Effects of quantization in digital filtering

Lab 9, SDP

Objective

Students should observe the effects of internal quantization errors on the output signal of a digital filter.

Theoretical notions

Exercises

1. Consider the system with the following difference equation:

$$y[n] = \frac{1}{2}y[n-1] + x[n]$$

   Compute the first 6 samples of the response to the input signal $x[n] = \left(\frac{1}{4}\right)^n$, in three different ways:

   a. Computations with infinite precision
   b. Computations with fixed-point 1S0I4F format, quantize by truncation
   c. Computations with fixed-point 1S0I4F format, quantize by rounding

2. In Octave, write a script file to study the quantization of an audio signal on a fixed point binary format with $N = 8$ bits.

   a. Load the audio signal `Sample.wav` and keep only 3 seconds of the audio signal. Name this vector `x`.
   b. Plot the signal, and figure out if a sign bit is needed or not in the binary format

c. Find the maximum absolute value of the signal, and figure out the number of bits required for the integer part. The remaining bits with be allocated to the fractionary part;
d. Use the provided function `cuant()` convert the signal `x` to the fixed-point data type, using all the three quantization methods;
e. For all the three quantization methods, visualize the quantized signal, the quantization error, and compute the total energy of the quantization errors. Which quantization method produces minimum errors?
f. Display the histograms of the quantization errors in all three cases.
g. Play the quantized signal. Can you hear the difference from the original signal? Expriment with smaller or larger number of bits until you start hearing distortions in the signal.

3. Consider an IIR filter implementation in Direct Form 2 provided in `filter_df2.m` (this file was written during Lab 4).

   a. Simulate the implementation of this filter using quantization on a fixed point binary format with $N = 8$ bits, by quantizing all values:

   - the input signal
   - the result of each multiplication and addition

   b. Apply the quantized system to the signal from exercise 2, not quantized. Also apply the original system, not quantized in `filter_df2.m`, not quantized. Display the output of the quantized system against the output of the original system.

   c. Compute the error between the two signals above, compute its energy, and display its histogram.

Final questions

1. TBD