NAM 96 Nordic Acoustical Meeting
Helsinki, 12 – 14 June 1996 E6

HRTF FILTER DESIGN BASED ON AUDITORY CRITERIA

Jyri Huopaniemi and Matti Karjalainen

Helsinki University of Technology
Laboratory of Acoustics and Audio Signal Processing
Otakaari 5 A, FIN-02150 Espoo, Finland
Jyri.Huopaniemi@hut.fi, Matti.Karjalainen@hut.fi

1 INTRODUCTION

Auralization and binaural technology is a strongly growing field in acoustics and audio signal processing [1]. Measurements and models of head-related transfer functions (HRTF) of human subjects or dummy heads are the source of information for the research of auralization and spatial hearing (see, e.g., [2], for a good review on these subjects). The information contained in HRTFs measured at various azimuth and elevation angles is sufficient for synthesizing realistic three-dimensional sound events for headphone or loudspeaker listening. One of the problems in binaural synthesis is the computational load of accurate HRTF calculation. To overcome this problem, both computationally efficient and perceptually relevant digital models of HRTFs have to be created.

In this paper, new methods for HRTF filter design using auditory criteria are presented. New classes of digital filters, namely warped FIR (WFIR) and IIR (WIIR) structures and their application to binaural processing are studied. The bilinear transform is used to warp the frequency scale to a nonuniform resolution, e.g., to match the critical bands of the human ear. Examples of warped filter design are presented. The implementation of WFIR and WIIR filters in real-time binaural applications is discussed. Results of the comparison between warped and traditional HRTF filter design methods are given. The paper is organized as follows. In Section 2, an overview of the properties and modeling strategies of HRTFs is given. Frequency warping and the application of warped filter structures to HRTF modeling are presented in Section 3. Finally, conclusions are drawn and directions for future research are given in Section 4.

2 PROPERTIES AND MODELING OF HRTFS

The HRTF is defined as a free-field transfer function from a point in a space to a point in the listener´s outer ear, normally the entrance to the ear canal [2]. This suggests that the HRTF includes all the necessary information for synthesis of static human spatial hearing properties including the main localization cues, the frequency- and direction-dependent interaural time difference (ITD) and interaural amplitude difference (IAD)

[3]. HRTFs are typically measured on human subjects or dummy heads using probe or miniature microphones placed at the entrance, or at some point in the subject's ear canal [2].

The sound transmission in an HRTF measurement includes characteristics of many subsystems that are to be compensated in order to achieve the desired response. The transfer functions of the driving loudspeaker, the microphone and the ear canal (if the measurement position was inside an open ear canal) may thus have to be equalized. If, however, a more general database of HRTFs is desired, we should consider other equalization strategies like free-field equalization or diffuse-field equalization [3] [4]. Free-field equalization is achieved by dividing the measured HRTF by a reference measured in the same ear from a certain direction. The reference direction is typically chosen as 0° azimuth and 0° elevation, that is, from the front of the listener. Diffuse-field equalization is derived using the power transfer functions of the measured HRTFs. Diffuse-field HRTFs are estimated by power-averaging all HRTFs from each ear and taking the square root of this average spectrum. Equalized HRTFs are obtained by dividing the original measurement by the diffuse-field HRTF of that ear. This leads to the fact that the factors that are not incident-angle dependent, such as the ear canal resonance, are removed.

In many cases further preprocessing of the measured HRTF data is required before the actual filter design stage. An attractive approach for HRTF smoothing is to apply a variable-size window function to the power frequency response to approximate, for example, the critical-band resolution of the human ear [5] [6] [7]. This smoothing applies only to the magnitude response, so we have to assume that the phase can be calculated by minimum-phase reconstruction. In our filter design investigations we have considered three HRTF specifications to be optimized; 1) raw HRTF impulse responses (measurement system equalized), 2) diffuse-field equalized HRTFs, and 3) HRTFs equalized for a specific headphone type.

2.1 Properties of HRTFs

HRTF impulse responses are the output of a linear and time-invariant system, that is, the diffraction and reflections around the human head, the outer ear, and the torso. Thus the impulse responses can directly be represented as finite-impulse response (FIR) filters. There are often computational constraints that lead to the need of HRTF impulse response approximation. This can be carried out using conventional digital filter design techniques. It is, however, necessary to note that the filter design problem is not a straightforward one. We should be able to design arbitrary-shaped mixed-phase filters that meet the set criteria both in the amplitude and phase response.

An attractive property of HRTFs is that they may be modeled as minimum-phase structures [8] [9]. The excess phase that is the result of subtracting the original phase response from its minimum-phase counterpart has been found to be approximately linear. This suggests that the excess phase can be separately implemented as an allpass filter or a simple delay line. In the case of binaural synthesis, the ITD part of the two

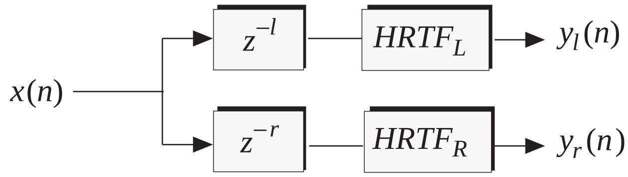

Figure 1. Implementation of HRTFs using minimum-phase HRTF approximation and pure delays for realization of ITD.

HRTFs may be modeled as a separate delay line, and minimum-phase HRTFs may be used for synthesis (see Fig. 1). According to Kistler and Wightman [9], minimum-phase reconstruction does not have any perceptual consequences. Furthermore, with this method the designed FIR filters are of the shortest possible length for that magnitude response (property of minimum-phase filters).

2.2 Digital filter approximation of HRTFs

The most straightforward way to approximate HRTF measurements is to use the windowing FIR filter design method. A filter of the desired order is obtained by windowing the measured impulse response with, e.g., a rectangular window. The use of a rectangular window may be motivated because it is the closest approximation to the original frequency response in the least-square sense [10].

Some papers on IIR design methods for HRTFs have been published in recent years. A comparison of FIR and IIR filter design methods was presented by Sandvad and Hammershøi [10]. The FIR filters were designed using rectangular windowing. The IIR filters were generated using a modified Yule-Walker algorithm that performs least-squares magnitude response error minimization. The low-order fit was enhanced a posteriori by applying a weighting function and discarding selected pole-zero pairs at high frequencies. Listening tests showed that an FIR of order 72 equivalent to a 1.5 ms impulse response was capable of retaining all of the desired localization information, whereas an IIR filter of order 48 was needed to achieve the same localization accuracy. Blommer and Wakefield [11] have proposed an IIR design method based on minimizing the squared log-magnitude difference. New approximations for the log-magnitude error minimization were also defined.

It is known that the human ear processes the auditory information according to a nonuniform critical band frequency resolution [12]. This suggests that modeling of HRTFs should also be carried out in the same manner. There are two possible approaches to approximate a non-linear frequency resolution. One possibility is to use weighting functions that allow more error at higher frequencies and demand a better fit at lower frequencies. The other possibility is to use a non-linear frequency resolution in the filter design. This is often referred to as *frequency warping*.

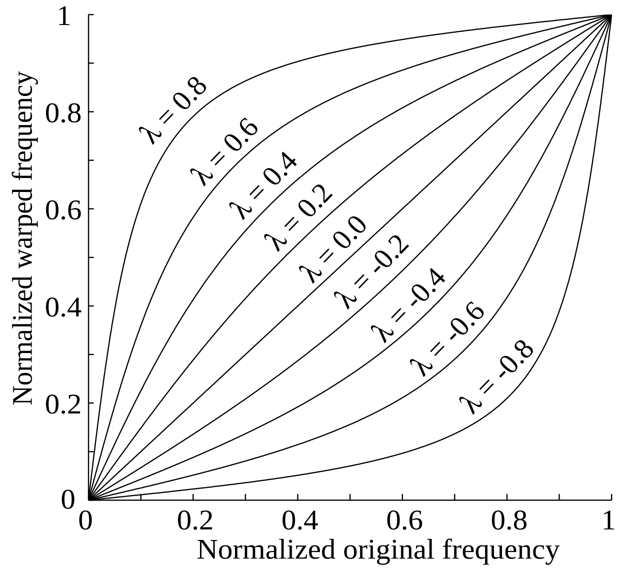

Figure 2. Frequency warping characteristics of the first-order allpass section $D_1(z)$ for different values of λ. Frequencies are normalized to the Nyquist rate.

3 WARPED HRTF FILTER DESIGN

Approximations of HRTFs using auditory criteria have not been extensively studied. Jot *et al*. [5] have proposed a method where the HRTFs are preprocessed using auditory smoothing and the IIR filter design using a standard Yulewalk algorithm is carried out in the warped frequency domain. In the following, we will present theory for frequency warping, and various filter design aspects.

3.1 Frequency warping

Some filter design and model estimation methods allow for an error weighting function in order to control the varying importance of different frequencies (e.g., [10], applied to HRTFs). Here we take another approach, the frequency scale warping, that is in principle applicable to any design or estimation technique. The most popular warping method is to use the bilinear conformal mapping. The bilinear warping is realized by substituting unit delays with first-order allpass sections

$$z^{-1} \Leftarrow D_1(z) = \frac{z^{-1} - \lambda}{1 - \lambda z^{-1}} \tag{1}$$

This means that the frequency-warped version of a filter can be implemented by such a simple replacement technique. It is easy to show that the inverse warping can be achieved with a similar substitution but using $-\lambda$ instead of λ. Figure 2 illustrates the frequency warping properties of Eq. 1 for different values of parameter λ. Frequency axis warping is easier to comprehend by noticing that an allpass filter $D_1(z)$ is a frequency-dependent delay, i.e. sinusoidals of different frequencies pass it in different speeds. The magnitude response $|D_1(z)| \equiv 1$.

The usefulness of frequency warping in our case comes from the fact that, given a target transfer function $H(z)$, we may find a lower order warped filter $H_w(D_1(z))$ that is a good approximation of $H(z)$. $H_w(D_1(z))$ should be designed in a warped frequency domain so that using allpass delays $D_1(z)$ instead of unit delays maps the design to a desired transfer function in the ordinary frequency domain. For an appropriate value of λ, the bilinear warping can fit the psychoacoustic Bark scale, based on the critical band concept [12], relatively accurately. An approximate formula for the optimum value of λ as a function of sampling rate is given in [13]. For a sampling rate of 44.1 kHz this yields $\lambda = 0.7233$ and for 22 kHz $\lambda = 0.6288$.

3.2 Warped FIR structures

The principle of a warped FIR filter (WFIR) may be interpreted as an FIR where the unit delays have been replaced by first-order allpass filters:

$$B_w(z) = \sum_{i=0}^{M} \beta_i \left[D_1(z)\right]^i \tag{3}$$

A more detailed filter structure for implementation is depicted on the left side of Fig. 3 [14]. It can be seen that a warped FIR is actually recursive, i.e., an IIR filter with M poles at $z = \lambda$, where M is the order of the filter. A straightforward method to get the tap coefficients β_i for a WFIR filter is to warp the original HRTF impulse response and to truncate by windowing the portion that has the amplitude above a threshold of interest.

3.3 Warped IIR structures

Having the warped HRTF response at hand, it is a straightforward operation to apply, for example, Prony's method to yield a warped pole-zero model of the form

$$H_w(z) = \frac{\sum_{i=0}^{M} \beta_i \left[D_1(z)\right]^i}{1 + \sum_{i=1}^{R} \alpha_i \left[D_1(z)\right]^i} \tag{4}$$

This cannot be implemented directly due to delay-free recursive propagation through $D_1(z)$ units [14]. By proper mapping of α_i coefficients to new σ_i coefficients the warped IIR filter can be implemented as shown in Fig. 3b [15].

3.4 Filter design examples

The application of warped filter design methods to HRTF modeling was tested using measurements carried out on a dummy head (Brüel&Kjaer 4100). HRTFs were measured for both ears at 5° intervals in an anechoic chamber using the DRA Laboratories' MLSSA system. Additional measurements were carried out to compensate for the measurement equipment and to obtain headphone correction filters (four headphone types were used). The internal sampling rate of the MLSSA system was 75.47 kHz and the results were bandpass filtered from 20 to 25000 Hz.

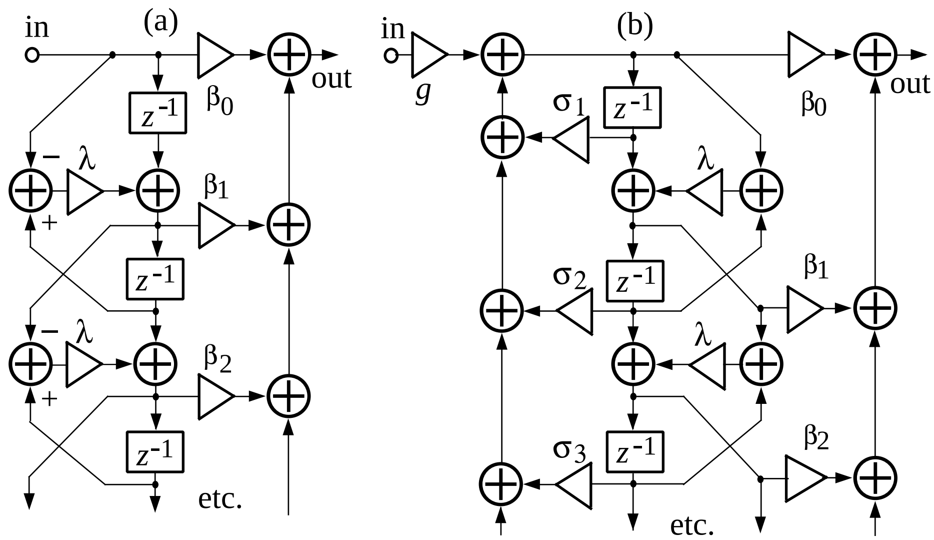

Figure 3. *Structures of warped filters. a) Warped FIR (WFIR), b) Warped IIR (WIIR).*

HRTF measurements were preprocessed according to the principles presented in Chapter 2. These stages included: 1) minimum-phase reconstruction, and 2) smoothing of the magnitude response data. Frequency warping was performed after preprocessing. In the warped frequency domain, different FIR and IIR filter design methods were compared. We ended up using a time-domain IIR design method, Prony´s method, because it outperformed other tested methods (i.e. the Yulewalk method and discrete least-squares fitting) especially in low-order approximations. In Fig. 4a, an example modeling result is illustrated. The example HRTF was measured at 0° elevation, 30° azimuth, and the frequency response of the measurement loudspeaker was deconvolved. It can be seen that a warped Prony design easily outperforms a linear Prony design of equivalent order. In this case, the order of the IIR filters was 44 compared to the FIR designs order of 90. Rectangular windowing was used both in the WFIR and FIR designs. The better performance of a WFIR compared to non-warped FIR of the same order is quite noticable. In Fig. 4b, the filter orders have been reduced to 22 and 12. The value of λ = 0.65 was used, which is slightly lower than for approximative Bark-scale warping.

3.5 Filter implementation on signal processors

The transfer function expressions of warped filters may be expanded (dewarped) to yield equivalent IIR filters of traditional form, such as direct form II filters. Such implementations have been reported in the literature [5] [6]. An alternative strategy is presented in Fig. 3, where implementation is carried out directly in the warped domain. As seen in Fig. 3, the warped filters are computationally more complex than these expanded forms. The question remains, what are the advantages and drawbacks of warped structures.

The first advantage of warped forms over traditional filters is that in many cases the warping by allpass sections results in filters less critical from the point of view of computational precision needed. Another desirable feature found in WFIR structures is

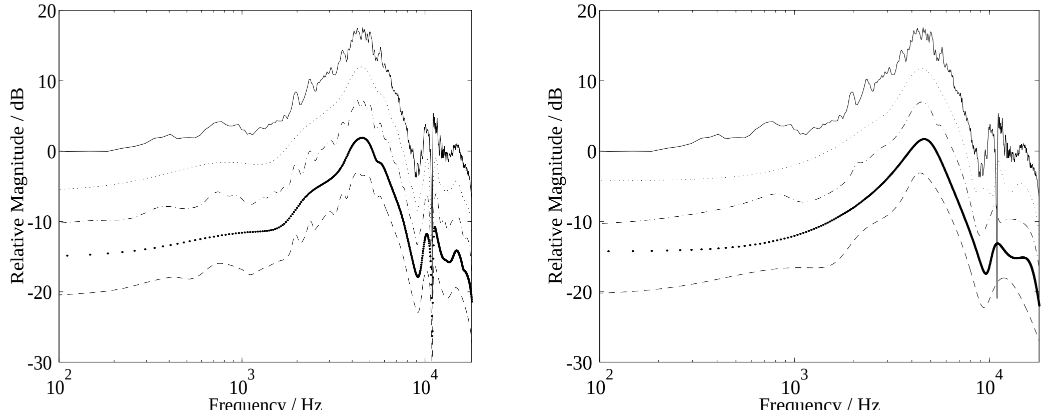

Figure 4. *Warped filter design compared with traditional FIR and IIR design.*
a) Solid line: original measurement, double-dotted line: 90-tap windowing FIR, dash-dot line: 90-tap WFIR, dotted line: Yulewalk IIR design, order 44, dashed line: Warped Yulewalk IIR design, order 44.
b) Solid line: original measurement, double-dotted line: Yulewalk IIR design, order 22, dash-dot line: Warped Yulewalk IIR design, order 22, dotted line: Yulewalk, order12, dashed line: Warped Yulewalk, order 12.

that for variable filters the coefficients are not inside recursive loops so that transients due to changing coefficients are effectively minimized. This feature may be attractive, e.g., in dynamic interpolation of HRTFs, where nonrecursive structures have been found to perform better.

The efficiency of warped vs. non-warped filters depends on the processor that is used. For Motorola DSP56000 series signal processors a WFIR takes three instructions per tap instead of one for an FIR. For WIIR filters four instructions are needed instead of two for an IIR filter. In custom design chips the warped structures may be optimized so that the overhead due to complexity can be minimized.

4 RESULTS AND DISCUSSION

The novelties presented in this paper are a theory for auditory-based filter design using FIR and IIR structures and more specifically, its application to binaural technology. Warped filter structures are designed to model HRTFs in a perceptually relevant manner, which is a major goal in binaural technology. In the future, listening tests will be performed, and more thorough comparisons of different HRTF filter design methods will be carried out.

The authors would like to express their thanks to Mr. Klaus Riederer for conducting accurate dummy head HRTF measurements at a very short notice, and to Mr. Antti Järvinen for advice on the measurement equipment. Mr. Esa Piirilä and Mr. Aki Härmä are acknowledged for their work on warped filter implementations and mapping of realizable WIIR coefficients.

5 REFERENCES

1. Kleiner, M, Dalenbäck, B.-I., Svensson, P. 1993. Auralization — An overview. *J. Audio Eng. Soc.*, vol. 41, no. 11, pp. 861-875.

2. Møller, H, Sørensen, M., Hammershøi, D., Jensen. C. 1995. Head-related transfer functions of human subjects. *JAES*, vol. 43, no. 5, pp. 300–321.

3. Blauert, J. 1983. *Spatial Hearing*. MIT Press.

4. Møller, H. 1992. Fundamentals of binaural technology. *Applied Acoustics*, vol. 36, pp. 171–218.

5. Jot, J.-M., Larcher, V., Warusfel, O. 1995. Digital signal processing issues in the context of binaural and transaural stereophony. *98th Audio Engineering Society Convention*, Paris, France, preprint no. 3980.

6. Smith, J. O. 1983. *Techniques for digital filter design and system identification with application to the violin*. Ph.D. dissert., CCRMA, Stanford Univ.

7. Takala, T., Hänninen, R., Välimäki, V., Savioja, L., Huopaniemi, J., Huotilainen, T., Karjalainen, M. 1996. An integrated system for virtual audio reality. *100th AES Convention,* Copenhagen, preprint no. 4229.

8. Mehrgardt, S., and Mellert, V. 1977. Transformation characteristics of the external human ear. *J. Acoust. Soc. Am.*, vol. 61, no. 6, pp. 1567–1576.

9. Kistler, D. J., and Wightman, W. L. 1992. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *J. Acoust. Soc. Am.*, vol. 91, no. 3, pp. 1637–1647.

10. Sandvad, J., and Hammershøi, D. 1994. Binaural auralization. Comparison of FIR and IIR filter representation of HIRs. *96th Audio Engineering Society Convention*, Amsterdam, preprint no. 3862.

11. Blommer, M. A., and Wakefield, G. H. 1994. On the design of pole-zero approximations using a logarithmic error measure. *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 3245–3248.

12. Zwicker, E., and Fastl, H. 1990. *Psychoacoustics*. Springer-Verlag.

13. Smith, J.O., and Abel, J. 1995. The bark bilinear transform. *Proc. 1995 IEEE ASSP Workshop,* Mohonk, New Paltz, New York.

14. Strube, H. W. 1980. Linear prediction on a warped frequency scale. *J. Acoust. Soc. Am.*, pp. 1071–1076.

15. Karjalainen, M. *et al*. 1996. Unpublished manuscript.