nederlands akoestisch genootschap
NAG
journaal
nr. 104
september 1990

SIGNAL PROCESSING FOR BINAURAL APPLICATIONS: ADAPTING LOUDSPEAKER SIGNALS FOR HEADPHONE REPRODUCTION

Ing. P.M. Boers and Ing. R.M. Aarts
Philips Research Laboratories
P.O. Box 80 000, 5600 JA Eindhoven

SIGNAL PROCESSING FOR BINAURAL APPLICATIONS: ADAPTING LOUDSPEAKER SIGNALS FOR HEADPHONE REPRODUCTION

Summary: Reproduction of loudspeaker signals via headphones leads to in-head localization. By convolving these signals with the impulse responses measured from loudspeakers to ear canals and deconvolving them with impulse responses measured from headphones to ear canals, binaural signals are generated which simulate a loudspeaker set-up and make better localization possible.
In general the inverse of the system describing an outer ear and a headphone is not stable and a modification is necessary to make implementation in a digital signal processor possible.
In this contribution a new technique is described to modify the measured functions and allow for implementation with minimal errors in the magnitude of the functions.
First the discrete-time impulse response of a measured transfer function is transformed into the z-domain. All zeros are then determined. The positions of the zeros outside the unit circle show deviations from a minimum phase system, indicating the frequency regions where errors in the phase characteristic are introduced when the function is inverted and made stable. It will be shown that when stabilizing the quotient of loudspeaker to ear and headphone to ear functions some of these errors may be avoided. Advantages of analysis and processing using z-transform are discussed.

1 INTRODUCTION

Binaural signals contain directional information and are intended to be reproduced in left and right ears to create a natural hearing event. Using an artificial head directional information of sound sources is included in the recorded signals.

However, when headphones are used to reproduce signals intended for a standard stereo loudspeaker set-up, then directional information is not transferred and as a consequence all sound sources are localized inside the listener's head. In much the same way as in a binaural mixing console [1] binaural correction functions containing directional information and compensating for the headphone to ear transfer may be generated electronically and used to improve headphone sound reproduction.

This paper describes a method to design the binaural correction functions with minimal errors in the magnitude characteristics and which can be implemented in a digital signal processor. We assume that our hearing is more sensitive to magnitude errors of the binaural correction functions than to errors in the phase. For this reason methods which trade magnitude errors against phase errors are not considered in this paper.

First a transfer function is measured from a sound source to a point in the ear canal (for instance $H_{Lr}(\omega)$ from left loudspeaker to right ear). This function contains the directional information of one sound source. A second transfer function $H_{Hr}(\omega)$ is measured from the headphone to the same point in the ear canal, the inverse of which must be used to compensate for the headphone when the binaural signals are reproduced. The application of directional information $H_{Lr}(\omega)$ and compensation for the headphone ($H_{Hr}^{-1}(\omega)$) is done in one step using a binaural correction function $H'_{Lr}(\omega) = H_{Lr}(\omega)\, H_{Hr}^{-1}(\omega)$. Four of these binaural correction functions are used to add directional information to left and right stereo loudspeaker signals L and R as shown in figure 1. Because of the similarity in the calculation of the four binaural correction functions only the processing of one of them ($H'_{Lr}(\omega)$) will further be discussed.

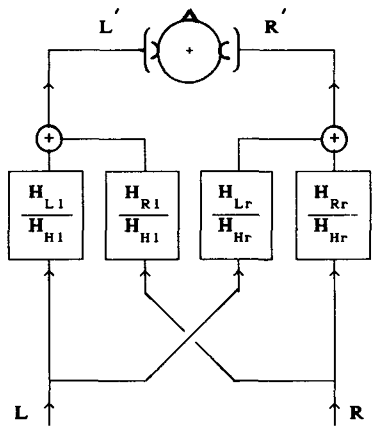

Figure 1 Scheme to change left (L) and right (R) loudspeaker signals to binaural signals L' and R' using four binaural correction functions. The transfer functions are denoted by a capital subscript for the sound source (L loudspeaker, H headphone) and a second subscript for the receiving ear (l left, r right). The binaural correction functions each consist of a quotient of two transfer functions. The numerator contains directional information of one loudspeaker to one ear and the denominator compensates for the headphone to ear transfer.

When reproducing the processed signals L' and R' via headphones the direct sound from a loudspeaker set-up in an anechoic environment is simulated. Therefore this simulation creates no impression of spaciousness, unless another processing step is performed where the reflection pattern of a listening room is added. This step will not be discussed in this paper.

In section 2 the measurements of the outer ear transfer functions of loudspeakers and headphones to ear canal are described. In section 3 a new method to determine the binaural correction functions in the z-domain is described. A discussion follows in section 4.

2 TRANSFER FUNCTIONS FOR LOUDSPEAKER SIMULATION

In a standard stereo set-up two loudspeakers are positioned in +30° and –30° directions at 2.5 m distance from the listener. The transfer functions from loudspeakers to ears can be measured in an anechoic room by positioning a subject and a loudspeaker in a corresponding position. A probe microphone is placed at a reference point in the ear canal to measure the signals radiated by the loudspeaker [2]. We thus obtain the function

$$H_{Lr}(\omega)\, M(\omega) \quad , \tag{1}$$

where $H_{Lr}(\omega)$ is the transfer function from the left loudspeaker to the right ear canal and $M(\omega)$ is the acoustic disturbance due to the microphone. Likewise, using the headphones as a source a measurement is performed to the same reference point in the ear canal

$$H_{Hr}(\omega)\, M'(\omega) \quad . \tag{2}$$

Here $H_{Hr}(\omega)$ is the transfer function from headphones to ear canal. $M'(\omega)$ is the disturbance by the microphone if the headphone is present. If a source signal $S(\omega)$ is reproduced using the binaural correction function then the signal at the reference point is equal to

$$S'(\omega) = S(\omega)\, H_{Hr}(\omega) \left(\frac{H_{Lr}(\omega)\, M(\omega)}{H_{Hr}(\omega)\, M'(\omega)} \right) \quad . \tag{3}$$

We call the bracketed term in (3) the binaural correction function. Schröter et al. proved [3] that if:

- the microphone is small enough not to change the acoustic input impedance of the ear canal, or
- the headphones have an open structure and do not change the radiation impedance of the ear canal,

then $M'(\omega) = M(\omega)$. Thus for the two loudspeaker directions +30° and –30° the transfer functions are measured at both ears and from these measurements and the two headphone to ear transfer functions four binaural correction functions are calculated.

3 CALCULATION OF THE CORRECTION FUNCTIONS

A correction function is determined by multiplication of the complex functions $H_{Lr}(\omega)$ and $H_{Hr}^{-1}(\omega)$. The convolution of the source signal with the impulse response of the correction function is carried out in a finite impulse response filter [4, 5]. From a number of binaural correction functions determined for various ears and headphones it appears that due to the limited number of coefficients in a practical implementation a considerable truncation error has to be made.

One reason is that the inversion of $H_{Hr}(\omega)$ increases the length of its impulse response from the length determined by the measurement of the function to infinite.

The second reason is that if $H_{Hr}(\omega)$ does not describe a minimum phase then the inverted function $H_{Hr}^{-1}(\omega)$ belongs to an unstable system. The truncation errors have a disturbing influence on the frequency response. We found that, if the function to be inverted is made minimum phase avoidable errors may be introduced in the correction function, as will be shown below.

Unlike the conventional methods, which perform processing in the time domain or in the frequency domain, we prefer proceeding in the z-domain. The z-transform is a useful technique for representing and manipulating discrete-time sequences. Given an impulse response h(k), which is zero outside the interval $0 \le k \le N-1$, its z-transform [4, 5] is

$$H(z) = \sum_{k=0}^{N-1} h(k)\, z^{-k}$$

$$= h(0) \prod_{k=1}^{N-1} \left(1 - z_k\, z^{-1}\right) \quad . \tag{4}$$

In this way a system with an impulse response of length N is completely described by the N–1 zeros z_k. The system described has a real impulse response and consequently zeros appear in complex conjugate pairs or are real. Therefore their position in the z-plane has a mirror symmetry with respect to the real z-axis. From the distribution of zeros in the z-plane we can make some observations.

- If a zero is located outside the unit circle the function is not minimum phase and after inversion the function becomes unstable,
- if in $H_{Lr}(\omega)$ and $H_{Hr}(\omega)$ similar resonances exist with coinciding zeros outside the unit circle these zeros will cancel in the quotient in (3) and will not give rise to unstability,
- resonant dips with a high quality factor which may appear due to a special headphone construction in combination with the outer ear can be recognized from a zero close to the unit circle.

The correction function can now be constructed in a few steps. First we invert $H_{Hr}(\omega)$ by replacing all zeros by poles, which are then combined with the zeros of $H_{Lr}(\omega)$. Figure 2 shows this configuration in the z-plane.

If any poles are situated outside the unit circle, as is the case shown in figure 2, then additional steps are necessary. The function must be modified to obtain a stable system, the impulse response of which must not be longer than the available finite impulse

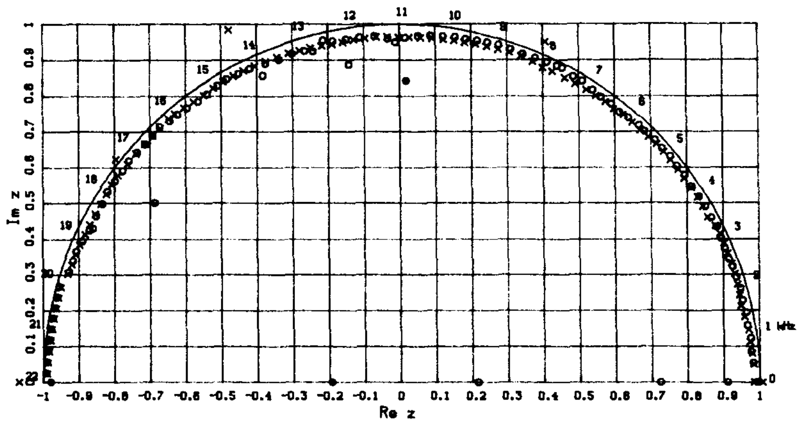

Figure 2 Complex z-plane with 200 zeros (o and •) of the transfer function from a Quad ESL–63 loudspeaker in direction of +30° to the nearest ear of the artificial head KU81i. Real frequencies from 0 to half the sampling frequency are plotted along the unit circle. The filled marks show zeros outside the unit cycle but drawn mirrored with respect to the unit circle. The inverse of the transfer function from a Stax–λ headphone to the same ear is described with 200 poles (x). This function is unstable due to poles outside the unit circle near frequencies of 0, 8.2, 14.2, 17.5 and 22 kHz.

response filter. If the magnitude function is to be left unchanged the only displacement of zeros and poles allowed is a mirroring with respect to the unit circle. Thus the following options exist

1. position all zeros and poles inside the unit circle to make the correction function minimum phase and stable,
2. position all poles inside the unit circle to make the correction function stable,
3. position only the poles that do not coincide with zeros inside the unit circle to make the correction function stable.

The first method changes the phase characteristic but not the magnitude. A disadvantage is that also all zeros are mirrored introducing unnecessary errors in the phase characteristic. However, this method results in a correction function with an impulse response of minimal length. If the system described by all zeros has an impulse response longer than the available finite impulse response filter, then this method limits the errors in the magnitude of the correction function, which would be caused by truncation during the implementation.

The second method gives better results by preventing errors in the phase characteristic due to displacement of zeros as mentioned above.

The third method shows an advantage of the description with zeros and poles. Now the correction function is made stable by mirroring with respect to the unit circle only those poles which lie outside the unit circle and are not compensated by zeros. All superfluous errors in the phase characteristic are avoided and the magnitude remains unchanged. Stabilizing the correction function this way therefore is preferable.

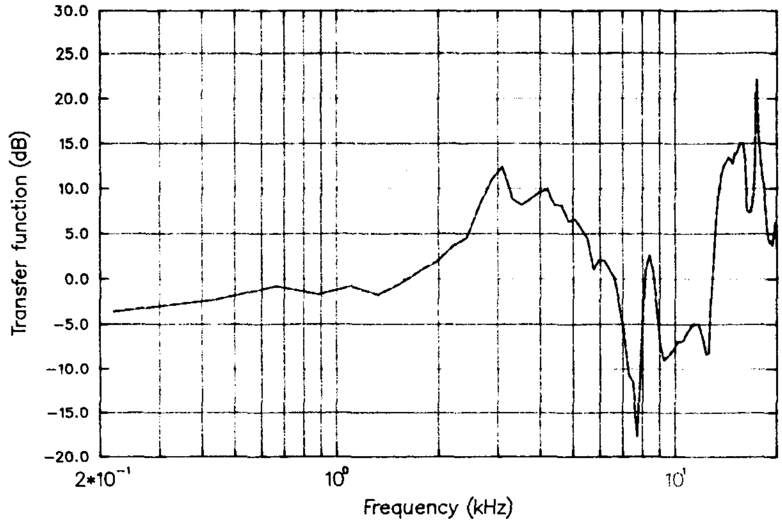

Figure 3a The magnitude of the unstable function $H_{R_r}(\omega)\ H_{Hr}^{-1}(\omega)$ as obtained from the same measurements as figure 2 (___) and of the processed version to be implemented in a finite impulse response filter of 200 taps (---). The negligible difference between the two curves is due to stabilizing the polynomial.

Figure 3 shows the complex quotient of a loudspeaker and a headphone measurement, prior to any processing (solid lines) and the correction function obtained using the third method (dashed lines). The magnitude characteristics which are equal before and after processing are shown in figure 3a; as the figure shows the two curves are hardly distinguishable. Figure 3b shows the errors appearing in the group delay characteristic due to the mirroring of some of the poles. The group delay is proportional to the derivative of phase with respect to frequency and shows processing errors more clearly. When the pole at $z = -1.07$ is mirrored with respect to the unit circle we see that its influence does not extend into the audible frequency region. Three poles at 8.2, 14.4 and 17.5 kHz are mirrored with respect to the unit circle to make the correction function stable. The difference between the two curves is due to stabilizing the polynomial. The

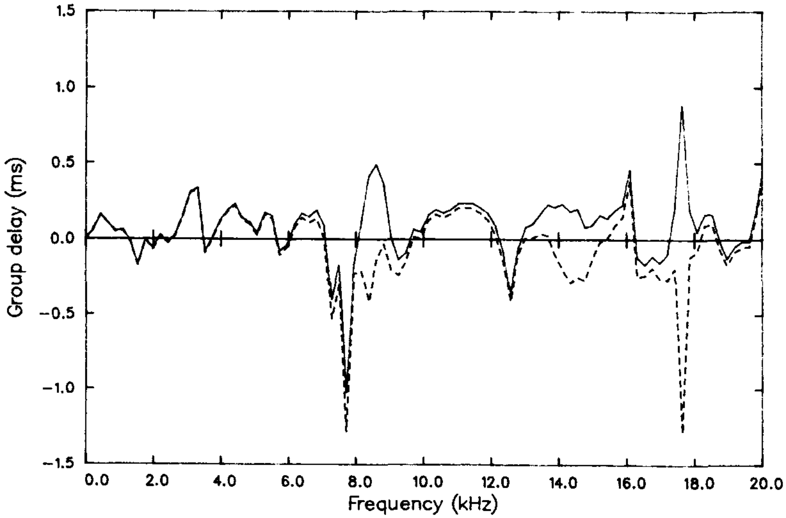

Figure 3b The group delay of the same functions as shown in figure 3a. The difference between the two curves is due to stabilizing the polynomial. This can be seen in the frequency regions where a pole has been mirrored to make the function stable (8.2, 14.2 and 17.5 kHz). The pole outside the unit circle at $z = 1.02$ (fig. 2) was not mirrored to prevent the introduction of phase errors in the low frequency region.

influence of the mirroring of the poles can be seen at the mentioned frequencies. One pole at $z = 1.02$ is compensated by a zero at the same position and not mirrored. Thus neither in the phase nor in the magnitude characteristic an error at low frequencies is introduced. Another advantage of the zeros and poles description appears when we have to deal with resonances of high quality factor. These give rise to sharp dips or peaks in the frequency characteristic and thus must have a zero or pole close to the unit circle. If some damping of this resonance is desired, this can be realized easily by moving the corresponding zero or pole away from the circle.

Another problem sometimes arises when the loudspeakers reproduce low frequencies better than the headphones. In this case the correction function will rise towards low frequencies to boost the headphone sound reproduction. This may result in an unacceptable distortion. Moving a zero or pole in this frequency region allows us to shape the low frequency characteristic and thus decrease the boosting.
After the processing in the z-domain we obtain the correction function from the zeros and poles by substituting in the system function [5]:

$$H(z) = \frac{\prod_{k=1}^{N-1}\left(1 - z_k z^{-1}\right)}{\prod_{k=1}^{M-1}\left(1 - p_k z^{-1}\right)} , \tag{5}$$

for z_k and p_k the calculated N zeros and M poles and $z = e^{j\theta}$. Here θ is the real angular frequency normalized to the sampling frequency. Using a discrete Fourier transformation we then find the impulse response which can be used in a finite impulse response filter to perform the convolutions.

4 DISCUSSION

A method is described based on the z-transform to analyze and manipulate functions consisting of the quotient of two transfer functions the result of which must be implemented in a finite impulse response filter. Following this method a binaural correction function is found which can be used to simulate a loudspeaker set-up when a normal stereo signal is reproduced via headphones. Another application could be the equalization of an artificial head for a sound source in a specific direction or for a specific headphone.

We assume that our hearing is more sensitive to errors in the magnitude of the binaural correction function than to errors in the phase characteristic. For this reason we describe a method resulting in minimal errors in the magnitude characteristic and trading of magnitude errors against errors in the phase are not considered.

Using the z-transform a function can be described with zeros only. One zero finding algorithm used is the method of Bairstow [6] in combination with a high precision number representation.

A description using zeros is useful to gain insight in functions to be processed. The configuration of zeros in the z-plane shows in what frequency regions the function is not minimum phase and where errors in the phase characteristic are introduced if the function has to be inverted and made stable. Inversion of a function is easily performed by replacing all zeros by poles and vv. When poles exist outside the unit circle then a function is made stable by mirroring all poles with respect to the unit circle.

When an implementation in a finite impulse response filter is desired then truncation errors will occur. It was shown that poles in the inverted function which are outside the circle and coincide with zeros should not be mirrored when the function is made stable. This will lead to minimal errors in the magnitude characteristic and only limited errors in the phase characteristic.

LITERATURE

1. Blauert, J., 1983. *Spatial Hearing*, MIT Press, Cambridge, Massachusetts, p. 366.
2. Mehrgardt, S. and Mellert, V. 1977. Transformation characteristics of the external human ear, J. Acoust. Soc. Am., **61**, pp. 157–167.
3. Schröter, J., Spikofski, G. und Theile, G., 1986. Messung des Diffusfeldübertragungsmaßes von Kopfhörern am Kunstkopf, Acustica, **60**, pp. 105–116.
4. Rabiner, L.R. and Gold, B. 1975, *Theory and application of digital signal processing*, Prentice-Hall, New Jersey, p. 28.
5. van den Enden, A.W.M. and Verhoeckx, N.A.M. 1989. *Discrete-time signal processing an introduction*, Prentice Hall, New Jersey, p. 96.
6. Fröberg, C., 1965. *Introduction to Numerical Analysis*, Addison-Wesley Publishing Company, Reading-Massachusetts, p. 28.