Subband Simulation in MATLAB

VLADISLAV SKORPIL, ABDULHAKIM ABUZAHU
Department of Telekommunications,
Brno University of Technology,
Purkynova 118, 612 00 Brno,
CZECH REPUBLIC,
http://www.vutbr.cz/

Abstract: - In the start of the flow chart, there are two steps, which are prepared for new input samples. Firstly, all the samples in the buffer (512 samples) are down shifted and then put new 32 samples at the end of the buffer, which is required for further calculations. But in this work different algorithm, which contain both in only one step is used. One more variable known as "ofs" is present, which shows where is the end of the previous data read, because 32 samples into 512 samples buffer is read, it is different number every time.

Key-Words: - Subband, sample, buffer, algorithm

1 Introduction

Instead of shifting in each run of flowchart (it is 12 times in one frame) only offset is shifted. It gives more time for other algorithms. When 32 new samples are added, it gives 512 together, which is all defined length for buffer for windowing. In next step is windowing with coefficients take from standard. Figure 1 shows analysis of filter bank. These are coefficients for matrixing in previous flowchart. These coefficients are calculated by the following formula:

$$Mik = \cos\left[(2i+1)(k-16)\frac{p}{64}\right] \quad (1)$$

2 Evaluation of Signal Quality

Signal quality can be evaluated using various criteria; it can essentially divide these into subjective and objective criteria. Subjective parameters of an acoustic signal are determined on the basis of the psycho acoustic properties of the hearing equipment. The objective parameters of the signal are the physical parameters. These include parameters in the temporal domain (middle value, variance, slope and peak coefficients, correlation functions, etc.) and the frequency

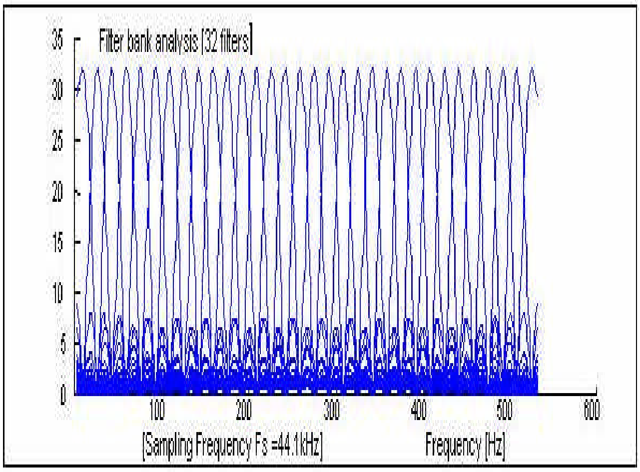

Fig.1 Filter bank analysis

domain (spectrum or the spectral output density of the signal). Research on the subjective quality of acoustic signals both speech and music has received a lot of attention, as no objective criterion capable of fully replacing subjective listening tests exists at present, a detailed methodology for performing listening tests exists under e.g. ITU-R.. In this work, I use, for listening tests, an evaluation scale based on the recommendations of ITU-R; see Tab. 1. Since listening tests place high demands on equipment that I cannot meet, and since I cannot perform tests on large groups of people and then statistically process them, my results may, however, differ from those acquired precisely according to recommendations, and are for information only. The median quadratic difference between the original signal and the compressed signal is an important objective criterion for judging signal quality. This criterion can in some cases also convey the subjective quality of a signal. In the case of the MPEG codec, psychoacoustic model guarantees (within the limits of its precision) that distortion caused by compression will be as inaudible as possible and, to some degree, will be remedied by the listener.

Description of Worsening of Quality	ITU-R Degree
Unnoticeable	5
Noticeable, but not disturbing	4
Mildly disturbing	3
Disturbing	2
Very disturbing	1

Table 1 Five Degree scale fro the worsening of Signal Quality (ITU-R)

3 Scale factor Calculating

A scale technique is used in the MPEG-1 coding scheme, which provides an effective overall dynamic range of 120 dB/subband with a resolution of 2dB/scale factor class. The calculation of the scale factors is made for every 12 subband samples. The maximum value of these samples is determined. The lowest value, which is larger than this maximum value is used as scale-factor. Scale-factor can be calculated after psychoacoustic model according to standard. In fact, in MATLAB program, the scale-factor was chosen before psychoacoustic model. The algorithm contains to looking for maximum value of samples and after searching minimum appropriate scale-factor.

The calculated scale-factor is presented by 6 bits word. It is transmitted only if a non-zero number of bits have been allocated to the subband. Coding of scale-factors is presented in MATLAB. When minimal scale-factor is find out, then information about it is stored in *scfi(i)* variable. The data about scale-factors are then given to decoders through this global vector of scale-factor indexes.

4 FFT Analysis

The masking threshold is derived from an estimate of the power density spectrum that is calculated by a 512-point FFT. The FFT is calculated directly from the input PCM signal, windowed by a Hann window (Equation 2) displayed in Figure 2.

$$h(i)=\sqrt{\frac{8}{3}\times 0.5\left(1-\cos\left[\frac{2pi}{N}\right]\right)} \qquad 0\leq i\leq N-1 \tag{2}$$

For a coincidence in time between the bit allocation and the corresponding subband samples, the PCM samples entering the FFT have to be displayed. The delay of the analysis subband filter in 256 samples, corresponding to 5.3 ms at the 48 kHz sampling rate. A window shift of 256 samples is required to compensate for the delay in the analysis subband filter. In addition, the Hann window must coincide with the subband samples of the frame. For Layer-1, this amount to an additional window shift of 64 samples.

When computing is at the end of 512 energy samples, values are reindexed. It means that calculated energy is shifting (by indexing) in corresponding position. New indexes are taken by *butterfly* algorithm as shown in Figure 3. It introduced 8-point butterfly for taking computing principle of FFT.

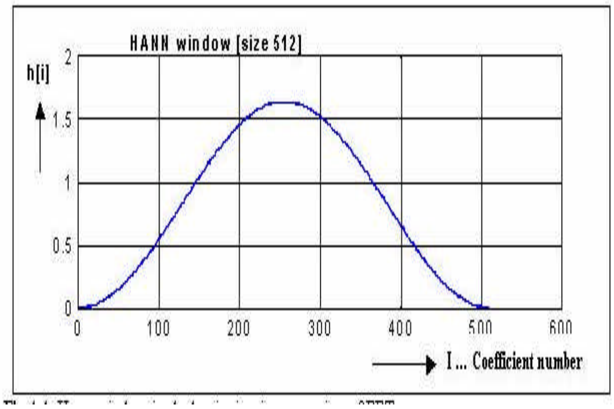

Fig. 2 Hann window in the beginning in computing of FFT

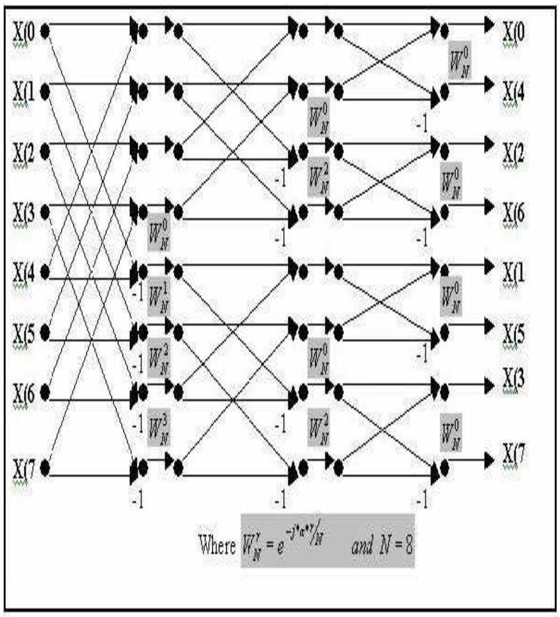

Fig. 3 Flow graph of-point FFT computation

Computation of FFT in MATLAB program can be calculated by only one instruction (procedure), but principle of computing is not possible to see. Here we use just multiplication, addition and subtraction in MATLAB.

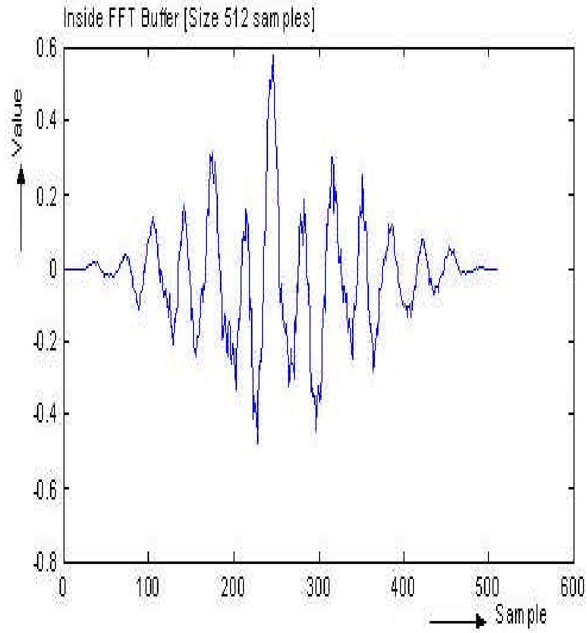

Fig.4 Test signal inside buffer after windowing by Hann window

This part has several loops. Shift of 64 samples is recommended by standard to keep coincidence in time, but before buffer is filled by 384 new samples. For correct mathematical calculations, it is necessary to prepare 512 samples large buffer for FFT. The rest of buffer is filled by zeros. When new samples are put in buffer, is used addressing with offset. This principle is similar as in first step of encoding. After windowing by Hann window as shown in Figure 4, variables are initialized on zeros values. Thus everything is ready for starting of calculating FFT output samples, keeping energy value in each subband. It is 512 point FFT, it gives 256 values to 32 subbands.

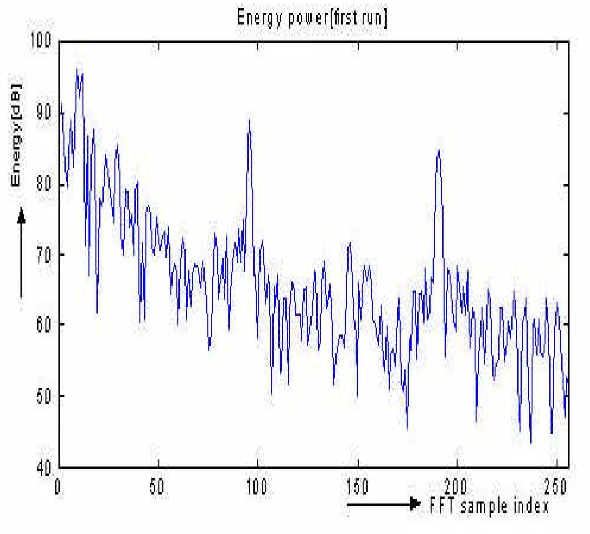

Fig. 5 Energy power calculated in first run

If input signal is applied, MATLAB gives energy spectrum picture as presented in Figure 5. In first run 384 samples are taken, but FFT buffer have positions for 512 equal to the size of FFT. The rest of free positions will be filled by zeros. And after second, third, fourth, and other runs buffer will be full. After first run, the signal will be still the same, so energy in the next runs will remain the same as shown in (Figure 6). This shows spread of power in frequency spectrum.
This is good position in program to start analysis of the input signal. User can see how does input signal looks like in the frequency domain. But for MPEG algorithm, this step is the most important, because further computations are based on the FFT output (energy).

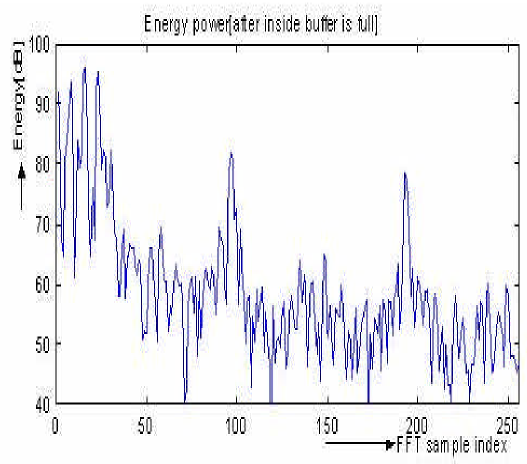

Fig. 6 Energy power calculated after sigth run

5 MATLAB Computing of the Sound Pressure Level

The more powerful version was chosen for the determination of the sound pressure level in program in MATLAB. A maximum is found from all spectral lines in each subband and one pressure level gives by scalefactors. In fact, it means two loops, first is making a searching through spectral lines in each subband, second is searching just through subbands comparing to an appropriate pressure level and maximum from first loop.
In this section of program, only one loop is possible to see. It is just first step from both. second is comparing at the end of the psychoacoustic model, which is given together with calculating *SNR* (signal-to-noise ratio). It is more powerful as one loop is spared.
The tonality of a masking component has an influence on the masking threshold. For this reason, it is worthwhile to discriminate between tonal and non-tonal components. For calculating the global masking threshold, it is necessary to derive the tonal and non-tonal components from the FFT.
So this step was divided into two parts. First is aimed to find tones and second to find noise with high power, which is necessary to transmit. In fact, it is named noise, but it should be many tones closer together. In addition, there is one more part called *'labeling of local maxima'*, but it is concluded in the part for tones.
This step of standard starts with labeling of local maxima. It then extracts tonal components (sinusoids) and calculates the intensity of the non-tonal components within a bandwidth of a critical band. The bandwidth of the critical band varies with the frequency with a bandwidth of about 0.1 kHz at low and with a bandwidth of about 4 kHz at high frequencies. It is known from psychoacoustic experiments that the ear has a better frequency resolution in the lower frequency region than in the higher ones [3]. To determine if a local maximum is examined, the frequency range *df* is given by standard for all sampling frequencies of MPEG-1.

6 Conclusion

For further information about df for different frequencies and different layers is better to study the Reccomendationn just one possibility for Layer-1 and 44.1 kHz as sampling frequency. Choosing it in the beginning of this work causes this aim. Moreover, sampling frequency 44.1 kHz is the most popular throughout the world. The second sampling frequency used very often for digital audio is 48 kHz. First is based mostly on commercial using and second on professional devices means studios.

References:

[1] ISO/IEC 13818-7 Information Technology – Generic coding of moving pictures and Associated Audio Information – part 7: Advanced Audio Coding (AAC), 1996, 131s.

[2] BOSI M. and RANDENBURG, K. ISO/IEC MPEG-2, Advanced Audio Coding. Journal of the Audio Engineering Society, Vol.45, 1997, No. 10,PP. 789-814

[3] SERANTES, C. PENA A., Prelcic N. A Fast Noise-scaling Algorithm for Uniform Quantization in Audio Coding schemes. IEEE Proceedings, 1997 , PP 339-342

[4].ABUZAHU,A. Digital Audio Compression. Ph.D. Thesis, BUT, Brno 2003

Acknowledgement:

This research was supported by the grants: No 102/03/0434 Limits for broad-band signal transmission on the twisted pairs and other system co-existence. The Grant Agency of the Czech Republic (GACR)

No 102/03/0260 Development of network communication application programming interface for new generation of mobile and wireless terminals. The Grant Agency of the Czech Republic (GACR)

No 102/03/0560 New methods for location and verification of compliance of quality of service in new generation networks. The Grant Agency of the Czech Republic (GACR)

No MS 1850022 Research of communication systems and technologies (Research design)

Grant 2811 F1 Advanced Technology of Transport Networks in Education (grant of the Czech Ministry of Education, Youth and Sports)

Grant 3112 F1 Inovation of Education of Last Mile Data Transmission (grant of the Czech Ministry of Education, Youth and Sports)