SUBBAND CODING OF SPEECH USING M-BAND PARALLEL QUADRATURE MIRROR FILTERS

H. Gondokusumo and T.S. Ng

Department of Electrical and Computer Engineering
The University of Wollongong

SUMMARY

In this paper, the design of both analysis and synthesis filters for a 3-band parallel sub-band coder with perfect reconstruction properties is given. An example of this design using FIR filters of 17th order is implemented in an IBM PC/AT. Experimental results using this subband coder on segments of speech will be demonstrated in the presentation.

INTRODUCTION

Subband coding of speech has received considerable attention in the past few years since its introduction (Croisier, Esteban & Galand, 1976 ; Esteban & Galand, 1977) because of the possible reduction in transmission rate without degrading speech quality. In subband coding, the full band signal is split into smaller subbands by using a set of analysis filters. These subband signals are then decimated, encoded (to suit the selected transmission rate) and transmitted over a channel. At the receiver end, the subband signals are decoded, interpolated and recombined through a set of synthesis filters.

In general, the reconstructed signal would experience three types of errors: aliasing distortion, amplitude distortion and phase distortion. The two band QMF bank (Esteban & Galand, 1977) can be designed so that the interband aliasing distortion, which occurs with sample decimation in the analysis filter outputs, is cancelled during the synthesis process. Once the aliasing distortion is cancelled, the other distortions (amplitude and phase) can then be minimised.

A recent study (Smith & Barnwell, 1984) has shown that the analysis and synthesis filters for splitting the full band signal into two overlapping subbands can be designed to eliminate all these distortions. These results have been further extended to the M-band parallel QMF (Vaidyanathan, 1987) based on the concept of lossless transfer matrix. In this case, the system is said to exhibit a perfect reconstruction property and the reconstructed signal is simply a delayed version of the input signal.

The M-band parallel QMF divides the full band signal into M arbitrary subbands and the reconstructed signal is delayed by the length of the filter. By comparison, most previous designs of the QMF split the full band signal into two overlapping subbands, with a tree structure being used to achieve the desired number of subbands. However, this implies that the number of subbands are limited to 2^n and the delay of the reconstructed signal is much longer than n times the filter length.

The organisation of the paper is as follows : in section 2, the design of both the analysis and synthesis filters for a 3-band parallel subband coder will be given. In section 3, results of experimental studies using these filters in speech signal will be presented and finally in section 4, some concluding remarks will be made on our experience with parallel M-band subband coder.

THREE BAND PARALLEL SUBBAND CODER

Fig. 1 shows a three band parallel subband coder (M=3) where $H_0(z)$, $H_1(z)$ and $H_2(z)$ represent the analysis filters and $F_0(z)$, $F_1(z)$ and $F_2(z)$ represent the synthesis filters. The input signal x(n) is divided into three frequency subbands by the analysis filters. Each subband signal is decimated (down sampled) by a factor of three, encoded and then transmitted. At the receiver end, each subband signal is decoded, interpolated by a factor of three, filtered by the synthesis filters and then added together to form the reconstructed signal y(n). The problem at hand is to find $H_i(z)$ and $F_i(z)$, i=0,1,2, such that perfect reconstruction of the input signal at the receiving end results.

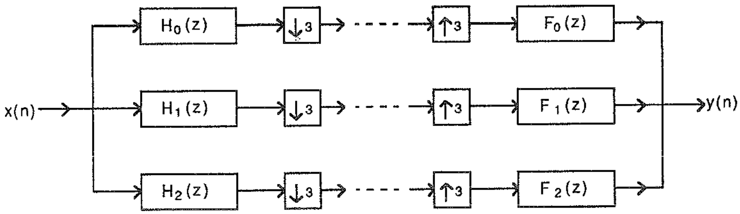

Fig. 1. 3 - band parallel subband coder

One method of designing perfect reconstruction M-channel QMF banks with $H_i(z)$ restricted to FIR filters has been described in Vaidyanathan (1987). The basic idea is to construct an MxM unitary transfer matrix **E**(z) on the unit circle such that $H_i(z)$ and $F_i(z)$ are given by :

$$H_i(z) = \sum_{j=0}^{M-1} z^{-j} E_{ij}(z^M) \qquad i = 0,\ldots,M-1 \tag{1}$$

$$F_i(z) = z^{-rM} \sum_{j=0}^{M-1} z^{-(M-1-j)} E_{ij}(z^{-M}) \qquad i = 0,\ldots,M-1 \tag{2}$$

where r is an integer dependent on filter length of $H_i(z)$ and is chosen such that $F_i(z)$ is causal. The matrix **E**(z) is unitary on the unit circle of the z-plane, if

$$\mathbf{E}^*(z)\mathbf{E}(z) = c\mathbf{I}, \qquad z = e^{jw} \tag{3}$$

where the superscript * denotes complex conjugate transpose, c is a constant and **I** is the identity matrix. The parallel subband coder of Fig. 1 can now be represented by Fig. 2.

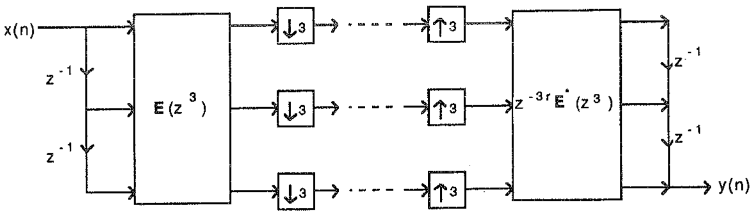

Fig. 2. 3-band subband coder using unitary transfer matrix **E**(z).

The unitary matrix **E**(z) can be constructed by a cascade of L-1 constant orthogonal matrices and delays as shown in Fig. 3.

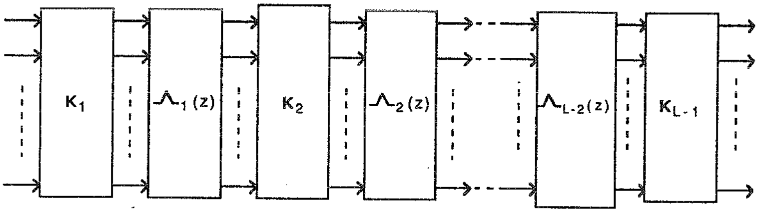

Fig. 3. Structure of Unitary Transfer Matrix $E(z^M)$

where

$$\Lambda_i(z) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & z^{-1} \end{bmatrix} \qquad i = 1, ..., L-2 \tag{4}$$

and

$$K_i = \begin{bmatrix} \cos\theta_{1i} & \sin\theta_{1i} & 0 \\ \sin\theta_{1i} & -\cos\theta_{1i} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{2i} & \sin\theta_{2i} \\ 0 & \sin\theta_{2i} & -\cos\theta_{2i} \end{bmatrix} \tag{5}$$

where $\theta_{1,i}$ and $\theta_{2,i}$, i = 1,....,L-1 are obtained by optimisation of an objective function such that the stopband attenuation of Hi(z) is maximised. The choice of L depends on the FIR filter length N. Their relationship is given by :

$$N-1 = 3(L-2) + 2 \tag{6}$$

For example, if L=7, then orders of $H_i(z)$ are N-1=17. If L=32, then orders of $H_i(z)$ are N-1 = 92.

<u>Example</u>

Let L = 7, Λ_i and K_i be given by (4) and (5) respectively. The objective function to be minimised is given by :

$$J = \int_{\pi/3+\varepsilon}^{\pi} |H_0(e^{jw})|^2\, dw + \int_{0}^{2\pi/3-\varepsilon} |H_2(e^{jw})|^2\, dw$$

$$+ \int_{0}^{\pi/3-\varepsilon} |H_1(e^{jw})|^2\, dw + \int_{2\pi/3+\varepsilon}^{\pi} |H_1(e^{jw})|^2\, dw \tag{7}$$

Equation (7) minimises the stopband energy which is equivalent to maximising the stopband attenuation. The quantity ε depends on the desired stopband edges and was chosen to be 0.21. Table 1 shows the resulting coefficients of the orthogonal matrices, $\cos\theta_{1i}$, $\sin\theta_{1i}$, $\cos\theta_{2i}$ and $\sin\theta_{2i}$, i = 1,...,6. The corresponding FIR filter coefficients for $F_0(z)$, $F_1(z)$ and $F_2(z)$ are listed in Table 2.

TABLE 1

The values of $Cos\theta_{1i}$, $Sin\theta_{1i}$, $Cos\theta_{2i}$, $Sin\theta_{2i}$, $i = 1, \ldots, 6$

i	$Cos\theta_{1i}$	$Sin\theta_{1i}$	$Cos\theta_{2i}$	$Sin\theta_{2i}$
1	.2095947	.9777884	.1008526	-.9949014
2	-.4870450	-.8733769	.2450896	-.9695004
3	.9947151	-.1026737	-.9408118	.3389294
4	-.9371068	-.3490428	.4194380	-.9077840
5	-.9255171	-.3787057	.5303817	-.8477590
6	.4858138	-.8740623	.9465415	-.3225820

TABLE 2

Coefficients of the Optimised Analysis FIR filter banks

n	$h_0(n)$	$h_1(n)$	$h_2(n)$
0	-.52049530-001	.16783203	.12282664-001
1	-.10507174-002	.34460984-001	.53925703-002
2	.10365229-001	-.33995428	-.53197184-001
3	-.32608347-001	-.58284356-001	.31758852-001
4	-.16272900	.36316290	.50320528-002
5	-.30521705	.44988898-001	-.18150451-001
6	-.34810946	-.20676505	-.60027940-001
7	-.24405733	-.15044861-001	.21051382
8	-.62170856-001	.19710893-002	-.33630757
9	.93439547-001	-.27676854-001	.32698902
10	.11474099	.82608642-001	-.19322715
11	.42681986-001	.84796866-002	.14209009-001
12	-.32002524-001	-.17787369-001	.10743386
13	-.17837198-001	.10563926-001	-.10491556
14	-.87032305-002	-.27615064-002	.12511858-001
15	.00000000	.00000000	.00000000
16	.82318275	.45753478-002	-.27634601-001
17	.83445598-003	.46380017-003	-.28013048-002

EXPERIMENTAL RESULTS

The 3-band parallel subband coder with the analysis and synthesis filters described in the last section has been implemented using C in an IBM PC/AT with a floating point co-processor. Experiments were performed on segments of speech using different transmission rates by varying the number of quantization levels in each subband. These results were recorded on cassettes and will be demonstrated in the presentation. The following comparisons have been made :

1. Original digitized speech.
2. Reconstructed speech.
3. Reconstructed speech with 32 kbs transmission rate.
4. Reconstructed speech with 16 kbs transmission rate.
5. Reconstructed speech with 13 kbs transmission rate.

For demonstration purposes, the three subband signals were also recorded separately at the sampling frequency.

CONCLUSION

The purpose of this paper has been to demonstrate the parallel subband coder which has been shown to have perfect reconstruction properties. An example of a 3-band parallel subband coder has been designed with filters restricted to 17th order FIR and was implemented on an IBM PC/AT. Experiments using this coder has been conducted on segments of speech with a number of different transmission rates. The results are encouraging and will be illustrated during the presentation.

REFERENCES

Croisier A., Esteban D. and Galand C.(1976), "Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques," *International Conf. on Information Science and Systems*, Patras, Greece.

Esteban D. and Galand C. (1977), "Application of quadrature mirror filters to split-band voice coding schemes," *Proc. of the IEEE International Conf. ASSP*, Hartford, Connecticut, pp. 191-195, May 1977.

Smith M.J.T. and Barnwell T.P. (1984), "A procedure for designing exact reconstruction filter banks for tree structured subband coders," *IEEE International Conf. on ASSP*, San Diego, California.

Vaidyanathan P.P. (1987), "Theory and design of M-channel maximally decimated quadrature mirror filters with arbitrary M, having the perfect-reconstruction property," *IEEE Trans. on ASSP*, vol. 35, 476-492.