1. Abstract

Reduction in bit rate requirement for high quality audio has been an attractive proposition for multimedia applications and digital broadcasting. In this project, we have used the Wavelet Transform as an effective tool for data compression and achieve the high quality low complexity scalable audio coding. A 6-levels of Wavelet decomposition was carried out in order to obtain 22 critical bands. An auditory masking model was used to determine the number of bits required for each critical band and still maintaining a good quality of the audio signal. The compressed audio signal was reconstructed using the Inverse Wavelet Transform.

2. Implementation of the Audio Coder

Our Project coder structure of a Wavelet Packet based audio coder is shown in Figure 2.1. The sampling frequency of the audio signal was 16 kHz and 256 samples per frame. A six-level of wavelet decomposition was carried out thus resulting in a 64 band WPT.

In our project implementation, the 64 band were grouped together in a particular manner to obtain 22 critical bands (see Figure 2.2) and an auditory masking model was then directly applied in the wavelet domain. The maximum signal energy and the masking threshold in each band was calculated. The masking model output was used to determine the bit allocation per subband for perceptually lossless quantisation. The samples were then scaled quantised according to the subband allocation. Figure 2.3 show a Two-channel Subband Coder using Quadrature Mirror Filtering (QMF).

3. Implementation of the Auditory Masking Model

The auditory model used in this project is similar to the one used by Black and Zeytinoglu (1995). The steps involved in calculating the masking threshold per critical band are as follows:-

- Calculate the maximum power per critical band.
- Calculate the centre frequency in Barks.
- Identify the masker in a critical band and calculate the amount of masking it introduces other critical bands.
- Calculate the value of self masking.

4. Conclusion

We had achieved a bit rate reduction from 256 kbits/sec to 112 kbits/sec with a compression ratio of 2.3. A subjective listening test was conducted in order to evaluate the performance of the coder/decoder. 20 people participated in this test and a mean opinion score of 4.7 was achieved. As the critical band decreases, the audio quality also degrades. Figure 4.1 and Figure 4.2 show the original and wavelet reconstructed audio signal using 22 critical band respectively.