Audio

CSC 790

WAKE FOREST UNIVERSITY

Department of Computer Science

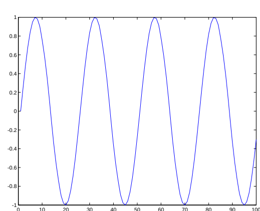
Fall 2009

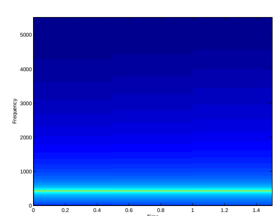
Audio

- Continuous wave that is a disturbance of mechanical energy
  - Properties include amplitude, frequency and phase

    What is the result of changing each?

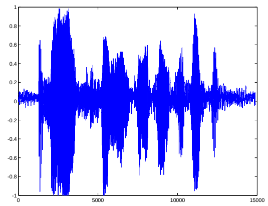
- Can represent the wave in the time or frequency domain

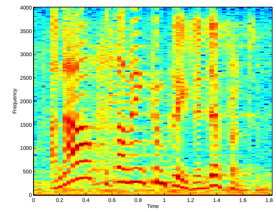




Complex Audio Waves

- Using Fourier analysis determine component waves (frequencies)



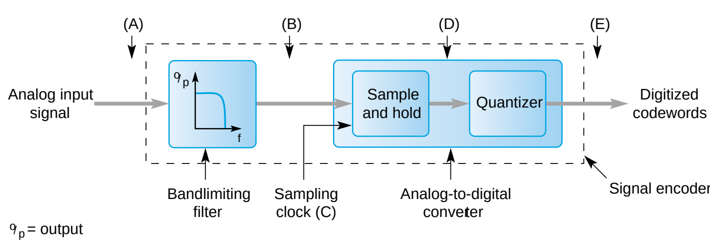


  - Complex waves formed from multiple simple sinusoidal waves
- Range between the smallest and largest frequency is bandwidth

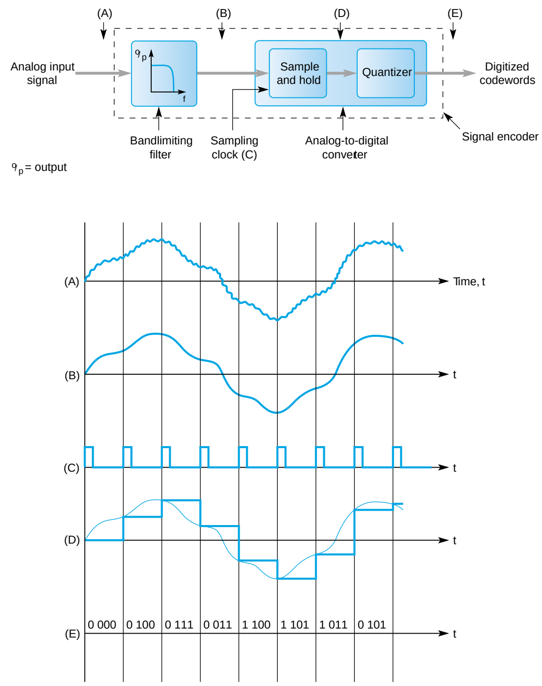
	Min. Frequency	Max Frequency
Speech	20 Hz	10 kHz
Telephony	200 Hz	3.4 kHz
CD Audio	15 Hz	20 kHz

Encoder Design

- Encoder converts the analog signal to a digital form



- Consists of the following components
  - Bandwidth limiting filter
  - Sample and hold circuit, samples analog signal at regular times
  - Quantizer converts sample amplitude into a codeword



Sampling Rate

- Analog wave sampled at regular intervals for digital representation
  - *What is the minimum sampling rate?*
- Nyquist (Fulp) sampling theorem states that

> *In order to obtain an accurate representation of a time-varying analog signal, its amplitude must be sampled at a minimum rate that is equal to or greater than twice the highest sinusoidal frequency component present in the signal*

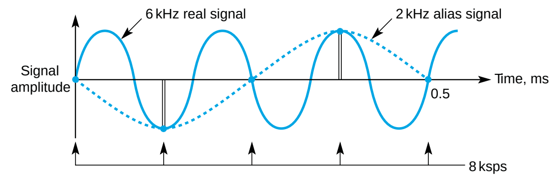
  - This rate is known as the **Nyquist rate** which is measured in frequency or samples per second
- Sampling below this rate causes frequency components to be generated that were not originally present

PCM Speech

- For telephony the bandwidth is *historically* 200 to 3400 Hz
  - Sampling rate is actually 8000 samples per second
    Nyquist only requires 6800 samples per second? What?
- 7 bits per sample (US and Japan), 8 bits per sample (Europe)
  - Most systems have converged to 8 bits per sample
    Performance implications?
- Digitization process is called **Pulse Code Modulation**
  - Standard for digital telephony is ITU-T Recommendation G.711

Alias

- Sampling below the Nyquist rate can cause **alias signals**
  - Consider the 6 kHz wave sampled at 8 ksps



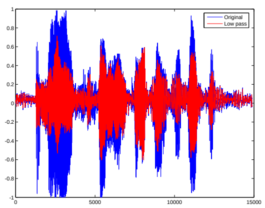
  - Result is a 2 kHz wave, which is an alias
- Filter can remove frequencies above a predetermined Nyquist rate
  - Called a antialiasing filter, avoids under-sampling

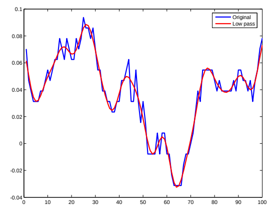
Filtering

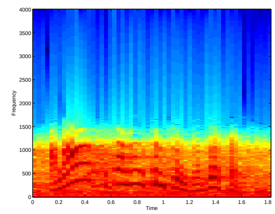
- Band pass filter allows only frequencies below/above threshold
  - Low and high pass filters are common

    Filter before the sampling? Why?

- Consider the following after 1000 Hz low pass Butterworth filter

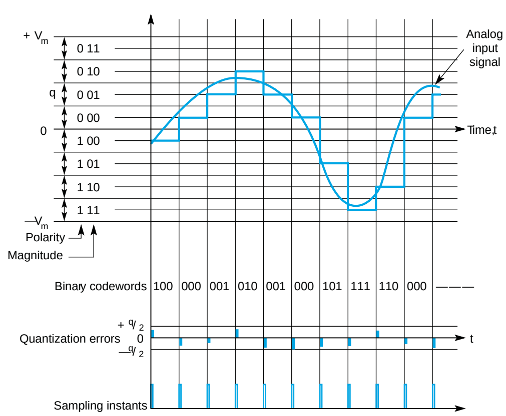






Quantization

- Quantization converts an analog sample into digital form



- Must match analog to a discrete codeword (finite number)

Quantization Intervals

- Assume v_{max} is the maximum amplitude (positive and negative) and n bits are used for the codeword
  - The **quantization interval**, q is

$$q = \frac{2v_{max}}{2^n}$$

  - Previous diagram used 3 bits which gives 8 levels (codeword)
- A signal sampled anywhere within the same interval is represented by the same codeword (interval level)
  - Codeword represents the nominal amplitude (interval center)
  - Actual signal level may differ by up to $\pm\frac{q}{2}$, an error

Quantization Error and Dynamic Range

- Difference between analog and digital value is **quantization error**
  - Normally varies randomly, is called **quantization error**
  - More levels reduces the quantization error
- Another factor in the number of levels is the **dynamic range**
  - Dynamic range, d, is ratio of peak to minimum amplitude (dB)

$$d = 20\log_{10}\frac{v_{max}}{v_{min}}$$

  - Must ensure the level of quantization noise relative to the smallest signal amplitude is acceptable

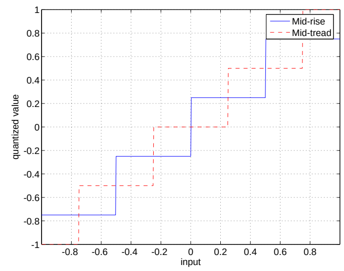
    An analog signal has a dynamic range of 40 dB, what is the magnitude of the quantization noise relative to the minimum signal amplitude if the quantization uses 6 or 10 bits?

Linear Quantization

- Consider a uniform quantizer, $q()$
  - Let the input x be a real value between -1 and 1
- *Mid-rise* quantizer that uses m bits of precision for each index

$$q(x) = \frac{\lfloor 2^{m-1}x \rfloor + 0.5}{2^{m-1}}$$

  - Where there are 2^m intervals between -1 and 1
  - Mid-rise quantizers do not include zero



Companding

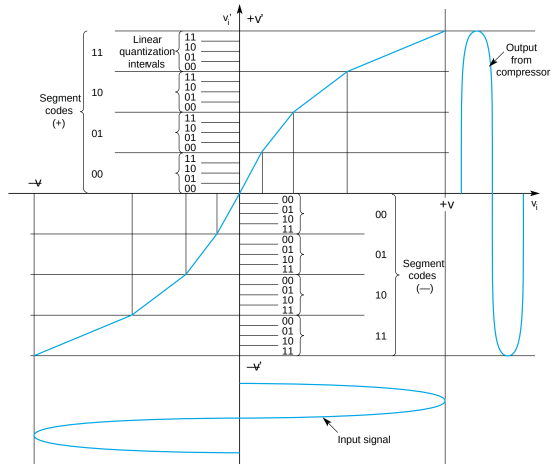
- If all quantization levels are same size then linear quantization
  - For any input magnitude same quantization noise is produced
  - Human hearing is more sensitive to noise at low amplitudes (noise is more noticeable at quite levels, *shhhhhhhhhh*)
- Non-linear quantization uses unequal length intervals
  - Smaller intervals at low amplitudes, large intervals for high
  - Reduces the effect of quantization when only 8 bits available
  - Called μ-law coding for the US and Japan, A-law for Europe

$$f(x) = \text{sign}(x)\frac{\ln(1 + \mu|x|)}{\ln(1 + |x|)}, \quad \mu = 255$$

  - Performance is comparable to 14 bit linear companding
    Is this compression?

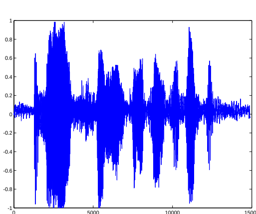
- The reverse, μ-law expansion is

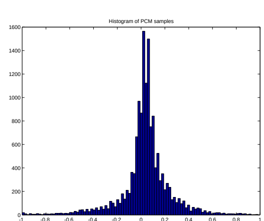
$$f^{-1}(y) = \text{sign}(y)\frac{1}{\mu}\left[(1+\mu)^{|y|} - 1\right]$$

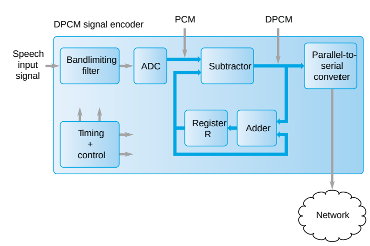


Differential Pulse Code Modulation

- Differential Pulse Code Modulation (DPCM) is derivative of PCM
  - Exploits small difference between consecutive samples







- Previous digitized sample is stored in a register
  - Difference computed between current and previous
  - Register updated with difference + previous sample
  - *Errors can be cumulative with this method*

    So what, how does differences help?

Lossless Predictive Coding

- Basic idea is to determine the differences and transmit
  - Predict next sample is equal to current + some error
- Assume a system produces a series of integer samples
  - Let f_i be i^{th} actual sample
  - Let f_i' be the i^{th} predicted sample
  - Let e_i be the i^{th} error
- Then if $f_i' = f_{i-1}$, then $e_i = f_i - f_i' = f_i - f_{i-1}$ (*just the difference between consecutive samples*)
- A very simple system, **too bad it does not work**

  What?

LPC Range

- Suppose the sample range is $0 \ldots 255$
  - As a result the difference could be $-255 \ldots 255$
  - The *dynamic range* has increased by a factor of 2
- A simple solution is to add two new *codewords*
  - Denote SU as shift up by 10 and SD as shift down by 10
  - Also need codewords for the values $0 \ldots 9$
- Any value outside $0 \ldots 9$ can be transmitted as a series of shifts followed by a value within the range $0 \ldots 9$
  - For example 19 would be $\{\text{SU}, 9\}$
- Given this strategy LPC works and is lossless

Better Predictions

- Previous LPC examples were based on the previous sample
  - For example, $f'_i = f_{i-1}$ and $e_i = f_i - f'_i$
- We can make predictions based on the last m samples
  - Predicted value is $f'_i = \sum_{k=1}^{m} a_k f_{i-k}$
  - Where a_i is a weight, $\sum_{k=1}^{m} a_k = 1$

    What is the value of e_i?
  - The weights could change dynamically
- Using multiple previous samples for the prediction *may provide a better predictor*

  Any ideas what the improvement is?

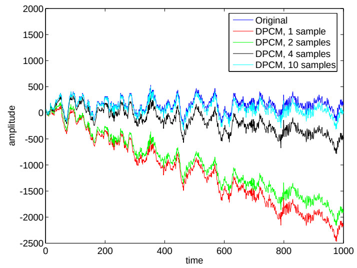
Differential Pulse Code Modulation

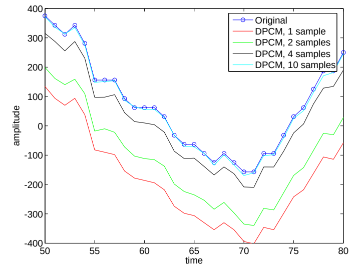
- DPCM is the same as predictive coding, just quantized
  - Therefore the difference transmitted is quantized

    So what is the problem we are trying to solve?
- The system operates as
  - Let $f'_i = \sum_{k=1}^{m} a_k \hat{f}_{i-k}$
  - Error is $e_i = f_i - f'_i$
  - Quantize the error and transmit, $q(e_i)$
  - Reconstructed value is $\hat{f}_i = f'_i + q(e_i)$
- **Achtung!** the error is based on the reconstructed quantized version of the signal, $\hat{f}_i$

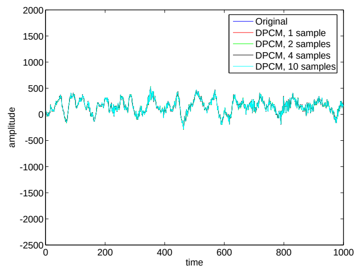
Original and Reconstructed

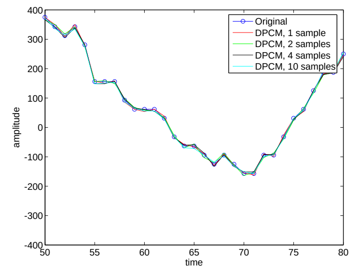
- Using the original samples





- Using the reconstructed samples





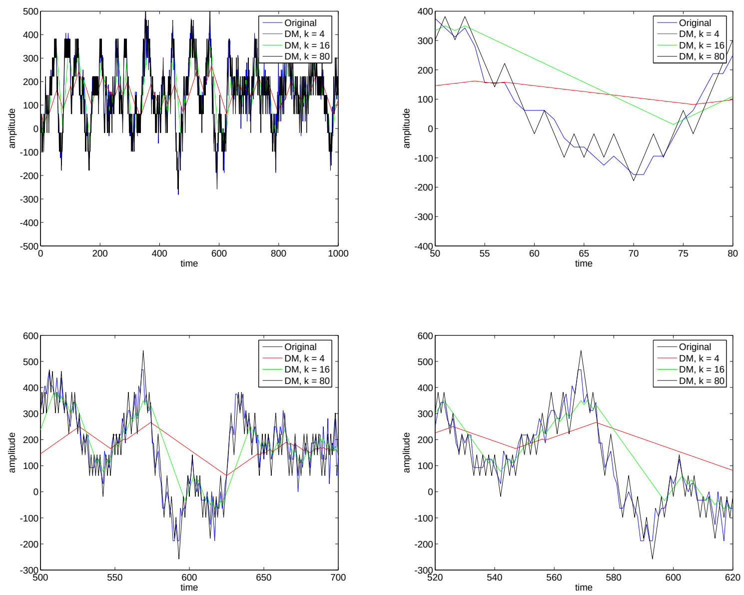
Delta Modulation

- A simple DPCM encoder/decoder
  - One example is *uniform-delta* DM
- Uniform-delta uses only one quantized error value
  - Let $f'_i = \hat{f}_{i-1}$
  - Error is $e_i = f_i - f'_i = f_n - \hat{f}_{i-1}$
  - Quantize the error and transmit, $q(e_i)$

$$q(e_i) = \begin{cases} +k & \text{if } e_n > 0\text{, where } k \text{ is a constant} \\ -k & \text{otherwise} \end{cases}$$

  - Reconstructed value is $\hat{f}_i = f'_i + q(e_i)$

Example Delta Modulation



Adaptive DM

- If the signal slope is too high, DM cannot keep up
  - *Examples of this behavior on the previous slide*
- A simple approach is the adjust k dynamically
  - The value of k depends on the current properties of the signal