

Publication IV

Localization of a moving virtual sound source in a virtual room, the effect of a distracting auditory stimulus

M. Gröhn

Reprinted from *Proceedings of the International Conference on Auditory
Display (ICAD 2002)*, pages 394-402, Kyoto, Japan, 2.-5. Jul. 2002
©2002, author.



LOCALIZATION OF A MOVING VIRTUAL SOUND SOURCE IN A VIRTUAL ROOM, THE EFFECT OF A DISTRACTING AUDITORY STIMULUS

Matti Gröhn

Telecommunications Software and Multimedia Laboratory
Helsinki University of Technology,
P.O.Box 5400, FIN-02015, HUT, Finland
matti.grohn@hut.fi

ABSTRACT

An audio localization test of moving virtual sound sources was carried out in a spatially immersive virtual environment, using loudspeaker array with vector based amplitude panning for reproduction of sound sources. Azimuth and elevation error in localization was measured. In this experiment the main emphasis was to explore the effect of a distracting auditory stimulus.

Eight subjects accomplished a set of localization tasks. In these tasks they perceived the azimuth more accurately than the elevation. The distracting auditory stimulus decreased the localization accuracy. There was large variation between the subjects. The median error in azimuth for the most inaccurate subject was approximately twice as much as for the most accurate subject. The amount of the localization blur was dependent on angular distance from virtual sound source position to the nearest loudspeaker. The localization blur increased while the angular distance increased. Results of this experiment were compared with the results achieved in our previous experiment without the distracting stimulus.

1. INTRODUCTION

Immersive virtual environments provide an integrated system of three dimensional (3D) auditory and visual display. Generally, models explored in virtual environments have parts of specific interest. So far the most common method to emphasize them has been to highlight them visually. The interesting parts of the models can also be accentuated by using auditory beacons [1] or other auditory stimuli. In a dynamic representation it is important that the user is able to follow the location of the moving sound source.

The purpose of this research is to find out, how much the additional stimuli will affect on the localization accuracy of a moving virtual sound source in a virtual room. We have previously accomplished an experiment on localization of a single moving virtual sound source [2]. In this article the results of a multi stimuli experiment are described and compared with localization results achieved in our previous experiment. Both of these experiments were accomplished without visual stimulus.

Auditory localization of static 3D sound sources has been tested in several experiments previously. Most of these tests have used headphone reproduction [3, 4, 5, 6]. As a part of his localization experiment Sandvad [7] has measured the localization accuracy in a direct loudspeaker reproduction. The localization accuracy in panned (amplitude interpolated) loudspeaker reproduction has been reported for example by Pulkki [8, 9].

1.1. Virtual room

Localization experiments were accomplished in the virtual room¹ of the Helsinki University of Technology [10] (Figure 1). Typically there are simultaneously multiple users in a virtual room. In multiuser situation the loudspeakers are a more convenient sound reproduction method than headphones.

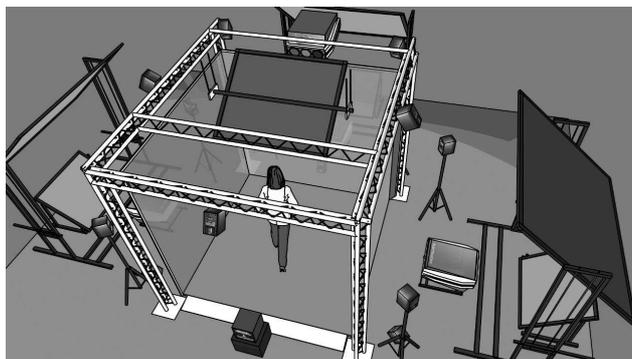


Figure 1: The schematic drawing (courtesy of Seppo Äyräväinen) of the virtual room of the Helsinki University of Technology. Localization experiments were accomplished without visual stimulus.

For the multichannel sound reproduction we use vector based amplitude panning (VBAP) [11]. It is less sensitive to listening position than Ambisonics [12], which is a benefit in a multiuser situation. In addition, to get an optimal result with Ambisonics the loudspeakers should be in a symmetric layout, which is hard to achieve in a virtual room, because the visual display system limits the possible loudspeaker locations. Due to the visual display, it is practically impossible to implement a wave field synthesis (WFS) [13] speaker array in a virtual room.

In a virtual room the screen between loudspeaker and listener has an effect on perceived signal. In our virtual room it has been measured that high frequencies of the direct sound are attenuated more than 10 dB. Impulse responses from each loudspeaker to the listening position were measured. Compensation filters were fitted to provide a sufficiently uniform timbre across the whole listening area. 10th order IIR filters are used for practical implementation of spectral compensation. More about the compensation of screen

¹<http://eve.hut.fi>

Loudspeaker	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Azimuth	-90	-30	30	90	180	-120	-68	-24	24	68	120	-110	0	110
Elevation	36	36	36	31	30	0	0	0	0	0	0	-34	-39	-34

Table 1: Azimuth and elevation angles of loudspeakers presented from the listening position.

damping and other implementation details of our audio environment are covered in another article [14]. Currently we are using 14 Genelec 1029A loudspeakers for sound reproduction. Their setup is presented in Table 1.

1.2. Direction indication methods

In localization experiments it is important to use a direction indication method to provide information about the perceived location. Researchers have used several different direction indication methods like graphical response screen [15], adjusting a reference sound [16, 9], pointing on a schematic drawing of the loudspeaker setup [17], using a head mounted laser pointer for pointing [18] or pointing with tracked toy gun [7].

Djelani et al [19] have compared three different direction indication methods: the Bochum-Sphere technique (also known as GELP) finger pointing and head pointing. In Bochum-Sphere technique the position of the auditory event is indicated on a sphere representing auditory space. According to their results finger pointing and head pointing were superior to the Bochum-Sphere technique. In the experiment described in this article, subjects pointed the perceived location of the sound source with tracked baton (figure 2).

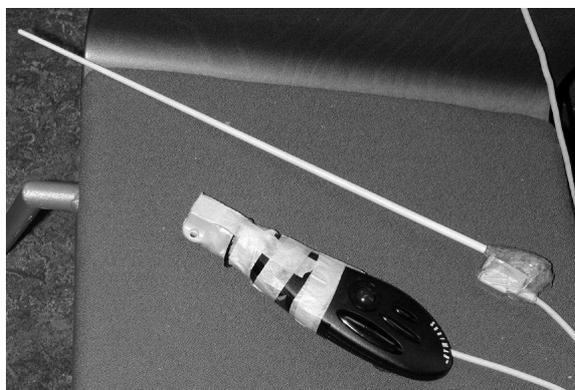


Figure 2: The devices used in our experiments are our wandlike device, and a tracked baton.

2. METHOD

The task of the subjects was to point to the direction of the perceived location of the target sound source. The azimuth and elevation values for perceived location were recorded as well as the azimuth and elevation values for sound source location. During the experiment subjects did not get any feedback about their localizing accuracy.

In our environment we have concentrated on loudspeaker reproduction. Our previous experiment [2] indicated, that in our virtual room a moving virtual sound source is localized as accurately as a static amplitude panned virtual sound source (not located in loudspeaker positions).

For this article an experiment with moving sources with additional distracting sounds was accomplished .

2.1. Subjects

In this experiment there were eight non-paid volunteers. Each of them reported to have normal hearing, although this was not verified with audiometric tests. There were six male subjects and two female subjects. Six of the subjects were the same as in our previous experiment.

2.2. Stimuli

To utilize both main binaural cues (ITD and ILD), the sound signal should have enough energy at low frequencies (below 1.5 kHz) and at high frequencies (above 1.5 kHz). There are also other factors in stimulus affecting the localization accuracy like temporal structure (see for example [20, 21, 22]). It has been found [23] that frequencies near 6 kHz are important for elevation perception.

In the experiment there were three different stimuli: pink noise (one minute long sample), music (2 minutes 45 seconds long excerpt from The Wall by Pink Floyd) and frog croak (0.5 second long). These signals have different kind of spectral content. Pink noise covers the whole audible frequency range, but temporal information is missing. Music is a broadband signal that also has a clear temporal structure. The croak sound has most of its energy below 2 kHz. In the experiments the stimuli were played continuously in a loop.



Figure 3: One of the subjects carrying out the experiment.

2.3. Procedure

During the experiment the subjects could freely move and turn they head. Each test task started with target sound in a static position. The subjects could freely locate the starting point of the target sound. The subjects pointed the perceived location of the starting point with a tracked baton (see Figure 3) and indicated that by clicking a wand button (Figure 2). After the subjects clicked the button of the wand, they heard a signal which indicated that sound source started to move. Simultaneously the additional distracting sound was started. The task of the subjects was to follow the movement of the target sound by pointing it with the baton. The end of the task was indicated with an end signal and silence. Then there was a short pause before the next task

In this experiment the distracting stimulus was always the same as the target sound but with different timing and gain. The gain of the distracting sound was 10 dB less than the gain of the target sound.

In this experiment there were three different target trajectories, one static distracting sound position and two different distracting sound trajectories and three different stimuli. For each subject each possible combination was presented once, which equals 27 tasks per subject. The tasks were presented in randomized order to avoid learning.

Each trajectory was 18 seconds long and the indicated sound source position was measured at a 20 Hz sampling rate. Azimuth and elevation values of the perceived location were recorded. In addition, virtual sound source azimuth and elevation values, the time from start, stimulus index, trajectory index, and distracting sound source azimuth and elevation values were recorded. The recording started while the user clicked the button of the wand, and ended when sounds were muted.

Most of the subjects reported that they pointed at the end position after the sound has already been muted. Unfortunately this information was not recorded. In the future, recording should be continued few seconds after the muting.

	Azimuth	Elevation
Pink noise	6.6	17.5
Pink Floyd	8.3	12.1
Frog croaks	7.8	13.5
All signals	7.9	15.0

Table 2: Median values of absolute azimuth and elevation errors for starting points (the target sound was not moving and there was no distracting stimulus).

	Azimuth	Elevation
Pink noise	17.4	28.3
Pink Floyd	18.5	24.9
Frog croaks	15.3	22.8
All signals	17.0	25.4

Table 3: Median values of absolute azimuth and elevation errors for dynamic trajectories for signals (moving sound with distracting stimulus).

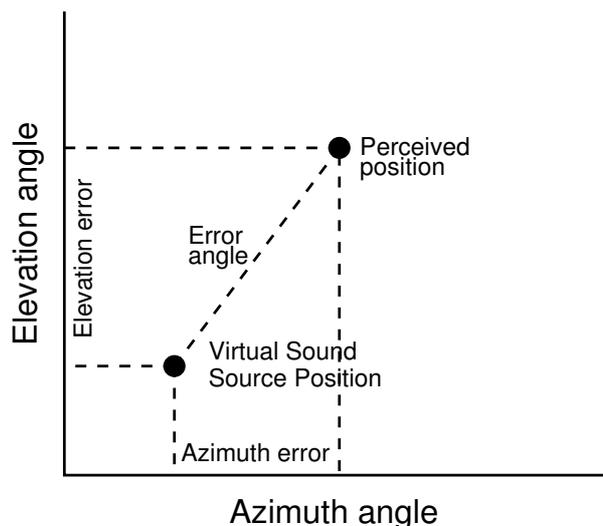


Figure 4: Positions are defined using azimuth and elevation angle from the listening position.

2.4. Results of the experiment

Virtual sound source position and perceived position are defined using azimuth and elevation angles from the listening position. The azimuth and elevation error, are defined as angular difference between source position and perceived position as shown in figure 4. The error angle is the shortest angular distance between the sound source position and perceived position.

The median values of absolute azimuth and elevation errors² in starting points (static sound) are shown in Table 2. The median azimuth localization error for starting points was 7.9 degrees and in elevation 15.0 degrees. These are in line with the overall median accuracy achieved in our previous static experiment [24].

The median values of absolute azimuth and elevations errors for moving sound trajectories are shown in Table 3. It is a median of all measurements. As expected the error increased due to movement of the sound source. The median azimuth error for trajectories was 17.0 degrees and median elevation error was 25.4 degrees. The error in elevation was larger than the error in azimuth. This is in agreement with our previous dynamic experiment [2] without distracting auditory stimulus.

	Azimuth	Elevation
Minimum	11.6	17.2
Maximum	22.6	36.7
All subjects	17.0	25.4

Table 4: Minimum and maximum median values of absolute azimuth and elevation error for subjects.

The median absolute errors for each subject are presented in figure 5. There was a remarkable difference both in azimuth and elevation accuracy between the subjects. In table 4 the minimum and maximum of subjects median errors are represented. The maximum azimuth error was almost twice as large as the minimum.

²absolute azimuth error = abs(perceived azimuth - source azimuth)

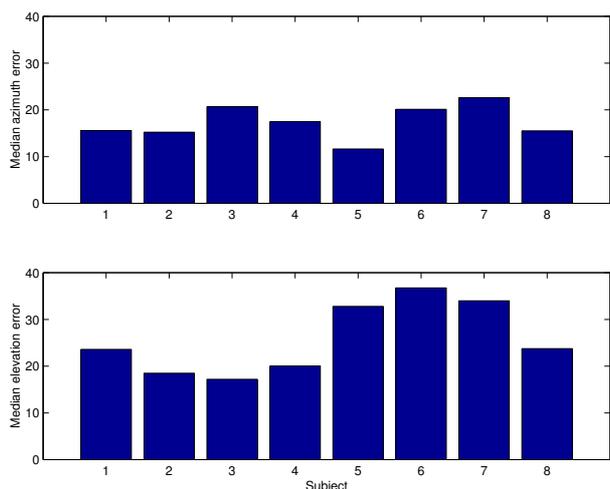


Figure 5: Median of absolute azimuth and elevation errors for all test subjects.

Also the maximum elevation error was approximately twice as large as minimum. An interesting phenomenon is, that the most accurate subject in azimuth (subject number 5) is among the worst ones in elevation accuracy. On the other hand the most accurate subject in elevation (subject number 3) is among the worst ones in azimuth accuracy.

For the figures 6, 7 and 8 the most representative examples were chosen. The difference between subjects is easily seen in figures 6 and 7. These figures are a combination of three different plots. On the left, there is an azimuth-elevation plot displaying the trajectories of target sound, distracting sound and measured pointing values for each signal. In the middle figure, the time dependent change in azimuth is presented, and on the right the time dependent change in elevation is presented. Subject number 4 (in Figure 6) perceived the change in elevation while subject number 5 (in Figure 7) failed to notice the change. On the other hand subject number 5 accurately followed the azimuth position of the target sound.

In most of the cases the distracting stimulus only decreased the accuracy of localization. In ten percent of the cases there occurred a confusion, in which the subject pointed at the distracting stimulus instead of the target. An example of the situation where subject pointed consistently for approximately five seconds at the static distracting sound instead of the target sound is shown in figure 8.

Although there were large differences between subjects, figures 6, 7 and 8 indicate that subjects were consistent in their perceptions.

Due to the VBAP reproduction the angular distance between the target sound and the nearest loudspeaker had an influence on accuracy. The error angle (see figure 4) increased while the angular distance between the target sound and nearest loudspeaker increased. In figure 9, the x-axis presents the angular distance (in degrees), and y-axis presents the error angle. In this plot the distance values were stratified using one degree accuracy. For each group the median and standard deviation were computed. The median is represented with a line and the standard deviation with error bars. The correlation coefficient value for the medians is 0.91.

The angular distance between the target sound and the dis-

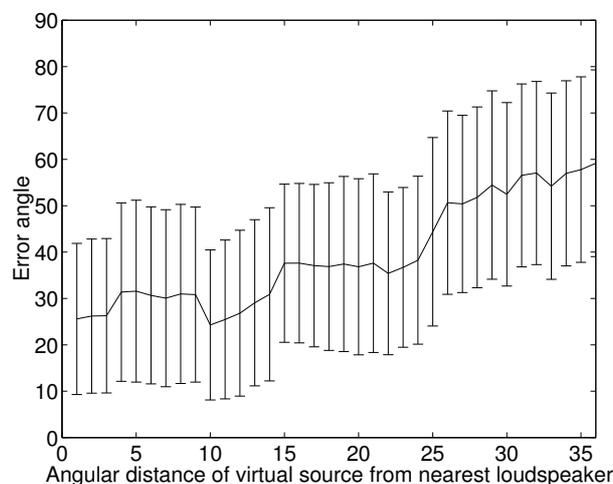


Figure 9: Dependency between error angle and angular distance of target sound and nearest loudspeaker.

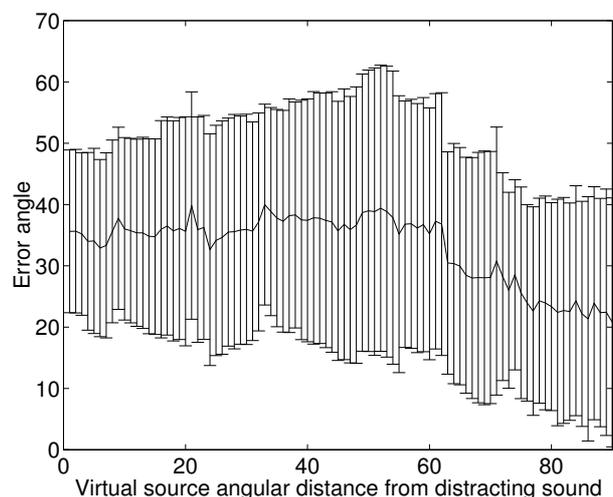


Figure 10: Dependency between error angle and angular distance of target sound and distracting sound.

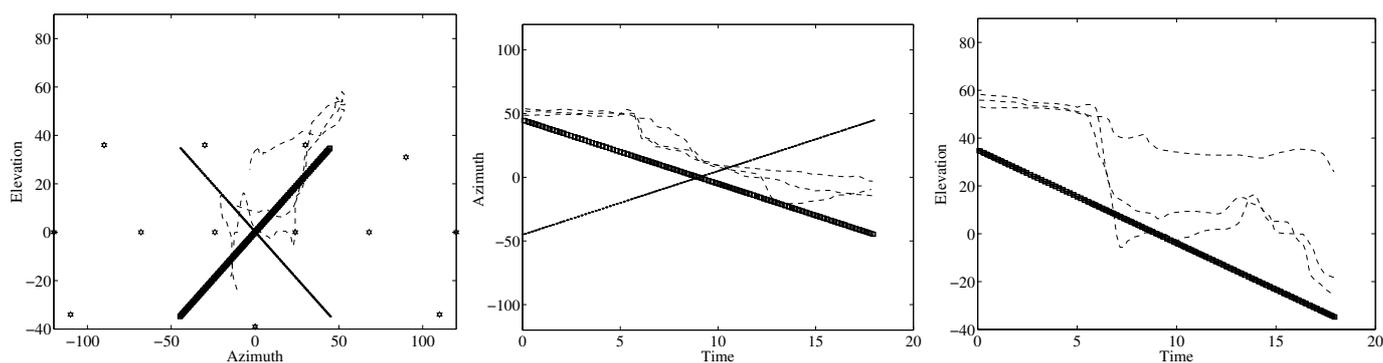


Figure 6: Trajectory plots for subject number 4 for one target sound and distracting stimulus combination. The measured locations for all three signals are presented using dashed lines. The position of the target sound source is indicated with thick line and the positions of loudspeakers are indicated with star-sign. In azimuth-elevation plot the locations of loudspeakers are indicated with star-sign. The starting point of the target sound is located in 45 degrees in azimuth and 35 degrees in elevation. The end point of the target sound is located at -45 degrees in azimuth and -35 degrees in elevation. The starting point of the distracting stimulus is located at -45 degrees in azimuth and 35 degrees in elevation. The end point of the distracting stimulus is located at 45 degrees in azimuth and -35 degrees in elevation.

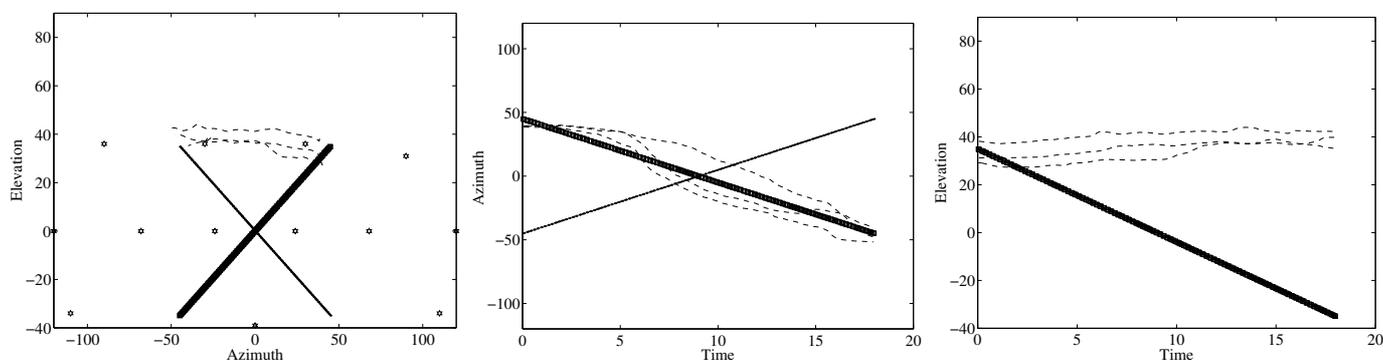


Figure 7: Trajectory plots for subject number 5 for the same target sound and distracting stimulus combination as in figure 6.

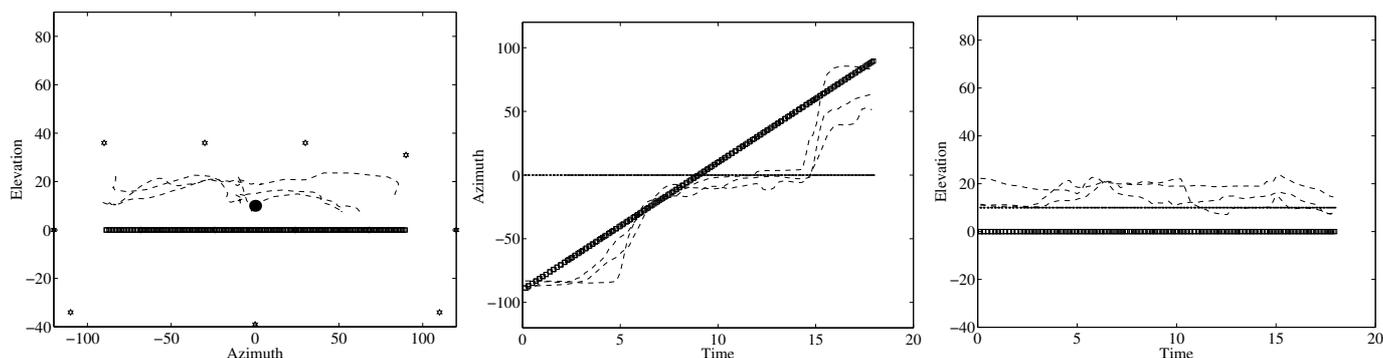


Figure 8: Trajectory plots for subject number 2 for another target sound and distracting stimulus combination. The starting point of target sound is located at -90 degrees in azimuth and 0 degrees in elevation. The end point of the the target sound is located at 90 degrees in azimuth and 0 degrees in elevation. The distracting stimulus is statically located at 0 degrees in azimuth and 10 degrees in elevation (black dot in azimuth-elevation plot).

tracting sound also had influence on accuracy, though the effect was smaller and the correlation not as strong (correlation coefficient for the medians -0.71). In figure 10, the x-axis presents the

angular distance and the y-axis presents the error angle. The error angle decreased while the angular distance increased.

2.5. Comparison of results with previous experiment

As expected, the error levels in this experiment at starting points (Table 2) were in line with the error levels in our previous experiment. For all signals the absolute median azimuth error in our previous experiment 6.8 degrees, and in this 7.9 degrees. In elevation the error in our previous experiment was 15.3 degrees and in this 15.0. At the starting point the distracting stimulus was not yet presented to the subjects.

In the moving sound trajectories the distracting sound had influence on measured azimuth accuracy. The median azimuth error increased almost five degrees. Without distracting stimulus the median error was 12.5 degrees (Table 5) and with distracting stimulus 17.0 degrees (Table 3). In elevation the difference was not remarkable (24.1 degrees without distracting stimulus and 25.4 with distracting stimulus).

	Azimuth	Elevation
Pink noise	13.8	26.6
Pink Floyd	12.5	22.8
Frog croaks	11.1	22.7
All signals	12.5	24.1

Table 5: Median values of absolute azimuth and elevation errors in trajectories in our previous experiment without distracting stimulus

The decreased accuracy is seen in figure 11. On the left are the measurements without distracting stimulus and on the right with distracting stimulus. Especially the perception of the elevation has degenerated. In the azimuth-elevation plot without the distracting stimulus, the shape of the triangle is roughly recognizable in dots although the measured trajectories are bent towards the loudspeaker positions. On the right, the shape of the triangle is more blurred in the dot cloud.

3. DISCUSSION

According to Blauert [6] the localization blur in azimuth is approximately one degree in optimal conditions. The localization blur for elevation is more signal dependent and it can vary from four degrees (white noise) to seventeen degrees (continuous speech by unfamiliar person). Familiarity with the signal also plays a role in elevation perception.

With interfering noise the localization blur is dependent on signal levels and frequencies[6]. If the level of the target signal is about 10 - 15 dB above the interfering noise, localization blur is in the same level as it is without the interfering noise. On the other hand, in an article by Tuyen and Letowski [25] it has been mentioned that a 6 dB signal-to-noise ratio is appropriate for tasks requiring accurate frontal localization. In the experiment described in this article the distracting stimulus increased the localization blur, although the gain difference between the target sound and distracting sound was 10 dB.

In our environment there was more localization blur than in optimal conditions. That is natural because there are several factors degrading the localization in a virtual room as listed in Table 6. On the other hand, localization blur in direct loudspeaker reproduction [24] in our environment was in line with the accuracy that Sandvad [7] achieved in an anechoic chamber.

In precedence effect research [6] it has been found, that in addition to auditory cues, cues from other sensory modalities and prior knowledge are taken into consideration. It can be assumed that non-auditory cues and prior knowledge could be effective also in situation where the precedence effect is not taking place. In the experiment described in this article the same stimulus was used as a target sound and as a distracting stimulus. Subjects were expected having more confusions with sounds than they had. Due to the method used in this experiment (target sound was played before the distracting sound is played) the subjects were concentrated on the target sound and the effect of the distracting sound was diminished.

Suzuki [16] and his colleagues have explored the influence of noise on sound localization. They found out, that for a signal with frequency below 1 kHz the perceived sound location shifts away from the distracting noise. For higher frequencies the localization accuracy was decreased, but no common tendency among subjects was found. In the similar kind of test [6] localization blur with 3 kHz sinusoidal signals was higher than with 500 Hz sinusoidal signals, which is in agreement with results above. All the signals used in the experiment described in this article were broadband signals having frequencies also above 1 kHz. With distracting sound the localization accuracy was more inaccurate than without distracting sound. This is in line with results by Suzuki and his colleagues.

Pulkki [9] has found in his listening tests that perceiving elevation of a virtual source is highly individual in VBAP reproduction. Our results in previous experiment and results in this experiment support his findings.

Grantham [26] has defined a minimum audible movement angle (MAMA), that should be exceeded before it is possible to perceive the direction of the moving sound source. Under the optimal circumstances (slowly moving sound presented directly in front of the subject) the MAMA is between two to five degrees (for the azimuth changes in horizontal plane). MAMA is one of the reasons for the delays in the beginning of the trajectories as seen for example in figures 6 and 8.

Ballas et al. [27] explored the effect of auditory rendering on perceived movement. According to their experiment increasing the number of loudspeakers in VBAP reproduction enhanced the accuracy in perceived movement. That is in line with results achieved in this article: error angle increases when the angular distance between the target sound and nearest loudspeaker increases. Increasing the number of loudspeakers will decrease the maximum distance to the nearest loudspeaker. The problem in a multimodal virtual environment is that due to a visual display configuration, there are lots of areas where one cannot set loudspeakers. The visual display configuration defines the lower limit for maximum distance to nearest loudspeaker. In our environment it might be possible to add still a few loudspeakers.

4. CONCLUSIONS

Due to several factors (see Table 6) there was more localization blur in our virtual room than in optimal conditions. Compared to results achieved without a distracting auditory stimulus it is clear that the distracting auditory stimulus decreased especially the azimuth accuracy. The distracting stimulus had practically no influence on elevation accuracy.

There were remarkable differences between the subjects. The most accurate subjects had a median error level, that was half of the median error level of the most inaccurate subjects. Even though

Reproduction	Environment	Perception	Pointing
non-optimal loudspeaker positioning use of VBAP	acoustics of the virtual room screens diffuse the signals screens low-pass filter the signals	localization blur MAMA	inaccuracy of pointing

Table 6: Factors degrading the localization accuracy in a virtual room.

there was this large variation between the subjects, each subject was consistent and in most cases more or less the same trajectory for all the different signals was perceived .

As expected, the angular distance between target sound and the nearest loudspeaker had an effect on error angle. The error increased while the angular distance increased. The angular distance between the target sound and the distracting sound had less influence on error angle than was expected. Although in this experiment the same signal was used as target and distracting sound, the subjects typically did not get confused.

The consistency of the results with previous experiments suggests that the test method is reliable.

Further experiments will be needed to examine the effects of different gain levels between the target sound and distracting stimulus. In addition, different target sound and distracting stimulus combinations should be explored.

The long term goal is to use auditory stimuli in immersive visualization tasks like visualization of building services or scientific visualization. Therefore, the effect of simultaneous visual stimuli in localization of auditory stimulus should also be examined.

5. ACKNOWLEDGEMENTS

I would like to thank Mr. Tommi Ilmonen for his work for our audio software and hardware. I would also like to thank Mr. Tapio Lokki, Prof. Lauri Savioja and Prof. Tapio Takala for their support and valuable comments. Ms. Ursula Holmström has helped me with the language.

This work has been partly financed by the Helsinki Graduate School in Computer Science and Engineering (HeCSE, <http://www.cs.helsinki.fi/hecse>), HPY Research Foundation, and KAUTE Foundation.

6. REFERENCES

- [1] G. Kramer, *Auditory Display: Sonification, audification and auditory interfaces.*, Addison-Wesley, Reading, MA., 1994.
- [2] M. Gröhn, T. Lokki, and T. Takala, "Static and dynamic sound source localization in a virtual room," in *Proc.AES 22nd Int. Conf. on Virtual, Synthetic and Entertainment Audio*, Espoo, Finland, June 15-17 2002.
- [3] F.L. Wightman and D.J. Kistler, "Localization of virtual sound sources synthesized from model HRTFs," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'91)*, New Paltz, NY, 1991.
- [4] E.M. Wenzel, "Localization in virtual acoustic displays," *Presence: Teleoperators and Virtual Environments*, vol. 1, no. 1, pp. 80-107, 1992.
- [5] E. Wenzel, M. Arruda, D. Kistler, and S. Foster, "Localization using non-individualized head-related transfer functions," *Journal of Acoustic Society of America*, vol. 94, pp. 111-123, 1993.
- [6] J. Blauert, *Spatial Hearing, The psychophysics of human sound localization.*, The MIT Press, Cambridge, MA, 1997.
- [7] J. Sandvad, "Dynamic aspects of auditory virtual environments," in *the 100th Audio Engineering Society (AES) Convention*, Copenhagen, Denmark, May 11-14 1996, preprint no. 4226.
- [8] V. Pulkki, "Localization of amplitude-panned virtual sources I: Stereophonic panning," *Journal of the Audio Engineering Society*, vol. 49, no. 9, pp. 739-752, Sept. 2001.
- [9] V. Pulkki, "Localization of amplitude-panned virtual sources II: Two- and three-dimensional panning," *Journal of the Audio Engineering Society*, vol. 49, no. 9, pp. 753-767, Sept. 2001.
- [10] J. Jalkanen, *Building a spatially immersive display - HUT-CAVE. Licenciate Thesis*, Helsinki University of Technology, Espoo, Finland, 2000.
- [11] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456-466, June 1997.
- [12] D.G. Malham and A. Myatt, "3-d sound spatialization using ambisonics techniques," *Computer Music Journal*, vol. 19, no. 4, pp. 58-70, 1995.
- [13] A.J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *Journal of the Acoustic Society of America*, vol. 93, May 1993.
- [14] J. Hiipakka, T. Ilmonen, T. Lokki, M. Gröhn, and L. Savioja, "Implementation issues of 3D audio in a virtual room," in *Proc. SPIE*, San Jose, California, Jan 2001, vol. 4297B.
- [15] E. Wenzel, "Effect of increasing system latency on localization of virtual sounds," in *Proc. AES 16th Int. Conf. on Spatial Sound Reproduction*, Rovaniemi, Finland, April 10-12 1999, pp. 42-50.
- [16] Y. Suzuki, T. Yokoyama, and T. Sone, "Influence of interfering noise on the sound localization of a pure tone," *Journal of Acoustic Society of Japan*, vol. 14, no. 5, pp. 327-339, 1993.
- [17] P. Minnaar, S.K. Olesen, F. Christensen, and H. Møller, "Localization with binaural recordings from artificial and human heads," *Journal of the Audio Engineering Society*, vol. 49, no. 5, pp. 323-336, May 2001.
- [18] R.L. Martin, K.I. McAnally, and M.A. Senova, "Free-field equivalent localization of virtual audio," *Journal of the Audio Engineering Society*, vol. 49, no. 1/2, pp. 14-22, Jan./Feb. 2001.

- [19] T. Djelani, C. Pörschmann, J. Sahrhage, and J. Blauert, "An interactive virtual-environment generator for psychoacoustic research II: Collection of head-related impulse responses and evaluation of auditory localization," *ACUSTICA acta acustica*, vol. 86, pp. 1046–1053, 2000.
- [20] F.L. Wightman and D.J. Kistler, "Factors affecting the relative salience of sound localization cues," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R.H. Gilkey and T.R. Anderson, Eds., pp. 1–23. Lawrence Erlbaum Associates Inc., 1997.
- [21] R.O. Duda, "Elevation dependence of the interaural transfer function," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R.H. Gilkey and T.R. Anderson, Eds., pp. 49–75. Lawrence Erlbaum Associates Inc., 1997.
- [22] J.C. Middlebrooks, "Spectral shape cues for sound localization," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R.H. Gilkey and T.R. Anderson, Eds., pp. 77–97. Lawrence Erlbaum Associates Inc., 1997.
- [23] V. Pulkki, *Spatial Sound Generation and Perception by Amplitude Panning Techniques*, Ph.D. thesis, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, report 62, 2001.
- [24] M. Gröhn, "Is audio useful in immersive visualization?," in *Proc. SPIE*, San Jose, California, Jan 2002, vol. 4660B.
- [25] T.V.Tuyen and T. Letowski, "Evaluation of acoustics beacon characteristics for navigation tasks," *Ergonomics*, vol. 43, no. 6, pp. 807–827, 2000.
- [26] D.W. Grantham, "Auditory motion perception: Snapshots revisited," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R.H. Gilkey and T.R. Anderson, Eds., pp. 295–313. Lawrence Erlbaum Associates Inc., 1997.
- [27] J.A. Ballas, D. Brock, J. Stroup, and H. Fouad, "The effect of auditory rendering on perceived movement: Loudspeaker density and HRTF," in *Proc. ICAD 2001*, Espoo, Finland, Jul 2001, pp. 235–238.

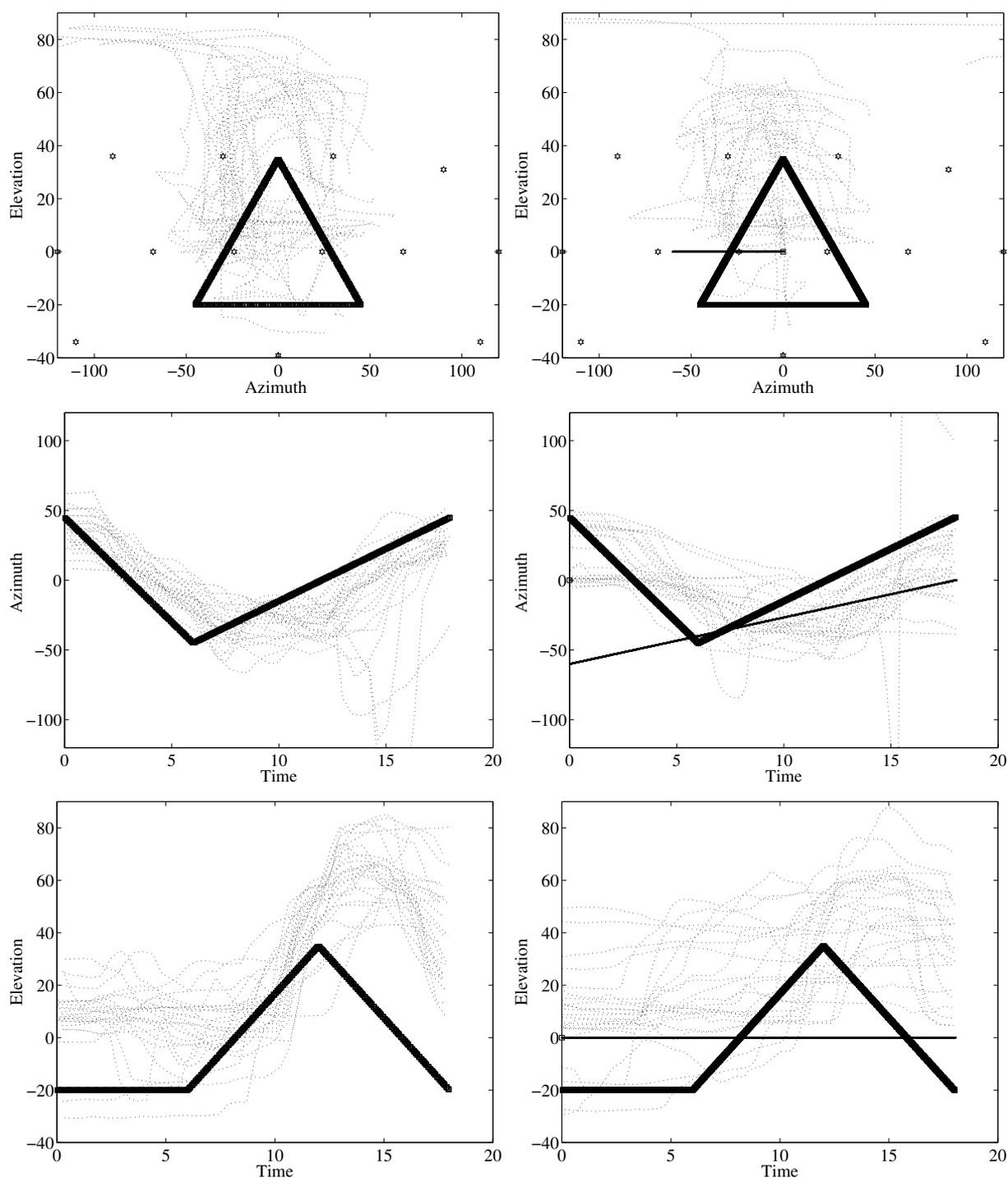


Figure 11: On the left azimuth-elevation, time-azimuth and time-elevation plots for trajectory without distracting sound and on the right the same plots for trajectory with distracting sound. In azimuth-elevation plots the locations of loudspeakers are indicated with star-sign. There are 24 measured trajectories in each figure. The starting point of the target sound is located in 45 degrees in azimuth and -20 degrees in elevation. The first turning point is located in -45 degrees in azimuth and -20 degrees in elevation. The second turning point is located in 0 degrees in azimuth and 35 degrees in elevation. The end point of the target sound is the same as the starting point. On the right, the starting point of distracting stimulus is located in -60 degrees in azimuth and 0 degrees in elevation. The end point of distracting stimulus is located in 0 degrees in azimuth and elevation