Exercise Multimedia Technology
WS 2003/2004

Sheet 5 (November 20th, 2003)

Exercise 5.1 Digitizing, coding and audio

In order to be digitized an audio signal is first captured by a microphone. The resulting amplitude oscillates between -3*U_0 and 3U_0 volt. U_0 denotes the size of a quantization interval in volt while t_0 denotes the quantization interval in the time domain.

1. How many bits are needed to store a single sample?

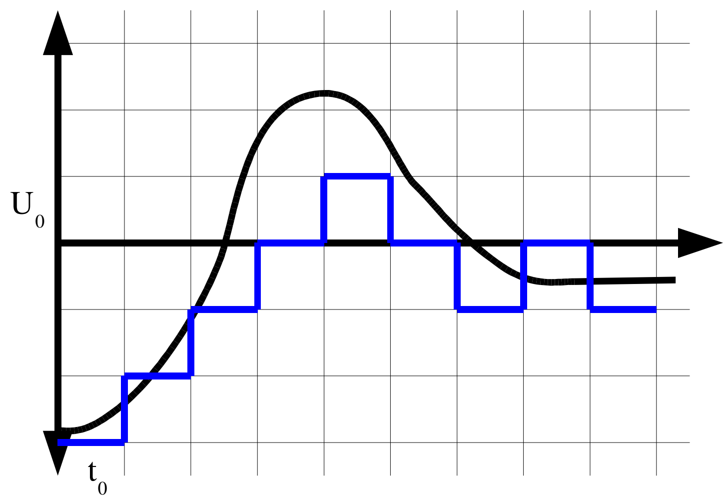
   The signal ranges from -3 to 2. In this case 3 bits are sufficient to code 6 states.

2. How many bits are used in the delta modulation?

   Only one bit. 0 = signal falling by one quantization unit / 1 = signal rising by one quantization unit

3. Code the signal from the following diagram using delta-modulation (draw right into the figure).

   Assuming that the signal is at -3 in the first time interval, the blue curve will result in the least possible error which delta modulation allows.



4. What are the disadvantages of delta coding?

   Whenever the first derivation of the signal (within a time unit) is greater than one, delta modulation will be too slow to follow the signal, otherwise it will be too fast. Another artifact is that a constant signal can only be coded as a small oscillation between two values. This introduces some noise into silent parts of the signal. As human perception is most sensitive to silent signals, this is especially disturbing.

5. Provide an optimal code which reproduces the signal as well as in (1) but that uses less bits.

    This is the right place to be creative. No standard solution was expected here.

    (1) Try Huffman encoding of the absolute signal values.

    (2) Do delta modulation with as many bits as needed and again, use huffman coding.

    (3) Apply some kind of predictive coding, for example LPC

6. Another coding method is pulse-code modulation using logarithmic quantization (or unequal quantization). What are the advantages does that kind of coding as opposed to equal quantization?

    Signal values with great magnitudes fall into larger quantization intervals when logarithmic quantization is used. Small magnitudes of signal values benefit from a finer quantization. This can be interpreted as an attempt to move bits or precision into ranges where they are more useful.

    Think of a concert with an orchestra. At the maximum loudness flutes can hardly be perceived so that the resolution can be coarser. The saved bits can be moved into dynamic areas that are important when the flutes play soli.

7. Explain the difference between quantization and discretization.

    In this context discretization refers to the time domain and means that samples are taken within fixed time intervals. Quantization refers to the values of each sample. In contrast to discretization, quantization involves assigning a value to the border of the nearest quantization interval.

8. Human beings can perceive frequencies between 16 Hz and 20 kHz. What does that tell us about the temporal resolution of the ear?

    The highest frequency, in this case 20kHz, takes 0.05 ms for one period. Thus 0.05 ms is the smallest temporal resolution.

9. How large is the bitrate of an audio CD (44100 Hz sample rate and 16 bit quantization, stereo)?

    44 100 samples * 2 byte/sample * 2 channels = 176 400 bytes/second

Exercise 5.2 Linear predictive coding

The idea of LPC (linear predictive coding) is to linearly combine the last i-k samples in order to predict sample i.

1. Explain the idea of LPC (linear predictive coding).

    In LPC a sample at time n is approximated as a linear combination of the preceding k samples. The coefficients for that linear combination are chosen such that the difference between the original signal and the approximation (the residual) is as small as possible.

2. What coefficients are needed to predict the following pattern without a residual (with no error): -1, 1, -1, 1, -1, ... and so on

    $s_n = (-1) s_{n-1}$ Only the last sample is needed to predict the current one. The coefficient is (-1) in this case.

3. Find out the 4 coefficients with which the following pattern can be losslessly encoded. The first four numbers are given, the rest should be predicted: 32, 32, 32, 32, -32, 0, -48, 24, -20, 46, -21, 31.5, ...

   Solve an equation with four unknowns:

$$32a_4+32a_3+32a_2+32a_1=-32$$
$$32a_4+32a_3+32a_2-32a_1=0$$
$$32a_4+32a_3-32a_2+0a_1=-48$$
$$32a_4-32a_3+0a_2-48a_1=24$$

$$32a_4+32a_3+32a_2+32a_1=-32$$
$$0a_4+0a_3+0a_2-64a_1=32$$
$$0a_4+0a_3-64a_2-32a_1=-16$$
$$0a_4-64a_3-32a_2-80a_1=56$$

$$a_1=-0{,}5$$
$$a_2=0{,}5$$
$$a_3=-0{,}5$$
$$a_4=-0{,}5$$