IMPLEMENTATION OF SUBBAND CODING ON FARSI SPEECH LANGUAGE

Razie Mohammadi
M.Sc, Azad Islamic University, Najafabad Branch, Najafabad, Isfahan, Iran
E-mail: r_m_kojidi@yahoo.com

Rassoul Amirfattahi
Assistant professor, Department of Electrical and Computer Engineering
Isfahan University of technology, Isfahan, 84156-83111, Iran
E-mail: fattahi@cc.iut.ac.ir

Key words: Subband coding, Farsi Speech Language, QMF filter.

ABSTRACT

In this paper, we implemented subband coding (SBC) technique on Farsi speech language. The goal is determining QMF filter effect on the Farsi speech language in SBC method. Results: symmetric FIR filterbank has some advantages in the coding of the Farsi speech. The IIR filterbank has also proper response but it is possible that reconstruction IIR filterbank becomes unstable.

I. INTRODUCTION

Subband coding is a frequency domain coding technique in which the input signal is decomposed into a number of subbands so that each of these frequency bands can be encoded separately. This technique was originally proposed by Crochiere, Webber and Flanagan [1] as a means to reduce the effect of quantization noise due to coding and therefore to improve the quality of speech coding systems.

The subband coding concept is base on the split frequency spectrum of original signal into some bands. The distributed energy in these bands are not equal over all frequencies. The energy of the low-frequency band has more than high-frequency one in the audio signals. In practice, it is not necessary to split the speech signal in to many subbands. In subband coding (SBC) the signal is divided into four to eight subbands and the waveform signal in each subband is encoded separately [3].

According to this method most of bits for coding the signals is specified to the lower band as energy that is distributed in that band [2].

Therefore in the QMF filter banks (FIR & IIR) each of them that have more energy in the lower bands is the best filter bank.

II. HALFBAND FILTER

Halfband filter, as seen in figure 1, can be used as a building block in a tree structure filter bank to divide the signal into more subbands, [1].

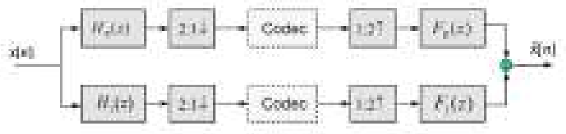


Figure 1.The Halfband Filter

The relation between input $x[n]$ and $\hat{x}[n]$ in the Z-domain is defined as: [2]

$$\hat{X}(z)=\frac{1}{2}[H_0(z)F_0(z)+H_1(z)F_1(z)]X(z) + \frac{1}{2}[H_0(-z)F_0(z)+H_1(-z)F_1(z)]X(-z) \quad (1)$$

To achieve perfect reconstruction, the first term must be pure delay and second term or aliasing term, must be canceled completely, then

$$\frac{1}{2}[H_0(z)F_0(z)+H_1(z)F_1(z)]=kz^{-m_0} \quad (2a)$$

$$\frac{1}{2}[H_0(-z)F_0(z)+H_1(-z)F_1(z)]=0 \quad (2b)$$

Therefore

$$F_0(z)=\frac{2kz^{-m_0}H_0(z)}{H_0^2(z)-H_0^2(-z)} \quad (3a)$$

$$F_1(z)=\frac{-2kz^{-m_0}H_0(-z)}{H_0^2(z)-H_0^2(-z)} \quad (3b)$$

In order to splitting the speech signal into two subbands, we have implemented QMF filter, then we choose:

$$H_1(z)=H_0(-z) \quad (4)$$

So, if we design filter $H_0(z)$, we are able to design other filters according to equations (3a), (3b) and (4). These filters can be in the form of FIR or IIR.

In the filter bank tree, original signal is divided into two Low and High bands, and in the second stage each of them is divided to two Low and High bands, and it will be addressed as Low-Low, Low-High, High-Low and High-High. Sequentially this tree will be expanded in later stages.

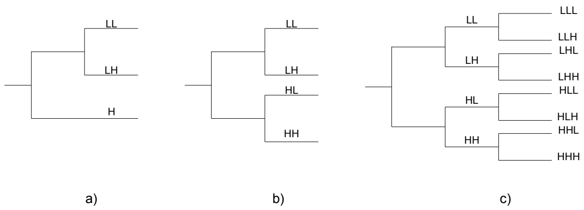


Figure 2. a) Non-uniform tree structure band b) Uniform tree structure band c) Octave tree structure band

III. RESULTS

As we know in the QMF filter bank, the condition of $H_1(z) = H_0(-z)$ must be satisfied, but in our method we perform both type of filter banks with and without this equal condition.

III.1. LINEAR-PHASE FIR FILTER WITH $H_1(z) = H_0(-z)$

We design $H_0(z)$ and $H_1(z)$ for a four channel subband system like Figure 2 when $H_1(z) = H_0(-z)$ is satisfied. These filters are designed in second order. Because the design complexity of the second order is fewer than high order filters, we select this type for compression between filter banks.

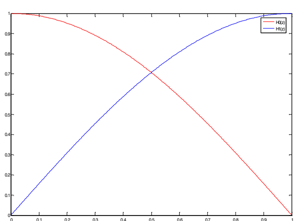
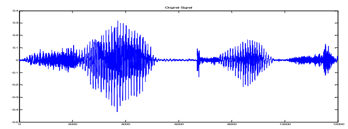


Figure 3. Frequency response of $H_0(z)$ and $H_1(z)$

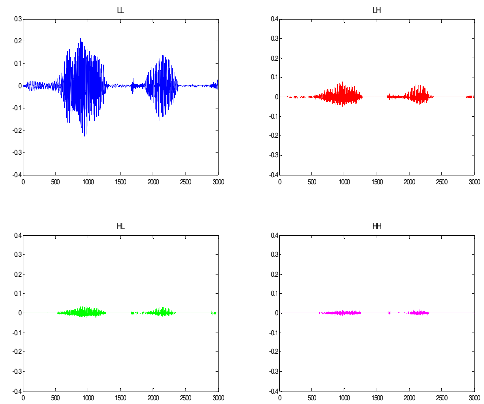
A Farsi speech string is passed through this filter bank. After two stages the original signal is divided to four bands (LL, LH, HL and HH). Most of the energy of the input signal is in the LL band. For example the computed energy in LL band of a typical string is near to %92 of the original and energy of LLL for eight band systems after three stages and for the same original is %85. For encoding it is specified maximum bits to this band and minimum of the bits to another three remained bands.

Some of the Farsi characters after are affected trough this filter, lose their contrast in the sentence. This characters are naturally is generated in high frequency and when affected by this filter bank, there archive minimum bit for encoding.

Figure 4 is demonstrate input signal and affected signal through filter bank.



a)



b) $H_1(z) = H_0(-z)$

Figure 4. a) Original signal b) Output of a two stages subband system by implementation of linear-phase FIR filter on the Farsi statement.

III.2. LINEAR-PHASE FIR FILTER WITHOUT $H_1(z) = H_0(-z)$

We relax this conditional and then redesign a new filter bank in the second order. With the previous Farsi sentence after two stages there are four channels that are obtained from the original signal.

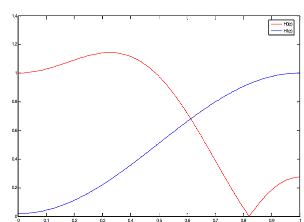
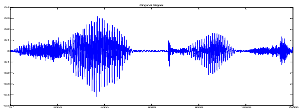


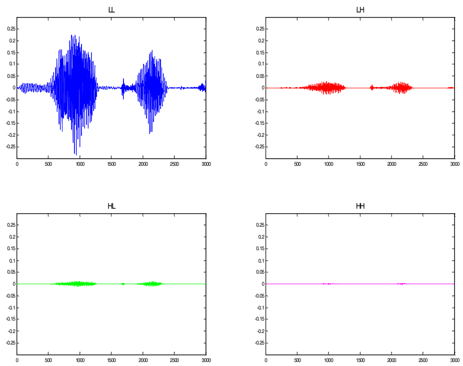
Figure 5. Frequency response of $H_0(z)$ and $H_1(z)$

In computation of the energy that is distributed in each channel, there is seen that the energy of the LL band after two stages is near to %99 and the energy of the LLL band after three stages is %97.

In experimental test, there is seen that the same Farsi characters are affected, but according to the computed value of energy in this filter bank in comparison with the previous filter bank (sec. 3.1) and with the same specified bits, there is better results. (Because remember that we specified more bits for the lower band).



a)



b) $H_1(z) \neq H_0(-z)$

Figure 6. a) Original signal b) Output of a two stages subband system by implementation of linear-phase FIR filter on the Farsi statement.

III.3. IIR FILTER WITH $H_1(z) = H_0(-z)$

In this section we discuss the possible application of IIR filters in filter banks, assuming that the condition $H_1(z) = H_0(-z)$ holds. The main difficulty of this approach is that although stable $H_0(z)$ and $H_1(z)$ can be easily designed; it is possible that the reconstruction filter bank be unstable.

One method overcome the difficulty, is to design $H_0(z)$ containing one or more parameters and then try to find the parameters such that reconstruction filter bank be stable. We use from this method for a second order IIR filter bank, [2]. This method for designing a digital filter is to first design an analog counterpart, then using bilinear transform to obtain the discrete version of that analog filter. The bilinear transform that is defined by equal (5) has the advantage that it maps the inside of unit circle to the left half of s-plan.

$$s = \alpha \frac{1 - z^{-1}}{1 + z^{-1}} \tag{5}$$

This property guarantees that under the bilinear transform, a stable analog filter results a stable digital filter. We use the Butterworth analog filters as the bases of our design.

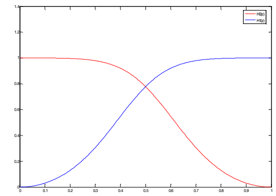
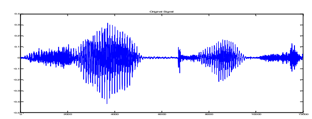
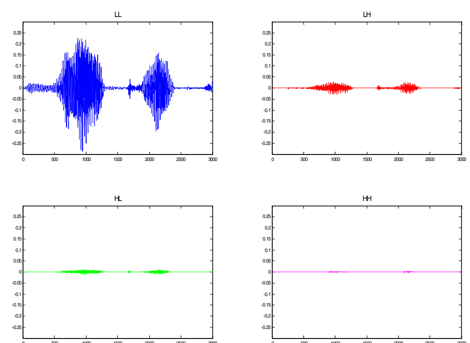


Figure 7. Frequency response of $H_0(z)$ and $H_1(z)$

Then, this filter is implemented to the same Farsi sentence and gets the result as following: %98 of the energy is specified in LL band of the subband system with two stages and %96 of the energy is specified in LLL band in the subband system with three stages.



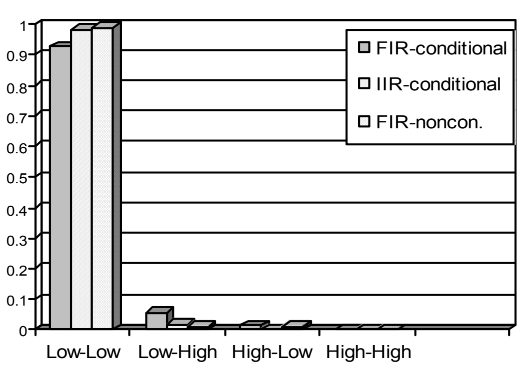
a)



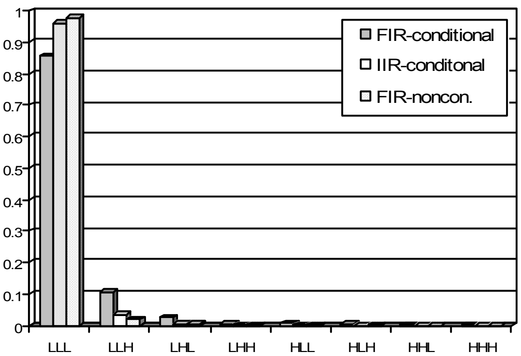
b) $H_1(z) = H_0(-z)$

Figure 7. a) Original signal b) Output of a two stages subband system by implementation of IIR filter on the Farsi statement.

Figure 8 is demonstrating the comparison chart between all of the bands in three previous types of filter banks.



a)



b)

Figure 8. a) Comparison energy for four channels b) Comparison energy for eight channels

IV. ADAPTIVE BIT ALLOCATION

n_i is the number of bits for input signal. The compression [4] gain in two stages subband system is calculated by (6).

$$\frac{n_{LL}+n_{LH}+n_{HL}+n_{HH}}{4n_i} \tag{6}$$

Since almost overall energy is allocated in the lower band, we consider n_1 bits for LL band and n_2 bits for the other bands, therefore compression gain is calculated in (7).

$$\frac{n_1+4n_2}{4n_i} \tag{7}$$

Also compression gain for the three stages subband system is

$$\frac{n_{LLL}+n_{LLH}+n_{LHL}+n_{LHH}+n_{HLL}+n_{HLH}+n_{HHL}+n_{HHH}}{8n_i} \tag{8}$$

Therefore in the eight-band system LLL band not contain all of the energy and another sidebands have a few energy that there are not dispensable and we must consider the energy on that side bands. It means few bit allocation for all of the remained seven side bands, will be generating a few errors in coding of the signal.

For compensation of the error we use adaptive bit allocation for coding. We consider three levels n_1, n_2 and n_3 bits which n_3 have less and n_1 have most of the bits. The methodology for bit allocation in subbands is that at the first the energy of the eight subbands is calculated. If each of calculated energy is more than %75, we use n_1 for this band and if each of them are between %5 to %75 we use n_2 else n_3 is used. With this method we implement an optimum adaptive bit allocation that not only has low bit rate but also has better quality response.

V. CONCLUSIONS

According to the sections 3, 4 and implementation of that filter banks on the Farsi speech sentences and also high bit allocation for lower band and low bit allocation for other bands, we achieve advantages that non-conditional FIR filter bank ($H_1(z) \neq H_0(-z)$), in comparison with other two types has better performance and high grade in compression. Also in experimental few of the Farsi characters like "ch" and "sh" are affected trough this filter bank but another characters remain non-distortion.

REFERENCES

[1] A.Alexandrou, "Design of Filterbanks for Subband Coding Systems", Department of Electrical Engineering McGill University Montreal, Quebec, June 1985.

[2] A.Komaee and A.Sepehri "Filter Bank Design and Subband Coding", Department of Electrical and Computer Engineering University of Maryland.

[3] J.R.Deller and J.H.L.Hansen and J.G.Proakis "Discrete-Time Processing Of Speech Signals" The Institute of Electrical and Electronics Engineers Inc, New York.

[4] = B.Francis and S.Dasgupta "Signal Compression by Subband Coding" July10, 1998 Electrical and Computer Engineering, University of Toronto of Canada and University of Iowa of USA.