

SPEECH SIGNAL ANALYSIS FOR LINEAR FILTER BANKS OF DIFFERENT ORDERS

Somashekhar Shankarayya Pujar¹, Shridhar S. Kuntoji², Mahiboob³

¹*Department of Electronics and Communication Engineering, Basaveshwara Engineering College Vidyageri, Bagalkot-587103, India*

²*Department of Electronics and Communication Engineering, Basaveshwara Engineering College Vidyageri, Bagalkot-587103, India*

³*Department of Electronics and Communication Engineering, Basaveshwara Engineering College Vidyageri, Bagalkot-587103, India*

Abstract

In speech signal processing using of filter banks is very important. The critical requirement is the sum of all the frequency responses of the band-pass filters of the filter bank i.e. composite frequency response be flat with linear phase. This paper deals with design and implementation of linear-phase FIR digital filters based filter bank flat with flat composite frequency response. The design is based on special properties of FIR filters by which excellent frequency response can be achieved.

Keyword: Speech signal, FIR Filter, Composite frequency response.

1. INTRODUCTION

In a variety of speech processing systems filter banks are used to perform short time spectrum analysis. To cover a desired portion of the speech band a set of band-pass filters is designed. By simply adding together the outputs of the band-pass filters, it is possible to approximate the spectrum of the input speech signal.

It is possible to characterize the behavior of filter banks by considering the composite frequency response. Since, ideally, the combined output should be equal to the input, then we desire that the composite frequency response have constant magnitude and linear phase over the desired band of frequencies. Earlier work emphasizes the relative phases between channels which is important in achieving a flat composite frequency response. Using this method excellent overall response can be obtained for both IIR and FIR digital filters in filter banks in which the center frequencies are uniformly spaced. This method is not easily extendable to non-uniformly spaced filter banks.

2. SPEECH ANALYSIS

Speech signals are basically partitioned into voiced speech segments and unvoiced speech segment. A voiced speech segment is also known as pitch of voiced speech. It has high energy content and is periodic in nature. The unvoiced part of the speech looks like a random noise with no periodicity. Some parts of the speech that are neither voiced nor unvoiced are called transition segments. The speech can be analysed either by Time domain method and Frequency domain method.

3. EXAMPLE OF BPF DESIGN:

$$H_d(\omega) = \begin{cases} e^{-j\omega T} & \omega_1 \leq \omega \leq \omega_2 \\ 0 & \text{otherwise,} \end{cases}$$

Cutoff frequency of $\omega_1 = \frac{\pi}{2}$, $\omega_2 = \frac{\pi}{4}$ rad/samples. Length of the filter $M = 25$

$$h_d(n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H_d(\omega) e^{j\omega n} d\omega$$

$$h_d(n) = \frac{1}{2\pi} \left[\int_{-\omega_2}^{-\omega_1} e^{j\omega n} e^{-j\omega T} d\omega + \int_{\omega_1}^{\omega_2} e^{j\omega n} e^{-j\omega T} d\omega \right]$$

$$= \frac{1}{2\pi} \left[\int_{-\omega_2}^{-\omega_1} e^{j\omega(n-T)} d\omega + \int_{\omega_1}^{\omega_2} e^{j\omega(n-T)} d\omega \right]$$

$$= \frac{1}{2\pi} \left[\frac{e^{j\omega(n-T)}}{n-T} \Big|_{-\omega_2}^{-\omega_1} + \frac{e^{j\omega(n-T)}}{n-T} \Big|_{\omega_1}^{\omega_2} \right]$$

$$h_d(n) = \frac{\sin \omega_2(n-T) - \sin \omega_1(n-T)}{\pi(n-T)} \quad \text{for } n \neq T$$

when $n = T$ equation will become as

$$h_d(n) = \frac{1}{2\pi} \left[\int_{-\omega_2}^{-\omega_1} d\omega + \int_{\omega_1}^{\omega_2} d\omega \right]$$

$$h_d(n) = \frac{\omega_2 - \omega_1}{\pi}$$

$$h_d(\omega) = \begin{cases} \frac{\sin \omega_2(n-T) - \sin \omega_1(n-T)}{\pi(n-T)} & \text{for } n \neq T \\ \frac{\omega_2 - \omega_1}{\pi} & \text{for } n = T \end{cases}$$

$$T = \frac{M-1}{2} \quad \text{where } M=25$$

$$T = \frac{25-1}{2}$$

$$T=12$$

$$h_d(\omega) = \begin{cases} \frac{\sin \omega_2(n-12) - \sin \omega_1(n-12)}{\pi(n-12)} & \text{for } n \neq 12 \\ \frac{\omega_2 - \omega_1}{\pi} & \text{for } n = 12 \end{cases}$$

n=1,2,3,4,5,.....25.

Result

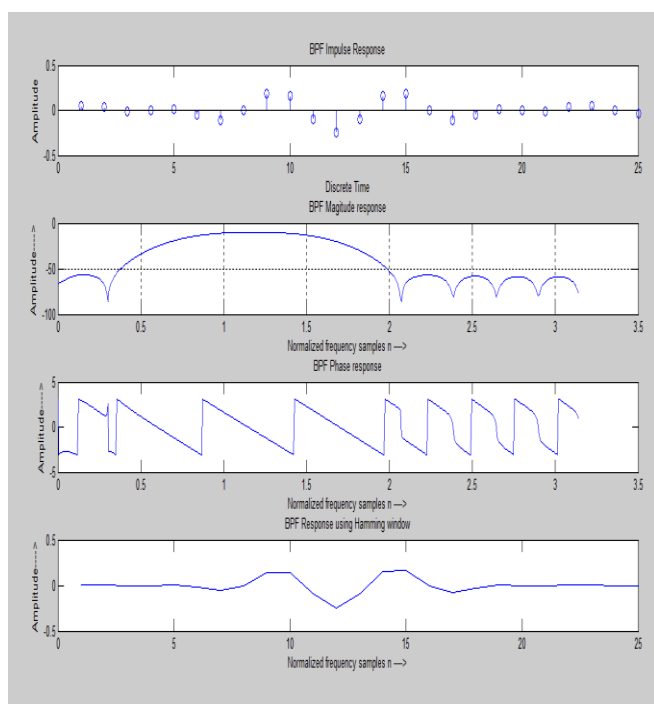


Fig: 3. Band Pass Filter Output

4. FILTER BANK DESIGN

To design a filter bank using FIR filters, we must first determine the range of frequencies to be covered by the composite response. Let us assume that these are denoted ω_{min} and ω_{max} , where $\omega_{max} \leq \frac{\pi}{T}$. Now, if there are a total of N_f filters, we must choose the bandwidths and center frequencies so that the entire range of frequencies $\omega_{min} \leq \omega \leq \omega_{max}$ is covered. This is depicted in Fig 2.3.1 for the case $N_f = 3$. This figure shows the ideal responses for each bandpass filter;

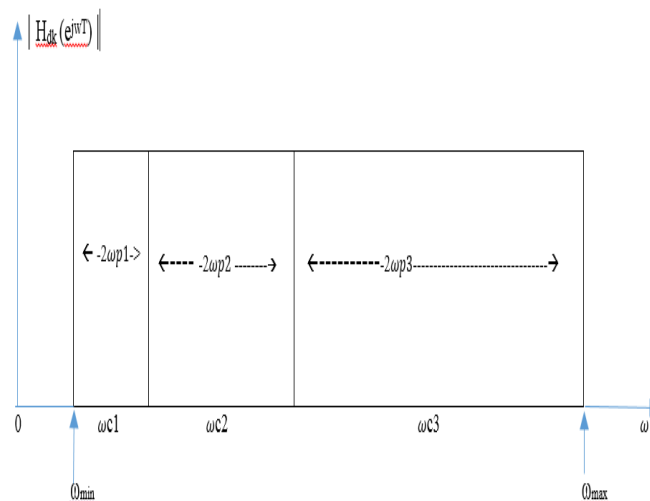


Fig:4 A typical nonuniform filter bank.

4.1 Design Examples

In this section, we illustrate the use of the principles with examples of both uniform and non-uniform filter banks. For all the examples, the sampling rate is assumed to be 8 kHz.

Example 4.1

Suppose that we wish to design a bank of 15 equally spaced filters that covers the range 200 to 3200Hz. Then we find that the cutoff frequency for all the low-pass filters is

$$f_o = \frac{\omega_o}{2\pi} = 100\text{Hz}$$

The center frequencies are

$$f_{ck} = \frac{\omega_{ck}}{2\pi} = 100(2k+1)\text{Hz} \quad k=1,2,3, \dots, 15.$$

Table: 4.1 Values of center frequencies

K	f_{ck}		
1	0.09375	8	0.53125
2	0.15625	9	0.59370
3	0.21875	10	0.65625
4	0.28125	11	0.71875
5	0.34375	12	0.78125
6	0.40625	13	0.84375
7	0.46875	14	0.90625
		15	0.96875

4.1.1 The specifications are,

1. Attenuation= 60dB.
2. Cutoff frequency= 100Hz.
3. The widest transition band that is reasonable is 200 Hz.
4. Order must N= 150.
5. Adjacent channels cross at 0.5 amplitude.
6. Filters will merge together.
7. The deviation is from unity.

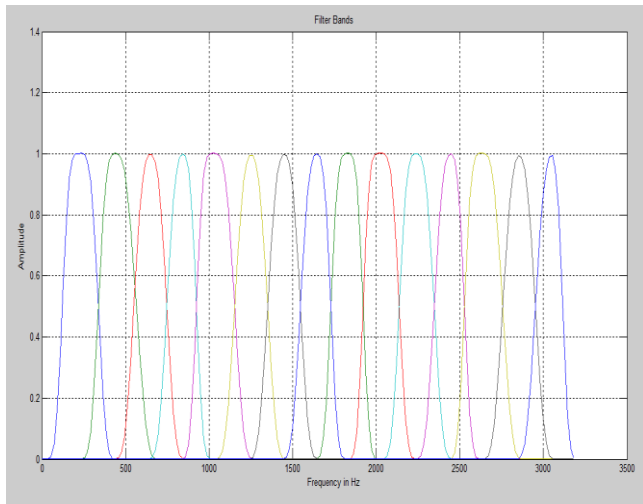


Fig: 4.1.1 FIR Filter Banks

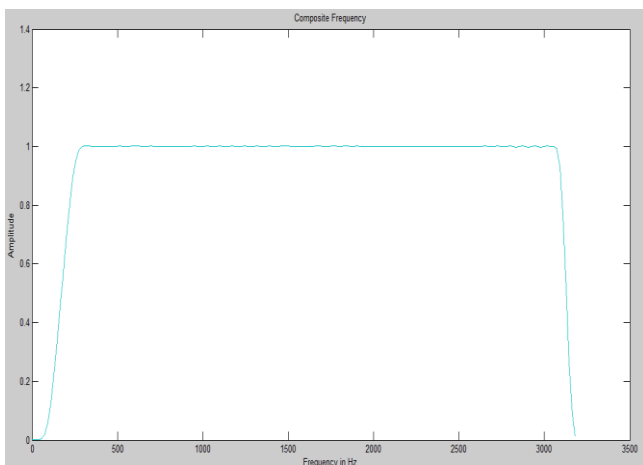


Fig: 4.1.2 Composite Frequency Response

Figure 4.1.2 shows the composite response of the filter bank. It is clear that the filters merge together very well at the edges of the frequency bands. Indeed, the deviation from unity is less than or equal to the peak approximation error, $\delta = 0.001$.

Example 4.2

A non-uniform spacing of the filters is often used to exploit the ear's decreasing frequency resolution with increasing frequency. Suppose that we wish to cover the same range 200 to 3200 Hz as in Example 4.1, but we wish to use only four octave band filters. That is, each successive filter will have a bandwidth twice the bandwidth of the previous filter. We find that the lowest frequency channel has cutoff frequency.

$$f_o = \frac{\omega_c}{2\pi} = \frac{3200-200}{2(2^4-1)} = 100\text{Hz}$$

In general, the cutoff frequencies of the prototype low-pass filters are

$$f_{pk} = \frac{\omega_{pk}}{2\pi} = 2^{k-1}f_o \quad k= 1, 2, 3, 4$$

Table: 4.2 Values of center frequencies

k	f_{ck}
1	0.03125
2	0.0625
3	0.125
4	0.25

4.2.1 The specifications are,

1. No. of filters=4.
2. Spacing = Twice the bandwidth of the previous filter.
3. Coverage range = 200 to 3200 Hz.
4. Order N= 150.

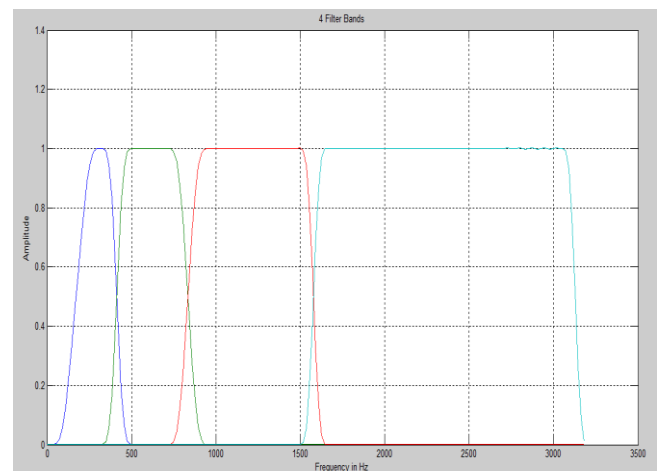


Fig: 4.2.1, 4 Non-uniform FIR Filter Banks for N = 150

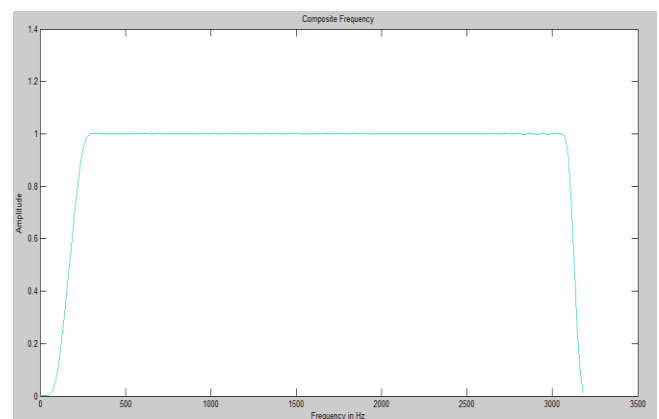


Fig: 4.2.2 Composite Frequency Response

Example 4.3

Suppose that all the parameters remain the same as in Example 4.2 except that we require narrower transition regions. This means that a larger value of N is required. Figure III.4.1 shows the filter bands corresponding to the parameters of Example 4.2 except that N =301 and $\Delta_f = 0.012082$ (transition width is 116 Hz). The sharper transitions are apparent in Fig. 4.3.1 shows that the composite response

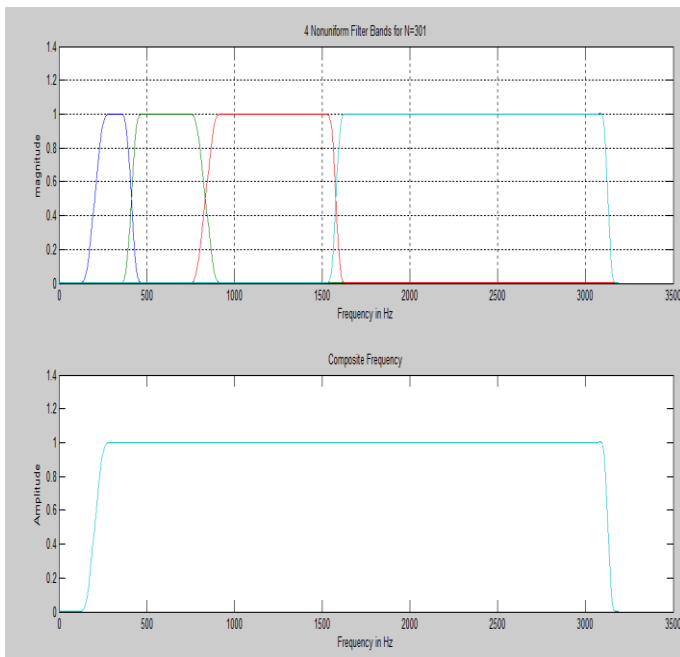


Fig: 4.3.1, Four Non-uniform FIR Filter Banks for N = 301 and Composite frequency.

Example 4.4

We have assumed throughout that the transition width was less than twice the smallest lowpass cutoff frequency. In our examples, this constraint required that N be at least 150. In this case, all the parameters were the same as in Examples 2.2 and 2.3, except in the case of N=101 and $\Delta_f=0.0362465$. The transition width is 348 Hz, which is much greater than twice the cutoff frequency of the first low-pass filter. It is clear that reasonable filters are obtained for the wider bandwidth filters however, the lowest filter does not attain unity response anywhere in its pass-band. The preceding examples make it abundantly clear that, for sufficiently long impulse responses, the composite filter bank response can be very flat.

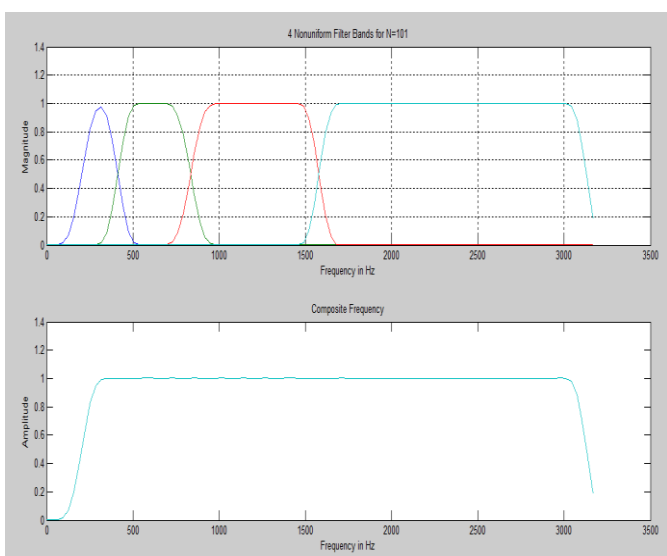


Fig: 4.4, Four Non-uniform FIR Filter Banks for N = 101 and Composite frequency.

Where design techniques for IIR filter banks were discussed, the best results achieved for the composite response were approximately 1-dB peak-to-peak ripple for uniform bandwidths and about 2.5-dB peak-to-peak ripple for non-uniform bandwidths. This is in contrast to the results of the examples of this section, where the peak-to-peak ripple in the composite response was about 0.0274 dB for all the filter banks independent of how the bandwidths were chosen. This, together with the precise linear phase that is easily achieved, makes the FIR filter banks superior to what can be achieved for IIR filter banks. The price that is paid for this is that rather large values of N are required to achieve sharp transitions. However, the values of N used in the previous examples are certainly not unreasonable if the filters are implemented by FFT convolution methods or in special purpose hardware.

5. RESULTS AND DISCUSSION

In this paper totally four filter banks were designed using FIR filter. One is linearly spaced and remaining three are non-linearly spaced. These filter banks are tested with some known signals and monotone signals, such as cosine signals with certain value of frequencies and also vowels.

When 1000Hz cosine signal is applied to all the filter banks. The output is obtained in only one filter bank which is having the bandwidth in the range of 1000Hz (860-1243 Hz). The output can be found by observing the output signals of all the FIR filter banks. The signal which is having more periodic as compared to other signals will be our required output. And in other way is by calculating the energy of all the filter outputs. The filter output which is having the higher energy is our required output. Later monotone signals are applied to filters such as vowels (a, e, i, o, u) and calculated the bandwidths of those signals.

5.1 Calculating the Bandwidths of monotone and vowels signals.

Table 5.1.1 Result obtained when the input to the filter bank (Linearly spaced filters) is sine signal of different frequencies.

Sl/No	Input Frequency Hz	Bands	Energy
1	100	37- 420	0.00028855
2	500	457-841	0.00022262
3	1000	660-1006	0.00015226
4	2000	1847-2213	0.0030
5	3000	2871-3190	0.0055

Table 5.1.2 Result obtained when isolated vowels are input to the filter bank.

Sl/No	Vowels of speaker 1	Bands	Energy
1	A	256-659	662.0616
2	E	37-420	43204873
3	I	256-659	950.5838
4	O	37-420	2456.2
5	U	37-420	69.7421

Table 5.1.3 Result obtained when isolated vowels are input to the filter bank.

Sl/No	Vowels of speaker 2	Bands	Energy
1	A	660-1006	61.6245
2	E	37-420	80.2513
3	I	37-420	99.8089
4	O	457-841	130.80
5	U	37-420	30.0643

Table 5.1.4 Result obtained when isolated vowels are input to the filter bank.

Sl/No	Vowels of speaker 3	Bands	Energy
1	A	731-1646	104.2574
2	E	731-1646	96.9293
3	I	731-1646	175.4988
4	O	330-932	486.1503
5	U	55-495	26.9864

5.2 Listening sounds of all the FIR filters of the filter bank when isolated vowels are input it is observed that, Only filters of lower bands are able to produce the given clear speech stimuli as output. It means the output speech is same as the input speech. Whereas in the higher bands the output does not convey the actual speech. This is true for all higher band filters. It is because vowels are low pass signals.

6. CONCLUSION

We have discussed a design method for filter banks composed of linear digital filters. The method exploits the linear-phase properties obtainable for such filters, as well as the symmetry of the transition region that results from the windowing method of design. We summarized this method of design for the Kaiser window and illustrated the filter bank design method with several examples. These examples show the proposed design method has a great deal of flexibility and that excellent response characteristics can be achieved.

ACKNOWLEDGMENTS

This paper is made possible through the help and support from the guide Dr. K. Sridhar for his guidance, understanding, and patience during my M.Tech course at B.E.C Bagalkot and during the all phases of this paper.

REFERENCES

- [1] J. L. Flanagan, *Speech Analysis, Synthesis and Perception*, Second Ed., New York: Springer-Verlag, 1972.
- [2] J. L. Flanagan and R. M. Golden, "Phase Vocoder," B.S.T.J., 45, No.9 (November 1966), pp. 1493-

- 1509.
- [3] R. W. Schafer and L. R. Rabiner, "Design and Simulation of a Speech Analysis Synthesis System Based on Short Time Fourier Analysis," IEEE Trans. Audio and Electroacoustics, *AU-21*, No.3 (June 1973), pp. 165-174.
- [4] J. L. Flanagan and R. C. Lummis, "Signal Processing to Reduce Multipath Distortion in Small Rooms," J. Acoust. Soc. Am., 47, part 1, June 1970, pp. 1475-1481.
- [5] R. W. Schafer and L. R. Rabiner, "Design of Digital Filter Banks for Speech Analysis," B.S.T.J., 50, No. 10 (December 1971) pp. 3097-3115.
- [6] J. F. Kaiser, "Digital Filters," *System Analysis by digital Computer*, Ch. 7, F. F. Kuo and J. F. Kaiser, eds., New York: John Wiley, 1966, pp. 228-243.
- [7] J. F. Kaiser, "Nonrecursive Digital Filter Design using the Is-Sinh Window Function," Proc. 1974 IEEE Int. Symp. on Cir. and Syst., San Francisco, 1974, pp. 20--23.