Defining an optimal set of decimation ratios corresponding to signal and applying in audio signal subband coding

Nguyen Xuan Truong
Department of Technological and Scientific Research
The Audio-Video Broadcasting College I
Phu ly town, Ha nam province
The Voice of Vietnam (VOV)
Fax: +84 351 854383
Tel: +84 913067429
Email: nguyenxuantruong@vov.org.vn

Key words: SBC, MFB, SRE

Abstract: The SBC (subband coding) using the multirate filter bank was proposed in 1980, this coding method splits full audio spectrum into subbands, and assigns more bits to each subband with higher energy. If the power spectral density (PSD) of any subbands is approximately zero, then those subbands need not be assigned any bits at all. While reconstruction, the subband signals are synthesized to the original signal. The most important advantage of SBC is data reduction. The multirate filter bank (MFB) must obtain requirement of perfect reconstruction and requirement of maximal decimation. This paper proposed two new methods to find out sets of decimation ratios, and an optimal set is chosen for designing MFB, satisfying above requirements. Because the optimal set is suitable with PSD distribution of signal, the achieved result does not help only to decrease data rate but also to reduce total error on the reconstructed output signal. An example where an audio signal is the input signal is included.

1. Introduction

Generally, the PSD distribution of any signal concentrates in low frequecies and decreases to higher frequencies. The bandwidth of signal is finite from 0 to F_{max}. To have higher rate of data reduction, SBC using multirate filter bank must have unequal subbands and the lower frequency domain, the more narrow bandwidth of subbands (decimation ratios decrease along frequency axis). The number of subbands is as many as possible, but more complicated and more costly.

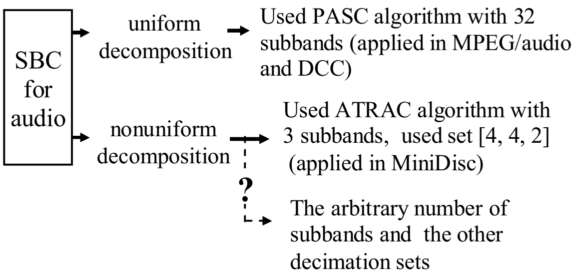

Figure 1: Summation in subband coding for audio signal

The SBC using Wavelet decomposition was applied in audio signal processing: precision adaptive subband coding (PASC) and adaptive transform acoustic coding (ATRAC). In PASC, the 32 subbands are equal. But in the ATRAC, 2 out of 3 subbands are equal [1]. Because PSD distribution of any signal varies as frequency, the widthbands of subbands must be unequal so that highly data reduced rate is obtained. This paper proposed the ways to overcome the above problems. Two new methods are proposed to find out sets of decimation ratios. They are more general than Wavelet decomposition. The summation in subband coding for audio signal is shown in figure 1.

1. 1 Introduction in multirate filter bank

The block scheme of SBC is shown in figure 2. The MFB with M channels consists of the analysis filter bank and synthesis filter bank.

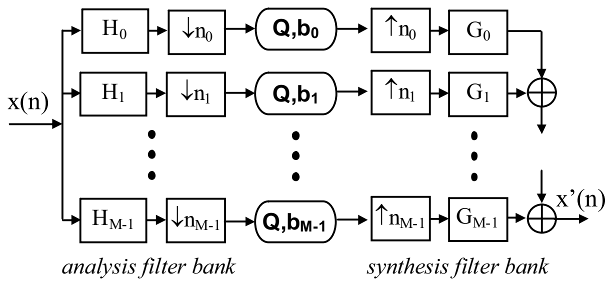

Figure 2: The SBC scheme using MFB with M channels

Where H_0 and G_0 are a lowpass filters (LPF), from H_1 to H_{M-2} and G_1 to G_{M-2} are a bandpass filters (BPF), H_{M-1} and G_{M-1} are highpass filters (HPF). The multirate filter bank must obtain 2 requirements as follows [2], [3]:

(i) The requirement of perfect reconstruction.

(ii) The requirement of maximal decimation, that is Eq.(1).

$$\sum_{i=0}^{M-1} \frac{1}{n_i} = 1 \qquad (1)$$

Where n_i are integer and positive number. If $n_0=n_1=n_2=\ldots=n_{M-1}=M$, the filter bank is uniform decomposition, its decimation ratios always obtain the requirement of maximal decimation. The others is nonuniform decomposition.

The signal processing consists of quantization Q, coding with the number of bits b_i for i^{th} channel, transmission (or storage), and then decoding to reconstruct original signal. The subband signals are quantized using different number of bits. The importance of SBC is the process of bit allocaton. The success of SBC is highly dependent on the way the bit allocation is performed. An intuitive way to accomplish this task is to assign more bits to the channel with higher energies. In the extreme case, some of the channels are identically zero, then those channels need not be assigned any bits at all. The bit allocation will not be shown in this paper.

To estimate quality of SBC, that is a distortion between the input x(n) and the reconstructed output x'(n), we use the mean-squared error [4] as follows:

$$\varepsilon = [\ |x(n)-x'(n)|^2] \qquad (2)$$

With b_i and n_i are number of bits and decimation ratio of i^{th} channel, the average bit rate of SBC with M channels is as follows:

$$b = \sum_{i=0}^{M-1} \frac{b_i}{n_i} \qquad (3)$$

For SBC using unifrom decomposition with M channels, the average bit rate is as follows:

$$b = \sum_{i=0}^{M-1} \frac{b_i}{M} \qquad (4)$$

Hence an optimal subband coder is the one that minimizes the error as the Eq. (2) for fixed average bit rate b bits/sample as Eq. (3) and the Eq. (4). Optimal subband coding involves both designing the filter bank and doing the allocation of bits in an optimal fashion. If $\sigma^2_{x_i}$ is variance of the subband signal in i^{th} channel and the coefficient c depends on the statistical property of signal resource and the quantizers operate at high bit rates, then the mean-squared error is as the Eq. (5) [5]:

$$\varepsilon = c.2^{-2b}\left(\prod_{i=0}^{M-1}\sigma^2_{x_i}\right)^{\frac{1}{n_i}} \qquad (5)$$

1. 2 Two methods to define decimation ratios

Two methods are proposed in this paper to find out the different sets of decimation ratios. They are more general than the Wavelet decomposition.

a- The multistage method

The first method is called a multistage method as shown in figure 3, at first the original point is divided into some branchs and a coefficient on each branch equals the number of branchs. Do also like that for each branch and so on. When stopping, each ratio is multiplication of all coefficients on their branchs. This method offers sets of decimation ratios obtaining requirement of maximal decimation.

The ratios in every branch must be integer, positive number and at least 2. The Wavelet decomposition is the individual case of multistage method with the ratio 2 of all branchs.

b- The method of consecutive plus the end

The second method is consecutive plus the end. This method uses below the Eq. (6):

$$\frac{1}{1.2}+\frac{1}{2.3}+\frac{1}{3.4}+\ldots+\frac{1}{(M-1).M}+\frac{1}{M}= \qquad (6)$$
$$=\frac{1}{1}-\frac{1}{2}+\frac{1}{2}-\frac{1}{3}+\ldots+\frac{1}{M-1}-\frac{1}{M}+\frac{1}{M}=1$$

Where M is a positive integer and at least 2. Therefore, the set [2, 6, 12, …(M-1).M, M] also obtains requirement of maximal decimation. The Wavelet decomposition is also the individual case of this method with M=2.

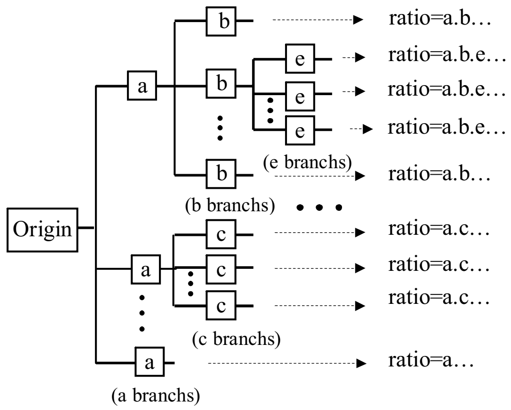

Figure 3: The scheme of the multistage method

Assume that k is the number of ratios in the set, there are k! available sets by permutation. As the method of consecutive plus the end, three ratios are (2, 6, 3), there are 3!=1.2.3=6 sets of decimation ratios, they are [2, 3, 6], [3, 2, 6], [2, 6, 3], [6, 3, 2], [6, 2, 3] and [3, 6, 2]. An optimal set is chosen and designing SBC by MFB with the optimal choosed set of decimation ratios is implemented.

2. The subband coding with the set [6, 3, 2]

2. 1 The reason of choosing the set [6, 3, 2]

At first, one specific audio file is recorded in computer. As (1.2) section, all nine found sets for SBC 3 subbands are [3,3,2]; [2,4,4]; [4,4,2]; [6,3,2]; [6,2,3]; [3,6,2]; [2,3,6]; [2,6,3]; [3,2,6]. They are shown in table 1. By calculating the reconstruction errors for all the 9 sets with the same audio file, we will choose the optimal set. If reconstruction error as Eq. (2) is minimum, the set is optimal. For given b bit rate and without loss of generality ignoring the coefficient $c.2^{-2b}$, as the Eq. (5) the scaled reconstruction error (SRE) is as following Eq. (7).

$$\text{SRE} = \left(\prod_{i=0}^{M-1}\sigma^2_{x_i}\right)^{\frac{1}{n_i}} \qquad (7)$$

The SRE depending on the sets are shown in the figure 4. We do same operation for many different audio files, the results are just the same. So the set [6, 3, 2] is optimal.

Because the PSD distribution of any signal (audio or video) concentrates in low frequecies and it is decreased to high frequencies, for optimal coding we split lower frequecy band into more narrow subbands, so the bigger decimation ratios is in lower frequency band. Thus the set [6, 3, 2] is optimal for SBC with M=3 channels.

Table 1: The available decimation sets with M=3.

N^0	Sets of decimation ratios			Decompo-sition	Method decimation
	n_0	n_1	n_2		
1	3	3	3	Unifrom	Multistage
2	2	4	4	Nonunifrom	Wavelet (as multistage)
3	4	4	2	Nonunifrom	Wavelet (as multistage)
4	6	3	2	Nonunifrom	Consecutive plus the end
5	6	2	3	Nonunifrom	Consecutive plus the end
6	3	6	2	Nonunifrom	Consecutive plus the end
7	2	3	6	Nonunifrom	Consecutive plus the end
8	2	6	3	Nonunifrom	Consecutive plus the end
9	3	2	6	Nonunifrom	Consecutive plus the end

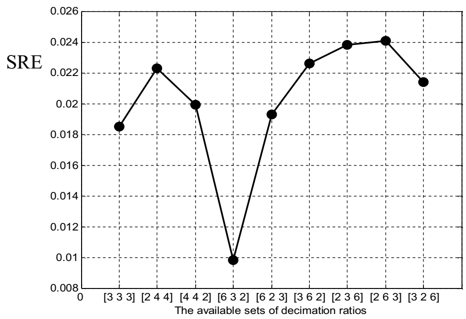

Figure 4: Variation of SRE depending on the sets

2. 2 The signal processing in SBC with the set [6, 3, 2]

The block scheme of the SBC using multirate filter bank with nonunifrom decomposition M=3 channels is shown in figure 5. Where H_0 filters out 1/6 of bandwidth of input signal in lowest frequency domain, H_1 filters out 1/3 of bandwidth of input signal in intermediate frequency domain, H_2 filters out 1/2 of bandwidth of input signal in highest frequency domain. The filters H_0 and G_0 are low pass filters with cutoff frequency 1/6. The filters H_1 and G_1 are the band pass filters with cutoff frequecies 1/6 and 1/2. The filters H_2 and G_2 are the high pass filters with cutoff frequency 1/2. Note that, this whole paper uses the scaled frequency axis Fn, that is as follows the Eq. (8):

$$\mathrm{Fn} = \frac{2.f\,(\mathrm{Hz})}{\mathrm{Fs}} = \frac{\omega\,(\mathrm{rad/s})}{\pi.\mathrm{Fs}} \qquad (8)$$

For anti-aliasing, we change ratios of decimation blocks and interpolation blocks from [6, 3, 2] into [6, 2, 2]. That is very effective way to obtain both the requirement of maximal decimation and the requirement of perfect reconstruction. Assume that the input signal x(n) is sampled from analog signal with sampling frequency $Fs=F_{Nyq.}$, $F_{Nyq.}$ is Nyquist frequency, $F_{Nyq.}=2.F_{max}$ (F_{max} is the highest frequency of the input signal). Thus, the spectrum of the signal x(n) is shown in figure 6a. All signal processing in frequency domain is shown in figure 6 [1], [6].

At first, the H_0, H_1 and H_2 split full spectrum of input signal x(n) into 3 subbands, the bandwidths of them are respectively equal 1/6, 1/3, and 1/2 the full spectrum of input signal. The spectrums of them are shown respectively in figure 6b. When passing decimation block, the spectrum of signal is extended by the decimation ratio times. Their spectrums are shown in figure 6c. And then, the signals of subbands are quantized Q and coded with different number of bits b_0, b_1, b_2. Finally the digital signals are stored or transmited.

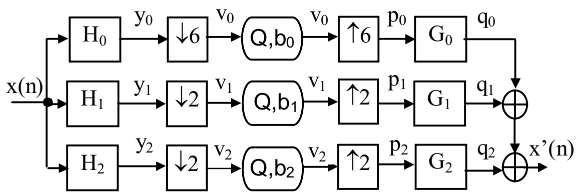

Figure 5: The block scheme of multirate filter, nonunifrom decomposition, 3 channel

The signal reconstruction contraries to the signal analysis. The signals of subbands pass the interpolation blocks and the synthesis filter bank. In frequency domain, the spectrum of signal passing the interpolation block is shortened by the interpolation ratio times [1]. The spectrums of 3 signals passing 3 interpolation blocks with 3 ratios 6, 2, 2 are shown in figure 6d.

Then, the signals pass the filters to eliminate all images after interpolation, the outputs only take out the useful spectrum components. They are shown in figure 6e. Finally, the useful spectrum components are synthesized to reconstruct full spectrum of original signal. The spectrum of the reconstructed signal x'(n) is shown in figure 6f. The spectrum of reconstructed signal x'(n) is identical to the spectrum of original signal x(n).

The SBC with set of [6,3,2] is designed and apply in audio signal coding. It obtains the requirement of perfect reconstruction and requirement of maximal decimation.

2. 3 Simulation with programme MATLAB 7.0

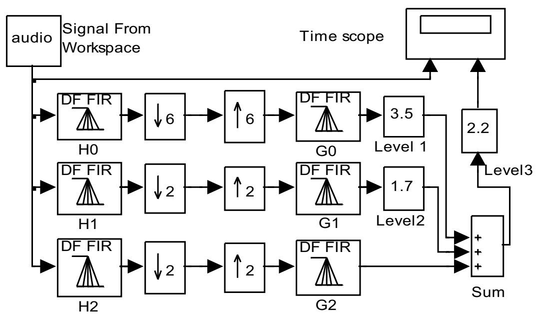

Figure 7: The simulated scheme of SBC using multirate filter bank with decimation set [6, 3, 2] for audio signal

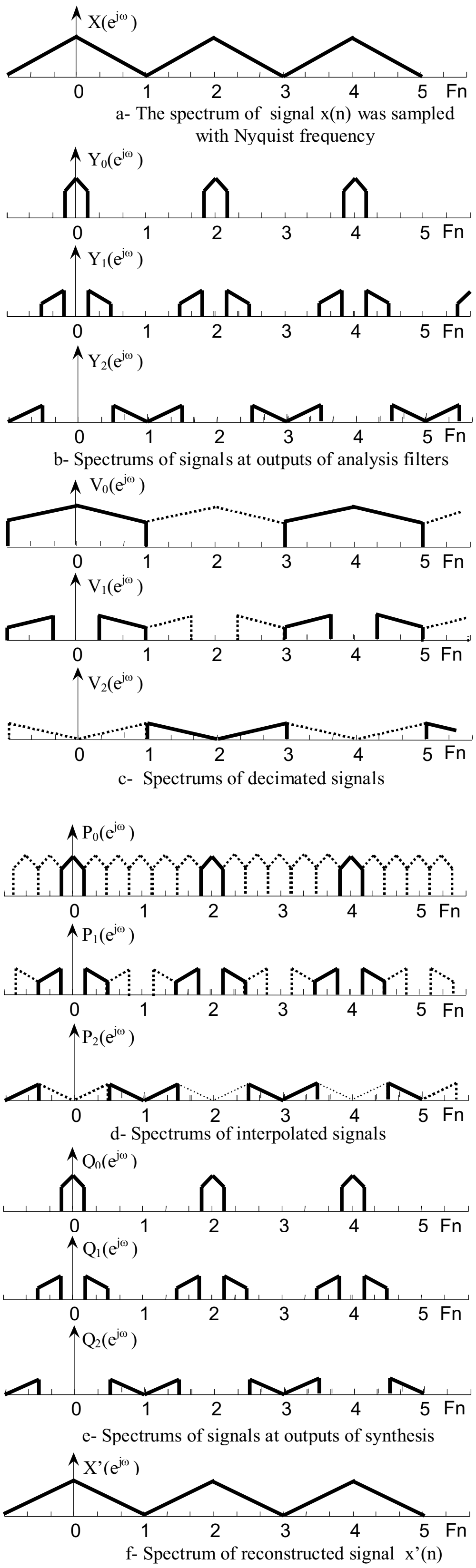

Figure 6: The spectrums of signals in processing

The MATLAB 7.0 scheme simulating results is shown in figure 7. The audio file is recorded from micro and saved in computer. The time scope is used to indicate the input signal and the reconstructed signal, they are shown in figure 8a and 8b, respectively. They are the almost identical waveforms.

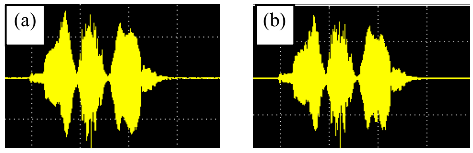

Figure 8: The waveform of original input signal (a) and of reconstructed output signal (b)

3. Conclusions

The two proposed new methods can find out many other sets of decimation ratios. Moreover, on the results offered in this paper we can code signal by SBC with increasing number of subbands. More subbands achieve higher data reduction with acceptable reconstruction error. The above methods are combined with the threshold of hearing and masking effect of human ears, audio coding is more effective. On this results, we can design SBC using multirate filter bank more channels. The two important results are obtained in this paper as follows:

(i) The methods of finding decimation sets for multirate bank and the method of choosing the only one optimal decimation set of them, it is suitable with specific input signal.

(ii) The method of signal processing, that changes the decimation set [6, 3, 2] into the decimation set [6, 2, 2]. When this method obtains not only the requirement of maximum decimation but also the requirement of and pecfert reconstruction. This method is more general than the Wavelet decomposition.

References

[1] Ken C.Pohlmann: "*Principles of digital audio*". 3rd edition by McGraw-Hill Book Company, 1995.

[2] Quoc Trung Nguyen: "*The digital signal processing and digital filter*". Printed in the scientific anh technical publisher Hanoi, Vietnam, 2001.

[3] Robert G.Winch, "*Telecommunication transmission systems*", published by McGraw-Hill Book Company.

[4] Dag Stranneby: "*Digital signal processing and applications*". 1st Edition in Great Britain, 2001.

[5] Mischa Schwartz, "*Information transmission modulation and noise*", 3rd edition by McGraw-Hill Book Company, 1980.

[6] John Watkinson, "*An introduction to digital audio*", first published in Great Britain, 1994.