Available online at www.sciencedirect.com
SCIENCE DIRECT®
Applied Acoustics 66 (2005) 1219–1234
applied acoustics
www.elsevier.com/locate/apacoust

Spatial sound resolution of an interpolated HRIR library

Jaka Sodnik *, Rudolf Sušnik [1], Mitja Štular [2], Sašo Tomažič [3]

University of Ljubljana, Faculty of Electrical Engineering, Laboratory for Communication Devices, Tržaška 25, 1000 Ljubljana, Slovenia

Received 16 June 2004; received in revised form 7 March 2005; accepted 8 March 2005
Available online 19 May 2005

Abstract

The paper evaluates the human directional resolution of virtual sound sources synthesised with the aid of a generalised head related impulse response (HRIR) library, i.e., an HRIR library measured using a dummy head and torso. The original HRIR set is first expanded using linear interpolation, and then directional resolution measurements are performed for playback through headphones. These results are compared to the results obtained using loudspeakers as sound sources in an anechoic chamber. Directional resolution is the ability of listeners to distinguish two closely-spaced sound sources alternately playing the same signal. Experiments show that two sound sources with insufficient spacing appear as a single source to the listener. Directional resolution for small azimuth changes is relatively high for both virtual and real sound sources. Most test subjects have no problem resolving two sound sources only 5° apart. Compared to real sound sources, detecting changes in elevation of virtual sound sources is much less accurate, which may be the main drawback of using a generalised HRIR library.

Keywords: Virtual sound sources; HRIR; HRTF; Acoustical resolution

* Corresponding author. Tel.: +386 1 4768 494; fax: +386 1 4768 266.
E-mail addresses: jaka.sodnik@fe.uni-lj.si (J. Sodnik), rudolf.susnik@fe.uni-lj.si (R. Sušnik), mitja.stular@fe.uni-lj.si (M. Štular), saso.tomazic@fe.uni-lj.si (S. Tomažič).
[1] Tel.: +386 1 4768 894; fax: +386 1 4768 266.
[2] Tel.: +386 1 4768 432; fax: +386 1 4768 266.
[3] Tel.: +386 1 4768 432; fax: +386 1 4768 266.

0003-682X/$ - see front matter
doi:10.1016/j.apacoust.2005.04.003

1. Introduction

Everyday life is full of different sounds, each carrying some kind of information about our surroundings. We are able to hear sounds and determine their location with different degrees of accuracy. This 3-D sound detection is a major factor in orienting in space, driving a car, communicating with several people at once, etc. The brain detects the position of sound sources by processing electrical impulses generated by sound waveforms impinging on the eardrums. As a sound waveform travels from the source to the eardrum, it is influenced by many factors, e.g., obstacles in space, shape of the body and head, size of the earlobes, shape of the ear canal, etc. A measure of transfer characteristics can be expressed as Head Related Impulse Response (HRIR) or Head Related Transfer Function (HRTF), i.e., the Fourier transform of the HRIR [1]. These representations enable generation of synthetic spatial sounds, which can then be played back through headphones. The HRIR depends on the spatial position of the sound source and should be measured for different spatial positions.

The HRIRs also contain some indirect information about the listener, e.g., the shape of his/her head or ears. This implies the need for personalised HRIRs for each listener to convey accurate spatial information. However, most applications are intended for a wider audience, and frequently, generalised HRIRs are used. They are usually measured using dummy heads and bodies. This introduces a certain deviation in reproduction and diminishes the ability of the listeners to accurately pinpoint the location of a virtual sound source.

These kinds of generated spatial sounds are most commonly used in computer games, military aircraft dogfight simulations and navigation systems for the visually impaired, with the intended representation of a visual picture of the space with sound. In this type of application, spatial resolution is very important.

Acoustic resolution is defined as the ability of a listener to differentiate nearby sound sources in space. Very important is the minimal angle between two sound sources where the listener can still differentiate between them.

Our previous measurements with the MIT Media Lab HRIR library [2–4] have shown that about 87% of test subjects can distinguish two virtual sound sources at the same elevation (0°) and 5° of difference in azimuth. Outside of the central azimuth, this resolution drops off, with only about 50% of subjects still able to differentiate a 5° difference at an azimuth of 45°. Resolution in the vertical plane is much lower, which is a drawback of using generalised HRIRs.

It turns out that resolution is heavily influenced by the choice of source sound signals [5,6]. The key factors are bandwidth and duration of the signal. These earlier results led to an idea of improving azimuth resolution using a more precise HRIR library. Using the new CIPIC HRIR library [7,8] and interpolation, functions with only a 1.3° difference in azimuth were calculated. There are four reasons for choosing the CIPIC library:

- it is very well documented,
- it contains three different non-personified libraries,

- impulse responses are previously compensated for limitations of the loudspeaker and microphones employed, and
- it suggests linear interpolation for acquisition of impulse responses for new directions.

For comparison, measurements were repeated with real sound sources loudspeakers in an anechoic chamber.

2. HRIR

HRIRs contain all key factors needed to reproduce a spatial sound source. The direction of the sound source is defined by its azimuth (left-right) and elevation (up-down). The key factors for perception of sound source azimuth are the Interaural Time Difference (ITD) and the Interaural Level Difference (ILD). These phenomena are illustrated in Figs. 1 and 2. The ITD is the delay between the arrival of a sound waveform to the first, and then the second, eardrum [9]. The ILD is the amplitude difference of these two waveforms. The incident sound wave for the shaded ear (the ear on the opposite side of the head, relative to the sound source) is attenuated in comparison to the direct sound wave.

When detecting the elevation of a sound source, the key factor is the spectral content of the sound signal at the eardrum, which depends on the spectral content of the source signal and on the HRTF [10]. HRTF magnitudes for two different elevations

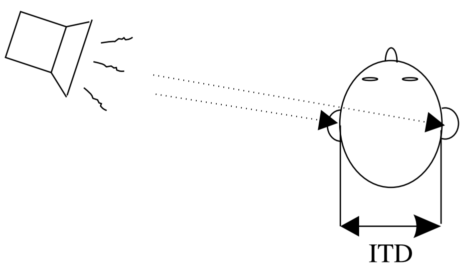

Fig. 1. Time difference between incident waveforms for the direct and the shaded ears (ITD).

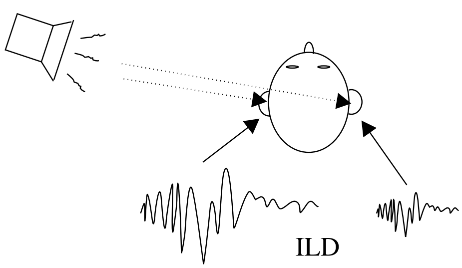

Fig. 2. The amplitude difference between incident waveforms for the direct and the shaded ears (ILD).

are shown in Fig. 3. HRTF is influenced by the shape of the shoulders, head, and ears.

Generalised HRIRs are usually measured in an anechoic chamber using a dummy head and torso. Two small microphones are inserted into the ears, representing the eardrums. A sound source is placed a certain distance away, which is then detected by the microphones in the ears. The dummy is attached to a rotating table, and the speaker is attached to a variable height support, enabling different azimuth and elevation settings.

HRIR is measured simply as the response of the system to a unit impulse. The other possibility is to perform the measurement in the frequency domain and then calculate the response using the inverse Fourier transform. A suitable filter can be used to compensate for the characteristics of the measuring equipment (speaker, microphones, reproduction unit, etc.). The impulse responses for different azimuths and elevations for each ear are then represented as finite impulse response filters. Two different responses are shown in Fig. 4.

The impulse responses are used for generation of virtual spatial sound in the headphones. The source signal is shaped separately for each ear. The left and right ear signals are produced using convolution of the source signal with the impulse response [11]:

$$x_{l,\phi,\theta}[n] = \sum_{k=0}^{N-1} x[n-k]h_{l,\phi,\theta}[k], \tag{1}$$

$$x_{r,\phi,\theta}[n] = \sum_{k=0}^{N-1} x[n-k]h_{r,\phi,\theta}[k], \tag{2}$$

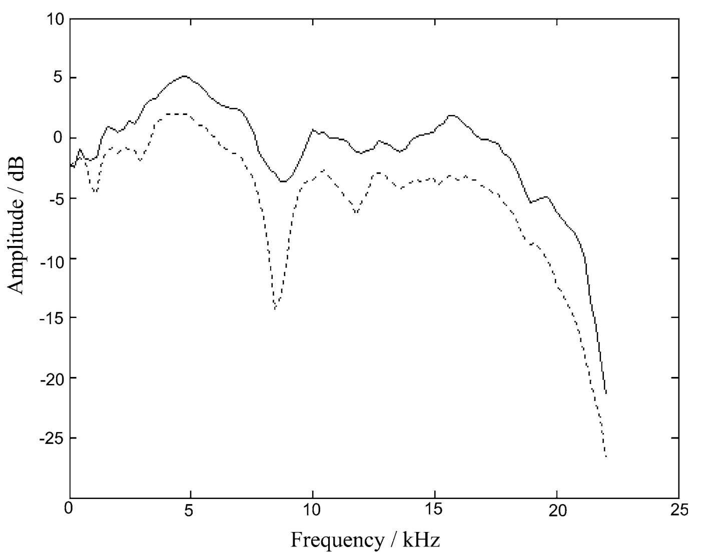

Fig. 3. HRTF magnitudes.

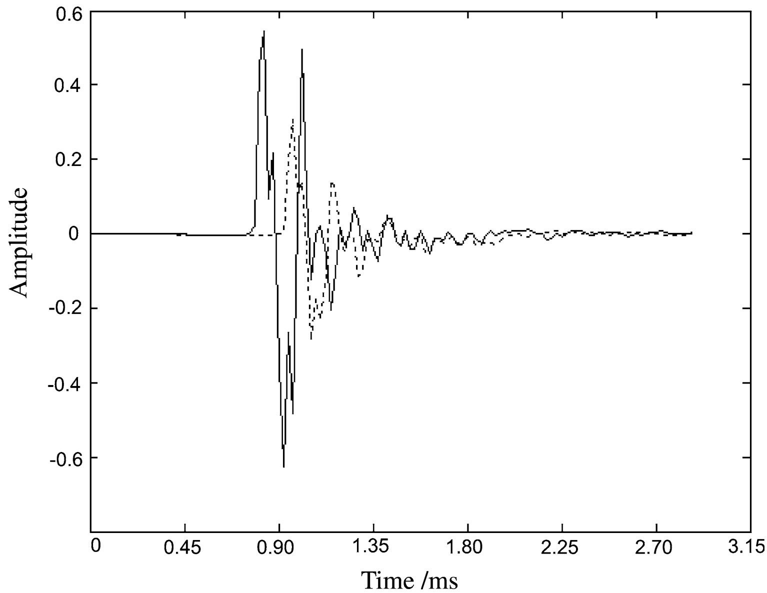

Fig. 4. Head-related impulse responses.

where $x_{l,\phi,\theta}[n]$ and $x_{r,\phi,\theta}[n]$ denote the signals at the left and right eardrum at azimuth ϕ and elevation θ, and $h_{l,\phi,\theta}[m]$ and $h_{r,\phi,\theta}[m]$ denote appropriate impulse responses at these azimuths and elevations. N denotes the number of samples for each response.

Generalised functions introduce deviations that must be taken into account. The most common problem for listeners is their inability to determine whether the source is behind them or in front of them ("front-back" confusion). In some cases, source elevation is also a problem, with the directions "up" and "down" being interchanged.

3. Description of the problem

The purpose of our experiments was to evaluate the resolution of a generalised HRIR library, i.e., to determine the smallest angle between two sources where the listener can still differentiate between two virtual sound sources. The sound signals of the two sources were the same, which made differentiation of the sources more difficult. Azimuth and elevation resolution were measured separately. In two separate experiments we first reproduced the spatial sounds through headphones, using generalised HRIRs, and subsequently, real sound sources in an anechoic chamber. The same source signal was used in both cases.

4. Methods

4.1. Original HRIR library

The library used for measurements was the CIPIC HRIR library. The original library contains HRIRs for 1250 locations in space. HRIRs are given for elevations from $-45°$ to $230°$ with $5.625°$ spacing. The spacing in azimuth is $5°$ for the range of $0°$–$45°$, and $10°$ or less in the range of $45°$–$80°$.

From our previous experiences with generalised HRIRs, we concluded that the spacing in elevation was small enough for our purposes; however, the spacing in azimuth needed to be decreased.

4.2. Expanded HRIR library

To increase the azimuth resolution of the original CIPC library, we can expand it using linear interpolation. HRIR libraries have been interpolated in many different ways by the Itakura team [12–14] and simple linear interpolation proved to have the largest signal-to-deviation ratio (SDR) [12], an objective means to evaluate the effectiveness of different interpolation methods. Simple linear interpolation has been shown to be appropriate for interpolating functions where azimuth spacing is less than $10°$, which in our case holds for azimuths from $0°$ to $45°$.

The initial step in the interpolation process is to align two neighbouring HRIRs on the time axis, i.e., to determine the relative delay between the two responses. As the two neighbouring HRIRs are very similar to each other, the position of the maximum of the cross-correlation function [11] corresponds to this delay [15,16]. The sample rate used in the original HRIR library is too low to allow alignment with sufficient accuracy. HRIRs of the original library are thus first over-sampled by a factor of 10, using low-pass interpolation [17]. Cross-correlation between responses at azimuths ϕ_1 and ϕ_2 and elevation θ is then:

$$R_{\phi_1,\phi_2,\theta}[m] = \frac{1}{K}\sum_{k=1}^{K-1} h_{\phi_1,\theta}[k]h_{\phi_2,\theta}[k+m], \tag{3}$$

where $h_{\phi 1,\theta}$ and $h_{\phi 2,\theta}$ denote over-sampled HRIRs, and K denotes the new number of HRIR samples. The delay Δ is equal to the value of m that maximizes the above expression. The HRIR between two neighbouring HRIRs is then equal to

$$h_{\frac{(\phi_1+\phi_2)}{2},\theta}\left[k+\frac{\Delta}{2}\right] = \frac{h_{\phi_1,\theta}[k] + h_{\phi_2,\theta}[k+\Delta]}{2}. \tag{4}$$

The intermediate HRTF thus obtained is then decimated back to the original sample rate. The interpolation is illustrated in Fig. 5.

If the interpolation procedure is repeated, an HRIR library with $1.3°$ accuracy in the frontal area is obtained.

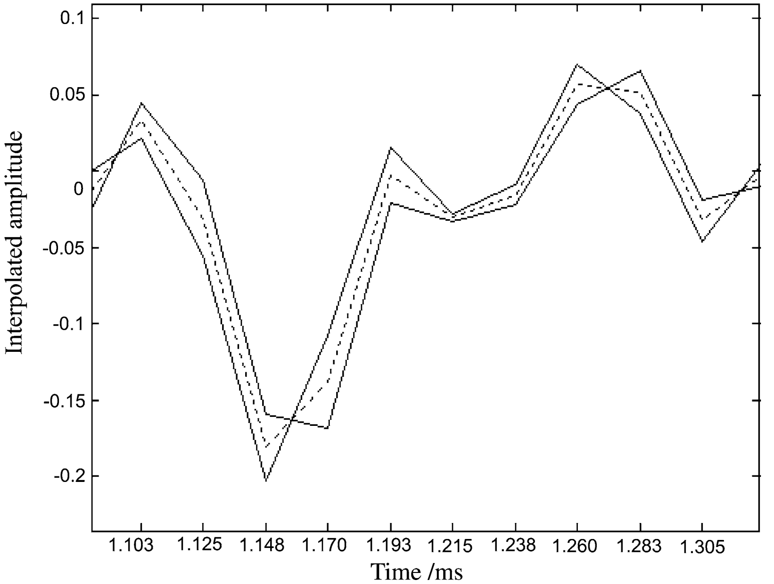

Fig. 5. Linear interpolation of impulse responses for intermediate azimuth values.

5. Test subjects

There were 32 volunteer test subjects, 15–50 years old, not aware of any sight or hearing problems. None of the test subjects had any previous experiences with virtual spatial sounds played through headphones. The test subjects reported that they found the experiment interesting, which assured their high concentration.

6. Sound signal and measuring equipment

According to experiences from our previous experiments [5], the best stimulus for resolution measurements proved to be a white noise sequence 100 ms long. The sequence was generated in a frequency domain to assure a flat spectrum with random phase, and the signal in the time domain was obtained using the Inverse Fourier Transform. The bandwidth of the signal was 22 kHz.

The CIPIC HRIR library was originally compensated to correct for limitations of the loudspeakers employed. Furthermore, we measured headphones impulse responses using Knowles Electronics FG-3329-C miniature microphones. These impulse responses were used to create an inverse filter to equalize the stimulus before its being played.

The measurements were performed on an IBM ThinkPad T30 notebook computer with a Digigram VX Pocket 440 soundcard, Sennheiser Control HD270 headphones and Genelec 8030APM loudspeakers.

Measurements with the headphones were taken in a laboratory with an average noise level of 40 dB. Sennheiser HD270 headphones are studio headphones with excellent attenuation of ambient noise (−10 dB to −15 dB).

Measurements with loudspeakers were made in an anechoic chamber with an A-weighted noise level of 28 dB and signal-to-noise ratio (SNR) of 50 dB. Noise levels and SNRs were measured with a Lutron SL-4012 sound level meter.

7. Measurements

A test environment, developed in Matlab 6.5, was able to play back a sound sequence from different virtual spatial sources. A coordinate grid on the display enabled playback of the sound from any coordinate point with a mouse-click.

The entire measurement sequence consisted of five individual experiments.

The first experiment was an initialisation period, designed to familiarise the test subjects with listening to spatial sounds through headphones. The listeners were given the freedom to explore the environment freely and to listen to various spatial sounds: white noise (100 ms long), pink noise (100 ms long), speech signals, etc. This phase lasted 2 minutes for each test subject.

The second experiment consisted of playing a sound stimulus (white noise sequence 100 ms long) with systematic changes of source position, using the following trajectories:

- Elevation 0°:
  Azimuth: −90°:5°:90°
- Elevation 33.750°:
  Azimuth: −90°:5°:90°
- Azimuth 0°:
  Elevation −45°:5.6250°:45°

This gave the test subjects a feeling for individual directions and increased the sensitivity to azimuth and elevation changes of the virtual sound sources. This phase lasted 5 minutes for each test subject.

In the third experiment, the central position (azimuth 0°) in the horizontal plane (elevation 0°) was chosen, as shown in Fig. 6.

Two different procedures, consisting of three trials, were performed. In the first procedure, the azimuth differences between sound sources, relative to the listener, were decreased systematically: 20°, 10°, 7.5°, 5°, 2.5° and 1.3°. For each azimuth difference, a repeating stimulus, consisting of 100 ms of white noise and 200 ms of silence, was played. The stimulus alternated from one virtual source to the other, never played from both sources at the same time. Test subjects were asked to decide if they heard the stimulus coming from two different directions or just a single direction. Possible answers were:

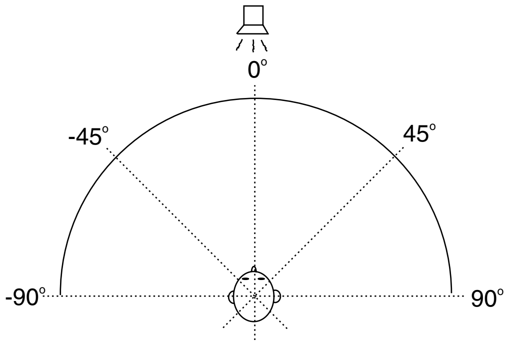

Fig. 6. The central position used to measure directional resolution.

- Two directions (positive)
- One direction (negative)

Subjects had 5 seconds to decide. ''Don't know'' answers were considered as negative. The minimum detectable difference in azimuth (the last one with a positive answer) was considered as the result of the trial.

In the second procedure, azimuth differences between sound sources were increased systematically: 1.3°, 2.5°, 5°, 7.5°, 10° and 20°. Other parameters remained identical to the first procedure.

In the fourth experiment, resolution of elevation was measured. The same central position in the horizontal plane was selected. Elevation differences between the two sources (56°, 45°, 28°, 23°, 11.6° and 6°) were increased and decreased. Other parameters were held constant, as in the second and third experiments.

In the fifth experiment, measurements were repeated with loudspeakers in a large anechoic chamber. Test subjects were situated on a chair with their eyes covered, so that they were unable to see the loudspeaker set-up in the room. The distance *d* between the test subjects and the loudspeakers was fixed at:

- 850 cm for azimuth resolution measurements
- 300 cm for elevation resolution measurements

Table 1
Distances between loudspeakers

Azimuth resolution measurement						
φ_{az}	20°	10°	7.5°	5°	2.5°	1.3°
x_{az}	295 cm	148 cm	111 cm	74.1 cm	37.1 cm	19.3 cm
Elevation resolution measurement						
φ_{el}	56°	45°	28°	23°	12°	6°
x_{el}	273 cm	250 cm	145 cm	120 cm	63 cm	31.5 cm

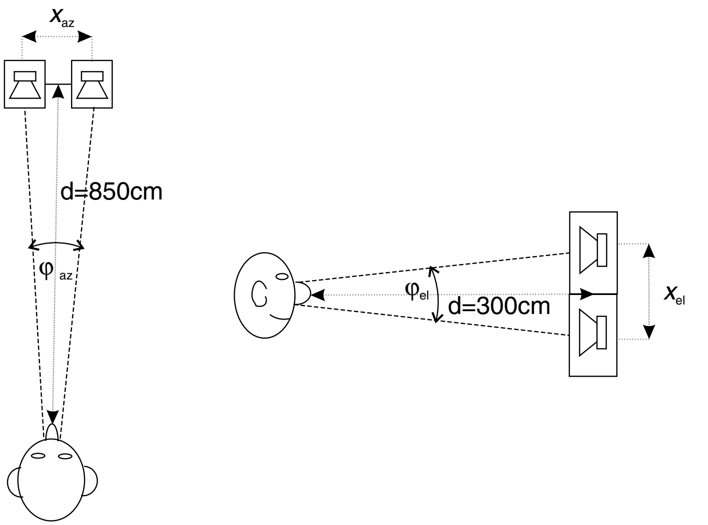

Fig. 7. Loudspeaker set-up for azimuth and elevation resolution measurements.

All parameters remained the same as in the first and second parts of the measurements with headphones. To set up different azimuth angles φ_{az} or elevation angles φ_{el} relative to the listener, distances x_{az} and x_{el} between loudspeakers were changed according to Table 1.

The loudspeaker set-up for both measurements is shown in Fig. 7.

The listeners were again asked to determine whether they could hear one or two directions of the sound sources.

8. Results and discussion

Our first important finding was that the accuracy of the test subjects (their ability to differentiate separate sound sources) was highly dependent on the duration of measurement. When the initialisation phase was longer, the test subjects showed better resolution. The experimental results showed that test subjects needed 3–6 min to achieve constant localisation performance. Therefore, we performed two individual (2 and 5 min) initialisation periods (described in Section 7) to eliminate dependency of measurements on their duration.

When locating virtual sound sources in space played back through headphones, most subjects frequently mistook front for back, and vice versa (front-back confusion). This problem has already been discussed in the introduction and is caused by the use of generalised HRIRs [18].

In experiments 3–5, six results for each test subject were obtained, three when decreasing the angle between the sources and three when increasing the angle. The mean value, the dispersion and confidence interval are calculated for each test subject. The confidence interval μ is defined as [19]

$$\mu = \bar{x} \pm s, \tag{5}$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{6}$$

is the mean value,

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{7}$$

is dispersion, x_i is the result of an individual trial and n is the number of test subjects.

The average minimal azimuth and elevation differences with appropriate confidence intervals for each test subject are presented in Figs. 8–11. Figs. 8 and 9 compare azimuth resolution of virtual sound sources (played through headphones) and real sound sources (played through loudspeakers). Figs. 10 and 11 compare elevation resolutions.

The averages and dispersions for all test subjects are shown in Table 2.

A cursory overview of results and figures offers some important insights. The highest resolution, or human ability to differentiate sound sources in near proximity, is in the horizontal plane. The very noticeable characteristic of Figs. 8 and 10 is the wide confidence intervals, indicating that individual measurements of each test subjects differ considerably. Therefore, it is hard to establish precise resolution when dealing with virtual sound sources (using headphones). Additionally, a significant difference between azimuth resolution and elevation resolution can be substantiated.

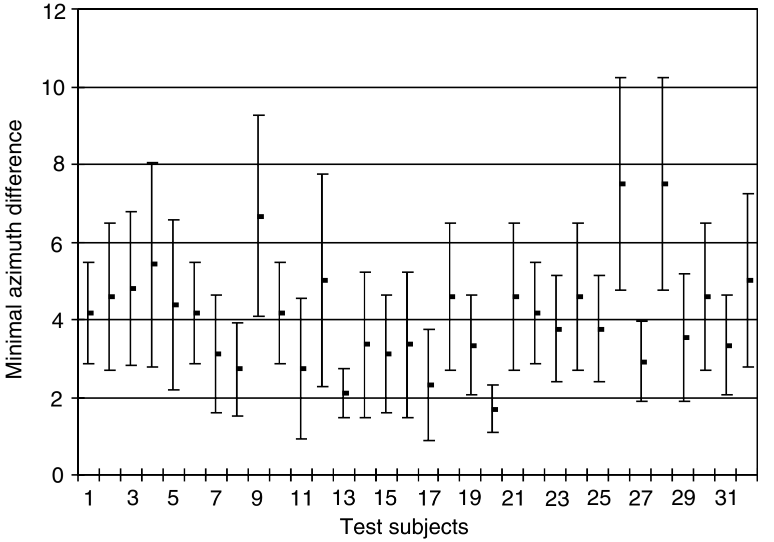

Fig. 8. Azimuth resolution of virtual sound sources (headphones).

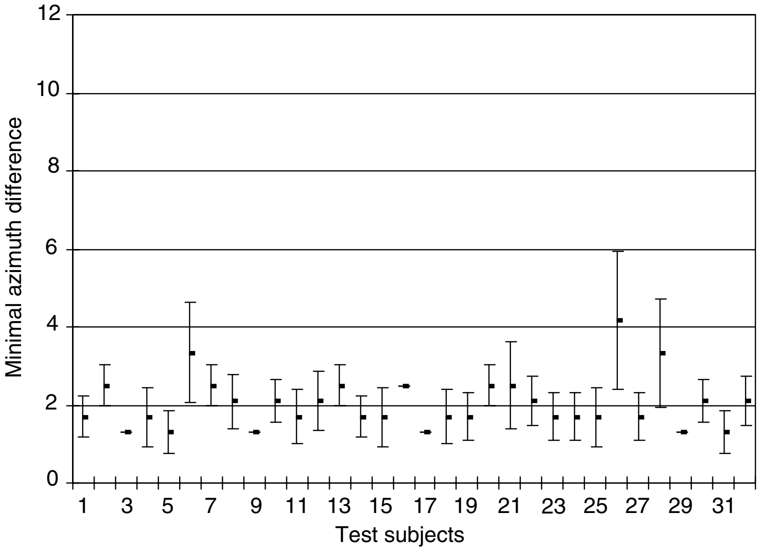

Fig. 9. Azimuth resolution of real sound sources (loudspeakers).

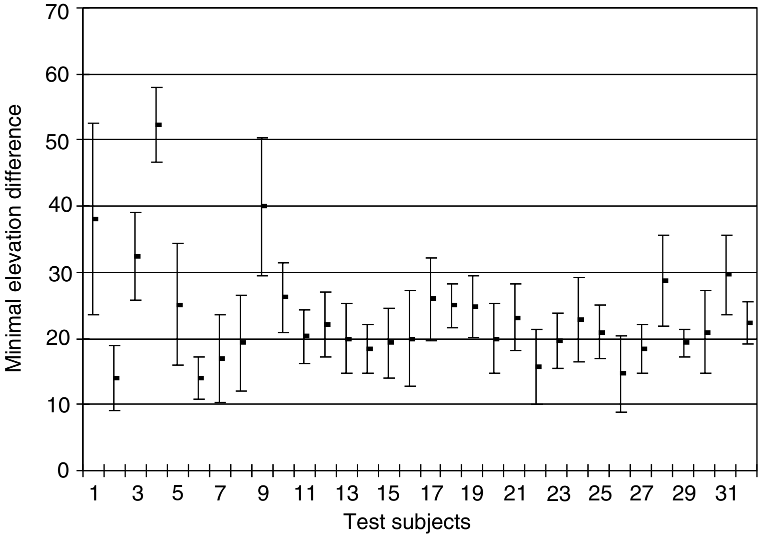

Fig. 10. Elevation resolution of virtual sound sources (headphones).

The latter confirms the limitations of the use of generalised HRIRs. Both azimuth and elevation resolutions are much higher when using loudspeakers; all resolutions are much more confident as well (Figs. 9 and 11).

Further, results of all test subjects are combined into five statistical classes, different for azimuth and elevation resolutions (Table 3):

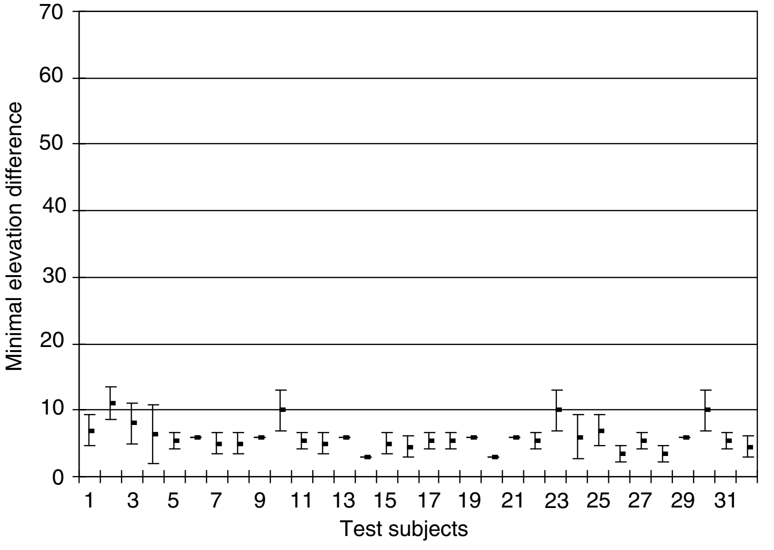

Fig. 11. Elevation resolution of real sound sources (loudspeakers).

Table 2
Average minimal differences and dispersions of all test population

	Headphones (deg)	Loudspeakers (deg)
Azimuth		
Average minimal difference	4.1	2.0
Dispersion	2.2	0.8
Elevation		
Average minimal difference	23.5	6.0
Dispersion	9.8	2.6

Table 3
Percentage of test population in individual statistical classes

Statistical class limits	Headphones (%)	Loudspeakers (%)
Azimuth resolution		
$\varphi_{az} < 2.5^0$	9	72
$2.5^0 \leqslant \varphi_{az} < 5^0$	72	28
$5^0 \leqslant \varphi_{az} < 7.5^0$	13	0
$7.5^0 \leqslant \varphi_{az} < 10^0$	6	0
$10^0 \leqslant \varphi_{az}$	0	0
Elevation resolution		
$\varphi_{el} < 6^0$	0	50
$6^0 \leqslant \varphi_{el} < 12^0$	0	50
$12^0 \leqslant \varphi_{el} < 18^0$	13	0
$18^0 \leqslant \varphi_{el} < 22^0$	41	0
$22^0 \leqslant \varphi_{el}$	46	0

The result in Table 3 confirms the fact that the ability to differentiate individual, closely spaced sound sources is much greater when loudspeakers are used. Our findings also correspond to similar results of previous experiments [20].

9. Conclusions

In our experiments we tried to determine several facts about human spatial sound resolution of virtual sources synthesised with general HRIR libraries. It has already been established that localisation performance drops dramatically when using general HRIRs instead of personified HRIRs. However, in an application developed for a large number of different people, measurement of HRIRs for each individual subject is impossible. Therefore, we dealt with the general HRIRs but focused on one specific part of localisation performance the ability to differentiate between sound sources in near proximity (resolution). In this paper, the measurements of azimuth and elevation resolution are described and compared with real sound sources (loudspeakers).

The results of the conducted experiments show that elevation resolution is much lower than azimuth resolution, even when dealing with loudspeakers. For the perception of azimuth and for high azimuth resolution, accurate ITDs and ILDs are most important, with the spectral content of the functions themselves being of secondary importance. When the azimuth is fixed and elevation is changed, ITD and ILD remain constant; only the spectral content of the HRIRs is of significance. As the personalised and general HRIRs differ primarily in their spectral content, the use of generalised HRIRs causes large errors in elevation perception. Therefore, the difference in azimuth resolution between virtual and real sound sources is not as obvious as the difference in elevation resolution.

Future plans in this field include further resolution measurements with blind test subjects. We expect to establish a much higher resolution because of the better-developed auditory system of the blind. We are also planning further study of the effect of the time component on listening to virtual sound sources. We believe that with some sort of training and continuous listening to virtual sound sources, both horizontal and vertical resolution can be improved, and, in addition, the confidence interval can be narrowed.

Acknowledgements

This work has been supported by the Ministry of Education, Science and Sport of Slovenia, within the program: Algorithms and Optimisation Methods in Telecommunications.

The authors would like to thank Andrea Alvarez Santangelo from Knowles Europe for providing free samples of FG-3329-C miniature microphones.

References

[1] Cheng CI, Wakefield GH. Introduction to head-related transfer functions (HRTF's): representations of HRTF's in time, frequency, and space (invited tutorial). J Audio Eng Soc 2001;49(4):231–49.

[2] HRTF Measurements of a KEMAR Dummy-Head Microphone: http://sound.media.mit.edu/KEMAR.html.

[3] Gardner B, Martin K. HRTF measurements of a KEMAR dummy-head microphone. MIT Media Lab Perceptual Computing-Technical Report #280, 1994.

[4] Sodnik J, Susnik R, Tomazic S. Acoustic signal localization through the use of Head Related Transfer Functions. In: Proceedings on international conference on computer, communication and control technologies, Orlando, FL 2003;(2):100–3.

[5] Susnik R, Sodnik J, Tomazic S. Sound source choice in HRTF acoustic imaging. In: HCI international adjunct proceedings, Heraklion, Crete 2003;101–2.

[6] Susnik R, Sodnik J, Tomazic S. Impact of sound cue length to the acoustic image resolution. In: International electrotechnical and computer science conference, Ljubljana 2003;91–4.

[7] The CIPIC Interface Laboratory Home Page: http://interface.cipic.ucdavis.edu/CIL_html/CIL_HRTF_database.htm.

[8] Algazi VR, Duda RO, Thompson DM, Avendano C. The CIPIC HRTF database. In: Proceedings of the 2001 IEEE workshop on applications of signal processing to audio and electroacoustics, Mohonk Mountain House, New Paltz, NY 2001; 99–102.

[9] Kuhn GF. Model for the interaural time differences in the azimuthal plane. J Acoust Soc Am 1977;62(1):157–67.

[10] Algazi VR, Avendano C, Duda RO. Elevation localization and head-related transfer functions analysis at low frequencies. J Acoust Soc Am 2001;109(3):1110–22.

[11] Oppenheim AV, Schafer RW. Discrete-time signal processing. Prentice-Hall International; 1989.

[12] Nishino T, Mase S, Kajita S, Takeda K, Itakura F. Interpolating HRTF for auditory virtual reality. In: The Third Joint Meeting ASA and ASJ, Hawaii, Sheraton Waikiki Hotel 1996;1261–6 (Proceedings of the papers submitted to the ASJ).

[13] Nishino T, Mase S, Kajita S, Takeda K, Itakura F. Interpolating head related transfer functions in the median plane. In: IEEE workshop on applications of signal processing to audio and acoustics (WASPAA'99), Mohonk Mountain House, New Paltz, NY, 1999.

[14] Nishino T, Ikeda M, Takeda K, Itakura F. Interpolating head related transfer functions. In: The seventh Western Pacific regional acoustics conference (WESTPRAC VII), Mielparque KUMAMOTO, Kumamoto 2000;(1A-1):293–6.

[15] Wenzel EM. Effect of increasing system latency on localization of virtual sounds with short and long duration. In: Proceedings of the 2001 international conference on auditory display, Espoo, Finland 2001.

[16] Roman N, Wang D. Binaural tracking of multiple moving sources. In: Proc ICASSP2003, (5), Hong Kong 2003;149–52.

[17] Tomažič S, Leonardis S. Digitalni signali in sistemi. Fakulteta za elektrotehniko, 2004.

[18] Wenzel EM, Arruda M, Kistler D, Wightman FL. Localization using nonindividualized head-related transfer functions. J Acoust Soc Am 1993;94(1):111–23.

[19] Kreyszig E. Advanced engineering mathematics. 7th ed.. Wayne Anderson; 1993.

[20] Bronkhorst AW. Localization of real and virtual sound sources. J Acoust Soc Am 1995; 98(5):2542–53.

Jaka Sodnik received the B.Sc. degree in telecommunications from the University of Ljubljana, Slovenia, in 2002. He is currently working on his Ph.D. thesis. He is a Junior Researcher at the Faculty of Electrical Engineering. His research interests are acoustics, signal processing and telecommunication networks.

Rudolf Sušnik received the B.Sc. degree in telecommunications from University of Ljubljana, Slovenia, in 2001. He is currently working on his Ph.D. thesis. He is a Junior Researcher at the Faculty of Electrical Engineering. His research interests are acoustics, information theory and telecommunication networks.

Mitja Štular received the B.Sc. degree in 1994, the M.A. degree in 1997 and the Ph.D. degree in 2000, all from University of Ljubljana. He was a Junior Researcher at the Faculty of Electrical Engineering. Currently he is in charge of UMTS project at the largest Slovenian mobile operator Mobitel.

Sašo Tomažič received the B.Sc. degree in 1979, the M.A. degree in 1981 and the Ph.D. degree in 1991, all from University of Ljubljana. He is a professor at the Faculty of Electrical Engineering, University of Ljubljana, where he teaches telecommunications and signal processing. He has already worked as the telecommunications advisor for the Ministry of Defence, information technology advisor for the Ministry of Education and the national coordinator of the telecommunications research in Slovenia. Currently he is the head of the Telecommunications Department and the head of the Laboratory of Communication Devices at the Faculty of Electrical Engineering.