Unit 5

Data Compression

Discussion Topics

- Storage space requirement by multimedia
- Coding requirements
- Source, entropy and hybrid coding
- JPEG
  - Introduction to JPEG
  - Basic steps in JPEG compression process
  - JPEG with Lossy sequential DCT-based mode
  - JPEG with Expanded lossy DCT-based mode
- MPEG

Storage space required by multimedia

- Graphics, audio, and video data require large storage space compared to text data.
- Similarly, such data also requires high bandwidth to transmit over network.
- To address above problem, data compression is used to minimize the size of data.
- Data compression is often referred to as coding
- JPEG and MPEG are very popular compression techniques for single pictures and for audio/video respectively.

We will use the terms “compression” and coding interchangeably.

Following is illustrative example of how much storage is required for video.

The storage for a video frame with 640 x 480 pixel resolution is 7.3728 million bits, if we assume that 24 bits are used to encode the luminance and chrominance components of each pixel.
Assuming a frame rate of 30 frames per second, the entire 7.3728 million bits should be transferred in 33.3 milliseconds, which is equivalent to a bandwidth of 221.184 million bits per second.
That is, the transport of such large number of bits of information and in such a short time requires high bandwidth.

Coding requirements?

- Coding of multimedia information is subject to certain constraints.
  - *The quality of data before/after coding/decoding should be as good as possible*
  - *The compression technique itself should be efficient. i.e. the coding/decoding process should be reasonably fast.*
  - *Compression technique should support different data formats.*
  - *Compression technique should be compatible in different platforms.*

Lossy vs lossless compression

	Lossy Compression	Lossless Compression
1.	Lossy compression is the method where part of the original data is lost.	Lossless Compression is the method in which original data is perfectly preserved.
2.	In Lossy compression, the data file can not be restored or rebuilt in its original form. We can only approximate the original data	In Lossless Compression, A file can be restored in its original form.
3.	In Lossy compression, Data's quality is compromised.	Lossless Compression does not compromise the data's quality.
4	Algorithms used in **Lossy** compression are: Discrete Cosine Transformation, etc.	Algorithms used in **Lossless** compression are: Run Length Encoding, Huffman Coding, Arithmetic encoding etc.
5.	Lossy compression is used in Images, audio, video.	Lossless Compression is used in Text, images, sound.
6.	Lossy compression is used to compress multimedia data (audio, video, and images), especially in applications such as streaming media and internet telephony.	Lossless data compression is used in many applications. For example, it is used in the ZIP file format and in the GNU tool gzip. Some image file formats, like PNG or GIF, use only lossless compression. Some image file formats, like PNG or GIF, use lossless compression.
7.	Lossy compression is used in cases when it is not necessary to perfectly preserve the original data or where small change in original data is acceptable. E.g. audio, video and images. Lossy compression exploits the perception capability of end-user where certain part of the data is not noticed and can be eliminated.	Lossless compression is used in cases when we need the original and decompressed data be same or where small change in original data is unacceptable. E.g. executable programs, text documents, and program source code. Lossless audio formats are most often used for archiving or production purposes

Source, Entropy and hybrid coding

- Compression techniques fit into different categories.
- For multimedia system, the techniques are categorized in three classes.

1. *Entropy coding:*
   - The term entropy encoding refers to all those coding and compression techniques which do not take into account the nature of the information to be compressed.
   - In other words, entropy encoding techniques simply treat all data as a sequence of bits and ignores the semantics of the information to be compressed.
   - Examples: Run-length coding, Huffman coding.
2. *Source coding:*

- The source encoding takes cognizance of the type of the original signal.
- For example, if the original signal is audio or video, source encoding uses its inherent characteristics in achieving better compression ratio.
- Examples: DCT
- Normally, it is expected that source encoding will produce better compression compared to entropy encoding techniques. But, it depends on the semantics of the data itself and may vary from application to application.

3. *Hybrid coding:*
   - In practical systems and standards, both entropy encoding and source encoding are usually combined for better effect in compression.
   - Examples: JPEG, MPEG etc

Compression Techniques

Entropy Encoding	Run-length Coding	
	Huflfman Coding	
	Arithmetic Coding	
Source Coding	Prediction	DPCM
		DM
	Transformation	FFT
		DCT
	Layered Coding	Bit Position
		Subsampling
		Sub-band Coding
	Vector Quantization	
Hybrid Coding	J PEG	
	MPEG	
	H.261	
	DVI RTV, DVI PLV	

Some basic entropy coding techniques:

1. run-length coding
2. Huffman coding
3. Arithmetic coding

1. Run-length coding:

- Images, audio and video dat streams often contain sequences of the same bytes.
- By replacing those repeated sequences with number of occurrences, data can be reduced.
- This is called run-length coding.
- Example:

  Original data: ABCCCCCCCCDEFGGG

  Compressed data: ABC!8DEFGGG

- Here, we use special character ! That is not the part of data itself and indicates the repetition of a symbol.
- We have assumed that, we compress only when a symbol(byte) occurs more than 4 times.

3. **Arithmetic coding**

- Arithmetic coding is a type of entropy encoding utilized in lossless data compression.
- It is similar to Huffman coding in that "frequently occurring characters are stored with fewer bits and not-so-frequently occurring characters are stored with more bits, resulting in fewer bits used in total.
- However, unlike Huffman coding, arithmetic does not encode symbols separately; each symbol is coded by considering the prior data.

2. **Huffman Coding**:

- Huffman coding is a data coding/compression technique in which the characters in a file are given unique binary codes, thereby reducing the size of the file.
- Huffman codes are used to compress text data by representing each alphabet by unique binary codes in an optimal way.
- Huffman coding is based on the idea that “different characters do not have to be coded with fixed number of bits.
- In Huffman coding,
  - frequently occurring characters are coded with shorter codes.
  - Seldom-occurring characters are coded with longer codes.
- To Compute Huffman code, a tree is computed where every leaf represents the characters to be encoded. Each internal node contains frequency of characters belonging to the subtree.

Huffman coding algorithm

- A greedy algorithm to find optimal prefix codes.

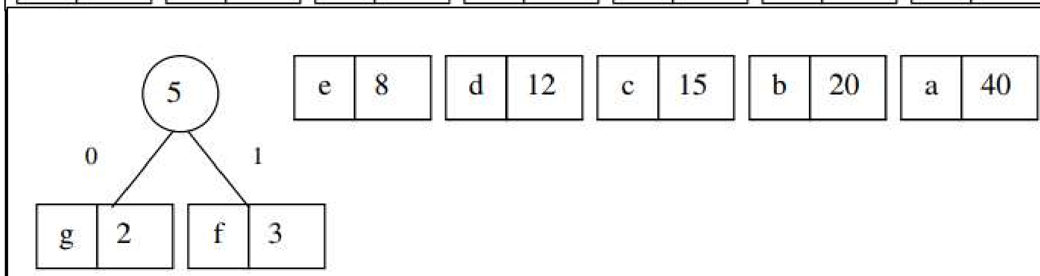
Basic Idea:

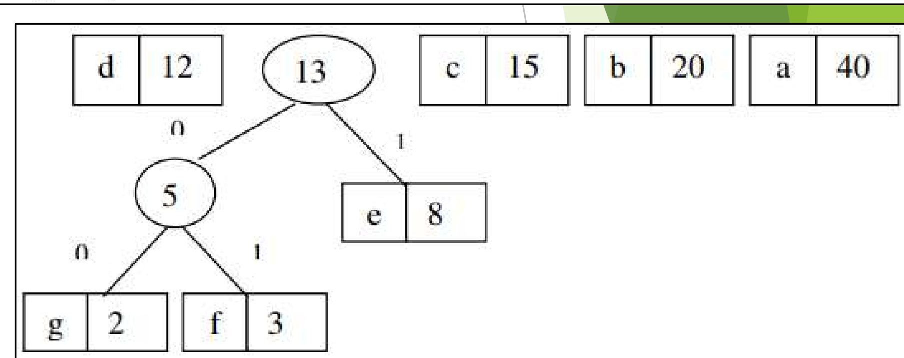
1. Let C is a set of distinct characters in a data file.
2. Let f_c be the frequency of for each character $c \epsilon S$.
3. Construct a priority queue of the characters based on frequencies
4. Take two minimum elements and constructs a tree with root as sum of the frequencies and children as the elements themselves.
5. Label the left edge as 0 and right edge as 1.
6. Insert the root into priority queue.
7. Repeat step 4 to 6 until the queue contains only one element.
8. Construct the codes by following the links from root to leaves.

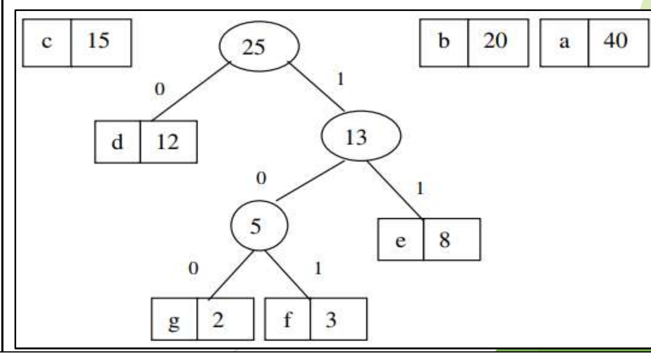
Example: $C = \{a, b, c, d, e, f, g\}$ $f(c) = 40, 20, 15, 12, 8, 3, 2$

Initial priority queue is

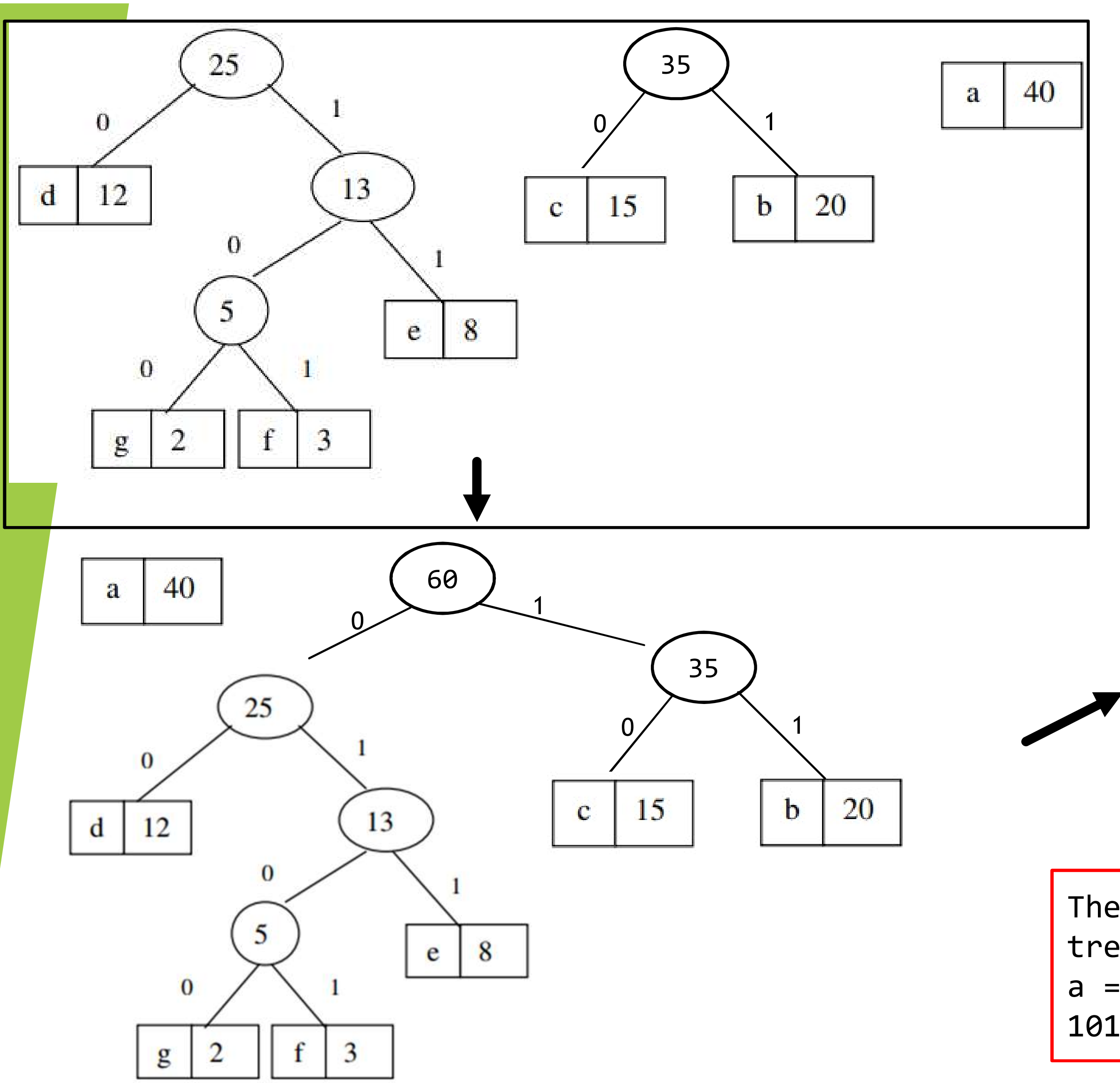
g	2	f	3	e	8	d	12	c	15	b	20	a	40

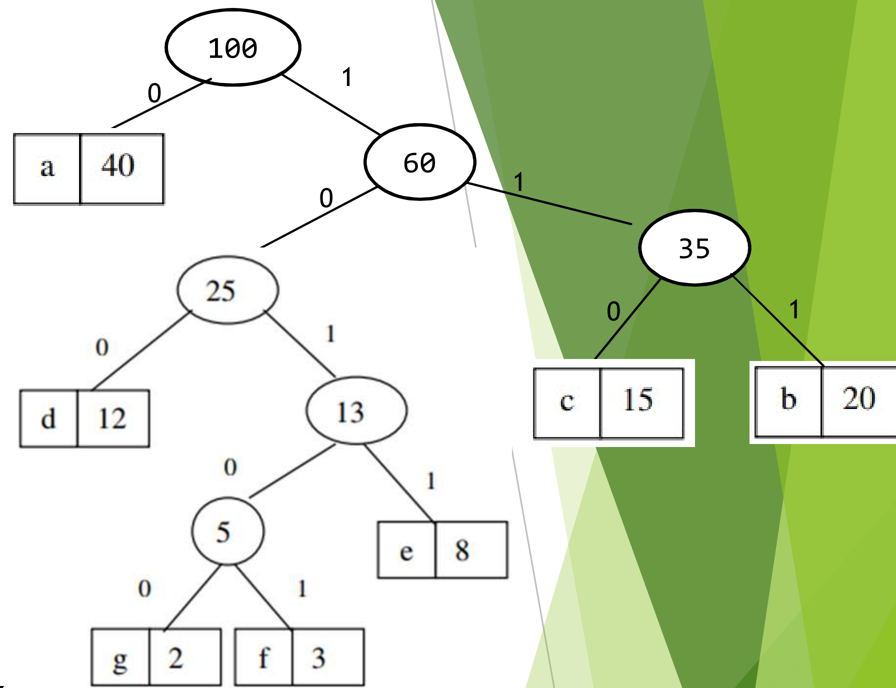






Continue the process.. .





The prefix codes are obtained by traversing the tree from root to each leaf.
a = 0, b = 111, c= 110, d = 100, e = 1011, f = 10101, g = 10100

Prepared by: Shiv Raj Pant

- The Huffman code table can be used to compress data without any loss.
- The same Huffman table must be available for encoding and decoding.
- Huffman coding can be used to compress any kind of data file.
- E.g. for image, each pixel of an image can be represented in a bit stream by coding pixels individually and reading them line by line.
- For videos, Huffman coding can be used for a single sequence of images, for a set of scenes or even an entire video clip.

Transformation coding techniques

- In transformation techniques, data is transformed into another mathematical domain which is more suitable for compression. An example of transformation is Fourier transformation, which transforms data from time domain to frequency domain.
- For this technique to work, the inverse transformation must exist and be known for compression and decompression.
- The most effective transformations for image compression are DCT (Discrete cosine transformation) and FFT (Fast Fourier transformation)

DCT:

A discrete cosine transform (DCT) expresses a finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies.

The DCT, first proposed by Nasir Ahmed in 1972, is a widely used transformation technique in data compression. It is used in most digital media, including digital images (such as JPEG), digital video (such as MPEG), digital audio (such as MP3), digital television (such as SDTV, HDTV) etc.

The discrete cosine transform (DCT) represents an image as a sum of sinusoids of varying magnitudes and frequencies. It computes the discrete cosine transform (DCT) of an image. The DCT has the property that, for a typical image, most of the visually significant information about the image is concentrated in just a few coefficients of the DCT. For this reason, the DCT is often used in image compression applications.

JPEG (Joint Picture Experts group)

- It’s a standard image format for containing lossy and compressed image data.
- Despite the huge reduction in file size JPEG images maintain reasonable image quality.
- This unique compression feature allows JPEG files to be used widely on the Internet, Computers, and Mobile Devices.

Note: jpeg and Jpg are same things.

- The first JPEG standard was issued in 1992 by ISO (International Standards Organization)
- JPEG can be applied to color and grayscale images. It an also be used for video sequences known as “motion JPEG”

- To facilitate distribution and application, JPEG standard has imposed following requirement:

  i. *The JPEG implementation should be independent image size.*

  ii. *... Should be applicable to any image and pixel aspect ratio*

  iii. *color representation should be independent of particular implementation*

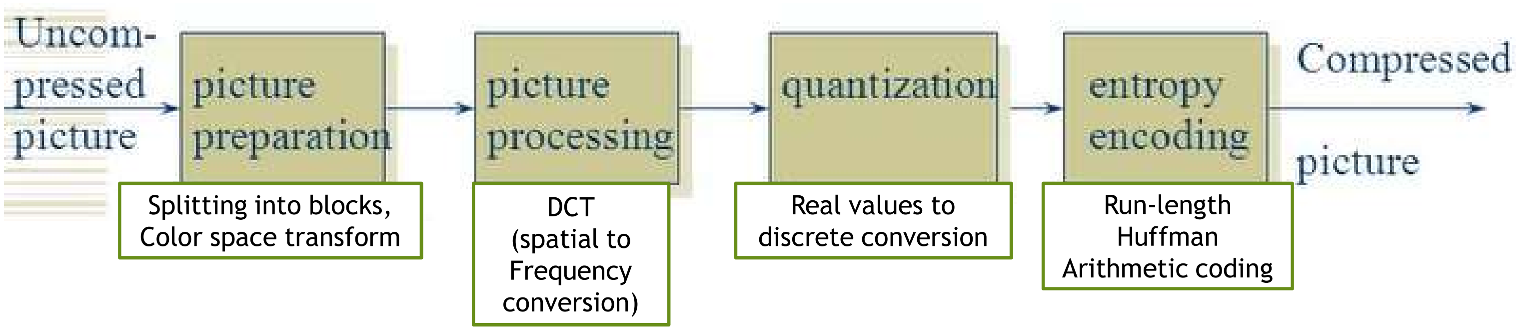
  iv. *Image content may be of any complexity with any statistical characteristics.*

  v. *The processing complexity of an application must be achievable by software running on standard processors as possible.*

Steps of JPEG compression process:

The general steps in JPEG compression are shown in figure:

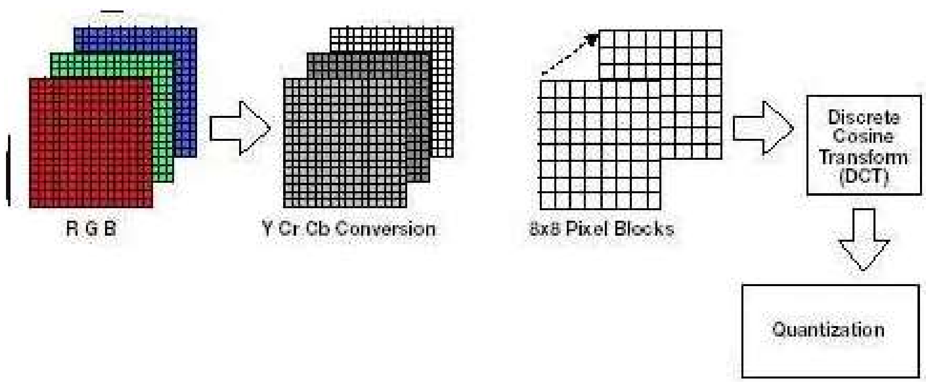
1. **Picture preparation**
2. **Picture processing**
3. **Quantization**
4. **Entropy coding**



1. **Picture preparation:** In this step, the image is prepared for compression. Basically, it involves two tasks: Color space transform(conversion) and splitting:

   We convert R, G, B to Y, Cb, Cr model. Here Y is for brightness, Cb is color blueness and Cr stands for Color redness. Because the human eye is more sensitive to luminance than chrominance, image bandwidth can be reduced by subsampling the (Cb,Cr) fields.

   The image is split into the blocks of 8*8 size. If the image dimensions are not multiples of 8, the last row and/or column of the image is duplicated as needed.



2. **Picture processing:** We apply Discrete cosine transform on each block. The DCT stage of JPEG exploits the fact that the human eye favors low-frequency image information over high-frequency details. The 8×8 DCT transforms the image from the spatial domain into the frequency domain. The discrete cosine transform (DCT) represents an image as a sum of sinusoids of varying magnitudes and frequencies.

3. **Quantization**: After DCT has been performed on the 8*8 image block, the real-number values from previous step are mapped to discrete values. Since, the human eye is more attuned to low frequency info. than high-frequency details. Therefore, small errors in high-frequency representation are not easily noticed, and eliminating some high-frequency components entirely is often visually acceptable. The JPEG quantization process takes advantage of this fact to reduce the amount of DCT information that needs to be coded for a given 8×8 block.

4. Entropy coding:

- Finally, the data is further compressed using entropy coding methods such as run-length coding, Huffman coding and/or Arithmetic coding.
- This is the final lossless compression applied during JPEG compression process.

Modes of JPEG image processing

- After image preparation, the further processing for compression can be done in different modes.

1. *Lossy sequential DCT-based mode*
2. *Expanded (Extended) Lossy DCT-based mode*
3. *Lossless mode*
4. *Hierarchical mode*

1. *Lossy sequential DCT-based mode*

- In this mode,
  - Image samples are grouped into units of 8x8 pixels
  - Each pixel is encoded using 8 bits.
  - the DCT used is similar to Discrete Fourier transformation (DFT). Which represents the data in 2D frequency.

2. *Expanded (Extended) Lossy DCT-based mode*
   - This mode differs from 1 in terms of the number of bits per sample
   - It uses 12 bits per sample instead of 8.
   - Also, in this mode, JPEG uses progressive encoding. i.e, in the first run, a very rough image appears which looks out of focus and is refined during successive steps.

3. *Lossless mode*
   - The lossless mode uses data units of single pixels for image preparation.

- It may use anywhere between 2 to 16 bits per pixel
- In this mode, the image processing and quantization use a predictive technique instead of transformation(DCT).

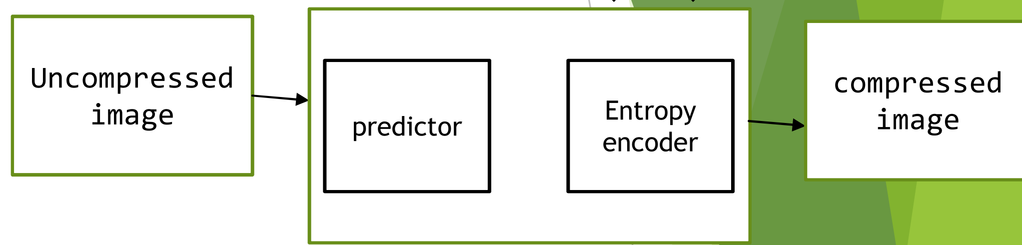


Fig: lossless mode of image processing in JPEG

4. *Hierarchical mode*

- This mode uses either DCT-based algorithm or alternative lossless compression technique
- The main feature of this mode is encoding of image at different resolutions.
- Hierarchical encoding requires more storage. But the compressed image is quickly available at different resolutions. (less processing power required to display)

MPEG(Moving Picture Coding Experts Group)

- it describes a whole family of international standards for the compression of audio-visual digital data. The most known are MPEG-1, MPEG-2 and MPEG-4.
- The MPEG-1 Standard was published 1992 and its aim was it to provide:
  - VHS(video home system) quality with a bandwidth of 1.2 Mb/s, which allowed to play a video in real time from a CD-ROM.

    VHS is a standard for consumer-level analog video recording on tape cassettes.
  - a fast forward and backward search and a synchronization of audio and video.
  - a stable behaviour, in cases of data loss
  - low computation times for encoding and decoding
- In 1994 MPEG-2 was released, which allowed a higher quality with a slightly higher bandwidth.
- Later, MPEG-2 was also used for High Definition Television (HDTV) and DVD, which made the MPEG-3 standard disappear completely.
- MPEG-4 was released 1998 and it provided lower bit rates (10Kb/s to 1Mb/s) with a good quality. It was a major development from MPEG-2 and was designed for the use in interactive environments, such as multimedia applications and video communication.

The MPEG Compression

The MPEG compression algorithm encodes the data in 5 steps:

1. *Reduction of the Resolution (color-subsampling)*
2. *Motion Estimation*
3. *Discrete Cosine Transform*
4. *Quantization*
5. *Entropy coding*

1. Reduction of the Resolution (color-subsampling)

- The human eye has a lower sensibility to color information than to dark-bright contrasts.
- A conversion from RGB-color-space into YUV color components help to use this effect for compression.
- Each image is divided into areas called macro blocks.
- Each macro block is partitioned into 16X16 pixels for luminance.
- The chrominance components U and V can be reduced (subsampling) to half of the pixels in horizontal direction (4:2:2), or a half of the pixels in both the horizontal and vertical (4:2:0)

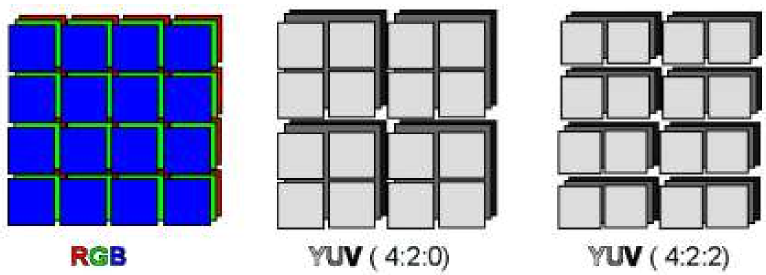


Fig: During subsampling, 2 or 4 pixel values of the chrominance channel can be grouped together.

2. Motion estimation

- An MPEG video can be understood as a sequence of frames.
- Because two successive frames of a video sequence often have small differences (except in scene changes),the MPEG standard offers a way of reducing this temporal redundancy.
- It uses three types of frames:
  - **I-frames (intra-coded)** :The I-frames are “key-frames”, which have no reference to other frames. An I-frame is treated as a still image. MPEG makes use of JPEG for I-frames. The I-frame compression is lowest in MPEG process. I-frames are points of random access in MPEG video stream.
  - **P-frames (predictive-coded):** The P-frames can be predicted from an earlier I-frame or P-frame.. P-frames cannot be reconstructed without their referencing frame, but they need less space than the I-frames, because only the differences between two successive images are stored.
  - **B-frames (bidirectionaly predictive-coded):** The B-frames are a two directional version of the P-frame, referring to both directions (one forward frame and one backward frame). B-frames cannot be referenced by other P- or B-frames, because they are interpolated from forward and backward frames.
- The references between the different types of frames are realized by a process called *motion estimation*. The correlation between two frames in terms of motion is represented by a motion vector.

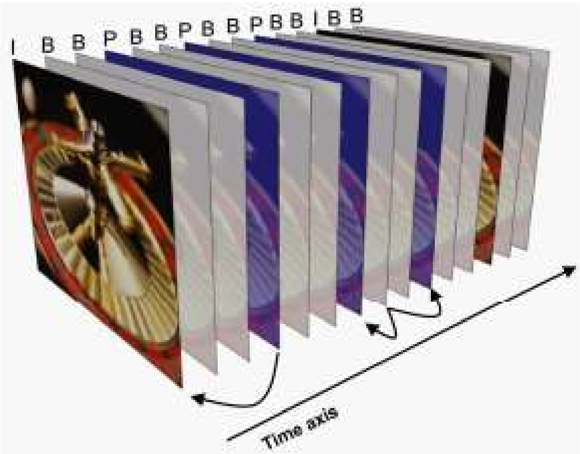


Fig: An MPEG frame sequence with two possible references: a P-frame referring to a I-frame and a B-frame referring to two P-frames

3. Discrete Cosine Transform

- DCT allows, similar to the Fast Fourier Transform (FFT), a representation of image data in terms of frequency components.
- So the frame-blocks (8x8 or 16x16 pixels) can be represented as frequency components to allow lossy compression.

4. Quantization

- After DCT has been performed on image block, the real-number values from previous step are mapped to discrete values.
- The human eyes are more reactive to low frequencies than to high ones. Higher frequencies end up with a zero entry after quantization and the data is reduced significantly.
-

5. Entropy coding:

- In this step, Run Length Encoding and Huffman coding are used to further compress the data stream.

MPEG Audio Encoding

- MPEG/audio is a generic audio compression standard.
- Instead, the coder exploits(utilizes) the perceptual limitations of the human auditory system.
- Much of the compression results from the removal of perceptually irrelevant parts of the audio signal. Removal of such parts is not perceived by human ear.
- This MPEG/audio compression is useful when the audio is meant to be heard by the human ear.
- If we are compressing audio to preserve it for future (archiving), then we should use lossless compression to keep the original quality. In this case we should not use MPEG because it is lossy.

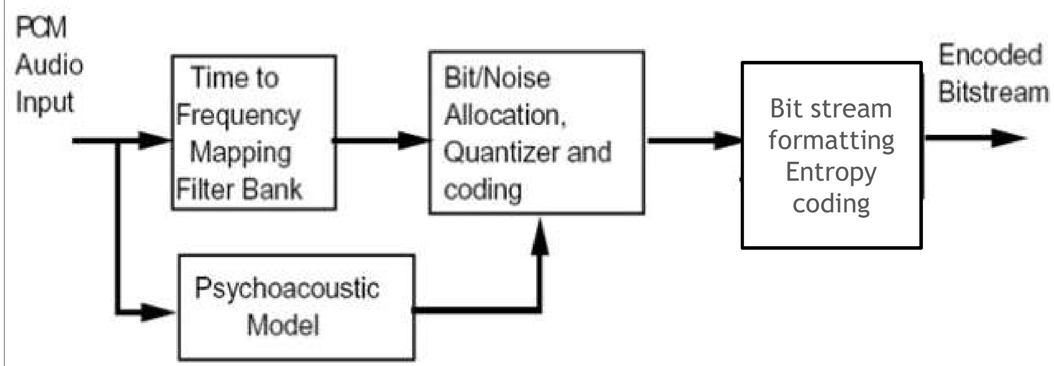


Fig: Basic steps of MPEG audio encoding

Steps:

1. Fast fourier transformation is used to map from time domain to frequency domain.
2. Psychoacoustic model used to determine which frequency components are usful. Those frequencies are removed which are not perceived by human ear. And later low bits are given to frequencies less perceived.
3. The quantized data is formatted and entropy codded for further compression.

End of unit 5

Thank you!