Compression and Image Formats

Compression

- Reduce amount of data used to represent an image/video
  - Bit rate and quality requirements
- Necessary to facilitate transmission and storage
- Required quality is application dependent
  - Medical vs. entertainment
- Data is "information"
  - Bits per second (bps), bits per pixel (bpp)

Compression is *Necessary*

- Example of a normal TV picture over a telephone network[1]:
  - Capacity of network: 56, 000 bps
  - Signal:
    - Image is 288 x 352
    - RGB color, 8 bits each channel
    - 30 frames per second
    - Data need: 288 x 352 x 8 x 3 x 30 = 72, 990, 720
      - 1289 times greater than capacity!
- Current networks are faster, but videos are larger

1. From: "Image and Video Compression" by Shi and Sun

Lossy vs. Lossless Compression

- Lossless Compression
  - No information is lost
  - Original image/video can be completely restored
- Lossy Compression
  - Some information is lost
  - Reduction in "quality" of image/video
  - Generally higher compression rates

Reason for hope

- Not all the data is required for a believable image
  - There is redundancy

Statistical Redundancy

- Interpixel Redundancy
  - Groups of pixels are not independent
  - Related in space and time
- Spatial Redundancy
  - For most images, consecutive rows (or columns) will be highly correlated
  - Same for rows slightly further a part, but this decreases as the separation gets larger

$\longrightarrow$ Can predict pixel intensity from neighbor

Statistical Redundancy

- Temporal Redundancy (Interframe Redundancy)
  - Pixels do not change much from frame to frame in a sequence
  - Observation from videophone-like signal:
    - Less than 10% of pixels change by more than 1% from frame to frame

  ⟶ Can predict pixel intensity from previous frame

Statistical Redundancy

- Coding Redundancy
  - Some values will occur more frequently in an image than others
    - e.g. Some colors are rare
  - Use less bits for the common colors and more for the uncommon ones
    - Reduces the total number of bits
    - e.g. Huffman codes
  - Better coding schemes can more efficiently represent the data
  - Compare index color and RGB for a three color image

Psychovisual Redundancy

- Image must read correctly to the human visual system (HVS)
  - Complicated and nonlinear
  - Tune to what people perceive
  - Some differences are much more important than others

Masking

- How sensitive the eye is to stimulus depends on the presence of another stimulus

Psychovisual Redundancy

- Luminance Masking
  - If background is bright, larger difference in intensity is needed to distinguish an object from the background
  - Suggests that noise will be more visible in a dark area than a light one
  - Nonuniform quantization can be more effective

Psychovisual Redundancy

- Texture Masking
  - Discrimination threshold increases with picture detail
  - i.e. Errors will be more noticeable in uniform/smooth areas of the image

Psychovisual Redundancy

- Frequency Masking
  - Human eye acts like a low-pass filter
  - Less sensitive to high frequency noise
- Temporal Masking
  - It takes time for the visual system to adjust after a rapid change in the image
    - Lower sensitivity during this time

Psychovisual Redundancy

- Color Masking
  - People are most sensitive to green, then red and last blue
    - Can allocate data (bits) based on this
  - Luminance (intensity) and chrominance (hue and saturation) can be a better representation than RGB
    - Can work in luminance space without distorting color (e.g. bring out shadow details with histogram equalization)
  - People are more sensitive to luminance than chrominance
    - Use more compression for chrominance than luminance

Image Formats

Common Image Formats

- JPEG (jpg)
- PNG
- GIF
- TIFF
- Bitmap

JPEG (.jpg)

- Became an international standard in 1992
- Different modes
  - Lossy
    - Uses Discrete Cosine Transform (DCT)-based coding
      - Beyond the scope of this course
      - **Image is divided into 8x8 blocks**, DCT run on each block
      - Coefficients of DCT are stored with image
  - Lossless
    - Based on predictive coding (also beyond scope)
    - Three neighboring pixels are used to predict current pixel
    - Huffman or arithmetic coding is used to store prediction difference

JPEG

- Different modes
  - Hierarchical
    - Image is spatially down sampled into a pyramid of progressively lower resolution images
      - e.g. an 4x4 can be sampled to a 2x2 can be sampled to 1 pixel
    - Can transmit progressively, lower resolution first and then add higher resolution detail
    - Can use either a lossy or lossless coding scheme

JPEG 2000 (.jp2, .jpx)

- Uses wavelet transform instead of DCT
- Provides excellent coding efficiency and good quality
- Wavelet transform also used in MPEG-4

More on (lossy) JPEG

- Can control amount of compression
  - Tradeoff between quality and image size
- **Every time you save an image, it will be recompressed and there will be a loss of quality**
  - Do not repeatedly edit and save lossy jpeg files
- 8-bit gray scale images
- 24-bit color images (8 bit each for RGB)

TIFF

- Lossless (in practice)
- Large file sizes
- 1 to 48 bit color

GIF

- Old format, developed by Compuserve
- 8-bit indexed color
  - Table of 256 colors (8 bits)
  - Each pixel stores a table index
  - All the colors that can be displayed in the image
    - **Image can only contain 256 colors**
    - 24 bit color gives 16 million colors
    - Huge reduction in color space
  - Bad for photographs, may work for images with limited colors
  - Lossless for those 256 colors

PNG

- Designed as open-source successor to GIF
- 8, 24 or 48-bit color
- Lossless
  - Image files can be large
  - No loss of quality
  - Good format for working with images
- Compression based on patterns in image
- Does well with large, uniformly colored areas

Read More

- <u>Image and Video Compression for Multimedia Engineering: Fundamentals, Algorithms, and Standards</u>, Yun Q. Shi and Huifang Sun, CRC Press, 2008

Analog vs. Digital Transmission

Transmission of Signals

- Goal of analog and digital transmission is different
- Analog signals:
  - Goal is to exactly reconstruct the original signal
  - Errors lead to degradation
  - Diagram on board

Transmission of Signals

- Digital signals:
  - Goal is to reconstruct the pattern of 0's and 1's encoded in signals
  - Signal may be noisy, but no loss in quality as long as the 0's and one's can be detected
  - Checksums to verify transmission
  - Diagram on board
  - With digital, it is possible to make an *exact* copy
    - Not true with analog