Chapter 5

Data Compression

Why Compression

- Multimedia data such as image, audio, video requires enormous amount of storage.
- Point to point communication of such data requires high bandwidth
- Real time communication such as video conferencing becomes impossible
- To make feasible communication and cost effective solutions, multimedia data is compressed

Why Compression

Uncompressed audio signal of telephone quality:

- sampled at 8 kHz, quantized with 8 bits per sample ➔ 64 kbits to store one second of playback

Uncompressed stereo audio signal of CD quality:

- sampled at 44.1 kHz, quantized with 16 bits per sample ➔ **705.6 kbits** to store one second of playback

PAL video format:

- image resolution of 640x480 pixels, 24 bits = 3 bytes/pixel to encode luminance and chrominance components,
- 25 frames/second ➔ **175.78 Mbits** to store one second of video

Why Compression

- Cd quality of stereo with 16 bits depth 44.1khz sampling rate, size of 10 in is
- Sampling rate(hz) * depth(bits)*no. Of channel* time(sec)
- =100.93MB

Compression techniques

Jpeg for still images

H.261(px64) for video conferencing, additionally for speech and music in ISDN and mobile communications

MPEG for video and audio

DVI for still images as well as for continuous media

Compression techniques

- Processing of uncompressed video requires storage in the range of giga bytes.
- They also requires data rate as high as 140Mbits/sec
- Compression techniques reduces the high data transfer rate requirements
- Quality of compressed and decompressed data must be as good as possible.

Multimedia Applications

- Can be run in two different modes
- Dialogue mode
- Retrieval mode
- Dialogue mode is interaction between human via multimedia information.
- Retrieval mode retrieves multimedia information from multimedia database.

Dialogue Mode

- End to end delay introduced by compression and decompression should not exceed 150ms
- 50ms delay should be achieved in face to face dialogue applications
- Other delay is introduced by communication protocol processing and data transfer from input devices

Retrieval mode

- Fast forward and backward data retrieval with simultaneous display
- Fast search for information in multimedia databases
- Random access to single images and audio frames with an access time less than 0.5 second
- Decompression without a link to other data units for random access and editing

Both mode requirements

Supporting scalable video in different systems

- format independant of frame size and video frame rate

Support of various audio and video data rates

- data rates can be adjusted

Synchronistaion of audio and video data

Economical solution: coding should be realized in

- software: cheap, but low quality
- hardware: at first expensive, but high quality

Compatiblity

- Exchange multimedia data using communictaion networks
- Programs available on CD can be read on different systems

Compression classification

Entropy encoding

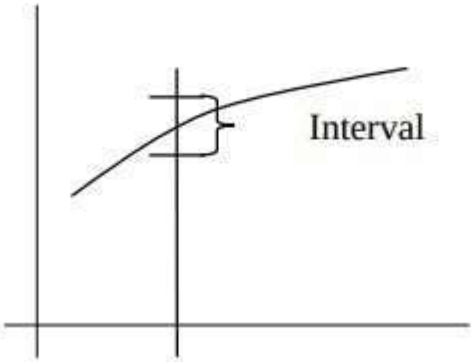
- Lossless coding, decrompession process regenerates the data completely
- Used regardless of the media´s specific characteristics
- Data stream is considered to be a simple digital sequence
- Examples: Run-length encoding, Huffman encoding, Arithmetic encoding

Source encoding

- Lossy coding
- Semantics of the data are considered
- Degree of compression depends on the data contents
- Examples: (Differantial) Pulse Code Modulation (DPCM) as content prediction technique, Discrete Cosine Transformation (DCT) as transformation technique

Hybrid encoding

- Used by most multimedia systems
- Combination of entropy and source encoding
- Examples: JPEG, MPEG, H.261, DVI



General Compression

Uncompressed Picture

↓

Picture Preparation

Analog-to-digital conversion
Generates an appropriate digital representation
Image division into blocks of *8x8* pixels (or more general *NxN*)
Fixed number of bits/pixel

↓

Picture Processing

First step of the compression process:
- Transformation from time to frequency domain using DCT
- Motion vector in the case of motion video compression

↓ Feedback Loop ↑

Quantization

Mapping of real numbers into rational numbers (approximation)
Result: reduction of precision

↓

Entropy Encoding

Compress a sequential digital data stream without loss

↓

Compressed Picture

ASCII encoding

- Suppose, we want to send message “go go gophers”
- The message are send in the form of sequence of bits according to their ASCII value.
-

ASCII coding

char	ASCII	binary
g	103	1100111
o	111	1101111
p	112	1110000
h	104	1101000
e	101	1100101
r	114	1110010
s	115	1110011
space	32	1000000

ASCII encoding

- ASCII encoded of “go go gophers” is
- 1100111 1101111 11000000 1100111 1101111 1000000 1100111 1101111 1110000 1101000 1100101 1110010 1110011
- ASCII uses 7 bits per character. Thus, 13 character requires 91 bits
- ASCII encoding may be very big if the message is very large.

Fixed length encoding

- Since message does not contain all the characters, we can encode with our own code.
- We can encode 8 characters with 3 bits.

3-bit coding

char	code	binary
g	0	000
o	1	001
p	2	010
h	3	011
e	4	100
r	5	101
s	6	110
space	7	111

Fixed length encoding

- 3 bit encoding requires 39 bits for message and
- This shows that fixed length code is not efficient if frequency is not high enough

- Content Dependent compression
- Lossless compression
- Replace sequence of the same bytes with their number of occurrences
- A special flag indicates the number of occurences
- Special flag may be !
- Uncompressed data: ABCCCCCCCCDEFGGG
- Compressed data: ABC!8DEFGGG

Huffman Encoding

- Variable length encoding
- Determines the optimal code using minimum number of bits
- Frequently occurring characters are coded with shorter bits than seldom occurring characters
- Binary Tree is constructed to determine huffman code

Huffman Encoding

- Leaf nodes represent characters
- Non-leaf nodes is frequency of characters
- First characters are sorted in the increasing order of their frequency
- Then two smallest frequencies are added to form commutative frequency node.
- This is done repeatedly till no frequencies are left to be added

Huffman Encoding

- Left edge is given 0 code and right edge gets 1.
- Code for each character is determined by combining the code for arc from root to leaf node.
- Lossless encoding
- In the similar way, we can encode image and video

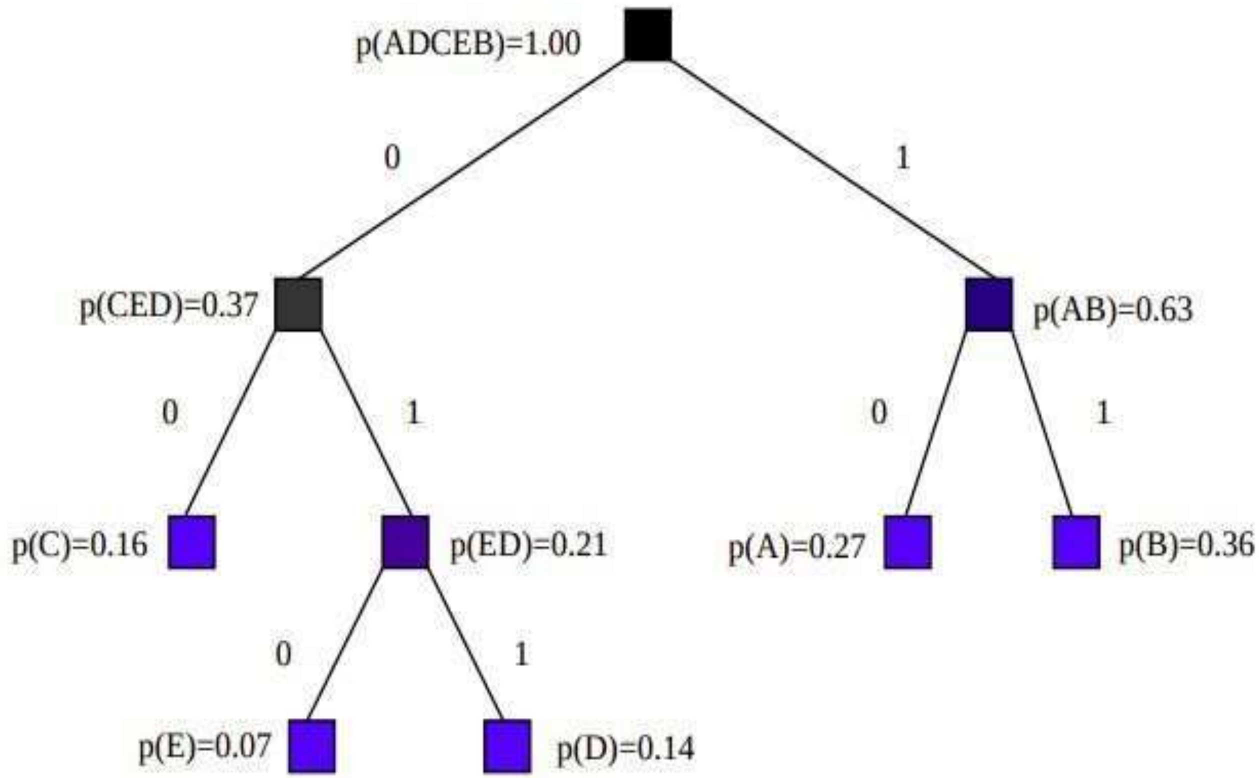
Huffman Encoding

Suppose that the characters A, B, C, D and E have the probabilities
$p(A) = 0.27, p(B) = 0.36, p(C) = 0.16, p(D) = 0.14, p(E) = 0.07$

Further assumption: Charcaters are independent of each other

Resulting Code:

x	$w(x)$
A	10
B	11
C	00
D	011
E	010



Huffman Algorithm

ASCII value of A ?

65 (1000001)

Suppose we want to write 3 A then it will have 21 bit. But if we somehow able to denote character bit as per its frequency let say for A (01) then we just require 6 bit to denote AAA.

Huffman coding is a lossless data compression algorithm. In this algorithm, a variable-length code is assigned to input different characters. The code length is related to how frequently characters are used. Most frequent characters have the smallest codes and longer codes for least frequent characters.

There are mainly two parts. First one to create a Huffman tree, and another one to traverse the tree to find codes.

For an example, consider some strings “YYYZXXYYX”, the frequency of character Y is larger than X and the character Z has the least frequency.

Prefix Codes, means the codes (bit sequences) are assigned in such a way that the code assigned to one character is not the prefix of code assigned to any other character. This is how Huffman Coding makes sure that there is no ambiguity when decoding the generated bitstream.

Let us understand prefix codes with a counter example. Let there be four characters a, b, c and d, and their corresponding variable length codes be 00, 01, 0 and 1. This coding leads to ambiguity because code assigned to c is the prefix of codes assigned to a and b. If the compressed bit stream is 0001, the de-compressed output may be “cccd” or “ccb” or “acd” or “ab”.

The Huffman coding or Huffman algorithm takes as input the frequencies (which are the probabilities of occurrences or repetitions known in advance) of symbols in a string of text and produces as output a prefix code that encodes the string using the fewest possible bits, among all possible binary prefix codes for these symbols.

Given symbols and their frequencies, the goal is to construct a rooted binary tree where the symbols are the labels of the leaves.

Input is an array of unique characters along with their frequency of occurrences and output is Huffman Tree.

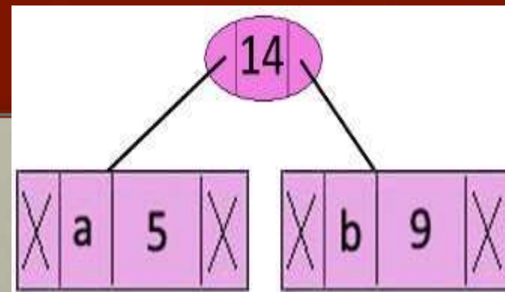
1. Create a leaf node for each unique character and build a min heap of all leaf nodes (Min Heap is used as a priority queue. The value of frequency field is used to compare two nodes in min heap. Initially, the least frequent character is at root)

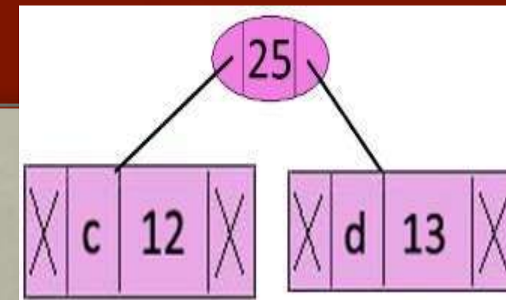
2. Extract two nodes with the minimum frequency from the min heap.

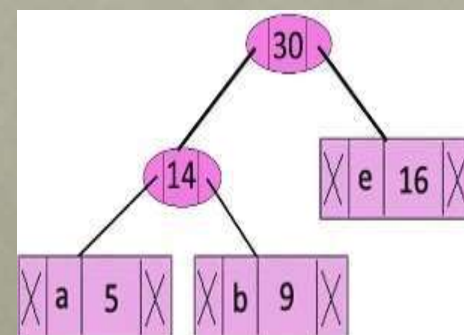
3. Create a new internal node with a frequency equal to the sum of the two nodes frequencies. Make the first extracted node as its left child and the other extracted node as its right child. Add this node to the min heap.

4. Repeat steps#2 and #3 until the heap contains only one node. The remaining node is the root node and the tree is complete.

character	Frequency
a	5
b	9
c	12
d	13
e	16
f	45



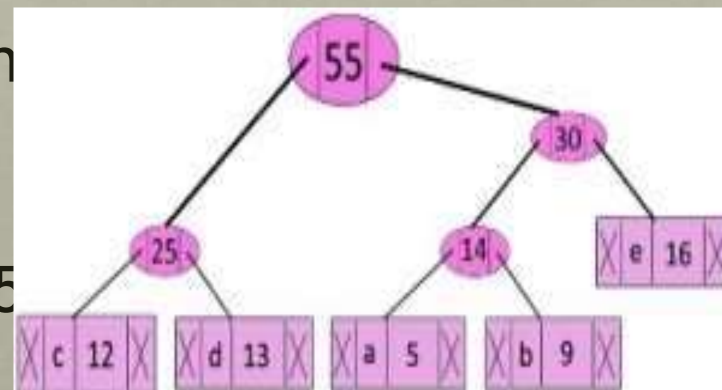


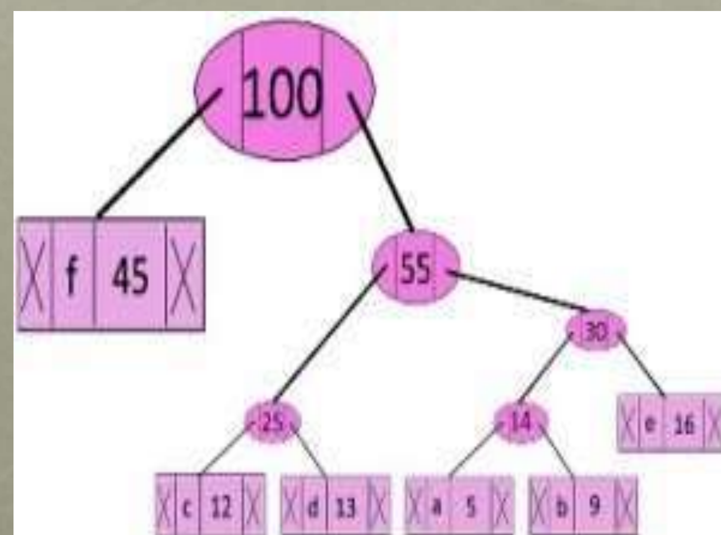


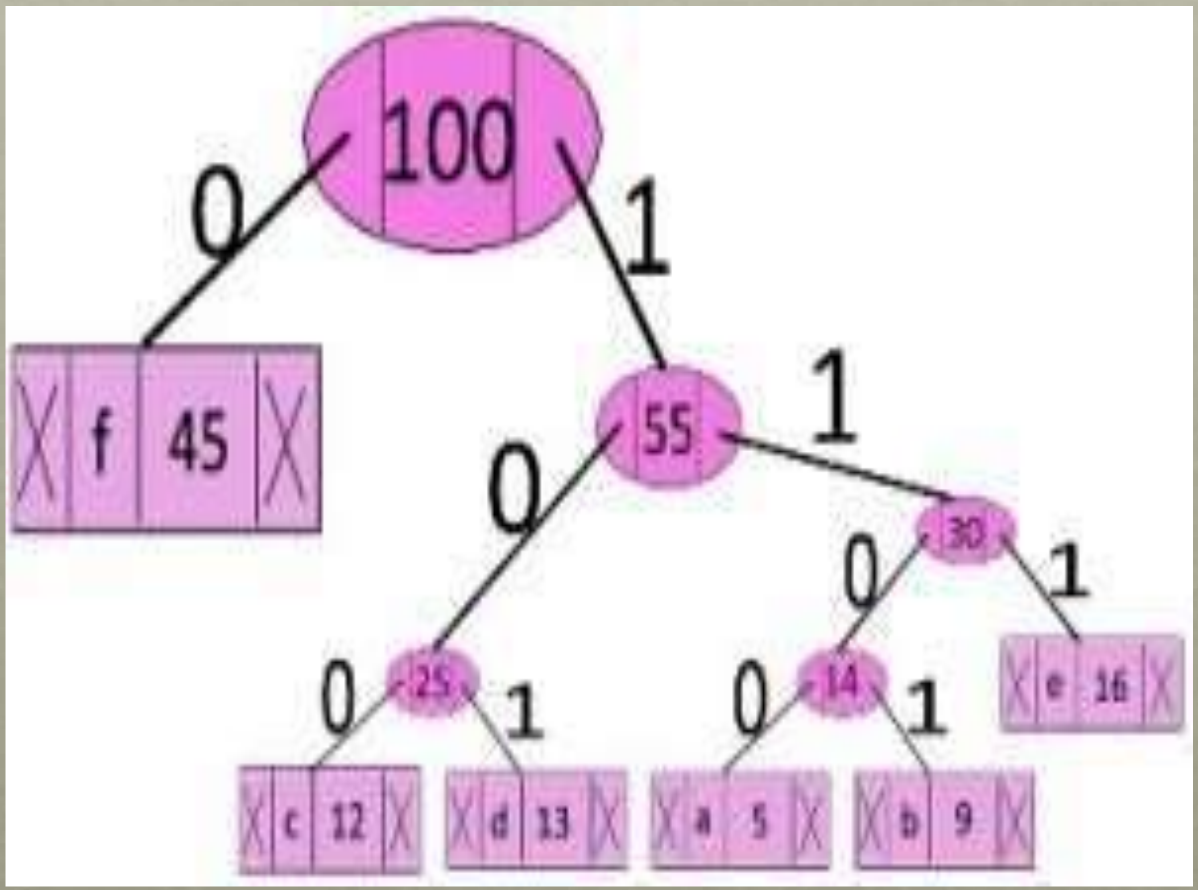
character	Frequency
c	12
d	13
Internal Node	14
e	16
f	45

character	Frequency
Internal Node	14
e	16
Internal Node	25
f	45

character	Frequen
Internal Node	25
Internal Node	30
f	45







100
0
1
f
45
55
0
1
30
0
1
e
16
25
0
1
14
0
1
c
12
d
13
a
5
b
9

Characteristics of Huffman Algorithm

1. Provides loss-less data compression (i.e. exact original data can be reconstructed from the compressed data). Note: A lossy data compression technique means that only an approximation of the original data can be reconstructed, however the compression rate is better in this case. Used in PNG, TIFF file formats and also some audio/video formats.
2. Provides variable-length encoding to the characters.
3. Bit-length of the code for a character depends upon its frequency. The most frequent character gets the smallest bit-length code and the least frequent character gets the largest bit-length code.
4. The generated codes for each character are prefix codes, which means that the codes (bit-string) assigned to one character is not prefix of code assigned to another character. Thus there is no ambiguity when decoding the generated bit stream.

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.

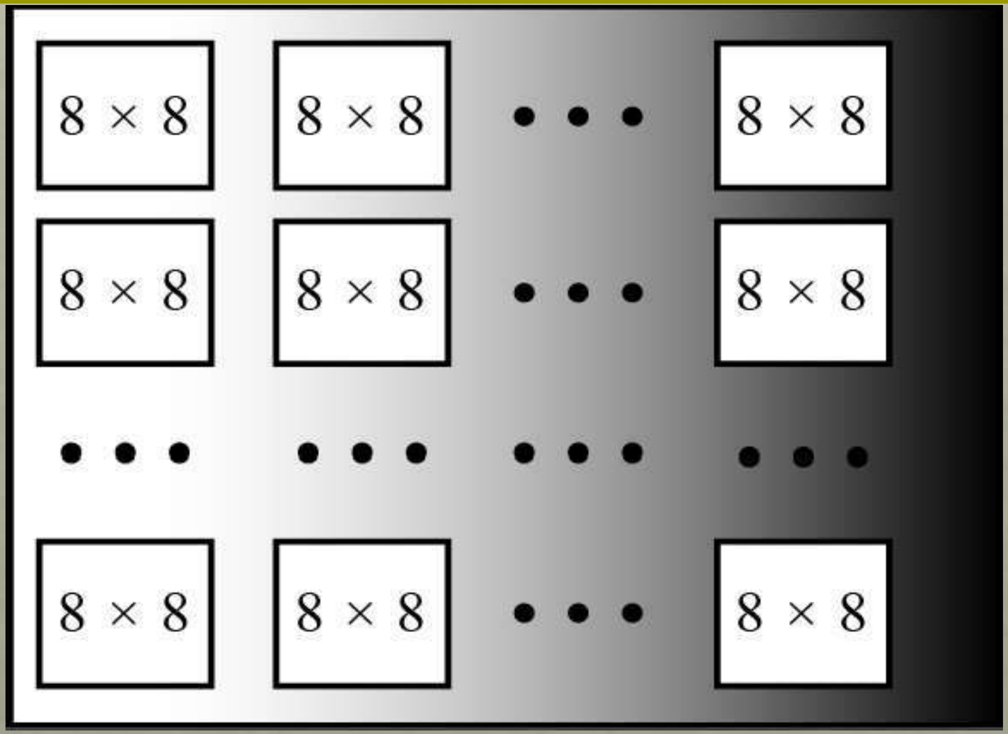


Figure 2 JPEG grayscale example, 640 × 480 pixels

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.

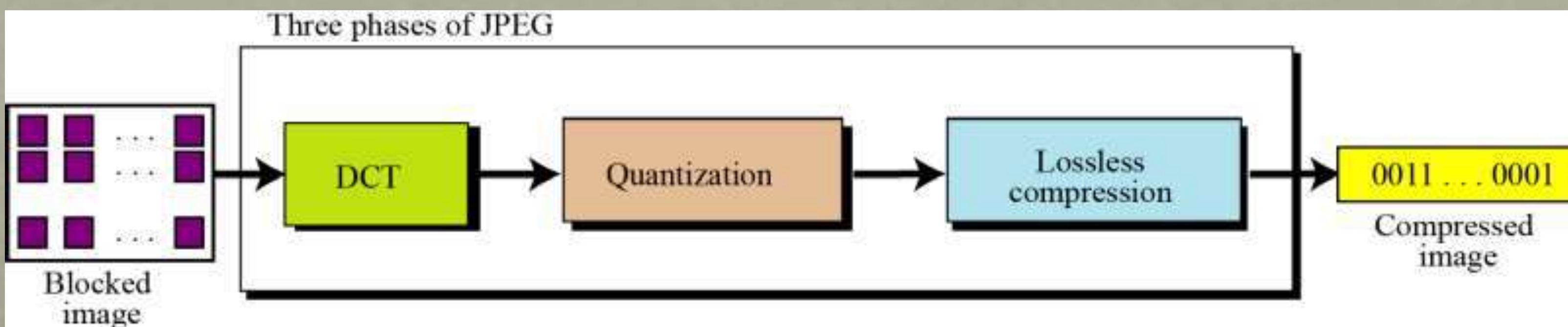


Figure 3 The JPEG compression process

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.

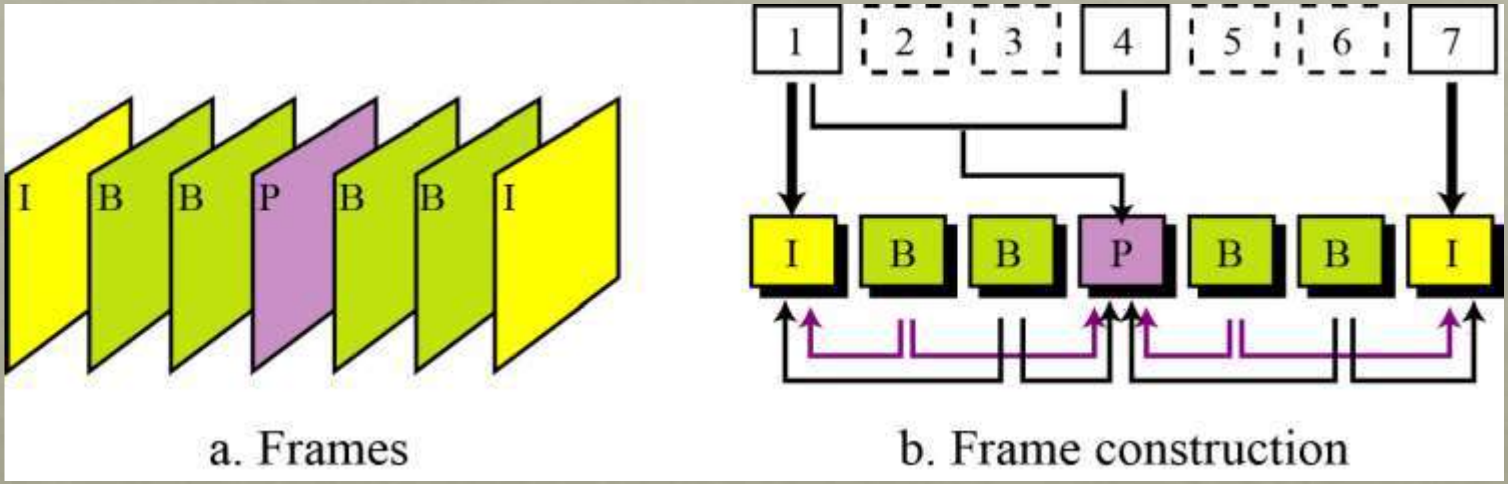


Figure 4 MPEG frames

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 $2 \le p \le 12$.**

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**
- 1920 1152 6

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**