AY 2019/2020
EE6403 Distributed Multimedia Systems

Part 3 Video Compression & Standards, Digital Audio, Media Storage & Display

Dr Yap Kim Hui

Room: S2-B2b-53
Tel: 6790 4339
Email: ekhyap@ntu.edu.sg

Section II
Media Compression and Standards

This Section

- Video compression basics
- Motion estimation and compensation
- MPEG standard: MPEG-1, MPEG-2
- Video standards
- Digital audio
- Media storage
- Media display

Video

Video Data Redundancy

Recap: Spatial Redundancy

- Spatial redundancy refers to the statistical correlation between pixels within an image (or more specifically, within a small image neighborhood).
- It is also called intraframe redundancy.

Recap: Temporal Redundancy

- Temporal redundancy refers to the statistical correlation between pixels from successive frames in a video sequence.
- Therefore, it is also called interframe redundancy.

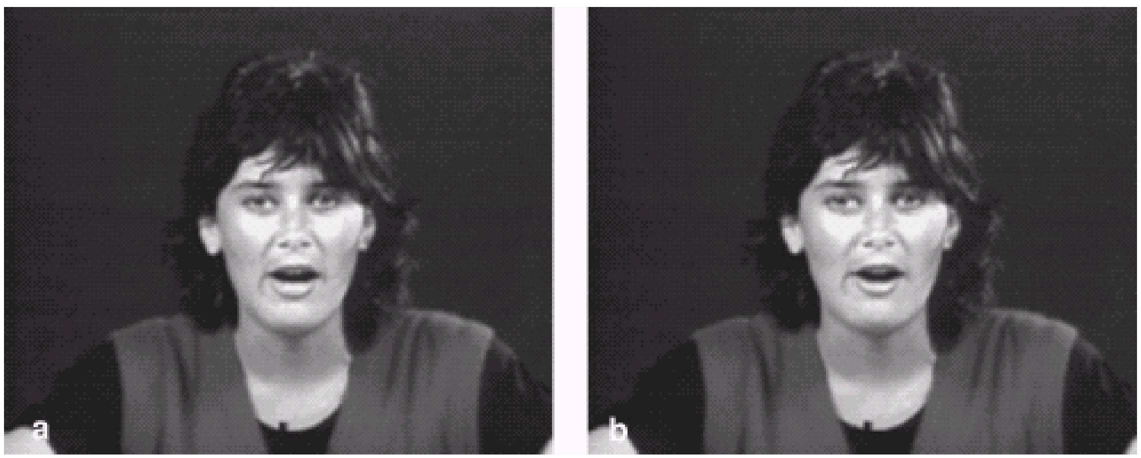


FIGURE 1.5 (a) The 21st frame, and (b) 22nd frame of the "Miss America" sequence.

Figure source:
Yun Q. Shi and Huifang Sun, Image and video compression for multimedia engineering: fundamentals, algorithms, and standards, CRC Press, 2000

Recap: Psychovisual Redundancy

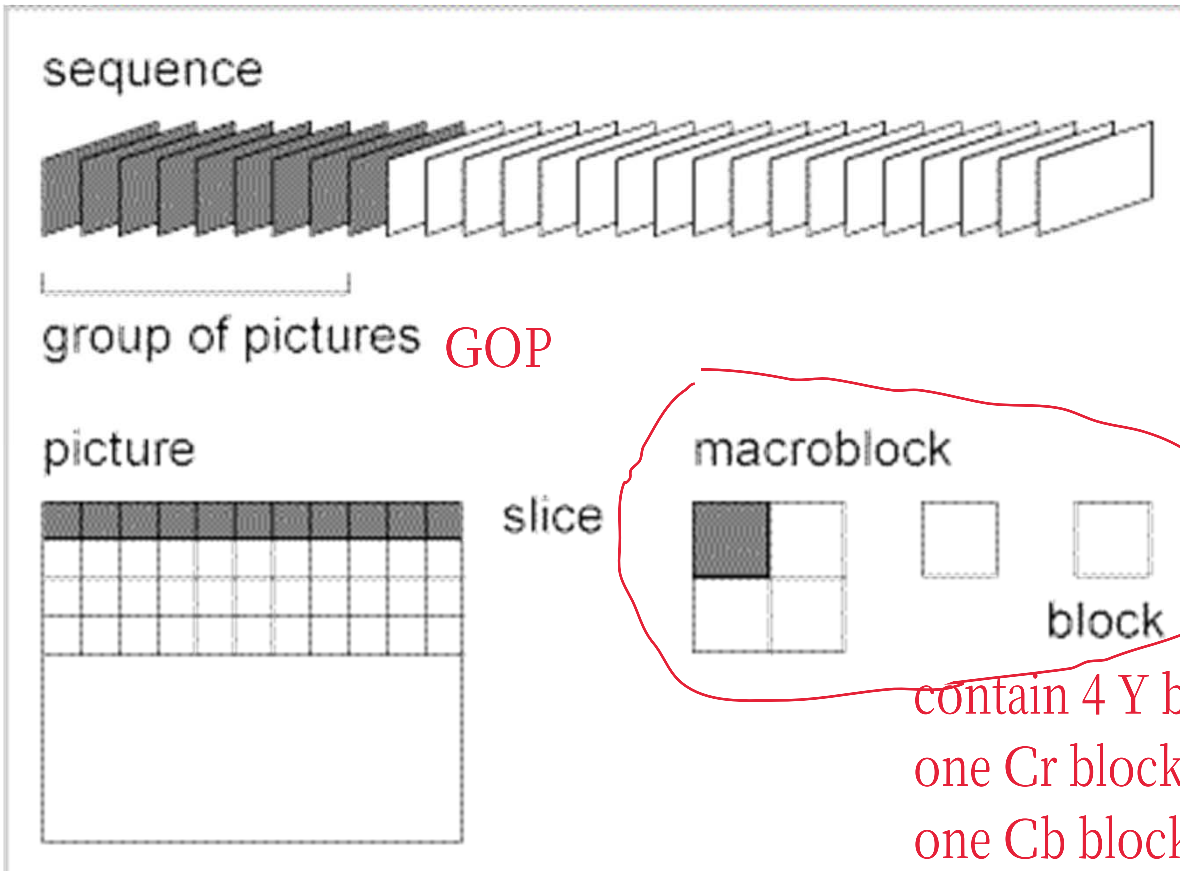
- Frequency masking
  - Human is less sensitive to noise or distortion in high frequency components, and vice versa.
- Color masking
  - Human is more sensitive to luminance (brightness) components than chrominance (color) components.

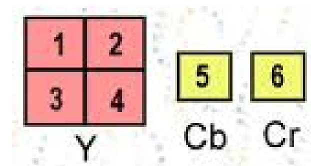
Video Compression Basics

Video Compression Basics

- Video (moving picture):
  - A sequence of pictures/images/frames, with audio.
- Video compression utilizes the following to perform compression:
  - Spatial redundancy
  - Temporal redundancy
  - Psycho-visual redundancy
- Main idea in video compression:
  - Motion estimation:
    prediction of motion vector for macroblocks.
  - Motion compensation: best matched block
    encode small difference between predicted and actual macroblocks.

Video Structure





contain 4 Y block,(luminance)
one Cr block(8 by 8)
one Cb block
eyes are less sensitive to latter two,
so need only 1 bl9ck

Macroblock

- A macroblock contains a 16x16 Y component and spatially corresponding 8x8 C_b and C_r components.
- A macroblock has four luminance blocks and two chrominance blocks and it is the basic unit for DCT operation and motion compensation.

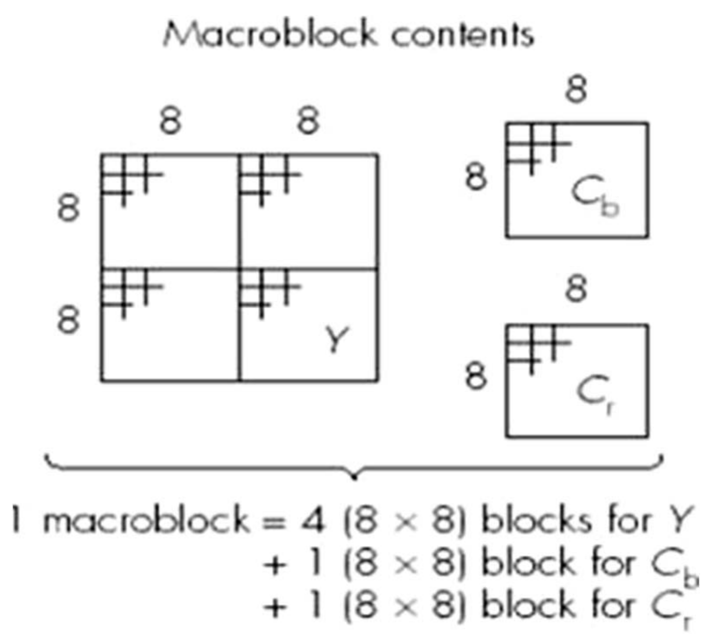
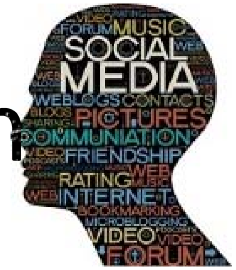


Figure source:
Fred Halsall, Multimedia communications: applications, networks, protocols and standards, Addison-Wesley, 2001

Intraframe & Interframe Compression



- Many video compression methods such as MPEG use both intraframe and interframe compression.
- Intraframe compression
  - Consider each video frame as still image (e.g. I-frames in MPEG).
  - Reduce spatial redundancy.
  - Employ JPEG-based compression.
- Interframe compression
  - Utilize correlation along temporal domain (e.g. P & B-frames in MPEG).
  - Reduce temporal redundancy.
  - Employ motion estimation and compensation.

Motion Estimation & Compensation

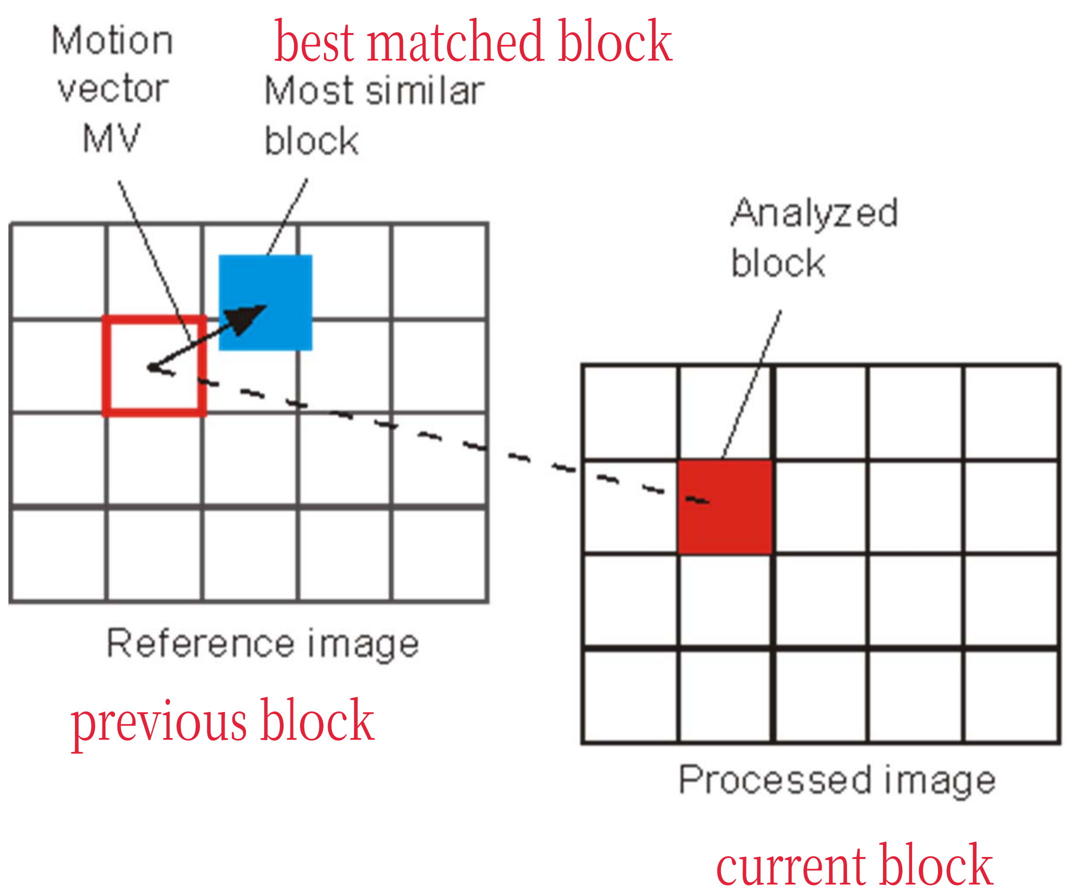
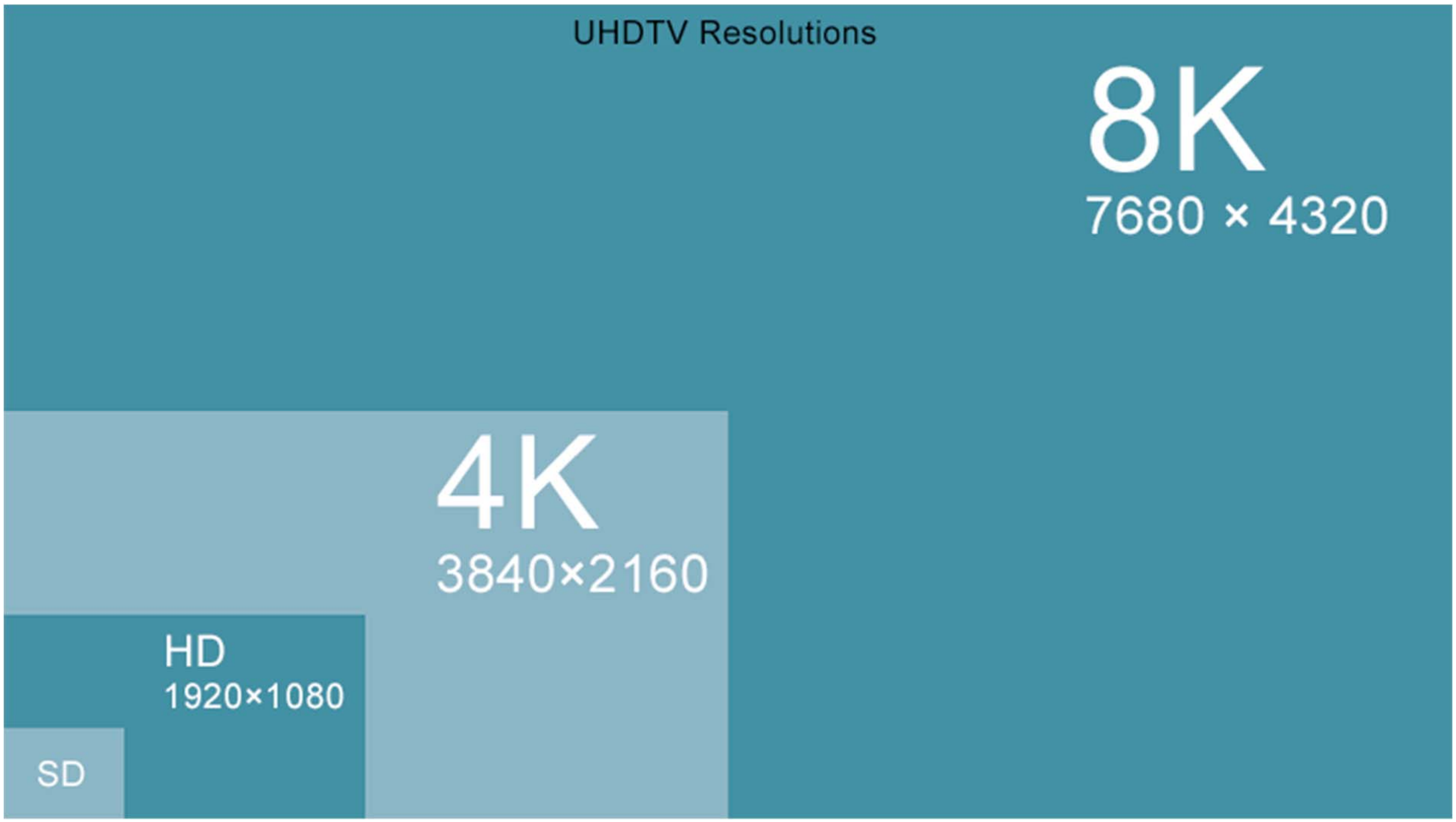


Figure source: http://www.multimedia.edu.pl

Video Resolution

- Standard definition (SD)
  - NTSC (480i): 720 x 480, 640 x 480 480 indicate the number of rows
  - PAL (576i): 720 x 576
- High definition (HD)
  - 720p: 1280 x 720 progressive
  - 1080i: 1920 x 1080 interlaced
  - 1080p: 1920 x 1080 progressive (Full HD)
- Ultra high definition (UHD)
  - 4K ~ 4 times FHD
  - 8K ~ 4 times 4K

Video Resolution

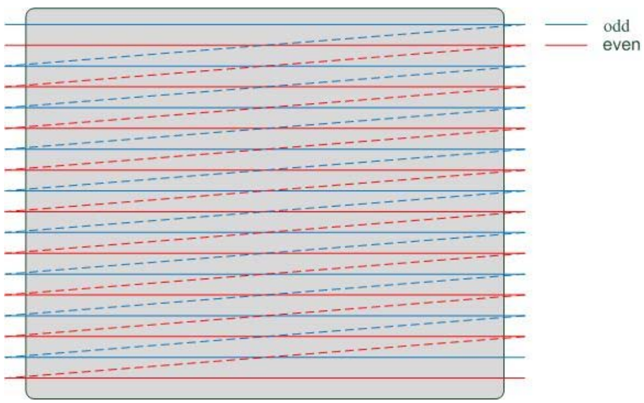


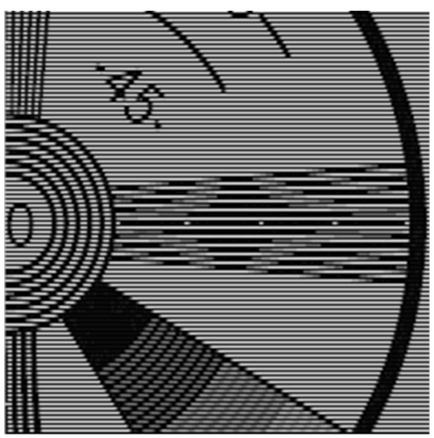
Scan Type

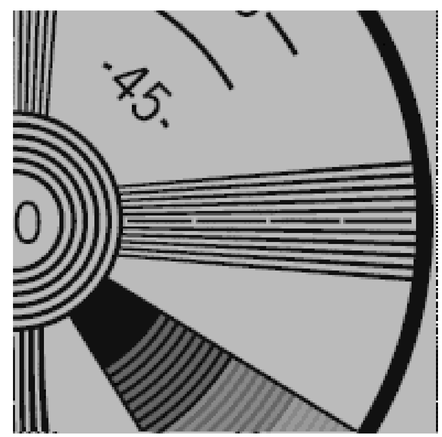
scan odd lines first, then all even lines, bandwidth is half

- **Interlaced**: 2 fields (odd and even lines) are scanned separately.
- **Progressive**: all the lines are scanned in sequence.

frame bits need to transmit is doubled







less sensitive on it

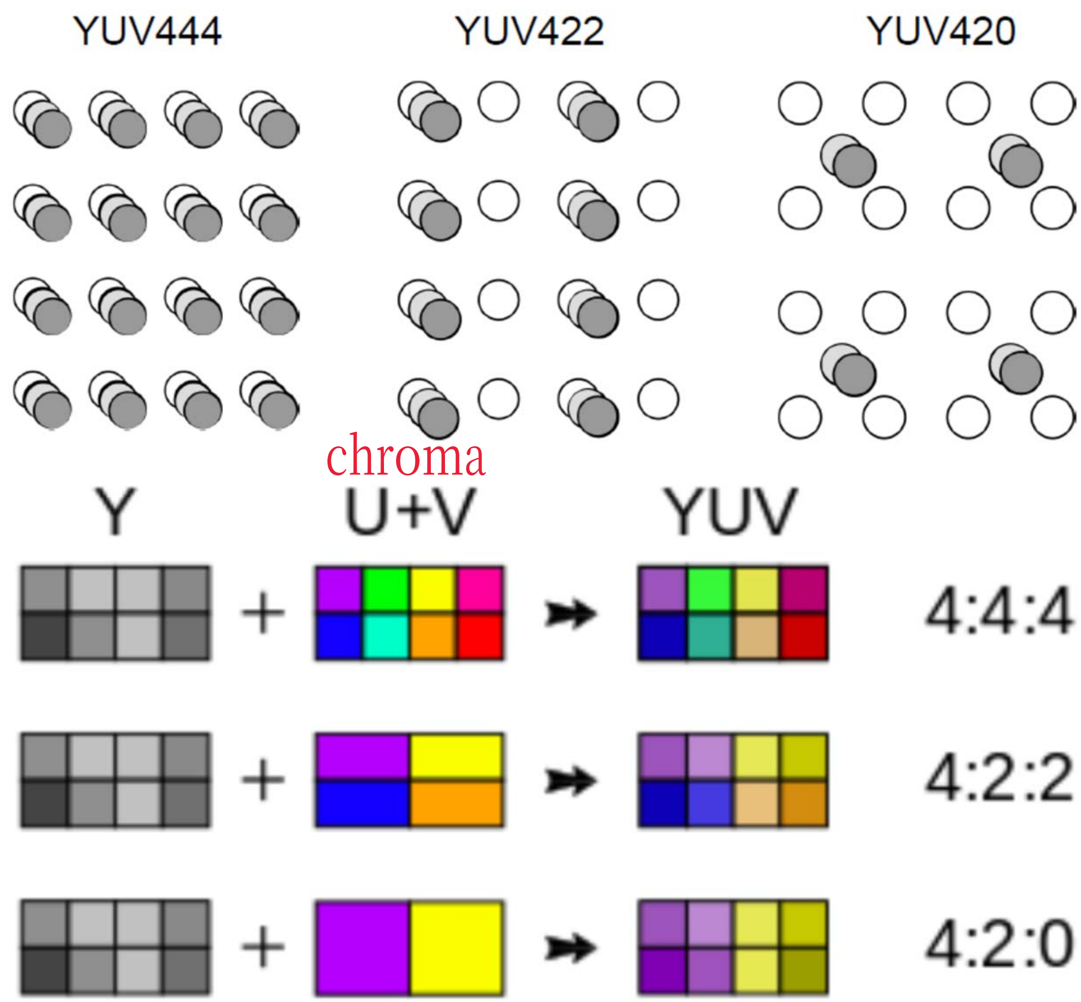
Chroma Subsampling

- Subsampling formats:
  - 4:4:4

    No subsampling
  - 4:2:2

    horizontal subsampling
  - 4:2:0

    horizontal and vertical subsampling



Chroma Subsampling

- Human is more sensitive to luminance component than chrominance components.
- Hence, to achieve compression, chrominance components are subsampled.
- 4:2:2 sub-sampling

  For every two horizontal Y samples, 1 Cr and 1 Cb are sampled.
- 4:2:0 sub-sampling

  For every 2x2 Y samples, 1 Cr and 1 Cb are sampled.

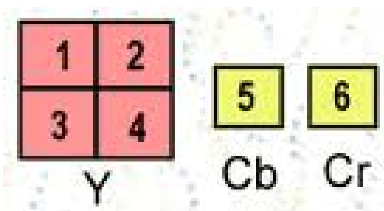
MPEG

MPEG Overview

- MPEG: Moving Pictures Experts Group, was established in 1988 for the development of digital video.
- Define only a compressed bitstream that implicitly defines the decoder.
- The compression algorithms, and thus the encoders, are completely up to the manufacturers.

MPEG

- Objective: to encode/compress video signals.
- MPEG standard is based on DCT coding and motion estimation & compensation.
  - DCT coding: remove intraframe redundancy.
  - Motion estimation & compensation: remove interframe redundancy.
- A color video source has three components
  - a luminance component (Y)
  - two chrominance components (C_b *and* C_r) in usually 4:2:0 subsampling format.



MPEG Video Structure/Hierarchy

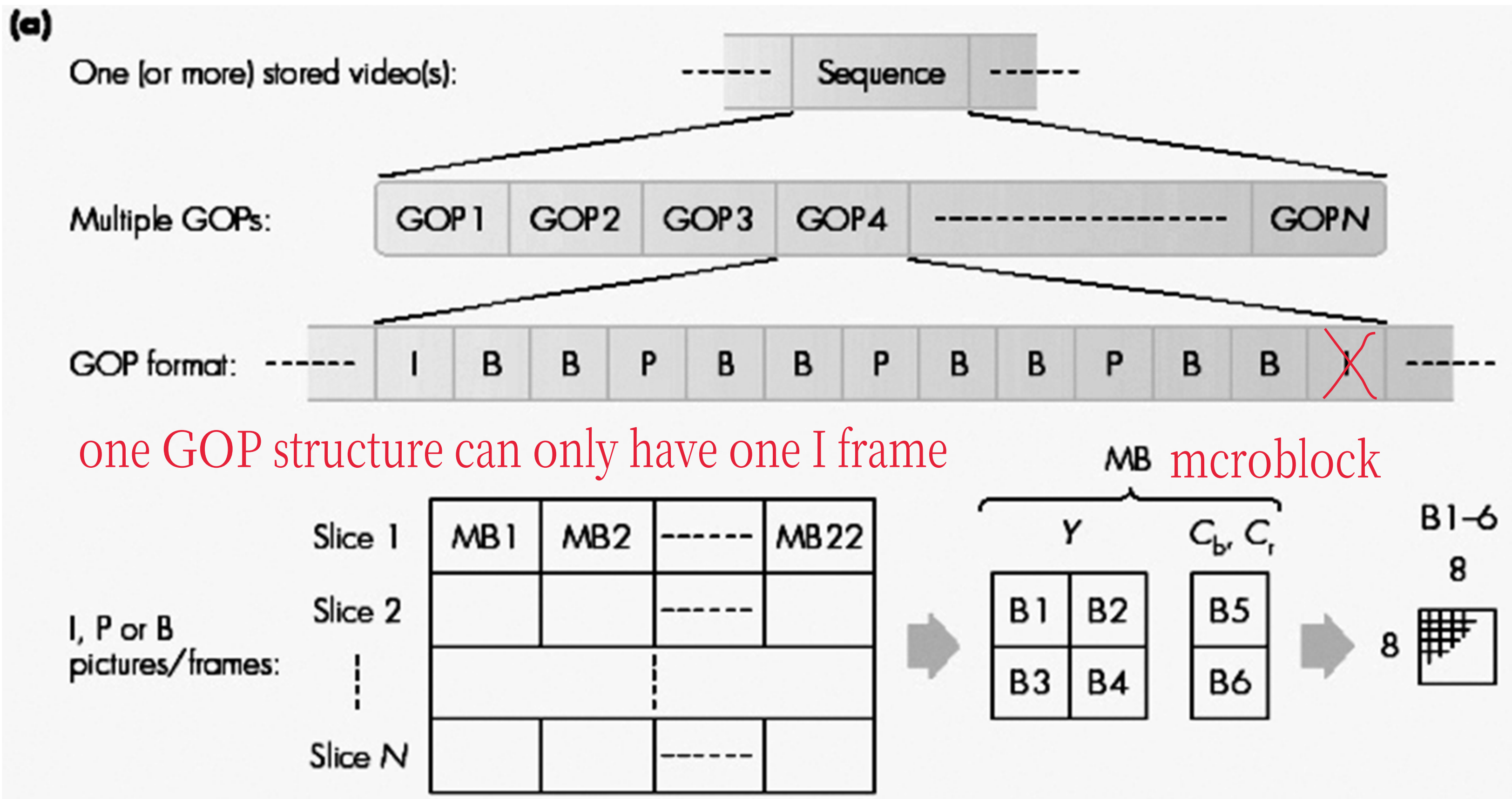


Figure source:
Fred Halsall, Multimedia communications: applications, networks, protocols and standards, Addison-Wesley, 2001

Group of Pictures (GOPs)

I need no reference

another GOP structure

N=9 total frame number

M=3 distance between 2 nearest reference frame (I P frame)

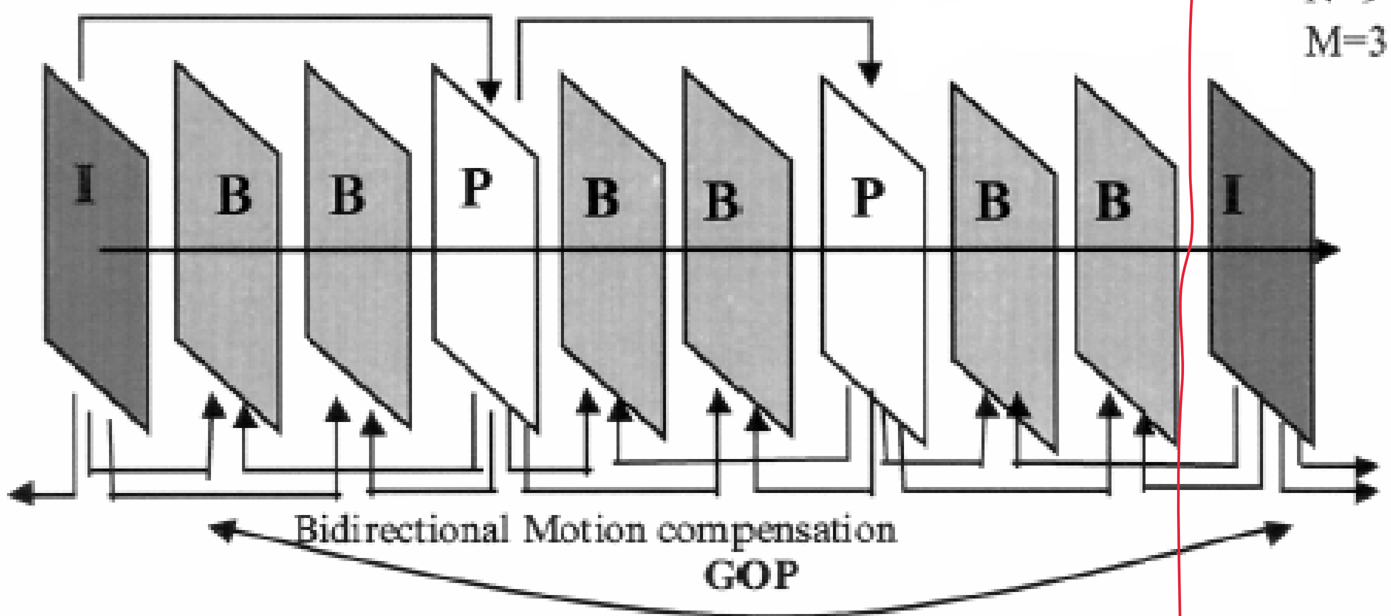


FIGURE 16.1 A group of pictures of video sequence in display order.

Figure source:
Yun Q. Shi and Huifang Sun, Image and video compression for multimedia engineering: fundamentals, algorithms, and standards, CRC Press, 2000

GOP

- The video sequence is divided into group of pictures (GOPs).
- Each GOP may include three types of pictures:
  - I-frames (Intra-coded frames)
  - P-frame (Predictive-coded frame)
  - B-frame (Bidirectional-coded frame)

I-frames, P-frames, B-frames

- I-frames
  - coded without reference to any other frames.
  - Y, C_b, C_r blocks are encoded independently using JPEG algorithm.

- P-frames and B-frames
  - known as intercoded or interpolation frames.
  - P-frames are coded using motion estimation and compensation from previous anchor I-frame or P-frame.
  - B-frames are coded using predictions from either past, future, or both anchor frames.

Parameters: M and N

- Size of GOP (N):
  - Distance between two nearest I-frames.
  - Typical values: 12-15
- Parameter M
  - Distance between two nearest anchor frames.
  - Typical values: 1-3
- Parameters N, M
  - User-selectable parameters, which are chosen during encoding.
  - Trade-off between coding performance and error propagation (drift).

Exercise 3.1: GOP Structure

p24 tu

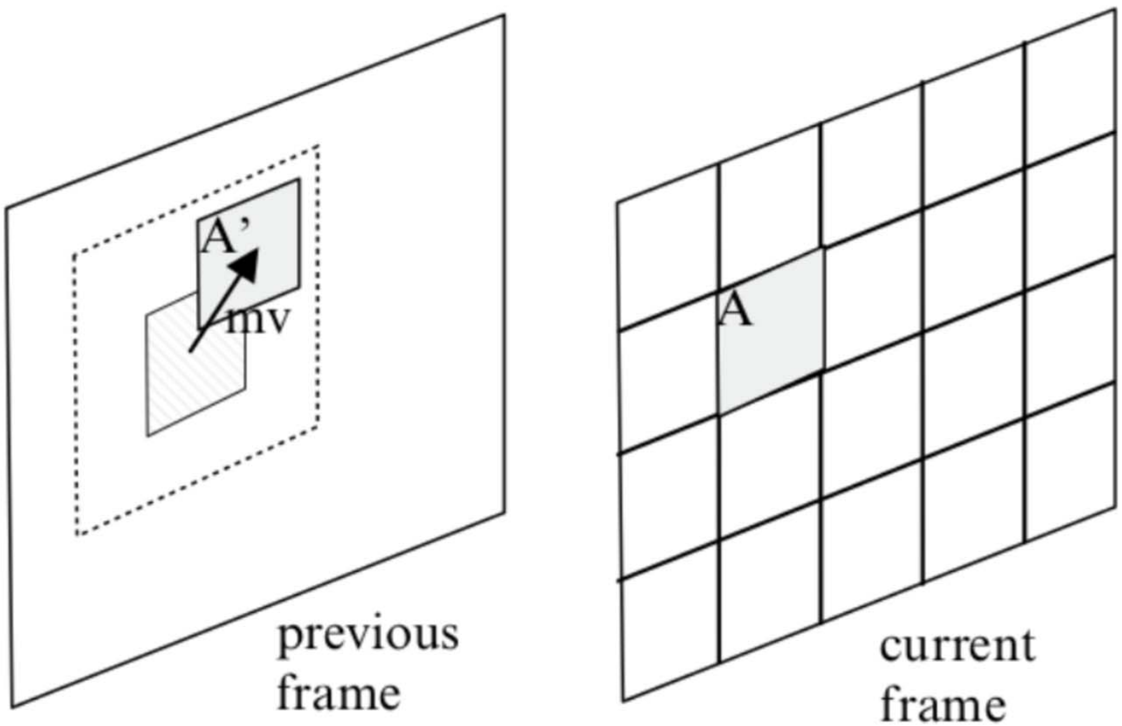
(a) With the aid of a diagram, explain the Group-of-Pictures (GOP) structure in the MPEG-1 standard. Explain briefly the effect of varying the size of GOP (parameter N), and discuss the dependency between different frame types during the encoding process.

p26 contents

(6 marks)

Motion Estimation & Compensation

Motion Estimation & Compensation



现在的指向过去的

Motion Estimation

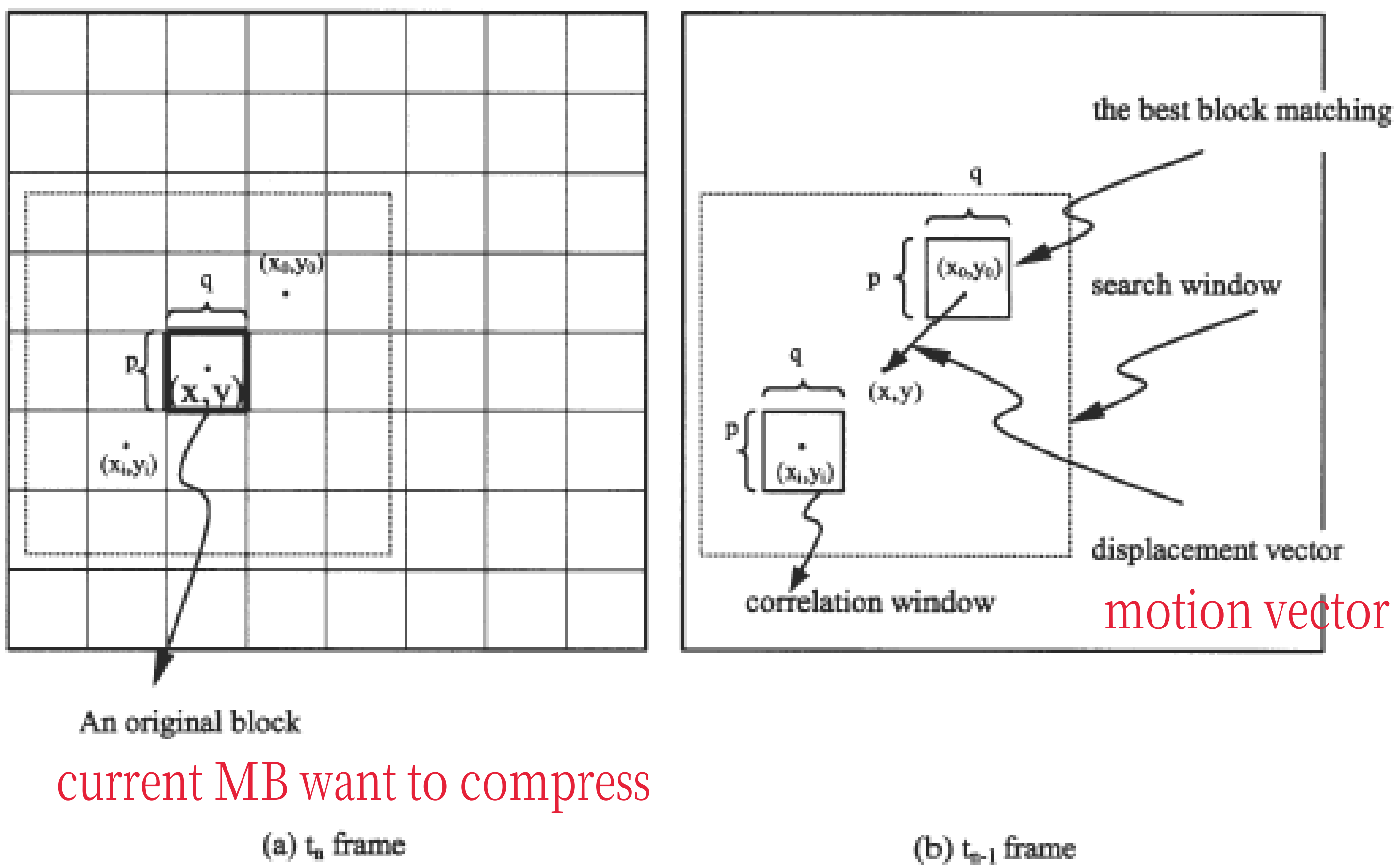


FIGURE 11.1 Block matching.

Figure Source:

Ze-Nian Li and Mark S. Drew, Fundamentals of Multimedia, Pearson Prentice Hall, 2004

Motion Estimation Methods

- Full Search
- Three-step Search
- 2D-Log Search
- Diamond Search
- Cross Search
- Etc…

Full Search

- Search every point within the search window.
- Pros: accurate, best compression ratio.
- Cons: slow, computationally intensive.

Three Step Search

whole is a search window

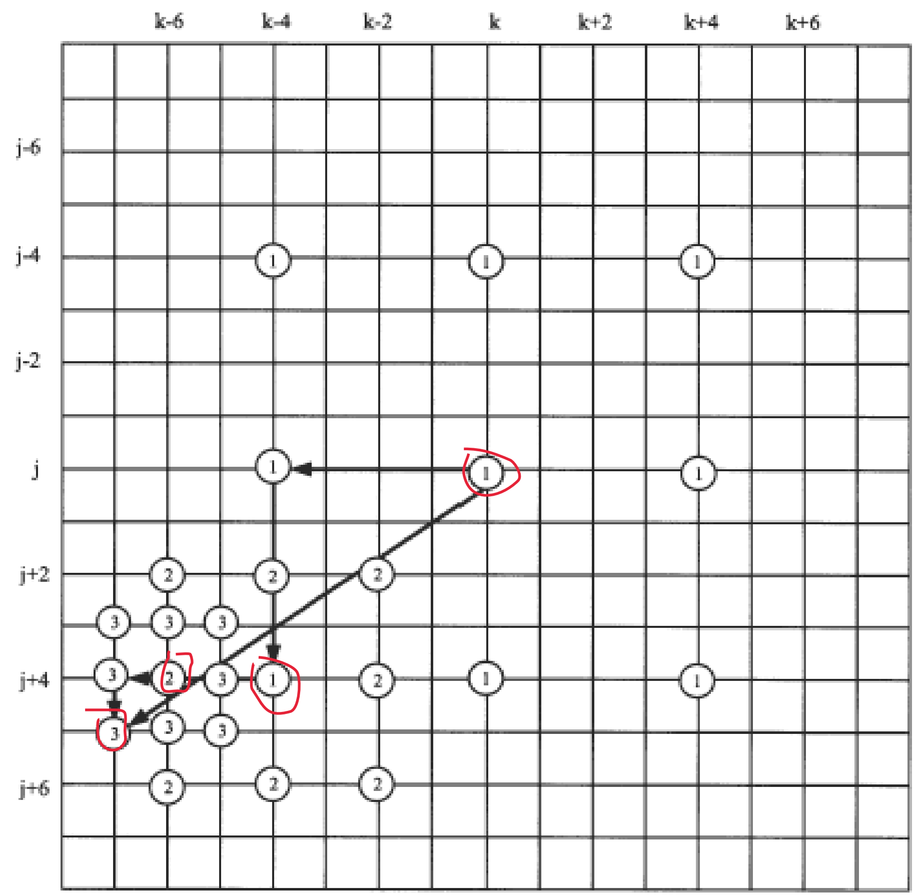


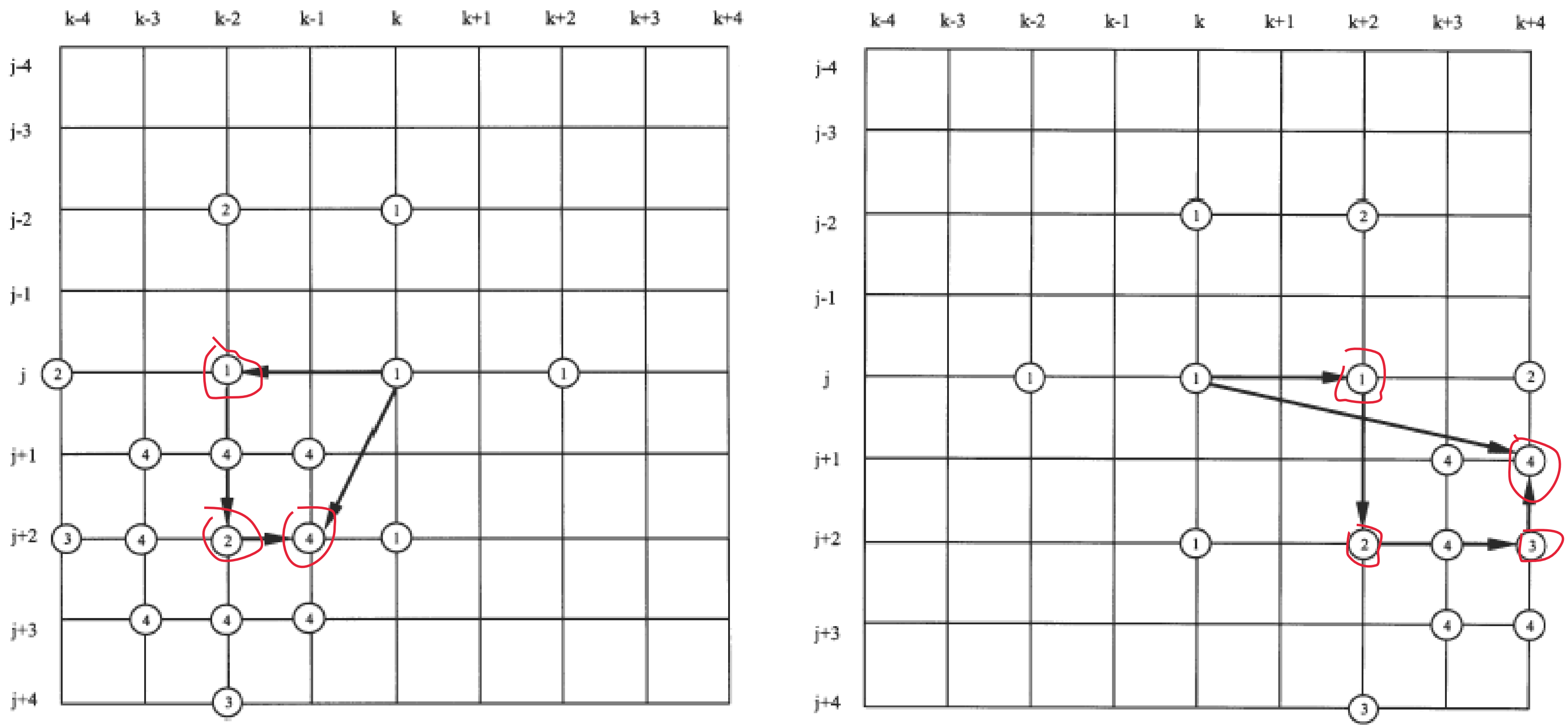
FIGURE 11.4 Three-step search procedure. Points (j+4, k-4), (j+4, k-6), and (j+5, k-7) give the minimum dissimilarity in steps 1, 2, and 3, respectively.

Figure source:
Yun Q. Shi and Huifang Sun, Image and video compression for multimedia engineering: fundamentals, algorithms, and standards, CRC Press, 2000

Three Step Search

- Consist of 3 steps only.
- Step 1: Search 9 points for minimum prediction error.
- Step 2: Centered on the found position in Step 1, halve the search width, search 9 points for minimum error.
- Step 3: Centered on the found position in Step 2, halve the search width, search 9 points for minimum error. The final position is the best position found.
- Pros: Fast
- Cons: less accurate, lower compression ratio.

2D Logarithmic Search



(b) (a)

FIGURE 11.3 (a) A 2-D logarithmic search procedure. Points at (j, k+2), (j+2, k+2), (j+2, k+4), and (j+1, k+4) are found to give the minimum dissimilarity in steps 1, 2, 3, and 4, respectively. (b) A 2-D logarithmic search procedure. Points at (j, k-2), (j+2, k-2), and (j+2, k-1) are found to give the minimum dissimilarity in steps 1, 2, 3, and 4, respectively.

2D Logarithmic Search

- Step 1: Search 5 points for minimum prediction error.
- Step 2: Centered on the found position in previous iteration,
  - (i) If the found position is the central position or located at the border, halve the search width; else, retain the search width.
  - (ii) (a) If the search window is not 3x3 yet, search 5 points for minimum error and repeat Step 2.

    (b) Else, if the search window is 3x3, search 9 points for minimum error and return best found position.
- Speed: Between full search and 3 step search.
- Accuracy: Between full search and 3 step search.

I, P & B Frame Encoding

I-Frame Encoding

treat like JPEG compression

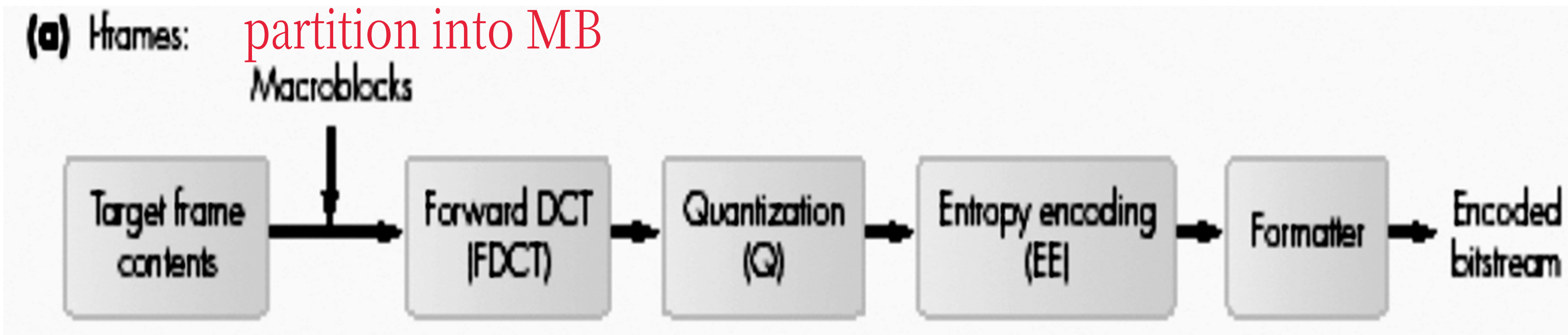


Figure source:
Fred Halsall, Multimedia communications: applications, networks, protocols and standards, Addison-Wesley, 2001

P-Frame Encoding Schematics

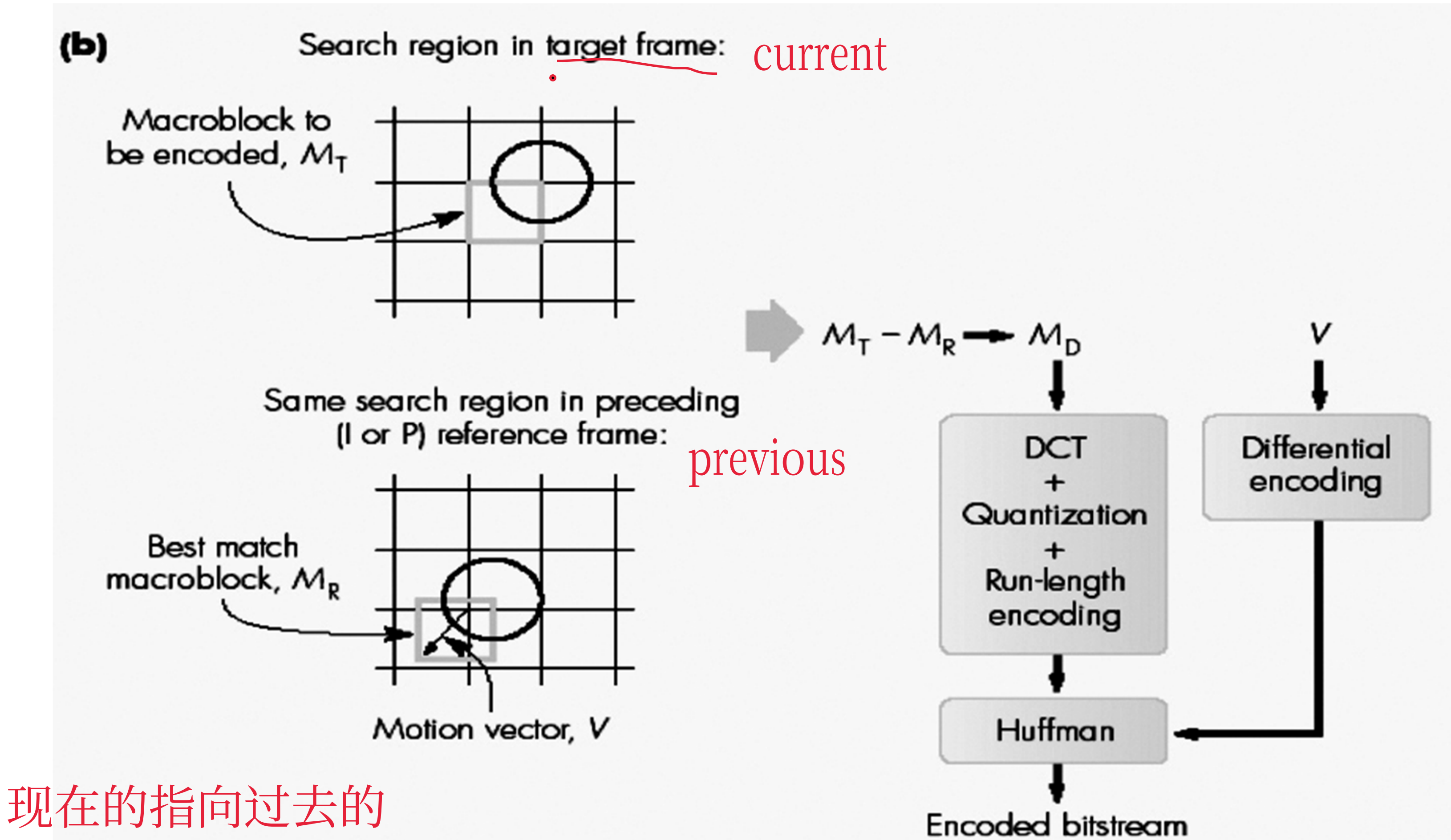


Figure source:
Fred Halsall, Multimedia communications: applications, networks, protocols and standards, Addison-Wesley, 2001

P-Frame Encoding Flowchart

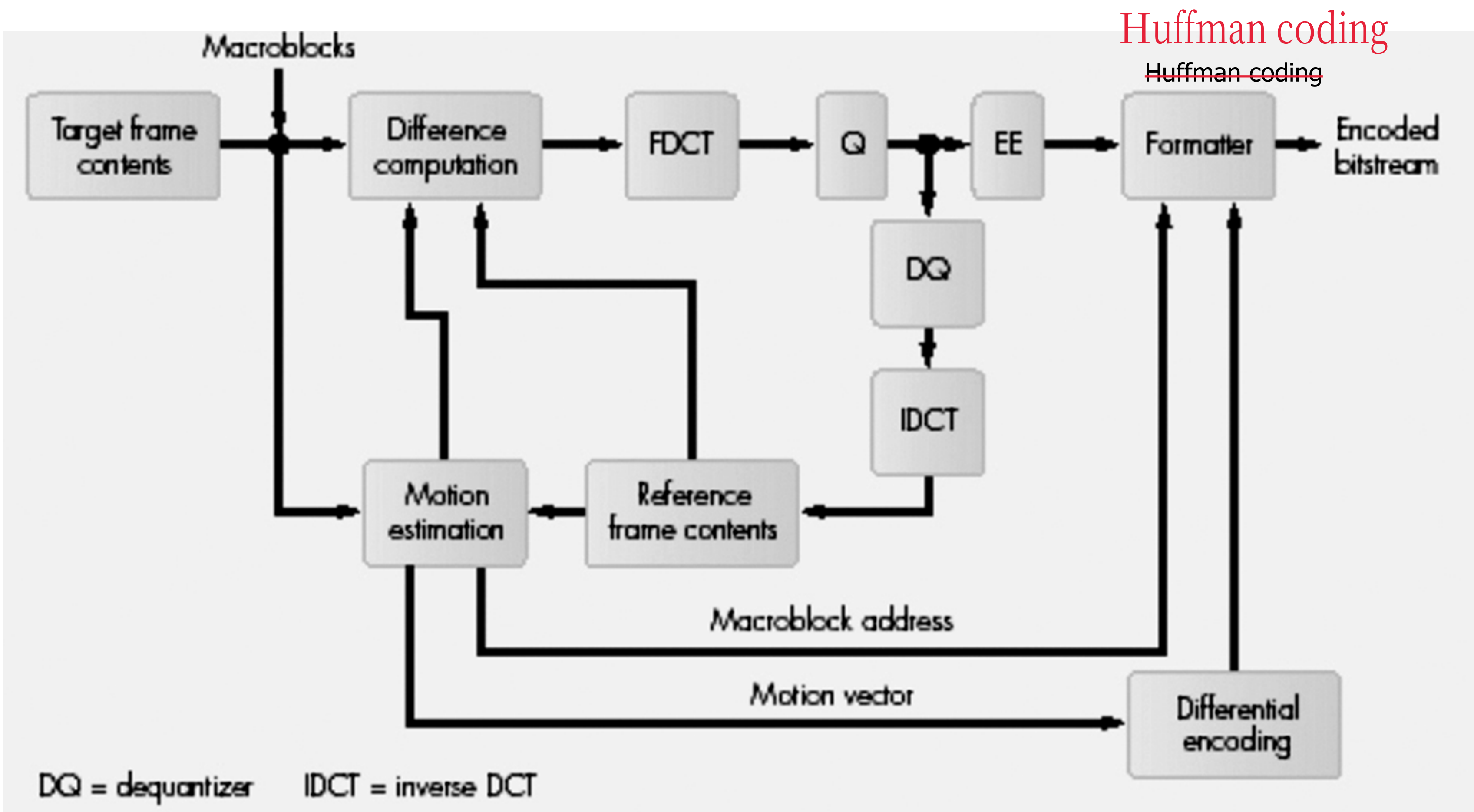


Figure source:
Fred Halsall, Multimedia communications: applications, networks, protocols and standards, Addison-Wesley, 2001

P-Frame Encoding (I)

- Perform motion estimation and compensation for each macroblock in the target frame.
- Each macroblock in the target frame is compared with the preceding I or P reference (anchor) frame.
- For each macroblock, a search is performed centered on a search window in the reference frame.
- A match is found if the Sum of Absolute Difference (SAD) is less than a threshold.
- If a match is found, two parameters are determined: motion vector and prediction error (matrix).

P-Frame Encoding (II)

- Motion vector measures the offset or displacement vector between the location of targeted macroblock with the location of the best-matched macroblock in the reference frame.
- Prediction error measures the difference (matrix) between the targeted macroblock and the best-matched macroblock in the reference frame.
- The motion vectors are encoded using differential encoding, and the resulting symbols are Huffman coded.
- The difference matrices for prediction error (Y, C_r, C_b) are encoded using the same steps as for I-frames.
- If no match is found, the macroblock is encoded independently as macroblocks in an I-frame.

B-Frame Encoding Schematics

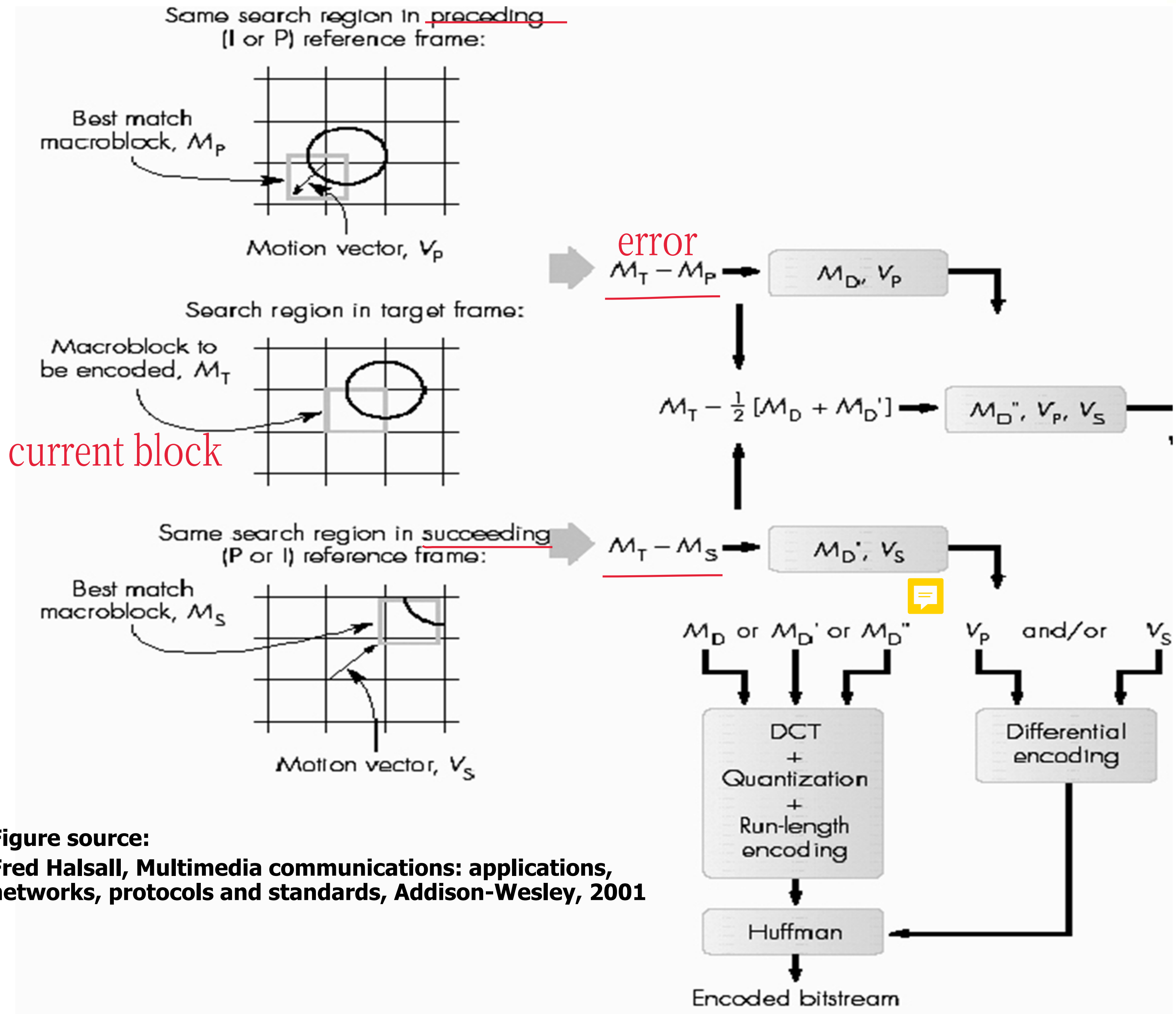


Figure source:
Fred Halsall, Multimedia communications: applications, networks, protocols and standards, Addison-Wesley, 2001

B-Frame Encoding Flowchart

choose smallest error

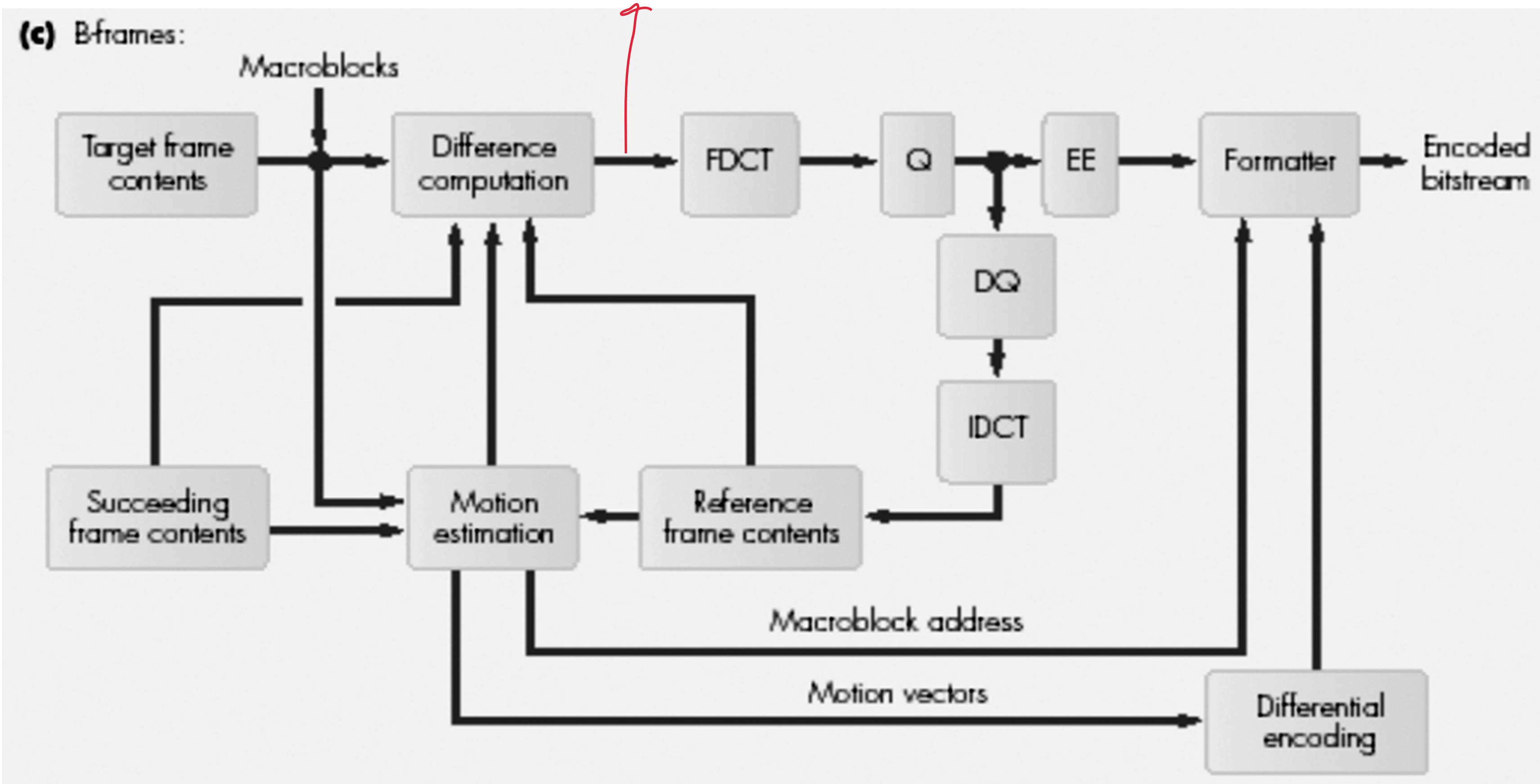


Figure source:
Fred Halsall, Multimedia communications: applications, networks, protocols and standards, Addison-Wesley, 2001

B-Frame Encoding

- For each macroblock in the target frame, motion estimation and compensation is performed with respect to both preceding (previous) and succeeding (following) reference (anchor) I- or P-frame.
- This produces 2 sets of (motion vector, difference matrix) pairs, one based on preceding anchor frame, the other based on succeeding anchor frame.
- A third difference matrix is computed using the target macroblock and the mean of preceding and succeeding best-matched macroblocks.
- The set with the smallest prediction error is chosen, and it is encoded in the same way as P-frame.

Implementation Issues in Encoding

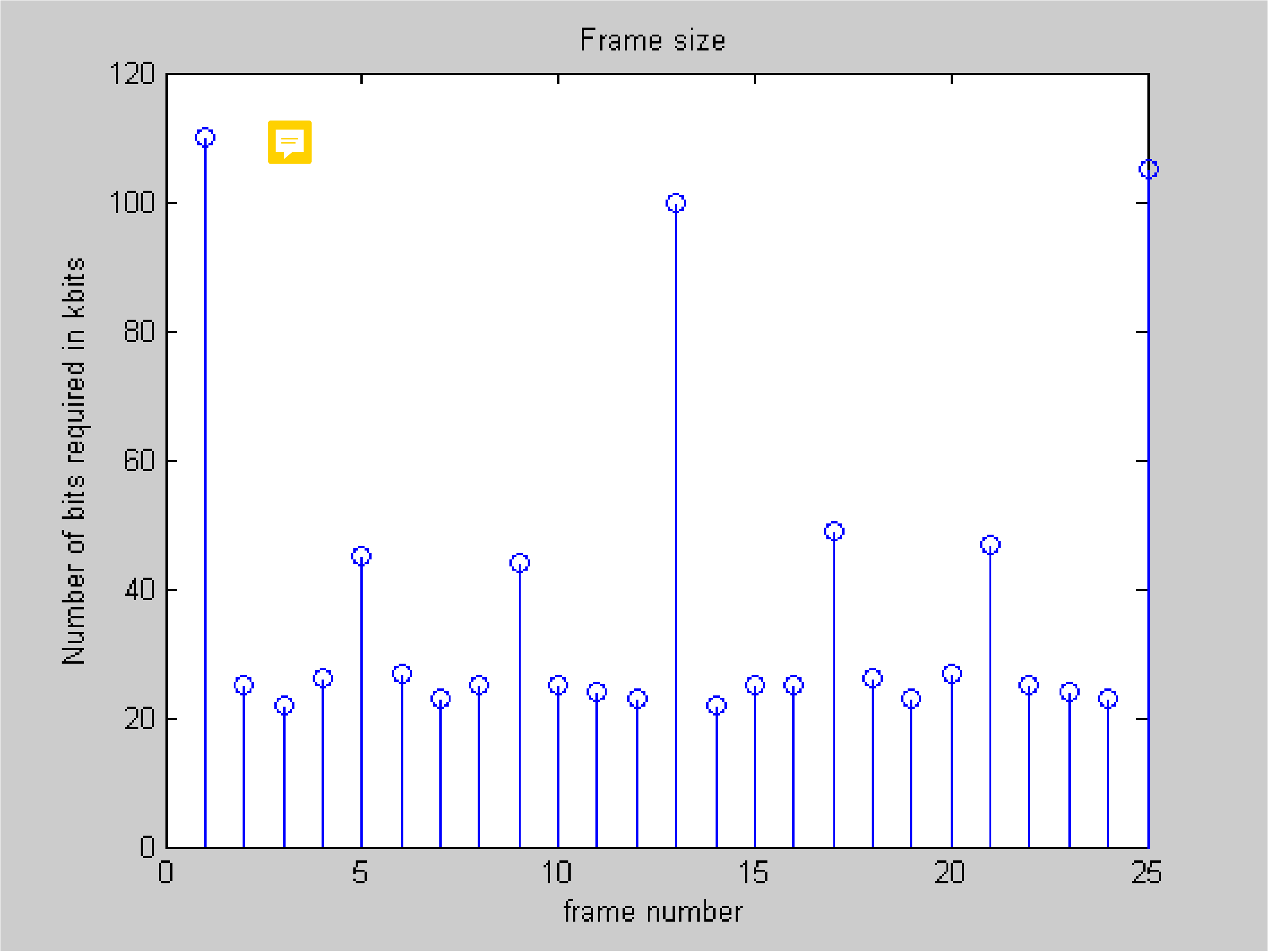
- Speed
- Complexity
- Compression ratio
- Software-based or hardware-based
- Application consideration

Typical Compression Ratios

- I-frames: 10:1-20:1
- P-frames: 20:1-30:1
- B-frames: 30:1-50:1

Frame Data Rate & GOP Structure

GOP structure is :IBBBPBBBPBBBI…



Exercise 3.2: MPEG Coding

(a) A digital video which is compressed using the MPEG-1 standard has the following display format:

Resolution of the luminance plane: 352× 240 pixels (8 bits/pixel)
Chroma subsampling format: 4:2:0
Frame rate: 30 frames/second

The number of bits required to encode the first 20 frames of the video using the MPEG-1 scheme is given in the following table:

Frame number	Number of bits required (kbits)
1	100
2	25
3	22
4	44
5	28
6	26
7	52
8	24
9	30
10	110

Frame number	Number of bits required (kbits)
11	22
12	25
13	44
14	28
15	26
16	48
17	23
18	27
19	90
20	22

Exercise 3.2: MPEG Coding

(i) Roughly sketch the line plot of the number of bits required to encode each frame against the frame number. Hence, determine the GOP structure that is used to compress this video. From the table, compute the average compression ratios of the I-frame, P-frame, and B-frame, respectively, for this video, if the frame types exist in the GOP structure.

(ii) If the compression ratios of the I-frame, P-frame, and B-frame of the video calculated in part 2(a)(i) are assumed to remain constant throughout the whole video, compute the effective compression ratio for this video. Calculate the storage requirement for a 120-minute video compressed using the scheme.

(12 marks)

Solution

MPEG-1

- MPEG-1 adopts CCIR601 digital TV format also known as SIF (Source Input Format)
- MPEG-1 supports only non-interlaced video.
- Normal picture resolution:
  - 352 x 240 for NTSC video at 30 fps
  - 352 x 288 for PAL video at 25 fps
  - uses 4:2:0 chroma subsampling
- MPEG-1 standard is also referred to as ISO/IEC 11172. It has five parts: 11172-1 Systems, 11172-2 Video, 11172-3 Audio, 11172-4 Conformance, and 11172-5 Software.

MPEG-2

- Key features of MPEG-2 as compared to MPEG-1:
  - Higher quality video at a bit-rate > 4 Mbps
  - Different profiles and levels
  - Support interlaced video
  - Support scalable coding
  - More resilient to bit-errors
  - Support 4:2:2 and 4:4:4 chroma subsampling
  - More flexible video formats

Profiles and Levels

Table 11.5: Profiles and Levels in MPEG-2

Level	Simple Profile	Main Profile	SNR Scalable Profile	Spatially Scalable Profile	High Profile	4:2:2 Profile	Multiview Profile
High		*			*		
High 1440		*		*	*		
Main	*	*	*		*	*	*
Low		*	*				

Table 11.6: Four Levels in the Main Profile of MPEG-2

Level	Max Resolution	Max fps	Max Pixels/sec	Max coded Data Rate (Mbps)	Application
High	$1,920 \times 1,152$	60	62.7×10^6	80	film production
High 1440	$1,440 \times 1,152$	60	47.0×10^6	60	consumer HDTV
Main	720×576	30	10.4×10^6	15	studio TV
Low	352×288	30	3.0×10^6	4	consumer tape equiv.

Figure Source:
Ze-Nian Li and Mark S. Drew, Fundamentals of Multimedia, Pearson Prentice Hall, 2004

Scalable Coding (I)

- Scalable coding is useful for MPEG-2 video transmitted over networks with following characteristics:
  - Networks with different bitrates.
  - Networks with variable bit rate (VBR) channels.
  - Networks with noisy connection.

Scalable Coding (II)

- Scalable coding consists of
  - a base layer and
  - one or more enhancement layers
- The base layer can be independently encoded, transmitted and decoded to obtain basic video quality.
- The encoding and decoding of the enhancement layer is dependent on the base layer or previous enhancement layer.
- Enhancement layer(s) will improve the video quality when decoded together with base layer.

Scalable Coding (III)

- MPEG-2 supports the following scalabilities:
  - SNR Scalability: enhancement layer provides higher Signal-to-Noise ratio (SNR) quality.
  - Spatial Scalability: enhancement layer provides higher spatial resolution.
  - Temporal Scalability: enhancement layer provides higher frame rate.
  - Hybrid Scalability: combination of any two of the above three scalabilities.
  - Data Partitioning: quantized DCT coefficients are split into partitions. Lower AC coefficients are used to encode base layer and higher AC coefficients are used to encode enhancement layer. relate to high freq.

Exercise 3.3: MPEG-2 Coding

(b) In a video streaming application, a user would like to stream an MPEG-2 video over a 5-Mbps channel, with as little buffering at the receiver as possible. The average compression ratios for the I-frame, P-frame, and B-frame are 10:1, 20:1, and 40:1, respectively. The MPEG-2 video sequence has the group of picture (GOP) structure of IBBBPBBBPBBB, and a frame rate of 30 frames per second (fps). The video can support the following resolutions and chroma subsampling formats:

can have 4 possible combination, give 4 kind of bit rate,which is less than 5Mbps

Resolution for luminance plane:
720×480 pixels (8 bits/pixel)
352×240 pixels (8 bits/pixel)

Chroma subsampling formats
4:2:2
4:2:0

Determine the most appropriate settings (resolution and chroma subsampling format) to encode the highest-quality video under these conditions. Your answer should include all proper justifications and calculations.

(8 marks)

Solution

Video Coding Standards

- *MPEG Standards*
  - *MPEG-1*
  - *MPEG-2 (H.262)*
  - *MPEG4 AVC/H.264*
  - *HEVC/H.265*

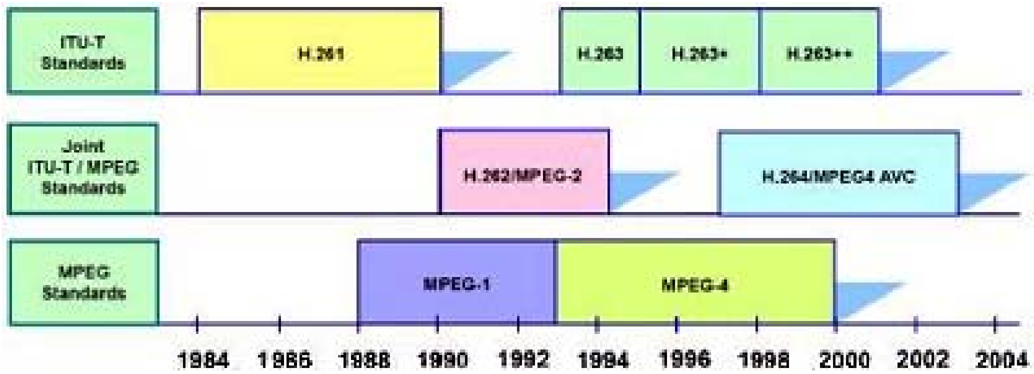
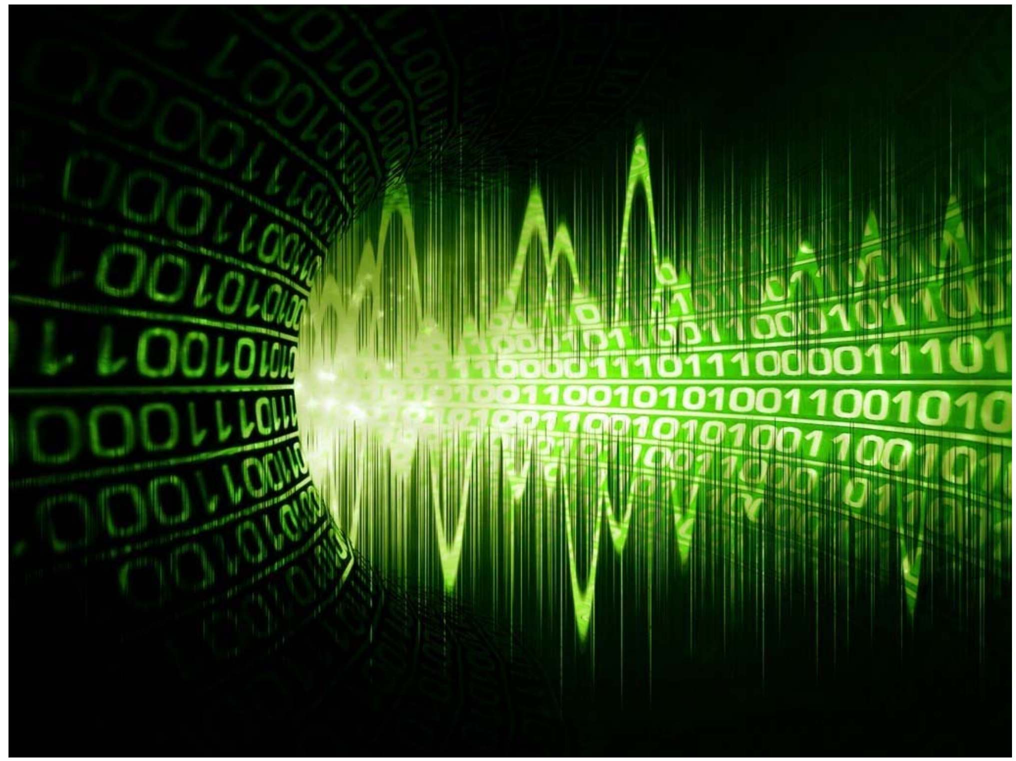


TABLE 15.2
Video/Image Coding Standards

Name	Completion Time	Major Features
JPEG	1992	For still image coding, DCT based
JPEG-2000	2000	For still image coding, DWT based
H.261	1990	For videoconferencing, 64Kbps to 1.92 Mbps
MPEG-1	1991	For CD-ROM, 1.5 Mbps
MPEG-2 (H.262)	1994	For DTV, 2 to 15 Mbps, most extensively used
H.263	1995	For very low bit rate coding, below 64 Kbps
H.263+ (version 2)	1998	Add new optional features to H.263
MPEG-4	1999	For multimedia, content-based coding
MPEG-4 (version 2)	2000	Adds more tools to MPEG-4
H.263++	2000	Adds more optional features to H.263+
H.26L	2000	Functionally different, much more efficient
MPEG-7	2001	Content description and indexing

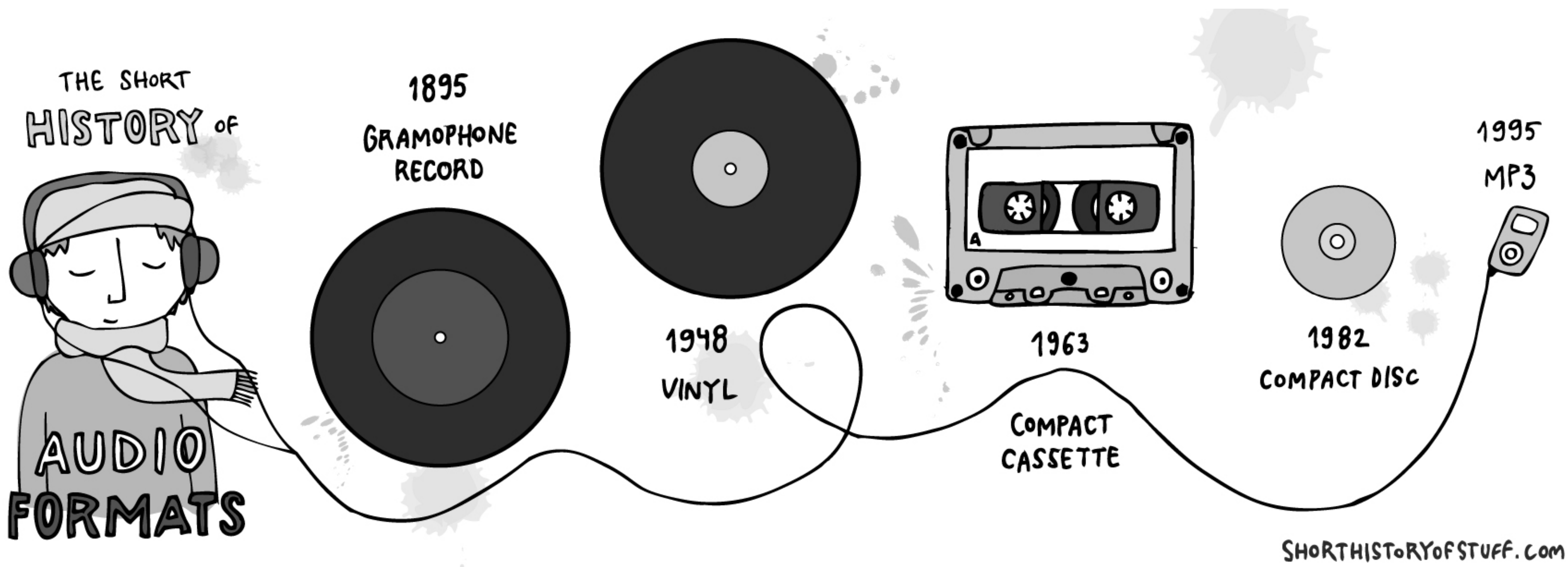
Compression Ratios

- JPEG: 10:1 for general images.
- MPEG-2: 30:1 for general videos.
- H.264/MPEG-4 AVC: 50:1 for general videos.



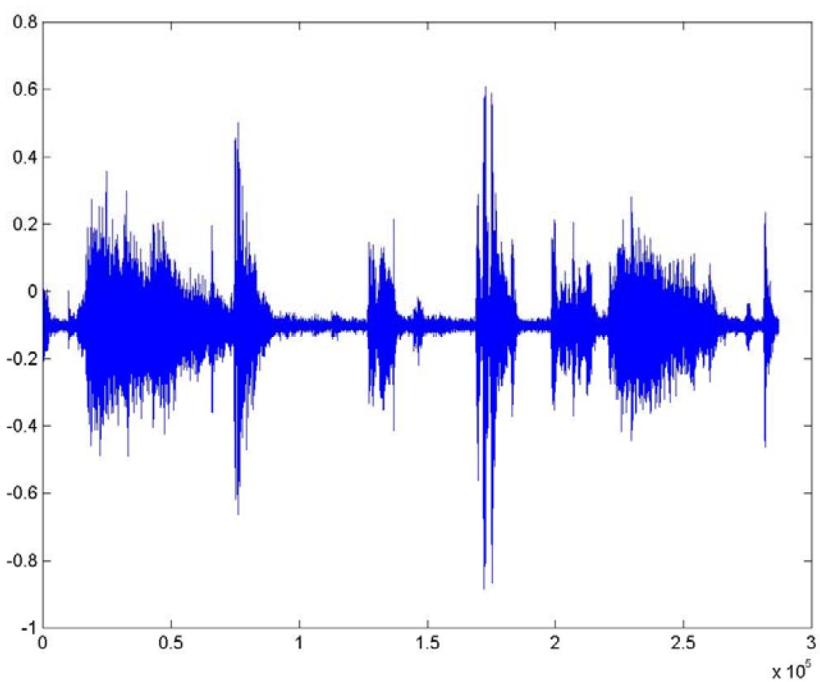
Digital Audio

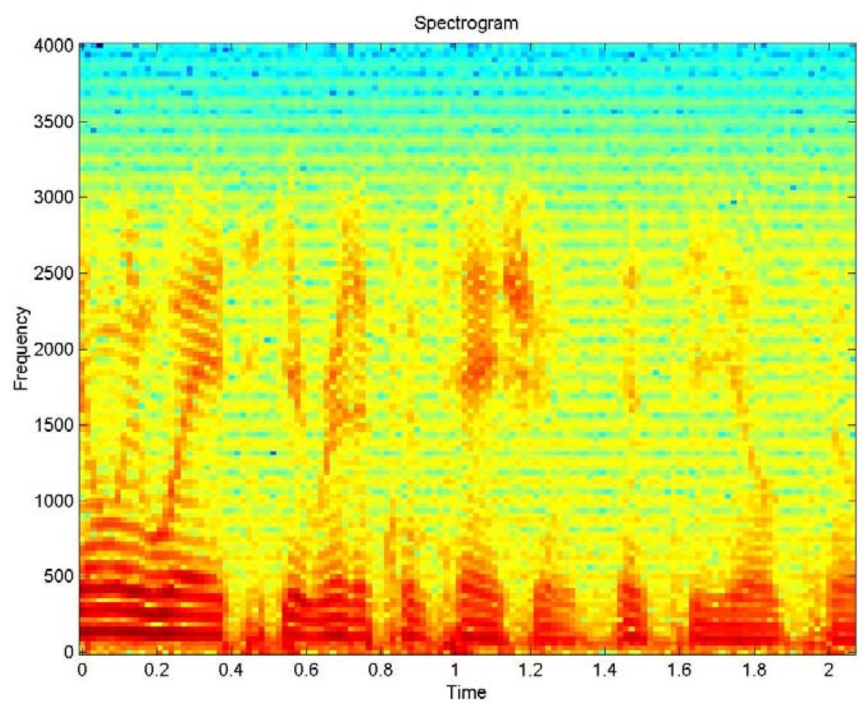
Short History of Audio



What is an Audio Signal?

- Audio signals are signals or sounds that are audible to humans.
- Frequencies in the range of 20 to 20,000 Hz.





Analog Audio

- Natural audio signals are analog.
- Disadvantages:
  - Data fidelity in reproduction
  - Durability

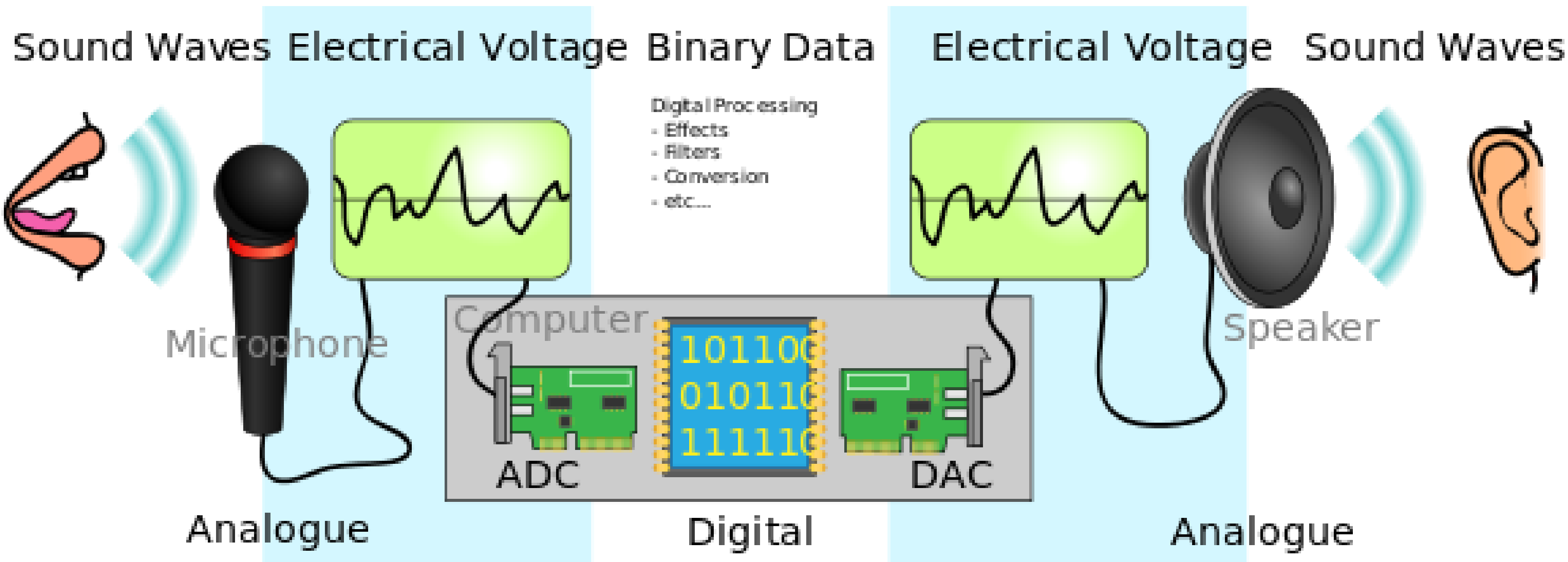
Advantages of Digital Audio over Analog Audio

- Wider dynamic range
- Increased resistance to noise
- Easy and exact reproduction
- Durability
- Easy to edit, transfer, distribute and store
- Provide error correction

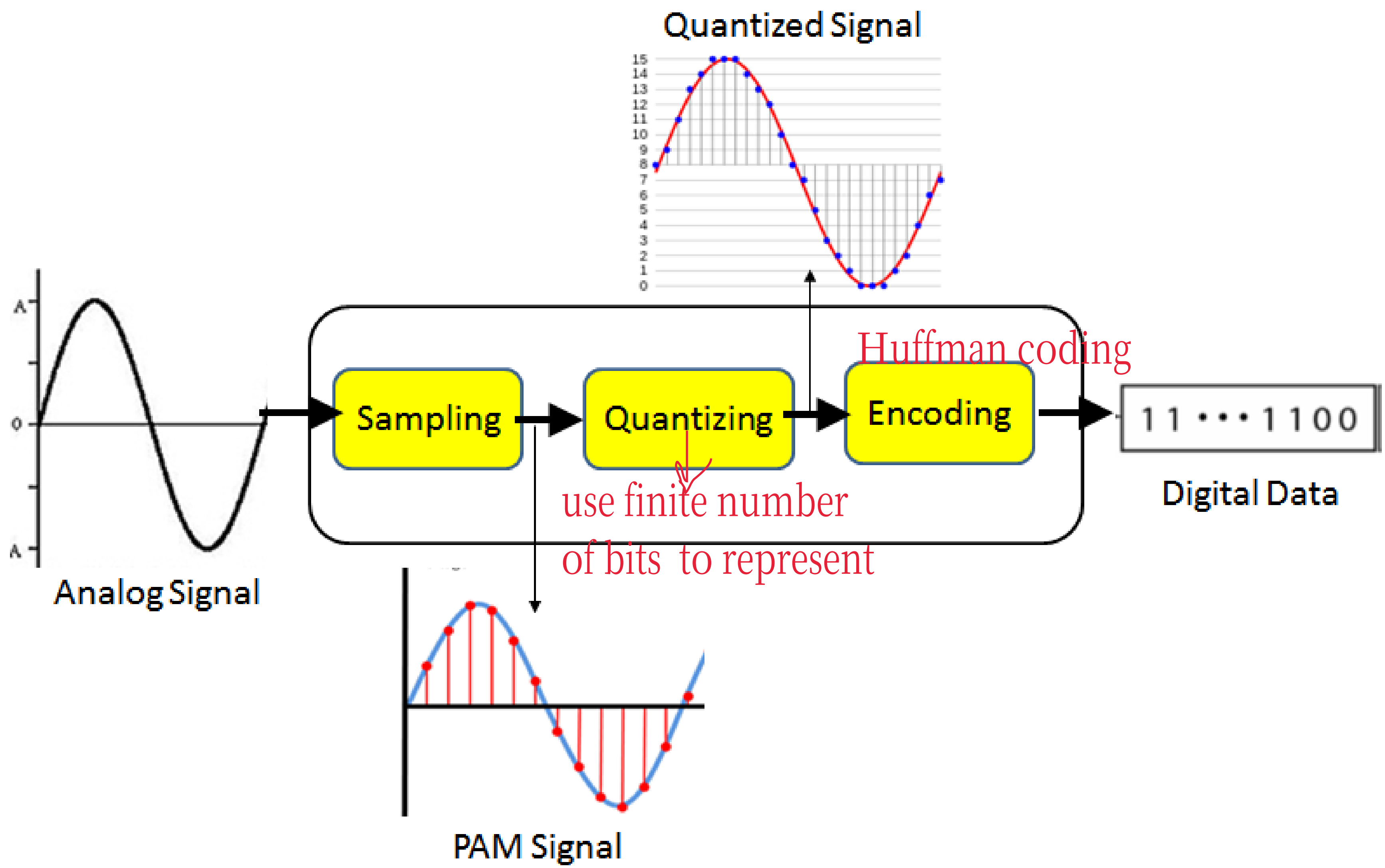
Digital Audio

- Digital audio aims to create audio signals with high fidelity and strong reproduction.
- Two types of audio signals:
  - Speech
  - Music
- Digital audio is produced by digitizing an analog signal generated by a sound source.

Digital Audio Conversion



Analog-to-Digital Conversion (ADC)



Audio Characteristics

- Typical bandwidth:
  - Speech (50Hz-10kHz), sampled at f>20KHz
  - Music (15Hz-20kHz), sampled at f>40kHz
- The number of bits per sample should be chosen so that the quantization noise generated at the sampling process is acceptable.
  - Speech: 12 bits/sample
  - Music: 16 bits/sample

expect higher quality

Audio Quality vs Data Rate

high quality, high data rate

Table 6.2: Data rate and bandwidth in sample audio applications

Quality	Sample Rate (KHz)	Bits per Sample	Mono/ Stereo	Data Rate (uncompressed) (kB/sec)	Frequency Band (KHz)
Telephone	8	8	Mono	8	0.200-3.4
AM Radio	11.025	8	Mono	11.0	0.1-5.5
FM Radio	22.05	16	Stereo	88.2	0.02-11
CD	44.1	16	Stereo	176.4	0.005-20
DAT	48	16	Stereo	192.0	0.005-20
DVD Audio	192 (max)	24 (max)	6 channels	1,200.0 (max)	0-96 (max)

Figure Source:
Ze-Nian Li and Mark S. Drew, Fundamentals of Multimedia, Pearson Prentice Hall, 2004

Nyquist Sampling Theorem

- Nyquist Sampling Theorem:

  A bandlimited continuous-time signal can be sampled and perfectly reconstructed from its samples if the waveform is sampled over twice as fast as its highest frequency component.
- If a signal is band-limited with maximum frequency f_{max}, then the sampling rate should be at least $2f_{max}$
- Nyquist rate = $2f_{max}$
- Aliasing will happen if the sampling rate is less than $2f_{max}$
- Antialiasing filter: low-pass filter that restricts the frequency content of the source so that it become band-limited.

No Aliasing vs Aliasing

time domain

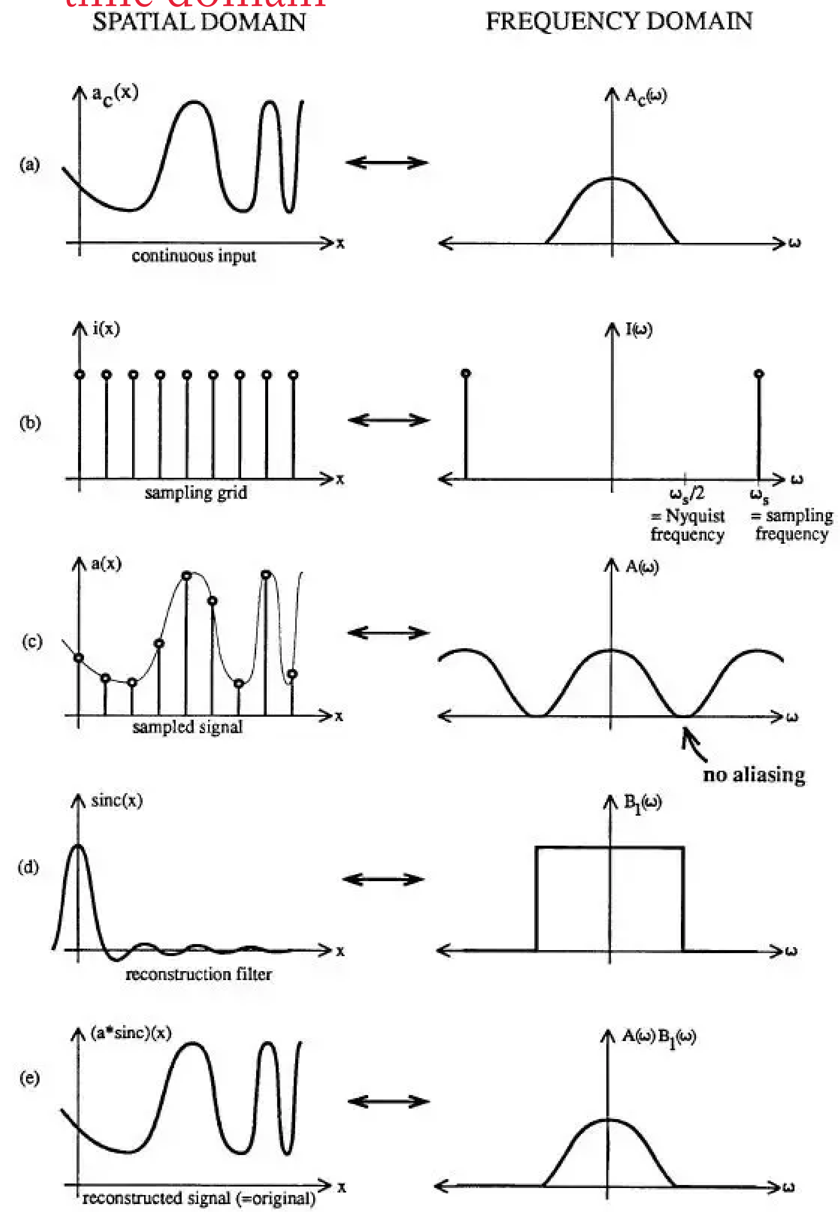


Figure 3.4: Sampling above the Nyquist rate (no aliasing).

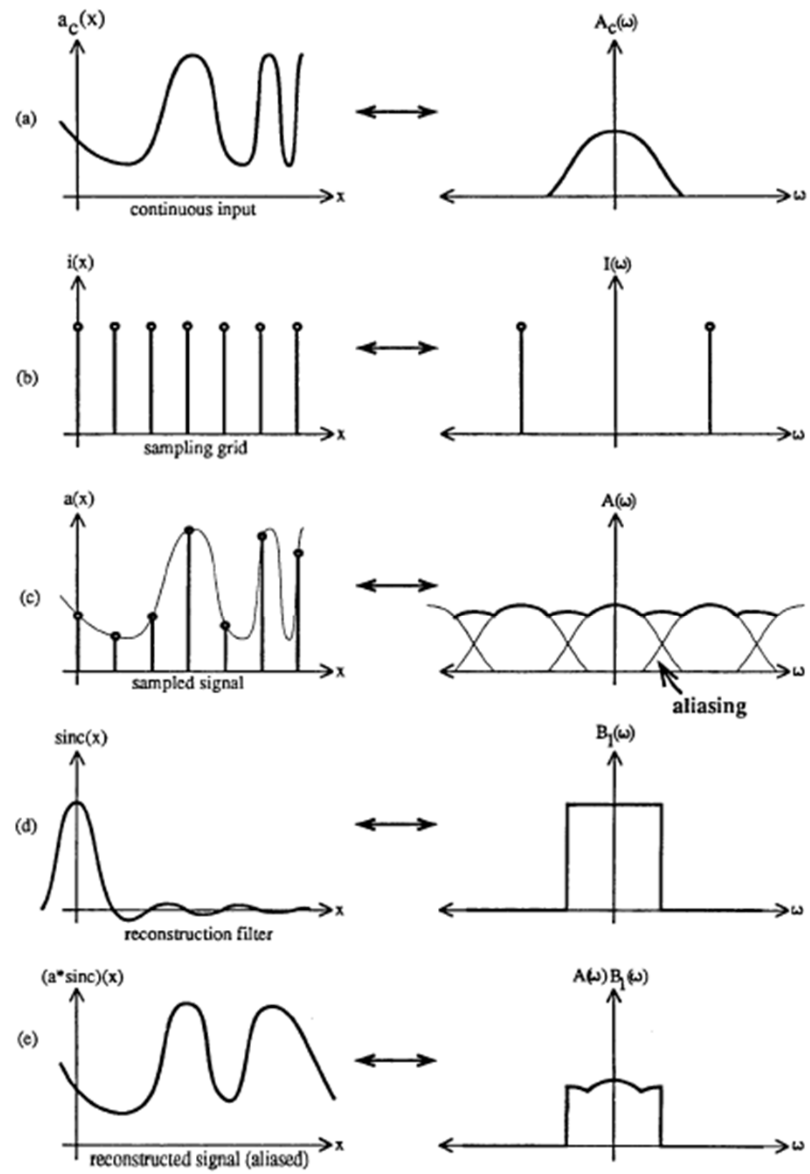


Figure 3.5: Sampling below the Nyquist rate (aliasing).

Pulse Code Modulation (PCM)

- Bandwidth of speech in public switched telephone network (PSTN): 200Hz-3.4kHz.
- Use a Nyquist sampling rate of 8kHz>6.8kHz.
- Each sample is quantized into 8 bits, giving rise to 64kbps.
- The digitization procedure is known as pulse code modulation (PCM) defined in ITU-T Recommendation G.711.

Types of Audio File Formats

- Uncompressed
  - E.g. WAV, AIFF, etc.
- Lossless compression
  - E.g. FLAC, ALAC (.m4a), MPEG-4 SLS, MPEG-4 ALS, MPEG-4 DST, Windows Media Audio Lossless (WMA Lossless), etc.
- Lossy compression
  - E.g. MP3, AAC, ATRAC and Windows Media Audio Lossy (WMA lossy), etc.

MP3

- MPEG-1 and/or MPEG-2 Audio Layer III, commonly called MP3, is an audio coding format for digital audio.
- Designed by the Moving Picture Experts Group (MPEG).
- Basic idea: use perceptual coding (psychoacoustic models) to approximate certain parts of sound that are beyond auditory resolution of most people.
- Lossy data compression
  - Typical compression ratio from 4:1 to 20:1

Audio Applications

Audio Applications

- Audio compression
- Speech recognition & natural language processing
- Virtual assistant - Siri, Google Assistant, Cortana
- Language translation
- Emotion recognition
- Music retrieval

Media Storage

Magnetic Data Storage

- Conventional magnetic data storage devices:
  - hard disks, magnetic tape
- Prone to demagnetization, wear and tear.

退磁场现象，质量受损

Optical Storage

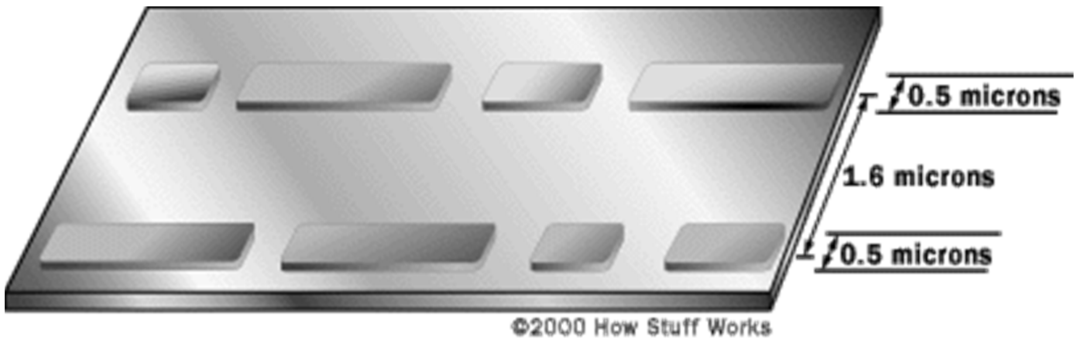
- Optical data storage:
  - Catalyst for the development of multimedia technology.
- Optical storage:
  - CD-DA, CDROM, VCD, DVD, Blu-ray Disc.
- Property:
  - Good storage density at lower cost.
  - Convenient.

History of Optical Storage

- 1982: Compact Disc Digital Audio (CD-DA)
- 1985: Compact Disc Read Only Memory (CD-ROM)
- 1991: Compact Disc Recordable (CD-R)
- 1995: Compact Disc Read Write (CD-RW)
- 1996: Digital Versatile Disc (DVD)
- 2006: High-definition DVD (HD-DVD), Blu-ray (BD)

Principle

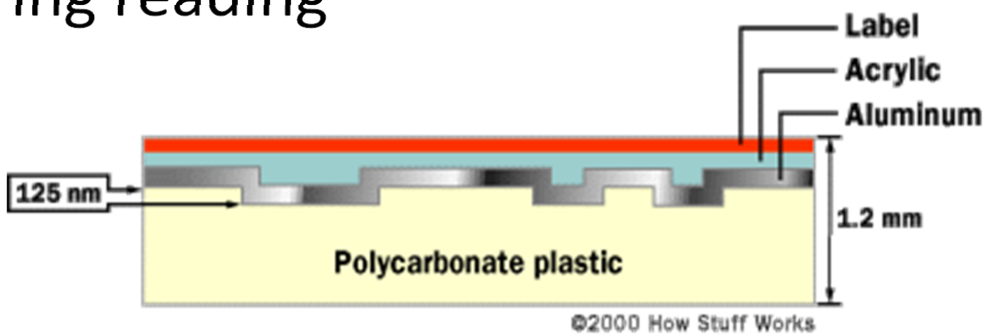
- An optical storage (e.g. CD) has a spiral track of data.

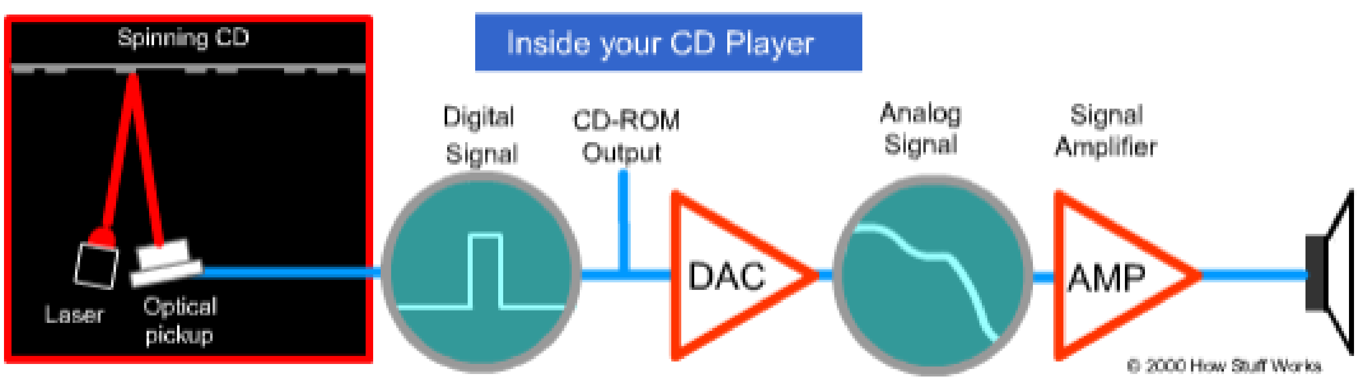


Source: electronics.howstuffworks.com

Principle

- Optical disc consists of a sequence of pits and lands within a track. The pits and lands represent data on the surface.
- Information is interpreted using the intensity of laser light reflected during reading





CDROM

- CD-ROM was conceived as a storage medium for general computer data, in addition to uncompressed audio data.
- Specified by Philips and Sony Corporation, and later accepted as ECMA standard

CD-ROM Mode 1

- Used to stored computer data
- Capacity: 650 Mbytes
- Data rate: 150 kBps

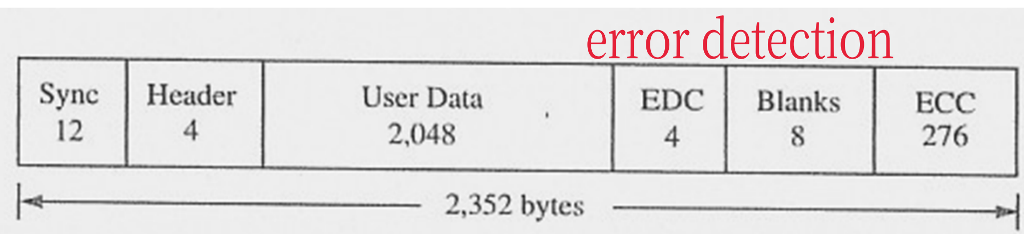


Figure 8-8 CD-ROM mode 1 block (sector) layout according to the *Yellow Book*.

Figure Source:

Ralf Steinmetz, Klara Nahrstedt, Multimedia Fundamentals, Volume 1: Media Coding and Content Processing, Prentice Hall, 2nd Edition, 2002.

CD-ROM Mode 2

- Used for storage of other media
- Capacity: 741 Mbytes
- Data rate: 175 kBps

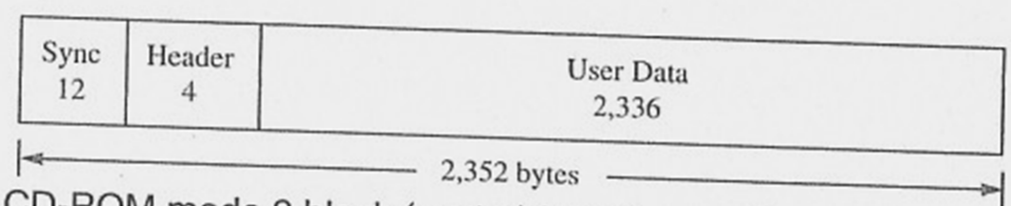


Figure 8-9 CD-ROM mode 2 block (sector) layout according to the *Yellow Book.*

Figure Source:
Ralf Steinmetz, Klara Nahrstedt, Multimedia Fundamentals, Volume 1: Media Coding and Content Processing, Prentice Hall, 2nd Edition, 2002.

What is VCD?

- Video CD (VCD) is a standard format for storing video on a Compact Disc.
- The first format for distributing films on standard 120 mm optical discs.
- The VCD standard was created in 1993 by Sony, Philips, Matsushita, and JVC and is referred to as the White Book standard.

Video CD Basics

- Codec: MPEG-1
- Resolution:
  - NTSC: 352x240
  - PAL: 352x288
- Frame rate:
  - NTSC: 30 fps
  - PAL: 25 fps
- Bitrate:
  - 1,150 kilobits per second
  - Constant bitrate

DVD

- In 1996, the DVD Consortium passed the first DVD standards.

Video	ITU-T H.262/ISO-IEC 13818-2 (MPEG-2 VIDEO) ISO/IEC 11172-2 (MPEG-1 VIDEO)
Audio	ISO/IEC 13818-3 (MPEG-2 AUDIO) ISO/IEC 11172-3 (MPEG-1 AUDIO) Dolby AC-3-Standard
System	ITU-T H.222 / ISO/IEC 13818-1 (MPEG-2 System) program/ only PES Stream (no Transport Stream)

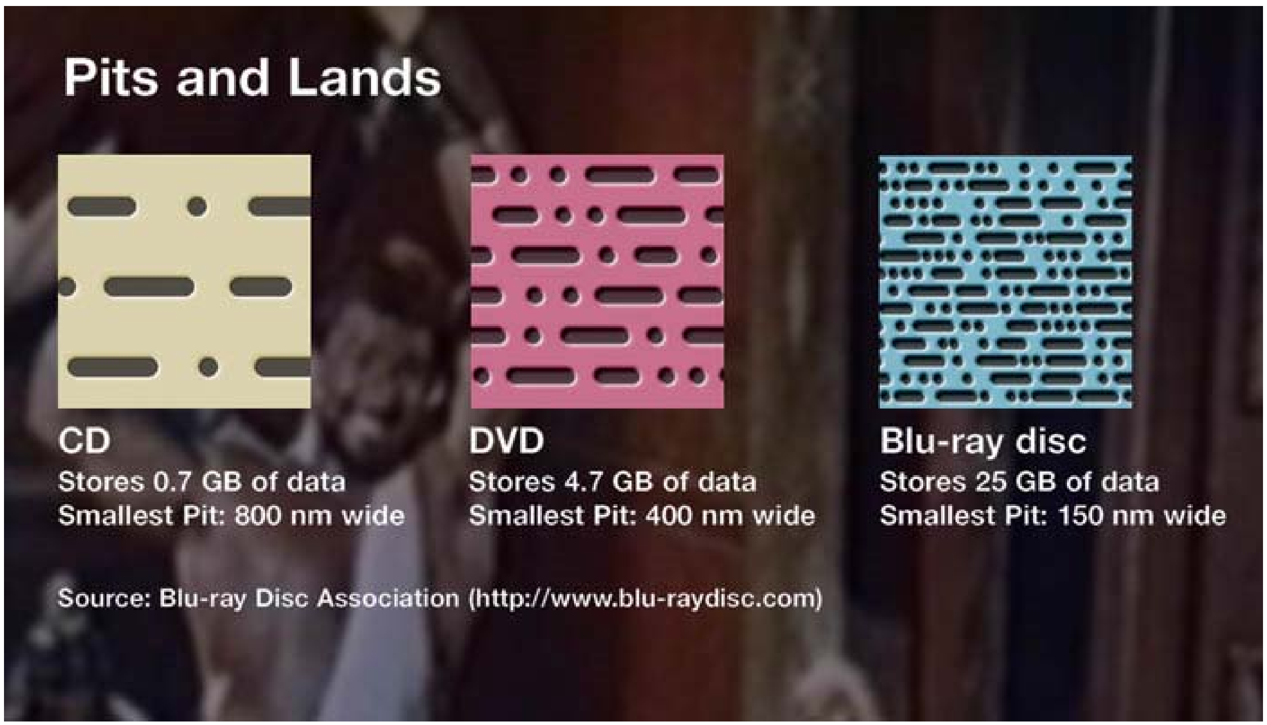
Table 8-7 DVD standards.

Figure Source:

Ralf Steinmetz, Klara Nahrstedt, Multimedia Fundamentals, Volume 1: Media Coding and Content Processing, Prentice Hall, 2nd Edition, 2002.

Why DVD Can Store More than CD?

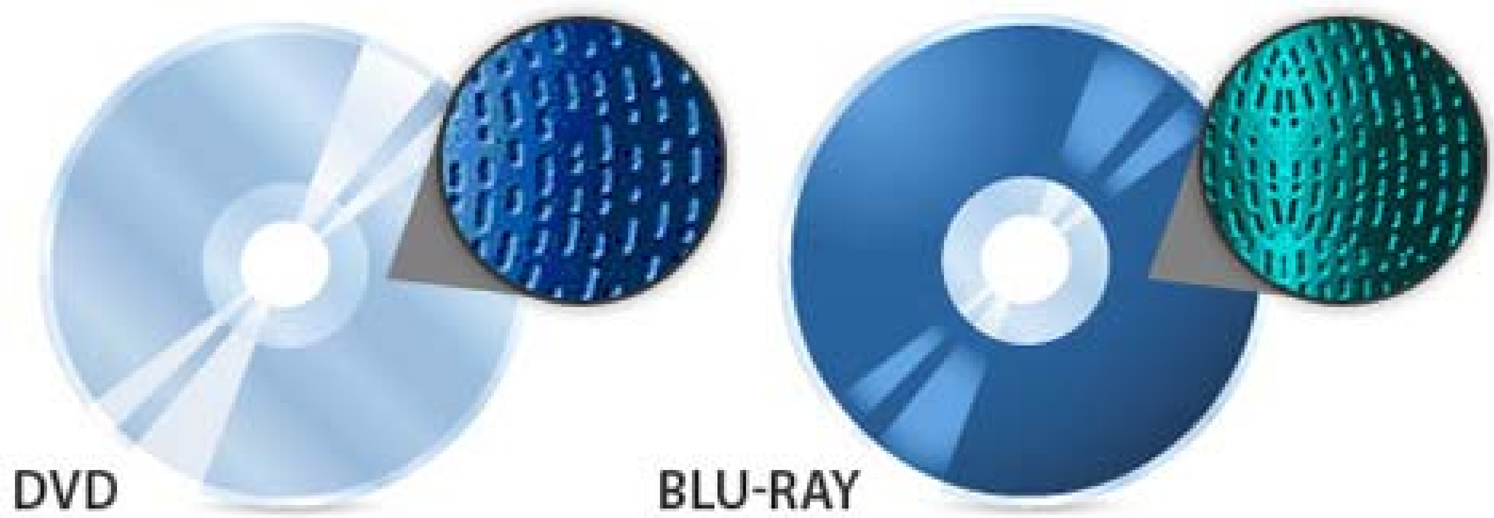
- DVD achieves higher data storage capacity than CD-ROMs by using smaller pits (which yields higher track density).



DVD-Video Basics

- DVD-Video is a consumer video format used to store digital video on DVD discs.
- Codec: MPEG-2
- Bitrate:
  - Up to 9.8 Mbps
- Frame rate:
  - PAL: 25 fps, NTSC: 30 fps
- Resolutions supported:
  - For 25 fps:
    - 720 × 576, 704 × 576, 352 × 576, 352 x 288 pixels
  - For 30 fps:
    - 720 × 480, 704 × 480, 352 × 480, 352 x 240 pixels

Why Blu-ray Disc?



Blu-ray Disc

- Media type: high-density optical disc.
- Encoding: MPEG-2, MPEG-4 AVC(H.264), and VC-1.
- Capacity: 25 GB (single layer), 50 GB (dual layer).
- Developed by: Blu-ray Disc Association.
- Usage: Data storage, high-definition video and games.

CD, DVD, Blu-ray Comparison

	CD	**DVD**	**Blu-Ray Disc**
Typical storage capacity (12cm disc) Single-sided.	800 MB	4.7GB SL	27GB SL
	NA	8.5GB DL	54GB DL
Track pitch	1.6 um	0.74 um	0.32 um
Minimum pit length (μm)	0.5 um	0.4 um	0.15 um
Wavelength of laser	780nm Red	650nm Red	405nm Blue
Data transfer rate/speed	153.6 kBps	1350 kBps	4.5 MBps
Thickness of Disc	1.2mm	1.2mm	1.1mm
Thickness of surface layer	1.2mm	0.6mm	0.1mm
Codec	MPEG1	MPEG2	MPEG2/H.264

Flash Memory

- Flash memory is a type of electrically erasable programmable read-only memory (EEPROM) chip.
- Applications of removable flash memory cards:
  - Memory card
  - Thumb drive
- Advantages of flash memory over hard disk:
  - Quiet
  - Smaller in size
  - Lighter
  - No moving parts
  - Flexibility in data transfer

Media Streaming

- Streaming media is media that is continuously received by and presented to end-user while being delivered by a provider.
- A client end-user uses media player to play the data file (e.g. a movie or song) before the entire file has been transmitted.
- Live streaming refers to Internet content delivered in real-time as events happen (e.g. Olympics).
- Become popular due to:
  - Faster networks
  - Better compression technologies
  - Consumer acceptance

Video Streaming Services

Media Contents …

TV Display

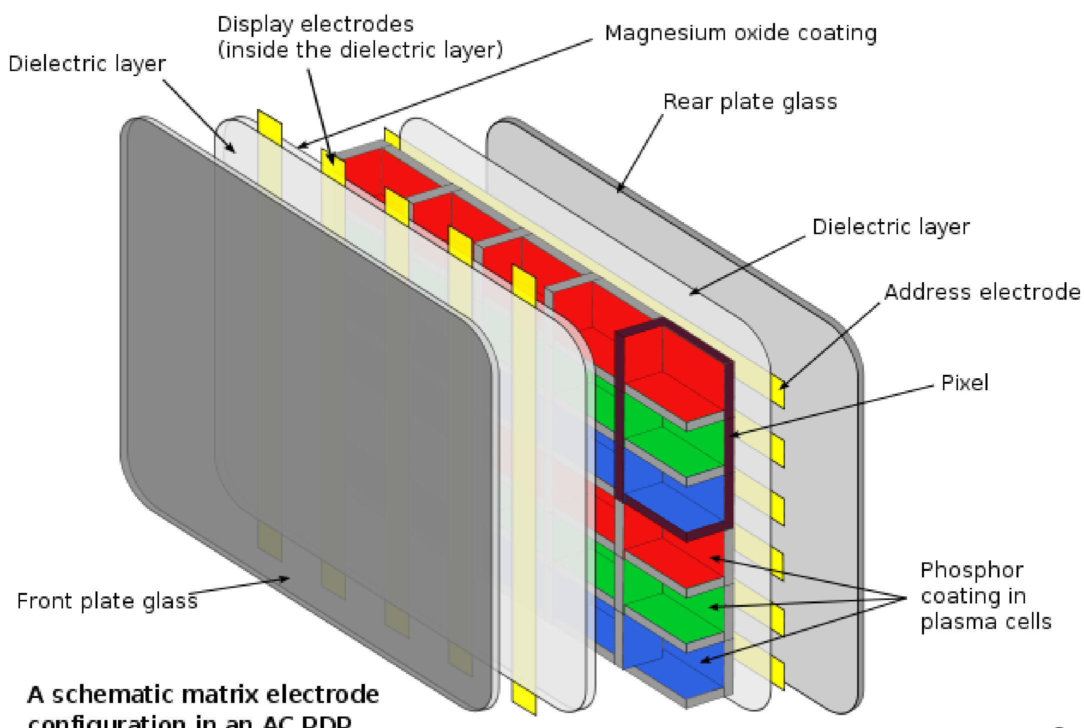
Display Technologies

- Flat-Panel TV
  - Plasma
  - LCD
  - LED
  - OLED

Plasma TV

Principle & Structure

- The basic idea is each subpixel is a microscopic fluorescent lamp that emits one primary color - red, green, or blue. By varying the intensity of the light from these three subpixels, a multitude of tints can be displayed. Certain gases (e.g. neon and xenon gas) become plasma and emit colored light when subject to electric voltage.



Source: wikipedia

Plasma Formation

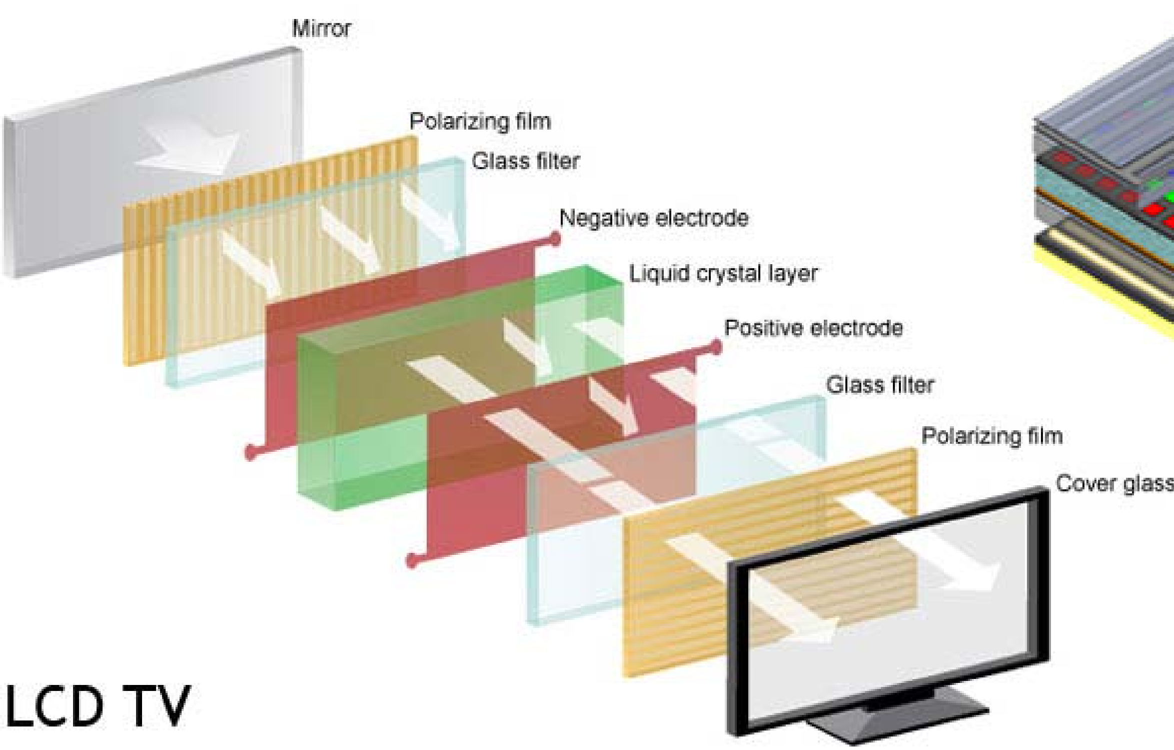
- When voltage is applied, electrons are torn off from the gas to form plasma.
- When voltage is stopped, electrons will combine with plasma to form gas again. Ultraviolet photons are released as a result.
- Ultraviolet radiation excites the phosphor coating inside the cell to give off visible light.

How Plasma TV Works?

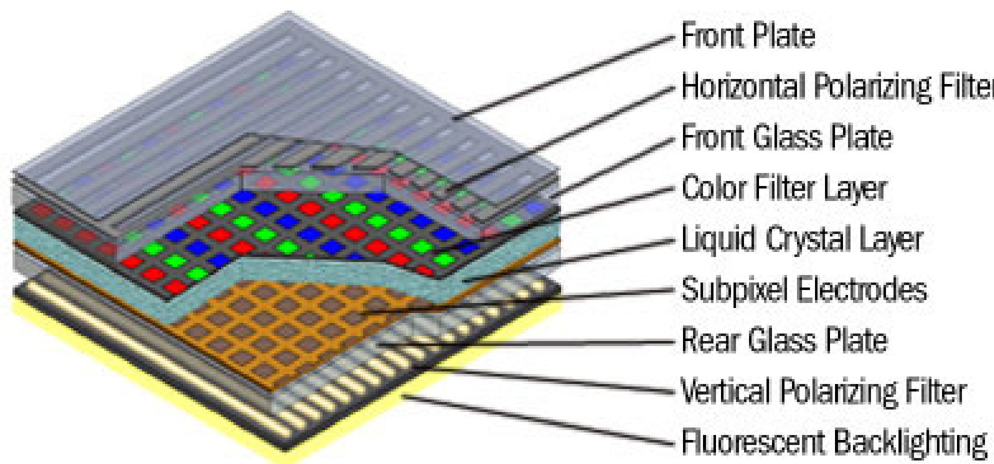
Liquid Crystal Display (LCD) TV

Principle & Structure

- Liquid crystal display (LCD) TV operates when light from behind the screen is shone through a succession of polarizing filters, electrodes and matrix of tiny liquid crystal cells. Different voltages at each subpixel control different amount of colour to pass through.

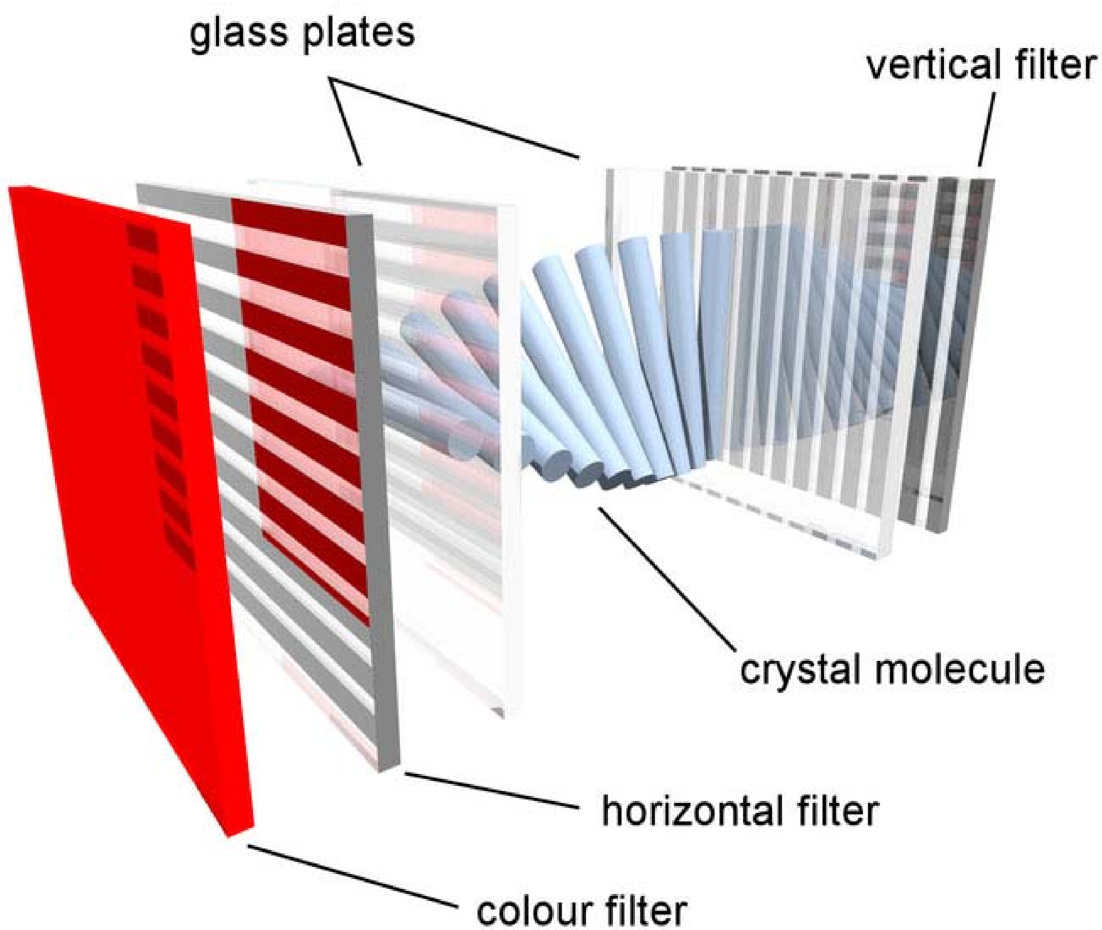


LCD TV



Source: which.co.uk

How LCD TV Works?



Source: clearvisionsystems.co.uk

How LCD TV Works?

- Two pieces of polarized glass with polarizing film at 90°.
- Light is polarized by the first filter.
- As the light passes through the liquid crystal layer, the crystal's molecules “bend” the light, or change the light's polarization plane. The alignment of liquid crystal can be adjusted by the electrode voltage.
- The color filter layer consists of numerous subpixels (red, green, blue). Each subpixel can be controlled independently by electrode voltage to yield different colors for each pixel.
- The polarized light will then pass through the second polarized filter.

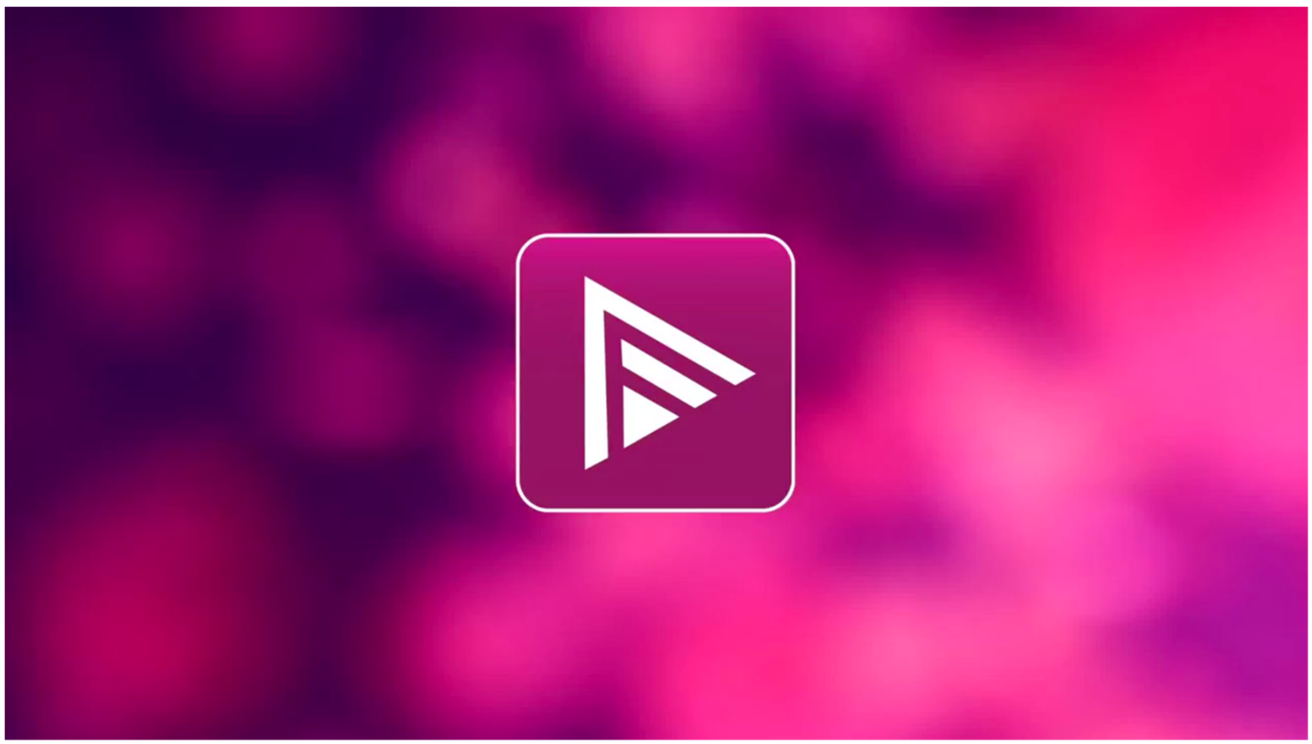
How LCD TV Works?

LED-Backlight TV

- Can be considered to be a better modification of LCD TV.
- Difference: use LED array to light up the LCD display (LCD TV uses fluorescent backlight to light up the LCD screen).

Organic Light-Emitting Diode (OLED) TV

- OLED is a LED in which the emissive electroluminescent layer is a film of organic compounds which emit light in response to an electric current.



Summary

- This section covers the following:
  - Video compression basics
  - Motion estimation and compensation
  - MPEG standard: MPEG-1, MPEG-2
  - Video standards
  - Digital audio
  - Media storage
  - Media display