SPECTRAL ANALYSIS OF SUBBAND FILTERED SIGNALS

B. Tang, A. Shen, G. Pottie, and A. Alwan
UCLA Dept. of Electrical Engineering
405 Hilgard Ave., Los Angeles, CA 90024
email: tang@ee.ucla.edu

ABSTRACT

In this study, a methodology for transform-based spectral analysis of subband filtered signals is developed. The methodology is based on performing the analysis on subband samples instead of on the input signal directly. Aliasing due to decimation is eliminated by including the effects of the adjacent subband in the analysis of frequencies near the filterbank transition regions. The frequency resolution and spectral leakage is nearly the same as if the transform had been performed on the input directly. In an M band filterbank, the analysis block length is reduced by a factor of M. This reduces the complexity of source compression techniques based on subband decomposition and spectral analysis, such as the high quality speech coder we are developing [1] and the ISO/MPEG audio codec[2].

1. Introduction

Various speech and audio compression algorithms utilize properties of the human auditory system to minimize perceptual distortion [1]-[4]. Techniques employing transform coders or subband filters separate the source signal's spectrum into multiple bands and use masking properties to spectrally shape the distortion thus optimizing perceptual quality of the coded signal.

Spectral analysis is typically done by performing a transform of the source signal, but we will show that this can be done more efficiently by operating on the subband samples instead.

The methodology presented here uses the anti-aliasing properties of QMF filterbanks to eliminate the errors introduced in the transition region, thus making subband spectral analysis feasible. The resulting analysis is nearly identical to performing spectral analysis on the source signal directly and is achieved at a greatly reduced complexity.

2. Methodology

Spectral analysis can be viewed as the correlation of the sampled input source signal with an analysis sequence to determine coefficients as follows:

$$Y_k = \sum_{n=-\infty}^{\infty} x(n)a_k^*(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega})A_k^*(e^{j\omega})d\omega$$

where $x(n)$ is the input signal, $a_k(n)$ is the k-th analysis sequence, $X(e^{j\omega})$ and $A_k(e^{j\omega})$ are their Fourier transforms respectively, and Y_k is the k-th analysis coefficient.

The properties of $A_k(e^{j\omega})$ determine the spectral analysis characteristics. Ideally, as in a DFT analysis, $A_k(e^{j\omega})$ should be a function of the form $\delta(\omega-\omega_k)$, so that for $\omega \neq \omega_k$, Y_k is independent of $X(e^{j\omega})$. In practice, however, $a_k(n)$ must be a finite length (windowed) sequence so that $A_k(e^{j\omega})$ approximates the delta function by having a narrow main lobe with attenuated side lobes.

Consider now the 2 band QMF subband filterbank with the structure shown in Figure 1.

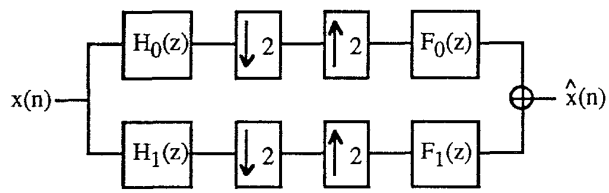

Fig. 1. 2 Band QMF Filterbank

The I/O relationship of the QMF filterbank is given by:

$$\hat{X}(z) = \tfrac{1}{2}\left[H_0(z)F_0(z) + H_1(z)F_1(z)\right]X(z) + \tfrac{1}{2}\left[H_0(-z)F_0(z) + H_1(-z)F_1(z)\right]X(-z)$$

The amplitude and phase distortion introduced by the QMF is given by the first term, the distortion transfer function, defined by:

$$T(z) = H_0(z)F_0(z) + H_1(z)F_1(z)$$

The aliasing introduced by the decimation is given by the second term, so the alias canceling constraint requires that:

$$H_0(-z)F_0(z) + H_1(-z)F_1(z) = 0$$

In subband spectral analysis, the analysis is done on the decimated subband samples, as depicted graphically in Figure 2.

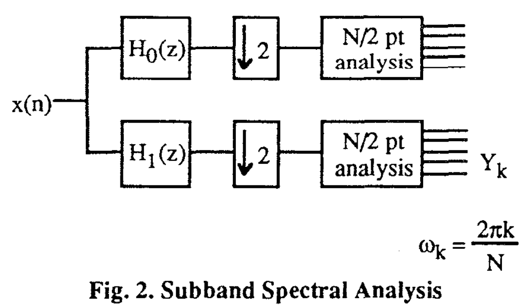

Fig. 2. Subband Spectral Analysis

The analysis coefficient Y_k is given by:

$$Y_k = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} x(m)h_i(n-m)w_k^*(n)$$
$$= \sum_{m=-\infty}^{\infty} x(m) \sum_{n=-\infty}^{\infty} h_i(n-m)w_k^*(n)$$

so the analysis sequence is now given by

$$a_k^*(n) = w_k^*(n) * h_i(-n)$$

where $w_k(n)$ is the decimated k-th analysis sequence and $h_i(n)$ is the subband impulse response of the band.

The analysis response $A_k(e^{j\omega})$ is the product of the decimated analysis transfer function consisting of two aliased "main" lobes located at ω_k and $\omega_{k'}=\pi+\omega_k$, and the subband conjugate response. One of the lobes corresponds to the analysis frequency and the other corresponds to the aliased frequency. Ideally, the subband filter response would leave one of the lobes unaltered and would reject the second lobe.

However, given the finite transition regions and out-of-band rejection, errors are introduced in the analysis. If the frequency to be analyzed is near a transition region, amplitude error is introduced due to attenuation in the passband edge and aliasing error is introduced due to the limited attenuation in the stopband edge. For this reason, subband spectral analysis has not been considered to produce acceptable results[2].

However, the subband spectral analysis methodology can be modified such that the results are acceptable. For analysis frequencies not near the transition region, the errors may be acceptable since there is less attenuation of the analysis frequency in the passband and more attenuation of the aliasing frequency in the stopband. Near the transition region, the properties of the QMF filterbank can be used to cancel out the errors as follows.

Let $h_i(n)$ and $f_i(n)$ be the impulse responses and $H_i(e^{j\omega})$ and $F_i(e^{j\omega})$ the frequency responses of the analysis and synthesis filterbank, respectively, for subband i. Consider now the impulse response

$$g_k(n) = F_i(e^{j\omega_k})h_i(n) + F_j(e^{j\omega_k})h_j(n)$$

with frequency response

$$G_k(e^{j\omega}) = F_i(e^{j\omega_k})H_i(e^{j\omega}) + F_j(e^{j\omega_k})H_j(e^{j\omega})$$

If the filterbank is designed to cancel aliasing and has no amplitude distortion, then $|G_k(e^{j\omega})| = 1$ for $\omega=\omega_k$, and $|G_k(e^{j\omega})| = 0$ for $\omega=\omega_{k'}$, where ω_k is the analysis frequency in subband i and $\omega_{k'}$ is the aliasing frequency in subband j.

Note that even in the presence of amplitude distortion and aliasing in the subband filterbank design, the errors in the subband spectral analysis are minimized by minimizing the amplitude and aliasing errors in the filterbank.

This observation is useful in extending this methodology to the M band filterbank case. Rather than considering all M bands, only the adjacent band needs to be considered. For non-adjacent bands, the stopband attenuation ensures that the attenuation and aliasing error contribution from that band is minor.

For subband spectral analysis, the aliasing and amplitude distortion introduced at $\omega_{k'}$ can be canceled by modifying the analysis sequence to

$$a_k^*(n) = w_k^*(n) * g_k(-n)$$
$$= F_i(e^{j\omega_k})(w_k^*(n) * h_i(-n)) + F_j(e^{j\omega_k})(w_k^*(n) * h_j(-n))$$

This method of modified analysis is denoted alias-canceled subband spectral analysis. Figure 3 depicts the operations to evaluate Y_k, the k-th coefficient of an N point analysis in an M band filterbank example.

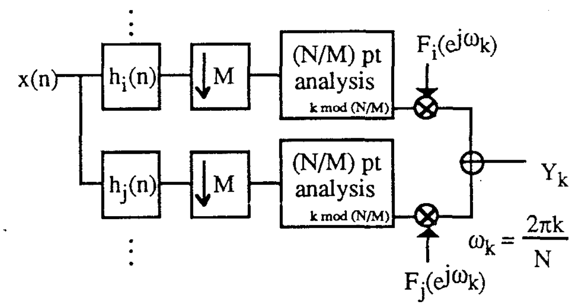

Fig. 3. Alias-Canceled Subband Spectral Analysis

To summarize, in order to perform subband spectral analysis with no aliasing errors, the only additional operation required is a linear combination of the coefficients at ω_k and $\omega_{k'}$ with the non-modified analysis sequences. In this manner, the analysis coefficients can be obtained by performing a shorter block analysis on the subband samples rather than on the input signal directly. Windowing and various transform based spectral analysis methods such as the DFT and the DCT are easily adaptable to subband analysis with this method.

3. Design Example and Results

To illustrate the effectiveness of this method, we compared the analysis response of a 256 point rectangular window DFT, the response of a 32 point subband DFT based on an 8 band, zero amplitude distortion, tree structured IIR QMF filterbank [6], and the response of the alias-canceled subband DFT for the same filterbank structure.

The structure and frequency response for the filterbank are shown in Figures 4 and 5. Note that decimation results in spectral inversion of the highpass band. This is represented by the crossovers in the block diagram shown in Figure 4. The tree structure of the QMF filter also leads to differences in the transition regions, with the least sharp transition at $\pi/2$.

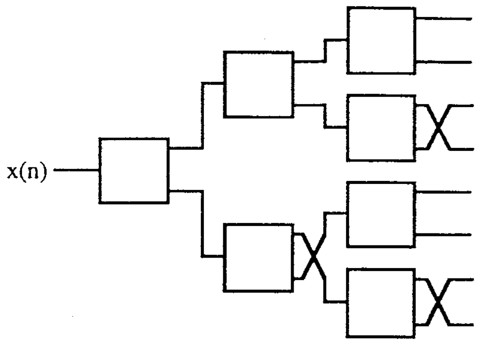

Fig. 4. 8-band QMF tree structure

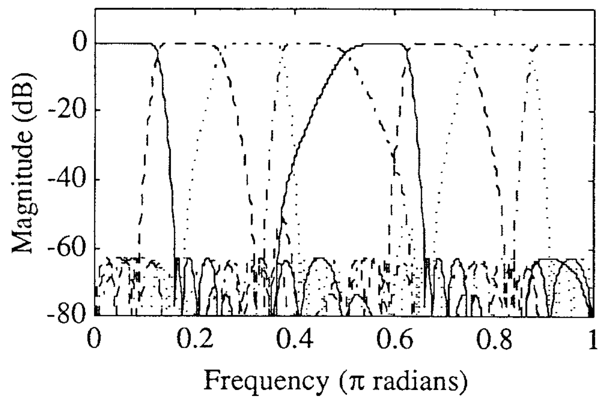

Fig. 5. 8-band QMF Frequency Response

The spectral analysis response was computed as a power sum of the analysis at ω_k and $2\pi-\omega_k$ (i.e., $|A_k(e^{j\omega})|^2 + |A_k(e^{j(2\pi-\omega)})|^2$). This makes the alias frequency more easily noticeable in the frequency response since it then simply appears as the image of the analysis frequency. In Figures 6-9, the analysis frequency is $\pm 63\pi/128$ and alias frequency is $\pm 65\pi/128$ (k=63, analysis subband i=3, aliasing subband j=4, N=256).

The non alias-canceled subband DFT analysis response (Figure 6) results in both amplitude distortion (the main lobe at $63\pi/128$ is at -3 dB) and considerable leakage (the adjacent lobe at $65\pi/128$ is at -6dB). Note that there is very little leakage from the other bands due to the stopband response of the subband filter.

The alias-canceled subband DFT analysis response (Figure 7) corrects both amplitude and aliasing distortion. The main lobe is at 0 dB and the adjacent lobe is at -13 dB, with a null at $\omega = 65\pi/128$. Compared to the direct DFT analysis response (figure 8), there is little difference in amplitude and aliasing errors. In addition, there is rejection of spectral components in the other subbands, since they have been "filtered" out of the subband samples. In practice, this is only a minor advantage, since the spectral leakage is dominated by frequencies closer to the analysis frequency.

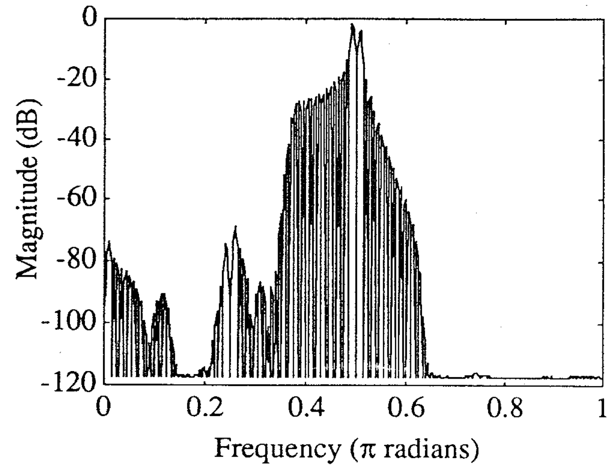

Fig. 6. Non Alias-Canceled Subband Analysis Response

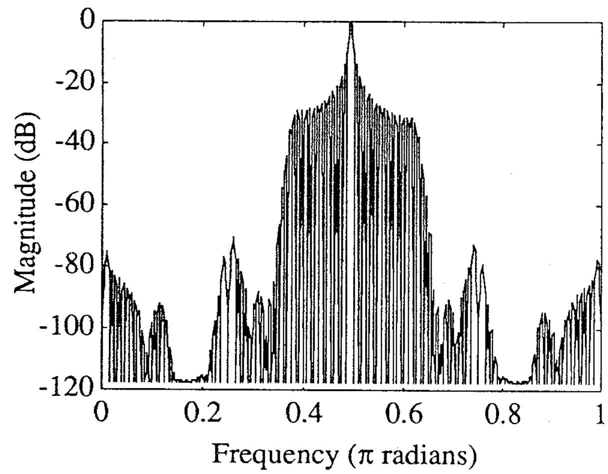

Fig. 7. Alias-Canceled Subband Analysis Response

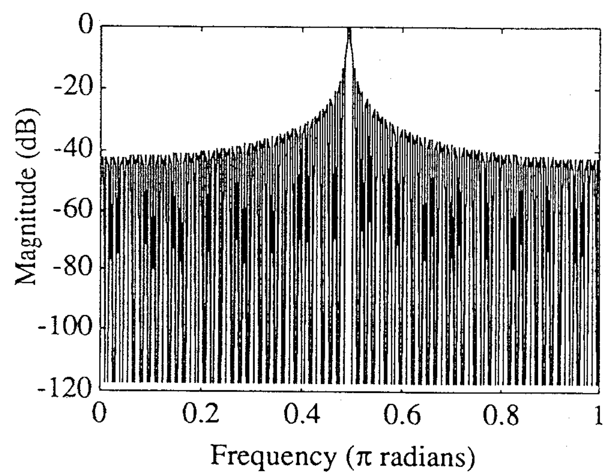

Fig. 8. Direct DFT Analysis Response

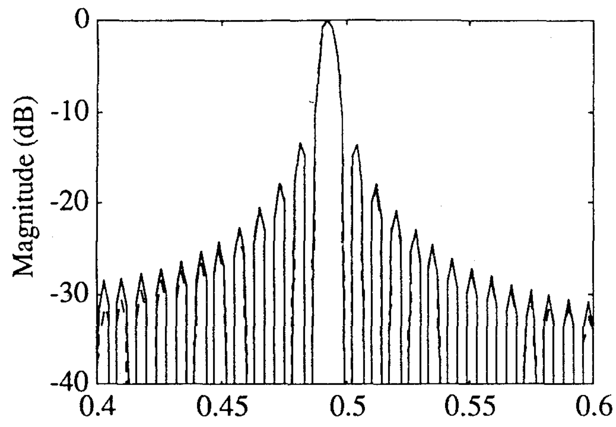

Fig. 9. Comparison of 256 pt direct DFT (dashed) and Alias-Canceled 32 pt subband DFT (solid) Analysis Responses

Figure 9 shows a more detailed comparison of the alias-canceled subband DFT and the direct DFT, illustrating that the analysis results are essentially identical for both methods.

This example demonstrates how a simple methodology involving a linear combination of coefficients obtained by a DFT of subband samples achieves essentially the same results as obtained from a DFT of the input samples directly.

4. Conclusion

A methodology for performing transform-based subband spectral analysis has been presented. The methodology uses the properties of QMF filterbanks to eliminate the alias and amplitude distortion errors resulting from subband spectral analysis.

If our method is not used, the recourses are to either perform the spectral analysis on the source sequence directly or perform the analysis on the subband samples and accept the distortion introduced in the transition region. The former requires higher computational complexity, while the latter leads to inadequate results due to the errors introduced in the analysis.

Our method allows one to achieve the desired accuracy at reduced complexity. With a fast transform such as the FFT, the complexity is reduced from N/2 log N to N/2M log N/M for an M band coder. The reduction in complexity is greater if analysis is required only over a limited interesting region, ignoring one or more of the subbands.

Transform-based spectral analysis plays an important role in speech and audio compression algorithms. They are applied widely in subband coders, transform coders, and hybrid subband/transform coders, such as the perceptually-based coders proposed by Brandenburg [2], Johnston [3], and Dolby Laboratories [4]. The large transforms required in these coders can benefit from our method. For example, Johnston's transform coder for 15 KHz audio uses a 2048 point FFT in 60 ms frames. Dolby AC-2, a transform coder for 20 KHz audio, uses 128 and 512 point MDCT with adaptive block lengths. The ISO/MPEG Audio Codec, a hybrid subband/transform coder for 15-22 KHz, uses 512 or 1024 point FFTs in 12 ms frames.

In the MPEG audio codec layers 1 and 2, the 512 point FFT is used to estimate the power spectrum and auditory masking effect to determine appropriate bit allocations for the 32 different subbands. Significant reduction in complexity can be achieved by utilizing the approach proposed in this paper. If the entire spectrum is computed, the complexity is reduced by 55%. If only one of the subbands needs to be analyzed then the complexity is reduced by 98.5 %.

The computational and storage requirements for implementing these large-size transforms may be a significant burden on the system, leading the designer to pursue less complex, lower performance techniques. Therefore, reducing the computational complexity is of significant importance in making the next generation, improved performance compression algorithms for applications such as Personal Communication Systems a reality.

5. References:

[1] A. Shen, B. Tang, A. Alwan, G. Pottie, "A Robust Variable Rate Speech Coder", ICASSP 1995.

[2] K. Brandenburg and G. Stoll, "The ISO/MPEG Audio Codec: A Generic Standard for Coding of High Quality Audio", 92nd AES Convention, 1992.

[3] J. D. Johnston, "Transform Coding of Audio Signals Using Perceptual Noise Criteria", IEEE JSAC, vol. 6, No. 2, February 1988.

[4] G. Davidson and M. Bosi, "AC-2: High Quality Digital Audio Coding for Broadcasting and Storage", Dolby Technical Papers Publication No. S92/9433, 1992.

[5] P. P. Vaidynathan, Multirate Systems and Filterbanks, Prentice Hall, 1993.

[6] Z. Jiang, A. Alwan, and A. Willson, "High Performance IIR QMF Banks for Speech Subband Coding", IEEE ISCAS, June 1994.

This work was supported in part by ARPA/CSTO contract J-FBI-93-112 and by NSF grant IRI-9309418.