COMSATS Institute of
Information Technology

ECI750 Multimedia Data Compression

Lecture 1

Introduction

Dr. Shadan Khattak

Department of Electrical Engineering

COMSATS Institute of Information Technology - Abbottabad

Introduction

- Without data compression:
  - It would not be practical to put images, let alone audio and video, on websites.
  - Cellular phones would not be able to provide communication with increasing clarity.
  - The advent of digital TV would not be possible.
- Data compression is ubiquitous!
- You are using data compression when you are:
  - Making a call on your cell phone.
  - Surfing the internet.
  - Listening to music on your MP3 player.
  - Watching a DVD.

Introduction

- What is Data?
  - Characters in a text file
  - Numbers that are samples of speech or image waveforms
  - Sequences of numbers that are generated by other processes
- What is Data Compression?
  - Art or science of representing information in a compact form.
  - Compact representations are made by identifying and using structures that exist in the data.

The Need for Compression

- Expansion of video data
  - Evolution of video resolutions
  - Higher bit-depths and frame rates (10-bit, 12-bit, 16-bit; 48fps, 120 fps)
  - Multiple number of views (e.g. for 3D video, Free View-point Video)
- A number of features are often combined
  - E.g. High resolution + 2 views + 10 bits
- Representation formats

Why Compress? Evolution of Video Resolutions

- In terms of pixel count,
  8K ≈
  16x FHD
  36x HD
  100x SD
- NHK plans to broadcast 2020 Tokyo Olympics in 8K UHD resolution (branded *"Super Hi-Vision"*).

Why Compress? Evolution of (not just) Video Resolutions

Format	Resolution (width X height)	Frame Rate (fps)	Data Rates (MBps)	Storage Capacity/Hour (GB)
SDTV (NTSC, 4:2:2, 8-bit)	720 X 480	~30	31	112
HDTV (1080p, 4:2:2, 8-bit)	1920 X 1080	24	149	537
Digital Cinema 2K (4:2:4, 10-bit) YUV	2048 X 1080	24	199	716
Digital Cinema 4K (4:4:4, 12-bit) YUV	4096 X 2160	48	1,910	6,880
Digital Cinema 8K (4:4:4, 16 bit)	7680 X 4320	120	23,890	86,000

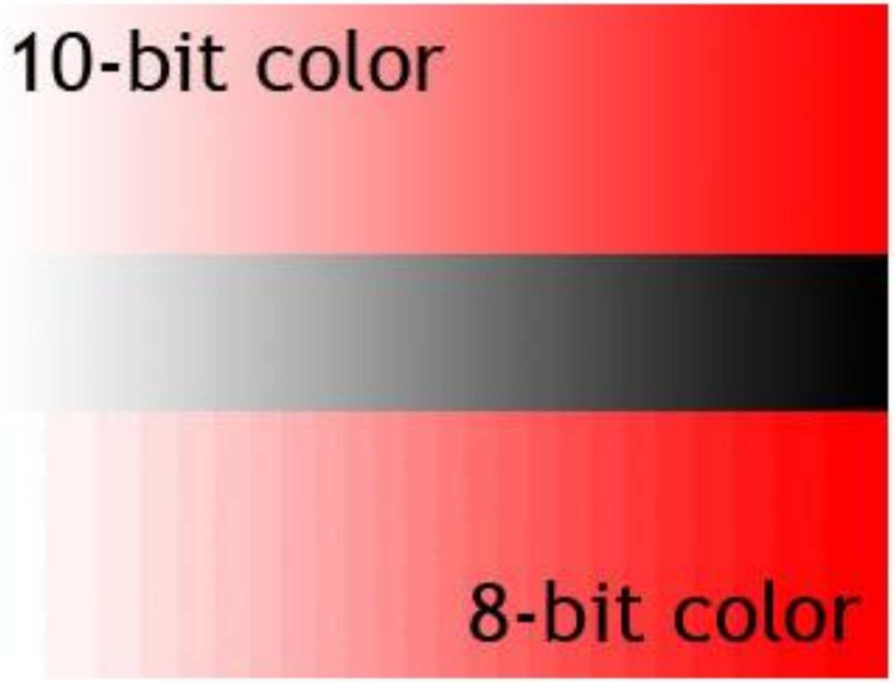
Higher bit-depths

8K Ultra-HD may use more than 100X capacity of HD!

8K X 4K based upon "Super Hi-Vision" Video Parameters for Next Generation Television, SMPTE Motion Imaging Journal, May/June 2012, P. 63-68

Why Compress? Higher bit-depths

- 8-bits:
  - 8-bit per channel
  - 256 shades of colour per channel
  - Suffers from *‘banding’ i.e., clear divisions between different shades of colour*
- 10-bits:
  - 10-bit per channel
  - 1024 shades of colour *(~4 times as many levels of colour for each channel)*
  - Smoother transition between different shades of colour
  - *Roughly 20% increase in file size*



Dr. Shadan Khattak
Department of Electrical Engineering
CIIT - Abbottabad

Why Compress? Higher frame rates

- Frame rates for movie content are increasing from the historical 24 frames per second (fps) to 48 fps (e.g., The Hobbit: The Desolation of Smaug, 2012) and 60 fps and may eventually be as high as 300 fps.
- In terms of pixel count,

48 fps = 2x 24 fps

120 fps = 5x 24 fps

"….. It blew me away, it's seriously the only way to view the movie. The difference between 24 and 48 FPS is goddamn amazing. Not only did everything look more realistic, it even made the CGI look more realistic ….. like you could just reach through the screen and touch Middle-earth. It seriously looked like a play, like the movie screen was not there and the film was actually happening IRL before your eyes ….. the difference is night and day ….. the panoramic shots and the CGI battle scenes are truly magnificent and real looking. It really improved the experience as a whole."

Why Compress? Multiple Views

- Advanced video applications require more than just one view:
  - *3D Video:* provides depth impression
  - *Free View-point Video:* provides interactive viewpoint selection *(the look-around effect)*
  - *3D+FVV:* provides the impression of depth as well as viewpoint selection

640 × 480 - research.microsoft.com

Why Compress? Multiple Views

- Both 3DV and FVV require the scene to be captured using more than one camera.
- 2 views *(=2x the data compared to 2D video)* are sufficient for 3D video

- The flexibility of FVV depends on the number of available views *(i.e. more views = more flexibility = more data)*

Dr. Shadan Khattak
Department of Electrical Engineering
CIIT - Abbottabad

The Need for Compression: In a nutshell

- Consider the following:
  - A 16000x8000 pixel resolution, 24 bits/pixel, 300 fps raw video content
    - *~115 GB/s (or 414 TB/hour)*
  - In case of FVV, if 4 cameras were used
    - *~1.66 PB/hour*
  - The bandwidth and capacity requirements to work with future rich media formats are staggering.

More Examples

- **Example 1:** consider the following video:
  - Length: 1 second
  - Resolution: 1920 x 1080 pixels (HD)
  - Frame rate: 30 frames per second
  - No. of channels: 3 (RGB)
  - No. of bits per channel: 8
  - Without compression, it represents:
    - 1920 x 1080 pixels x 3 colour channels x 8 bits/colour sample x 30 frames
    - = 1,492,992,000 bits = 186,624,000 bytes = 182,250 KB = 177.9 MB
  - Consider the number of seconds in a movie and you will be convinced that we need compression.

More Examples

- **Example 2:** consider the following audio:
  - Length: 1 second
  - Sampling rate: 44,100 samples per second (CD quality)
  - Bits per sample: 16
  - No. of channels: 2 (stereo)
  - Without compression, it represents:
    - 2 channels x 1 second x 44100 samples/sec x 16 bits per sample
      - = 1,411,200 bits = 176,400 bytes = 172.26 KB
  - Downloading music from internet (using slow connections) at these rates would take a long time.

More Examples

- **Example 3:**
- Assume that a PSTN modem can operate at a maximum bit rate of 56,600 bits per second.
- Assume:
  - Resolution: 288x352 (CIF)
  - No. of channels: 3 (RGB)
  - No. of bits per channel: 8
  - Frame rate: 30 frames per second (fps)
  - Without compression, it represents:
    - 288x352x3x8 = 72,990,720 bps
    - This is $\frac{72990720}{56600} \cong 1289$ more data than the modem can handle.

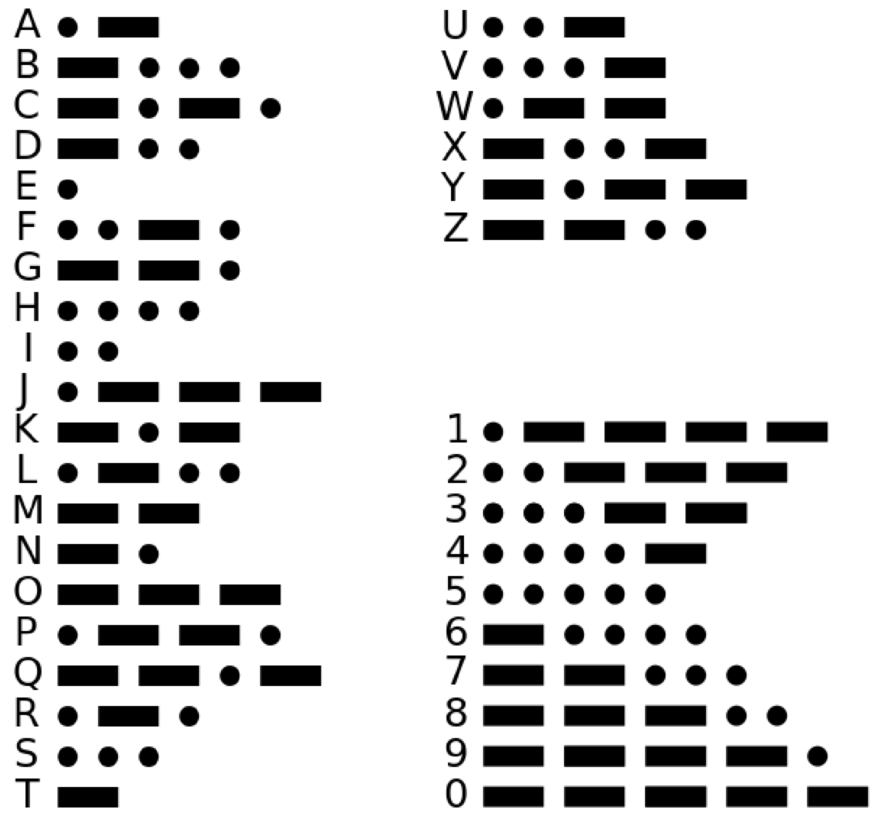
Dr. Shadan Khattak
Department of Electrical Engineering
CIIT - Abbottabad

Introduction

- So what if the data is huge?
  - Limited transmission capacity
  - Limited storage capacity
- So why don’t we just increase the transmission and storage capacities?
  - They are indeed increasing but the need for mass storage and transmission increases at least twice as fast as storage and transmission capacities improve.

Introduction

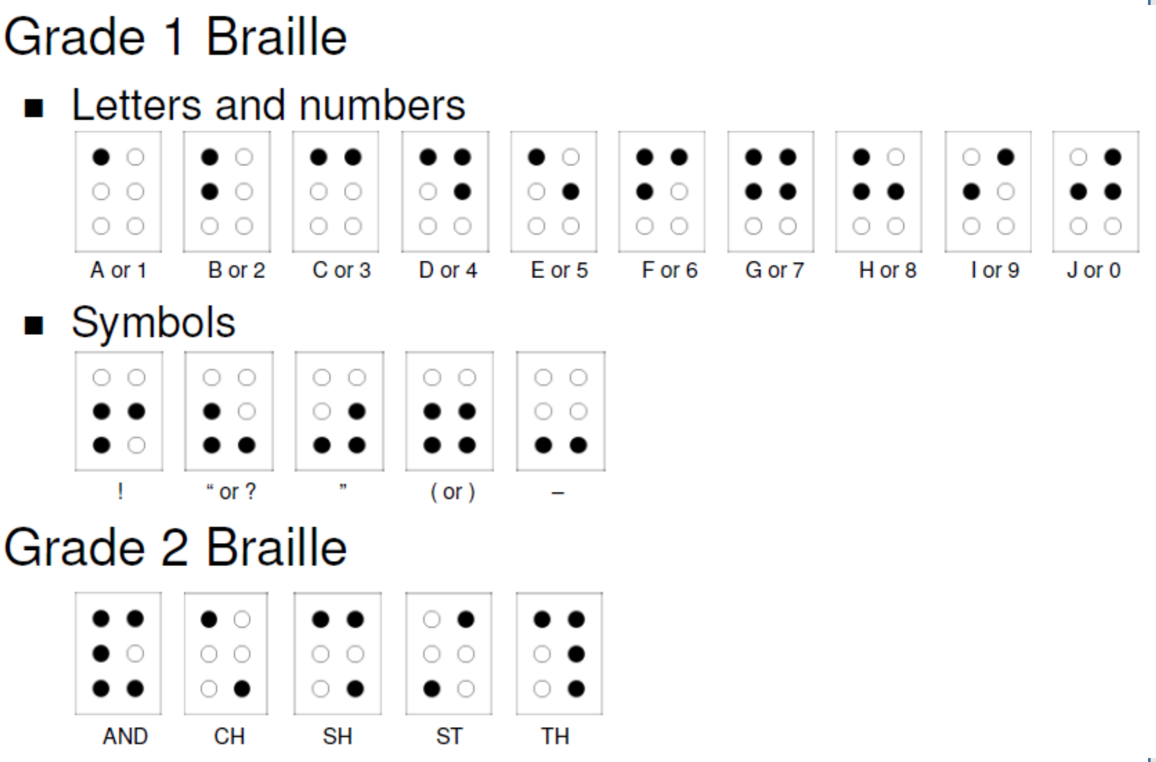
- Data Compression – Early Examples
- Example 1: Frequency Based Coding (Morse Code)
  - Developed by Samuel Morse in the 19th century
  - Letters sent by telegraph are encoded with dots and dashes
  - Certain letters occur more often (e.g., a, e, I, t etc.)
  - To reduce the average time required to send a message, Morse assigned shorter sequences to these letters and longer sequences to letters that occur less frequently.
  - This idea was latter adopted in Huffman Coding (which we are going to study later)



Dr. Shadan Khattak
Department of Electrical Engineering
CIIT - Abbottabad

Introduction

- Data Compression – Early Examples (2)
- Example 2: Context Based Coding (Braille Code)
  - Developed by Louise Braille (at the age of 15) in the 19th century.
  - Based on frequency of occurrence of words *(unlike Morse code which was based on frequency of occurrence of letters)*
  - 2x3 arrays of dots are used to represent text
  - Dots are either raised or flat.

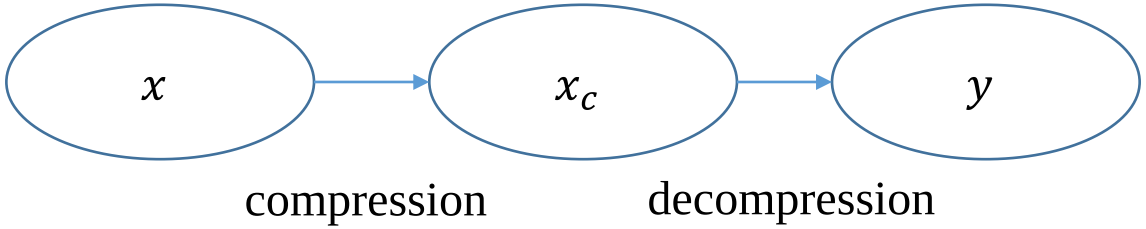


Introduction

- Both of our examples (Morse code, Braille code) are based on statistical structures.
- Apart from using structures, compression can also be obtained using the characteristics of the intended user.
  - Perceptual limitations of humans
    - Audible frequency range (20 Hz – 20,000 Hz)
    - Human visual system (less sensitive to high frequencies)
- For example, if a part of our data cannot be perceived by humans, do we really need to keep it?
  - No, we should get rid of it and hence reduce (compress) the data.

Compression Techniques

- General Compression Algorithms
  - x - original data, x_c- compressed representation, y – reconstructed data



  - Lossless compression: when y is equal to x
  - Lossy compression: when y is not equal to x

Compression Techniques

- Lossless compression:
  - The original data can be recovered exactly from the compressed data.
  - Useful for applications that cannot tolerate any difference between the original and reconstructed data
  - For example, text compression, computer files holding bank records, radiological images, data from the satellites etc.

Compression Techniques

- Lossy compression:
  - The original data cannot be recovered exactly from the compressed data.
  - Much higher compression ratios than is possible with lossless compression
  - Useful when exact reconstruction is not a problem
  - For example: speech, video etc.

Compression Techniques

- How do we measure the performance of a compression scheme?
- Possible aspects of the algorithm that can be measured:
  - Complexity
  - Memory requirement
  - Amount of compression
  - How closely the reconstruction resembles the original

Compression Techniques

- Measuring the amount of compression:
- Normally, *“compression ratio”* is used in the context of storage
- While, *“compression rate”* is used in the context of transmission

$$Compression\ Ratio = \frac{number\ of\ bits\ required\ to\ represent\ the\ data\ before\ compression}{number\ of\ bits\ required\ to\ represent\ the\ data\ after\ compression}$$

- Example:
  - Original image = 65,536 bytes
  - Compressed image = 16,384 bytes
  - $Compression\ Ratio = \frac{65,536}{16,384} = \frac{4}{1} = 4:1$
  - Compression ratio can also be presented in percentages
  - E.g., here ¾ (or 75%) of the data has been discarded so compression ratio = 75%

Compression Techniques

- Measuring the amount of compression:
- Compression Rate or Compressed Bit Rate represents the average number of bits required to represent a single sample.
- Example:
    - Original image = 8 bits/pixel
    - Compressed image = 2 bits/pixel
    - Compression Rate or Compressed Bit Rate (or bits per pixel) = 2 bits/pixel
    - Other units include: bits/sample, bits/character, bits/second.

Compression Techniques

- How similar (or different) are the reconstructed and the original (uncompressed) data in lossy compression?

$$\boxed{Distortion = difference\ between\ the\ original\ and\ the\ reconstructed\ data}$$

  - If the distortion is small, we say the “quality” or “fidelity” is high.
  - Common evaluation metrics include
    - Mean Squared Error (MSE)
    - Peak Signal-to-Noise Ratio (PSNR)