WARPED FILTERS AND THEIR AUDIO APPLICATIONS

Matti Karjalainen, Aki Härmä, Unto K. Laine, and Jyri Huopaniemi

Helsinki University of Technology
Laboratory of Acoustics and Audio Signal Processing
P. O. BOX 3000, FIN-02015 HUT, Finland
matti.karjalainen@hut.fi, aki.harma@hut.fi, unto.laine@hut.fi, jyri.huopaniemi@hut.fi

ABSTRACT

An inherent property of many DSP algorithms is that they tend to exhibit uniform frequency resolution from zero to Nyquist frequency. This is a direct consequence of using unit delays as building blocks; a frequency independent delay implies uniform frequency resolution. In audio applications, however, this is often an undesirable feature since the response properties are typically specified and measured on a logarithmic scale, following the behavior of the human auditory system. In this paper we present an overview of *warped* filters and DSP techniques which can be designed to better match the audio and auditory criteria. Audio applications, including modeling of auditory and musical phenomena, equalization techniques, auralization, and audio coding, will be presented.

1. INTRODUCTION

The idea of *warped DSP* is not a new one. FFT on a warped frequency scale was first introduced by Oppenheim et al. [1] and warped linear prediction was published by Strube [2]. A recursive warped filter structure was introduced by Steiglitz [3]. Generalized methods using FAM functions have been developed by Laine et al. [4]. The idea of warped transversal filters has been systematically studied also under the concepts of Laguerre and Kautz filters; for good introductions see [5] and [6].

In addition to the warped FFT mentioned above, there have been some practical applications of warped designs such as modeling the body of the violin [7]. Recently we have applied the principles to several new applications that will be discussed below. Warped signal processing and digital filtering principles remain, however, widely unknown. Especially warped IIR filters have not been studied in detail although they reveal interesting potential for applications. In this paper we first discuss the idea of warped filters and describe the basic filter structures. Then, a survey of recent applications and some new ideas is presented.

2. BASICS OF WARPED FILTERING

The idea of warped filters is best illustrated using the FIR-like structures in Figure 1. If each unit delay of an FIR filter is replaced with a new delay element so that each new delay is frequency dependent (dispersive), the filter can be designed and realized on a warped frequency scale.

The design of warped filters may be based on any pair of functions of complex variable, $\tilde{z} = f(z)$ and $z = g(\tilde{z})$, so that functions $f(\cdot)$ and $g(\cdot)$ are one-to-one mappings of the unit disc onto itself, and $z = g\{f(z)\}$, i.e., they are inverse mappings. There exists only one rational function type that meets the requirement, the bilinear

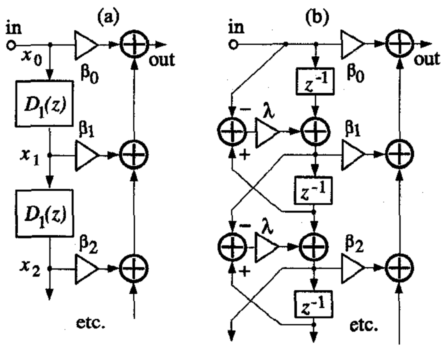

Figure 1: The principle of warped filters as an FIR structure: a) with allpass delay elements and b) as a computationally efficient version.

conformal mapping [8], which corresponds to the first order allpass filter

$$\tilde{z}^{-1} = D_1(z) = \frac{z^{-1} - \lambda}{1 - \lambda z^{-1}} \quad (1)$$

where λ, $-1 < \lambda < 1$, is a warping parameter and $D_1(z)$ is a dispersive delay element. A proper value of λ yields a good match to the psychoacoustic Bark scale [9].

We may derive the design of the WFIR filter in Fig. 1 in the following way. The desired impulse response $h(n)$ and its z-transform $H(z)$ must be equal to the impulse response $\tilde{h}(k)$ and its z-transform $\tilde{H}(\tilde{z})$ in the warped domain, i.e.,

$$H(z) = \sum_{k=0}^{\infty} \tilde{h}(k)\,\tilde{z}^{-k} \quad and \quad \tilde{H}(\tilde{z}) = \sum_{n=0}^{\infty} h(n)\,z^{-n} \quad (2)$$

Mappings between sequences $h(n)$ and $\tilde{h}(k)$ are linear but not shift-invariant. The first form specifies the WFIR realization (= synthesis) structure yielding

$$H_{WFIR}(z) = \sum_{n=0}^{M} \tilde{h}(n)\,\tilde{z}^{-n} = \sum_{n=0}^{M} \beta_n \{D_1(z)\}^n \quad (3)$$

and the second form of (2) yields a method to compute the WFIR coefficients (= analysis). It is easy to show from (1) that both forms of (2) may be computed with the same warping structure but using coefficient λ for synthesis and $-\lambda$ for analysis.

Notice that both forms of (2) yield responses of infinite length even if the sequence to be mapped is of finite length[1]. Since the coefficient sequence β_i must in practice be of finite length, we have to approximate $\tilde{h}(i)$, e.g., by truncation (3) or by windowing.

[1] Warped FIR filters have infinite impulse responses since allpass elements are internally recursive. Thus the term WFIR is somewhat contradictory but describes well the structural analogy to transversal FIR filters.

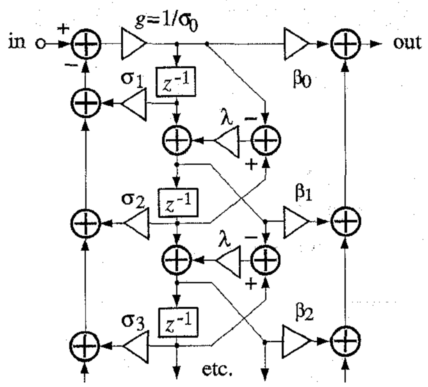

Figure 2: A realizable WIIR structure with first-order allpass delays and a single unit delay.

2.1. WIIR Filtering

A general form for the transfer function of a warped IIR (WIIR) filter is

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

Direct implementation of (4) using allpass delays is not possible since the feedback loops will contain lag-free paths when $\lambda \neq 0$. There exist several solutions that make WIIR filters realizable. Strube [2] proposed a structure where lowpass sections are used instead of allpass delays. Unfortunately, this works in practice only for low-order filters with moderate warping. A more robust realizable WIIR structure that works also for high filter orders was proposed by Steiglitz [3].

We have discussed the problem of WIIR filter design in [10] where we introduced a new filter structure, shown in Fig. 2, which is built of allpass elements only. The first delay is a unit delay and the other ones are first-order allpass sections. The recursive feedbacks α_i in (4) are mapped to coefficients σ_i in Fig. 2 which feed back from the outputs of the unit delays of the allpass sections in order to avoid lag-free loops. The following recursive mapping is derived in [10].

$\sigma_{R+1} = \lambda \alpha_R; \quad S_R = \alpha_R;$
for $i = R, R-1, ..., 2$
  $S_{i-1} = \alpha_{i-1} - \lambda S_i;$
  $\sigma_i = \lambda S_{i-1} + S_i;$
end
$\sigma_1 = S_1; \quad 1/g_0 = \sigma_0 = 1 - \lambda S_1;$

The second form of (2), the 'prewarping' of the target response, can be applied to the design of warped filters, both WFIRs and WIIRs. If the impulse response to be modeled is given, it can be mapped to warped domain and any filter design technique may be applied to yield a WFIR or a WIIR filter which can then be implemented as shown in Figures 1 and 2. If the target response of the desired filter is specified in the frequency domain, such as a magnitude response, it can be mapped to the warped frequency domain for filter design.

There are two potential advantages of using warped structures instead of traditional filters. The order of the filter will be reduced through warping in two cases: a) the inherent frequency resolution of the response to be implemented will be mapped to uniform resolution using a proper warping or b) the non-uniform frequency resolution of the human auditory system can be utilized through warping mapping. Another substantial advantage is the improved robustness and lowered precision requirements in warped filtering. This is achieved if the distribution of poles and zeros is made more uniform through the warping.

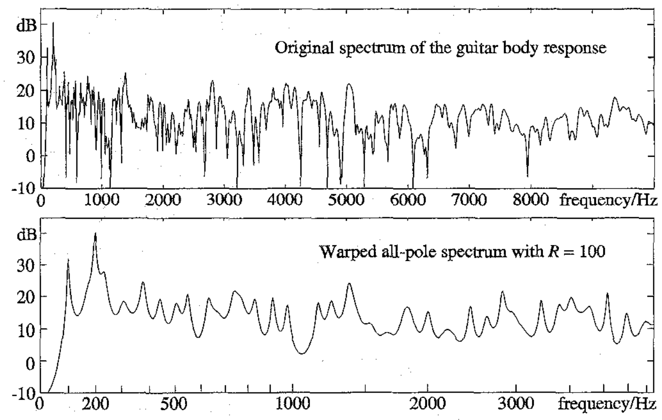

Figure 3: Modeling of the guitar body response: a) original magnitude response, b) magnitude response using WIIR modeling (notice the warped frequency scale).

The obvious disadvantage of warped filters is the increased computational complexity, see Figures 1 and 2. This extra cost depends on the DSP hardware used [10]. In some applications the tradeoff favors the warped structures. Below we will discuss some applications where we have applied warped filters successfully.

3. GUITAR BODY MODELING

The first example concerns model-based synthesis of the acoustic guitar. The modeling and real-time synthesis of string vibration is well mastered and the body can be simulated efficiently by commuted synthesis (body response as excitation) but simulation of the body as a digital filter is computationally very expensive [11].

A typical magnitude response of the acoustic guitar body is shown in Fig. 3a. For a sample rate of 22 kHz, simple FIR implementation requires an order of about 2000–5000 for good result since the lowest resonances are sharp and they decay slowly. On the other hand, the high-frequency modes are much boader by bandwidth and thus they decay faster. FIR modeling is not well suited and an IIR model fits better. Using linear prediction an all-pole model of order 500–1000 works relatively well.

Warped FIR filters of order 500 are comparable to those mentioned above. WIIR filters yield the lowest order so that a denominator order of 100–200 and numerator order of 50–100, designed using Prony's method in the warped domain, are comparable. Since the warped structures are inherently more complex, a small efficiency advantage over traditional filters remains when implementing the warped structures using typical DSP processors. In this body modeling case the frequency warping has a double match to the problem. Firstly, physically, the warping means balancing of the resonance Q values so that in the warped domain the low sharp peaks will be broadened to be more similar to the high frequency resonance peaks (Fig. 3b). Secondly, the warping has a natural match to the auditory resolution and Bark scale so that the filter order which is needed is minimized.

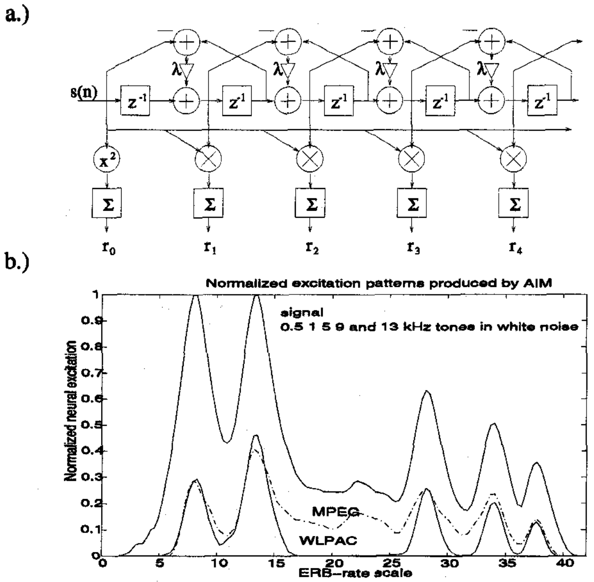

Figure 4: a) Warped autocorrelation network, b) 'Neural excitation pattern' for the original signal (upper solid curve), error signal produced in WLP codec and error signal produced in an MPEG I layer 3 codec (dashed-dotted curve). The signal is a set of tones in white noise.

4. WARPED LINEAR PREDICTION

In audio coding it is advantageous to use auditory frequency resolution at all stages of processing. Hence, frequency warped signal processing techniques appear to be a promising starting point in developing new wideband audio codecs. Most of the current audio coding schemes use subband decomposition or transform coding. The authors have made several attempts to develop methods of warped subband decomposition where *perfect reconstruction* would be available even if the subchannels were *critically downsampled.* One such solution will be published in [12].

In [13] and [14] an audio coding technique based on *Warped Linear Prediction* (WLP) was shown to be a feasible core for a new audio codec. The codec is based on the use of WFIR and WIIR filters (see Fig. 1 and 2). The coefficients of the filters are estimated in frames using warped autocorrelation analysis. The autocorrelation network is shown in Fig. 4a. The WFIR filter is applied in encoding to produce a whitened *residual* signal. After quantization this is transmitted to the decoder where a WIIR is used to reconstruct a degraded version of the original signal. The degradation of a set of tones measured in terms of an error signal, as seen at the output of a simplified auditory model [14], see Fig. 4b, shows that the auditory spectrum of the error signal (lower solid curve) follows quite closely the characteristics of the error signal produced in an *MPEG I, layer 3* codec (dash-dotted curve). The main difference between the two techniques is that in MPEG codecs there is a complex separate auditory model that controls the quantization process but in the present codec there is no separate auditory model. To some extent, the WLP codec works automatically in an auditorily convenient way.

However, in the present early phase of development, the bit rate of the WLP codec is approximately 2 times higher than that of the MPEG codec (56 kbit/s). A new highly integrated stereo codec based on a complex-valued WLP technique is proposed in [15].

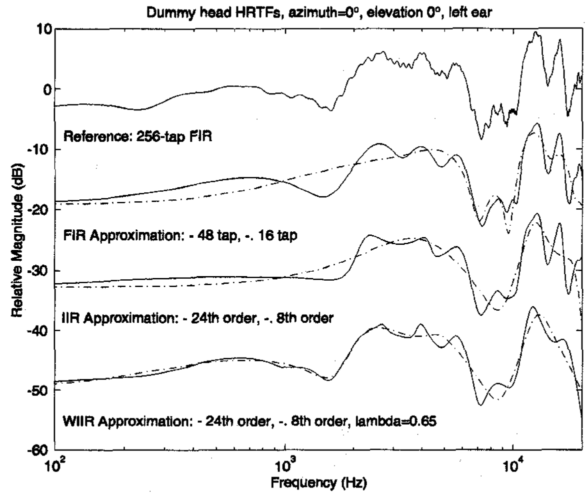

Figure 5: HRTF filter design for dummy head measurement. Frequency-sampling FIR design, Pronys method IIR and WIIR designs.

5. HRTF FILTERING

Real-time digital modeling of human spatial hearing cues is often referred to as 3-D sound spatialization or auralization. The static cues of spatial hearing are contained in head-related transfer functions (HRTF). Traditionally, HRTF filters have been created using minimum-phase reconstruction and different FIR and IIR design methods (see [16] for detailed summary). We have investigated the use of warped filters in binaural and transaural filter design. The use of a psychoacoustically based frequency scale is well motivated, and considerable reduction of filter order can be achieved using warped designs. 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 [17]. An alternative strategy is presented in [16], where implementation is carried out directly in the warped domain using warped FIR and IIR structures.

Our theoretical and empirical investigations have shown that dewarped WIIR structures outperform traditional FIR and IIR design methods. In Fig. 5, a comparison of design methods is illustrated. A dummy head HRTF was used, and two filter orders (48 and 16 taps for FIR, orders 24 and 8 for IIR and WIIR) were tested. It can clearly be seen from the results that the lower frequency fit is enhanced in WIIR designs with a trade-off of reduced high frequency matching. According to the psychoacoustic theory, this can be tolerated.

6. LOUDSPEAKER EQUALIZATION

Loudspeaker response equalization by digital inverse filtering is a well known technique although relatively few commercial implementations exist. The most common method is FIR equalization but IIR filters have also been used. The equalization is applied either to magnitude response only or to both magnitude and phase. We have studied the applicability of different equalizer filter structures, including warped structures [18].

It turns out that FIR filters are very efficient at high frequencies. This is due to the fact that FIRs inherently yield a uniform frequency

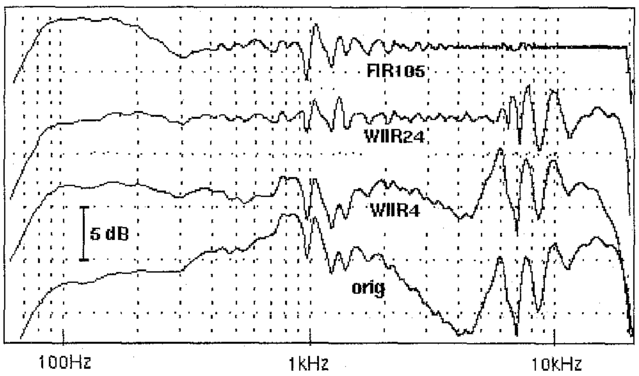

Figure 6: Loudspeaker equalization curves; *orig*: original magnitude response, *WIIR4*: warped IIR equalization, filter order 4, *WIIR24*: warped IIR order 24, *FIR105*: FIR filter equalization, filter order 105.

resolution while in audio the response specifications as well as response measurements are given on a logarithmic scale. Thus FIRs are particularly problematic to equalize at low frequencies. IIR filters avoid some of the problems with FIRs but they are more difficult to design and share the frequency resolution problem.

We have demonstrated [18] that WFIR and WIIR filter structures are useful competitors to traditional filters. Figure 6 shows a set of magnitude responses for a less-than-medium quality speaker including the original response and three equalized ones. The WIIR equalizer (inverse filter) design was based on warped Prony's method. Very low filter orders (less than 10) already show good overall equalization. Figure 6 depicts also a comparison to traditional FIR filter equalizer (order 105) which yields about the same degree of equalization than WIIR of order 24. Notice also that while the FIR filter does best job at high frequencies, WIIR filters work best at middle to low frequencies (depending on the amount of warping).

7. DISCUSSION

The applicability of warped digital filters has been shown in this article by discussing four application areas; modeling of musical instruments, warped linear prediction for audio coding, HRTF filtering for auralization, and loudspeaker response equlization. In all of these applications the advantage of frequency warping in order to match the human auditory system is utilized. In some cases the approach can be used also in the modeling of physical systems with similar frequency resolution properties.

Among some other applications we have experimented with are auditory modeling [19] and modeling of reflections and sound propagation in real and virtual acoustic spaces (to be published). We believe that there are numerous other applications where the concept of frequency warping can be found useful.

8. Acknowledgement

This work has been partially supported by the Academy of Finland.

References

[1] Oppenheim A. V., Johnson D. H., and Steiglitz K., "Computation of Spectra with Unequal Resolution Using the Fast Fourier Transform," *Proc. of the IEEE*, vol. 59, 1971.

[2] Strube H. W., "Linear Prediction on a Warped Frequency Scale," *J. Acoust. Soc. Am.*, vol. 68, no. 4 (1980).

[3] Steiglitz K., "A Note on Variable Recursive Digital Filters," *Proc. IEEE ASSP*, vol. 28, no. 1, Feb. 1980, pp. 111-112.

[4] Laine U. K., Karjalainen M., Altosaar T., "Warped Linear Prediction (WLP) in Speech and Audio Processing," *Proc. IEEE ICASSP -94*, pp. III-349-352, Adelaide, Australia, 1994.

[5] Oliveira e Silva T., *Laguerre Filters—An Introduction.* **URL:** *ftp://inesca.inesca.pt/pub/tos/English/ revdetua.ps.gz*

[6] Oliveira e Silva T., *Kautz Filters.* **URL:** *ftp://inesca .inesca.pt/pub/tos/English/ibm94e.ps.gz*

[7] Smith J.O., *Techniques for Digital Filter Design and System Identification with Application to the Violin.* Ph.D. dissertation, CCRMA Tech. Report STAN-M-58, Stanford Univ, 1983.

[8] Churchill R. V., *Complex Variables and Applications.* McGraw-Hill, New York, 1960.

[9] Smith, J.O., and Abel, J.S. "The Bark Bilinear Transform," *Proc. IEEE ASSP Workshop*, Mohonk, New Paltz, 1995.

[10] Karjalainen M., Härmä A., and Laine U.K., "Realizable Warped IIR Filters and Their Properties", *Proc. IEEE ICASSP-96*, Munich 1996.

[11] Karjalainen M., and Smith J.O., "Body Modeling Techniques for String Instrument Synthesis", *Proc. of ICMC-96*, Hong Kong, 1996, pp. 232-239.

[12] Laine U. K., "Critically Sampled PR Filterbanks of Nonuniform Resolution Based On Block Recursive FAMlet Transform", *Proc. EUROSPEECH'97*, Rhodos, 1997.

[13] Härmä A., Laine U.K., and Karjalainen M., "Warped Linear Prediction (WLP) in Audio Coding", *Proc. NorSig-96*, Sept. 24-27, 1996, pp. 447-450.

[14] Härmä A., Laine U.K., and Karjalainen M., "WLPAC – A Perceptual Audio Codec in a Nutshell", AES 102nd Convention preprint 4420, Munich, p. 33, 1997.

[15] Härmä A., Laine U.K., and Karjalainen M., "An Experimental Audio Codec Based on Warped Linear Prediction of Complex Valued Signals", *Proc. ICASSP'97*, Munich, April 1997.

[16] Huopaniemi J., and Karjalainen, M., "Review of digital filter design and implementation methods for 3-D sound," *102nd AES Convention*, Munich, March. 1997, preprint 4461.

[17] Jot J.-M., Larcher V., and Warusfel O., "Digital Signal Processing Issues in the Context of Binaural and Transaural Stereophony", 98th AES Convention, Paris, France, Feb. 1995, preprint 3980.

[18] Karjalainen M., Piirilä E., Järvinen A., and Huopaniemi J., "Comparison of Loudspeaker Equalization Methods Based on DSP Techniques", *102nd AES Convention*, Munich, March. 1997, preprint 4437.

[19] Karjalainen M., "A Binaural Auditory Model for Sound Quality Measurements and Spatial Hearing Studies", *Proc. IEEE ICASSP-96*, Atlanta, 1996.