Proceedings of ITC-CSCC 2000, Pusan, Korea

3D Acoustic Image Localization Algorithm by Embedded DSP

Wataru Kobayashi†, Noriaki Sakamoto††, Takao Onoye†††, Isao Shirakawa††

†Arnis Sound Technologies, Co. Ltd.
2-7-9 Kita-Senzoku, Ota, Tokyo, 145-0062 Japan
Phone: +81-3-5749-3621, Fax: +81-3-5749-3625, E-mail: kobachan@arns.com

††Dept. Information Systems Engineering, Osaka University
2-1 Yamada-Oka, Suita, Osaka, 565-0871 Japan
Phone: +81-6-6879-7807, Fax: +81-6-6875-5902, E-mail: {sakamoto, sirakawa}@ise.eng.osaka-u.ac.jp

†††Dept. Communications & Computer Engineering, Kyoto University
Yoshida-Honmachi, Sakyo, Kyoto, 606-8501 Japan
Phone: +81-75-753-4803, Fax: +81-75-753-4804, E-mail: onoye@kuee.kyoto-u.ac.jp

Abstract: This paper describes a real-time 3D sound localization algorithm to be implemented with the use of a low power embedded DSP. This algorithm first divides the audible frequency band into three, on the basis of the analysis of the sound reflection and diffraction effects through different media from a certain sound source to human ears, and then in each subband a specific procedure is devised for the 3D sound localization so as to operate real-time on a low power embedded DSP. This algorithm aims at providing a listener with the 3D sound effects through a headphone at low cost and low power consumption.

1. Introduction

In recent years, in the field of acoustic signal processing a number of approaches have been attempted to realizing 3D sound effects [1,2], which are based on the so-called Head-Related Transfer Function (HRTF) [3].

The frequency response characteristic of an HRTF can be measured by recording sound sources through dummy head microphones at left and right ears, as illustrated in Fig.1. As can be seen from Fig. 2, an HRTF characteristic measured by such binaural recording is so complicated in terms of peaks and dips that, in order to realize the accurate characteristic, we have to use a considerable number of digital filters with great flexibility in setting parameters of frequency, gain, and quality factor (Q).

To cope with this difficulty, this paper describes a novel real-time 3D sound localization algorithm, which attains low computational complexity by extracting primary factors necessary for perceiving the 3D localization through the human auditory sense.

A distinctive feature of this algorithm is that first the audible frequency band is divided into three, according to the sound reflection and diffraction effects through different media from a certain sound source to human ears, and then in each of the three subbands a specific procedure of the 3D sound localization is devised so as to operate real-time on a low power DSP.

2. Conventional 3D Sound Localization

A conventional approach to the 3D sound localization is summarized as follows [3,4]: First, calculate the HRTFs necessary for the 3D sound localization on the outside of the human head in the auditory senses, which can be attained with the use of two equations, one representing the signal output from the dummy head microphones for the sound transferred from a given sound source and the other expressing the signal output from the stereo-headphone (see Fig. 1). Then, a monaural input signal is processed through the use of these HRTFs in real time, and then the results are superposed to be reproduced. In other words, the 3D sound localization can be realized by the stereo sound which is obtained by processing the sound source of no localization through the use of a number of HRTFs.

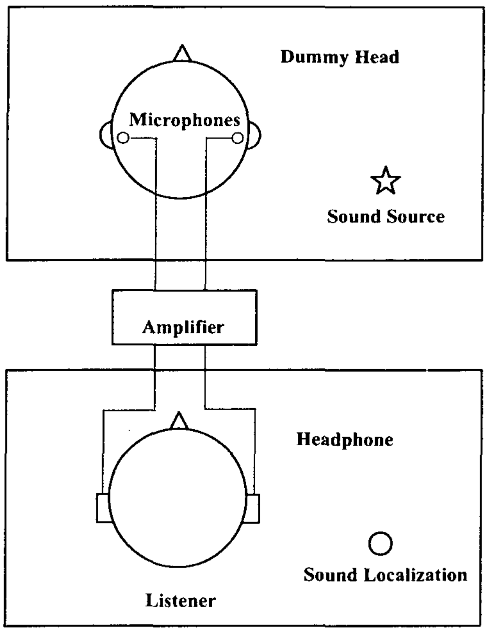


Fig. 1. Binaural recording and reproducing.

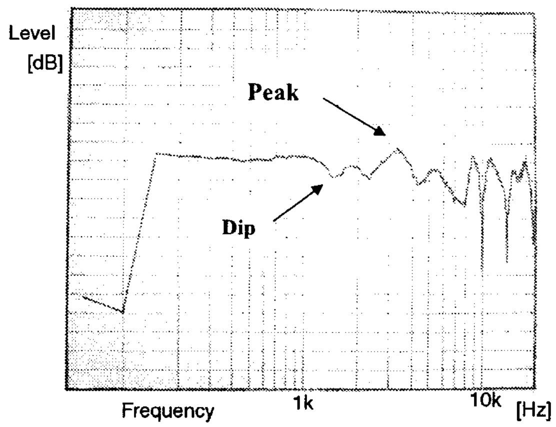


Fig. 2. An example of HRTF frequency characteristics.

Usually the frequency response characteristic of an HRTF is so complicated in terms of peaks and dips that a great number of digital filters are necessary to realize the characteristic accurately, and moreover, in the case when a parametric equalizer (PEQ) is employed for such a digital filter, all the parameters of frequency, gain, and quality factor (Q) must be taken into consideration for each filter.

Furthermore, frequency response characteristics at both ears for any localized sound source are different each other, and hence the structures of filters to realize these characteristics at both ears should be different.

Therefore, to implement the 3D sound localization a great amount of operations for numbers of parameters are necessary for the whole set of filters.

Accordingly, in order to realize the real-time 3D sound localization in line with the conventional 3D sound localization, there remains much room to revise the algorithm so as to reduce drastically the computational complexity.

3. A New Real-time Algorithm

To construct a new real-time algorithm of the 3D sound localization, we first consider the primary factors necessary for reducing the computational complexity.

To this end, we first investigate the phenomena of the sound reflection and diffraction through different media from a certain sound source to human ears.

A: Low Frequency Subband

First, consider the sound diffraction by a human head (see Fig, 3). Now, assume that the head is of spherical shape with diameter of 150~200mm, although there are small differences among individuals. If the half of a sound wavelength is larger than the head diameter, the effects of sound reflection and diffraction by the head are negligibly small. In this case, given an audio digital input signal, the 3D localization can be achieved only with the use of the difference between sound volumes and that between arrival times of the sound emanating from the sound source and entering into right and left ears.

Assuming that the sound velocity and the head diameter are given by $v = 340$m/s and $d1 = 150\sim200$mm, respectively, the boundary frequency is given by $f1 = v/(d1/2) = 850\sim1{,}100$Hz.

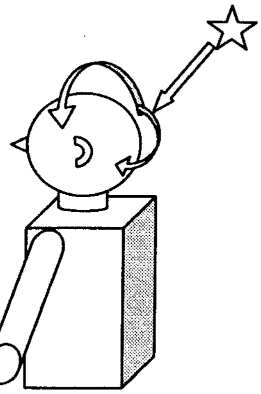


Fig. 3. Sound diffraction by human head.

B: High Frequency Subband

Next, consider the sound diffraction by a human pinna. Suppose that the pinna is a cone with base diameter of 35~55mm. If the half of a sound wavelength is shorter than the base diameter, the pinna has physically a great influence on the sound diffraction.

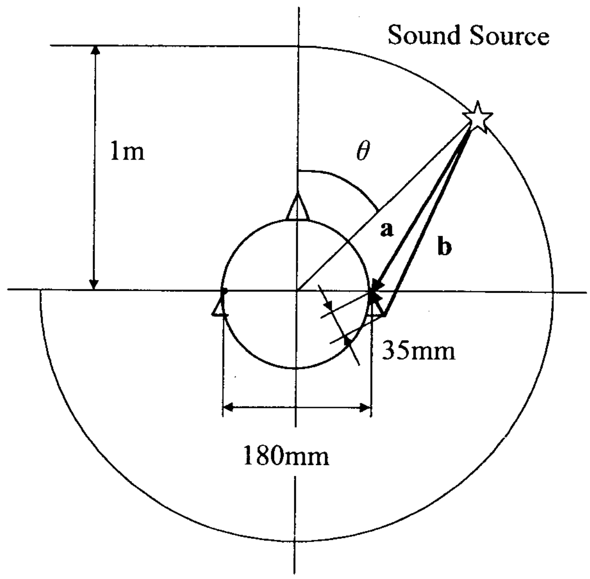


Fig. 4. Paths of major diffraction by pinna.

Assuming that the base diameter of a human pinna is given by $d2 = 35\sim55$mm, the boundary frequency is obtained as $f2 = 3\sim5$kHz.

It is observed through the measurement by dummy head microphones that the frequency response characteristic of any sound in this high frequency subband can be approximated with the use a comb filter. Moreover,, in case a comb filter is adopted, the boundary frequency f2 and the characteristic of a comb filter depend on the location of the sound source. In other words, the characteristics of sounds entering into ears are obtained by superposing those sounds transmitted through different paths from the source to ears, each of which can be approximated by a comb filter.

For example, consider the parameters of angle θ, and path lengths **a** and **b** in Fig. 4, where **a** denotes the length of the straight path and **b** that of the major reflection path, each from the source to the ear. Table 1 shows the relation among parameters of angle θ, difference **b-a** and boundary frequency f2.

TABLE I
CORRELATION BETWEEN DIFFERENCES
b-a AND BOUNDARY FREQUENCY f2

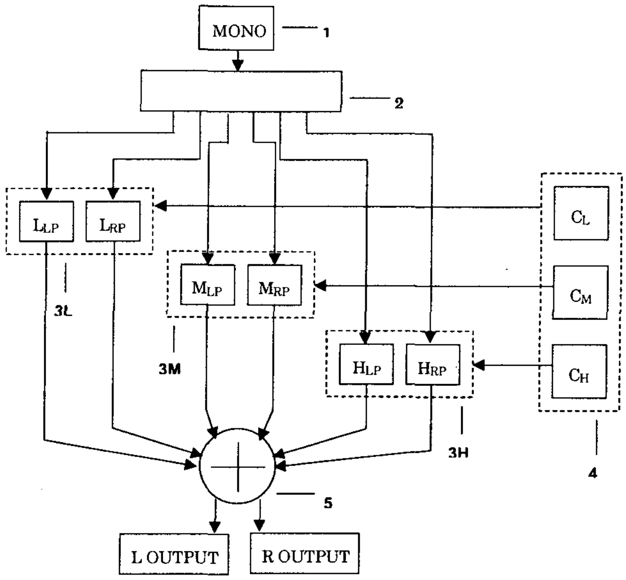
θ[°]	b-a[mm]	f2[Hz]
0	62.1	2737.0
5	60.1	2827.8
10	57.9	2935.1
15	55.5	3061.5
20	53.0	3210.2
25	50.2	3385.0
30	47.3	3591.2
35	44.3	3835.0
40	41.2	4124.7
45	38.0	4471.2
50	34.8	4888.7
55	31.5	5396.3
60	28.2	6020.1
65	25.0	6796.6
70	21.9	7777.8
75	18.8	9039.8
80	15.9	10697.9
85	13.1	12932.9
90	10.6	16042.1

C: Intermediate Frequency Subband

In the intermediate frequency subband between f1 and f2, the 3D sound localization is to be performed in the same way as in the conventional method, that is, it can be realized by a parametric equalizer (PEQ) by taking account of the sound diffractions by both the head and the pinna.

A scheme for the 3D sound localization in each subband is distinct, as described in A, B, and C. Since the frequency band of audio signals is divided into three so that a specific procedure for the 3D sound localization can be devised independently in each subband, this algorithm can greatly reduce the total computational complexity by employing a sophisticated scheme which is fitted to each subband.

An outline of this real-time algorithm of the 3D sound localization is shown in Fig.5.



1 MONO: monaural source
2 Filters dividing frequency subbands
3L LLP: a low frequency subband processing for a left ear
LRP: a low frequency subband processing for a right ear
3M MLP: an intermediate frequency subband processing for a left ear
MRP: an intermediate frequency subband processing for a right ear
3H HLP: a high frequency subband processing for a left ear
HRP: a high frequency subband processing for a right ear
4 CL: parameters for a low frequency subband processing
CM: parameters for an intermediate frequency subband processing
CH: parameters for a high frequency subband processing
5 mix and delay
L OUTPUT: processed sound for a left ear
R OUTPUT: processed sound for a right ear

Fig. 5. Block diagram of proposed real-time algorithm.

5. Conclusion

This paper has described a new real-time 3D sound localization algorithm to be implemented with use of a low power embedded DSP. A distinctive feature of this algorithm is that the audible frequency band is divided into three so that a specific procedure for the 3D sound localization can be exploited properly in each subband. This algorithm has reduced the total power dissipation to such an extent that our new algorithm can provide a listener with the facility of real-time 3D sound localization through a headphone at low cost and low power consumption. Development is continuing further on sophisticated low-power VLSI implementation.

References

[1] M. Okamoto, I. Kinoshita, S. Aoki, and H. Matsui: "Sound image rendering system for headphones," *IEEE Trans. Consumer Electronics*, vol. 43, pp. 689-693, Aug. 1997.

[2] An-Nan Suen, Jhing-Fa Wang, Jia-Ching Wang "VLSI Implementation of 3-D sound generator, " *ibid*, vol. 43, pp. 679-688, Aug. 1999.
[3] E. M. Wenzel, M. Arruda, D. J. Kistler, and F. L. Wightman: "Localization using nonindividualized head-related transfer functions," *J. Acoust. Soc. Amer.*, vol. 94, pp. 111-123, July 1993.
[4] John Sunier, "EARS WHERE THE MIKES ARE," *Audio* , pp.49-56, Nov. 1989.