We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

15

Virtual Moving Sound Source Localization through Headphones

Larisa Dunai, Guillermo Peris-Fajarnés,
Teresa Magal-Royo, Beatriz Defez and Victor Santiago Praderas
Universitat Politécnica de València
Spain

1. Introduction

Humans are able to detect, identify and localize the sound source around them, to roughly estimate the direction and distance of the sound source, the static or moving sounds and the presence of an obstacle or a wall [Fay and Popper, 2005]. Sound source localization and the importance of acoustical cues, has been studied during many years [Brungart et al., 1999]. Lord Rayleigh in his “duplex theory” presented the foundations of the modern research on sound localization [Stutt, 1907], introducing the basic mechanisms of localization. Blauert defined the localization as “the law or rule by which the location of an auditory event (e.g., its direction and distance) is related to a specific attribute or attributes of a sound event” [Blauert, 1997].
A great contribution on sound localization plays the acoustical cues, Interaural Time Difference ITD and Interaural Level Diference ILD, torso and pinnae (Brungart et al., 1999), [Bruce, 1959]. [Kim et al., 2001] confirm that the Head Related Transfer Functions (HRTFs) which represent the transfer characteristics of the sound source in a free field to the listener external ear [Blauert, 1997]), are crucial for sound source localization.
An important role in the human life plays the moving sound localization [Al’tman et al., 2005]. In the case of a moving source, changes in the sound properties appear due to the influence of the sound source speed or due to the speed of the used program for sound emission.
Several research have been done on static sound localization using headphones [Wenzel et al., 1993], [Blauert, 1997] but few for moving sound source localization. It is well known that on localization via headphones, the sounds are localized inside the head [Junius et al., 2007], known as “lateralization”. Previous studies [Hartmann and Wittenberg, 1996] in their research on sound localization, showed that sound externalization via headphones can be achieved using individual HRTFs, which help listeners to localize the sound out in space [Kulkani et al., 1998], [Versenyi, 2007]. Great results have been achieved with the individual HRTFs, which are artificially generated and measured on a dummy head or taken from another listener. Due to those HRTFs, the convolved sounds are localized as real sounds [Kistler et al., 1996], [Wenzel, 1992].
This chapter presents several experiments on sound source localization. Two experiments are developed using monaural clicks in order to verify the influence of the Inter-click interval on sound localization accuracy.
In the first of these experiments [Dunai et al., 2009] the localization of the position of a single sound and a train of sounds was carried out for different inter-click intervals (ICIs). The

initial sound was a monaural delta sound of 5ms processed by HRTFs filter. The ICIs were varying from 10ms to 100ms. The listeners were asked to inform what they listened, the number and the provenience of the listened sound and also if there was any difference between them, evaluating the perceived position of the sound ("Left", "Right" or "Centre"). It was proven that the accurateness in the response improves with the increase of the length of ICI. Moreover, the train of clicks was localized better than the single click due to the longer time to listen and perceive the sound provenience.
In the second study (Dunai et al., 2009), the real object localization based on sensory system and acoustical signals was carried out via a cognitive aid system for blind people (CASBliP). In this research, the blind users were walking along a 14m labyrinth based on four pairs of soft columns should localize the columns and avoid them. The average time of sound externalization and object detection was 3,59 min. The device showed no definitive results due to the acoustical signal speed, which required improvements.

2. Experiment

2.1 Experiment 1. A pair of sounds and a train of sounds source localization

In the Experiment 1, the localization of the static sound source was studied; the saltation perception on the inter-click presence was also analyzed. The experiment is based on monaural click presented at different inter-click intervals (ICI), from 10ms to 100ms. Two types of sounds single click and train of clicks are generated and thereafter tested at different inter-click intervals. At short inter-click intervals, the clicks were perceived as a blur of clicks having a buzzy quality. Moreover, it was proven that the accurateness in the response improves with the increase of the length of ICI.
The present results imply the usefulness of the inter-click interval in estimating the perceptual accuracy. An important benefit of this task is that this enables a careful examination of the sound source perception threshold. This allows detecting, localizing and dividing with a high accuracy the sounds in the environment.

Sound sample

Sound source positions used for stimulus presentation in this experiment were generated for a horizontal frontal plane. A sound of 5ms duration was generated with Above Audition software.
In the first case, the generated sound with duration of 5ms was used as spatial sound and in the second case; the sound was multiplied by six, becoming a train of sound with duration of 30ms.
The sound has been convolved using Head Related Transfer Functions (HRTFs). It is known that the HRTFs are very important for sound localization, because they express the sound pressure at the listener eardrum over the whole frequency range. In the present study, the HRTFs were generated at 80dB at a frequency of 44100 Hz and processed by a computer for the frontal plane, for a distance of 2 m, with azimuth of 64º (32º at the left side of the user and 32º at the right side of the user).
In the experiments the sound were presented randomly in pairs Left-Right and Right-Left, delivered using Matlab version 7.0, on an Acer laptop computer.

Test participants

Ten volunteers, 4 females and 6 males, age range 27-40 years, average 33,5 participate in this experiment. Each subject reported to have normal hearing, they did not reported any hearing deficiencies. All of them were supposed to other acoustical experiments with computer and acoustical mobility devices.

Procedure

The experiment was carried out in a single session. The session consisted of two runs, one for a single sound and one for a train of sound. Each run was based on six sounds. Fig.1 shows the schematic presentation of the sound: a) shows the monaural sound in which, the click comes from (Left) L→R (Right) and R→L, with randomly varying ICIs; b) shows the train of sound, where the presentation procedure is the same as for the single sound, the sound come from L→R and R→L, with randomly varying ICIs. Different interclick intervals (ICI), from 10 ms to 100 ms were used (10ms, 12ms, 25ms, 50ms and 100ms).

Localization test were carried out in a chamber of 4,8m x 2,5m x 12m, where external sounds were present.

Since the experiments described in this chapter were focused on examining the perception in human listeners, it was important to be able to measure spatial capabilities in an accurate and objective way. For the localization test, subject localized auditory sound presented in the headphones, telling the direction of the listened sound. In both cases the experiment begins with various exercises where the subjects are able to hear the sound and train of sound, separately, firstly the left one and afterwards the right one, continuing with the six sounds delivered by the program randomly. Afterwards the subject completed the all six sounds, the new exercises were presented of the combination "Left-Right" and "Right-Left".

For the localization tests, listeners were sitting comfortably in a chair in front of a computer. Before starting the test, the listeners received written and oral instructions and explanations of the procedure. They were asked to pay especial attention and to be concentrated on the experiment.

Before localization experiments, subjects had a training protocol to become familiar with the localization. This protocol included the speech pointing techniques, which requires that the subject verbally informs the evaluator about the perceived localization of a sound. During the experiment, since the subject had not access to the computer screen, the tendency of capturing the sound with the eyes was eliminated.

During the test, the subjects were supposed to listen through the headphones, model HD 201, twelve pairs of sounds; six pairs of single sound and six pairs of trains of sound "Left-Right" and "Right-Left" at different ICIs, from 100 ms to 10 ms in a decreasing succession.

The sounds were delivered in a random position. The sound used in the experiment was the same sound used in the testing procedure. The sound duration was brief enough, so that listener could not make head movements during the sound presentation. Between each two consecutive pair of sound, the decision time (Td) was computed; this was the time needed for evaluating the sound (see Fig. 1).

The subjects were asked what they listened, the number and the provenience of the listened sound and also if there was any difference between them. The subjects where allowed to repeat them, if necessary, after they had evaluated the perceived position for each sound, classifying them as "Left", "Right" or possible "Centre". Once the subject had selected a response, a next pair of sound was presented. Each trial lasted approximately 2 min. The average time per subject for all experiment was around 35 min.

Some distraction cues as: environmental noises, draw away seeing or hearing someone- since the subject remained with opened eyes influenced on the experimental sound source perception and results. Because of this reason, the subjects were allowed to make judgments about the source location independently.

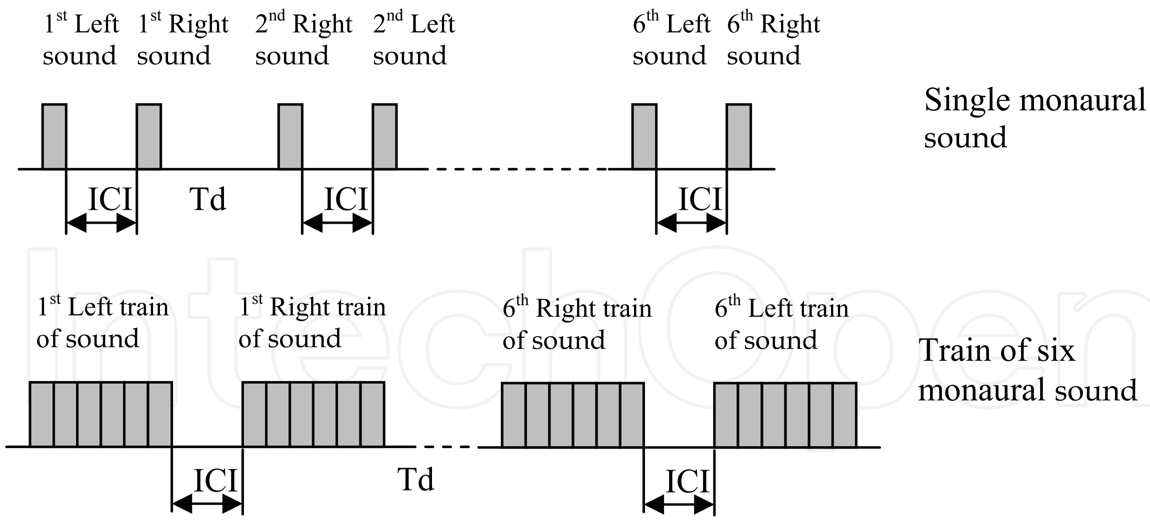


Fig. 1. Schematic presentation of the sound. In both situations the sound is of 5ms. In the first case, the sound has been listened at the different interclick intervals ICI separated by a decision time Td. In the second case, the sound has been substituted by a train of six sound.

The results were collected by the evaluator and introduced manually into a previously prepared table. After the test, localization performances were examined using the analyses described in the following section.

Results

The results from the Experiment 1 were collected for data analysis. Localization performances summary statistics for each subject are listed in Table 1. The graphical user interface was generated by Excel in linear standard model. Subject response was plotted in relation to the Inter-click Interval. The main data for all subjects is presented in Fig. 2 with an error of 5%.

The perception of the single and train of sound and the perceived position of the sound pairs “Left-Right” and “Right-Left” were analyzed. Both factors as well as the interaction with the ICIs were significant.

Fig. 2 shows that the perception of the sound source position decreases when ICIs does. For avoiding errors, the tests results were registered up to an ICI of 10ms. Because ICI was enough short, the sound were perceived as a single entity moving from one ear to another or from one ear to the centre having a buzzing quality.

In the case of the single pair of sound at ICI of 12ms, because the length of the sound and the length of the ICI were too short, the subjects could not distinguish clearly the sound corresponding to the pairs “Left-Right” and “Right-Left”.

When comparing the perception of the single sound with the perception of the train of sound Fig. 2 a), a great continuity of the sound position across almost the entire range of ICIs was detected. In other words, the perception of the sound position was stronger for the train of sound. This effect may be a result of the better localization associated with the sound.

$$\sqrt{\frac{\sum (x-\bar{x})^2}{(n-1)}} \tag{1}$$

For ICIs between 25 and 10ms, the subjects perceive the “Right-Left” pair of sounds with a higher precision than that of pairs “Left-Right” for single sound and train of sound.

In other case, for ICIs of 50ms, the perception of the pair of single sound "Right-Left" is higher than the perception of the pair Left-Right. In the case of the train of sound, the perception results are equivalent for both pairs Left-Right and Right-Left.
When trying to explain the sound source perception threshold, we perceive the perception of the saltation illusion. With shorter ICIs, a blur of sound were perceived, in contrast with the individual sound at longer ICIs. As the psychologist Gestalt noted, the perceptual system scrambles for the simplest interpretation of the complex stimuli presented in the real world. Therefore, the studies were based on analyzing and proving that, grouping the sound, the sound source is better perceived and localized.
For longer ICIs, this procedure is not so important, since each sound can be identified and localized. The present results demonstrate the usefulness of the inter-click interval in estimating the perceptual accuracy. A possible benefit of this task is enabling a careful examination of the sound source perception threshold. This allows detecting, localizing and dividing with high accuracy the sounds in the environment.

Sound perception in %			Train of sound perception in %		
interclick ms	Azimuth -30º	azimuth 30º	interclick ms	Azimuth -30º	azimuth 30º
100	100%	100%	100	100%	100%
50	90%	86%	50	100%	100%
25	80%	90%	25	88%	96%
12	83%	95%	12	76%	79%
10	88%	86%	10	75%	86%
8	100%	95%	8	100%	96%
6	100%	95%	6	85%	93%
5	100%	92%	5	100%	95%
1	100%	100%	1	100%	100%

Table 1. Localization performance summary statistics for all subjects (P1-P9) in frontal field. The percentage of the perception experiment is calculated on the basis of the six delivered sounds.

2.1 Experiment 2. The influence of the inter-click interval on moving sound source localization tests

In the Experiment 2, an analysis of moving sound source localization via headphones is presented. Also, the influence of the inter-click interval on this localization is studied. The experimental sound consisted of a short delta sound of 5ms, generated for the horizontal frontal plane, for distances from 0,5m to 5m and azimuth of 32º to both left and right sides, relative to the middle line of the listener head, which were convolved with individual HRTFs. The results indicate that the best accurate localization was achieved for the ICI of 150ms. Comparing the localization accuracy in distance and azimuth, it is deduced that the best results have been achieved for azimuth. The results show that the listeners are able to extract accurately the distance and direction of the moving sound for higher inter-click intervals.

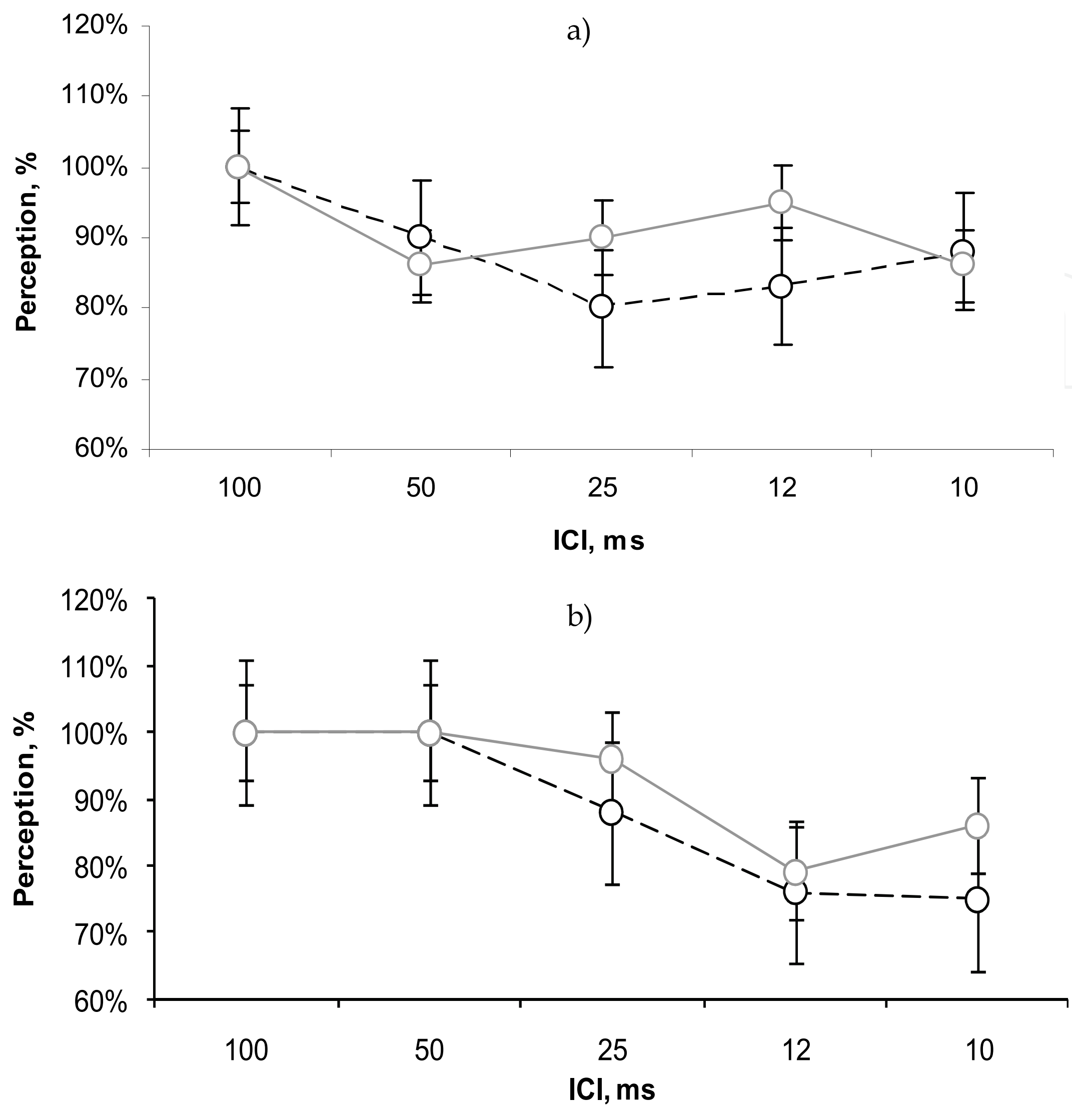


Fig. 2. Mean estimation of the click location: a) shows the sound perception at -30º (left side) and +30º (right side); b) corresponds to the train of sound perception at -30º (left side) and +30º (right side)

Subjects

Nine young subjects students with ages between 25 and 30 years and different gender, all of them had normal vision and hearing abilities, were involved in the experiments. All participants had normal distance estimation and good hearing abilities. They demonstrate a correct perception of the sounds via headphones. A number P1-P9 identified the subjects.
All subjects participated in previous auditory experiments in the laboratory. Each participant received a description of what was expected of him/her and about all procedure. All participants passed the localization training and tests described below.

Stimuli and signal processing

A delta sound (click) of 2048 samples and sampling rate of 44.100 Hz was used. To obtain the spatial sounds, the delta sound was convolved with Head-Related Transfer Function (HRTF) filter measured for each 1º in azimuth (for 32º left and 32º right side of the user) at

each 1cm in distance. The distance range for the acoustical module covers from 0,5m to 5m, an azimuth of 64º, and 64 sounding pixels per image at 2 frames per second.
Recording of Head-Relates Transfer Functions were carried out in an anechoic chamber. The HRTFs measurements system consist on a robotic and acquisition system. The robotic system consists of an automated robotic arm, which includes a loudspeaker, and a rotating chair on an anechoic chamber. A manikin was seated in the chair with a pair of miniature microphones in the ears. In order to measure the transfer function from loudspeaker-microphone as well as for headphone-microphone, the impulse response using Maximum Length Binary Sequence (MLBS) was used. The impulse response was obtained by taking the measured system output circular cross-correlation with the MLBS sequence.
Due to that the HRTF must be measured from the two ears, there is necessary to define the two inputs and output signals. Lets x_1 (n) be the digital register of the sound that must be reproduced by the speakerphone. Lets $y_1(n)$ be the final register recorded by the microphone placed in one of the acoustic channels of the manikin or man, corresponding to the response to $x_1(n)$. Similarly, let $x_2(n)$ be the sound to be reproduced through the headphone and $y_2(n)$ the answer registered by the headphone, respectively for the second ear. The location of the head in the room is assumed to be fixed and is not explicitly included in our explication.
In order to determine $x_1(n)$, it is necessary to generate a $x_2(n)$ such that the $y_2(n)$ is identical to $y_1(n)$. In that way, we achieve that an acoustic stimulus generated from the speakerphone and another generated by the headphones, produce the same results in the auditive channel of the user or manikin. Therefore we obtain the same acoustical and spatial impression.
In order to obtain these stimuli, a digital filter which transforms the $x_1(n)$ into $x_2(n)$ has been developed. In the transformed frequency domain, let be X_1 the representation of the $x_1(n)$ and Y_1 the representation of the y_2 (n).
Then Y_1, which is the registered response of the $x_1(n)$ reproduction, is:

$$Y_1 = X_1 LFM \tag{1}$$

In (1), L represents the grouped transfer function of the speakerphone and all audio reproduction system. F represents the transfer function of the environment situated between the speakerphone and the additive channel (HRTF) and M represents the set of functions composed by the microphone and the whole audio reproduction system.
The response registered by the microphone via headphones, when the $x_2(n)$ is reproduced, can be expressed as follows:

$$Y_2 = X_2 HM \tag{2}$$

where H represents the transfer function of the headphone and all reproduction system to the additive channel.
If Y_1=Y_2, isolating X_2 we obtain:

$$X_2 = \frac{X_1 LF}{H} \tag{3}$$

Then, for any measurement the digital filter will be defined as follows:

$$T = \frac{LF}{H} \tag{4}$$

Therefore, it will filter the signal $x_1(n)$ and the resulting signal $x_2(n)$ will be reproduced by the headphone; then the signal registered by the microphone, which is placed in the auditive channel must be $y_1(n)$. This signal must be equal to the signal $x_1(n)$, which is reproduced by the speakerphone.
The filter described by (4) describes the speakerphone for a single spatial position for only one ear. For both ears two filters are required for the simulation of each signal source for a determined spatial position.
Assuming that we measure the Y_1 and X_1 transfer functions for different spatial positions for both ears at the same time, the Transfer Function speakerphone-microphone (G_{LM}) is defined as follows:

$$G_{LM} = \frac{Y_1}{X_1} = L \cdot F \cdot M \tag{5}$$

Having the function given by (5) simultaneously for both ears, we measure both transfer functions Y_2 and X_2, on which the transfer functions headphone-microphone G_{HM}, are defined:

$$G_{HM} = \frac{Y_2}{X_2} = H \cdot M \tag{6}$$

The necessary filters for the sound simulation are obtained from the function speakerphone-microphone G_{LM} for each ear, as the reverse of the function headphone-microphone G_{HM} of the same ear (see (4)). So, for both ears:

$$T = \frac{G_{LM}}{G_{HM}} = \frac{L \cdot F \cdot M}{H \cdot M} = \frac{L \cdot F}{H} \tag{7}$$

For both transfer function speakerphone-microphone G_{LM} and headphone-microphone G_{HM}, the measurement technique of the impulse response Maximum Length Binary Responses MLBS was applied with later crossed correlation between the system answer and input of the MLBS.
The impulse response of the system can be obtained through circular crossed correlation between input MLBS of the system and the output answer. This is, if we apply to the system an MLBS, which will called s(n), and measure the output the signal y(n) during the time which MLBS lasts, the impulse response h(n) will be defined as follows:

$$h(n) = \Omega_{sy}(n) = s(n)\,\Phi\, y(n) = \frac{1}{L+1}\sum_{k=0}^{L-1} s(k) \cdot y(n+k) \tag{8}$$

where Φ represents the circular or periodic crossed correlation operation, corrupted by the aliasing time, and not a pure impulse response.
In the event that the sequence is enough long, then the resultant aliasing can be rejected. Due to that, the direct implementation of (8) for long sound sequences require high computational time, the equivalent between the correlation and periodic crossed correlation has been used. The obtained information was passed into the frequency domain, where the convolution operation is translated into a vector multiplication.
After this, the results were passed into the frequency domain, where the convolution operation is translated into a vector multiplication.

$$a(n)\Phi b(n) = \frac{1}{L+1} a(-n) * b(n) \tag{9}$$

where the inversion of the first sequence is circular, similar to the convolution. Nevertheless, the computational time results to be enough high, due to that the used Fast Fourier Transform (FFT) have a length of 2^k-1. In order to obtain an increasing performance in time processing the FFT length has to be $(2^k-1)^2$.
Finally, using the Fast Hadamard Transform (FHF), it was possible to reduce the computational time between the two magnitudes. The h(n) is then calculated as follows:

$$h(n) = \frac{1}{(L+1)s[0]} P_2 \left\langle S_2 \left\{ H_{L+1} \left[S_1 \left(P_1 y(n) \right) \right] \right\} \right\rangle \tag{10}$$

In this case to the system has been applied a MLBS s(n) with a length L, after what the result y(n) was registered. The matrix P is the permutation matrix, the matrix S is matrix of rescaling, the H_{L+1} is the matrix Hadamard of degree L+1. After the HRTFs were measured, with the equipment shown in figure 3.13, it was verified if the HRTFs are realistic and externalized. For this purpose, an off-line localization procedure was carried out.
The output signals (the HRTF) are sampled at 22050Hz and a length of 46ms (8192 bit).
The HRTFs were measured for the horizontal frontal plane at the ear level from 0,5 to 5m in distance and in azimuth between 32º left and 32º right with respect to the centre of the listener head (measurements at every 1º). Fig. 4 shows the graphical representation of the sound reproduction.

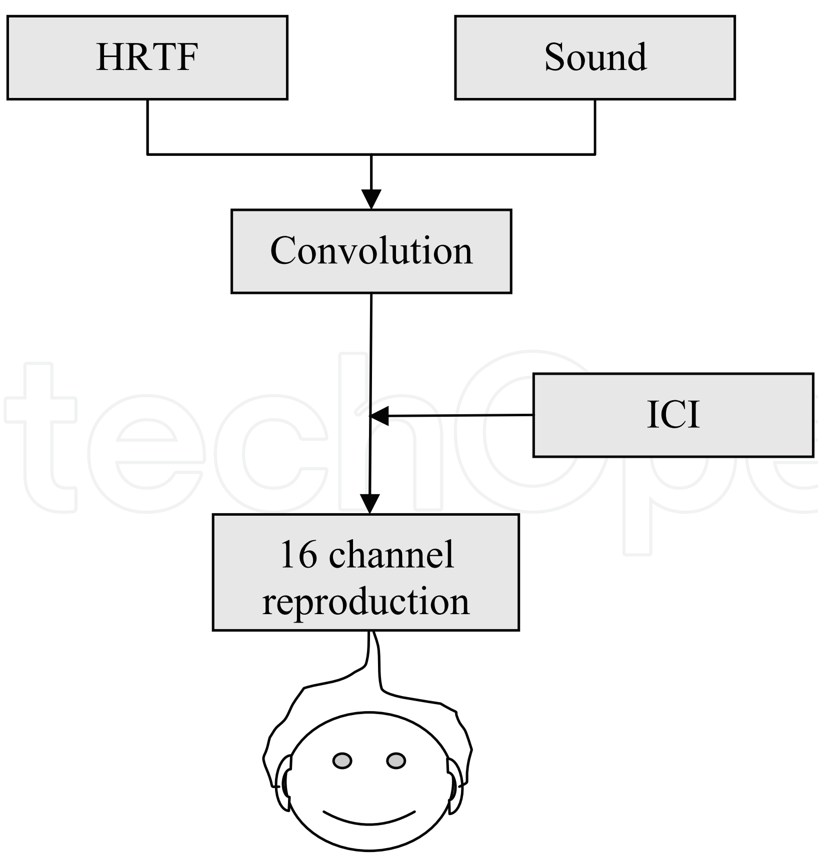


Fig. 4. Method for sound processing and reproduction

Equipment

A Huron system with 80 analogue outputs, eight analogue inputs and eight DSPs 56002, and a computer for off-line sound processing was used for the sound generation and processing. SENNHEISER headphones models HD 201 were used to deliver the acoustical information. MATLAB 7.0 was used as experimental software. The resultant graphical sound trajectory for each experiment was displayed on a separate window and saved for off-line processing. All experiments run on ACER Aspire 5610 computer.

Procedure

The goal of the experiments is to analyze the localization of a moving sound source via headphones and to see how the inter-click interval (ICI) influences the sound localization quality. The comparison between the localization performances enables to evaluate the importance of the inter-click interval parameter for its use in sound localization and acoustical navigation systems.

The movement of the sound source was achieved by switching the convolved sound for a frontal plane at the eyes level at increasing distances from 0,5 to 5m (1 cm increase) and for azimuth between 32º right and 32º left (1º increase) with respect the middle of the head. The sounds were delivered for five inter-click intervals [200ms, 150ms, 100ms, 75ms and 50ms]. Fig. 5 shows one of the trajectories the sound was running. Four different trajectories were created. The delivered trajectory was selected randomly by the computer when the experiment starts.

Before starting the experiment, the training exercises were carried out; the objective and the procedure of the experiment were explained to each individual participant. One sound was delivered for all five ICIs, where the participants were able to see graphically the listened sound trajectory (See Fig. 5). In order to proceed with the test and experiment, the

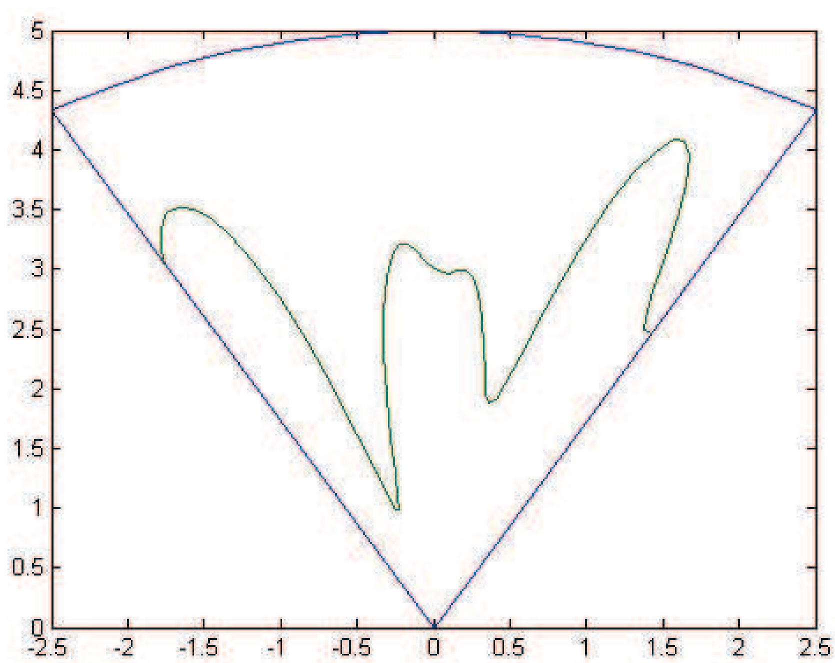


Fig. 5. Sound trajectory example, direction from left to right. The x axis represents the azimuth where the 0 is the centre of the head, which is 0º. The -2.5 is the -32º at left side of the head and 2.5 respectively is 32º at the right side of the head. The y axis represents the distance from 0 to 5m

participants were asked to seat comfortably in the chair in front of a computer. After reading and testing the training exercises, the participants were supposed to carry out the experiment.
A sound at a specific ICI was delivered by the computer via headphones. During the experiment, the participants were free to move. Nevertheless, they were required to move the less possible and to be concentrated on the sound, in order to create a plane of the sound route in the imagination. The test was performed both with open eyes and with closed eyes depending on the participant wishes. In the case of the closed eyes, there was a limitation of effects of the visual inputs. Due to this, the participant achieved a better interpretation of the trajectory image.
The participants were asked to carefully hear the sound and draw the listened trajectory in a paper. They were allowed to repeat the sound if it was necessary. All the participants asked to repeat the sound at least three times. Each participant was supposed to have five trials, one for each ICI. Only one sound trajectory was used per participant for all five ICIs. For all participants, the experiment started with the ICI of 200ms, decreasing it progressively up to 50ms.
After the experiment the participants commented the perceived sound trajectory and they compared the listened sound for each ICI.

Results

The moving sound source localization is an important factor for the navigation task improvement. The main variables analyzed in this paper were the moving sound source localization and the inter-click interval ICI [200, 150, 100, 75, and 50ms]. The study analyzes the interaction between these variables in measurements of distance and azimuth.
Generally, no significant differences on the results were registered between participants. However, great difference was found in the sound localization between higher and lowers inter-click intervals.
The maximum displacement in distance is 1,26m for an ICI of 50ms and the minimum displacement was 0,42m for an ICI of 150ms, the maximum displacement in azimuth was 11,4^o for an ICI of 50ms and the minimum 0,71^o for an ICI of 150ms.
Average results of sound localization in azimuth and distance as a function of the inter-click interval are shown in Fig. 6. Best results have been achieved for greater ICIs, due to the time needed by the brain to perceive and process the received information. Because the time between two sounds is higher, the sound is perceived as jumping from one position to another from left to right in equal steps. For the ICI of 200ms, the sound was not perceived as a moving sound, but rather as a jumping sound from location to location. However, for the ICIs lower than 100ms the sound was perceived as a moving sound from the left to right, but there was enough difference between the original sound trajectory and the perceived one. The participants had great difficulties to perceive the exact distance and azimuth, because the sound was delivered too fast. Moreover, when the sound trajectory had multiple turning points on a small portion of the space, the participants perceived this portion as one turn-return way. Fig. 7 represents a specific case, corresponding to one of the participants; it shows the moving sound localization at four ICIs. The red colour represents the listened sound trajectory drawn by the participant. The grey colour represents the real sound trajectory drawn by the computer. The x axis represents the azimuth where the 0 value is the centre of the head, the negative values are the values at the left side of the head, whereas the values at the right side of 0 represent the azimuth values at the right side of the human head. The -2.5 represents the 32^o at left side of the head and 2.5 the 32^o at the right side of the head. The y axis represents the distance from 0 to 5m.

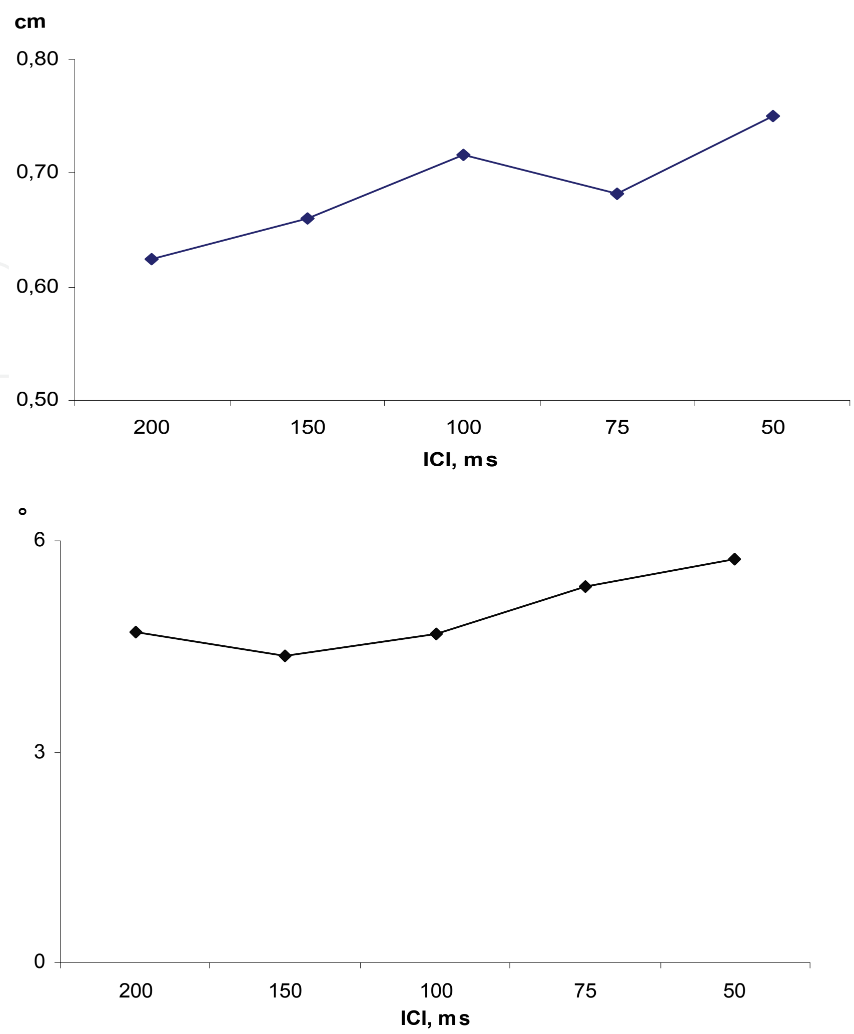


Fig. 6. Average displacements in azimuth and distance for all participants

In some cases, the participants perceived the sound trajectory as an approximate straight line when the inter-click interval was 50ms. Even repeating several times the experiment, the participants were confused regarding the localization of the moving sound. They commented "the sound moves too fast and I feel that it is running from left to right in a straight line". Despite listeners were not able to localize the moving sound source at lower inter-click intervals so well as they were able to localize the moving sound for greater inter-click intervals, they were able to judge about the sound position in azimuth and distance.

Various factors as drawing abilities (how the participants can accurately draw), sound interpretation (how the participants can interpret the heard sounds, by colours, by image etc.), the used hearing methods (with closed or opened eyes), the external noises, etc., influenced the experiment results. Despite all participants were informed about the use of one sound per participant for all ICIs, they draw the trajectories at different distances. This error appears because of the participant drawing ability; it is not so easy to interpret graphically what is listened or the image the brain creates if there is not practice on that. For some of participants, great concentration and relaxation was required, to be able to correctly perceive the sounds.

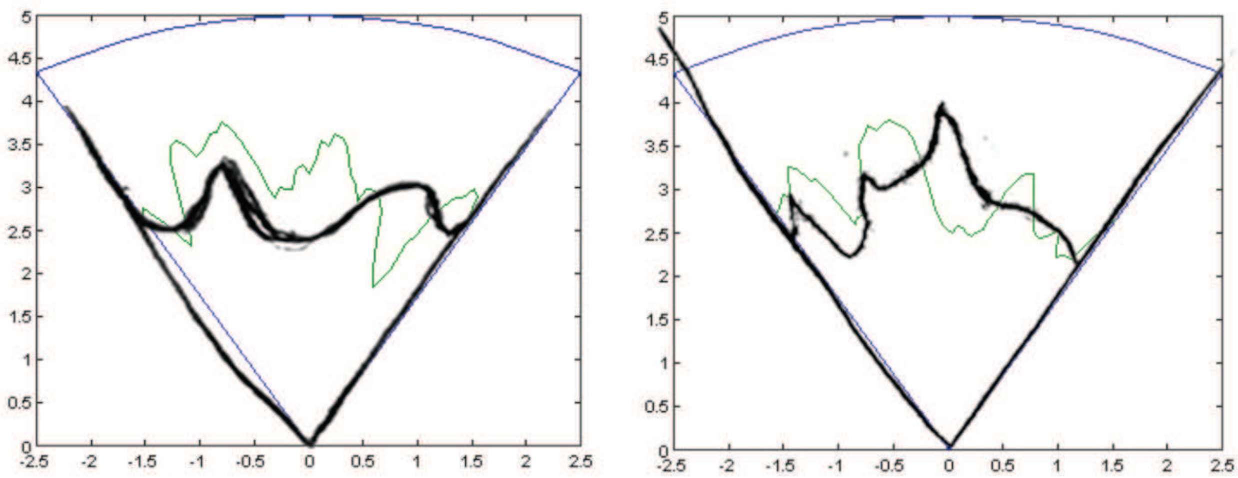


Fig. 7. Sound trajectory for one participant for the ICIs of 50ms and 100ms. The black colour represents the heard sound trajectory drawn by the participant; the green colour represents the real sound trajectory drawn by the computer. The x axes represent the azimuth, in which the 0 value is the centre of the head, the negative value are the values at the left side of the head and the values at the right side of 0 represent the azimuth values at the right side of the human head. -2.5 represents the 32° at left side of the head and 2.5 respectively the 32° at the right side of the head. The y axis represents the distance from 0 to 5m.

Multiple observations on training sound trajectory were given to participants about how to perceive the sound and to be confident of their answer. Two participants were excluded from the main analysis due to the difficulties in localizing the sound. The participants experienced the moving sound localization as a straight line for all inter-click intervals.

3. Conclusion

In the present chapter two sets of experiments are described according to the examined spatial performance involving simple broad-band stimuli. Both experiments measured how well single and train of static and moving sounds are localized in laboratory conditions. These experiments demonstrated that sound source is essential for accurate three-dimensional localization. The approach was to present sounds overlapped in time in order to observe the performance in localization, in order to see how time delay between two sounds (ICI inter-click interval) influences on sound source localization. From the first experiment it was found that better localization performance was achieved for trains of sounds at an ICI of 100ms. If analyzing the localization results at the left and right side of the human head, it must mention that improved results were obtained at the left side for the single click and at the right side for the train of clicks. At short inter-click intervals, the train of clicks was perceived as a blur of clicks. At short inter-click intervals the single clicks was perceived as one click, there were not perceived the difference between the first click and the second one. In this case only the first click was perceived, the second click was perceived as a week eco. Moreover, the sound perception threshold was studied. In the second study the localization of a moving sound source both in distance and azimuth was analyzed. The results demonstrate that the best results were achieved for an inter-click interval ICI of 150ms. When comparing the localization accuracy in distance and azimuth, better results were obtained in azimuth. The maximum error in azimuth is of 11,4° at the ICI of 50ms. The disadvantages of the results at short ICI's are due to that the total time of the sound run is very short, that prevent the user to perceive all the sound coordinates. Regarding the large ICI's, the saltation from one click to another don not allows the user to make the connection between the two clicks. From this motive the user perceive the sounds as diffuse. Spatial cues such as interaural time difference ITD and interaural level difference ILD play an important role in spatial localization due to their attribution on the azimuthal sound localization. They arise due to the separation of the two ears, and provide information about the lateral position of the sound.

4. References

Al`tman Ya.A.; Gurfinkel V.S.; Varyagina O.V.; Levik Yu.S. (2005). The effect of moving sound images on postural responses and the head rotation illusion in humans, Neuroscience and Behavioral Physiology, 35 (1), 103-106

Blauert J. (1997). Spatial Hearing: The Psychophysics of Human Sound Localization, revised edn, The MI Press, Cambridge, MA, USA,

Bruce H., Hirsh D. and I.J., (1959). Auditory Localization of Clicks J. Acoust. Soc.Am., 31(4), 486-492

Brungart D.S Nathaniel I., W. R. Raibiowitz. (1999). Auditory localization of nearby sources II. Localization of a broadband source, J. Acoust. Soc.Am. 106 (4), 1956-1968

[1]Brungart D.S., Rabinowitz W. M. (1999). Auditory localization of nearby sources. Head-related transfer functions, J. Acoust. Soc.Am. 106(3), 1465-1479

Dunai L., Peris F G.., Defez B. G., Ortigosa A.N., Brusola S F. (2009). Perception of the sound source position, Applied Physics Journal, (3), 448-451

[1]Dunai L., Peris F G.., Defez B. G., Ortigosa A.N., (2009). Acoustical Navigation Sysyem for Visual Impaired People, LivingAll European Conference

Dunai L., Peris Fajarnes G., Defez Garcia B., Santiago Praderas V., Dunai I., (2010), The influence of the Inter-Click Interval on moving sound source localization for navigation systems, Applied Physics Journal, (3), 370-375

Hartmann W.M., Wittenberg A., (1996). On the externalization of sound images, J. Acoust. Soc.Am. 99 (6): 3678-3688

Junius D., Riedel H., Kollmeier B., (2007). The influence of externalization and spatial cues on the generation of auditory brainstem responses and middle latency responses, Hearing Research 225, 91-104

Kim H.Y., Suzuki Y, Sh. Takane, Sone T. (2001). Control of auditory distance based on the auditory parallax model, Applied Acoustics 62, 245-270

Kistler D.J., Wightman F.L., (1996). A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction a, b), J. Acoust. Soc.Am. 91 (3), 1637-1647

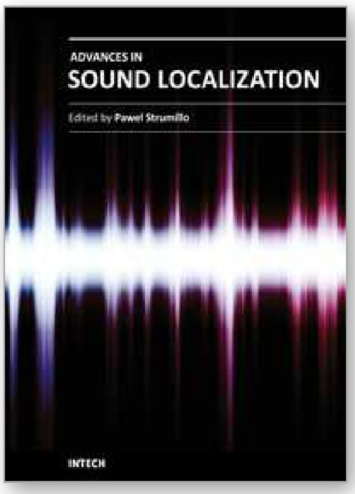
Kulkani A., Colburn S.H., (1998). Role of spectral detail in sound-source localization, Nature, 396, 747-749

Strutt J.W. (1907). On our perception of sound direction, Philos. Mag.; Vol 13, 214.232

Versenyi G. (2007). Localization in a Head-Related Transfer Function-based virtual audio sysnthesis using additional high-pass and low-pass filtering of sound sources, Acoust. Science & Technology, 28 (4), 244-250

Wenzel E., Arruda M., Kistler D., Foster S. (1993). Localization using non-individualized head-related transfer functions, J. Acoust. Soc.Am. 94, 111-123

Wenzel E.M., (1992). Localization in virtual acoustic display, Presence Telop Virt. Environ. 1, 80-107



Advances in Sound Localization

Edited by Dr. Pawel Strumillo

ISBN 978-953-307-224-1
Hard cover, 590 pages
Publisher InTech
Published online 11, April, 2011
Published in print edition April, 2011

Sound source localization is an important research field that has attracted researchers' efforts from many technical and biomedical sciences. Sound source localization (SSL) is defined as the determination of the direction from a receiver, but also includes the distance from it. Because of the wave nature of sound propagation, phenomena such as refraction, diffraction, diffusion, reflection, reverberation and interference occur. The wide spectrum of sound frequencies that range from infrasounds through acoustic sounds to ultrasounds, also introduces difficulties, as different spectrum components have different penetration properties through the medium. Consequently, SSL is a complex computation problem and development of robust sound localization techniques calls for different approaches, including multisensor schemes, null-steering beamforming and time-difference arrival techniques. The book offers a rich source of valuable material on advances on SSL techniques and their applications that should appeal to researches representing diverse engineering and scientific disciplines.