6 first order reflections and 18 second order reflections) are necessary to simulate first and second order reflections. Their positions are calculated in real-time. Each source is filtered one or two times (depending on the number of walls encountered), then delayed according to the difference between their trajectory lengths and the trajectory of direct sound. In order to reduce computational time due to binaural rendering, the 24 sources are spatialized using a third order ambisonic method rendered over 12 virtual loudspeakers. These virtual loudspeakers surrounding the subject are then spatialized with binaural synthesis at classic positions on a sphere (for more details, see [24, 25]). The resulting binural signal is then mixed with the binauralized direct sound signal. Fig. 2 (left) represents the impulse response of this metaphor effect for two different distances (0.6 m and 1.5 m).

2.3. Geiger Counter (GC)

One of the first sonification applications was the Geiger counter, invented by Hans Geiger in the early 1900's. It consists of increasing the rate of a generated "beep" in proportion to the intensity of non-visible radiation. This well-known metaphor has been successfully tested in a number of sonification applications, and has now become a part of everyday life, used for several commercial applications. For example, it is used on some vehicle reversing/parking aids, which are intended to avoid collisions when reversing a vehicle. As an obstacle comes closer, the warnings become more strident and insistent.

To increase the perception of distance, this effect consists of repeating the stimulus three times and varying the time interval between each repetition as a function of distance. Thus, the closer the target is, the faster the repetition.

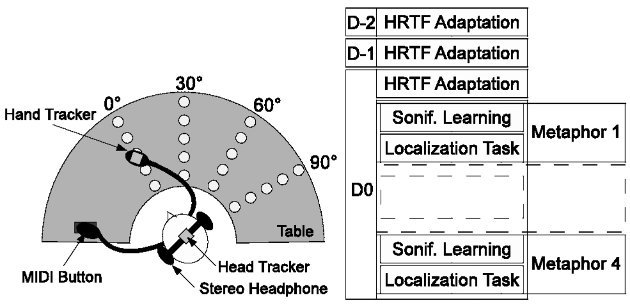
This mapping was chosen so as to avoid any overlap of sounds when the target is near the user, thus the variations were sufficiently noticeable. Time repetitions are therefore of 20 ms at 0.6 m and of 320 ms at 1.5 m, the evolution between these two distances is linear. The sound signal resulting from the application of this metaphor to a 10 msec burst for two different distances (0.6 m and 1.5 m) is presented in Fig. 2 (center).

2.4. Sliding Bandpass Filter (SBF)

Several studies have shown that the used of pitch in data sonification was easily understandable and efficient [26]. The idea of this metaphor is to transpose this sonification concept to an audio effect applicable to any type of sound.

This effect is created using a band-pass filter with a time sliding central frequency and a time varying bandwidth, such that so the quality factor $Q = \Delta f/f$ remains constant (where Δf is the bandwidth and f the central frequency). The initial central frequency of the filter (at T=0 sec, beginning of the sound) is fixed to 200 Hz regardless to the distance. The final central frequency of the filter (at T= sound length, end of the sound) increase proportionally with distance. With this effect, a noise burst will sound as a noisy chirp with a higher final frequency depending on the distance.

For the experiment, the quality factor was fixed to $Q = \Delta f/f = 2$, the final frequency was fixed to 1 kHz for a target placed at 0.6 m and to 8 kHz for a target at 1.5 m. The evolution of the final frequency according to the variation of the distance is linear. Fig. 2 (right) represents the spectrogram of the sound resulting from this effect applied to a white noise burst of 0.5 sec for two different distances (0.6 m and 1.5 m).



(a) (b)

Figure 3: (a) Experimental setup. Small circles = sound source positions (b) Timeline of the experiment.

3. METHODS

3.1. Participants

A total of 16 adult subjects not visually impaired (3 women and 13 men, mean age 28 $\pm$ 6) served as paid volunteers; An audiogram was performed on each subject before the experiment to ensure that their audition was normal (> 15 dB(HL)). All were naive regarding the purpose of the experiment and the sets of spatial positions selected for the experiment.

3.2. Apparatus

A diagram of the setup for the experiment is shown in Fig. 3, with a timeline of the experimental procedure. The first three stages consist of three adaptation sessions with the non-individual HRTFs, see sec 3.3. The next stages consist of the evaluation of each sonification condition with a localization task. During the localization sessions, subjects were seated on a swivel chair located at the center of a wooden circular table of 90 cm in diameter.

The subjects were equipped with a stereo open ear headphone (model Sennheiser HD570) tracked with a 6-DoF position/orientation magnetic sensor positioned on the top of the headphone. They held a position sensor in their dominant hand and interacted with the system using a MIDI button with their other hand. The position of the hand was calculated relative to the tracked center of the head. No headphone equalization was used.

The stimulus used was rendered via a set of non-individual HRTF measured on a KEMAR mannequin (describe in sec 3.3). It was brief to avoid head movement effects and consisted of a train of three, 40 ms Gaussian broadband noise bursts (50 – 20000 Hz) with 2 ms Hamming ramps at onset and offset and 30 ms of silence between each burst. This stimulus was chosen following Dramas *et al.* [27] where the effect of repetition and duration of the burst on localization accuracy was analyzed. Their results showed an improvement of the accuracy between three repeated 40 ms bursts and a single 200 ms burst. The overall level of the train was approximately 60 dBA measured at the ears for a binaural sound source rendered at 50 cm in front of the subject (0° in azimuth and 0° in elevation).

3.3. KEMAR HRTF

The HRTF of a KEMAR mannequin was measured at IRCAM's anechoic chamber. In order to render all the localization test's positions, it was necessary to measure the HRTF over the entire sphere. The set used contained measures from $-90°$ to $90°$ in elevation in steps of 5°, and from $-180°$ to $180°$ in azimuth in steps of 15°. These measures are more precise in elevation, otherwise they have the same characteristics as HRTFs of the LISTEN database [28].

In order to improve the localization performances of the subject with the binaural rendering using this non-individual HRTF, three adaptation sessions of 12 min were conducted according to the method proposed by Parseihian and Katz [29]. Briefly, this method consists of a training game allowing the subject to do a quick exploration of the spatial map of the virtual rendering by an auditory-kinesthetic process. These training sessions were performed three days in a row, twelve minutes per day, the last session being immediately followed by the main experiment.

3.4. Procedure

The experiment was divided into four blocks of 80 trials, each block lasting approximately 15 min. Each block corresponds to a different distance metaphor condition. In order to evaluate the improvement effect of each sonification metaphor, a block of trials without sonification (i.e. only binaural rendering) served as a reference for localization performance. The four blocks are called: *control* (for no sonification), *geiger counter (GC)*, *sliding band-pass filter (SBF)*, and *early reflection (ER)*. For each subject, the blocks were presented in a random order so as to counterbalance any potential task learning effect. Each block of trials began with a short learning session of the sonification metaphor during which the sound was repeated every two seconds. The aim of this learning session was to accustom the subjects to the distance metaphor by allowing them to interact with the distance with an auditory-kinesthetic process. First, for the subject to be aware of the distance ranges and the variations of the acoustic cues, he was asked to move his hand from the inside to the outside of the table and then return, thus two times for two different directions (frontal and lateral). Then, for a periode of one minute, the subject had total control of a virtual sound source spatialized at his hand position and was asked to freely explore the entire surface of the table.

The localization task consisted of reporting the perceived position of a static spatialized sound sample using a hand placing technique validated by a MIDI button. Each subject was instructed to orient himself straight ahead and to keep his head fixed, in a reference position at the center of the system, 0.65 m over the table, during the brief sound stimulus presentation. Before each trial, the subject's head position was automatically compared to the reference position and the subject was asked to correct his position if there was no concordance($\pm$5 cm for the position and $\pm 3°$ for the orientation). After presentation of the stimulus, each subject was instructed to place his hand on the table at the current position of the perceived sound source location and to validate the response with the MIDI button. The subjects were placed in the system in order to use their dominant hand. The perceived position was calculated between the initial head position/orientation when the stimulus was played and the final hand position when the listener validated the target. No feedback was given to the subject regarding the actual target position.

Condition	Control	ER	GC	SBF
Regression slope	0.14 (.09)	0.06 (.13)	0.64 (.29)	0.50 (.20)
Goodness-of-fit	0.66 (.26)	0.27 (.31)	0.96 (.05)	0.92 (.12)

Table 1: Mean linear regression analysis and goodness-of-fit criteria r^2 of the perceived distance. Variances shown in parentheses.

A total of 20 positions (5 different distances relative to the head: 0.73 m, 0.80 m, 0.88 m, 0.97 m, and 1.07 m and 4 azimuths: $0°$, $30°$, $60°$, and $90°$, see Fig. 3), were randomly presented with 4 repetitions each. Subjects had to localize a total of 80 targets and were naive with respect to the set of spatial positions selected for the experiment.

4. RESULTS

The contribution of the sonification metaphors on the perceived distance was analyzed by comparing the distance and azimuth errors of each metaphor (*geiger counter*, *sliding bandpass filter*, and *early reflection*) to those of the control reference condition without sonification (*control*). Because of validation problems with some participants, all trials with a hand position outside the table have been removed from the analysis. Some front/back confusion errors were noticed for rendered sources at $30°$ and $60°$. Since this paper is focused on distance perception, these confusions were corrected before data analysis.

4.1. Effect of the metaphors on the perceived distance

Fig 4 shows the average mean response of perceived source distance as a function of virtual source distance and the mean of linear regression for each condition. It highlights a tendency to overestimate sound distance for the two nearest rendered distances and to overestimate it for the others. It can also be noted that results for *control* and *early reflection* were poorer than those for *GC* and *SBF* conditions. A linear regression analysis was performed on these results. The mean and standard deviation across subjects of the slope of the regression line and goodness-of-fit criteria r^2 for each condition are shown in Table 1. Regression slope lines were far from the unity expected for a perfect distance perception of virtual sound for the *control* and *ER* conditions. For these two conditions there was no real perception of distance. The results for the *SBF* and the *GC* conditions were better with regression slopes nearer to unity but with larger inter-subject variability (highlighted by the large standard deviation).

These results are confirmed by the boxplot of relative distance error shown in Fig. 5. Indeed, the mean errors of the *GC* and the *SBF* conditions are approximately 5 cm lower than those of the *control* and the *ER* conditions. A repeated measures ANOVA was performed on the mean distance error, taking into account three within-subjects factors: metaphor condition (4 levels, fixed factor), rendered distance (5 levels, fixed factor) and rendered azimuth (4 levels, fixed factor). It showed a significant effect of the metaphor condition ($F(3, 42) = 19.76, p < 0.001$), the rendered distance ($F(4, 56) = 12.01, p < 0.001$) and the rendered azimuth ($F(3, 42) = 9.32, p < 0.001$). A Duncan test on categories showed significant differences between *control* and *GC* conditions ($p = 6.10^{-5}$) and between *control* and *SBF* conditions ($p = 2.10^{-4}$). The comparison of *control* and *ER* conditions showed no-significant effects ($p = 0.59$). For the rendered positions, a Duncan test on distance revealed significant differences between the farther distance and the others (highlighting poorer performances for farther distances), and a Duncan test on azimuth revealed significant differences between the lateral angle $90°$ and the others (highlighting better performance for lateral positions).

A thorough study of the perceived distance error while taking into account the effect of the rendered azimuth is shown for all conditions together in Fig. 6 and for each condition in the Table 2. The boxplot highlights better performance for distance perception for lateral sound sources than for frontal sound sources. Regarding Table 2, this slight improvement in performance for lateral sound sources almost appeared for the *control* and the *SBF* conditions

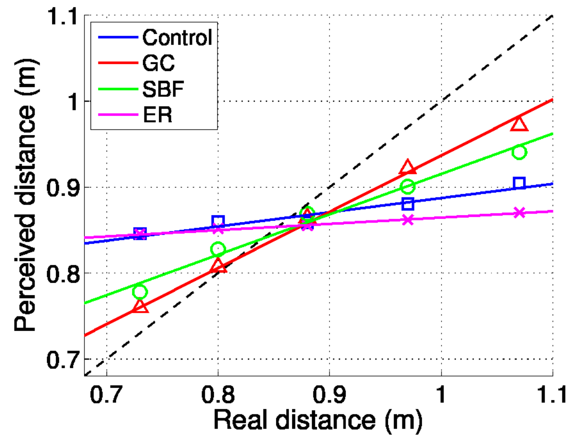


Figure 4: Perceived distances as a function of rendered distance for each sonification condition. «□, △, ∘, ×»: Mean under each condition. Lines: Mean of linear regression.

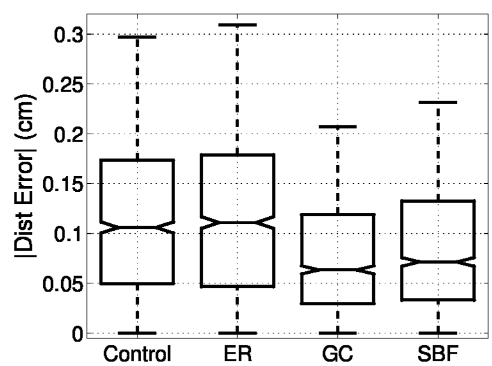


Figure 5: Boxplot of the relative distance error for each metaphor.

Angle	$0°$	$30°$	$60°$	$90°$
Control	0.13 (.10)	0.11 (.09)	0.11 (.08)	0.09 (.08)
ER	0.11 (.10)	0.10 (.09)	0.12 (.09)	0.10 (.08)
GC	0.07 (.08)	0.06 (.07)	0.06 (.07)	0.06 (.06)
SBF	0.09 (.09)	0.08 (.08)	0.07 (.07)	0.06 (.06)

Table 2: Mean distance error (in m) per angle and metaphor. Variances shown in parentheses.

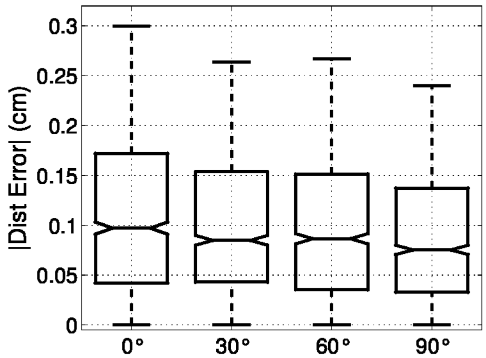


Figure 6: Boxplot of the relative distance for all conditions as a function of azimuth angle.

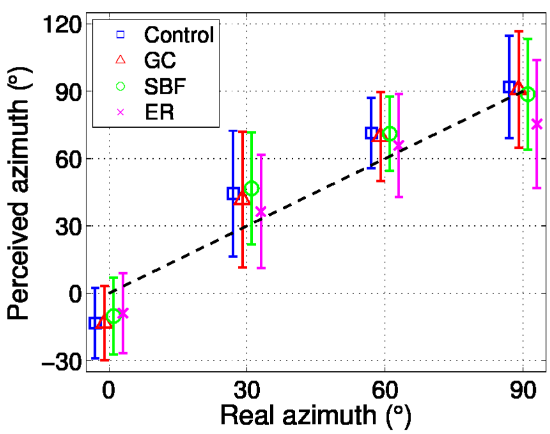


Figure 7: (a) Perceived azimuth as a function of rendered azimuth for each sonification condition. «□, △, ○, ×»: Mean for each condition. Vertical lines: Standard deviation for each modality. For the sake of readability, results corresponding to the different conditions have been slightly horizontally shifted.

(but with a large standard deviation).

4.2. Effect of the metaphors on the perceived azimuth

Although this was not the primary aim of this study, it is interesting to look at the effect of the sonification metaphors on the perceived azimuth angles. Fig. 7 shows the average mean response of perceived source azimuth as a function of virtual source azimuth for each condition. It highlights a large standard deviation mainly at 30° and 90°, and a shift of 10° for frontal sources. Regarding each condition, it appears that the metaphors did not affect the azimuth performances except for lateral sound sources with the *ER* condition.

The mean azimuth error was $20 \pm 15°$. Performing a repeated measurement ANOVA on the relative azimuth error for each metaphor, mixing all the positions, showed no significant effect on the metaphor condition ($F(3, 45) = 0.206, p = 0.89$).

5. DISCUSSION

Regarding the results, of the three designed metaphors, only two most were effective (the *geiger counter* and the *sliding bandpass filter* metaphors) than the control condition of pure binaural anechoic synthesis. Compared to the control condition without sonification (condition whose performances were almost zero for the rendered distances of the experiment), these two effect metaphors improved distance perception significantly. The superiority of the *geiger counter* metaphor over the *sliding bandpass filter* could be explained by their mapping parameters. Indeed, the mapping of the *sliding bandpass filter* metaphor was linear, whereas our perception of frequency is logarithmic. It seems that the variation range of the frequency was not wide enough to be sufficient for a complete rendering of the distances.

Contrary to what was expected, the *early reflection* metaphor failed to improve the distance perception and led to poorer performances than the *control* condition. Furthermore, directional localization at 90° was degraded by this metaphor, which was not the case with the other conditions. To explain this, several observations can be made. First, the chosen model with only early reflections of the first and the second order was too simple, and the absence of the reverberation tail may have affected perception by creating an abnormal situation. Second, all of the studies reporting an improvement of the perceived distance with early reflections were conducted with distances superior to one meter. These cues are perhaps not effective for the shorter distances used in this study.

For all the conditions, but mainly in *control* and *sliding bandpass filter* conditions, perceived distance performance was better for lateral sound sources (especially at 90° azimuth). This improvement, appearing in all conditions, seems to be specific to the binaural rendering. Indeed, in this experiment, distance was linked to elevation as the subjects were 0.65 m over the table. This results in an elevation of $-37°$ for the longest distance and of $-63°$ nearest source. For these elevations, the influence of the torso is more important for lateral sources than frontal sound sources. This probably influenced distance perception. These results are confirmed by the results of a study by Kopco and Shinn-Cunningham [30] that showed better performance for distance perception for lateral sound sources using real sound sources. This result is mainly explained by the variation of Interaural Level Difference (ILD) as a function of distance for lateral sources (due to the shadowing effects of the head) and by the absence of variation for frontal sources (since the ILD is equal to zero).

Regarding the results for directional localization, except for the condition *early reflection* at 90°, there was no effect of distance metaphor on the perceived azimuth. The directional errors were slightly poorer than results with real sound sources (for distance between 0.5 and 1 m, and elevation below $-20°$, Brungart *et al.* [10] obtained a mean azimuth error of 11°). With an average error of 20°, these performances are not so bad considering that the HRTF set used in the experiment contained azimuthal measures at 15° intervals, as well as being non-individualized.

Since the setup of this experiment differs from how previous studies have been organized, precise comparison is impossible. For localization of real nearby sound sources in anechoic environments, distance performance obtained by Brungart *et al.* [10] were from a regression slope of 0.3 for frontal sources to 0.8 for lateral sources. While simulating nearby sound sources with binaural room impulse responses recorded in a reverberant environment, Kopco and Shinn-Cunningham [30] obtained better performance with a mean of regression slope of 0.6 for frontal and 0.8 for lateral distance perception. In this study, with only binaural conditions there was no real perception of distance (regression slope of 0.14). This can be explained by the used HRTFs that were non-individualized and were actually measured at a distance of two meters, so they do not naturally contain near field binaural cues despite attempts to improve performance. In addition, the source positions used in this study, all being in the lower hemispphere, may bias results due to the potential difficulty in this region. With the *geiger counter* and the *sliding bandpass filter* metaphors (regression slope of 0.64 and 0.5), the results approach the performances obtained in [10], thereby highlighting the effectiveness of the adopted method for the sonification.

6. CONCLUSION

The aim of this study was to design and evaluate several metaphors of sound source distance sonification for virtual auditory display. In order to respond to user needs, the designed sonification needed to be independent of the actual sound as well as easy to learn. On the basis of these constraints, the concept of sound effect based sonification was introduced. This new sonification concept consists of the application of an audio effect, whose parameters are dependent on the data to sonify, to any type of sound. With this method, the information is contained in the audio effect and not in the sound. On this basis, three distance metaphors were created and evaluated with sound localization experiments. These experiments underline the contribution of these metaphors to distance perception compared to a control reference condition consisting solely of anechoic binaural rendering. The results highlight a significant improvement of the distance perception with two of the tested metaphors (the *geiger counter* and the *sliding bandpass filter*) in spite of only a short learning period (one minute). It would be interesting to explore the mapping of these metaphors in more detail and their effects on users performance.

The success of these two effect metaphors in improving near field distance perception shows the equivalence of the effect metaphor concept to the traditional parameter mapping sonification applied to sound synthesis. This is a positive result regarding user acceptance of the sonification, which often suffers from a lack of aesthetics.

Since this study was focused on the efficiency of the effect metaphors with "laboratory sounds" (noise burst), further experiments should now be carried out to validate their efficiency with "real sounds" (ecological, instrumental, or electronic sound) in order to approach the real situations and determine if it meets users requirements. Through further studies, it will be interesting to modify traditional parameter mapping sonification strategies into effect mapping sonifications. This will allow for expanded testing based on the findings of this emerging research field.

7. ACKNOWLEDGMENT

This work was supported in part by the French National Research Agency (ANR) through the TecSan program (project NAVIG ANR-08TECS-011) and the Midi-Pyrénées region through the APRRTT program. The authors would like to thank all of the subjects for participating in the experiment, and IRCAM for use of their measurement facilities regarding the acquisition of the KEMAR HRTF. Finally, thanks to Omnihead (http://www.univ-brest.fr/mstis/omnihead/accueil.htm) for the use of the KEMAR dummy head.

8. REFERENCES

[1] B. Katz, F. Dramas, G. Parseihian, O. Gutierrez, S. Kammoun, A. Brilhault, L. Brunet, M. Gallay, B. Oriola, A. Auvray, P. Truillet, M. Denis, S. Thorpe, and C. Jouffrais, "Navig: Guidance system for the visually impaired using virtual augmented reality," *Journal of Technology and Disability*, vol. 24, 2012 (in press).

[2] F. Dramas, S. Thorpe, and C. Jouffrais, "Artificial vision for the blind: A bio-inspired algorithm for objects and obstacles detection," *International Journal of Image and Graphics*, vol. 10, no. 4, pp. 531–544, nov 2010.

[3] B. Katz, E. Rio, and L. Picinali, "LIMSI Spatialization Engine," Inter Deposit Digital Number: F.001.340014.000.S.P.2010.000.31235.

[4] D. Begault, *3-D Sound for Virtual Reality and Multimedia.* Cambridge: Academic Press, 1994.

[5] J. Loomis, R. Klatzky, and R. Golledge, "Auditory distance perception in real, virtual and mixed environments," *Mixed Reality: Merging Real And Virtual Worlds*, 1999.

[6] P. Zahorik, D. Brungart, and A. Bronkhorst, "Auditory distance perception in humans : A summary of past and present research," *Acta Acoustica United with Acoustica*, vol. 91, no. February 2003, pp. 409 – 420, 2005.

[7] P. Coleman, "Failure to localize the source distance of an unfamiliar sound," *J. Acoust. Soc. Am.*, vol. 34, pp. 345–346, 1962.

[8] D. Mershon and L. King, "Intensity and reverberation as factors in the auditory perception of egocentric distance," *Attention, Perception, & Psychophysics*, vol. 18, no. 6, pp. 409–415, 1975.

[9] J. Blauert, *Spatial Hearing.* Cambridge: MIT Press, 1996.

[10] D. Brungart, N. Durlach, and W. Rabinowitz, "Auditory localization of nearby sources. II. localization of a broadband source," *J. Acoust. Soc. Am.*, vol. 106, no. 4, pp. 1956–1968, 1999.

[11] B. G. Shinn-Cunningham, S. Santarelli, and N. Kopco, "Tori of confusion: Binaural localization cues for sources within reach of a listener," *J. Acoust. Soc. Am.*, vol. 107, no. 3, pp. 1627–1636, 2000.

[12] D. Devallez, F. Fontana, and D. Rocchesso, "Linearizing auditory distance estimates by means of virtual acoustics," *Acta Acustica united with Acustica*, vol. 94, no. 6, pp. 813–824, Sept 2008.

[13] F. Fontana, D. Rocchesso, and L. Ottaviani, "A structural approach to distance rendering in personal auditory displays," in *IEEE International Conference on Multimodal Interfaces (ICMI 2002)*, 2002.

[14] G. Kramer, *Auditory Display: Sonification, Audification and Auditory Interfaces.* Perseus Publishing, 1993.

[15] T. Hermann, A. Hunt, and J. Neuhoff, Eds., *The Sonification Handbook.* Berlin, Germany: Logos Publishing House, 2011. [Online]. Available: http://sonification.de/handbook

[16] B. N. Walker and G. Kramer, "Mappings and metaphors in auditory displays: An experimental assessment," in *Proceedings of the 3rd International Conference on Auditory Display (ICAD96)*, S. P. Frysinger and G. Kramer, Eds., 1996.

[17] C. Sikora, L. Roberts, and L. Murray, "Musical vs. real world feedback signals," in *Conference companion on Human factors in computing systems*, ser. CHI '95. New York, NY, USA: ACM, 1995, pp. 220–221.

[18] D. Brungart and B. Simpson, "Design, validation, and in-flight evaluation of an auditory attitude indicator based on pilot-selected music," in *Proceedings of the International Conference on Auditory Display (ICAD2008)*, Paris, France, 2008.

[19] L. Brunet, "Étude des besoins et des stratégies des personnes non-voyantes lors de la navigation pour la conception d'un dispositif d'aide performant et accepté (Needs and strategy study of blind people during navigation for the design of a functional and accepted aid device)," Master's thesis, Departement of Ergonomics, Universitée Paris-Sud, Orsay, France, 2010.

[20] S. Nielsen, "Auditory distance perception in different rooms," in *Audio Engineering Society Convention 92*, March 1992.

[21] D. Begault, "Perceptual effects of synthetic reverberation on three-dimensional audio systems," *J. Audio Eng. Soc*, vol. 40, no. 11, pp. 895–904, 1992.

[22] G. Kearney, M. Gorzel, H. Rice, and F. Boland, "Distance perception in interactive virtual acoustic environments using first and higher order ambisonic sound fields," *Acta Acustica united with Acustica*, vol. 98, no. 1, pp. 61–71, 2012.

[23] J. Allen and D. Berkley, "Image method for efficiently simulating small-room acoustics," *Acoustical Society of America Journal*, vol. 65, pp. 943–950, Apr. 1979.

[24] A. McKeag and D. S. McGrath, "Sound field format to binaural decoder with head tracking," in *Audio Engineering Society Convention 6r*, 1996.

[25] M. Noisternig, A. Sontacchi, T. Musil, and R. Holdrich, "A 3d ambisonic based binaural sound reproduction system," in *Audio Engineering Society Conference: 24th International Conference: Multichannel Audio, The New Reality*, 2003.

[26] L. M. Brown, S. Brewster, R. Ramloll, M. Burton, and B. Riedel, "Design guidelines for audio representation of graphs and tables," in *Proceedings of the 9th International Conference on Auditory Display (ICAD2003)*. Boston, USA: Boston University Publications Production Department, 2003, pp. 284–287.

[27] F. Dramas, B. F. Katz, and C. Jouffrais, "Auditory-guided reaching movements in the peripersonal frontal space (poster)," in *Acoustics 08, Paris, 29/06/2008-04/07/2008*, vol. 123. J. Acoust. Soc. Am., 2008, p. 3723.

[28] "IRCAM LISTEN HRTF database," http://recherche.ircam.fr/equipes/salles/listen/.

[29] G. Parseihian and B. Katz, "Rapid head-related transfert function adaptation using a virtual auditory environment," *J. Acoust. Soc. Am.*, vol. 131, no. 4, 2012.

[30] N. Kopco and B. Shinn-Cunningham, "Effect of stimulus spectrum on distance perception for nearby sources," *J. Acoust. Soc. Am.*, vol. 130, no. 3, pp. 1530–1541, June 2011.