



NORDIC ACOUSTICAL MEETING
12 - 14 JUNE 1996 HELSINKI

WARPED FILTER DESIGN FOR THE BODY MODELING AND SOUND SYNTHESIS OF STRING INSTRUMENTS

Matti Karjalainen

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

1 INTRODUCTION

It is characteristic to natural sound sources that low frequencies show sharper resonances along with slower temporal decay while high frequencies show the opposite, i.e. broader bandwidths and faster decay. Thus it is only natural that the human auditory system is also matched to this general trend and follows a similar tradeoff between time and frequency resolution for different frequencies.

The body of string instruments, such as the acoustic guitar, is a good example of systems with the characteristics described above. In this paper we show how the transfer function of such bodies can be formulated in terms of warped digital filters. First we analyze typical body responses to show their properties both in the time and the frequency domain. Next the principle of warped filters is introduced and some interesting filter structures are discussed. Based on this approach, body filters can be designed and implemented directly in the warped frequency domain. Computational efficiency of warped filters is compared with unwarped ones which shows that warped filters are more optimal for the purpose.

2 BODY RESPONSE CHARACTERISTICS

The acoustics of string instrument bodies and soundboards is a relatively widely studied topic, see [1] and further references in it. From the point of view of real-time sound synthesis, all traditional simulation methods by computer are too slow. Instead, we need signal processing models that are as efficient as possible. Since the transfer function from string vibration to radiated sound can in most cases be accurately modeled as a linear and time-invariant system, a practical implementation for sound synthesis is by means of digital filtering [2]. As was shown in [3] and [4], even more efficient way is to aggregate the body response into a string excitation wavetable. However, as an alternative with ability to control the model parametrically, we will discuss efficient filter-based models in this paper. Filter-based body models have been studied earlier, e.g., by

Smith [5] for the violin, by Karjalainen for the guitar [6], and by Laroche [7] for the piano.

A natural starting point for body modeling is to measure or to compute the body impulse response. Figure 1 shows the first 100 milliseconds of the impulse response from the body of an acoustic guitar of classical design. The response was measured by tapping the bridge vertically with an impulse hammer (strings were damped) and by measuring the response with a microphone located one meter in front of the sound hole. Figure 2 depicts the magnitude behavior in the frequency domain for the full impulse response.

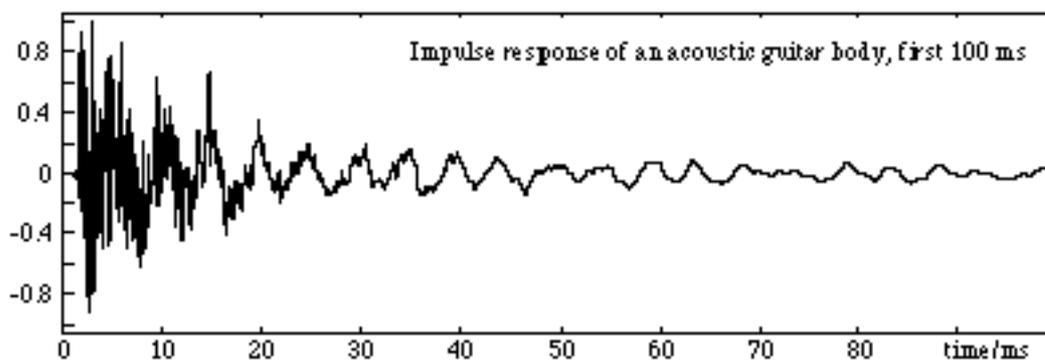


Figure 1. An example of body impulse response for an acoustic guitar.

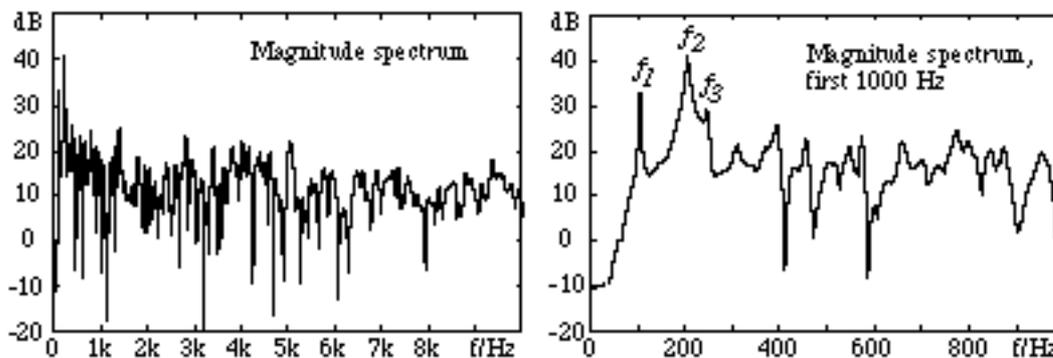


Figure 2. Magnitude spectrum of the impulse response shown in Figure 1.

Figure 1 suggests that the impulse response of the body is a combination of exponentially decaying sinusoids, i.e., signal components corresponding to the resonance frequencies of Figure 2. In the frequency domain, depending how the harmonics of a string signal are located in relation to peaks or valleys, the signal will be spectrally colored. It is obvious that both the magnitude spectrum shape and the temporal structure of the impulse response are important from the point of view of auditory perception.

A more comprehensive look at the body response characteristics may be achieved by a proper time-frequency representation. Figure 3 illustrates a short-time Fourier spectrum plot for the first 70 milliseconds of the response. A general relation between the decay rate and the frequency of a mode can be easily observed. We may also notice that resonance modes do not follow simple exponential decay that should be linear curves on the dB scale. This ripple can be explained by multiple resonances that interact since they are more closely located than the spectral resolution of the analysis. (At low frequencies

the alignment of window in relation to signal cycles may also be a source of ripple. Actually, there is no good window size for this analysis since low frequencies require better frequency resolution and high frequencies better temporal resolution.)

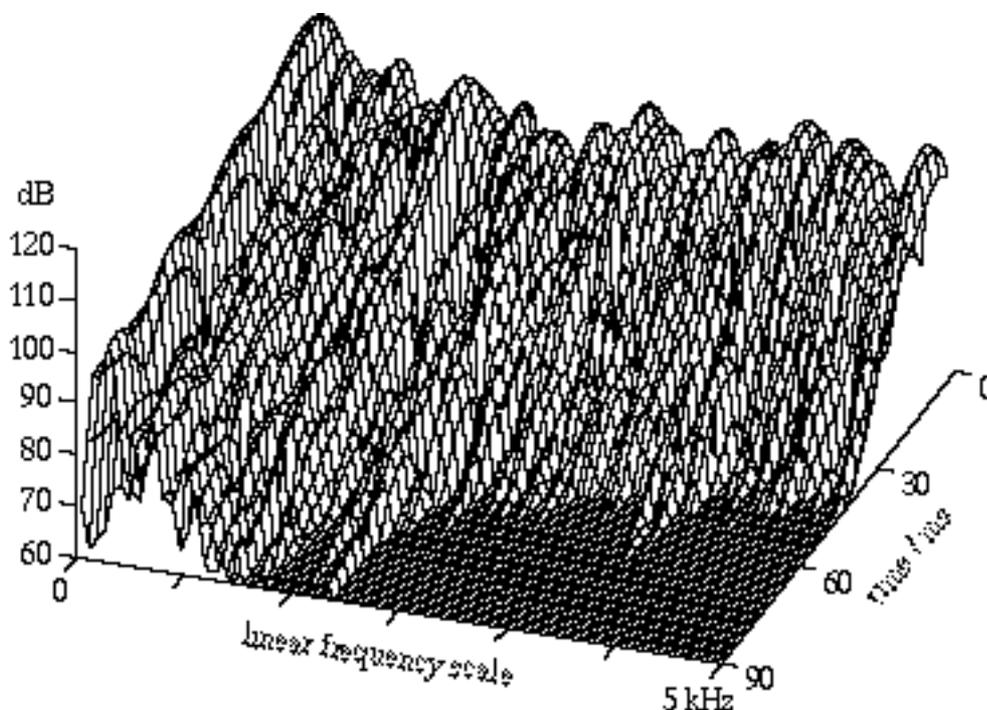


Figure 3. Time-frequency plot of the guitar body response using short-time Fourier analysis. A Hamming window of 12 ms was used with a 3 ms hop size.

3 TRADITIONAL DIGITAL FILTERS AS BODY MODELS

The signal transfer properties from strings to radiated sound can be considered to be linear and time invariant (LTI) in most string instruments. Thus the most efficient way of modeling the body or soundboard for sound synthesis purposes is by means of digital filtering. Here we first study the use of traditional filter structures—FIR and IIR filters—as direct implementations of body impulse responses. Then we introduce warped filter techniques and their application to body modeling.

3.1 FIR and IIR filters as body models

A discrete-time LTI system may be represented using z-transform by

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{i=0}^N b_i z^{-i}}{1 + \sum_{i=1}^P a_i z^{-i}} \quad (1)$$

The most straightforward way to realize a known body response is to use the samples of a measured or computed impulse response as taps in an *FIR filter*, for which coefficients a_i in (1) are zero. This implements the desired convolution of the string output

and the body response yielding a full accuracy body model as far as the whole audible portion of the response is available free of noise and artifacts.

An obvious problem with FIR modeling is the filter length N and thus the computational expense of the method. In the current guitar example, in order to cover a period of a single decay time constant for the lowest mode, an FIR filter of $N = 5000$ taps is needed when a sampling rate of 22 kHz is used. For a 60 dB dynamic range and full audio bandwidth, an FIR order of about $N = 25000$ is required! In practice, a 100 ms slice of the response is found quite satisfactory, which means a 2200 tap filter for a 22 kHz sampling rate. Even this is computationally much more expensive than a model for six guitar strings, and may be more than a modern signal processor can do in real time. The conclusion is that FIR models are impractical unless very efficient FIR hardware is available.

The sharply resonating and exponentially decaying components of a body response imply that IIR filters are more appropriate for efficient synthesis models than FIR filters. In order to see how well straightforward all-pole modeling works, we may apply autoregressive (AR) modeling using the autocorrelation method of *linear prediction* (LP) [8] to the impulse response shown in Figure 1. This yields an *all-pole filter* model where coefficients b_i of (1), for $i = 2 \dots N$, are equal to zero, a_i are the predictor coefficients, and P is the order of the filter. Experiments with all-pole modeling [9], [10] have shown that in our example an order of $P = 500 \dots 1000$ is needed to yield a proper temporal response. Lower filter orders, although relatively good from a spectral point of view, make the lowest resonances decay too fast.

The next generalization with traditional digital filters is to model the impulse response with a pole-zero (or ARMA) model. We have tried this using the Prony's method [2]. The results show [9], [10] that this does not relax the requirements for the order P but addition of, e.g., 100 zeros to the model improves to fit the transient attack of the response. From the point of view of auditory perception, however, this has only a negligible effect.

4. WARPED FILTERS

Some filter design and model estimation methods allow for an error weighting function in order to control the varying importance of different frequencies [5]. 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 [11], [2] for the warping of impulse responses or transfer functions. Warped FFT was first introduced by Oppenheim & al., [12] and warped linear prediction was developed by Strube [13]. Generalized methods using the FAM functions have been developed by Laine & al., [14]. Smith has applied the bilinear mapping in order to design filter models for the violin body [5].

The bilinear warping is realized by substituting unit delays by first-order allpass sections, i.e.

$$z^{-1} \Leftarrow D_1(z) = \frac{z^{-1} - \lambda}{1 - \lambda z^{-1}} \quad (2)$$

This means that the frequency-warped version of a filter can be implemented by such a simple replacement technique. (Modifications are needed to make warped IIR filters realizable.) The transfer function expressions after the substitution may also be expanded to yield an equivalent IIR filter of traditional form. It is easy to show that the inverse warping can be achieved with a similar substitution but using $-\lambda$ instead of λ .

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 $H(z)$. For an appropriate value of λ , the bilinear warping can fit the psychoacoustic Bark scale, based on the critical band concept, relatively accurately [13]. For this purpose, an approximate formula for the optimum value of λ as a function of sampling rate is given in [15]. For a sampling rate of 44.1 kHz this yields $\lambda = 0.7233$ and for 22 kHz $\lambda = 0.6288$. When using the warping techniques, the optimality of λ in a specific application depends both on auditory aspects and the characteristics of the system to be modeled.

4.1 Warped FIR (WFIR) filters

The principle of a warped FIR filter (WFIR) is shown in Figure 4a, which may be written as

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

A more detailed filter structure for implementation is depicted in 4b. As the latter form shows, 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.

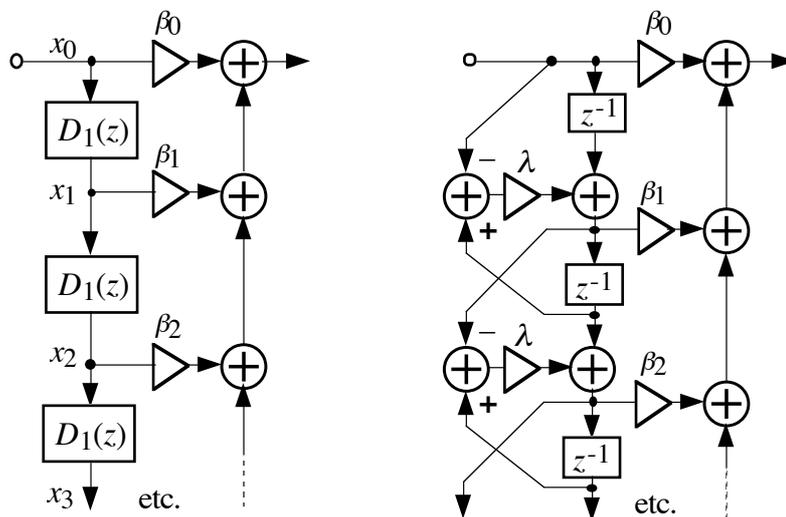


Figure 4. Warped FIR modeling: (a) general principle, (b) detailed filter structure for implementation.

A straightforward method to get the tap coefficients β_i for a WFIR filter is to warp the original impulse response and to truncate by "windowing" the portion that has amplitude above a threshold of interest. (Notice that the bilinear mapping of a signal by (2) is linear but not shift invariant [13]). There exist various formulations for computing a warped version of a signal [13], [5], [14]. An accurate and numerically stable method is

to apply the FIR filter structure of Figure 4a or 4b with tap coefficients being the samples of the signal to be warped. When an impulse is fed to this filter, the response will be the warped signal. Figure 5 shows the warped ($\lambda = 0.63$) guitar body response as a time-frequency plot for comparison with the original one in Figure 3. As can easily be noticed, the warping has balanced the average decay rates and resonance bandwidths for all frequency ranges.

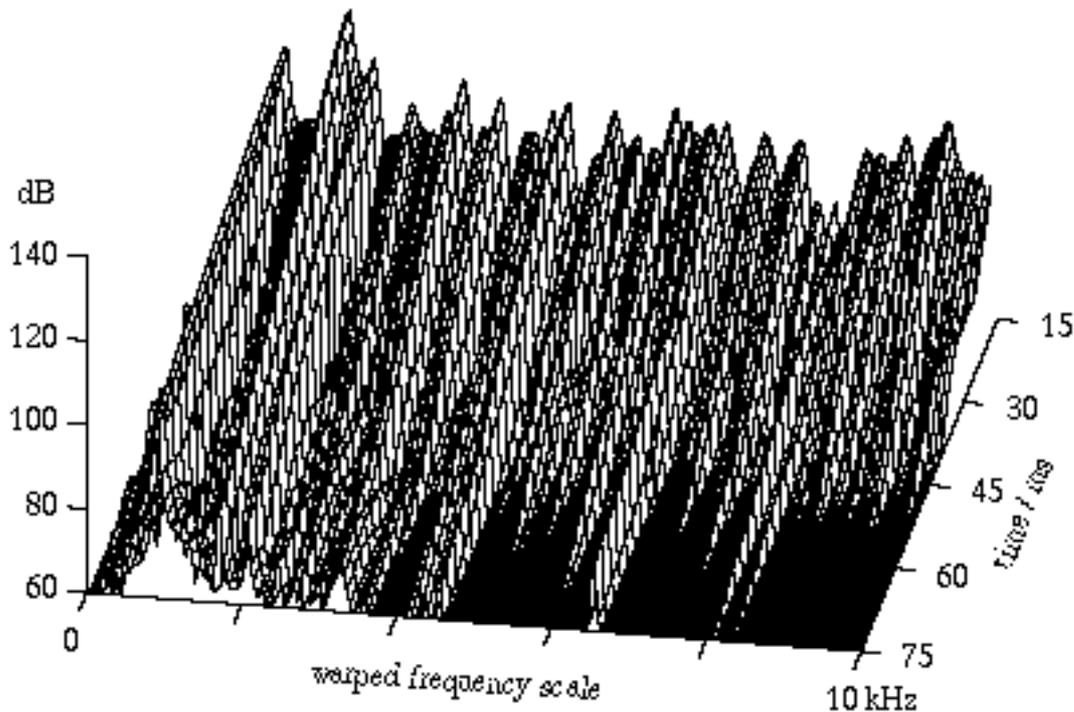


Figure 5. Time-frequency plot of warped guitar body response. A Hamming window of 24 ms^* was used with a 3 ms^* hop size, where $*$ denotes warped time.

4.2 Warped IIR (WIIR) filters

When linear prediction is applied to a warped impulse response it yields a warped all-pole filter. Other methods for warped LP analysis (WLP) are studied in [14], including an efficient way to compute warped autocorrelation coefficients $r_w(i)$ directly from the original signal. This is based on the warped delay-line structure of Figure 4a, whereby

$$r_w(i) = \sum x_0(n) x_i(n) \quad (4)$$

is summed over a time interval or window of interest. After that, the warped predictor coefficients α_i are achieved from warped autocorrelation coefficients as usual [8] to yield a filter model

$$H_w(D_1(z)) = \frac{G_w}{1 + \sum_{i=1}^R \alpha_i [D_1(z)]^i} \quad (5)$$

A somewhat surprising observation is that the filter structure of (5) cannot be implemented directly since there will be delay-free loops in the structure for $\lambda \neq 0$. (Of course the bilinear mapping, inherent in the filter structure, may be expanded at design time to

yield a normal IIR filter. This, however, leads to numerical instability if the filter order exceeds about 20...30.). Figure 6 depicts two realizable forms of WIIR filters. Strube [13] suggested a version where lowpass sections are used instead of allpass sections, see Figure 6a. In practice it works only for low orders and warping values λ . The version in Figure 6b is more general for warped pole-zero modeling but it has also a more complex structure. The coefficients σ_i can be computed from coefficients α_i of (5) using a recursion or matrix operation as shown in [9] and [10].

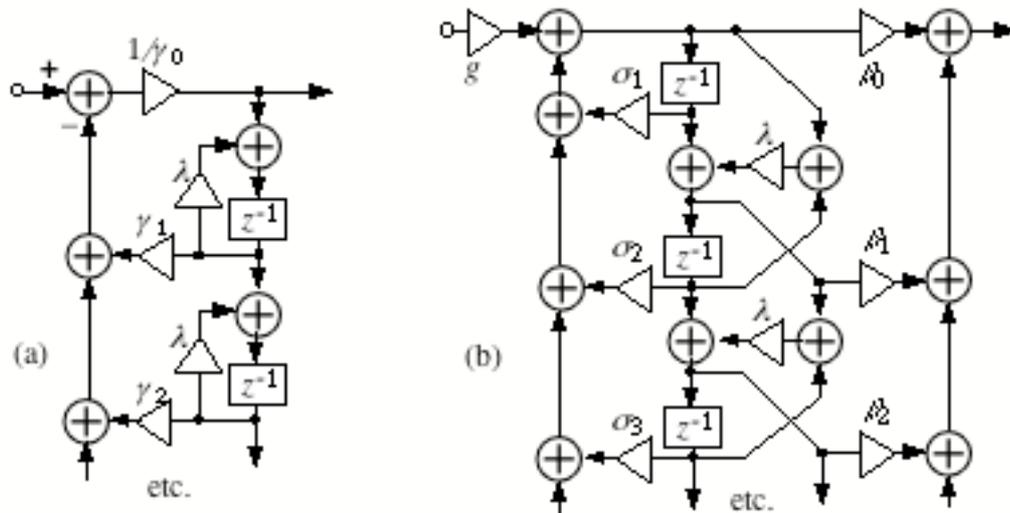


Figure 6. Filter structures for implementation of warped IIR filters: (a) lowpass structure that does not work with high orders, and (b) modified allpass structure (warped pole-zero filter).

4.3 Body model implementation with warped filters

The warped filter strategy using Bark scale warping yields a reduction in filter order of factor 5...10. This means that a WFIR of M less than 500 gives results similar to FIR of $N = 2000$. For warped all-pole filters order of about $R = 100$ is equivalent to normal IIR of $P = 500...1000$. A body filter, resulting from using the Prony's method for warped filter with $M = 50...100$ and $R = 100...200$, will represent both transient and decay properties relatively well, although a Bark warping ($\lambda = 0.63$, sampling rate 22 kHz) has a tendency to shorten the impulse response of the highest frequencies a bit too much.

The reduction in filter order due to warping is 'real' only if unit delays and allpass sections were computationally equally complex. In reality, many digital signal processors have hardware support to run ordinary FIR and IIR filters very efficiently. The complexity of the first-order allpass section used as a warped delay is much higher than that of a unit delay. Due to this complexity of realization, reduction in computational cost with warped all-pole and IIR structures remains smaller than indicated by order savings. In a typical case, for the TMS320C30 floating-point signal processor, a WIIR body filter model is only about two times faster than an equivalent normal IIR filter.

5 ACKNOWLEDGEMENT

This work was carried out while the author was a visiting scholar at CCRMA, Stanford University, U.S.A. The work was financially supported by the Academy of Finland.

6 REFERENCES

1. Fletcher, N., H., and Rossing, T., D. 1991. *The Physics of Musical Instruments*. Springer-Verlag, New York.
2. Parks, T., W., and Burrus, C., S. 1987. *Digital Filter Design*. John Wiley & Sons, New York.
3. Smith, J., O. 1993. "Efficient Synthesis of Stringed Musical Instruments," in *Proc. 1993 Int. Comp. Music Conf. (ICMC'93)*, pp. 64-71, Tokyo, Japan.
4. Karjalainen, M., Välimäki, V., and Jánosy, Z. 1993. "Towards High-Quality Synthesis of the Guitar and String Instruments," *Proc. 1993 Int. Comp. Music Conf. (ICMC'93)*, pp. 56-63, Tokyo.
5. Smith, J., O. 1983. *Techniques for Digital Filter Design and System Identification with Application to the Violin*. Ph.D. dissertation, CCRMA Tech. Report STAN-M-58, Stanford University, 260 p. Stanford.
6. Karjalainen, M., Laine, U., K., and Välimäki, V. 1991. "Aspects in Modeling and Real-Time Synthesis of the Acoustic Guitar," In *Proc. IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, New York.
7. Laroche, J., and Meillier, J-L. 1994. "Multi-Channel Excitation/Filter Modeling of Percussive Sounds with Application to the Piano," *IEEE Trans. Speech and Audio Processing*, vol. 2, no 2, pp. 329-344 (April).
8. Markel, J., D., and Gray, A., H. 1976. *Linear Prediction of Speech*. Springer-Verlag, New York.
9. Karjalainen, M. (ed.) 1996. *Modeling and Synthesis of String Instruments*. URL: <http://www.hut.fi/HUT/Acoustics/stringmodels/> (Word Wide Web).
10. Karjalainen, M., Smith, J., O. 1996. "Body Modeling Techniques for String Instrument Synthesis". To be published in *Proc. 1996 Int. Computer Music Conf. (ICMC'96)*, Hong Kong.
11. Churchill, R., V. 1960. *Complex Variables and Applications*. McGraw-Hill, New York.
12. Oppenheim, A., V., Johnson, D., H., and Steiglitz, K. 1971. "Computation of Spectra with Unequal Resolution Using the Fast Fourier Transform," *Proc. of the IEEE*, vol. 59, pp. 299-301.

-
13. Strube, H., W. 1980. "Linear prediction on a Warped Frequency Scale," *J. Acoust. Soc. Am.*, vol. 68, no. 4, pp. 1071-1076.
 14. Laine, U., K., Karjalainen, M., Altosaar, T. 1994. "Warped Linear Prediction (WLP) in Speech and Audio Processing," *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP -94)*, pp. III-349-352, Adelaide, Australia.
 - 15.. Smith, J., O., and Abel, J., S. 1995. "The Bark Bilinear Transform," *Proc. IEEE ASSP Workshop*, Mohonk, New Paltz.