CSc 461/561
Multimedia Systems
Lossy compression

Jianping Pan

Spring 2015

Compression

- Why compression?
  - there is (a lot of) redundancy!
- How to compress?
  - remove data *and* information redundancy
- Lossless compression
- Lossy compression
  - remove information redundancy *adequately*
  - information loss, but higher compression ratio!

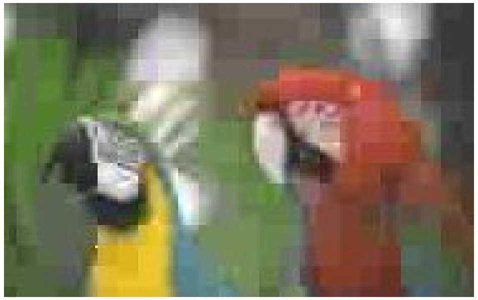
Lossy compression examples

- Why lossy compression is possible?
  - some information is more important than others for *human*
  - keep the important one

Original

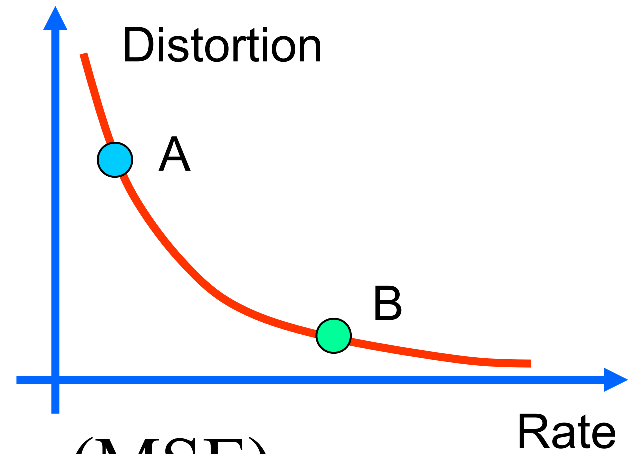
Compression Ratio: 7.7

Compression Ratio: 12.3



Compression Ratio: 33.9

Tradeoff: rate vs distortion



- Rate
  - # of bits per source symbol
- Distortion
  - one measure: mean square error (MSE)
  - x: original value; y: reconstructed value
  - $MSE = [(x_1-y_1)^2+(x_2-y_2)^2+\ldots+(x_N-y_N)^2]/N$
- Rate vs distortion
  - lower rate, higher distortion

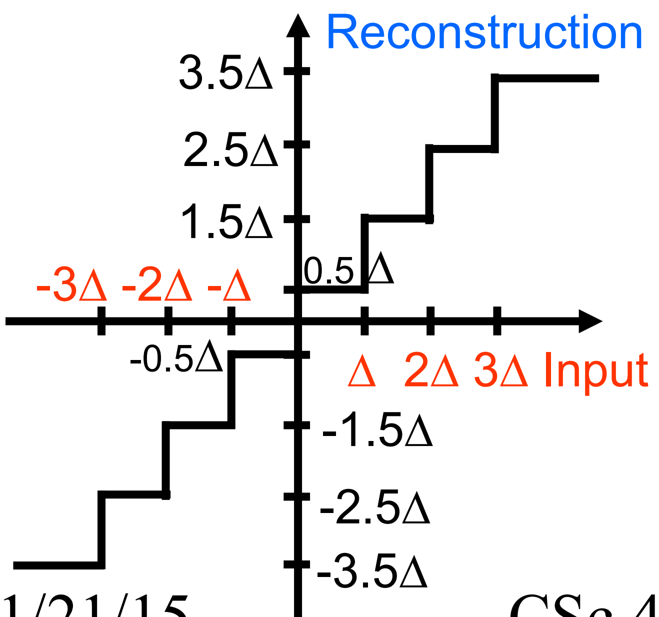
Quantization

- Quantization (recall audio A/D)
  - use a discrete value to represent a value range
  - information loss!
- The smaller range, the less distortion
  - granular distortion
- Quantization steps
  - uniform: all ranges have the same size
  - non-uniform: otherwise

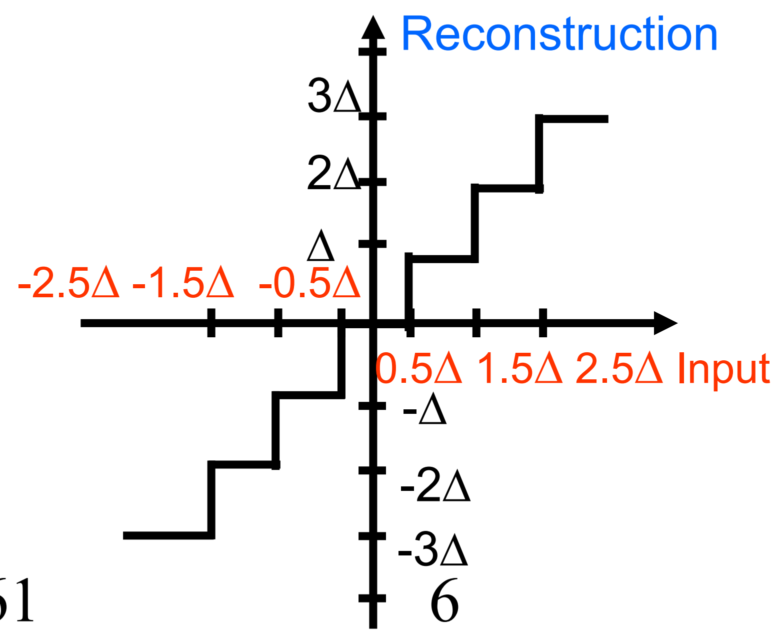
Uniform quantization

- Quantization step: uniform
- Two constructions: midrise, midtread

Uniform Midrise Quantizer



Uniform Midtread Quantizer



Signal-to-quantization-noise ratio

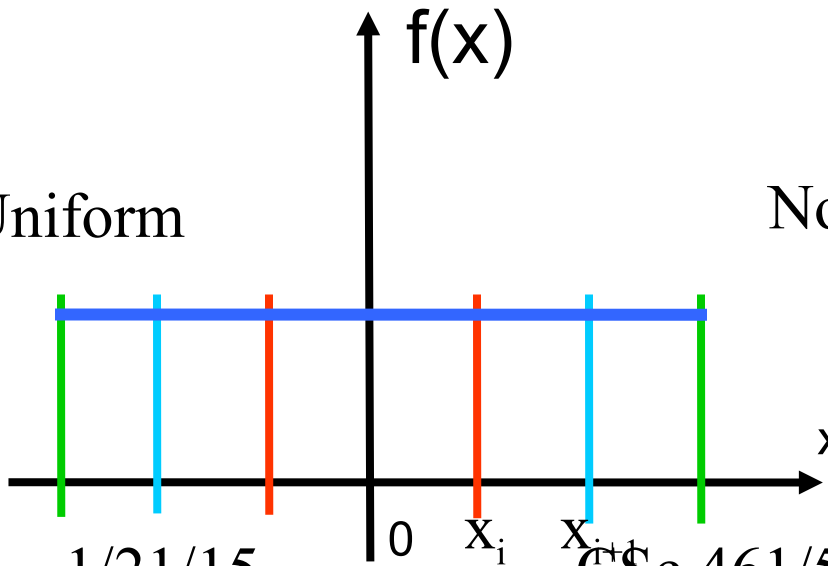
- Quantization
  - n bits; 2^n steps for $[-X_{max}, X_{max}]$
  - step size: delta = $2X_{max} / 2^n$
  - granular distortion: $\sigma^2_q = \int_{-\Delta/2}^{\Delta/2} (x-0)^2 \frac{1}{\Delta} dx = \frac{1}{12}\Delta^2$
- SQNR in dB
  - 10 $\log_{10}$ signal_energy / noise_energy

    $=10 \log_{10} [(2X_{max})^2/12]/[delta^2/12]=20n \log_{10}2$
- One more bit adds 6 dB to SQNR

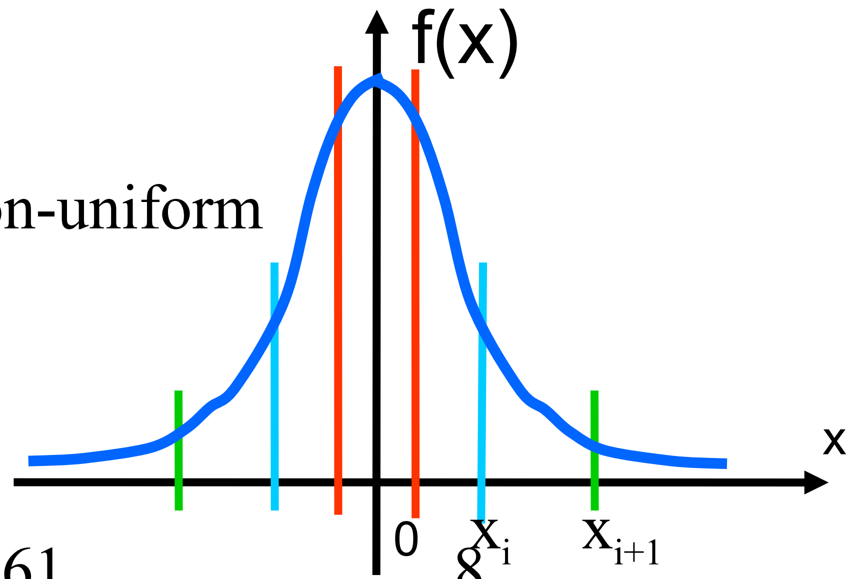
Non-uniform quantization

- Recall u-law or A-law voice compander
- How to choose quantization steps?
  - $\mathrm{Int}_{x_i}^{x_{i+1}} f(x)\, dx = 1/2^n$

Uniform

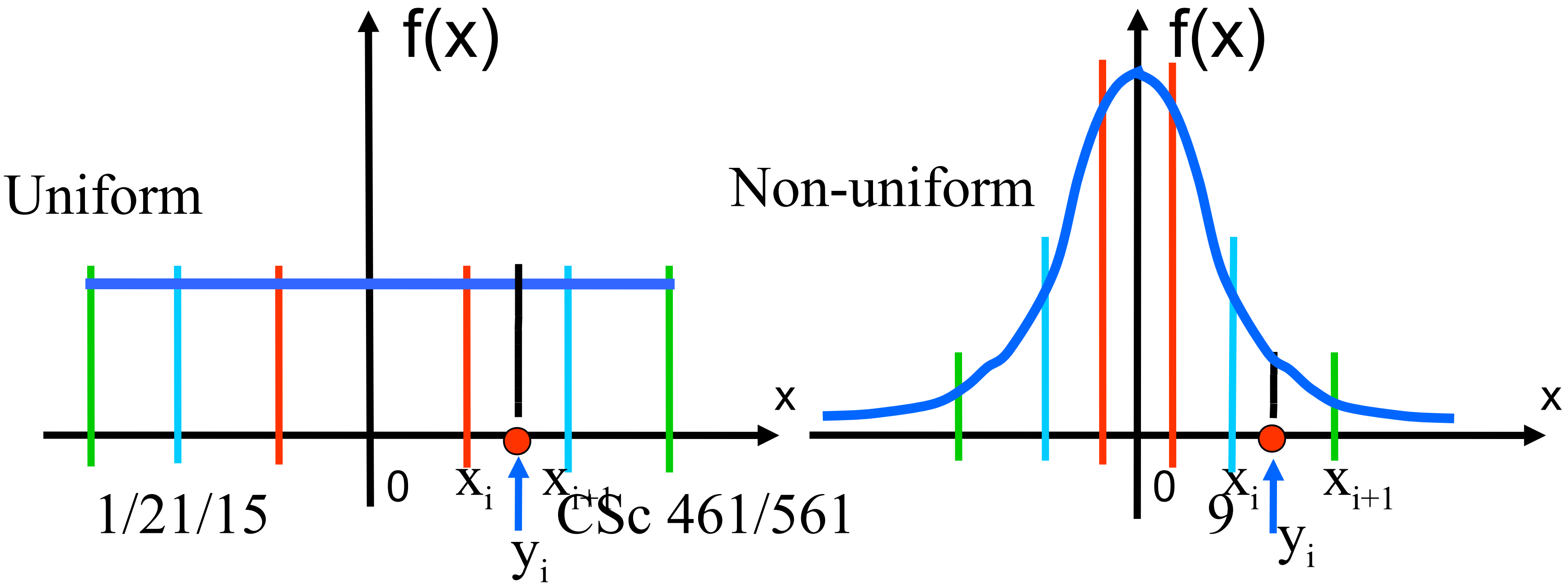


Non-uniform



Non-uniform quantization: more

- How to represent a range?
  - $\text{Int}_{x_i}^{y_i} f(x)\, dx = 1/2^{n+1}$
  - when uniform: $y_i=(x_i+x_{i+1})/2$



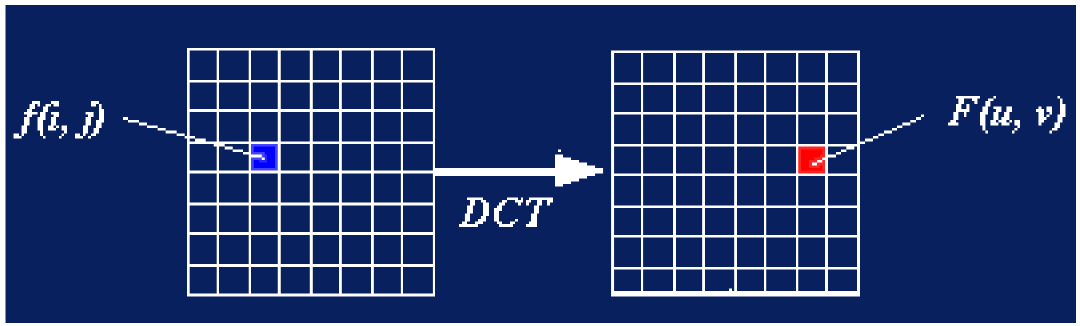
Transformation

- Transformation
  - represent information in anther space
    - identify and remove (hard-to-remove) correlation, i.e., redundancy, in the original space
    - information loss!
  - e.g., time/space => frequency (FFT)
- Inverse transformation
  - represent the info back in the original space

Discrete Cosine Transform

- Recall: a wave is of many *wave*s
- "Any signal can be expressed as a sum of multiple signals that are sine or cosine waveforms at various amplitudes and frequencies."
- Cosine transform: using cosine waveforms
- DCT: integer indexes
  - widely used in image compression (e.g., JPEG)

DCT: more



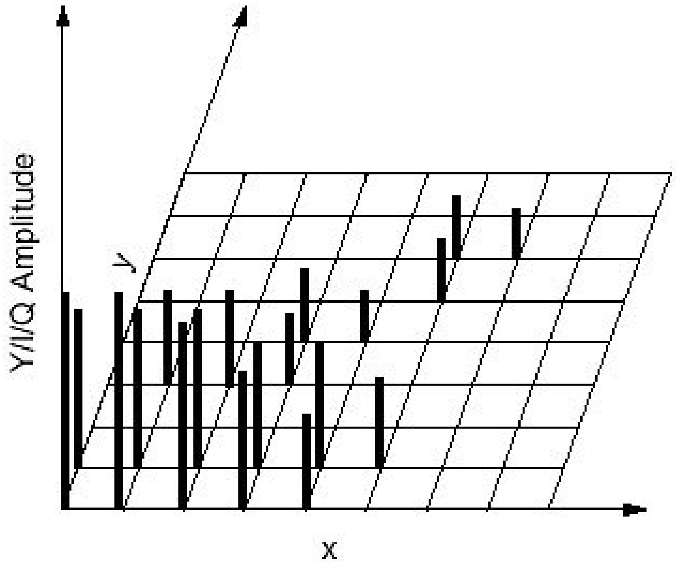
- 2-D DCT (8x8); C(x)=1/sqrt(2) when x=0

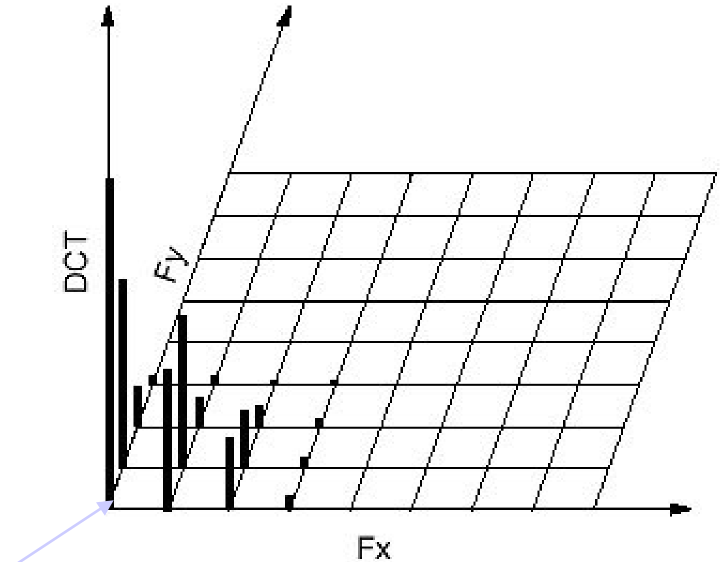
$$F(u,v) = \frac{1}{4} C(u)C(v) \left[\sum_{i=0}^{7} \sum_{j=0}^{7} f(i,j) \cos\frac{(2i+1)u\pi}{16} \cos\frac{(2j+1)v\pi}{16} \right]$$

- Inverse 2-D DCT (IDCT); C(x)=1

$$f(i,j) = \frac{1}{4} C(u)C(v) \left[\sum_{u=0}^{7} \sum_{v=0}^{7} F(u,v) \cos\frac{(2i+1)u\pi}{16} \cos\frac{(2j+1)v\pi}{16} \right]$$

DCT: examples





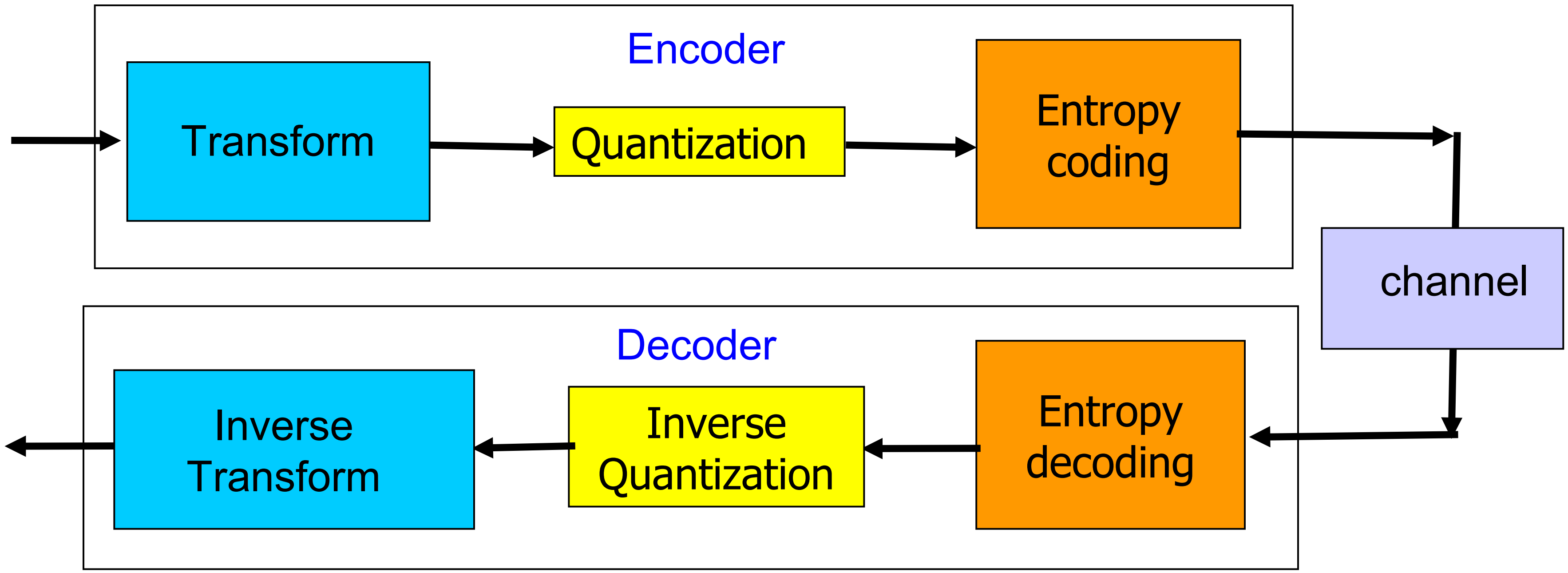
DC Component

Original values of an 8x8 block (in spatial domain)

Corresponding DCT coefficients (in frequency domain)

Lossy + lossless compression

- In a big picture...



This lecture

- Multimedia manipulation
  - lossy compression
    - rate vs distortion
    - quantization: uniform vs non-uniform
    - transformation: DCT
- Explore further
  - wavelet-based coding [Ref: Li&Drew 8.6.1]

Next lecture

- Multimedia manipulation
  - audio compression [Ref: Li&Drew Chap 13-14]
    - quick review on PCM, DPCM and ADPCM
    - examples: MPEG audio [14.1-2]