DISTANCE AND SHAPE AUDITORY ASSESSMENT IN NATURAL ACOUSTICS

TOWARDS THE DESIGN OF VIRTUAL SOUND SCULPTURES

CORALIE DIATKINE
Sonic Art Research Center, Queen's University Belfast
cdiatkine01@qub.ac.uk coraliediatkine@gmail.com

Abstract

This paper aims to provide a critical standpoint for the design of virtual sound sculptures. Few studies have been lead in that field, and some difficulties arise when combining creation and research. Firstly, perceptual and cognitive factors interact in sound shapes perception. This interaction is all the more complex as it occurs in ecologically valid conditions[1], that is in natural environments. Secondly, there is no obvious equivalency relationship or straightforwardness between the sense of proportions, depth, density, or distance in visual and in auditory perception, which have their own mechanisms. Thirdly, real life ranging depends on source and space properties that include spectral complexity and reverberation. We will particularly investigate distance and sound shape assessment, which relate to three-dimensional sound sculpture design through the review of recent study. This study will lead to spatial audio installation and experimental project. Elementary synthetic signals will be spatialised in a natural room acoustics, using decorrelated point sources. Combining binaural and ambisonics techniques will allow the internal and external exploration of auditory scenes in the close and far field. Measuring distance and shape assessment could allow the study of the influence of sound and space properties and auditory system, on perception of sound shapes. This could contribute to the design of perceptually consistent auditory scenes.

1. INTRODUCTION

Although loudness was long considered at a major cue in ranging, this role must be relativised. Reverberation especially takes an important part in distance assessment, and cognitive or perceptual cues should also be taken into account. Besides, interestingly, perceived distance in natural environments can be correlated to physical distance by an approximate psychophysical function. Sound shape perception is investigated from a more technical point of view, as current literature on the subject remains limited. We will describe the use of the Ambisonics spatialisation and decorrelation techniques to design bi-dimensional sound shapes, and review some conclusion regarding relationships between perceived and intended sound shapes. This review has allowed to conceive a framework for a spatial audio project of installation and experiments in the domain of the design of virtual sound sculpture, which is overviewed at the end of this paper.

2. ACOUSTIC AND PERCEPTUAL CUES FOR DISTANCE ASSESSMENT

Distance assessment, or ranging, proceeds from the *degradation of a signal during its propagation through space* [2]. This implies two things. Firstly, this process requires prior knowledge[1] of both *source and space characteristics, or attributes.* Secondly, from a cognitive stand point, space and source characteristics are *undissociated* in distance assessment: perceived sounds carry information about sources *and* the space they are located in. This is why, in *ecologically valid conditions* [1], there is no such thing as the neutral space of an anechoic room, which is precisely used for the independent study of signal properties. Hence, in order to provide an ecologically valid framework to this project, sound propagation should be considered in a real atmosphere and space, which implies a greater number of factors than in a non reverberant. Several signal alteration processes occur during propagation. These can be due to the sound propagation acoustic process itself, or to the environment's physical properties. Their interaction results in *overall* or *frequency-dependent* changes in the signal structure [2]. These processes are reviewed below.

2.1. Preliminary remarks

Prior knowledge of source and space characteristics is necessary to distance assessment. From a cognitive standpoint, this depends on the *type of cue* or *signals* that are apprehended by the listener. We are for instance able to range sources using the reverberation cue of a given space because our routine experience has exposed us to thousands of rooms impulse responses. And ranging accuracy is higher for familiar sounds, which we have an extended experience of.

Auditory assessment of distance can be *relative*: the comparison of the distances of several sources is used to provide information on the distance from one source. It can be a*bsolute*: the distance of one source is assessed without resorting to comparison. Relative assessment is considered to be an easier task than absolute assessment, and we do not use all acoustic cues the same way to perform one or the other task.

2.2. Energy attenuation and loudness

Loudness has long been a commonly accepted cue to distance. Nevertheless, its value as a distance cue varies with additional factors.

2.2.1. Energy attenuation in space

The energy attenuation of sound during their propagation results in an *overall loudness attenuation*. It is *independent from the propagation space properties*.

[1] In the field of cognition, knowledge is the result of a process of information storage achieved through experience. Learning can be almost immediate or require a more extensive experience of sources and space characteristics.

The relationship between distance and energy attenuation, is expressed by an *inverse square law*:

$$Lp = Lw - 20 \log r - K' \quad (1)$$

where Lp is the sound pressure level (dB), Lw is the sound power level (dB), r is the distance from the source (m) and K' a constant [14]. With a spherical attenuation, K'= -11, and with a hemispherical attenuation, K'=-8.

2.2.2. Loudness and perceived distance dynamics

Nielsen [12] compared the values obtained in several experiments on ranging carried in anechoic, or quasi anechoic conditions , with values obtained in his own test (Tab. 1). Dynamics in perceived distance *is much lesser than the dynamics in loudness*. Whereas a *6dB level difference corresponds to a doubling in loudness,* this ratio is multiplied by 3-3.5 in anechoic conditions, and by 7 in reverberant conditions.

	Rel. level	**Dist. factor**	**Space**	**Signal**
Gardner [13]	20 dB	2	anechoic	speech
Petersen	21 dB	2	anechoic	pure tone
Laws	20 phons	2	free field	unknown
Mershon	10 dB	1.6	anechoic	w. noise
Nielsen	17 dB	2	anechoic	speech
Nielsen	22 dB	2	listen. room	speech
Nielsen	41 dB	2	classroom	speech

Tab. 1. Loudness and perceived distance dynamics in anechoic and reverberant environments. At 0° azimuth in anechoic conditions, a 20dB decrease of loudness is necessary to increase the distance by a factor of 2. This value is doubled in reverberant conditions [12].

2.2. Relativity of loudness as a ranging cue

2.2.1. Intrinsic sound source amplitude

The judgement of humans or animals is less accurate for signals whose amplitude is likely to vary independently from distance, and especially if this is done independently from spectral cues variations. A lyric voice is trained to develop its *degrees of freedom* in timbre, loudness or pitch independently from each other[2]. Animals can modulate the amplitudes of the sounds they produce to deceive a rival, a prey or a predator. On the other hand, distance is easily assessed for signals with invariant amplitude, such as whistles, familiar bird songs, or speech [2].

2.2.2. Cognitive factors and interaction with other cues

The reliability of loudness is influenced by cognition, which associates familiar sound sources with a distance, a gesture, an intention, can lead to biased judgements in unnatural situations. Whispers and shouts are associated with proximity and remoteness[18]. The dominance of visual on auditory cues is also well known – with ventriloquism effect for instance.

Reciprocally, the weight of loudness in ranging can be regulated by other source or space acoustic cues, such as transient energy, duration, spectral envelope, spectral components, or reverberation. At equal loudness, we can discriminate the timbre of a piano when strings are hammered softly or strongly. We perceive differently the sound of a violin played with a mute in the very near-field from the sound of a violin played reasonably loud from backstage. In the first case, the sound carries a timbre effect. In the second case, it carries an evocation of distance and space.

2.5. Room attributes: reverberation

Loudness is considered as the main cue for distance assessment *in absence of reverberation*. But reverberation plays a farther greater role in ranging in natural spaces. It is a complex phenomenon, whose main characteristics will be recapitulated here.

2.5.1. Reverberation time and room impulse response

A sound travelling in a room also informs us on its properties : volume and materials. Rooms acoustics is characterised by a reverberation time (*Tr*) and a room impulse response (RIR). The *Tr* is function of the volume and total absorption area, which also depends on the absorption coefficients of the materials that compose the surfaces. *Tr* is calculated as follows

$$Tr = \frac{0.161 \cdot V}{A} \quad (3)$$

Where V is the volume of the room (m^3), A is the total absorption area (m^3). As frequencies are not absorbed equally, the *Tr* and the RIR depend on these absorption coefficients.

2.5.2. Reverberation as an absolute or relative cue

Reverberation being a space-specific cue – is doesn't characterise signals – listeners can extrapolate from a routine extended experience of reverberation and dry signals to instantly assess the distance of reverberated sounds [2]. The learning process of reverberation is unconscious and instant when entering a room.
Mershon, King and Bowers [4] [5], or Nielsen [12], have shown that reverberated sounds are judged more distant than sounds in anechoic environments, *even when subjects have no experience of the sound source or space characteristics*.

[2] The *degrees of freedom* of the sound source parameters depend on the mechanical characteristics of the instrument, and the technique of the instrumentalist. A change in amplitude can then have a more or less pronounced effect on the relative amplitudes of the components. And of course an expert listener will use both amplitude and spectral cues to range the distance of an instrumental source.

2.5.1. Direct sound to reverberated sound ratio as a cue

This ratio is function of the *distance to a source and of the sound source directivity and location*. It is *a very reliable relative or absolute cue[3] in distance assessment*. In a natural environment, ranging is quite accurately correlated with the distance from the sound sources, whereas there is no relationship between physical and perceived distance in anechoic environments[12] [13].

2.3. Scattering, atmospheric absorption, excess attenuation : spectrum and amplitude modifications

These factors are responsible for the *frequency-dependent attenuation of the signal*, affecting mostly the amplitude of high-frequencies, and thus the *timbre* of the perceived sound.
•Scattering depends on the relationship between wavelength and obstacles width.
•Atmospheric absorption, is function of air humidity, temperature and frequency at long distances [3].
Both affect high-frequencies.
•The global Excess Attenuation (A_E) results from a conjunction of factors in open-field conditions[4].
These factors might be of some interest for the creation of illusory open-air large spaces, or changes in spatial dimensions. As we are not dealing with complex spatial simulations and do not have sufficient material, they won't be treated in this paper.
Like amplitude, high frequencies attenuation in anechoic conditions doesn't allow an accurate ranging, and requires repeated presentations of signals to improve [6].

2.6. Ear selectivity in acoustic cues processing

The *overall effect* of acoustic cues on the signal can be measured with global methods such as *cross-correlation of the waveforms* at the source and near the listener. Nevertheless, the ear temporal and frequency selectivity allows to perceive and analyse these processes *separately,* and *selectively*.
This allows a selective apprehension of the various attributes that characterise open air or closed spaces – volume, materials, atmospheric effects–, and the relative preservation of the intelligibility of the signals, regardless of those space attributes. Let us consider reverberated bird sweeps (Fig. 2) [2].

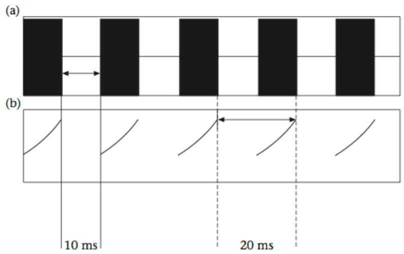

Fig 2. Waveform (a) and spectrogram (b) of 10 ms bird frequency sweeps separated by 10 ms gaps [2].

High frequencies are extended from the end of a sweep until the end of the next, without affecting the perception of low frequencies and the discrimination of the sweeps. The selective analysis of reverberation allows the auditory system to discriminate the pulses of the original signal.

2.7. Binaural cues

Binaural cues – interaural time (ITD), level (ILD) or envelope (IED) difference – vary with the location of sources in the median, horizontal and frontal plane relatively to the listener. Few studies about the role of these cues in ranging have been conducted. They should be reviewed for the continuation of our research.

2.7.1.Use of binaural cues in the near-field

The use of these cues in ranging is considered as controversial [2] by some authors. However, it seems these can inform us of the distance to a source in the near-field [8].

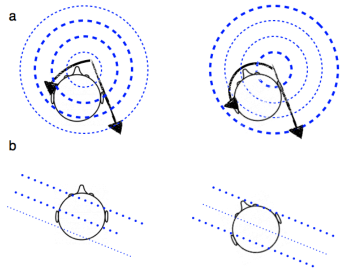

Fig. 3 Propagation of a high frequency sound wave for 30° and 90° azimuths, in the near-field (a), and in the far-field (b).

Below 1 m, the spherical propagation of sound waves results in binaural differences, that *are function of the radial distance to the source*. (Fig.3). In the near-field, with spherical waves, interaural differences are stronger than in the far-field, where waves can be considered as planar. Binaural differences increase from the median axis (0° azimuth) to the frontal/interaural axis (90° azimuth) .

2.7.2. Interaction of distance with frequency

Besides, the weight of the different binaural cues in localisation varies for frequencies below or above 1500 Hz, due to the relationship of the wavelength to the size of the head (Fig. 3).
This aspect of ranging is all the more so significant as *sound sculptures are related to the design of immersive auditory scenes,* which can exacerbate the sense of proximity with virtual sound sources. Sound creation is free from realism, and the design of closely surrounding sound patterns (or even internalized patterns we don't experiment in real conditions is of obvious interest.

2.7.3. Correlating HRTFS to distance estimates

More reviews regarding correlations of HRTFs and ranging in the near-field by psychophysical functions –in reference to[7] and [8]– need to be done. Such functions

[3] An absolute cue gives information about the properties of the sound at the first presentation. It requires past experience. A relative cue gives information on changes [12]

[4] These factors are the vegetation, obstacles, meteorological conditions, ground interaction,which are responsible for frequency-dependent reverberation, and amplitude fluctuations.

could be computed from data collected in ecologic conditions, and the weighting of acoustic parameters.

3. Auditory assessment of distance in ecological conditions : two experiments

The perception of distance under reverberant conditions has been seldom investigated. In laboratory conditions, auditory cues can be separately to assess distance, which leads to a relative degree of uncertainness. In natural conditions, the integration of of multiple cues increases the accuracy of judgement, and allows to compensate the absence of a specific cue. It also allows ranging in various types of spaces [2].

3.1. Perceived distance as a function of physical distance in reverberated conditions

Recent works by Zahorik [7] and [8] have allowed to establish the existence of a *strong bias in our estimation of distance*, *and furthermore, to approximate the relationship between perceived and physical distance to a psychophysical function*.

3.1.1 Experimental setup

Zahorik aimed to find out a *relationship between physical and perceived distance*, and to evaluate the weight of each acoustic parameters in ranging. The weighting of the acoustic factors is complex and is not exposed here. In a first experiment, the sound source consisted of one loudspeaker whose distance varied from 0.3 to 13.79 m exponentially for a 0° and 90° azimuth in a small auditorium[5]. In a second experiment, the sound source and the room acoustics were then simulated with individualised binaural room impulse response functions (BRIR) and headphone impulse-response (IR) measurements. Listeners were presented a noise burst, and a speech signal, and were asked to range the sources in familiar units[6]. The average individual ranging was plotted as a function of distance, signal and azimuth [7].

3.1.2. Analysis of the results

A psychophysical distance function was approximated by a *compressive power function* [8] :

$$r' = kr^{a} \qquad (2)$$

Where the exponent a varies from 0.15 to 0.70, depending on the listeners, source and direction.
This function gives a good approximation of the relationship between perceived and physical source distance, *regardless the type of signal or the azimuth* [8].

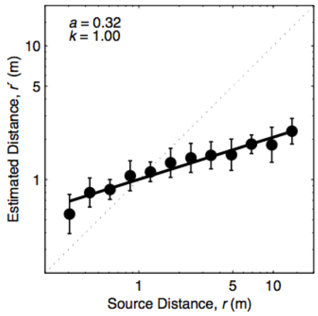

Fig. 4. Average ranging as a function of physical distance for one listener. Bars denote a standard deviation [7]. Stimulus and azimuth is not specified.

3.1.3. Conclusion : a compressed auditory space

Listeners overestimate distances inferior to 1 m, while they consistently underestimate distances superior to 1m. The auditory space seems compressed, with a center of gravity located at a further distance than immediate proximity. A distance of 1 m would be the limit of our safety zone, in terms or proxemics. Ranging below this threshold could be investigated more precisely.

3.2. Towards a mental toponymy of auditory spaces

Such results are meaningful for the design of volumetric acoustic structures or virtual auditory displays. Perceptually consistent three-dimensional sound patterns require a mapping of the targeted physical distance to its auditory representation.

3.2. Influence of level and distance on ranging in anechoic and reverberant conditions

Experiments lead by Nielsen [12] bring interesting results regarding the influence of level and reverberation on ranging.

3.2.1. Aim of the experiment

Nielsen studied the effect of four variables: reverberation time and direct sound to reverberated sound ratio, level, and azimuth. *Results in anechoic conditions differ from results in reverberant conditions.* Some conclusions seem to corroborate Zahoric's results.

3.2.2. Experimental setup

A dry listening room, a reverberant classroom and an anechoic room were used. Logarithmically equidistant loudspeakers were placed from 1m to 5m. Signals were played at 58, 68 and 78 phons, at 0°, 45°, 90° and 180° azimuth. Listeners were asked to mark the sound source position on a 8m radius circular plane. The loudspeakers were hidden by a curtain.

3.3. Reliability of reverberation vs. level

3.3.1. Results in reverberant conditions

In reverberant conditions, distance has the same effect, regardless reverberation time or room impulse response.

5 TR60 is displayed in [7]

6 According to the author, the method produces similar results as using paired-comparisons or walking to the perceived source location.

a)There is an obvious relationship between direct noise to reverberated noise ratio and perceived distance. Yet, it can't be approximated to a psychophysical function.
b)Ranging is *almost independent* from loudness, with a tendency to overestimation at low levels and more accurate judgement at high levels.
c)Loudness influenced ranging to a greater degree in the drier room.
d)The dispersion of answers doesn't vary, but it improves with the distance and the level, which would indicate a larger consensus and a greater reliability in distance estimates for a given threshold of information.

3.3.2. Notes

•The greater accuracy noted in b) can reveal a tendency to overestimate short distance and underestimate long distances, which would corroborate Zahoric's conclusions in [7] and [8].
•Similar results in reverberant rooms probably result from the listeners extended knowledge of reverberation: *Rt* is not a variable in reverberated environments.
•The relative role of loudness suggests that cues are combined and *weighted depending on the amount of individual information they carry.*

3.2.3. Results in anechoic conditions

Ranging is inaccurate, as it is only correlated to the level. *In absence of reverberation*, not surprisingly, loudness is the only criteria for distance assessment. This is in the agreement of the common idea regarding the role of loudness in ranging. This constitutes a strong bias for sources whose level is subject to vary.
As in a reverberant environment, dispersion improves with loudness.

3.3 Relative influence of binaural cues

3.3.2. Binaural cues

Binaural differences do not seem to contribute to the ranging accuracy, although the results for various azimuth could not be systematically compared. The statistic analysis shows that distance at 45° is 4% and 5% larger in the listening and class room than at 0°. Sources at 90° and 180° also seem more distant. Note that in the anechoic room, this effect is more pronounced : at 45°, sources appear 17% more distant.

Conclusion

These tests open perspectives for further experiments over a larger ranger of azimuth and elevation, which could hopefully lead to a tridimensional cartography of our mental representation of auditory space.

4. SOUND SHAPE PERCEPTION

Sound shape perception also relates to the design of virtual sound sculpture. To our knowledge, there exist almost no studies in this domain. Potard and Burnett [15] looked for a relationship between the intended and the *perceived source extent* in two dimensions : *height* and *width*, measuring the accuracy of sound source extent assessment in 3D audio environments, in two consecutive studies[15] [16]. We will review key notions regarding sound shapes properties and perception, and presents some of the authors conclusions.

4.1. Prior definition of sound shape

A perception-based approach to the design of three-dimensional sound sculptures should be highly dependent on the relationship between acoustic and perceptive cues.

4.1.1. Definition

We would like to suggest a possible definition for sound shapes in analogy with sound scenes [1], as a *coherent group of point sources*, whose arrangement in space is susceptible to change in time.

4.1.2. Sound shape attributes

These attributes can be considered as a set of relationships between a set of points in space. Height, width and depth represent the *spatial extent delineated by point sources* in the three dimensions of space. Density represent the *amount of point sources* perceived within these limits. Potard and Burnett [15] actually only mentions horizontal and vertical *wideness* – to specify that we are considering perceived dimensions – which suggests that for some reason, the study of shape perception is restricted to a bi-dimensional space. But it seems relevant to us to consider that all three dimensions define the *spatial extent*, or size of the sound source.

4.2. Acoustic and perceptive cues

In a first review [15], Potard and Burnett mentions a number of acoustic cues that influence the *wideness* of sources :

Signal duration [19]*:* more information on the subject need to be collected, but it is likely that long signals occupy a larger perceptual space than short signals

Frequency – Pitch : long waves need a greater distance to unfold. Low pitch sources occupy a larger space. The author doesn't specify whether sources are complex sounds or pure tones.

Distance–loudness [20]: a distant source seems narrower than a close source, as its loudness decreases proportionally with distance. The author seems unclear about the conditions in which such conclusions were drawn, and this assertion should be considered very carefully, given the relativity of loudness and the importance of other cues such as reverberation that previously highlighted.

Interaural Cross Correlation Coefficient [15] and [16] sound source extent seems based on two dimensions: vertical and horizontal wideness (azimuth and elevation). These depend on rom a binaural cue the inter-aural cross correlation coefficient (IACC). A particular stress is given on the IACC in this section.
Relative distance
According to Potard and Burnett, sound shapes seem reduced to a surface, a bi-dimensional space of azimuth and elevation, where point sources are equidistant from

the listener[7]. We draw the hypothesis that *source or scene depth*, which particularly interests us in the context of our study, more specifically relies on a third dimension : *relative distance assessment*.

4.2. Interaural Cross Correlation Coefficient

IACC measures the *relationship of the signals at the two ears*. It is correlated with the senses of diffuseness, narrowness and wideness of sound sources, or spaciousness of rooms. Highly correlated sounds are perceived as a single auditory event and [16 Blauert] with a center of gravity.

It is the maximum value of the IACC function defined as

$$IACC(\tau)=\frac{\int_{-\infty}^{+\infty} s_L\,(t-\tau)\,s_R(\tau)dt}{\sqrt{\int_{-\infty}^{+\infty} s_L^2\,dt \int_{-\infty}^{+\infty} s_R^2\,dt}}$$

Where *$s_L(t)$ and $s_R(t)$* are the let and right ear signals, τ is a time constant, *t* is the time.

A maximum value of 1 means that sounds will be localised in the same place. A nil value means that sounds will perceived as diffuse.

Due to signal properties and to unavoidable movements of the body, these coefficients can change in time. Dynamic changes of values can be computed in order to model a *running IACC function*, which will allow to implement a *time varying decorrelation* in the design of sound shapes. IACC coefficients can also be frequency dependent. This relationship can be measured in 1/3 octave sub bands.

4.3. Building sound shapes with decorrelated Point Sources

The use of decorrelation in spatial audio for the creation of auditory scenes has been largely descrived by Kendall [21]. Potard and Burnett [15] describes a technique using spatially distinct point sound sources decorrelated from one another to create broad sound sources. We cannot detail this technique here but it will be reviewed for further research.

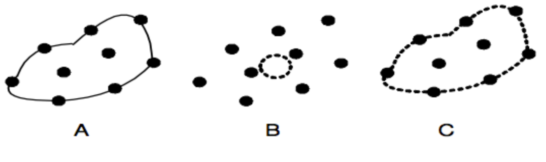

Fig.5 Decomposition of a source into point sources (A), high correlation creating a narrow sound image (B), low correlation creating a wide sound image (C) [15].

•***Time-invariant decorrelation***: a small delay time is introduced between the signals, but Potard and Burnett underlines undesired comb -filtering effects, and the limitated number of signals this technique can generate.

•***Time-varying decorrelation*** is achieved using all-pass filters with a randomized phase response calculated at every frame. It allows to simulated fluctuations caused by moving air, which is interesting with regards to the realism it can bring to a scene, or to create possible other effects.

•***Sub-band decorrelation*** allows to create sound sources whose spatial extent varies with frequency – this is known as Fourier decomposition effect.

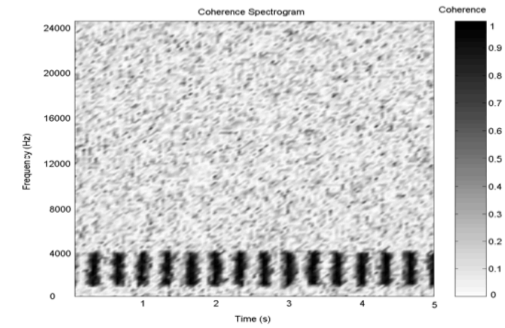

Fig. 6 Periodical correlation and uncorrelation of two signals within one common frequency band [16]

•***Real or virtual sources***: point sources can be real sources produced by loudspeakers, or virtual sources. Ambisonics, starting from first order. seems to be an efficient reproduction technique.

4.4. Assessment of sound shape extent

4.4.1. Experimental framework

For this purpose, six point sources fed with white noise were spatialised with fourth order ambisonics technique, on a 16 speaker dome array. The IACC coefficient of the point sources was 0. Subjects were placed at the centre of the dome, facing a 0° azimuth. They could move their head. They were asked to draw the shape they perceived on a top down view of the dome array.

49 sound patterns were designed : horizontal lines, vertical lines and squares.The extent of the sources varied from 40 to 180°.

4.4.2. Conclusions

Wide sources localisation is generally good, with a coarse approximation of spatial extent. Subjects can also easily discriminate vertical, horizontal and square sources with different extents. Narrower sources are perceived as more narrow than they are.

It appears that, more than actual sound shapes, it is possible to display *sound areas* using sound source extent.

A number of factors could improve sound shape perception: *trained subjects* would have a better assessment of sound shapes. The number of point sources or *density* of point sources, as well as spatially smaller loudspeakers would contribute to *sharpen the sound shape definition*. The author intends to lead further experiments to study the influence of these factors.

[7] Some anamorphosis might be necessary to map surfaces and volumes from the euclidian space to a spherical propagation space.

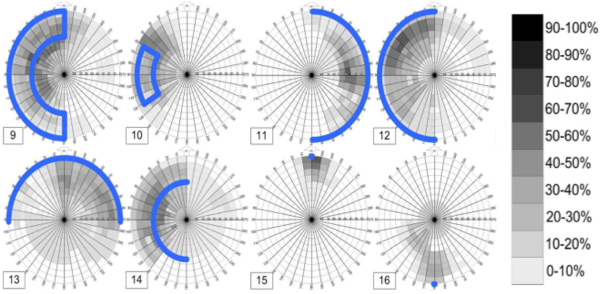

Fig. 7 Some of the density graphs generated from the areas drawn by the subjects. The mean perceived source extent is represented in shades of grey. The intended sound source extent is represented by a thick line [16].

4.5. **Notes**

Taking *depth* and *density* perception into consideration could allow a better understanding of three-dimensional shapes perception in the context of three-dimensional sound-shapes[8]. We don't know of any studies on density. We assume that depth perception then results from *relative distance assessment*. Few studies relate to the perception of depth. Kearney, Gorzel, Boland and Rice [17] have recently investigated perception of source depth using High Order Ambisonics (HOA). We have little knowledge of the results, but it seems that, for each order, the Ambisonics techniques allows a good reproduction of real sources distance.

5. SPATIAL AUDIO PROJECT OVERVIEW

We briefly expose the framework of a spatial audio installation which should be a first step towards a of three dimensional sound shapes design project. This installation takes on both an artistic and and experimental dimension. The perception of several dimensions of sound sources can be studied through this installation : ranging, extent, depth and density. An analogy between auditory scene and sound shape seems relevant as both objects imply a global perception of a coherent ensemble [1].We hope to study the influence of several variables on shape perception.

5.1. Ecologic validity

In order to provide an ecologically valid framework for this installation and for the experiment, a reasonably reverberant acoustics should be simulated.

5.2. Point source paradigm

In reference to Potard and Burnett [17], multiple point sources will be arranged in space to create three dimensional sound patterns. The resulting perceived shapes will result from multiple levels of *relationships* between these individual point sources. This choice is also based on the fact that relative ranging – and possibly localisation – are more precise than absolute ranging.

5.2. Auditory scene characteristics

5.2.1. Maximum broadcasting range

The broadcasting limits of the scene seems to have an influence on relative ranging. The Just Noticeable Difference (JND) in distance is a function of distance, frequency spectrum, and amplitude. With constant spectrum and amplitude, the ratio between JNDs for degradation and distance decreases with the distance of the source [2]. The JND in distance would determine a broadcasting limit.

5.2.2. Sound shape and listener's location

In a first set of tests, listeners will sit in the center of the plane. In a second set, they will sit outside the plane. In order to limit the number of variables that might interfere with the listeners assessment, elevation will not be considered. Shapes will be arranged with a mean height and around a 0° elevation will be used.
Ranging, extent and density assessment can be studied for various azimuth from 0° to 180° with a fixed width of the auditory scene. Changes in the positioning of the listener will not be treated in the tests. In the context of the installation, listeners will be free to scan the space to experience sound shapes form an internal immersive, or from an external standpoint.

5.3. Signals properties and production

5.3.1. Signals properties

Unfamiliar signals should be used for perceptual and functional reasons. The spectral characteristics of elementary synthesis sounds are easier to control than those of familiar recorded sounds. It was noted earlier that our ranging abilities differ for familiar and unfamiliar sounds. Perception of shape should be based on shape as a whole, and not on individual components whose familiarity might distract the listener. We will also invoke personal aesthetic taste.

We also assume that homogeneous signals seem necessary to facilitate the perception of a coherent sound shape. In order to get an even spread of each point source, signals should belong to a common frequency band for each sound pattern for a first set of experiments. Likewise, they should have homogeneous spectrum, duration and loudness, with randomised variations. We assume that these frequencies should be chosen in order to obtain wave whose length would be compatible with the target source extent, sharpening and density. The spectrum of the signals should be harmonic or quasi harmonic and remain limited in the frequency domain, to ease manipulations the analysis of the influence of the variable.

An interesting approach to get a sense of coherent volume could be based on the reproduction of the own resonance mode of the structure, based on the components frequency relationship.

Absolute / relative spectral centroid of sounds, which would depend on the modes of the membrane, can vary in function of the localisation of the excitation, or with the synthesis in MaxMSP. Prior experience with the

[8] Shape is a polysemic concept that includes bi and three-dimensional shapes.

manipulation of physical modelling has shown us that damping only characterises the timbre of percussive sounds and that pleasant results can be achieved with spectrums constituted by a single partial.

5.3.2. Signals production

The signals of the point sources used for the installation can be first created with resonant filter synthesis algorithms designed in Max MSP. Eventually, the projection of granular material into membranes, simulated by physical modelling should be used. These membranes delineate the extent of the scene or shape in the three dimensions and are a means to represent and simulate a *homogeneous sound structure with a physical coherence*. On the long term, it might be interesting to investigate the influence of acoustic characteristics' heterogeneity or changes in sound shapes perception. The signals can be decorrelated dynamically. Nevertheless, this technique was found to have a distracting effect, and provoke fatigue. It should hence be used with care. We believe that at this stage of our research, a sub-band decorrelation could constitute a bias in the experimental context of the project. This could affect the localisation of each point source, creating an auditory blur that could degrade the shape's definition. The simulation of randomised micro movements can also bring more liveness and organicity to the shape.

5.4. Spatialisation Techniques

Two different spatialisation techniques can be used in this installation.

Binaural technique is an efficient three dimension reproduction technique. It can enable internal exploration of sound patterns and spatialisation within 1m. Additionally, a head tracker combined with can contribute to simulate a steady sound scene while allowing head movements. Ideally, individualised HRTFs could improve the accuracy of sound shape perception. Non-individualised HRTFs can affect the directional localisation of sounds: individualised HRTFs would allow optimal ranging with varying azimuths [8]. The efficiency of Ambisonics technique has been demonstrated in various studies. Although consistent result were obtained with first-order ambisonics, the use of third of fourth order Ambisonics would allow the widening of the sweet spot of the scene and a sharpening of the sound shape definition.

5.5. Experimental Variables

The influence of different variables on sound shape assessment can be studied, but a choice will have to be made due to schedule constraints.

5.5.1. Signal properties

•The different acoustic properties of the signals can be considered individually, and relatively, that is from the relationship between the point sources – for a homogenous class of signals–. The influence of the degree of homogeneity of these properties on shape perception might be also an interesting factor to consider.

• Interaural Envelope Differences – spectrum differences between ipsi and contraleteral ears, ITD and ILD will interfere with ranging depending on the azimuth. The way they should be handled remains to be defined.
Interaural differences are significant in the near-field, and will mostly affect high frequencies. It would be interesting to study their impact on the assessment of sources characteristic in an immersive context.

5.5.2. Space properties

Various room impulse responses and reverberation time can be simulated. The reverberation time and the direct sound-early reflections ratio should then be taken into account in the analysis of the results.

5.5.3. Point sources arrangement

The influence of sound shape extent, types of patterns, but also the density and relative distance of point source on sound shape perception should be studied in priority.

5.6. Interpretation

Individual variations of the acoustic and space parameters may allow to look for psychophysical functions according to each parameter. Collective variations may be used to establish a weighting and hierarchy of these parameters. An accurate projection of sounds into space could then be achieved using inverse functions. Given the number of variables, the analysis of the results requires complex statistics calculation, which will probably limit the scope of the tests.

5.7. Assessment of source characteristics by the Listener

The sound shapes characteristics can be assessed individually, or globally.

5.7.1. Ranging, extent and depth

No visual cues should be provided to the listeners, as these dominate acoustic cues, affect auditory percepts, increase distance accuracy and improve judgement variability. Several strategies can be used to measure the listeners assessment. The precision of distance assessments by birds has been measured from the distance they had flown over to reach an imaginary rival. An equivalent procedure could be introduced, asking the listener to point at a close source or walk towards it. Accordingly to our review, other methods, such as free measurements using familiar units, yield satisfactory results.

5.7.2. Density

To our knowledge, the sense of spatial density hasn't been studied yet. An adequate semantic definition should be provided to the listeners prior to any tests, which would require a prior consensus between experts and non experts listeners. The listeners assessment could be measured by comparisons of sound patterns, or by rating on a scale.

5.7.3. Shape

Two types of drawings, placing points or outlines on the circular horizontal plane, and on the vertical plane, could allow the listeners to express a global assessment of the shapes extent in three dimensions. A free verbalisation of the impressions could provide a useful documentation that could contribute

Conclusion

A review of recent studies on distance and sound shape perception has underlined important differences in ranging in anechoic and reverberant environments, raising the importance of reverberation as a ranging cue, and the existence of a psychophysical function between perceived distance and physical distance. It was also shown that loudness, due to differences with ranging dynamics was relatively reliable depending on the types of signals or propagations conditions. More research should be made with respects to the role of binaural cues in near field ranging, as these might inform us on possible interactions between lateralisation and distance perception. Recent experiments have shown interesting possibilities of bi-dimensional sound shapes by the means of virtual acoustics and dynamic decorrelation. These techniques, completed by binaural techniques will be used in an spatial audio installation aiming at creating three-dimensional sounds shapes. The elements provided by our review constitute a framework for the definition of signals and space properties used in this installation. As this project takes on an ecological validity and an artistic dimension, the measurement of distance and sound shape assessment will be a complex task. Depth and density, two novel notions in that domain, will also have to be more thoroughly investigated, Strategic choices will have to be made with respects to the variables that will be studied and to the analysis of the listeners assessments.

References

[1] Rumsey, F., "Spatial Quality Evaluation for Reproduced Sound: Terminology, Meaning, and a Scene-Based Paradigm" J. Audio Eng. Soc., Vol. 50, No. 9, September 2002

[2]Naguib, M., Wiley, R.H., "Estimating the distance to a source of sound: mechanisms and adaptations for long-range communication", Animal Behaviour, No 62, pp. 825-837, 2001

[3] Bass, H.E., Sutherland, L.C., Zuckerwar, A.J., "Atmospheric absorption of sound : update", J. Acoust. Soc. Am.: Letters to the Editor, No 88(4), Oct 1990

[4] Mershon, D.H., King, L.E., "Intensity and reverberation as factors in the auditory perception of egocentric distance", Perception and Psychophysics, No18, pp. 409-415, 1975

[5] Mershon, D.H., Bowers, J.N., "Absolute and relative cues for the auditory perception of egocentric distance", Perception, No 8, pp. 311-322, 1979

[6] Little, A.D., Mershon, D.H., Cox, P.H., "Spectral content as a cue to perceived auditory distance". Perception, 21, pp. 405-416, 1992

[7] Zahorik, P., "Assessing auditory distance perception using virtual acoustics", J. Acoust. Soc. Am., Vol. 111, No.4, April 2002

[8] Zahorik, P., "Auditory Display of sound source distance", Proceeding of the 2002 International Conference on Auditory Display, Kyoto, Japan, July 2-5, 2002

[9] Coleman, P. D., "An analysis of cues to auditory depth perception in free space" Psychological Bulletin, No 60, 302-315

[10] Brungart D.S., Rabinowitz, W.M., "Auditory localisation of nearby sources. Head-related transfer functions" J. Acoust. Soc. Am., No 106, pp. 1465-79, 1999

[11] Duda, R.O., Martens, W.L., "Range dependence of the response of a spherical head model", J. Acoust. Soc. Am.: No 104, pp. 3048-3058, 1998

[12] Nielsen, S. H., "Auditory distance perception in different rooms", J. Audio Eng. Soc. Vol. 41, No 10, October 1993

[14] Lamancusa, J.S., *Outdoor sound propagation*, Lecture notes http://www.me.psu.edu/lamancusa/me458/10_osp.pdf

[15] Potard, G. and Burnett, I., "A study on sound source apparent shape and wideness", Proceedings of the 2003 International Conference on Auditory Display, 6-9 July 2003

[16] Potard and Burnett, G., Burnett, I., "Control and measurement of apparent sound source width and its applications to sonification and virtual auditory displays", Proceeding of ICAD 04 - Tenth Meeting of the International Conference on Auditory Display, 6-9 July 2004

[17] Kearney, G., Gorzel, M., Boland, F., Rice, H., "Depth perception in interactive virtual acoustic environments using higher order ambisonics soundfield", Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics, May 6-7, 2010

[18] Gardner, M. B., "Distance estimation of 0° or apparent 0° -oriented speech signals in anechoic space" J. Acoust. Soc. Am., **45**, 47-53

[19] Perrot, D., Musicant, A., Schwethelm, B., "The expanding image effect: The concept of tonal volume revisited", J. Auditory Research, vol. 20, pp 43-55, 1980

[20] Perrot, D., Buell, T., "Judgments of sound volume: Effects of signal duration, level, and interaural characteristics on the perceived extensity of broadband noise", J. Acoustical society of America, vol. 72(5), pp 1413-7, Dec. 1981

[21] Kendall, G., The decorrelation of audio signals and its impact on spatial imagery, Computer Music Journal Vol. 19, No. 4 (Winter, 1995), pp. 71-87