POLYNOMIAL APPROXIMATION OF BINAURAL IMPULSE RESPONSES FOR MOVING SOUND IMAGES

MItsuo Matsumoto, Chiba Institute of Technology
Hirofumi Yanagawa, Chiba Institute of Technology

1. INTRODUCTION

Advances in computers and computer networks have enabled the development of virtual auditory systems[1]. We also have developed a virtual auditory system, the Integrated Sound Field Network (ISFN), that renders many auditory events and describes these events using an artificial language, the auditory events modeling language[2].

Auditory events include not only static sound images, but also moving sound images. In virtual reality systems, the sound image is moved by updating coefficients of a digital filter convolution[3]. Changes in the coefficients may produce clicks. A moving sound image can be simulated using a series of static sound images. Therefore, moving a sound image smoothly and slowly requires a series of slightly different binaural impulse responses. However, measuring many binaural impulse responses is both difficult and time consuming.

Kistler and Wightman introduced a principal components analysis (PCA) that extracts the five basic magnitude functions and approximates head-related transfer functions as linear combinations of the basic functions[4]. They also introduced the minimum-phase component of a binaural impulse response. Extracting the minimum-phase component of a binaural impulse is similar to correcting for arrival time. However, the responses interpolated using minimum-phase components do not always result in minimum-phase responses. We have proposed an algorithm that moves a sound image[5]. This algorithm consists of a time-variant convolution and a method for interpolating binaural impulse responses that allows for arrival time correction[6].

For simulating a moving sound image, describing global changes in binaural impulse responses due to changes in the sound source direction should be better. Binaural impulse responses after arrival time correction gradually change due to changes in the sound source direction. Therefore, we approximate a set of the responses using a set of low degree functions. We evaluated the accuracy of the approximation using the error ratio. We evaluated six binaural signals generated using a rotating dummy head, or simulated using a conventional method and the approximation method.

2. METHOD FOR APPROXIMATING BINAURAL IMPULSE RESPNSES

2.1 MEASURING BINAURAL RESPONSES

Figure 1 shows the path of a sound source moving along an arc. Even when the center of the arc is centered on the listener, the distance, *l* in Fig. 1, between the left ear and the sound source changes as the sound source moves. Changes in the distance appear as changes in the arrival time of the binaural impulse responses.

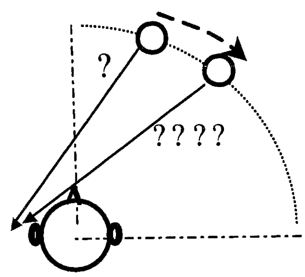


Figure 1 Changes in distance between one ear and sound source

We measured 37 binaural impulse responses in 5° steps, from left to right through the center of an arc using a B & K 4100 dummy head mounted on a turntable, as shown in Fig. 2.

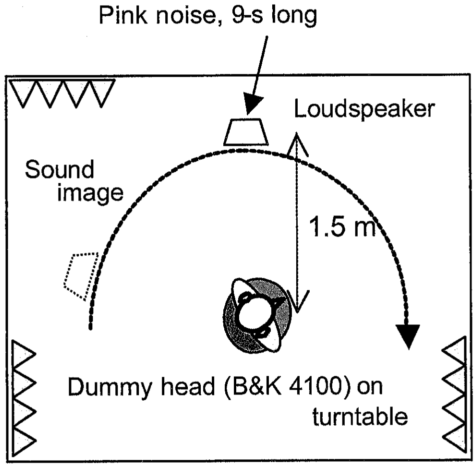


Figure 2 Experimental set-up

The measured responses for the left ear are shown in Fig. 3. The X- (horizontal) and Y- (depth) axes denote the azimuth of the sound source and the sampled time (Fs=352.8 KHz). The Z-axis denotes the instantaneous value of the responses as the sound source moved from left to right.

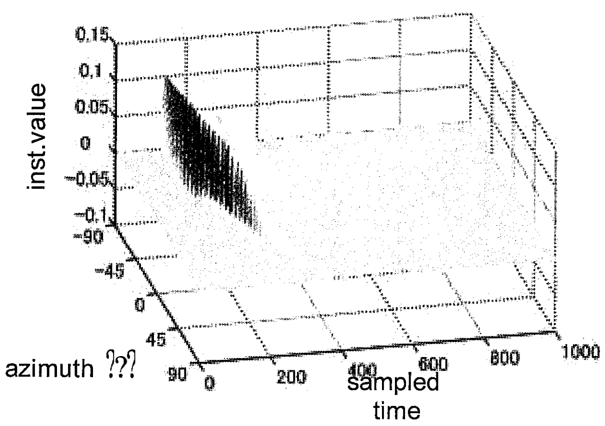


Figure 3 Measured binaural impulse responses – left ear

2.2 ARRIVAL TIME CORRECTION

Figure 4 shows changes in the arrival time of the responses. The X- and Y- axes denote the azimuth of the sound source and the sampled times. As the sound source moves from left to right, the arrival times of the responses for the left ear increased.

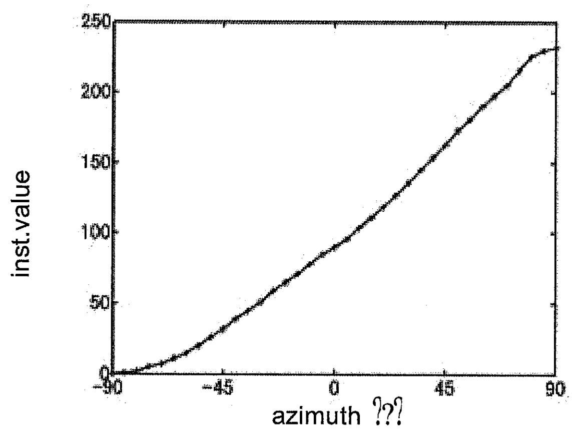


Figure 4 Changes in arrival times due to changes in azimuth of sound source

The responses after arrival time correction are shown in Fig. 5. Figure 5 also shows that the responses gradually changed due to changes in the azimuth of the sound source.

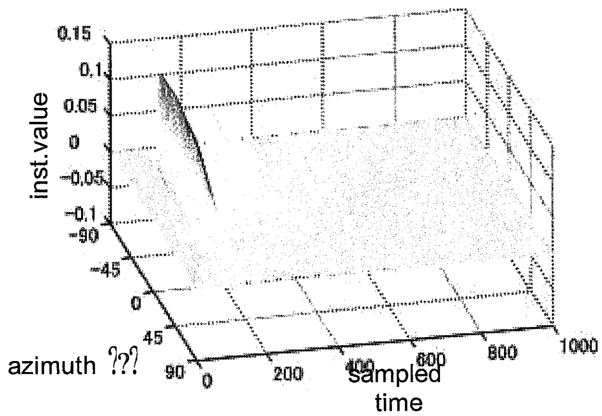


Figure 5 Binaural impulse responses after arrival time correction

2.3 APPROXIMATION OF CHANGES IN INSTANTANEOUS VALUE

The dotted line in Fig. 6 shows the changes in the instantaneous values at the 200th time sample of the measured responses in Fig. 3, due to changes in the azimuth of the sound source. The X and Y- axes denote the azimuth of the sound source and the instantaneous value. The 37 asterisks show the changes in the instantaneous values of the responses with corrected arrival-times for the same time sample. The plot of the asterisks changes gradually when compared to the dotted line. This suggests that changes in the instantaneous values of the binaural impulse responses with corrected arrival-times can be approximated using a low degree function.

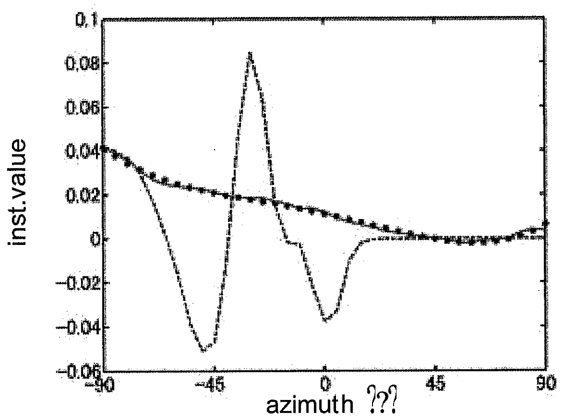


Figure 6 Changes in instantaneous value due to arrival correction and approximated values

In this research, low degree functions were applied to reduce the number of coefficients required to represent a set of binaural impulse responses. In this paper, 2nd-, 3rd-, 4th-, and 5th- degree functions were used to approximate the responses.

For example, the 4th-degree function

$$f_n ?? ?? a_n ? b_n? ? c_n?^2 ? d_n?^3 ? e_n?^4,$$

where n and ? denote the sampled time and azimuth of the sound source. The solid line plotted using the 4th-degree function approximates the asterisks on the least square error basis. Therefore, as shown in Fig. 7, a set of binaural impulse responses, $h_?(n)$, can be approximated using a set of low degree functions, $f_n(?)$.

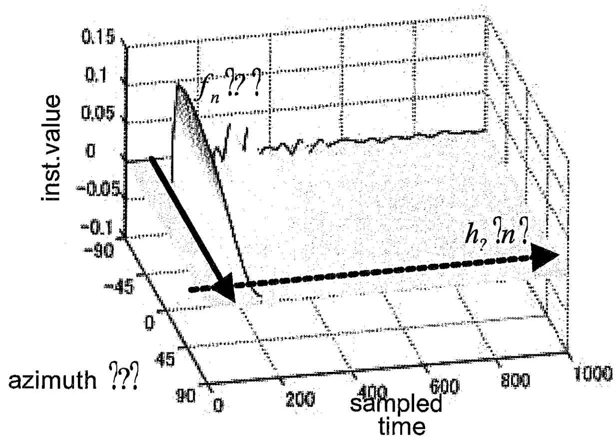


Figure 7 Set of binaural impulse responses and set functions for approximation

The set of binaural impulse responses shown in Fig. 3 were approximated using the set of functions shown in Fig. 8.

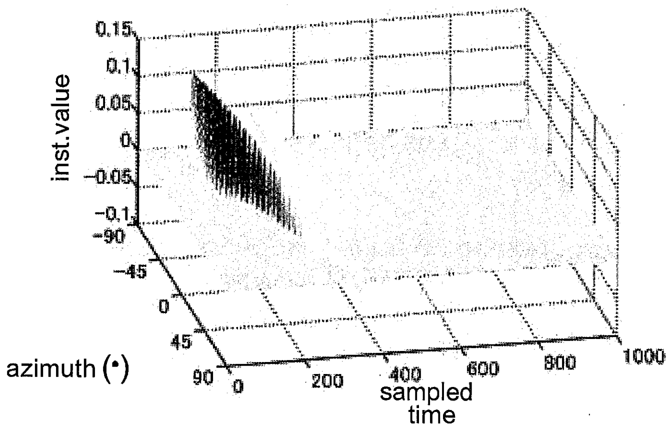


Figure 8 Binaural impulse responses approximated using 4th-degree functions

2.4 EVALUATION OF APPROXIMATION ACCURACY

The approximation accuracy was evaluated using error ratio *ER*, as shown below.

$$ER\ ?\ 10\log_{10}\frac{\underset{?}{?}\ \underset{n}{?}\ ?h_{approx}?n,??? \ h_{meas}?n,???^{2}}{\underset{?}{?}\ \underset{n}{?}\ h_{meas}{}^{2}?n,??},$$

where h_{meas} $(n,?)$ and h_{approx} $(n,?)$ denote the measured impulse response after arrival-time correction and the approximated impulse response. Table 1 shows the error ratio for degree of function.

Table 1 Changes in error ratio due to changes in degree of function

Degree of function	2	3	4	5
Error ratio (dB)	-15.1	-19.7	-21.0	-23.1

Table 1 indicates that 3rd- or 4th-degree functions are sufficient for approximating binaural impulse responses.

3. OBJECTIVE COMPARISON OF MOVING SOUND IMAGES

The binaural signals recorded using a rotating dummy head shown in Fig. 2, synthesized using cross fading (a well-known conventional method) and synthesized using the approximation method were compared.

3.1 SYNTHESIZING BINAURAL IMPULSE RESPONSES

For cross fading, we have shown that cross fading contains a linear interpolation of the binaural impulse responses without considering arrival time differences[5]. We interpolated responses in 1° steps using the measured responses in 5° steps without arrival time correction.

For the approximation method, binaural impulse responses in 1° steps were approximated using 2nd-, 3rd-, 4th-, and 5th- degree functions. Arrival times for the responses were linearly interpolated using the measured arrival times shown in Fig. 3.

Binaural signals for the cross fading and approximation methods were synthesized using time-variant convolutions of nine seconds of pink noise with the responses.

3.2 OBJECTIVE COMPARISON

Figures 9(a) to (f) show spectrograms of the six binaural signals for the left ear. The x-axis denotes the time transition from 0 to 9 sec. It also shows the changes in azimuth of the sound source. The y-axis denotes frequency up to 22.05 kHz, and the graduation shows the frequency amplitude characteristics.

Figure 9 (a) shows the spectrogram for the binaural signal recorded using the dummy head. Since the dummy head was rotated manually, the binaural signal was affected by time fluctuations. Since the sound source moved from the sunny side of the left ear to the shady side during the sound source transition, the frequency levels higher than 8 kHz decreased.

Figure 9 (b) shows the spectrogram for the binaural signal simulated using cross fading. As compared to the spectrogram in Fig. 9 (a), many vertical pin stripes are seen at regular intervals. Such stripes are perceived as ripples of the signal.

Figures 9 (c) through (f) show the spectrograms for the binaural signals simulated using the approximation method. The changes in the frequency amplitude characteristics higher than 8 kHz, shown in Fig. 9 (e) for 4^{th}- degree function, due to time transitions is the closest to that shown in Fig. 9 (a).

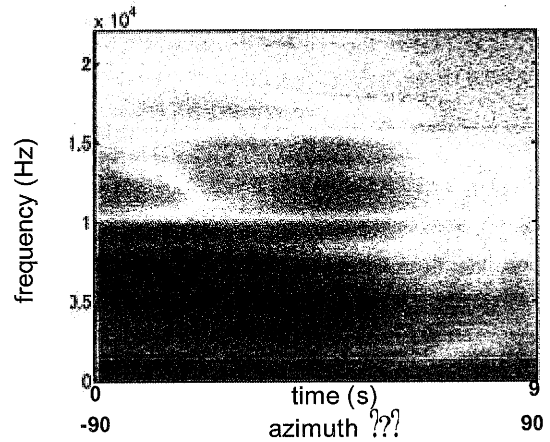
4. SUMMARY

We have proposed a method for approximating binaural impulse responses for moving sound images using low degree functions. Global changes in the binaural impulse responses due to sound source movement were approximated using a set of low degree functions. Approximation accuracy was evaluated using the error ratio. The error ratios were -15.1, -19.7, -21.0, and -23.1 dB for the 2^{nd}-, 3^{rd}-, 4^{th}-, and 5^{th}- degree functions, respectively. The set of approximated binaural impulse responses was similar to that of the measured responses.

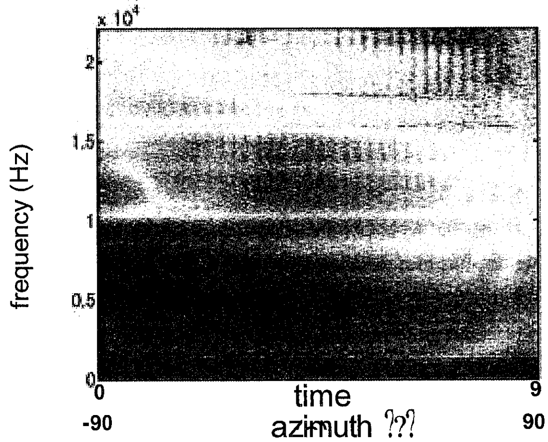
Furthermore, six binaural signals were compared and the binaural signal simulated using the approximation method of 4^{th}- degree function was the closest to that generated using the dummy head.

5. ACKNOWLEDGEMENT

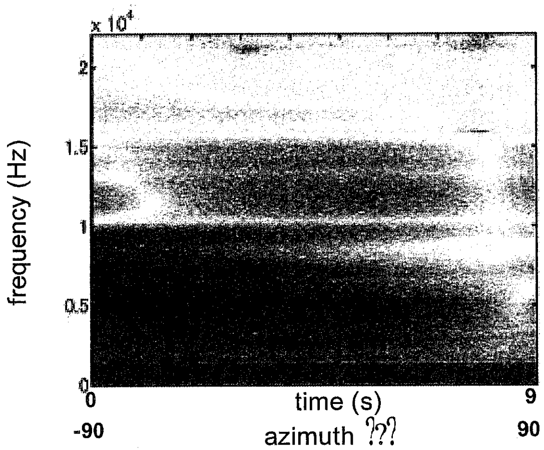
We thank Prof. M. Tohyama for his comments on this research work. We are grateful to Mr. H. Hotta and Mr. T. Inoue for their assistance with the Matlab programming and signal processing. This research work has been partly supported by KAKENHI (14380084).



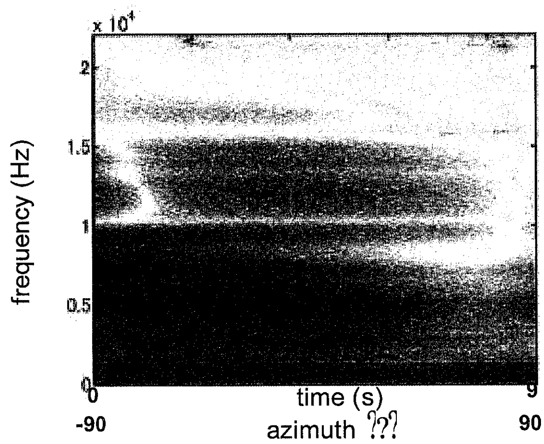
(a) Recorded using rotating dummy head



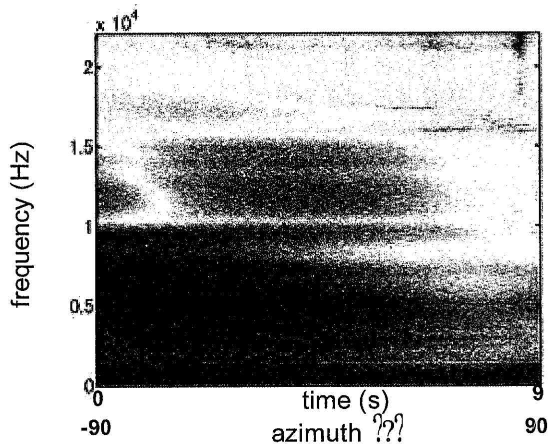
(b) Simulated using cross fading



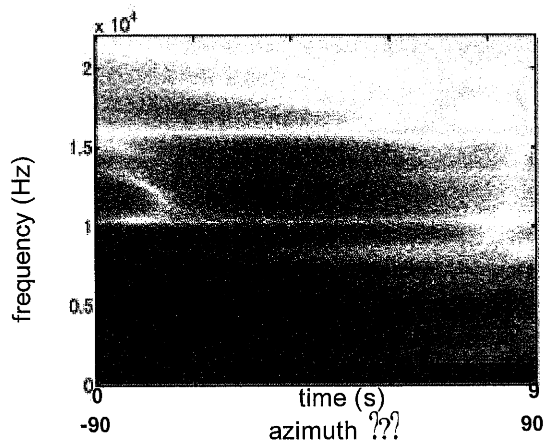
(c) Simulated using approximation method
2nd –degree function



(d) Simulated using approximation method
3rd –degree function



(e) Simulated using approximation method
4th –degree function



(f) Simulated using approximation method
5th –degree function

Figure 9 spectrograms for 6 binaural signals

6. REFERENCES

1 L. Savioja, J. Huopaniemi, T. Lokki and R. Vaananen, "Creating Interactive Virtual Environments," J. Audio Eng. Soc., vol. 47, pp. 675-705 (1999 September)

2 Y. Takahashi, M. Matsumoto and M. Tohyama, "An Acoustic Events Modeling Language (AEML), " Forum Acousticum Seville 2002, ELE-03-002-IP (2002)

3 J. Jot, V. Lacher and O. Warusfel, "Digital signal processing issues in the context of binaural and transaural stereophony," Audio Eng. Soc. 98th Conv., preprint #3980, 1995 February

4 D. J. Kistler and F. L. Wightman, "A model of head-related transfer function based on principal components analyses and minimum phase reconstruction," J. Acoust. Soc. Am. Vol. 91 (3), pp. 1637-1647 (1992 March)

5 M. Matsumoto, M. Tohyama and H. Yanagawa, "A method of interpolating binaural impulse responses for moving sound images," Acoust. Sci. & Tech. 24, 5 (2003)

6 M. Matsumoto, S. Yamanaka, M. Tohyama and H. Nomura, "Effect of arrival time correction on the accuracy of binaural impulse response," J. Audio Eng. Soc., vol. 52, pp. 56-61 (2004 January/February)