DATA COMPRESSION

Why Compress?

- To reduce the volume of data to be transmitted (text, fax, images)
- To reduce the bandwidth required for transmission and to reduce storage requirements (speech, audio, video)

Data compression implies sending or storing a smaller number of bits. Although many methods are used for this purpose, in general these methods can be divided into two broad categories: lossless and lossy methods.

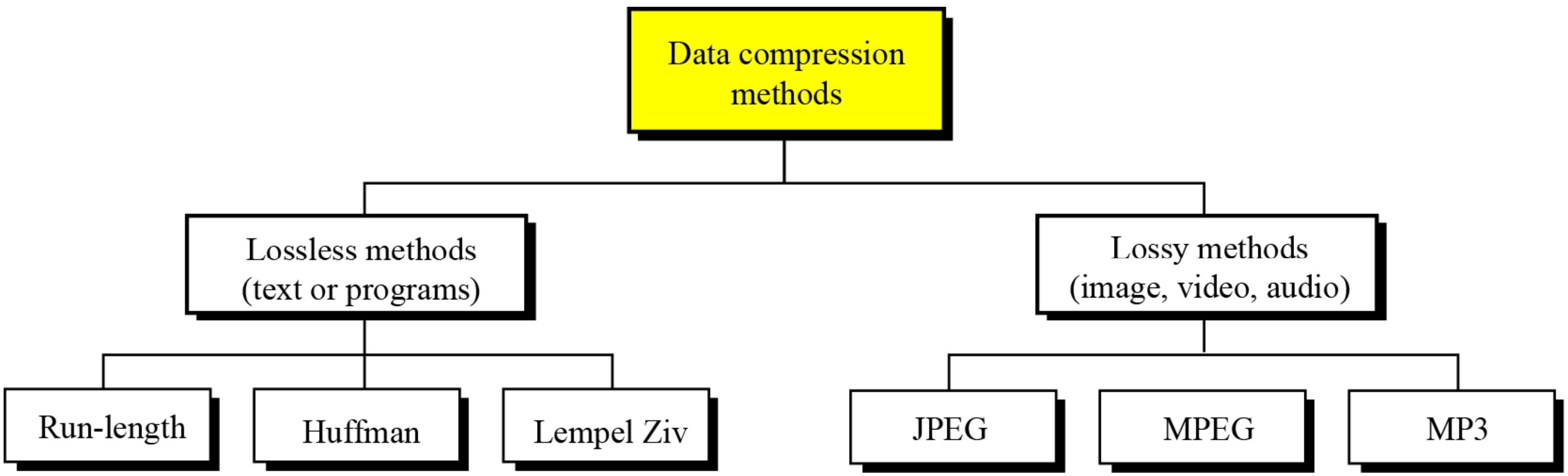

Figure1 Data compression methods

Compression

- How is compression possible?
  - Redundancy in digital audio, image, and video data
  - Properties of human perception
- Digital audio is a series of sample values; image is a rectangular array of pixel values; video is a sequence of images played out at a certain rate
- Neighboring sample values are correlated

Redundancy

- Adjacent audio samples are similar (predictive encoding); samples corresponding to silence (silence removal)
- In digital image, neighboring samples on a scanning line are normally similar (spatial redundancy)
- In digital video, in addition to spatial redundancy, neighboring images in a video sequence may be similar (temporal redundancy)

Human Perception Factors

- Compressed version of digital audio, image, video need not represent the original information exactly
- Perception sensitivities are different for different signal patterns
- Human eye is less sensitive to the higher spatial frequency components than the lower frequencies (transform coding)

Classification

- Lossless compression
  - lossless compression for legal and medical documents, computer programs
  - exploit only data redundancy
- Lossy compression
  - digital audio, image, video where some errors or loss can be tolerated
  - exploit both data redundancy and human perception properties
- Constant bit rate versus variable bit rate coding

Entropy

- Amount of information *I* in a symbol of occurring probability *p* : $I = \log_2(1/p)$
- Symbols that occur rarely convey a large amount of information
- Average information per symbol is called entropy H

  $H = p_i x \log_2(1/p_i)$ bits per codeword
- Average number of bits per codeword = $N_i p_i$ where N_i is the number of bits for the symbol generated by the encoding algorithm

Huffman Coding

- Assigns fewer bits to symbols that appear more often and more bits to the symbols that appear less often
- Efficient when occurrence probabilities vary widely
- Huffman codebook from the set of symbols and their occurring probabilities
- Two properties:
  - generate compact codes
  - prefix property

Run-length Coding

- Repeated occurrence of the same character is called a run
- Number of repetition is called the length of the run
- Run of any length is represented by three characters
  - eeeeeee7tnnnnnnnn
  - @e7t@n8

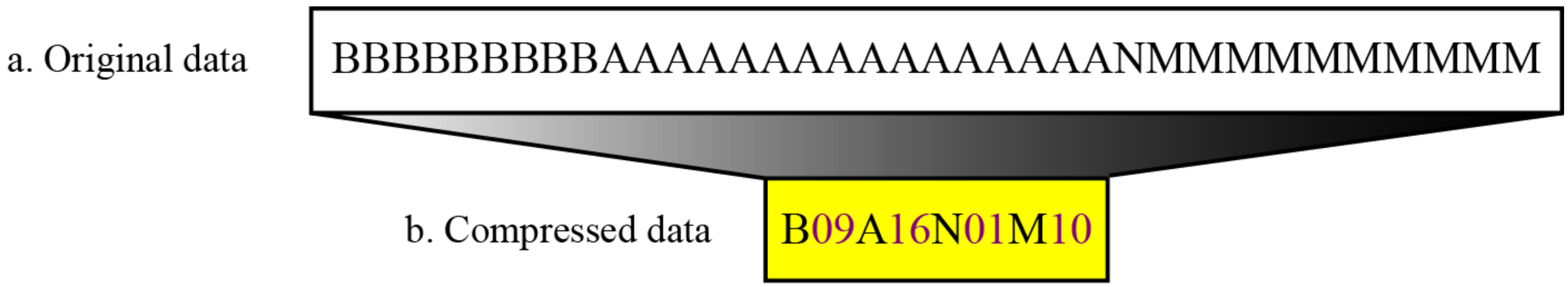

Figure 15.2 Run-length encoding example

Lempel-Ziv-Welch (LZW) Coding

- Works by building a dictionary of phrases from the input stream
- A token or an index is used to identify each distinct phrase
- Number of entries in the dictionary determines the number of bits required for the index -- a dictionary with 25,000 words requires 15 bits to encode the index

Arithmetic Coding

- String of characters with occurrence probabilities make up a message
- A complete message may be fragmented into multiple smaller strings
- A codeword corresponding to each string is found separately

Data Compression

Coding Requirements

- **Let us consider the general requirements imposed on most multimedia systems:**
- **Storage — multimedia elements require much more storage space than simple text. For example, a full screen true colour image is 640 x 480 X 3 = 921600 bytes**

Data Compression

- The size of one second of uncompressed CD quality stereo audio is 44.1kHz X 2 X 2 = 176400 bytes
- The size of one second of uncompressed PAL video is

  384X288X 3 X 25 frames = 8294400 bytes

Data Compression

- **Throughput** — continuous media require very large throughput. For example, an uncompressed CD quality stereo audio stream needs 176400 bytes/sec. A raw digitized PAL TV signal needs(13.5MHz + 6.75MHz + 6.75MHz) X 8bits

= 216 $X10^6$bits/sec

= 27 $X10^6$Bytes/sec

Data Compression

- **Interaction** — to support fast interaction, the end-to-end delay should be small. A ‘face-to-face’ application, such as video conferencing, requires the delay to be less than 50ms. Furthermore, multimedia elements have to be accessed randomly.
- **Conclusion**:
- Multimedia elements are **very large**.
- We need to reduce the data size using **compression**.

Data Compression

Kinds of coding methods

- **Lossless — the compression process does not reduce the amount of information.**

- The original can be reconstructed exactly

- **Lossy — the compression process reduces the amount of information.**

-**Only an approximation of the original can be reconstructed.**

Data Compression

Categories of Compression Techniques

- **Entropy coding is lossless**
- **Source coding and hybrid coding are lossy.**

Data Compression

Coding techniques

- **Vector Quantization — a data stream is divided into blocks of n bytes (where n > 1). A predefined table contains a set of patterns is used to code the data blocks.**
- **LZW —a general compression algorithm capable of working on almost any type of data. It builds a data dictionary of data occurring in an uncompressed data stream. Patterns of data are identified and are matched to entries in the dictionary. When a match is found the code of the entry is output.**

Data Compression

- **Since the code is shorter than the data pattern, compression is achieved. The popular zip application used this method to compress files.**
- **Differential coding — (also know as prediction or relative coding) The most known coding of this kind is DPCM (Differential Pulse Code Modulation). This method encodes the difference between the consecutive samples instead of the sample values. For example,**

Data Compression

- **PCM 215 218 210 212 208 . . .**
- **DPCM 215 3 -8 2 -4 . . .**
- **DM (Delta Modulation) is a modification of DPCM. The difference is coded with a single bit.**

Data Compression

Huffman coding

- **The principle of Huffman coding is to assign shorter code for symbol that has higher probability of occurring in the data stream.**
- **The length of the Huffman code is optimal.**

A Huffman code tree is created using the following procedures

- Two characters with the lowest probabilities are combined to form a binary tree.
- The two entries in the probability table is replaced by a new entry whose value is the sum of the probabilities of the two characters.
- Repeat the two steps above
- Assign 0 to be left branches and 1 to the right branches of the binary tree.
- The Huffman code of each character can be read from the tree starting from the root.

Data Compression

JPEG

- **JPEG (stands for Joint Photographic Experts Group) is a joint ISO and CCITT (*Comité Consultatif International Téléphonique et Télégraphique),* working group for developing standards for compressing still images**
- **The JPEG image compression standard became an international standard in 1992**
- **JPEG can be applied to colour or grayscale images**

LOSSY COMPRESSION METHODS

Our eyes and ears cannot distinguish subtle changes. In such cases, we can use a lossy data compression method. These methods are cheaper—they take less time and space when it comes to sending millions of bits per second for images and video. Several methods have been developed using lossy compression techniques. JPEG (Joint Photographic Experts Group) encoding is used to compress pictures and graphics, MPEG (Moving Picture Experts Group) encoding is used to compress video, and MP3 (MPEG audio layer 3) for audio compression.

Image compression – JPEG encoding

an image can be represented by a two-dimensional array (table) of picture elements (pixels).
A grayscale picture of 307,200 pixels is represented by 2,457,600 bits, and a color picture is represented by 7,372,800 bits.

In JPEG, a grayscale picture is divided into blocks of 8×8 pixel blocks to decrease the number of calculations because, as we will see shortly, the number of mathematical operations for each picture is the square of the number of units.

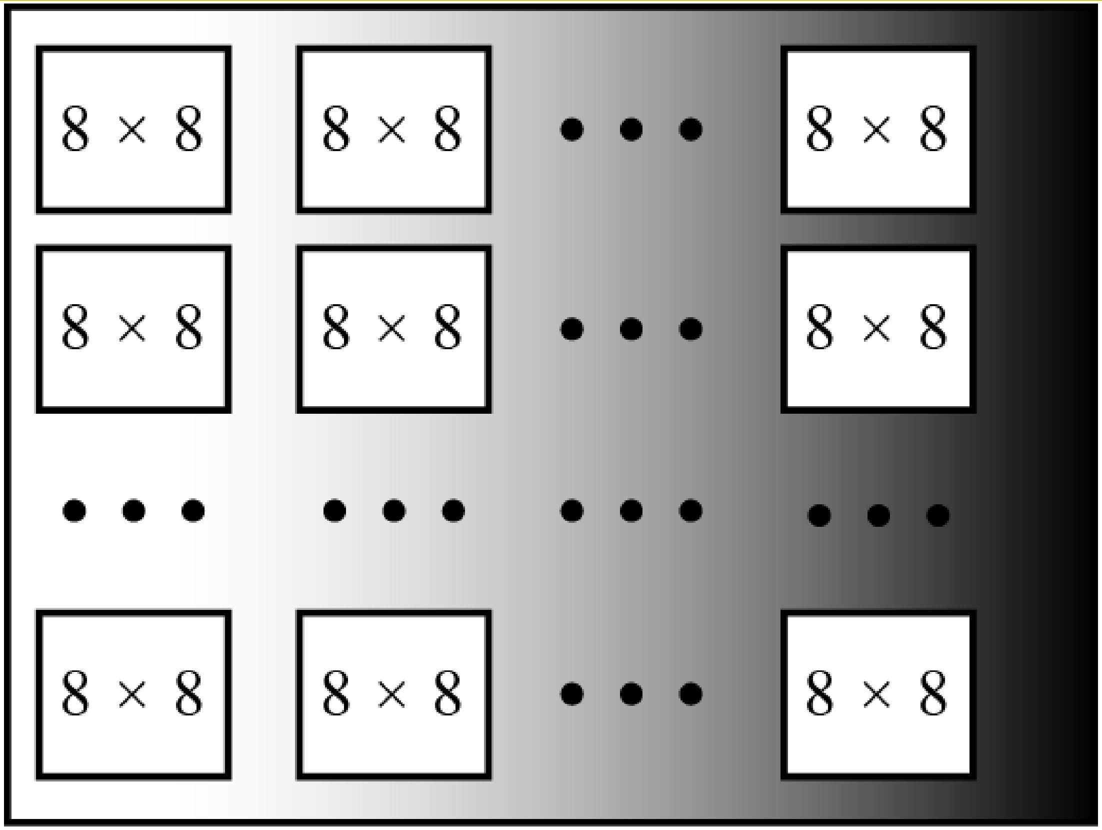

Figure 2 JPEG grayscale example, 640 × 480 pixels

The whole idea of JPEG is to change the picture into a linear (vector) set of numbers that reveals the redundancies. The redundancies (lack of changes) can then be removed using one of the lossless compression methods we studied previously. A simplified version of the process is shown in Figure 3.

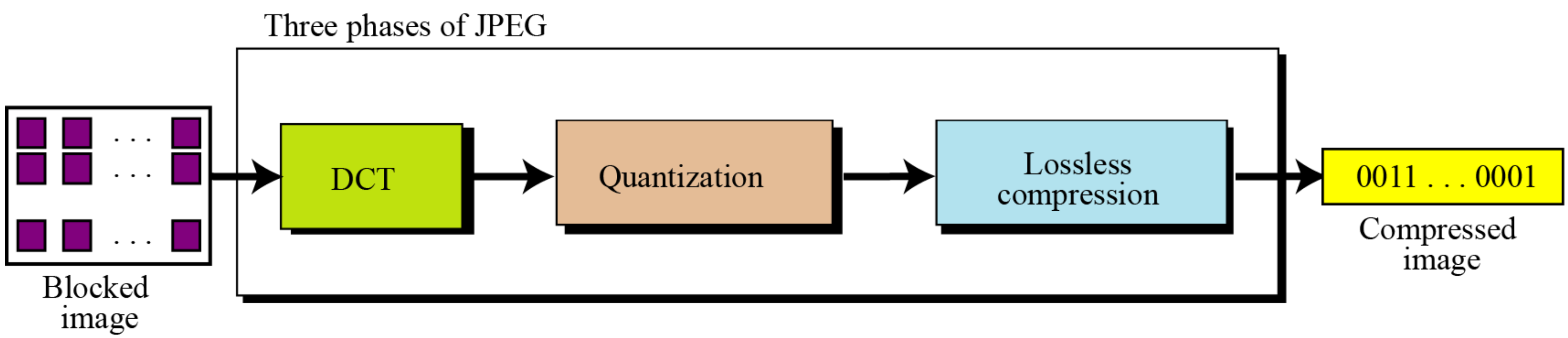

Figure 3 The JPEG compression process

Video compression – MPEG encoding

The Moving Picture Experts Group (MPEG) method is used to compress video. In principle, a motion picture is a rapid sequence of a set of frames in which each frame is a picture. In other words, a frame is a spatial combination of pixels, and a video is a temporal combination of frames that are sent one after another. Compressing video, then, means spatially compressing each frame and temporally compressing a set of frames.

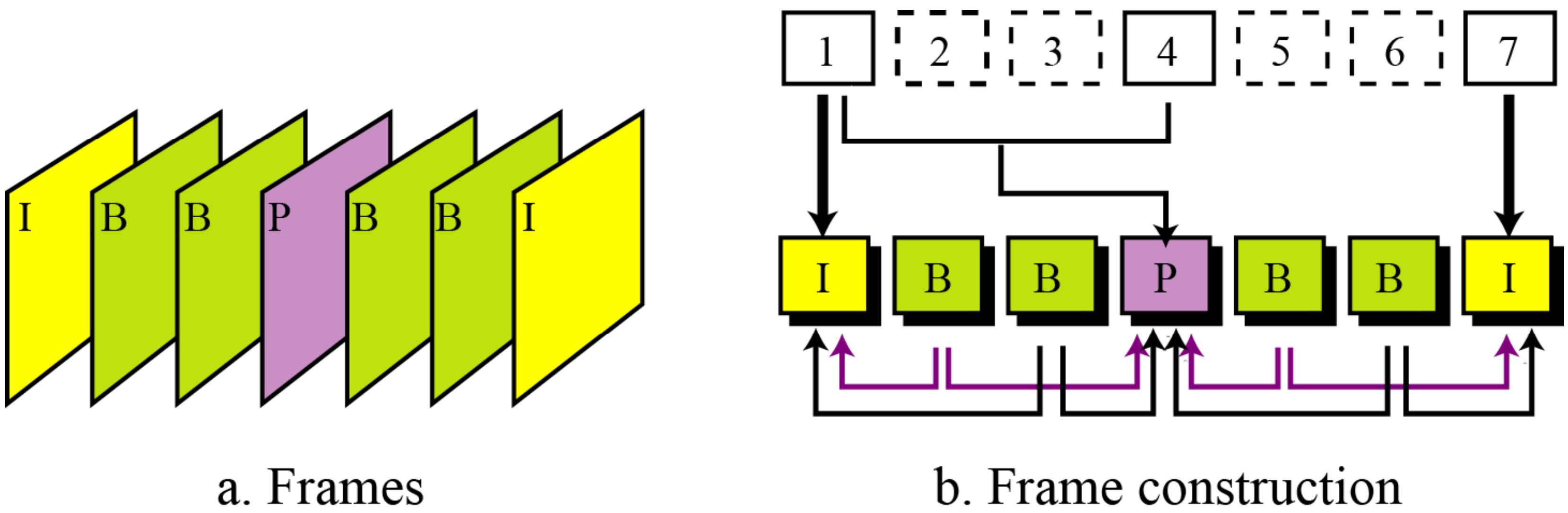

Figure 4 MPEG frames

Audio compression

Audio compression can be used for speech or music. For speech we need to compress a 64 kHz digitized signal, while for music we need to compress a 1.411 MHz signal. Two categories of techniques are used for audio compression: predictive encoding and perceptual encoding.

Data Compression

- **By changing appropriate parameters, the user can select**

– the quality of the reproduced image
– compression processing time
– the size of the compressed image

Data Compression

- The JPEG standard have three levels of definition as follows:
- Baseline system — must reasonably decompress colour images, maintain a high compression ratio, and handle from 4bits/pixel to 16bits/pixel.
- Extended system — covers the various encoding aspects such as variable length encoding, progressive encoding, and hierarchical mode of encoding.

Data Compression

- Special lossless function— ensures that at the resolution at which the image is compressed, decompression results in no loss of any detail the was in the original image.

Data Compression

- **JPEG — Preparation**
- **A source image consists of at least one and at most 255 planes.**
- **Each plane Ci may have different number of pixels in the horizontal (Xi) and vertical (Yi) dimension.**
- **The resolution of the individual plane may be different.**
- **Each pixel is represented by a number of bits p where $2 \leq p \leq 12$.**

Data Compression

- **The meaning of the value in these planes is not specified in the standard.**
- **The image is divided into 8 X 8 blocks.**

Data Compression

MPEG

- **MPEG (stands for Moving Picture Experts Group) is also a joint ISO and CCITT working group for developing standards for compressing still images**
- **The MPEG video compression standard became an international standard in 1993**
- **MPEG uses technology defined in other standards, such as JPEG and H.261**

Data Compression

- It defines a basic data rate of 1.2Mbits/sec
- It is suitable for symmetric as well asymmetric compression It follows the reference scheme that consists of four stages of processing:

1. Preparation
2. Processing
3. Quantization
4. Entropy Encoding

Data Compression

- In the preparation stage, unlike JPEG, MPEG defines the format of the images
- Each image consists of three components — YUV
- The luminance component has twice as many samples in the horizontal and vertical axes as the other two components (known as colour sub-sampling
- The resolution of the luminance component should not exceed 768 pixels

Data Compression

- for each component, a pixel is coded with eight bits

Data Compression

How MPEG encode the video stream

- **In order to achieve higher compression ratio, MPEG uses the fact the image on consecutive frames differ relative small. It uses a temporal prediction technique to encode the frame so that the storage requirement is greatly reduced.**
- **Common MPEG data stream consists of four kinds of frames:**

Data Compression

- **I-frame (Intra-frame) — it is a self contained frame, and it is coded without reference to any other frames.**
- **P-frame (Predictive-coded frame) — It is coded using the predictive technique with reference to the previous I-frame and/or previous P-frame.**
- **B-frame (Bi-directionally predictive coded frame) — It requires information of the previous and following I- and P-frames for encoding and decoding.**

Data Compression

- D-frame (DC-coded frame) Only the lowest frequency component of image is encoded. It is used in fast forward or fast rewind.

Data Compression

MPEG-2

- **MPEG-2 is a newer video encoding standard which builds on MPEG-1**
- **It supports higher video quality and higher data rate (up to 80 Mbits/sec)**
- **It supports several resolutions**:

pixels/line	line/frame	frames/sec
352	288	30
720	576	30
1920	1152	60

Data Compression

Summary

- **Compression methods — lossless vs. lossy**
- **Entropy coding — run-length encoding, Huffman encoding**
- **Source coding — prediction (DPCM, DM), transformation (DCT)**
- **hybrid coding — JPEG, MPEG**