OVERSAMPLED FILTER BANK EVALUATION FOR JOINT SUBBAND AUDIO PROCESSING AND CODING

David Hermann
AMI Semiconductor
Waterloo, Ontario, Canada

Robert D. Dony, Shawki M. Areibi
School of Engineering
University of Guelph
Guelph, Ontario, Canada

ABSTRACT

This paper presents an experimental evaluation of oversampled, modulated filter banks for joint subband audio processing and coding applications. Joint subband processing and coding may be useful in some wireless audio devices such as advanced wireless digital hearing aids. We examine the use of oversampled GDFT and cosine modulated filter banks and propose using single sideband (SSB) real-valued filter banks as a compromise which is ideal for this application. The SSB filter bank provides real-valued signals for coding which are free from any aliasing cancellation constraints and hence are also suitable for non-uniform subband processing such as subband gain adjustment. We support this conclusion with an experimental analysis of various filter bank designs for subband gain adjustment and subband audio coding.

Index Terms— channel bank filters, hearing aids, audio systems

1. INTRODUCTION

Oversampled filter banks allow for subband processing with greater flexibility than critically sampled filter banks. One key advantage for audio processing with oversampled filter banks is their ability to offer improved performance with lower delay [1, 2]. For a given group delay requirement, an oversampled filter bank can provide reduced levels of aliasing in each subband as compared to an equivalent critically sampled filter bank. This is because the filter requirements for an oversampled filter bank are less stringent than those for a critically sampled filter bank, and hence shorter filters can be used to yield a similar level of performance. This key advantage, coupled with the availability of low-complexity filter bank implementations, have made the oversampled filter bank a popular technology in low delay and constrained resource audio devices such as digital hearing aids [1, 3]. In hearing aids, the processing delay is particularly important because a perceptible delay between when an audio event occurs and when it is output by the hearing aid can be discomforting to the hearing aid user [4]. This low delay requirement can also be applicable to other personal audio devices such as assistive listening devices and wireless headsets for telecommunications applications.

The increasing use of wireless technology in digital hearing aids [1] and hearing aids [4] has made the oversampled filter bank a popular technology in low delay and constrained resource audio devices such as digital hearing aids [1, 3]. In hearing aids, the processing delay is particularly important because a perceptible delay between when an audio event occurs and when it is output by the hearing aid can be discomforting to the hearing aid user [4]. This low delay requirement can also be applicable to other personal audio devices such as assistive listening devices and wireless headsets for telecommunications applications.

The increasing use of wireless technology in digital hearing aids creates a need to combine subband audio processing and audio coding while still taking into account the low delay, low complexity and low power constraints of the application. In this paper, we evaluate a number of oversampled uniform, modulated filter banks for their performance in subband audio processing and subband audio coding. Various kinds of complex modulated filter banks (both critically sampled and oversampled) have been proposed for subband audio processing including subband adaptive filtering and subband gain adjustment applications such as spectral subtraction type noise reduction algorithms. Meanwhile, various kinds of cosine modulated filter banks including lapped transforms have become the standard in subband audio coding. However, the use of oversampled filter banks for coding has generally been avoided because oversampling implies an unnecessary increase in the data rate. Oversampling incurs a rate-distortion penalty, but this may be mitigated by techniques such as predictive coding and noise shaping [5, 6]. Some other potential benefits of oversampled filter banks in coding are also now being investigated, including their error robustness properties (see e.g. [7, 8]), but these properties are not considered in this paper. This evaluation will demonstrate that a near-perfect reconstruction (NPR) single sideband (SSB) oversampled filter bank can provide the benefits of NPR complex modulated filter banks while providing real-valued subband signals like cosine modulated filter banks. Thus, this filter bank is a good candidate for use in joint subband audio processing and coding. Joint audio processing and coding represents a novel use for this type of filter bank.

The remainder of this paper is organized as follows. In Section 2 we begin by briefly reviewing the concept of uniform, modulated filter banks and describe how oversampled and critically sampled filter banks have been used in subband audio processing and subband audio coding applications. In Section 3 we further review two classical filter bank designs and describe a type of single sideband modulated filter bank which combines the benefits of the two classical designs while avoiding their major disadvantages. In Section 4 we present prototype filter designs for the filter banks based on a simple windowed FIR filter design approach. In Sections 5 and 6 we perform an experimental evaluation of various filter bank designs for subband audio processing and subband audio coding, respectively. Finally, Section 7 summarizes the evaluation results and describes the directions for our future research.

2. BACKGROUND

In this section, we will briefly review the concept of a uniform, modulated filter bank, and its applications in audio processing and audio coding. Uniform, modulated filter banks with FIR prototype filters are particularly useful because of their ease of design and computational efficiency [9]. Also, low-power efficient implementations of these filter banks are already available for the target application of wireless digital hearing aids [1]. A block diagram view of a uniform modulated filter bank is shown in Figure 1. The ideal filterbank uses the analysis filters \( h_k(n) \) to divide the input signal, \( x(n) \), into \( N \) non-overlapping frequency regions. For a uniform filter bank, the nominal bandwidth of each analysis filter is \( \pi/N \). Because they are now band-limited, the outputs of the analysis filters can be downsam-
For an equivalent complex-valued filter bank, there are $N$ banks; this normally means there are exactly $N$ subbands of the input frequency spectrum.

Researchers have demonstrated improved results in psychoacoustic applications such as multi-band dynamic range compression, noise reduction and subband adaptive filtering for echo cancellation [16, 14]. The analysis and synthesis filters in a GDTF filter bank are defined by equations (1) and (2) respectively, where $k_0$ represents an arbitrary frequency offset, $n_0$ represents an arbitrary time offset, $h(n)$ is the analysis prototype filter and $g(n)$ is the synthesis prototype filter [22]. We select $k_0 = 0.5$ and $n_0 = 0$ to produce an odd stacking arrangement of the filters which corresponds to the frequency decomposition of a typical cosine modulated filter bank as described below. Although it has proven performance in subband audio processing applications, the oversampled GDTF filter bank is non-ideal for coding applications due to the complex-valued subband signals which are more complicated to quantize efficiently.

$$h_k(n) = h(n)e^{j\frac{2\pi}{N}(k+k_0)(n+n_0)}, \quad k = 0 \ldots 2N - 1 \quad (1)$$

$$g_k(n) = g(n)e^{j\frac{2\pi}{N}(k+k_0)(n+n_0)}, \quad k = 0 \ldots 2N - 1 \quad (2)$$

The second candidate filter bank to be considered is an oversampled version of the pseudo-quadrature mirror filter (PQMF) filter bank. The analysis and synthesis filters for this filter bank are expressed as cosine modulated replicas of a single prototype filter, $p(n)$, as described by equations (3) and (4). This filter bank employs adjacent band aliasing cancellation to achieve near-perfect reconstruction [9]. The traditional, critically sampled PQMF filter bank is used as part of the industry standard MP3 codec [23] and is a pre-cursor to more recent PR filter banks such as the modified discrete cosine transform (MDCT) based lapped orthogonal transform. By using an oversampled PQMF filter bank, we hope to gain the advantages of an oversampled filter bank but retain the real-valued subband signals which are amenable to sophisticated compression methods. However, as we will demonstrate in the evaluation section of this paper, the aliasing cancellation employed by this filter bank is still a disadvantage for subband audio processing.

$$h_k(n) = 2p(n)\cos\left(\frac{\pi}{N}(k + \frac{1}{2})(n - \frac{L}{2}) + (-1)^k \frac{\pi}{4}\right), \quad k = 0 \ldots N - 1 \quad (3)$$

$$g_k(n) = 2p(n)\cos\left(\frac{\pi}{N}(k + \frac{1}{2})(n - \frac{L}{2}) - (-1)^k \frac{\pi}{4}\right), \quad k = 0 \ldots N - 1 \quad (4)$$

**3. CANDIDATE FILTER BANKS**

In choosing candidate filter banks for use in a joint subband audio processing and coding system, we focus only on near-perfect reconstruction (NPR) filter banks which minimize the amplitude and phase distortion but do not necessarily provide perfect reconstruction (PR) through complete alias cancellation. When the filter bank analysis outputs are subject to subband signal processing involving substantial spectral modifications before re-synthesis, the PR property is inevitably destroyed and hence it is better to create an NPR design which will usually have lower in-band aliasing and better separation between the subbands.

The first candidate filter bank to be evaluated is the oversampled, complex modulated, generalized DFT (GDTF) filter bank. This filter bank has been shown to work well in many audio processing applications such as audio compression, perfect-reconstruction (PR) critically sampled filter banks are preferred because they provide the benefits of subband coding with minimal data rates and no aliasing distortion. There are only a few examples of subband audio compression schemes which use oversampled subbands, see e.g. [19, 20]. One potentially interesting benefit of oversampled subbands for audio coding is when using perceptual audio coding. Some researchers have demonstrated improved results in psychoacoustic models when there is less aliasing present in the frequency domain signals being used to generate the model [21].

![Fig. 1. General Model of a Uniform Modulated Filter Bank](image-url)
Our final candidate filter bank is the single sideband (SSB) real-valued filter bank, which is so-called because of its similarity to single sideband modulation in communications systems. We create this filter bank by combining the conjugate pairs of the filter bank channels in a GDF filter bank having twice the desired oversampling rate of the SSB filter bank. To simplify the discussion, we first consider a modified version of the GDF filter bank where a post-modulation is applied to the subband channels in order to ensure that every decimated subband signal is demodulated to DC in the decimated subband domain \[22\]. Such a filter bank channel therefore looks like Figure 3. It is interesting to note that the modulation terms \(e^{-j\frac{\pi}{2}m}\) correspond to simple multiplications by \(+1\), \(-1\), \(+j\) and \(-j\) so that the overall SSB conversion can be implemented efficiently using careful selection and negation of the real or imaginary parts of the modified GDF filter bank analysis outputs.

\[
\hat{X}_k(m) = X_k(m)e^{j\frac{\pi}{2}m} + X_{2N-k-1}(m)e^{-j\frac{\pi}{2}m} = X_k(m)e^{j\frac{\pi}{2}m} + \hat{X}_k^*(m)e^{-j\frac{\pi}{2}m} = 2\text{Re}\{X_k(m)e^{j\frac{\pi}{2}m}\} \tag{5}
\]

Another view of the SSB filter bank is shown in Figure 4 based on the GDF analysis filters. This figure shows how by modulating the GDF subband signals to \(\pm\frac{\pi}{2}\), the resulting real valued channel has maximum separation between the positive and negative frequency components and hence exhibits minimum in-band aliasing. This type of SSB filter bank is similar to classical SSB filter banks \[22\] and is mathematically equivalent to the SSB filter bank described in \[11\] for real-valued subband adaptive filtering with minimized in-band aliasing. We propose to use this SSB filter bank as a compromise between the GDF filter bank for subband audio processing and cosine modulated filter banks for subband audio coding.

One remaining aspect of the filter bank designs to be considered is the value of parameters such as the prototype filter length, number of channels and oversampling ratio. In this work, we are interested in low-delay (ideally \(< 10\) ms) filter banks with sufficient frequency resolution for signal adjustment in a typical hearing aid (i.e. frequency resolution equal to that of the smallest audiological frequency difference used in hearing aid fitting). We will therefore focus our evaluation on filter lengths corresponding to delays in the range of \(9 - 11\) ms and a number of channels \(N = 32\) which provides 250 Hz frequency resolution if we assume a 16 kHz sampling rate for hearing aid applications. Finally, we will consider a number of integer oversampling ratios in the experiments to follow.

\[4. \text{ PROTOTYPE FILTER DESIGN}\]

A single prototype filter design method is used for evaluation of all filter bank designs in this paper. The method we use is based on windowed symmetric FIR filter design and seeks to design a filter which provides minimal amplitude distortion, i.e. a filter which is power-complementary to modulated copies of itself \[24, 2\]. The resulting filters can be used for complex modulated or cosine modulated NPR filter banks and can be easily re-designed for a wide variety of filter bank configurations. An additional extension to the method from \[2\] was implemented for the purposes of generating windows in this evaluation. The Kaiser window was used to allow for an additional trade-off between transition bandwidth and stop-band attenuation of the filter through the window’s \(\beta\) parameter \[25\]. The original method optimizes over the windowed FIR filter’s cut-off frequency in order to find low-pass filter which is power-complementary with modulated copies of itself. The modification to this method performs this same optimization by adjusting both the filter’s cut-off frequency and the \(\beta\) parameter of the Kaiser window. The use of the Kaiser window and its adjustable parameter is motivated by its successful use in other, critically sampled NPR filter
ball design techniques [24].

The windows designed with this method for $N = 32$ and $L = \{128, 160, 192\}$ are shown in Figure 5. The cut-off points of $\pi / N$ and $\pi / 2N$ are also indicated in the figure by dashed lines. It is clear that the filters achieve faster roll-off and greater stopband attenuation for higher filter orders as expected. Another item of note is the substantial overlap between bands these filters generate. Although they reach a $-3$ dB cut-off at approximately $\pi / N$, the do not reach the first sidelobe until past $2\pi / N$ implying at least 50% overlap between adjacent bands. Therefore, to obtain reasonably alias-free processing of each band we need to consider oversampling ratios of at least two.

5. SUBBAND AUDIO PROCESSING EVALUATION

In order to evaluate the subband audio processing performance of the filter bank designs, we would like to compare the filter bank output to some reference desired output. In the case of subband gain adjustment for noise reduction or multi-band dynamic range compression, it is difficult to define this reference because the desired output signal is substantially different from the actual input signal. We propose to compare each filter bank performance’s to an equivalent non-oversampled (i.e. maximally oversampled) GDFT filter bank using the same prototype filter and number of channels. The non-oversampled filter bank output represents the “best” possible performance for a particular filter bank configuration and prototype filter design and can be evaluated under specific subband processing conditions (e.g. non-unity subband gains). Any additional distortion between the oversampled filter bank being evaluated and the non-oversampled filter bank for the same subband processing is a result of the the subsampling (i.e., aliasing and imaging). For simplicity, we propose that a global metric such as signal-to-noise ratio (SNR) calculated between the non-oversampled filter bank output and the oversampled filter bank output can provide a first-order assessment of the distortion. If the non-oversampled filter bank output for a finite-length signal is denoted by $y_o(n)$ for $n = 0 \ldots L - 1$, and the subsampled filter bank output is denoted by $y(n)$, then the SNR difference between the signals can be defined by equation (6).

$$SNR_{dB} = 10 \log_{10} \left( \frac{\sum_{n=0}^{L-1} \| y_o(n) \|^2}{\sum_{n=0}^{L-1} \| y_o(n) - y(n) \|^2} \right)$$

In this evaluation method, arbitrary target subband gains were used which alternated from $-6$ dB to $+6$ dB in a comb-filter type effect. Thus, the gain difference from band-to-band is $12$ dB, which may still be considered small relative to possible gains in hearing aids of $30 - 40$ dB [26]. The output SNR measurements of each oversampled filter bank relative to the non-oversampled reference filter bank using the same gains and a small range of power-of-two oversampling ratios are shown in Figures 6, 7 and 8. The first conclusion of note for this data is that the cosine modulated filter bank performs poorly for any filter length due to its failure to sufficiently cancel the aliasing between adjacent bands when subband gain adjustments are applied. That is, the aliasing cancellation property of the cosine modulated filter bank is lost when the subband signals are substantially modified. When looking at the GDFT and SSB results, we see similar results for most filter lengths and oversampling ratios. We can conclude from these measurements that the cosine modulated filter bank is not suitable for subband audio processing even when the subband signals are oversampled. Furthermore, the GDFT and SSB filter banks have comparable performance for subband gain adjustment. The choice of prototype filter length or oversampling ratio may depend on the specific target system requirements, but a prototype length of $L = 160$ and an oversampling ratio of $K = 4$ appears to provide a good tradeoff between delay, oversampling ratio and performance for this particular example.

6. SUBBAND AUDIO CODING EVALUATION

Ultimately, the subband coding performance is a function of the filter bank, the coding system used (e.g. the bit allocation) and the input signals probability density function (pdf) [27]. Therefore, it is difficult to definitively evaluate the performance of a particular filter bank design in terms of the coding performance. We propose experimentally evaluating the filter bank’s performance using synthetically generated signals. In particular we will use white noise filtered through a digital filter having a magnitude frequency response corresponding roughly to that of a long-term average speech spectra (LTASS). An auto-regressive, moving-average (ARMA) model for the LTASS magnitude response was generated based on experimental LTASS data from [28]. Simple uniform scalar quantization was used in each subband with a fixed bit allocation which is optimized for the LTASS frequency response. For simplicity, we applied optimal subband bit allocation principles from classic coding theory in order to find the bit allocation to use with the synthetic test signal (see e.g., [27]). Each candidate filter bank design was evaluated in terms of rate-distortion performance using this experimental subband coding system. When comparing the performance of the oversampled filter bank designs, we used an equivalent critically sampled, PQMF cosine modulated filter bank for reference. Since it is generally simpler to quantize using real-valued signals, we examined the performance of only the oversampled cosine modulated and oversampled SSB filter banks for the $L = 160$, $K = 4$ ($M = 8$) filter bank designs which were identified in the audio processing evaluation in Section 5. In this case, the oversampled cosine modulated filter bank provides another benchmark for coding performance but we are primarily interested in the performance of the SSB filter bank since it was demonstrated to be suitable for audio signal processing.

The results of an experimental rate-distortion evaluation using the synthetic input and test setup described above are shown in Figure 9. The reference, critically sampled filter bank clearly outperforms the oversampled filter banks as expected. It is important to note that the overall reconstruction SNR of the critically sampled case approaches a limit because this filter bank is still not PR. A
more important result is that the rate-distortion performance of the oversampled cosine modulated filter bank and the oversampled SSB filter bank are identical for moderate to high bit rates. The SSB filter bank actually outperforms the cosine modulated filter bank at very low bit rates. This can be explained by the cosine modulated filter bank failing to sufficiently cancel the aliasing when some subbands are no longer coded at all (i.e. when the optimal bit allocation is zero for some subbands). Our general conclusion from these experimental results is that the SSB filter bank should provide equivalent or better rate-distortion performance than an equivalent oversampled cosine modulated filter bank in subband coding applications.

7. CONCLUSION AND FUTURE WORK

In this paper, we have reviewed uniform complex and cosine modulated filter banks designs for their suitability in both subband audio processing (e.g. subband gain adjustment) and subband audio coding applications. A real-valued filter bank based on single sideband modulation of GDFT filter bank channels was described that provides real-valued subband signals with no aliasing cancellation constraints. Especially for applications requiring both subband coding and substantive spectral modifications, this filter bank provides the benefits of both the GDFT and cosine modulated filter bank designs while avoiding their major disadvantages.

A variety of filter bank designs were then evaluated for subband audio processing performance using arbitrary, non-uniform subband gain adjustment as an experimental use case. The results confirm that the complex-modulated GDFT filter bank provides good performance under subband gain adjustment due to its lack of aliasing cancellation constraints. The cosine modulated filter bank performs poorly, even when oversampled, due to the adjacent band aliasing cancellation property which is violated during subband adjustments. The SSB filter bank was shown to provide performance close to that of an equivalently oversampled GDFT filter bank.

The oversampled cosine modulated and SSB filter bank designs were then assessed for their rate-distortion performance in a subband coding experiment using an LTASS modelled input. The performance was compared to an equivalent, critically sampled cosine modulated filter bank. The critically sampled filter bank outperforms the oversampled filter banks substantially, as expected. The oversampled cosine modulated and SSB filter banks show similar rate-distortion performance, thereby supporting our choice of the SSB
filter bank as an ideal design for joint subband audio processing and coding.

With a filter bank design established, the next step in our research is to focus on subband coding methods which will help to mitigate the increased data rate in oversampling, and hence improve the overall rate-distortion performance of audio coding using oversampled filter banks. Predictive coding methods in oversampled filter banks have been previously studied and could be useful in this application [5, 6]. The goal is to develop a complete oversampled filter bank-based audio codec suitable for wireless digital hearing aids and other subband audio processing applications where processing or delay constraints do not allow for a separate critically sampled filter bank.

8. REFERENCES


