A WARPED BANDWIDTH EXPANSION FILTER

Marc A. Boillot

Motorola iDEN Group
Plantation, FL
emb034@email.mot.com

John G. Harris

Computational Neuro-Engineering Lab
University of Florida, Gainesville, FL
harris@cnel.ufl.edu

ABSTRACT

A warped filter is presented as a new speech enhancement method to adjust formant bandwidths on a critical band scale. The warped filter enhances perceived loudness without adding signal energy by exploiting the psychoacoustic nature of the auditory system. The critical band concept in auditory theory states that when the energy in a signal remains constant, loudness increases when the energy spreads beyond a critical bandwidth. A warped filter is proposed and developed to elevate the perceived loudness of clean speech by applying non-linear bandwidth expansion to the formant regions of vowels in accordance with the critical band scale. The filter has been inspired and motivated by the biological representation of loudness in the peripheral auditory system and the critical band concept of hearing.

1. BACKGROUND

Loudness is intimately related to the critical band concept of hearing. The critical band concept states that spectral components separated by frequency so as to fall into different auditory channels are processed separately. The critical band concept states that when the energy in a signal remains constant, loudness will increase when a critical band is exceeded. This provides a compelling motivation for a means to increase speech loudness without adding energy to the signal via formant bandwidth expansion. This would be a practical consideration for power limited devices such as cell phones or hearing aids with high audio output requirements.

Vowels are precipitated as candidates for formant bandwidth expansion since they are high energy, resonant, and spectrally smooth. Because vowels have formant bandwidths which increase with increasing frequency, the filter should elevate speech loudness by applying bandwidth expansion on a critical band scale. The LPC pole displacement technique is applied in the warped domain to increase formants on a critical band scale. The bandwidth adjustment technique has been used in spectral distortion measures [1], to sharpen formant bandwidths [2], as a postfilter [3], and recently to increase vowel loudness on a linear frequency scale [4]. The authors believe they are the first to apply the technique in the warped domain to adjust formants on a critical band scale to elevate perceived loudness [5]. The technique is used within the context of a Warped Linear Prediction Coefficient (WLPC) vocoder structure to provide the necessary degree of freedom for critical bandwidth adjustment [6]. In this paper we show how an off axis radius term is incorporated in the linear transformation of the warped coefficient set.

Funding for this research was provided by the iDEN Technology Group and Product Development Group of Motorola

2. POLE DISPLACEMENT MODEL

A technique used to alter formant bandwidth is shown in Eq (1) [7] and demonstrated in Fig(1). This provides a way to evaluate the Z transform on a circle with radius r greater than or less than the unit circle. For $0 < r < 1$ the evaluation is on a circle closer to the poles and the contribution of the poles has effectively increased, thus sharpening the pole resonance. Stability is of concern since $1/A(\tilde{z})$ may no longer be an analytic expression within the unit circle. For $r > 1$ (bandwidth expansion) the evaluation is on a circle farther away from the poles and thus the pole resonance peaks decrease and the pole bandwidths are widened. The poles are always inside the unit circle and $1/A(\tilde{z})$ is stable.

$$A(\tilde{z})|_{\tilde{z}=re^{jw}} = \sum_{k=0}^{p}(a_k r^{-k})\, e^{-jwk} \tag{1}$$

If the poles are well separated [8], the bandwidth increase is related to the new evaluation circle with radius $r > 1$ by

$$\triangle B \approx \ln(r) f_s/\pi \tag{2}$$

This follows from an s-plane result that the bandwidth of a pole in radians/sec is equal to twice the distance of the pole from the jw-axis when the pole is isolated from other poles and zeros [9]. The bandwidth adjustment technique simply requires a scaling of the LP coefficients by a power series of r, Eq(1).

3. WARPED LPC FILTER

Eq(1) shows how the Z transform can be evaluated on a circle of radius r given the LP coefficients. This operation is a function of the radius which determines the amount of bandwidth change, as given in Eq(2), which is constant for well separated poles. The elegance of the technique is that the coefficients can be scaled directly. It will be applied to expand pole bandwidths on a scale closer to that of the human auditory system using a WLPC filter design. Expanding beyond critical bands increases the perception of loudness and the intent is to increase the perceived loudness without adding signal energy. Warped filters have primarily been used for audio filter design to better model the frequency response to that of human hearing. Since warped filter structures are realizable, the linear bandwidth expansion technique can be used in this transformed space to achieve nonlinear bandwidth expansion. Warped linear prediction uses an all-pass filter of Eq(3) to achieve frequency warping with a phase response given by Eq(4). An all-pass factor of $\alpha = 0.47$ at $f_s = 10kHz$ provides an approximation to critical band warping [6].

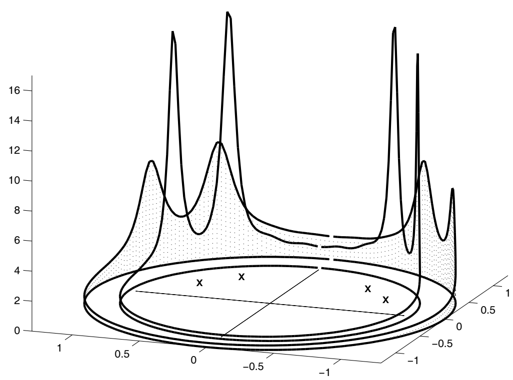

Fig. 1. Pole displacement model

$$\tilde{z}^{-1} \quad = \quad \frac{z^{-1} - \alpha}{1 - \alpha z^{-1}} \tag{3}$$

$$\tilde{w} \quad = \quad w + 2 \arctan \left[\frac{\alpha \, \sin w}{1 - \alpha \, \cos w} \right] \tag{4}$$

4. METHOD

The transformation is a one-to-one mapping of the z domain, and can be can be done recursively using the Oppenheim recursion [10]. The recursion can be applied to the autocorrelation sequence R_n, power spectrum P_n, prediction parameters a_p, or cepstral parameters. We used the Oppenheim recursion on the autocorrelation sequence for the frequency warping transformation. Once the warped autocorrelation function $\tilde{r}_k$ was determined, the warped coefficients $\tilde{a}_k$ were obtained by the Durbin algorithm [8]. Fig(2) shows the resulting frequency response of the Oppenheim recursion as applied to a synthetic speech segment with $\alpha = 0.47$. It can be seen that the method stretches the envelope rightwards. Critical bandwidths increase with increasing frequency. Since the warped spectrum is on a critical band scale, the large bandwidth high frequency regions of the original spectrum become compressed, and effectively result in a warped spectrum stretched towards the right. For $0 < \alpha < 1$ frequency warping stretches the low frequencies and compresses the high frequencies. For $-1 < \alpha < 0$ frequency warping compresses the low frequencies and stretches the high frequencies.

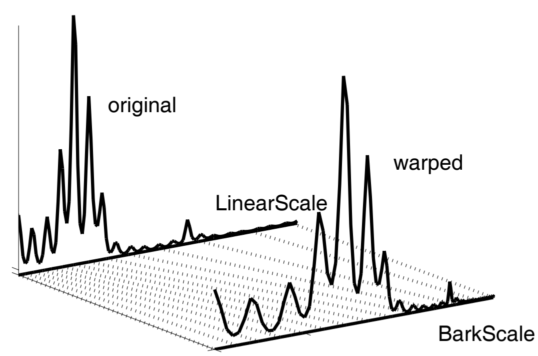

Fig. 2. Critical band frequency warping

5. FILTER DESIGN

The warped prediction coefficients $\tilde{a}_k$ define the prediction error (analysis) filter given by

$$\tilde{A}(z) = 1 - \sum_{k=1}^{p} \tilde{\mathbf{a}}_k z^{-k}(z) \tag{5}$$

and can be directly implemented as an FIR with each unit delay being replaced by an all-pass filter. However, the inverse (synthesis) IIR filter is given by $\tilde{A}^{-1}(z)$, and is not a straightforward unit delay replacement. The substitution of all-passes into the unit delay of the recursive IIR form creates a lag free term in the delay feedback loop. The lag free term must be incorporated into a delay structure which lags all terms equally to be realizable. Realizable warped recursive filter designs have been proposed to mediate this problem and elegant solutions exist [6][11][12].

One method for realization of the warped IIR form requires the all-pass delay elements to be replaced with first order low pass sections. The filter structure will be stable if the warping is moderate and the filter order p is low [12]. Eq(5) can be expressed as a polynomial in $z^{-1}/(1-\alpha z^{-1})$ to map the prediction coefficients $\tilde{a}_k$ to a coefficient set used directly in a standard recursive filter structure. In this manner the all-pass lag free-element is removed from the open loop gain and a realizable warped IIR filter is possible [6].

$$A(z) = \sum_{k=0}^{p} b_k \left[\frac{z^{-1}}{1-\alpha z^{-1}} \right]^k \tag{6}$$

The b_k coefficients are generated by a linear transform of the warped LPC coefficients $\tilde{a}_k$, using the binomial equations, or recursively [6]. The bandwidth expansion technique of Eq(1) can be incorporated into the warped filter of Eq(6). Modification of the binomial equation is necessary to include the bandwidth adjustment factor r in the calculation of the coefficients.

$$b_k = \sum_{n=k}^{p} C_{kn}\tilde{\mathbf{a}}_n,\ for\ C_{kn} = \binom{n}{k}(1-\alpha^2)^k(-\alpha)^{n-k}r^{-n}$$

Application of the linear expansion technique in the warped domain effectively bandwidth broadens the poles on a nonlinear scale given by α. For $\alpha = 0.47$ the technique broadens the pole bandwidths on an approximate critical band scale [5]. Figure 3 shows our canonic direct form of the WLPC filter with critical band expansion. The $\dot{b}_k$ coefficients (on the left) are the bandwidth expanded terms in the IIR structure ($r > 1$). The WLPC bandwidth expansion filter can be put in the same form as that of the general vocoder postfilter [3] of Eq(7).

$$H(z) = \frac{A(\tilde{z}/\gamma)}{A(\tilde{z}/\beta)} \tag{7}$$

Figure 4 shows the family of bandwidth expansion curves given a particular sampling frequency and evaluation radius. This figure characterizes the warped filter for an evaluation radius of $r = 1.02$. The α values specify the level of bandwidth expansion or compression. For $\alpha \neq 0$ the intersection of each curve with the $a = 0$ curve sets the crossover frequency. It can be seen from Figure 4 that at $\alpha = 0$ there is uniform bandwidth expansion across all frequencies and the bandwidth corresponds to $B = \log(r)fs/\pi$, which is $B = 50$Hz for $fs = 8kHz$ and $\alpha = 0$. The change in bandwidth is specified by the evaluation radius, sampling frequency, and α value. The

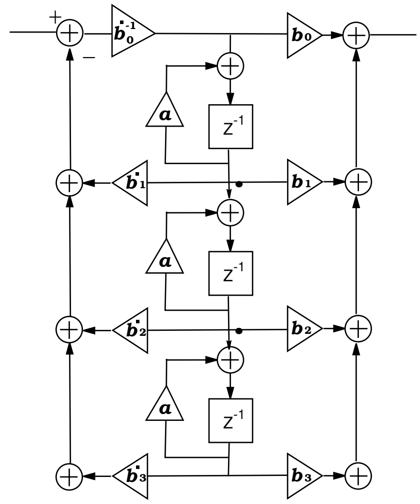

Fig. 3. Warped bandwidth expansion filter

bandwidth expansion is constant in the warped domain and given by Eq(2). The resulting linear bandwidth change can be evaluated from the warped bandwidth using the all-pass inversion property of Eq(4). The warped bandwidth expansion filter introduces a nonlinear frequency gain because it effectively increases pole bandwidth by decreasing pole resonance. This reveals the interdependency of pole bandwidth and pole gain as a function of frequency in a warped filter structure.

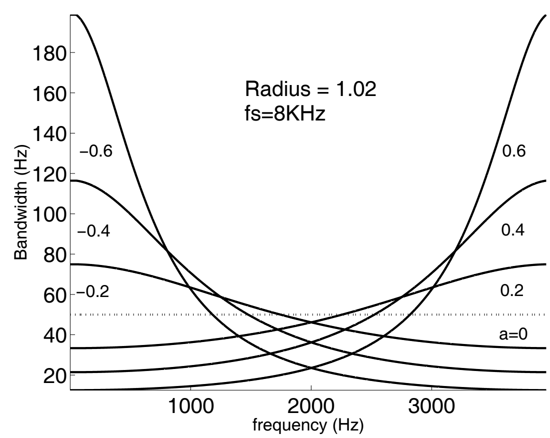

Fig. 4. Family of curves showing frequency dependent bandwidth expansion for an evaluation radius $r = 1.02$ for the warped filter with various values of the warping factor, $-0.6 < \alpha < 0.6$.

Figure 5 illustrates the change in formant bandwidths of a vowel processed with the warped filter of Eq(7). The filter demo uses a radius specified by the center slider with a warping factor of 0.34 set by the right hand slider which is

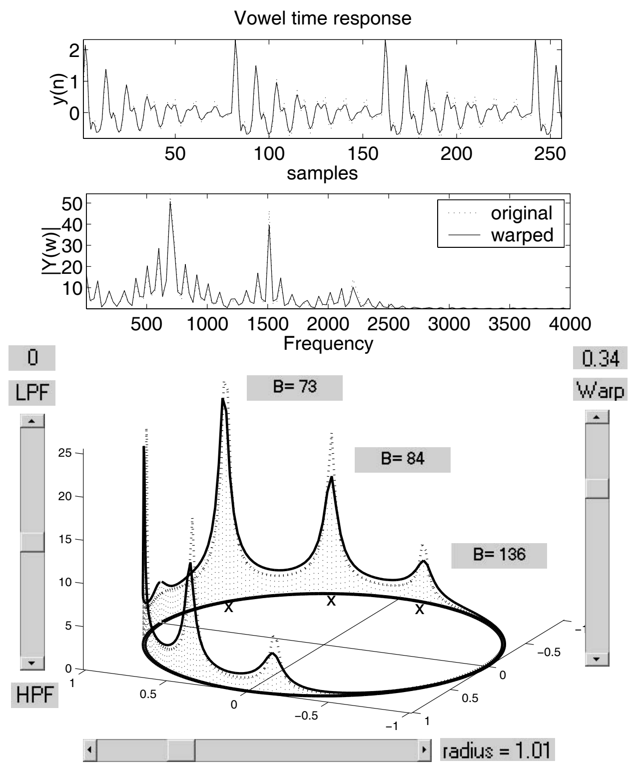

Fig. 5. Spectral envelope of a synthetic vowel shows an evaluation off the unit circle with warping results in a non-uniform bandwidth change for all formants. Original formant bandwidths (dotted) are all 50Hz.

adjustable. It also includes a first order high-pass filter of the form $1 - \gamma z^{-1}$ with γ set by the left slider. Results show the change in pole bandwidths are no longer constant as would be for the case of a linear pole displacement model.

6. CONCLUSIONS

The motivation of this work has been to design a real-time speech enhancement filter which adjusts formant bandwidths on a critical band scale. Our goal has been to exploit the psychoacoustic nature of the auditory system by increasing the perceptual loudness of speech without adding energy to the signal. A warped filter with pole displacement was proposed for this task. The design was inspired by the biological representation of loudness in the peripheral auditory system. The critical band concept of hearing provided a logical motivation to incorporate a warped pole displacement model. The design effectively allows for bandwidth expansion of the vowel formants on a critical band scale. Human listening tests also reveal the warped filter can enhance loudness without adding energy [5].

7. REFERENCES

[1] F.K. Soong and M.M. Sondhi, "A frequency Weighted Itakura Spectral Distortion Measure and its Application to Speech Recognition in Noise," *IEEE Trans. on Acoustic., Speech, and Signal Proc.*, vol. 36 (1), pp. 41–48, Jan 1988.

[2] D. Mansour and B.H. Juang, "A Family of Distortion Measures Based Upon the Projection Operation for Robust Speech Recognition," *IEEE Trans. on Acoustic., Speech, and Signal Proc.*, vol. 37 (11), pp. 1659–1671, Nov 1989.

[3] J.H. Chen and A. Gersho, "Adaptive Postfiltering for Quality Enhancement of Coded Speech," *IEEE Trans. on Speech and Audio Proc.*, vol. 3 (1), pp. 59–71, 1995.

[4] M. Boillot and J.G. Harris, "A Loudness Enhancement Technique for Speech," *ISCAS*, vol. INV-14-3424 May 23-26 Vancouver Canada, 2004.

[5] M. Boillot, *A Warped Filter Implementation for the Loudness Enhancement of Speech*, University of Florida, Ph.D. Dissertation 2002.

[6] H.W. Strube, "Linear Prediction on a Warped Frequency Scale," *J. Acoustic. Soc. Amer.*, vol. 68 (4), pp. 1071–1076, 1980.

[7] S. McCandless, "An Algorithm for Automatic Formant Extraction Using Linear Predictive Spectra," *IEEE Trans. on Acoustic., Speech, and Signal Proc.*, vol. ASSP-22, pp. 135–141, 1974.

[8] L.R. Rabiner and B.H. Juang, *Fundamentals of Speech Recognition*, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[9] J.D. Markel and A.H.Gray, *Linear Prediction of Speech*, Springer-Verlag, 1976.

[10] J.W Picone, "Signal Modeling Techniques in Speech Recognition," *Proceedings of the IEEE*, vol. 81 (9), pp. 1215–1247, 1993.

[11] A. Harma and U.K. Laine, "A Comparison of Warped and Conventional Linear Predictive Coding," *IEEE Trans. on speech and audio proc.*, vol. 9 (5), pp. 579–588, 2001.

[12] M. Karjalainen, A. Harma, and U.K. Laine, "Realizable Warped IIR Filters and Their Properties," *Proc IEEE ICASSO-96*, vol. Munich, pp. 2205–2208, 1996.