LOW-DELAY COSINE-MODULATED QMF BANK FOR MPEG AUDIO COMPRESSION

H. Xu†, *W.-S. Lu*‡, *and A.Antoniou*‡

† INRS-Telecommunications, 16 Place Du Commerce
Nun's Island, Verdun, Quebecc, Canada H3E 1H6
‡ Dept. of Electrical and Computer Engineering, University of Victoria
Victoria, B.C., Canada, V8W 3P6
Email: andreas@ece.uvic.ca

ABSTRACT

A 32-channel, cosine-modulated QMF bank whose reconstruction delay is about half that of the current MPEG filter bank is proposed. The low-delay QMF bank is then realized in terms of a fast polyphase structure.

1. INTRODUCTION

The Moving Picture Expert Group (MPEG) has proposed an international standard [1] for the compression of high-bit-rate moving pictures and associated audio signals. In layers I and II of the current MPEG audio compression scheme, a 32-channel cosine-modulated quadrature mirror-image filter (QMF) bank is used to realize the time-frequency mapping and a fast polyphase structure is available for implementing this QMF bank. Regardless of the coding process, the reconstruction delay of the filter bank is $512T$, where T is the sampling period.

In this paper we propose a 32-channel, low-delay, cosine-modulated QMF bank which has a reconstruction delay about half that of the current MPEG filter bank. We then realize the low-delay QMF bank in terms of a fast polyphase structure.

2. A 32-CHANNEL LOW-DELAY COSINE-MODULATED QMF BANK

A low-delay cosine-modulated QMF bank based on the multi-channel filter bank shown in Fig. 1 can be designed as follows [2]. The frequency response of the prototype filter P_L is $P_L(e^{j\omega}) = |P_L(e^{j\omega})|\, e^{-j\omega k_d/2}$ where $k_d/2 < (N-1)/2$ is the group delay and N is the filter length, and the impulse responses of the analysis filters H_k and the synthesis filters F_k are given by

$$h_k(n) = 2p_L(n)\cos\left[(2k+1)\frac{\pi}{2M}\left(n-\frac{k_d}{2}\right)-(2k+1)\frac{\pi}{4}\right]$$
$$f_k(n) = 2p_L(n)\cos\left[(2k+1)\frac{\pi}{2M}\left(n-\frac{k_d}{2}\right)+(2k+1)\frac{\pi}{4}\right]$$

for $0 \le n \le N-1$ and $0 \le k \le M-1$, where M is the channel number and $p_L(n)$ is the impulse response of the prototype filter P_L. The reconstruction delay is k_d which, in principle, could be any integer less than $N-1$.

An iterative method to design a low-delay filter bank with $M = 32$, $N = 513$, and $k_d = 255$ was proposed in [3]. The performance of the filter bank obtained is evaluated in terms of the peak-to-peak reconstruction error $E_{p-p} = \max\limits_{\omega}|T_0(e^{j\omega})| - \min\limits_{\omega}|T_0(e^{j\omega})|$ where $|T_0(e^{j\omega})| = |\sum_{k=0}^{M-1} H_k(e^{j\omega})F_k(e^{j\omega})|$; the aliasing error $E_a = \max\limits_{\omega} E(\omega)$ where $E(\omega) = \frac{1}{M}[\sum_{l=1}^{M-1}|T_l(e^{j\omega})|^2]^{1/2}$ and $T_l(e^{j\omega}) = \sum_{k=1}^{M-1} H_k(e^{j(\omega-2\pi l/M)})F_k(e^{j\omega})$; and the signal-to-noise ratio $\mathrm{SNR} = 10\log_{10}\{\sum x^2(n)/\sum[x(n)-\hat{x}(n+k_d)]^2\}$ where the input $x(n)$ is a ramp signal. For comparison, the performance of the current MPEG QMF bank is also evaluated using the same criteria and the results are listed in Table 1 where T represents the sampling period. The amplitude responses of the analysis filters of the low-delay QMF bank and the current MPEG QMF bank are illustrated in Figs. 2 and 3, respectively. It can be noted that the proposed low-delay QMF bank has reduced the reconstruction delay by 50%, and E_a and SNR are comparable with those of the corresponding MPEG QMF bank but E_{p-p} is larger.

Table 1. Comparisons between the low-delay QMF bank and current MPEG QMF bank

	k_d	E_{p-p}	E_a	SNR (dB)
Low-delay QMF bank	$255T$	8.9669×10^{-4}	4.6497×10^{-6}	88.14
MPEG QMF bank	$512T$	1.6328×10^{-4}	2.7128×10^{-6}	84.34

2.1. An efficient polyphase implementation

Like the current MPEG QMF bank, the proposed low-delay QMF bank can be implemented by an efficient polyphase structure [1].

2.2. Analysis subband filtering

For convenience, coefficient $p_l(512)$ is dropped as in the MPEG QMF bank. At the lth instant, the subsampled signal in the kth channel, $x_k(l)$, is given by

$$x_k(l) = \sum_{n=0}^{N-1} h_k(n)x(lM-n)$$

for $k = 0, 1, \ldots, M-1$ where $N = 512$, $M = 32$, and $x(n)$ represents the input audio signal. Let $n = 2\gamma M + \rho$, for $\gamma = 0, 1, \ldots, 7$, $\rho = 0, 1, \ldots, 63$. By substituting n in the above equation, we obtain

$$\begin{aligned} x_k(l) &= \sum_{\rho=0}^{63}\sum_{\gamma=0}^{7} h_k(2\gamma M+\rho)x(lM-2\gamma M-\rho) \\ &= \sum_{\rho=0}^{63}\sum_{\gamma=0}^{7} 2\cos\left[(2k+1)\frac{\pi}{2M}\left(2\gamma M+\rho-\frac{k_d}{2}\right)\right. \\ &\quad \left. -(2K+1)\frac{\pi}{4}\right]\cdot p_L(2\gamma M+\rho)x(lM-2\gamma M-\rho) \\ &= \sum_{\rho=0}^{63}\cos\left[(2k+1)\frac{\pi}{2M}(\rho-16-\frac{k_d}{2})\right] \\ &\quad \cdot\sum_{\gamma=0}^{7} 2(-1)^{\gamma}p_L(64\gamma+\rho)x(32l-64\gamma-\rho) \end{aligned}$$

where $k = 0, 1, \ldots, 31$. It follows that the analysis subband filtering can be performed by the following steps:

- Input 32 audio samples.
- Construct an input sample vector $\mathbf{x}$ of 512 entries. Shift in 32 samples of the signal at positions 0 to 31 with the most recent one at position 0 and shift out the 32 oldest samples.
- Compute $\mathbf{z}_i = 2\cdot(-1)^{int[i/64]}\mathbf{x}_i\cdot p_L(i)$ for $i = 0, 1, \ldots, 511$.
- Compute $\mathbf{y}_i = \sum_{j=0}^{7} \mathbf{z}_{i+64j}$ for $i = 0, 1, \ldots, 63$.
- Compute $\mathbf{s}_i = \sum_{k=0}^{63} \mathbf{M}_{ik}\cdot\mathbf{y}_k$ for $i = 0, \ldots, 31$, where

$$\mathbf{M}_{ik} = \cos\left[(2i+1)(k-16-k_d/2)\pi/64\right] \qquad (1)$$

  for $i = 0$ to 31 and $k = 0$ to 63.
- Output $\mathbf{s}_i$ for $i = 0, \ldots, 31$ as the subband samples.

2.3. Synthesis Subband Filtering

The upsampled signal $\tilde{x}_k(n)$ and the subsampled signal $x_k(n)$ in the kth channel are related by

$$\tilde{x}_k(n) = \begin{cases} x_k(n) & n = lM \text{ for integer } l \\ 0 & \text{otherwise} \end{cases}$$

for $k = 0, 1, \ldots, M-1$. Hence the reconstructed signal can be expressed as

$$\hat{x}(n) = \sum_{k=0}^{M-1}\sum_{i=0}^{N-1} f_k(i)\tilde{x}_k(n-i)$$

If we let $y(n,i) = \sum_{k=0}^{M-1} f_k(i)\tilde{x}_k(n-i)$, then

$$\hat{x}(n) = \sum_{i=0}^{N-1} y(n,i)$$

By assuming

$$\begin{aligned} \mathbf{Q} &= \begin{bmatrix} f_0(0) & f_0(1) & \cdots & f_0(N-1) \\ f_1(0) & f_1(1) & \cdots & f_1(N-1) \\ \vdots & & & \vdots \\ f_{M-1}(0) & f_{M-1}(1) & \cdots & f_{M-1}(N-1) \end{bmatrix} \\ &= [\, p_L(0)\mathbf{q}_0 \;\; p_L(1)\mathbf{q}_1 \;\; \cdots \;\; p_L(N-1)\mathbf{q}_{N-1} \,] \end{aligned}$$

where $\mathbf{q}_i$ for $i = 0, 1, \ldots, N-1$ are column vectors and

$$\begin{aligned} q_i(k) &= \cos\left[(2k+1)\frac{\pi}{2M}\left(i-\frac{k_d}{2}\right)+(2k+1)\frac{\pi}{4}\right] \\ &= \cos\left[\left(16+i-\frac{k_d}{2}\right)(2k+1)\frac{\pi}{64}\right] \end{aligned}$$

for $k = 0, \ldots, M-1$, we obtain $y(n,i) = \langle p_L(i)\mathbf{q}_i, \mathbf{w}_{n-i}\rangle$ where $\mathbf{w}_n = [\,\tilde{x}_0(n)\ \tilde{x}_1(n)\ \cdots\ \tilde{x}_{M-1}(n)\,]^T$. Hence

$$\hat{x}(n) = \sum_{i=0}^{N-1}\langle p_L(i)\mathbf{q}_i, \mathbf{w}_{n-i}\rangle \qquad (2)$$

Now consider a reconstructed signal vector $\hat{\mathbf{x}}_{lM}$ formed by the reconstructed signals at the (lM)th, $(lM+1)$th, ..., $(lM+31)$th instants. By (2) and noting that only $\mathbf{w}_{lM} \neq 0$ for integer l, we can write

$$\begin{aligned} \hat{\mathbf{x}}_{lM} &= \begin{bmatrix} \hat{x}(lM) \\ \hat{x}(lM+1) \\ \vdots \\ \hat{x}(lM+31) \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=l}^{l-15}\langle p_L((l-i)M)\mathbf{q}_{(l-i)M}, \mathbf{w}_{iM}\rangle \\ \sum_{i=l}^{l-15}\langle p_L((l-i)M+1)\mathbf{q}_{(l-i)M+1}, \mathbf{w}_{iM}\rangle \\ \vdots \\ \sum_{i=l}^{l-15}\langle p_L((l-i)M+31)\mathbf{q}_{(l-i)M+31}, \mathbf{w}_{iM}\rangle \end{bmatrix} \\ &= \mathbf{A}\mathbf{D}_{lM} \end{aligned}$$

where $\mathbf{A} = [\,\mathbf{A}_0\ \mathbf{A}_1\ \ldots\ \mathbf{A}_{15}]$ with $\mathbf{A}_i = \mathrm{diag}[p_L(iM)\ \cdots\ p_L((i+1)M-1)]$ for $i = 0, 1, \ldots, 15$, and

$$\mathbf{D}_{lM} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \mathbf{d}_3 \\ \vdots \\ \mathbf{d}_{15} \end{bmatrix} = \begin{bmatrix} [\mathbf{q}_0\,\mathbf{q}_1\cdots\mathbf{q}_{31}]^T\mathbf{w}_{lM} \\ [\mathbf{q}_{32}\,\mathbf{q}_{33}\cdots\mathbf{q}_{63}]^T\mathbf{w}_{(l-1)M} \\ [\mathbf{q}_{64}\,\mathbf{q}_{65}\cdots\mathbf{q}_{95}]^T\mathbf{w}_{(l-2)M} \\ \vdots \\ [\mathbf{q}_{480}\,\mathbf{q}_{481}\cdots\mathbf{q}_{511}]^T\mathbf{w}_{(l-15)M} \end{bmatrix}$$

Note that for an integer m,

$$\begin{aligned} q_{64m+i}(k) &= \cos[(16+64m+i-k_d/2)(2k+1)\pi/64] \\ &= (-1)^m \cos[(16+i-k_d/2)(2k+1)\pi/64] \\ &= (-1)^m q_i(k) \end{aligned}$$

for $i = 0, \ldots, 63$, and hence $\mathbf{q}_{64m+i} = (-1)^m \mathbf{q}_i$. Therefore, to form $\mathbf{D}_{iM}$, only $\mathbf{q}_0$, $\mathbf{q}_1$, $\cdots$ and $\mathbf{q}_{63}$ need to be calculated.

In summary, the synthesis subband filtering can be carried out by the following steps:

- Input 32 new subband samples $\mathbf{s}_i$ for $i = 0, \ldots, 31$ and initialize vector $\mathbf{v}$ of length 1024.
- Set $\mathbf{v}_i = \mathbf{v}_{i-64}$ for $i = 1023$ to 64.
- Compute $\mathbf{v}_i = \sum_{k=0}^{31} \mathbf{N}_{ik} \cdot \mathbf{s}_k$ for $i = 0, \ldots, 63$, where

$$\mathbf{N}_{ik} = \cos[(16+i-k_d/2)(2k+1)\pi/64] \qquad (3)$$

  for $0 \le i \le 63$, $0 \le k \le 31$.
- Construct vector $\mathbf{u}$ as follows
  for $i = 0$ to 7,
   for $j = 0$ to 31,
   { $\mathbf{u}_{64i+j} = \mathbf{v}_{128i+j}$
    $\mathbf{u}_{64i+j+32} = \mathbf{v}_{128i+96+j}$ }
- Compute $\mathbf{w}_i = 2 \cdot (-1)^{int[i/64]} \mathbf{u}_i \cdot p_L(i)$ for $i = 0, \ldots, 511$.
- Compute 32 samples $\hat{\mathbf{x}}_j = 32 \sum_{i=0}^{15} \mathbf{w}_{j+32i}$ for $j = 0, \ldots, 31$, and output $\hat{\mathbf{x}}_j$ as the reconstructed PCM samples.

3. SIMULATIONS

The polyphase implementation described in Sec. 1 is a modified version of the MPEG standard [1] in that the prototype filter with symmetrical impulse response is replaced by a filter with asymmetrical impulse response, and entries $\mathbf{M}_{ik}$ and $\mathbf{N}_{ik}$ are now calculated using (1) and (3), respectively.

To show the efficiency of the polyphase implementation, the numbers of multiplications (multi) and additions (add) per input and per output for the polyphase implementation, and those for a direct implementation [4], where each channel is implemented individually in a polyphase form, are listed in Table 2. It is observed that the operations required in a polyphase implementation are much less than those in the direct implementation.

To verify the polyphase implementation, we used a ramp signal as the system's input. The input and the outputs of the proposed low-delay QMF bank and of the current MPEG QMF bank are shown in Fig. 4. It is observed that, in agreement with the design specification, the low-delay QMF bank needs only half of the time required by the current MPEG QMF bank to reconstruct the input signal.

Table 2. Computation required by the polyphase and direct implementations

	per input		per output	
	multi.	add.	multi.	add.
Polyphase	80	77	80	77
Direct	512	511	512	511

4. CONCLUSION

A low-delay cosine-modulated QMF bank has been proposed for MPEG audio compression and an efficient polyphase structure has been presented. The simulation shows that the reconstruction delay in the low-delay QMF bank can be made half of that in current MPEG QMF bank.

ACKNOWLEDGEMENT

The authors are grateful to Micronet, Networks of Centres of Excellence Program and to the Natural Sciences and Research Council of Canada, for supporting this work.

REFERENCES

[1] ISO/IEC 11172-3, "Coding of moving pictures and associated audio for digital storage at up to about 1.5 Mbit/s - Part 3: Audio," ISO/IEC, Switzerland, Aug. 1993.

[2] H. Xu, W.-S. Lu, and A. Antoniou, "Efficient design of multi-channel cosine-modulated QMF banks by an iterative method", *IEEE Trans. on Signal Processing* (in press).

[3] H. Xu, *Efficient Design Methods for Multirate Filter Banks and Their Applications*, Ph.D. thesis, University of Victoria, Aug. 1995.

[4] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Englewood Cliffs, NJ: Prentice Hall, 1993.

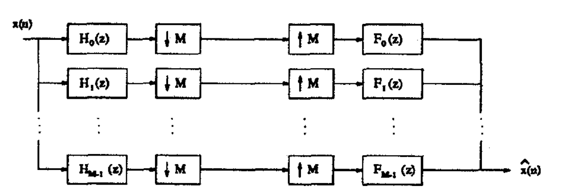

Figure 1. Multi-channel filter bank.

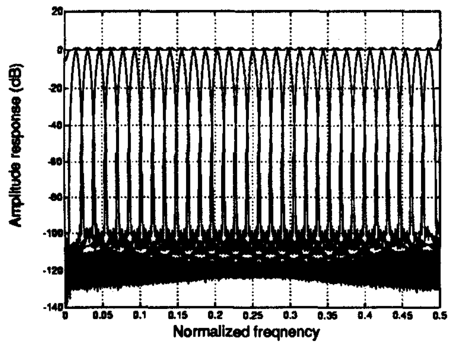

Figure 2. Amplitude responses of low-delay analysis filters.

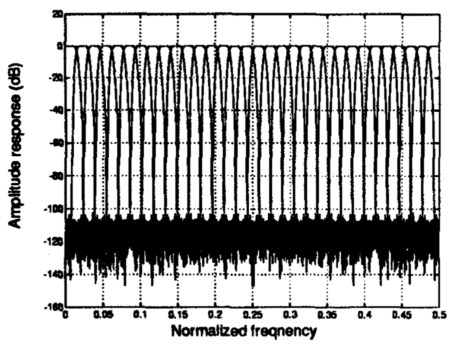

Figure 3. Amplitude responses of current MPEG analysis filters.

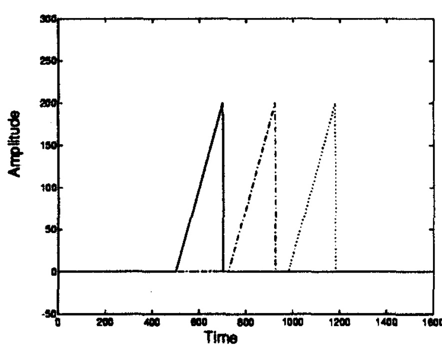

Figure 4. Input: solid line; output (low-delay QMF bank): dash-dot line; output (current MPEG QMF bank): dotted line.