Module 9

AUDIO CODING

Lesson 31

Format and encoding

Instructional Objectives

At the end of this lesson, the students should be able to:

1. Show the flow diagram for encoding an audio frame
2. State the role of bit allocation section of the encoder
3. State the role of the scale factor section of the encoder
4. Distinguish between layer-1 and layer-2 encoding.
5. Explain the bit allocation algorithm.

31.0 Introduction

The implementation of polyphase filters was described in lesson-30. We now discuss how the filter outputs are to be endoded before encoding into a bit stream.

The MPEG audio compression scheme uses an adaptive bit allocation policy where the scale factors and the number of bit per sample vary from frame to frame, as determined by the psychoacoustic model. MPEG-1 is based on three layers. In this lesson we shall cover the simplest, that is layer-1 followed by layer-2. The bit allocation algorithm will be illustrated with an example.

31.1 Layer –1 encoding scheme

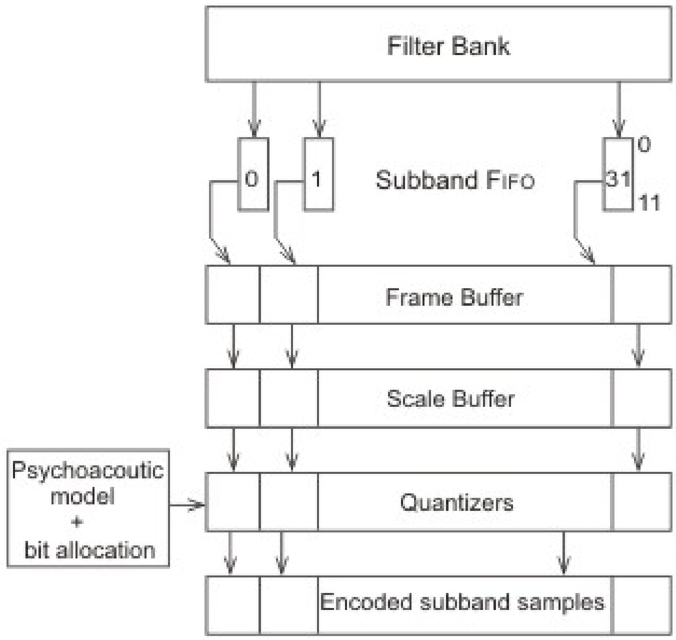

FIGURE 31.1 Flow - diagram for layer-1 encoding.

Fig 31.1 shows the flow diagram for processing the subband samples, obtained from the polyphase filters, as described in lesson-30.

In layer – 1 encoding, 12 samples from each of the 32 subbands, i.e. 384 samples constitute an audio frame. Each subband has a FIFO to store these 12 samples. For each of the subbands, the 12 samples are scaled, so that the maximum value is closest to, but does not exceed one. The psychoacoustic model and the target bit rate are used to compute the number of bits allocated per sample for each of the subbands. The samples are thereafter quantized to the required number of levels. The bit allocation, the scale factor and the encoded samples are placed in three designated areas of the bit stream.

31.2 Bit allocation section

The bit allocation section of the flow diagram shown in fig.31.1 specifies the word lengths (in bits) for encoding the samples belonging to a subband. This in turn decides the number of levels in the quantizer. Increasing the number of levels reduces the quantization noise, but increases the bit rate. The quantizer field has 4-bits assigned to each subband. However, the code 1111 is invalid, since four consecutive 1 s form a part of the synchronization code and hence, there are 15 different quanitizers to choose from.

31.2 Scale factor section

The scale factor section contains 32 six-bit scale factor values, one for each subband. Here too, the combination of all 1s is not permitted, leading to 63 different quantizer values. The scale factor value depends on the maximum value of the subband samples and the scale factor values are used to multiply the requantized sample values.

The MPEG-1 standard specifies an n-level mid–tread quantizer, where n is one less than the antilog to the base-2 of the number of bits assigned to the subband. For example, 3-bits per sample in a subband means a 7- level mid-tread quantizer, as shown in fig 31.2. Again, the code of all 1s is not permitted.

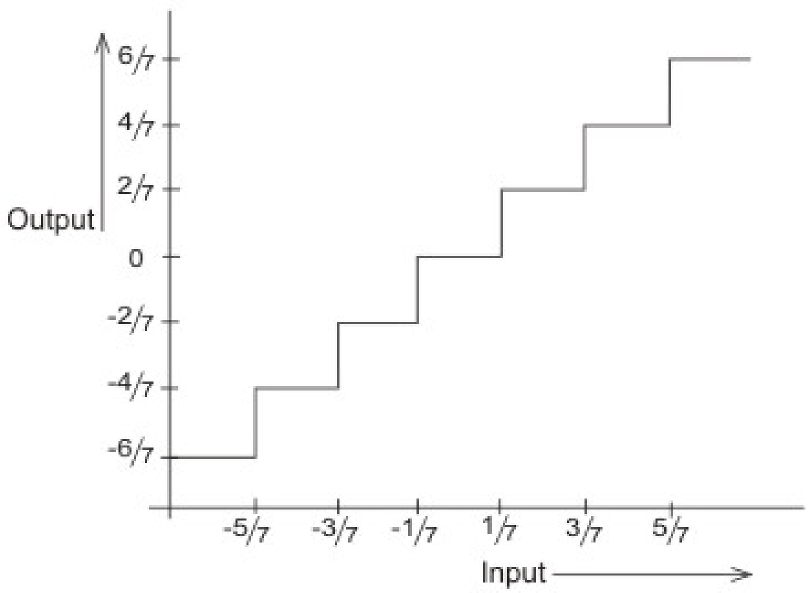

FIGURE 31.2 Example of 7-level midtread quantizer. The allowable encoded sample values are 000,001,............,110. The code 111 is not allowed.

31.3 Layer-3 encoding scheme

The layer-2 encoding uses the same psychoacoustic model as layer-1 and is an extended form of layer-1 encoding. The number of samples per subband in an audio frame is increased to 36 (i.e. 1152 samples per audio frame). This is divided into three parts, called part-0, part 1 and part-2. Each part contains 12 samples and can be treated like layer-1 encoding. All the three parts contains a single bit allocation, but the scale factors may be different. The scale factor section in layer-2 contains a 2-bit scale factor selection information (*sfsi*), which indicates whether one, two or three scale factors are transmitted, and how they are applied.

The samples are quantized and coded in a method similar to that used for layer-1. In layer-2, there is a provision to pack three consecutive samples in a single codeword for certain quantizers.

The bit allocation section is also modified in layer-2. Instead of always sending 4 bits per subband in order to specify the bit allocation choice, the numbers of bits vary from 0 to 4 as a function of subband number.

Fig 31.3 shows the frame formats for layer–I and layer-II bit streams

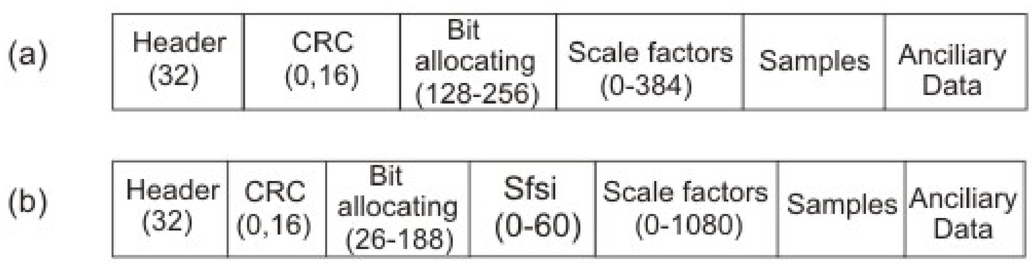

Figure 31.3 Frame format for (a) Layer - I and (b) Layer - II bit stream

31.4 Bit –allocation algorithm

MPEG-1 audio standard offers a compression ratio of 5:1, which roughly corresponds to 3-bits per sample. In terms of PCM, 3-bits/sample corresponds to a meager 18-dB of signal-to-noise ratio (SNR), yet MPEG-1 audio standard is able to provide noise suppression better than 60-dB, by effectively utilizing the masking characteristics and the psychoacoustic models. The psychoacoustic model determines from the input audio the maximum noise level, which would just be perceptible for each of the subbands. As the amount of quantization noise is directly related to the number of bits used by the qualtizer, the bit allocation algorithm assigns the available bits in a manner, which minimizes the audible distortion.

For layer-I and II, the bit allocation process starts by computing the mask-tonoise ratio (*MNR*) for each subbands

$$MNR = SNR - SMR$$

The psychoacoustic models provide the signal-to-mask ratio (*SMR*) and the *SNR* is given as a function of bit allocation tables provided in the standard.

The algorithm finds the subband with the lowest *MNR*, whose bit allocation has not reached the maximum limit. The bit allocation for that subband is increased by one level and the number of additional bits required is subtracted from the total number of bits available, after subtracting the header bits, CRC bits, bit allocation data and the anciliary data. With new bit allocation for the subbands, its *SNR* and hence *MNR* improves. Now, the subband with lowest *MNR* is searched once more and the bit allocation process is repeated until all the available bits are used or all the subbands have reached their maximum bit limits.