Applied Signal Processing

Laboratory Project 3: SUBBAND CODING OF SPEECH SIGNALS

T. McKelvey C. Brown M. Viberg

Oct 2003

1 Background

In modern telephone systems the connection between the caller and the called are realized using high speed *digital* transmission links. It is only the last few kilometers from the telephone switch station to the customers' premises which are transmitted in analog form (POTS – plain old telephony system). The exception to this of course is if the subscriber has ISDN service in which case the information is transmitted in digital form all the way to the telephone.

The standard format to sample and code a telephone signal is using a sampling frequency of 8 kHz and to quantize the analog signal to 8 bits. This gives an overall data capacity requirement of 64 kbits/s in each direction.

Your task is to design a subband coder which will decrease the original bitrate of 64 kbits/s while still preserving good telephony quality. In this lab we will not investigate sophisticated speech coding schemes – we will simply quantize the subband signals.

You can read about subband coding in examples 11.9 and 11.10 in the textbook [1].

2 Tasks

Start by designing a 2 branch analysis and reconstruction filter bank of quadrature mirror filters (QMF) with perfect reconstruction properties, see Figure 11.20 in the textbook [1]. A detailed description on how to do this is found in Section 5. A QMF bank has halfband filters which split the original signal into a low frequency and a high frequency subband. Figure 11.19 (b) in the textbook [1] shows the magnitude of the two filters in the QMF bank. After implementation, use these filters on a recorded speech sample and verify that the reconstructed signal is still of good quality. Now use this filter design in a tree type filterbank to split the signal into the desired subbands. See Figure 2 or Figures 11.13 and 11.16 in [1] for examples of tree type filter banks.

When finalizing the design the following questions need to be resolved:

- How should the subbands be selected? Uniform or non-uniform.
- How many subbands should be used?

- How many bits should be assigned to each subband?
- What bitrate does your subband coder achieve?

All these design questions are difficult to give exact answers to – the answers you come up with may depend on your subjective judgment (i.e. there is no single "right" answer). Experiment with some different designs. Explain why, in your opinion, certain choices are superior to others.

3 Hints

- Remember that each filter will give rise to a certain delay. In the reconstruction filter bank you need to make sure that all branches have the same delay, i.e., some branches might need to be explicitly delayed.

4 Matlab issues

4.1 Quantization

Use the MATLAB function `quant` to quantize the real valued signal. The code can be downloaded from the course homepage. Below you find the help text for the function.

```
function y = quant(x,nbits,maxAmp);
% function y = quant(x,nbits,maxAmp);
%
% Quantizes the signal x using nbits number of bits.
%
% The quantization is symmetric around 0 and
% gives 2^nbits levels between -maxAmp and +maxAmp
%
% If the input x is outside the -maxAmp to +maxAmp interval
% the function clips the signal to +maxAmp or -maxAmp depending on
% the sign.
```

4.2 Recording and playing audio

From within MATLAB you can both record and play sound. The commands `wavplay` and `wavrecord` will play and record sound respectively. These commands only work on a Windows platform. To store and retrieve sound files to and from MATLAB and Microsoft Wave files (.wav) use the commands `wavread` and `wavwrite`. On the course home page there are two .wav files you can use if you want.

5 Perfect reconstruction filter banks

5.1 Introduction

A filter bank (FB) consists of a set of parallel filters used to decompose a signal into different frequency components. For an introduction to the area, please read Sections 11.3-11.4 in the textbook [1]. For more details on multirate systems and filter banks, the interested reader is referred to, for example, [2] or [3].

There are two basic types of FBs, the uniform DFT FB and the tree structured FB. When analyzing audio signals, the tree structured FB is more useful because of its more flexible frequency resolution properties. The basic building block is illustrated on page 322 of [1], see also Figure 1. Here, $H_0(z)$ and $H_1(z)$ are the analysis low-pass and high-pass filters respec-

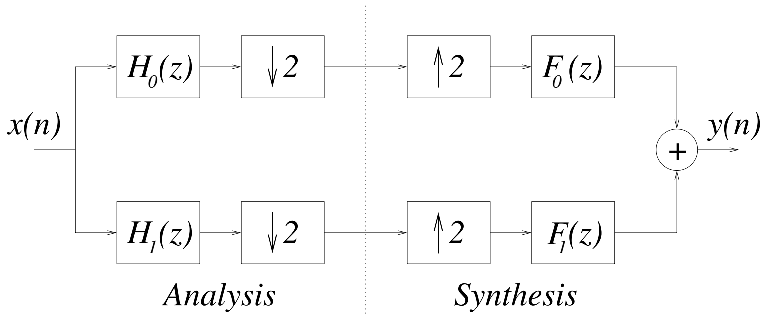

Figure 1: A two-channel analysis filter bank followed by a synthesis filter bank.

tively, whereas $F_0(z)$ and $F_1(z)$ are the corresponding synthesis filters. In practical systems, several such two-channel blocks are combined to obtain a desired frequency resolution. For example, repeatedly applying a two-channel FB to the low-pass output of the previous FB, while leaving the highpass part unchanged, leads to a frequency analysis as in Figure 2. This

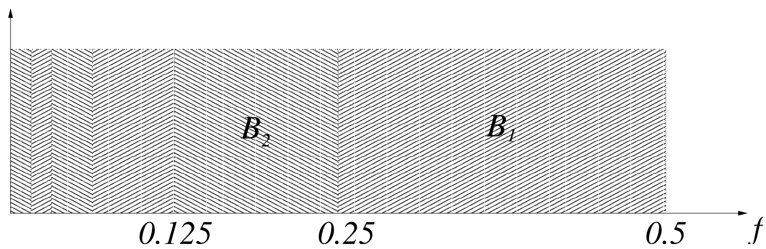

Figure 2: Octave band decomposition of the frequency axis.

so-called octave-band decomposition is closely related to the wavelet transform. Splitting a signal into frequency components is useful in a number of applications, including for example:

- Subband coding for audio and image/video
- Hearing aids and simulation of hearing impairments
- Classification of transient and non-stationary signals
- Time-frequency signal analysis

In this project you will apply the filter bank to decompose a speech signal into subbands to allow for more efficient coding.

5.2 Perfect Reconstruction Filter Banks

An important question when using filter banks if of course if the original signal can be reconstructed from its subband components. Below follows a short motivation for the procedure you will use to (nearly) achieve this. It is hard to avoid some mathematical details when dealing with filter banks.

Consider the basic two-channel analysis/synthesis pair of Figure 1. It can be shown that the z-transforms of the input and output signals are related by

$$\begin{aligned} Y(z) &= \tfrac{1}{2}X(z)\left\{H_0(z)F_0(z) + H_1(z)F_1(z)\right\} \\ &+ \tfrac{1}{2}X(-z)\left\{H_0(-z)F_0(z) + H_1(-z)F_1(z)\right\} \end{aligned} \tag{1}$$

The first term above represents the (possibly distorted) desired output, whereas the second term is due to aliasing. The aliasing is induced by the downsampling, since the filters are not ideal. However, we see that the aliasing can be canceled if we choose the filters to satisfy

$$H_0(-z)F_0(z) + H_1(-z)F_1(z) = 0\,. \tag{2}$$

Similarly, there is no distortion if

$$H_0(z)F_0(z) + H_1(z)F_1(z) = 2z^{-n_k}\,, \tag{3}$$

where n_k is the overall delay of the filter bank. If the above conditions are met, the reconstructed signal is $y(n) = x(n - n_k)$, and we have a Perfect Reconstruction (PR) Filter Bank (FB).

In a so-called paraunitary filter bank, the filters are constructed from a prototype $H_0(z)$ according to

$$H_1(z) = -z^{-(N-1)}H_0(-z^{-1}) \tag{4}$$

$$F_0(z) = z^{-(N-1)}H_0(z^{-1}) \tag{5}$$

$$F_1(z) = z^{-(N-1)}H_1(z^{-1})\,, \tag{6}$$

where N is the filter length, which is assumed equal for all filters. Provided N is even, it is easily verified that (4)–(6) imply (2), i.e. the aliasing is canceled. Also, the "distortion filter" in (3) reduces to

$$H_0(e^{j\omega})F_0(e^{j\omega}) + H_1(e^{j\omega})F_1(e^{j\omega}) = e^{-j(N-1)\omega}\left(|H_0(e^{j\omega})|^2 + |H_1(e^{-j\omega})|^2\right)\,. \tag{7}$$

Thus, to satisfy (3), the lowpass filter $H_0(z)$ and the high-pass filter $H_1(z)$ should be power-complementary. Now, (4) implies that

$$|H_1(e^{j\omega})|^2 = |H_0(e^{j(\pi-\omega)})|^2\,. \tag{8}$$

In other words, $|H_0(e^{j\omega})|$ and $|H_1(e^{j\omega})|$ are mirror images about $\omega = \pi/2$. They are therefore referred to as Quadrature Mirror Filters (QMF). Finally, by (3) and (7), a PR-QMF filter bank is obtained if $H_0(z)$ is chosen to satisfy

$$|H_0(e^{j\omega})|^2 + |H_0(e^{j(\pi-\omega)})|^2 = 2\,. \tag{9}$$

Let $P(z) = H_0(z)H_0(z^{-1})$ be the spectrum associated with $H_0(z)$. Then the above reads $P(z) + P(-z) = 2$, which holds if $P(z)$ takes the form

$$P(z) = h_{N-1}z^{N-1} + \cdots + h_3z^3 + h_1z + 1 + h_1z^{-1} + h_3z^{-3} + \cdots + h_{N-1}z^{-(N-1)}\,.$$

Such a filter, which has all even coefficients equal to zero is called a *half band filter*. To find a $P(z)$ with a "good" lowpass spectrum is non-trivial, and often requires trading one desired property against the other. In this project you will use the following steps to hopefully achieve a reasonable PR-QMF design:

- *Design a basic lowpass FIR filter*, $P(z)$, with cutoff frequency $\omega = \pi/2$. The filter length should be equal to $2N-1$, resulting in analysis and synthesis filters of length N, which must be even! You can use either `fir1` or `remez` in Matlab to design $P(z)$.
- Add a small number (like 0.002) to the middle coefficient. This is done to make sure that $P(e^{j\omega})$ is positive for all ω.
- *Make the filter a half-band filter* – i.e. zero out all even indexed coefficients (except the middle coefficient). These coefficients should already be close to zero, however it is still be desirable to set them exactly equal to zero.
- *Compute the zeros of* $P(z)$. This can be done using `roots` – e.g.
  `zz = roots(P);`.
- *Construct* $H_0(z)$*, i.e.,* $h_0(n)$*, from the* $N-1$ *zeros of* $P(z)$ *that are inside the unit circle.* If there are zeros exactly on the unit circle, use half of them in your construction of $H_0(z)$. You may choose to do this by finding the magnitude of the zeros, `sort`ing them and using the $N-1$ smallest zeros to construct $H_0(z)$ – e.g.
  ```
  [zabs,ind] = sort(abs(zz));
  zinside = zz(ind(1:N-1));
  ```
  Keep in mind that conjugate pair zeros should be kept together in a design. Form $h_0(n)$ using `poly`. To get the gain correct, use:
  ```
  h0 = sqrt(2)*h0/max(abs(fft(h0)));
  ```
- *Form* $h_1(n)$ *by time-reversing* $h_0(n)$ *and sign-changing the coefficients of the odd powers of* z.
- *Let* $f_l(n)$ *be* $h_l(n)$ *time-reversed, for* $l = 1, 2$.

Results from some of these steps are shown in Figure 3. When finished, plot $|H_0(e^{j\omega})|^2$, $|H_1(e^{j\omega})|^2$ and $|H_0(e^{j\omega})|^2 + |H_1(e^{j\omega})|^2$ in the same figure. Are you happy with the design?

References

[1] B. Mulgrew, P. Grant, and J. Thompson. *Digital Signal Processing: Concepts and Applications.* Palgrave, New York, NY, 2000.

[2] P.P. Vaidyanathan. *Multirate Systems and Signal Processing.* Prentice-Hall, Englewood Cliffs, NJ, 1993.

[3] S.K. Mitra. *Digital Signal Processing - A Computer-Based Approach.* McGraw-Hill, New York, 2 edition, 2001.

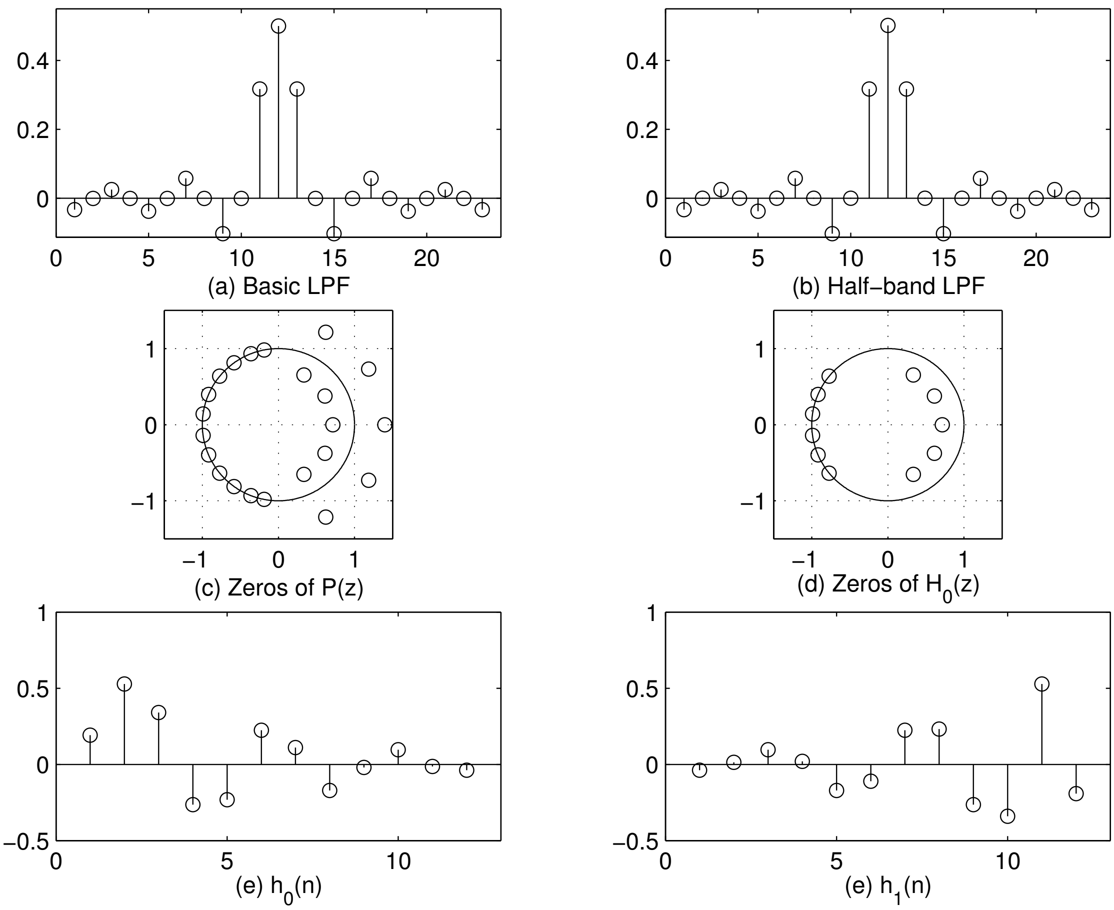

Figure 3: Some steps in construction of PR-QMF bank ($N = 12$).