

DISTANCE VARIATION FUNCTION FOR SIMULATION OF NEAR-FIELD VIRTUAL AUDITORY SPACE

Alan Kan, Craig Jin, André van Schaik

Computing and Audio Research Laboratory, School of Electrical and Information Engineering,
University of Sydney, Australia

ABSTRACT

We present a method for simulating a near-field sound source in virtual auditory space (VAS). The method scales individualised HRTFs, measured at a distance of 1m to arbitrary distances in the near-field. It uses a model of the acoustic scattering for a point-source on a rigid sphere to calculate a distance variation function (DVF) to apply to the HRTFs. A sound localisation experiment was conducted in VAS with three subjects to evaluate the acoustic spatial fidelity of this method. Results show that with the modified HRTFs directional localisation is generally maintained at different distances and there is reasonable correlation between the perceived distance and target distance for distances up to 50cm from the centre of the subject's head.

1. INTRODUCTION

A head-related transfer function (HRTF) characterises the pressure transformation of a sound source from its position in free space to the listener's ear. HRTFs are commonly measured acoustically for both ears and may be used to generate virtual auditory space (VAS) using headphones. The HRTFs are generally measured at a fixed distance, within 1-2m from the centre of the listener's head.

It has been shown that HRTFs are largely independent of distance in the far-field, i.e., beyond approximately 1m from the centre of the listener's head, since the waves emitted from sound sources at these distances can be approximated as plane waves when they reach the head. However, in the near-field, i.e., distances within 1m, HRTFs vary considerably with distance because the plane wave source approximation is no longer valid.

A number of difficulties in measuring HRTFs in the near-field have so far prohibited accurate reproduction of near-field sound sources in virtual auditory displays [1]. Firstly, a small, broadband acoustic point source is required for near-field HRTF measurements. Secondly, a subject's head position must be tightly constrained so that small movements during measurement do not lead to substantial parallax errors. Thirdly, an increased number of measurement points are required to cover a large range of

distances, increasing the duration of the measurement process.

To date, HRTF measurements in the near-field have been made on an acoustic manikin and the results presented in [2]. These HRTFs were used in an auditory localisation experiment reported in [3] and compared to results in free-field auditory localisation of broadband noise stimulus reported in [4]. The results showed that subjects could tell distance and location of virtual sound sources generated using the non-individualized HRTFs, but performance was poorer than localization of nearby sound sources in the free-field.

In this paper, we present a simple method for simulating sound sources in the near-field by calculating a distance variation function (DVF) that avoids the technical difficulties associated with measuring HRTFs in the near-field. The model was tested in a sound localization experiment that measured the ability of three subjects to judge the distance and direction of virtual sound sources.

2. METHOD

2.1. Near-field simulation

We use a rigid sphere model to give a reasonable first-order approximation for the changes in human HRTFs at different source distances. A model for the frequency response at the surface of a rigid sphere for a sound source located an arbitrary distance from the sphere is derived in [5, 6]. For a rigid sphere of radius a , the sound pressure at a location, $\vec{x} = (a, \theta_s, \varphi_s)$, due to a sinusoidal point-source of sound at a frequency, f , wave number, $k = 2\pi f/c$, and at a location, $\vec{q} = (r, \theta_k, \varphi_k)$, is given by:

$$p_s(a, \theta_s, \varphi_s; k, r) = -\frac{4\pi}{ka^2} \sum_{n=0}^{\infty} \frac{h_n(kr)}{h'_n(ka)} \sum_{m=-n}^n Y_n^m(\theta_k, \varphi_k) Y_n^{m*}(\theta_s, \varphi_s),$$

where c is the speed of sound, $h_n(kr) = j_n(kr) + jy_n(kr)$ is a modified spherical Bessel function of the first kind of order n , and $Y_n^m(\theta, \varphi)$ is a spherical harmonic function of degree n and order m . The pressure, $p_s(a, \theta_s, \varphi_s; k, r)$, at

the surface of the rigid sphere can be calculated for all frequencies of interest in order to determine a pressure transfer function at the surface of the sphere due to a point-source of sound at a specified distance, r .

To modify HRTFs measured at an initial distance D_I to simulate a sound source from the same direction at a target distance D_T , we calculate $S_I = p_s(a, \theta_s, \varphi_s; k, D_I)$ and $S_T = p_s(a, \theta_s, \varphi_s; k, D_T)$. The numerical value for a is determined by the size of the listener's head and can be pre-calculated from the set of HRTFs using Kuhn's model [7]. The numerical values for the azimuth and elevation angles (θ_s, φ_s) are determined by the location of the listener's ears on his/her head. The distance variation function (DVF) can then be calculated as $DVF = S_T/S_I$. This function can be used to scale the frequency response of the HRTF for the desired source distance and direction. A different DVF is calculated for each source distance and direction. Figure 1 shows the DVF for the left ear and right ear for a sound source to the right of the listener at different distances.

Using the rigid-sphere model, the DVF approximates the change in interaural level difference as a function of distance. It also accounts for low frequency parallax effects which would be present for a sound source in the near-field close to a human head. However, high frequency parallax effects are not accounted for as there are no pinnae on the rigid sphere. In addition, it is unclear whether calculating the angle of parallax and choosing the appropriate far-field HRTFs will give the correct high frequency effects.

2.2. Stimulus

A psychophysical experiment was conducted to evaluate localisation performance for near-field virtual sounds simulated using the method described above. Individualised HRTFs were measured in an anechoic chamber at a distance of 1m from the subject for 393 different directions at an 80 kHz sampling rate using a blocked ear canal method [8]. From the 393 measured directions, 76 directions were tested for this experiment, roughly equally spaced around the head between 40° and -40° elevation. The HRTFs for these directions were resampled to 48 kHz. An ear canal resonance, measured on a Brüel and Kjær Head and Torso Simulator (HATS) manikin, was added to the measured HRTFs to compensate for the missing ear canal resonance when using ER-2 in-ear tube phones for presentation. Because the HRTF measurements are at the limit of the noise floor for frequencies below 500 Hz, the measured HRTFs were compensated below this frequency for each direction according to the frequency response derived from the rigid sphere model in [6].

The DVF was calculated as described above for 512 frequency bins and the positions of the ears were assumed to be 100° on either side of the midline on the horizontal

plane. The appropriate HRTF was modified using the DVF and convolved with 400ms of white noise with a 5ms raised cosine onset and offset ramp. Near-field sound stimuli were generated evenly and randomly for distances within the following four ranges (in cm): [10-20], [25-35], [40-60], [60-100].

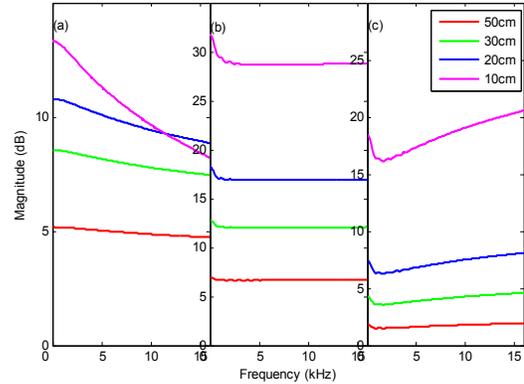


Figure 1 DVF for the left ear (a), and right ear (b) for a source at different distances directly to the right of the listener: (c) The difference between the right and left DVF.

2.3. Experimental Setup

A sound localization task was set up in which a subject was required to position the end of a ruler at the perceived sound source location. This is similar to the direct localization task described in [4]. At the beginning of a trial subjects begin by facing straight ahead in the calibrated start position which is indicated by an LED display. After pressing a push button, a sound stimulus is played over the in-ear tube phones and the subject points the end of the ruler to the perceived location of the sound source. The subject presses the pushbutton once again to indicate the completion of the localization task. At this point, the position of the end of the ruler is read by an electromagnetic tracking system (Polhemus Fastrak) and the subject returns to the calibrated start position for the beginning of the next trial.

The Polhemus Fastrak system is used in a two-sensor configuration to accurately measure the subject's perceived sound source location relative to the centre of the head. One sensor is mounted on top of a rigid headband worn by the subject and is used to accurately measure the location and orientation of the subject's head at the beginning of each trial. The second sensor is attached to the end of the ruler which is used for pointing and indicating the target location. The position of this sensor is read at the end of each trial.

The position tracking system is calibrated at the start of the experiment by measuring the location of the left and right ear relative to the sensor located on top of the head.

The position sensor at the end of the pointing ruler is used to make these measurements and the position of the midpoint along the interaural axis is identified as the centre of the subject's head. These position measurements ensure that an accurate position reading can always be taken of the centre of the subject's head during the experiment.

A handheld response pushbutton which is connected to the control computer is provided to the subject. The control computer is equipped with an RME Digiface connected to an Alesis AI-3 digital-to-analogue converter. The analogue sound signal was amplified using a TDT System3 HB7 headphone buffer and presented to the subject using Etymotic ER-2 in-ear tube phones. The experiments were conducted in an ordinary room with no acoustical treatment. The room was lit during the experiments so that the subject could see the sensor at the tip of the ruler. Apart from the Polhemus position-tracking system and the LED display, all other equipment were located outside the experiment room.

The sound stimuli were presented randomly in blocks of 76 trials with each trial consisting of the 76 test directions at varying distances. Each subject completed 12 blocks, consisting of a total of 912 stimuli. Three subjects participated in the experiment. All had prior training in auditory localisation experiments using head-pointing to indicated perceived source direction, but were new to the current localisation paradigm, using a ruler to indicate both absolute distance as well as direction.

3. RESULTS AND DISCUSSION

3.1. Directional Localisation

Positions on a sphere may be broken down into lateral and polar angle, which is a convenient coordinate system for describing localisation results. The lateral angle is the angle to a circle perpendicular to the interaural axis and varies from -90° to $+90^\circ$. The polar angle is the rotation angle around the interaural axis, with 0° representing the horizontal plane in front, 90° representing directly above, 180° behind, and 270° directly below. Figure 2 shows the localization data in terms of lateral angle for each subject. The results show that there are no major localization errors in terms of matching between the lateral angles of the target and response directions and that there is no major increase in lateral error at different simulated source distances. The mean lateral angle error across the three subjects for the 60-100cm region was 2.3° and for the 10-20cm region it was 3.3° .

The results, in terms of polar angles, are shown in Figure 3. In these trials, Subject 2 showed the least amount of front-back errors but had a tendency to localize sounds about 12° higher than their target polar angle location. The other two subjects do not show this bias but display a number of front-back errors when the target sound source was at the front. Figure 4 shows the percentage of front

back errors per subject. The average number of front-back errors for all subjects was around 11% for distances less than 50cm and 7% for distances greater than 50cm. This is comparable to the results presented in [9] for free-field localization of near-field broadband stimulus.

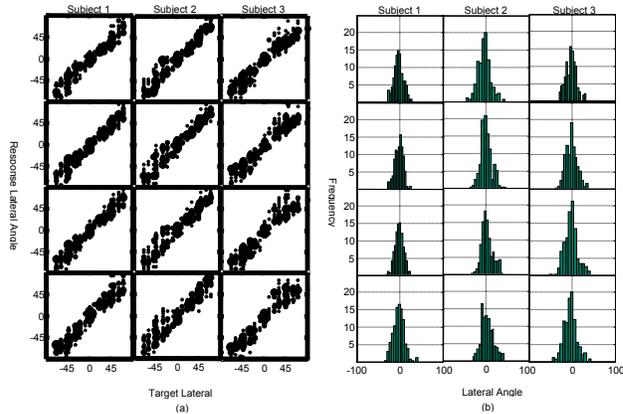


Figure 2 (a) The lateral angle results for each subject at the target distances of [60-100cm], [40-60cm], [25-35cm], and [10-20cm] is shown down the column. (b) Histogram of lateral angle errors per subject at the mean distances.

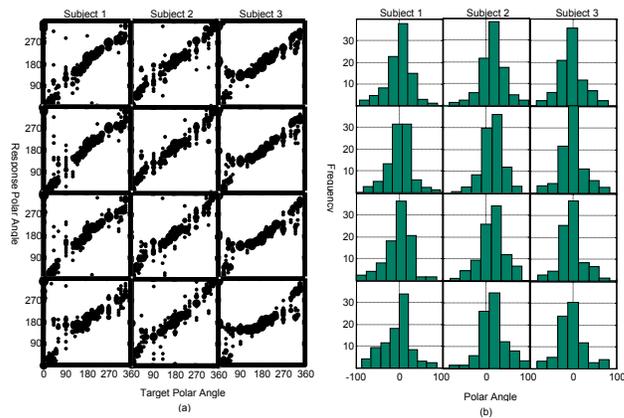


Figure 3 (a) The polar angle results for each subject at the target distances of [60-100cm], [40-60cm], [25-35cm], and [10-20cm] is shown down the column. (b) Histogram of polar angle errors per subject at mean distances.

3.2. Distance Localisation

The raw performance data for distance localization is shown in Figure 5. The solid line indicates the perfect response for each distance and the dashed line is the best linear fit to the data. The dotted line fits the data restricted to the range between 10cm and 50cm. From the data, there is reasonable

correlation between target and response distances between 10cm and 50cm, especially at the lateral positions ($60^\circ \leq |\text{azimuth}| \leq 120^\circ$), but subjects have a tendency to overestimate distances in their responses at these distances. At distances greater than 50cm, there is a greater variability in responses and subjects tend to underestimate the distance of the target. This is likely related to the small effect of the DVF at these distances (see Figure 1). These observations are similar to those presented in [4] for localisation of near-field sources in the free field, where the magnitude of distance errors tend to increase with distance and distance errors are greater at the front and back than more lateral locations.

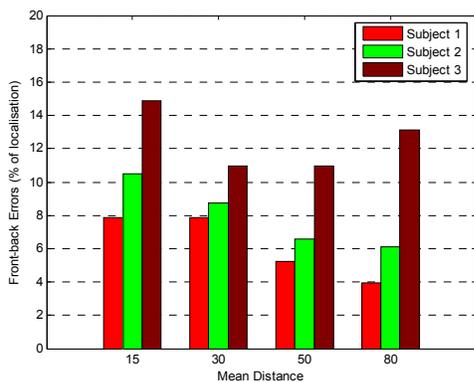


Figure 4 Percentage of front-back errors made by each subject.

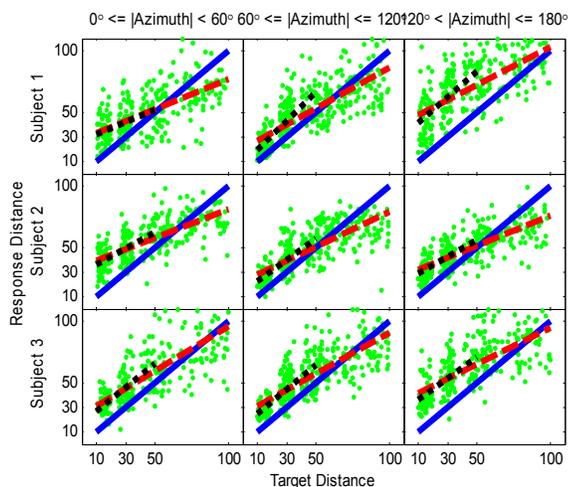


Figure 5 Distance localisation data for each subject is shown along each row. The solid line is the ideal response, the dashed line is the best linear fit of the data and the dotted line is the best linear fit for the data between 10cm and 50cm.

5. CONCLUSIONS

We have presented a simple method for presenting near-field sounds in virtual auditory space. Psychoacoustic

testing with three subjects demonstrates that directional localisation accuracy in the virtual space is comparable to that in the free-field at all distances. For distance localisation, there is reasonable correlation for distances up to 50cm from the centre of the head and is similar to the accuracy obtained in the free-field. However, there is greater variability in distance localization performance for distances greater than 50cm.

6. ACKNOWLEDGEMENT

The DVF technique described in this paper is patent pending and gratitude is given to VAST Audio Pty. Ltd. for permission to publish this work.

7. REFERENCES

- [1] D. S. Brungart, "Near-Field Virtual Audio Displays," *Presence*, vol. 11, pp. 93-106, 2002.
- [2] D. S. Brungart, "Auditory localization of nearby sources. Head-related transfer functions," *Journal of the Acoustical Society of America*, vol. 106, pp. 1465-1479, 1999.
- [3] D. S. Brungart and B. D. Simpson, "Auditory localization of nearby sources in a virtual audio display," presented at 2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, Oct 21-24 2001, New Paltz, NY, 2001.
- [4] D. S. Brungart, N. I. Durlach, and W. M. Rabinowitz, "Auditory localization of nearby sources. II. Localization of a broadband source," *Journal of the Acoustical Society of America*, vol. 106, pp. 1956-1968, 1999.
- [5] W. M. Rabinowitz, J. Maxwell, Y. Shao, and M. Wei, "Sound localization cues for a magnified head: Implications from sound diffraction about a rigid sphere," *Presence*, vol. 2, pp. 125-129, 1993.
- [6] R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model," *The Journal of the Acoustical Society of America*, vol. 104, pp. 3048-3058, 1998.
- [7] G. F. Kuhn, "Model for the interaural time differences in the azimuthal plane," *The Journal of the Acoustical Society of America*, vol. 62, pp. 157-167, 1977.
- [8] C. Jin, A. Corderoy, S. Carlile, and A. van Schaik, "Contrasting monaural and interaural spectral cues for human sound localization," *The Journal of the Acoustical Society of America*, vol. 115, pp. 3124-3141, 2004.
- [9] D. S. Brungart, "Auditory localization of nearby sources. III. Stimulus effects," *Journal of the Acoustical Society of America*, vol. 106, pp. 3589-3602, 1999.