Brad Erwin
Mary Lou Kesse
11/16/98

MPEG 1 – Layer III Audio Codec

Data Sheet

This document will elaborate on the MPEG-1 Layer III Codec system block diagram presented in the previous assignment and reproduced here as figure 1. Again, since our project is being done in a simulation phase and a hardware build phase, we will delineate the differences within the documentation. Generally, the encoder portion of the project will be simulation only, while the decoder portion will be both simulated and built in the lab.

Input & Encode

The PCM (pulse-code modulated) digital audio input signal shown in block 1 is the input to the system. This signal will be a monophonic, 16-bit signal sampled at 44.1KHz. For simulation purposes, we will run tests using WAV files as system inputs, making inspection of the system output convenient using the department's sound editor Cool Edit. The encode subsystem shown in blocks 2 through 8 will be implemented for simulation on MATLAB. These blocks decompose the digital data from the WAV file (block 1) into specific frequency subbands and quantize each of these according to the psychoacoustic model. The digital audio output is fed simultaneously into two paths, both of which will compress the data. The first of these paths is the filter bank (block 2). Each will be at least a 200th order FIR filter. These will dissect the signal into 32 subbands, each with a bandwidth of 750Hz. An 18-channel modified discrete cosine transform (MDCT) filter bank (block 3) is applied to break each of the 32 subbands into 18 separate frequency bands, for a total of 576 bands. Next, each subband is downsampled (Block 4) to 1.5KHz by retaining every 32nd data point and discarding all other data points (48Khz / 32 = 1.5KHz).

The second data compression path includes two subsystems – a 1024 point Fast Fourier Transform (FFT) and the psychoacoustic model (blocks 6 and 7, respectively). The FFT supplies the psychoacoustic model with the frequency components of the input signal. The model uses this information to identify components of the input signal that are imperceptible to the human ear. The quantizer (block 7) uses the spurious signal components detected by the psychoacoustic model to determine how each of the 576 subband signals should be encoded. Subbands containing spurious data below a certain threshold will be discarded or encoded with less accuracy than subbands with more perceptually important signals. The actual implementation of the psychoacoustic model is still being researched. There are many different approaches to this, including spreading convolution, time domain alias canceling, and signal-to-mask ratio threshold checking. Due to the limited signal processing capabilities of MATLAB with data arrays of this magnitude, we may not be able to simulate this block. The bitstream produced by the quantizer is formatted into standard MP3 frames in block 9. The structure of these frames is also quite complex. Each frame, in addition to samples and scale factors, requires a 32-bit header, a 16-bit checksum, and a maximum of 136,256 bits of ancillary data that we will most likely exclude from simulation. The samples and scale factors are coded according to a “bit reservoir” system that is implemented to further reduce bitstream size. Depending on signals within a specific band, this system will use the least number of bits to encode information. For example, a dense spectrum across a certain band will use many bits compared to a band containing only a quiet low frequency tone. Thus, the dense band “borrows” bits from the reservoir while the sparse band “refills” it. As with the psychoacoustic model, complete implementation of this block in MATLAB may not be feasible or necessary for our understanding of the project.

Decoding

Blocks 9 through 12 make up the decoder subsystem and will be used for both simulation and hardware phases of the project. The MPEG data frames are fed into the bitstream extraction logic (block 9), which recovers the audio bitstream in 576 bands from the data frames. The inverse modified discrete cosine transform (IMDCT, block 10) receives this data and reconstructs the signal into the 32 subbands. All of the subbands are summed (block 11) and delivered as PCM digital audio (block 12).

Simulation Phase

As previously mentioned, the simulation phase of the project will first utilize the department’s MATLAB and Cool Edit software packages. WAV files will serve as convenient inputs and outputs of the system. Currently, implementation of the bitstream formatting block, the bitstream extraction block, and the psychoacoustic model block has been disregarded. By comparing the input signal with various signals created within the system, we will be able to see which pieces are working properly. Cool Edit provides a spectrum view of the signals as well as a FFT function to inspect the frequency components of each file. Also, this program will play the output files so we can listen to the input compared to the output to find any audible differences between the real and compressed bitstreams. For ideal MPEG decoding, the output file should be 12 times smaller than the input signal, but sound just as good.

Hardware Phase

Since our physical deliverable is a standalone MPEG decoder, we will only be implementing the decoder subsystem in hardware. Its input will be any standard MP3 file – many of which are readily available for download from the Internet. The said encoded file will be stored in the external evaluation board memory. Blocks 9 through 11 will be implemented using a currently unspecified Texas Instruments DSP microprocessor. Decoding software will be written in C. This code will be tested on a DSP simulator before implementation on an evaluation board. The chip needs to operate fast enough to perform the necessary real-time decode from memory as well as create a 44.1KHz 16-bit parallel output. As in the simulation phase, the output of the decoder should sound like the MP3 file stored in memory.

Brad Erwin
Mary Lou Kesse
11/16/98

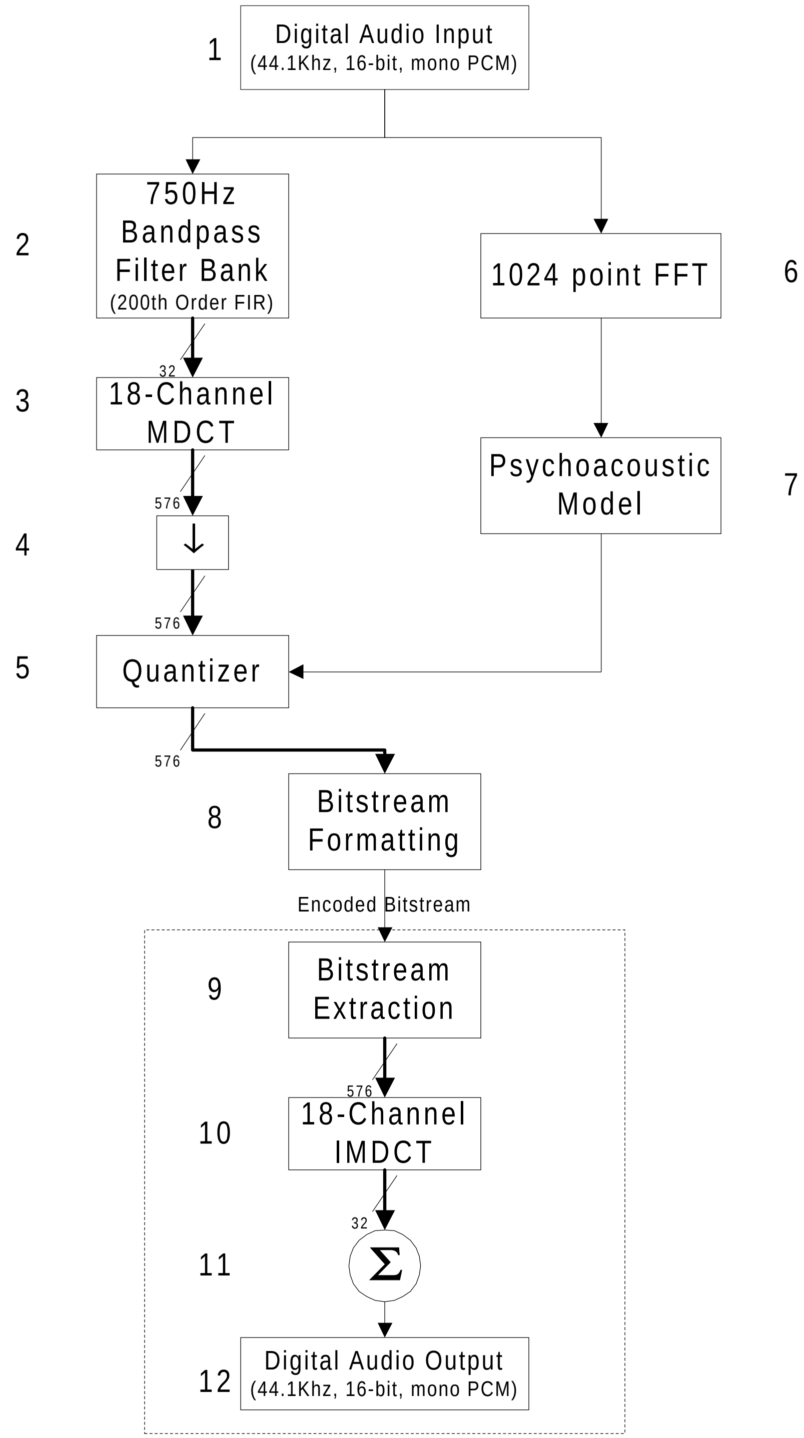

Figure 1
(Dashed Line indicates Hardware Phase of Project)