On the Use of Eye Movements in Acoustic Source Localization Experiments

Sascha Spors, Robert Schleicher, Dirk Jahn, and Robert Walter
Quality and Usability Lab, Deutsche Telekom Laboratories, Technische Universitt Berlin, Germany.
Email: {Sascha.Spors,Robert.Schleicher}@telekom.de

Introduction

The ability of virtual auditory environments and sound reproduction techniques to provide well localized virtual sources is a key feature. Experiments that judge the localization are hence amongst the first ones performed in evaluation. One difficulty in such experiments lies in reporting the perceived direction by the test subjects. Various exo- and egocentric methods have been developed and used in the past for this purpose. See e.g. [1, 2] for an overview. Applied exocentric methods are for instance marking the perceived direction on paper or using a graphical user interface. Egocentric methods are for instance hand-pointing, head-pointing [3] and eye-pointing [4]. Also combined head- and eye-tracking has been used to obtain the absolute gaze direction [2, 5]. We aim at a method that allows to perform source localization experiments with mostly untrained subjects that are restrained as little as possible in order to obtain results under natural conditions. Highly accurate and unobtrusive techniques for head tracking are commercially available. Most of the published studies based on eye-pointing used the scleral search coil technique [2, 4, 5] for eye movement recording. Besides its high accuracy this method requires to insert a contact lense containing the coil into the subjects eye which can create quite some discomfort. In this study we used a commercially available camera-based head-mounted eye tracker. It is hence less obtrusive than the scleral search coil techniques. The eye tracker provides the relative binocular gaze direction at high spatial accuracy and sampling rate after a short calibration. We combined the relative gaze direction of the eye tracker with a head tracker in order to obtain the absolute gaze direction.

In order to evaluate the accuracy and applicability of this setup, a basic localization experiment was conducted where the users had to report the perceived direction by fixation of the presumed source direction. This paper will present the general setup, as well as results concerning the accuracy of the method.

Experimental Setup

The experimental setup is illustrated by Figure 1. An array of seven active monitors (Genelec 8030APM) was placed in front of the test subjects. The loudspeakers were located on a semi-circle with a radius of 2 m. One loudspeaker was placed in front of the subject, the other six distributed equiangular with a spacing of 15° to the left and right. The angular range covered by the outmost loudspeakers is ±45°. The loudspeakers were hidden from the subjects by a grey acoustically transparent

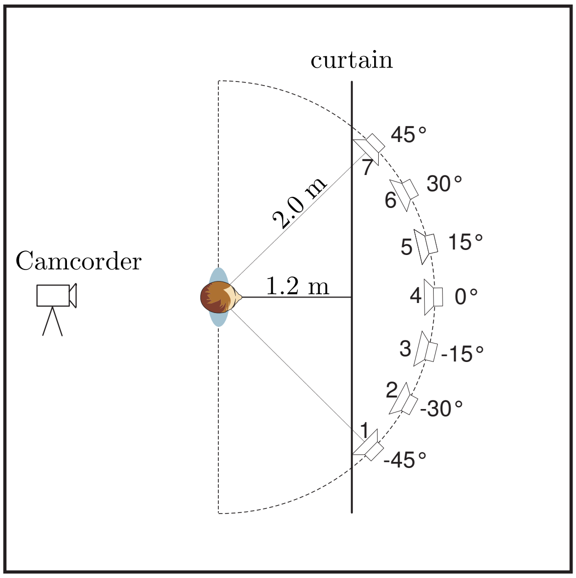

Figure 1: Setup used for the localization experiments.

curtain in order to avoid visual anchors. A white strip of 1 cm height with vertical lines and intermittent dots every 1 cm was attached to the curtain in order to facilitate stable fixation.

Binocular eye movements were tracked by a head mounted Eyelink II tracker from SR Research. The eye movements were recorded at a rate of 250 Hz using the internal log. The head orientation and position was tracked using a Fastrak 6-degrees of freedom motion tracker from Polhemus. Its sensor was attached to the eye tracker helmet. The head tracking data was transmitted to the eye tracker and stored also in its log. A standard bimanual game pad was used for user feedback.

The video from the helmet camera of the eye tracker, with an overlay of the detected gaze direction, was recorded for each experiment, as well as a video of the entire scene captured by a camcorder placed behind the subject. Playback of the audio material and also the entire experiment was controlled by a personal computer.

The experiment took place in an acoustically treated room of size $5.2 \times 5.3 \times 3$ m (w × l × h). The reverberation time of the room at low frequencies was $R_{\mathrm{T60}} \approx 450$ ms and considerably lower for mid to high frequencies. The noise level in the room was very low and the lights dimmed.

Stimuli and Procedure

A total of 70 test items were presented in random order to each subject per session. This resulted in 10 test

items per loudspeaker position. Each stimuli consisted of a series of four pink noise burst with each burst having 1000 ms duration followed by a 400 ms pause. The playback level was adjusted to 60 dBA sound pressure level at the position of the subject.
Upon arrival, the subject was seated and the height of the chair was adjusted to align the eye level with the reference strip on the curtain. The subject was introduced to the experimental procedure and made familiar with the stimulus. The eye tracker helmet was fitted and the eye tracker was calibrated by a short device-specific procedure. The experiment began after fulfilment of these steps. The subjects were asked to listen to the stimuli, look in its direction, fixate its presumed origin on the reference strip and press a button on the game pad. No initial fixation was requested. Only one trial per stimuli was allowed.
Two blocks of experiments were conducted: reporting the perceived direction (1) by gaze and (2) with a laser pointer. This paper will only report the results from the first block. Refer to [6] for results on reporting by a laser pointer and a comparison of both methods.
Ten normal hearing subjects (4m/6f) with age 20-30 participated in the experiment. The subjects were predominantly naive listeners with no experience in localization experiments.

Results

The test subjects reported no difficulties in performing the required tasks. The eye and head-tracker continuously recorded all eye and head movements at a high rate throughout the experiment. All experimental results were collected in the log of the eye tracker. The log has been exported for subsequent analysis of the data. Post-processing included the combination of the horizontal eye and head-tracker angles in order to derive the absolute gaze direction, the removal of invalid data, like e.g. resulting from lid movements and outliers. The tracked head position was used to correct the relative angle under which the subjects see the loudspeakers. The final gaze direction was determined by averaging over the last ten samples of the absolute gaze direction before the button of the game pad was pressed.
Data analysis revealed that the measurements of the gaze direction are reasonably well normal distributed around the true angle. No subject showed systematic outliers. This allows to use mean and standard deviation as measure for the accuracy of the evaluated method. Figure 2 shows the localization error for all 10 subjects for the seven different loudspeaker positions. Localization error is defined as the difference between the true loudspeaker position and the absolute gaze direction. The results show that the evaluated method provides very accurate results for head-unrestrained and untrained subjects. The absolute error is below 1.5 degrees with a variance not exceeding 0.5 degrees. The results also reveal a slight systematic tendency to the underestimate the true source angle. However, this could be corrected for in a localization experiment.

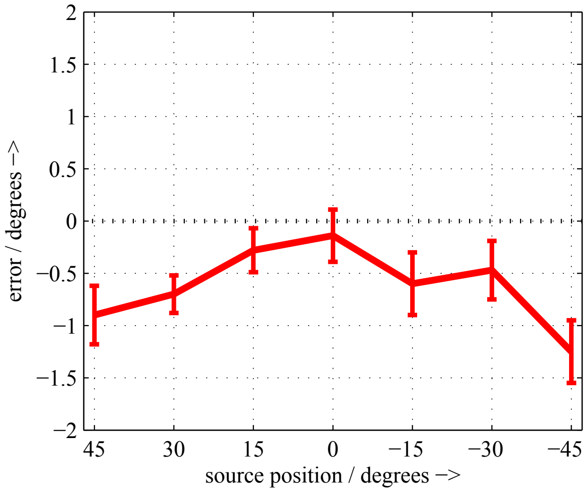

Figure 2: Mean and standard deviation of localization error for all 10 test subjects.

Conclusions and Outlook

The presented results confirm that a combination of head- and eye-tracking can be used to obtain the perceived direction of a sound source from test subjects. This study used a head-mounted camera-based eye-tracker in contrary to other studies based on e.g. the scleral search coil techniques. The used eye-tracker is less obtrusive for the test subjects than these methods. The main focus of the presented study was to evaluate the applicability and accuracy of the particular setup. The results revealed a high accuracy and the test subjects did not report any major difficulties. A more detailed discussion of this experiment, as well as a comparison using a laser pointer as reporting method has been published in [6].
Further analysis of the recorded data is planned. For instance, the separate recording of eye- and head-tracker data allows to analyze head movements and absolute gaze direction independently from each other.

References

[1] B. Seeber. A new method for localization studies. *ACUSTICA - acta acoustica*, 83(1-2):1–5, 1997.

[2] L.C. Populin. Human sound localization: measurements in untrained, head-unrestrained subjects using gaze as a pointer. *Experimental Brain Research*, 190(1):11–30, September 2008.

[3] A.W. Bronkhorst. Localization of real and virtual sound sources. *Journal of the Acoustical Society of America*, 98(5):2542–2553, November 1995.

[4] P.M. Hofman and A.J. van Opstal. Spectro-temporal factors in two-dimensional human sound localization. *Journal of the Acoustical Society of America*, 103(5):2634–2648, May 1998.

[5] H.H.L.M. Goossens and A.J. van Opstal. Influence of head position on the spatial representation of acoustic targets. *Journal of the American Physiological Society*, 81:2720–2736, 1999.

[6] R. Schleicher, S. Spors, D. Jahn, and R. Walter. Gaze as a measure of sound source localization. In *AES 38th International Conference on Sound Quality Evaluation*, June 2010.