Unit 4
Digital Video

CS 3570

Shang-Hong Lai

Video

- Video is made up of a series of still images (*frames*) played one after another at high speed
- This fools the eye into believing that it is observing a continuous stream

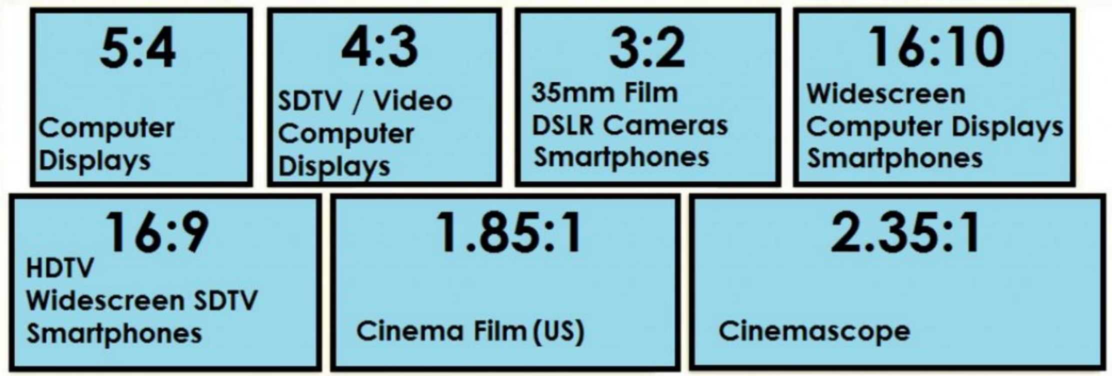
  - Video (real-world pictures)
  - Animation (Computer generated)

Video and Film

- ***Movie*** : a story told with moving images and sound
- The word ***film*** seems to imply a movie that is shot and/or stored on cellulose film.
- Film and video both rest on the same phenomenon of human perception, called ***persistence of vision*** – the tendency of human vision to continue to “see” something for a short time after it is gone.
- A related physiological phenomenon is ***flicker fusion***—the human visual system’s ability to fuse successive images into one fluid moving image.

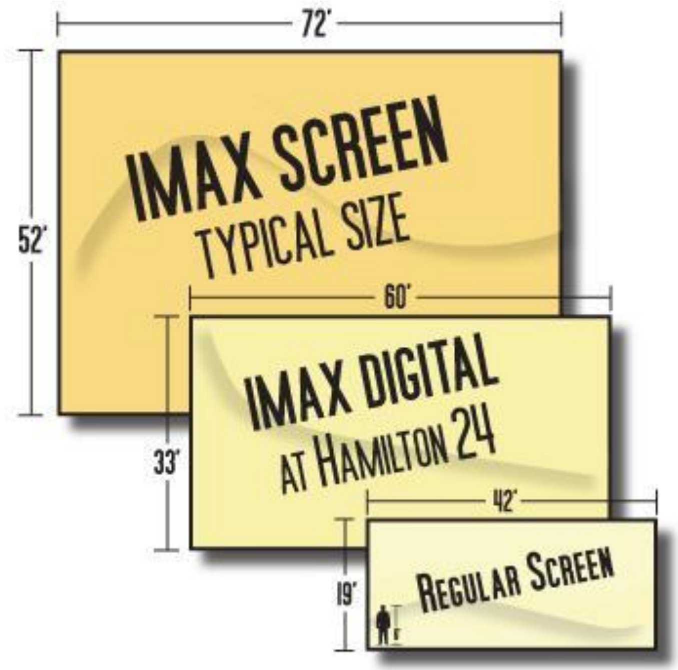
Standard film aspect ratios

- ***Aspect ratio*** is the ratio of the width to the height of a frame, expressed as *width : height*.
- Two common video graphic aspect ratios are **4 : 3**(1.33 : 1), the universal video format of the 20th century, and **16 : 9** (1.77 : 1), universal for high-definition television
- The most common aspect ratios used today in the presentation of films in cinemas are **1.85 : 1** and **2.35 : 1**



Film Development

- Silent movies and early sound movies were shot mostly on 16 mm film, introduced by Eastman Kodak in 1923
- **IMAX** movies are shot on 70 mm film with aspect ratio of 1.43:1.
- **IMAX** movies are displayed on very large screens, e.g. a standard IMAX screen is 22 m × 16.1 m (72 ft × 52 ft).



High-definition television

- In 1981, NHK began broadcasting what came to be known as ***high-definition television, HDTV.***
- The current definition of HDTV is television that has an aspect ratio of 16: 9, surround sound, and one of three resolutions: 1920 × 1080 using interlaced scanning(**1080i**), 1920 × 1080 using progressive scanning(**1080p**), or 1280 × 720 using progressive scanning (**720p**).
- Digital encoding is not part of this definition, and, historically, HDTV was not always digital.

Video and film displays

- Like film, video is created by a sequence of discrete images, called **frames**, shown in quick succession.
- Film is displayed at 24 frames/s. The standard frame rate for NTSC video is about 30 frames/s. The frame rate for PAL and SECAM video is 25 frames/s.
- A film frame is a continuous image. Video frames, in contrast, are divided into lines. Television has to be transmitted as a signal, line-by-line.
- Video is displayed (and recorded) by a process called ***raster scanning.*** The raster refers to a single frame.

Raster scanning

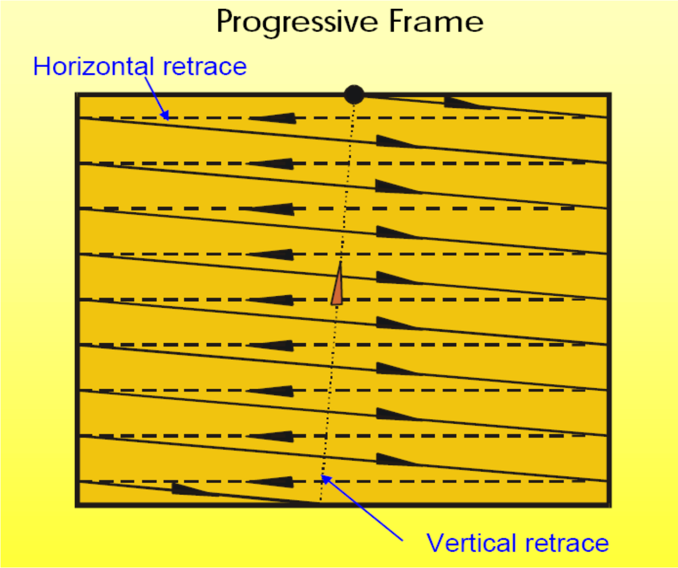
- The scanning process is a movement from left to right and top to bottom.
- When the scanner has finished with one line, it moves back to the left to start another in a motion called ***horizontal retrace***.
- ***Vertical retrace*** takes the scanner from the bottom of the monitor to the top again.
- In the case of video camera, the purpose of the scanning is to record the data that will be saved and/or transmitted as the video signal.

Raster scanning

- For many years, the dominant video display technology was the ***cathode ray tube*** (***CRT***). Most television sets were built from CRTs, as were the computer monitors.
- Scanning can be done by one of two methods: either **interlaced** or **progressive scanning.**
- Video standards are sometimes described in terms of ***field rate*** rather than frame rate.
  - For PAL analog video, 50 fields/s = 25 frames/s

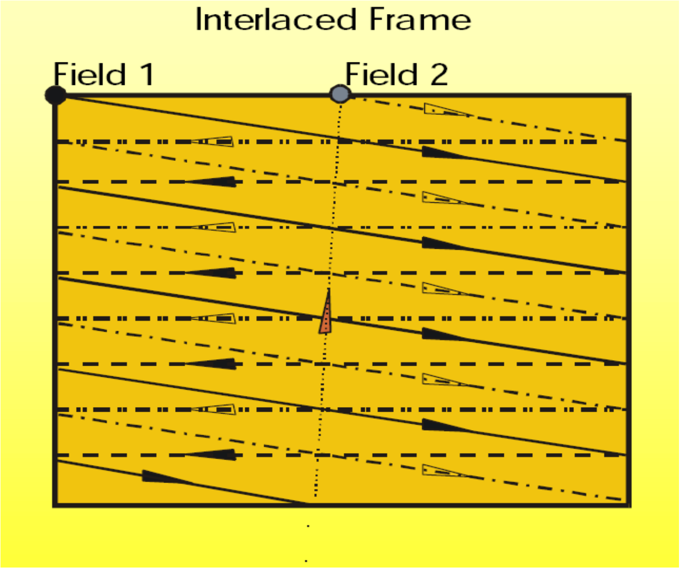
Progressive scanning

- In ***progressive scanning***, each frame is scanned line-by-line from top to bottom.
- For progressive scanning, the frame rate and field rate are the same because a frame has only one field.
- Computer monitors and many digital televisions use progressive scanning.

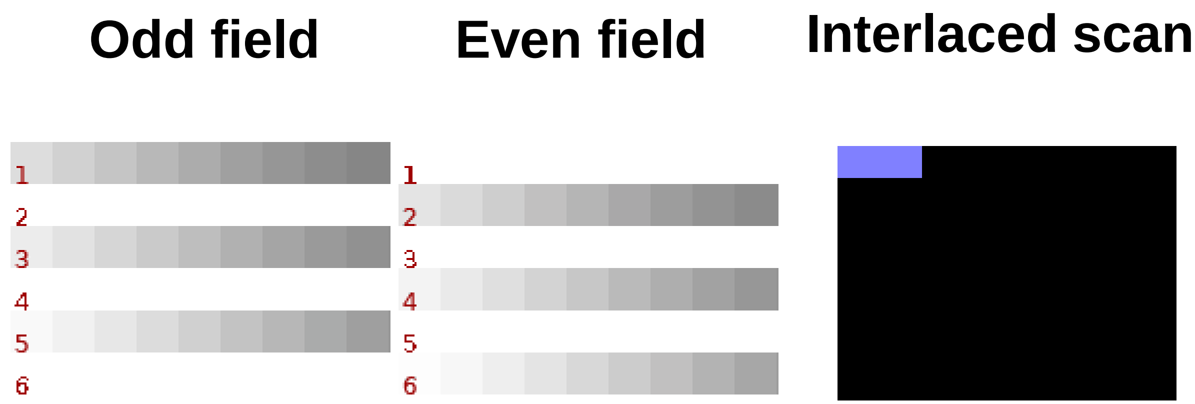


Interlaced scanning

- In ***interlaced scanning***, the lines of a frame are divided into **two** *fields*: The odd-numbered lines, called the ***upper field (odd field)***, and the even-numbered lines, called the ***lower field (even field).***

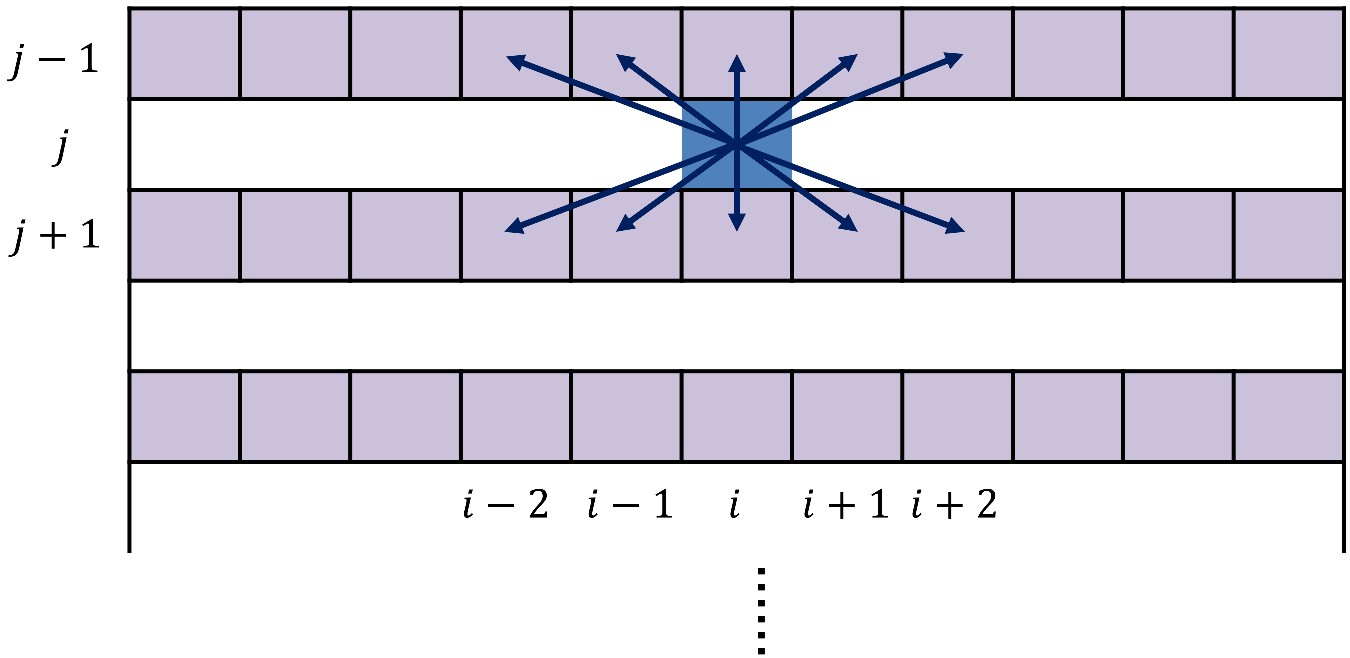


Interlaced scanning



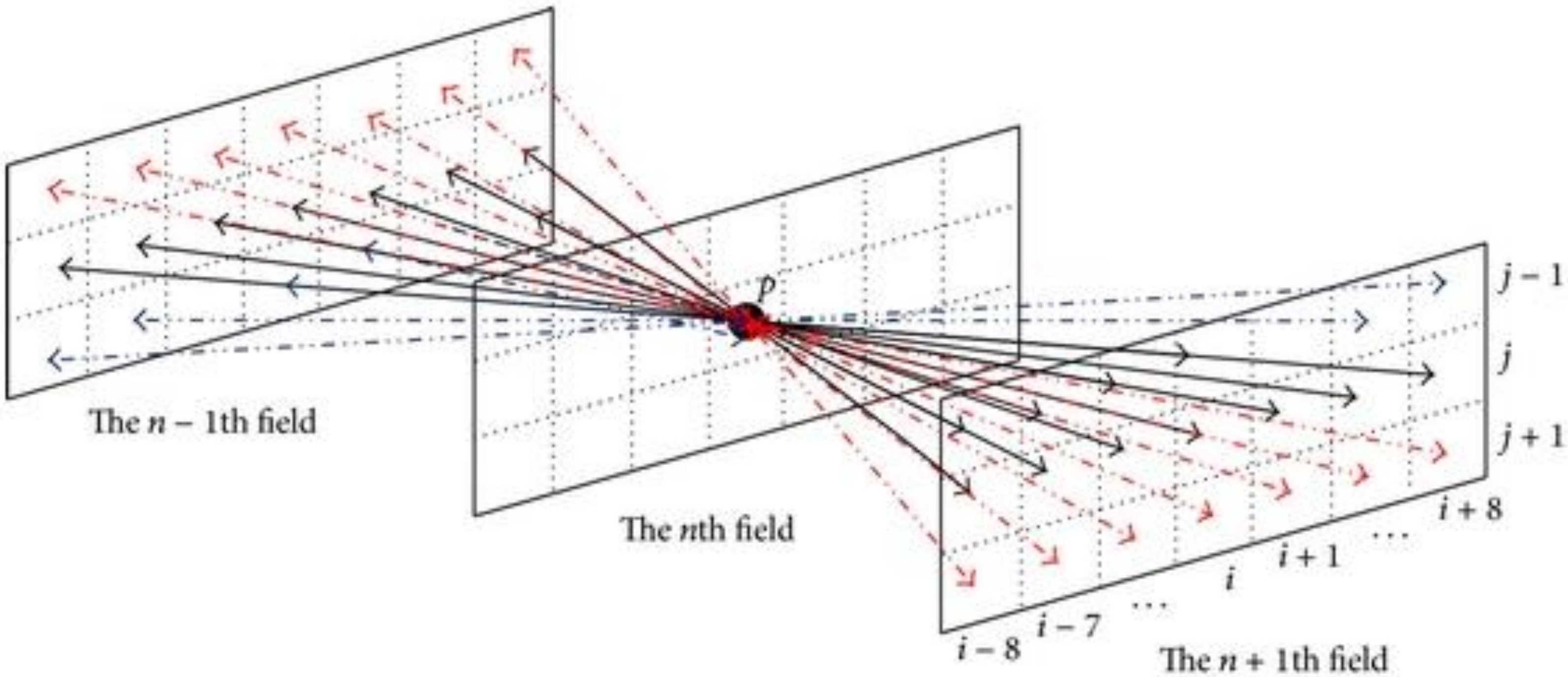
Deinterlacing

- **intra-field interpolation deinterlacing**



Deinterlacing

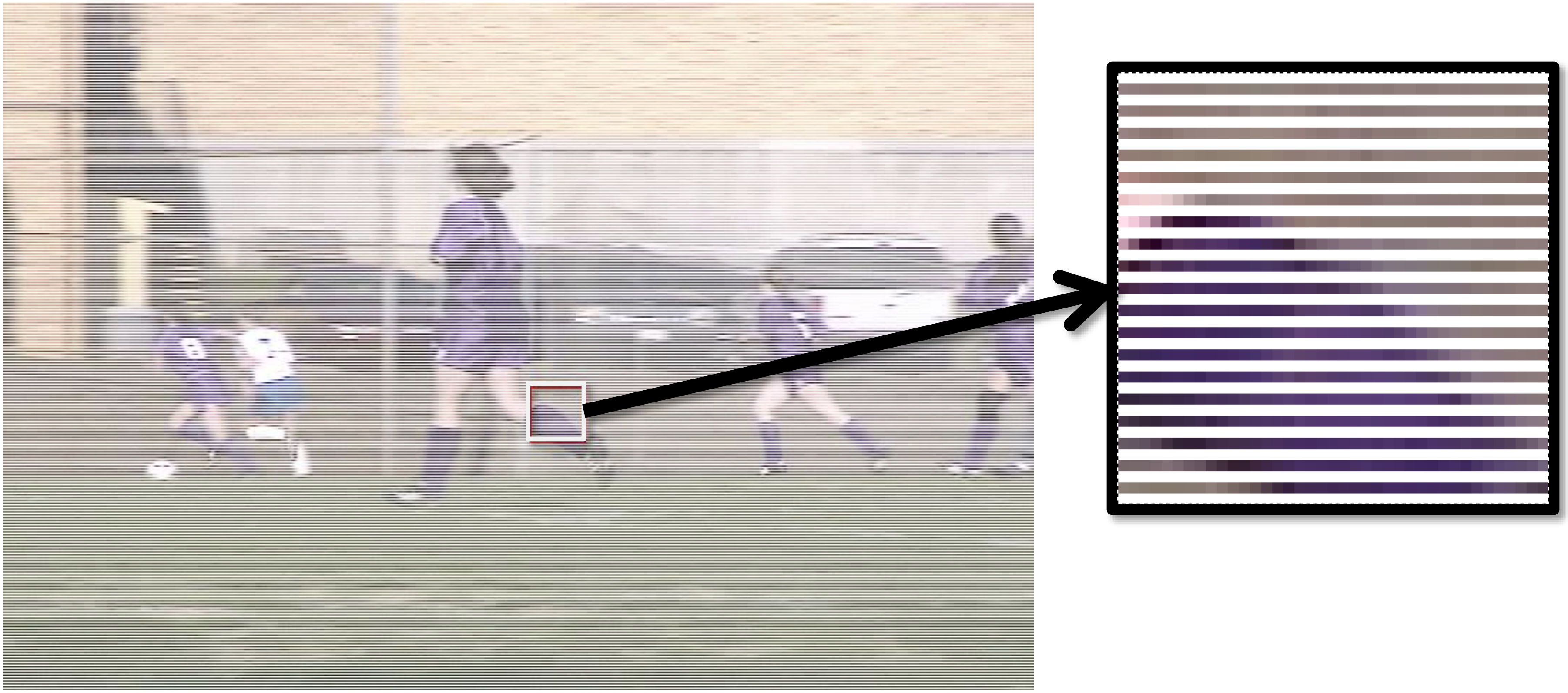
- **inter-field interpolation deinterlacing**



Deinterlacing

- **motion adaptive deinterlacing:** as areas that haven't changed from field to field don't need any processing, the regions apply intra-field interpolation. The other areas use inter-field interpolation.
- **motion compensated deinterlacing:** the exact original frames can be recovered by copying the missing field from a matching previous/next frame.

Interlacing



Deinterlaced Result

Progressive vs. Interlaced Scan

- Progressive
  - Computer monitors
  - Scans entire picture line by line
  - Eliminate flicker seen in interlaced
- Interlaced
  - Developed for **CRT** (Cathode Ray Tube) technology
  - Divides scans into odd and even lines
  - Alternately refreshes odd lines, then even lines
  - Slight delay between refreshes causes “jaggedness” or **interlace artifacts**
  - **Deinterlacing** can compensate somewhat

High Definition HDV Format

Picture Format		720/25p, 720/30p, 720/50p, 720/60p	1080/50i 1080/60i
Pixel Dimensions		1280 × 720	1440 × 1080
Frame Aspect Ration		16:9	
Pixel Aspect Ratio		1.0	1.33
Data Rate	Video data only:	approx. 19 Mbps	approx. 25 Mbps
Color Sampling Method		YUV 4:2:0	
Audio Setting	Sampling rate and bit depth:	48 kHz 16-bit	
	Bit rate after compression:	384 kbps	

HDV Picture Format Notation

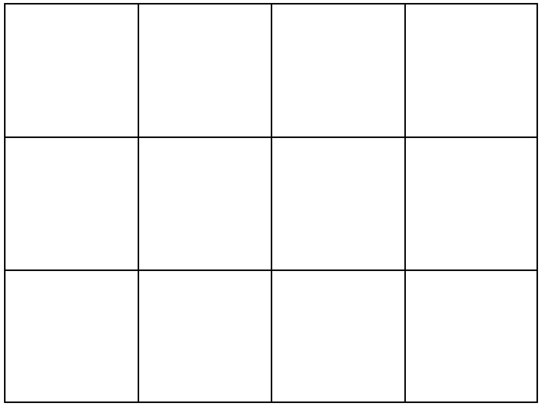
1080 / 60 i

"i": interlaced
"p": progressive

frame/field frequency

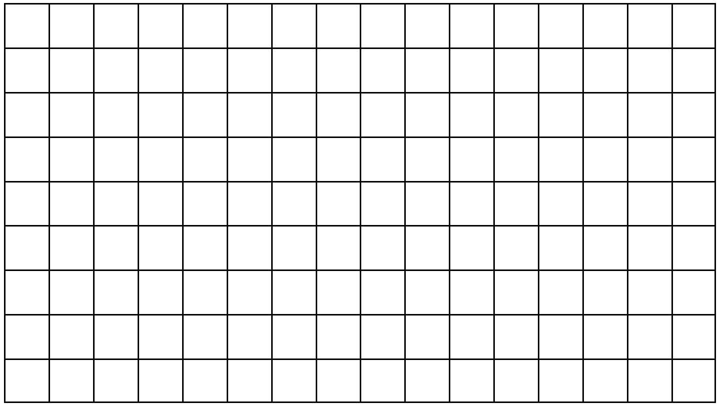
frame height

Frame Aspect Ratio Examples



4:3

Example:
- Standard definition NTSC standard format



16:9

Examples:
- Standard definition NTSC wide-screen format
- High definition digital video
- High definition TV

Comparison of 4: 3 and 16: 9 image aspect ratio

16:9 Aspect ratio

16:9 Aspect ratio displayed on a 4:3 screen (letter box)

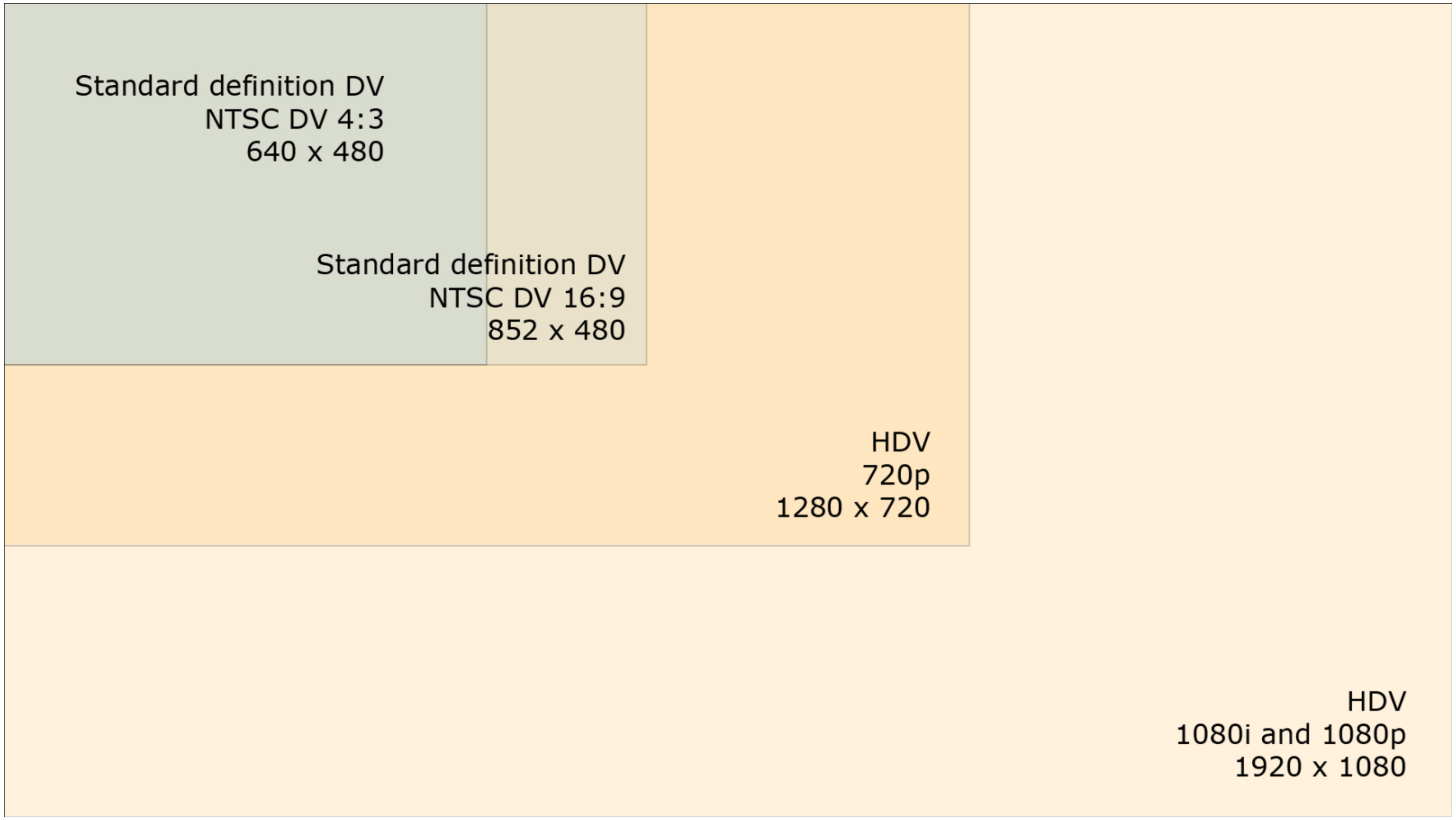
→ Letter box

4:3 Aspect ratio

4:3 Aspect ratio displayed on a 16:9 screen (pillar box)

→ Pillar box

Frame Size (Resolution) Comparison



By viewing frame size

Frame Size (Resolution) Comparison

A frame from a 1080i video

Frame Size (Resolution) Comparison

Same frame as 720p

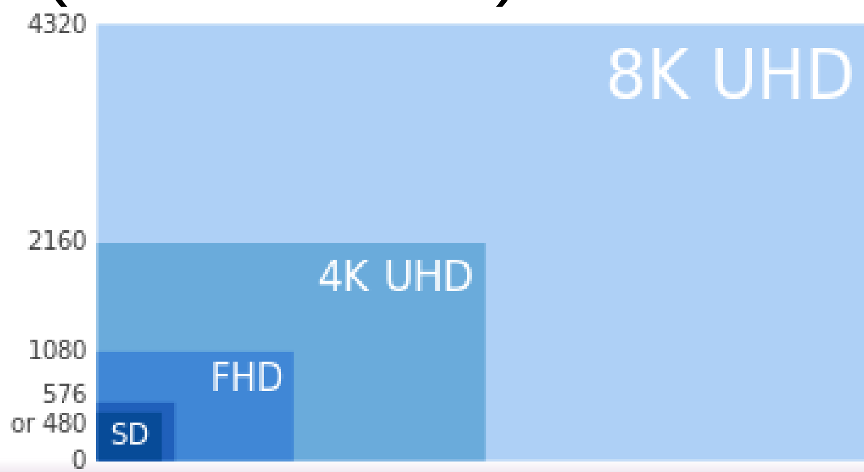
Higher Resolution Nowadays

There are 2 common standards on 4K videos:

- A 4K resolution, as defined by Digital Cinema Initiatives, is 4096 x 2160 pixels(1.9:1 aspect ratio)
- Ultra HD(UHD) 4K resolution is 3840 x 2160 (1.78:1 aspect ratio)

8K resolution or FUHD (Full Ultra HD) is 7680 x 4230.

4K TV (3840×2160)	
1080p (1920×1080)	1080p (1920x1080)
1080p (1920×1080)	1080p (1920×1080)

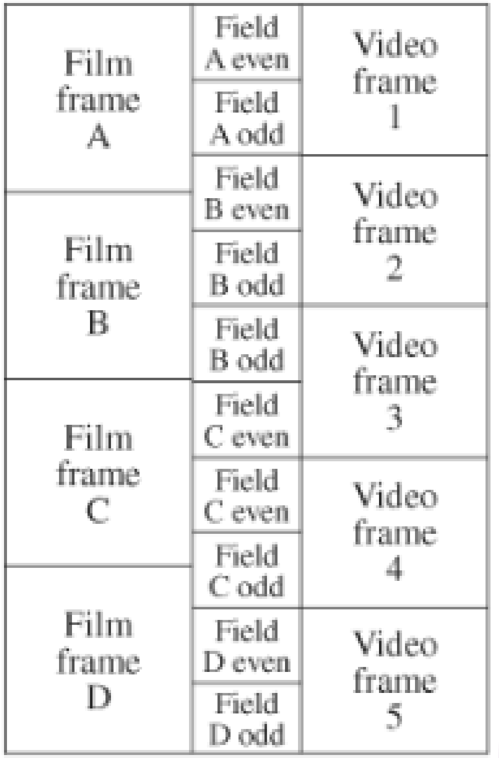


Telecine and pulldown

- The word ***telecine*** refers to both the process that transfers film to video and the machine that performs the process.
- The major difficulty is that film and television have different frame rates. A better way, one that creates smoother video and truer audio, is called ***pulldown.***
- ***Pulldown*** is a method for using interlaced fields more than once, across frames, to make up for a discrepancy in frame rates as film is translated to video.
- For NTSC video, it uses ***3:2 pulldown.*** The first step is to slow down the film by 0.1% so that we can get to an integer-based ratio of frame rates.

Telecine and pulldown

- If you multiply 24 * 0.999, you get 23.976. This gives you a ratio of 23.976/29.97, which is 4/5.
- Now these numbers are something we can deal with. For each four frames of film, we need to create five frames of video.
- The figure illustrates 3: 2 pulldown.
- The name reflects the pattern of how many fields are used from a frame .



Digital Video Distribution Media

TABLE 6.7 Digital Video Distribution Media

Format	VCD	SVCD	DVD	HD-DVD	Blu-ray
NTSC Resolution	352 × 240	480 × 480	720 × 480	1920 × 1080	1920 × 1080
Video Compression	MPEG-1	MPEG-2	MPEG-2	MPEG-2, MPEG-4 AVC, SMPTE-VC1	MPEG-2, MPEG-4 AVC, SMPTE-VC1
Audio Compression	MP1	MP1	PCM, DD, DTS Surround	PCM, DD, DD^{+}, DD, TrueHD, DTS, DTS-HD	PCM, DD, DD^{+}, DD, TrueHD, DTS, DTS-HD
Video Bit Rate	~1.2 Mb/s	~2 Mb/s	~10 Mb/s	~28 Mb/s	~40 Mb/s
Length (in time)	74 min. on CD	35–60 min. on CD	1–4 hours or more of SD	2 hours or more of HD, depending on the number of layers	2 hours or more of HD, depending on the number of layers

Properties of codecs

- Digital video files are very large. With no compression or subsampling, NTSC standard video would have a data rate of over 240 Mb/s; HD would have a data rate of about 1 Gb/s.
- Remove redundancies and extraneous information within one frame is called ***intraframe compression***. It also can be referred to as ***spatial compression.***
- There are two commonly used methods for accomplishing ***spatial compression***: transform encoding and vector quantization.
- ***Temporal compression*** is a matter of eliminating redundant or unnecessary information by considering how images change over time. it is also called ***interframe compression***.

Properties of codecs

- The basic method for compressing between frames is to detect how objects move from one frame to another, represent this as a vector.
- Determining the motion vector is done by a method called ***motion estimation.***
- Some codecs allow you to select either ***constant*** or ***variable bit rate encoding*** (***CBR*** and ***VBR***, respectively). Variable bit rate varies the bit rate according to how much motion is in a scene.
- Codecs are mostly ***asymmetrical***. This means that the time needed for compression is not the same as the time needed for decompression.

Compression Standards

Year	Standard	Publisher	Applications
1984	H.120	ITU-T	
1990	H.261	ITU-T	Videoconferencing Videotelephony
1993	MPEG-1 Part 2	ISO, IEC	Video-CD
1995	H.262/MPEG-2 Part 2	ISO, IEC, ITU-T	DVD Video, Blu-ray, Digital Video Broadcasting, SVCD
1996	H.263	ITU-T	Videoconferencing, videotelephony, video on mobile phones (3GP)
1999	MPEG-4 Part 2	ISO, IEC	Video on Internet (DivX, Xvid ...)
2003	H.264/MPEG-4 AVC	Sony, Panasonic, Samsung, ISO, IEC, ITU-T	Blu-ray, HD DVD, Digital Video Broadcasting, iPod Video, Apple TV, videoconferencing
2013	H.265	ISO, IEC, ITU-T	Not popular yet

MPEG compression

- MPEG compression was developed in two lines.
  - The first was the work of ITU-T and their subcommittee, the **Video Coding Experts Group**. We know this line of codecs as the H.26* series
  - The second line emerged from the **Motion Picture Experts Group**, from which we get the name MPEG
- The revolutionary advance in MPEG-4 compression is the use of object-based coding.
- ***MPEG-4 AVC*** (Advanced Video Coding) and equivalent to ***H.264***, is an improved MPEG-4 version introduced in 2003 that quickly achieved wide adoption for DVD; videoconferencing; videophone…

MPEG General Information

- Goal: data compression 1.5 Mbps
- MPEG defines video, audio coding and system data streams with synchronization.
- MPEG information
  - Aspect ratios: 1:1 (CRT), 4:3 (NTSC), 16:9 (HDTV)
  - Refresh frequencies: 23.975, 24, 25, 29.97, 50, 59.94, 60 Hz

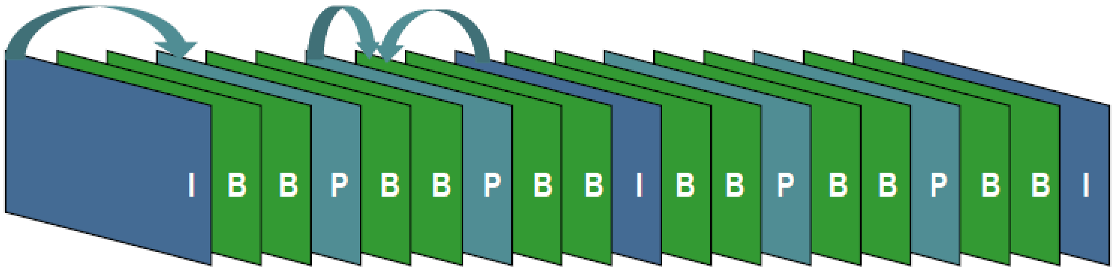
MPEG Image Preparation - Blocks

- Each image is divided into **macro-blocks.**
- Macro-block : 16x16 pixels for luminance; 8x8 for each chrominance component.
- Macro-blocks are useful for Motion Estimation.
- No MCUs which implies sequential non-interleaving order of pixels values.

H.264

Main steps of MPEG compression

- **Step 1.** Divide the sequence of frames into GOPs, identifying I, P, and B frames.
  - A ***GOP*** is a ***group of pictures***, that is, a group of *n* sequential video frames
  - ***I frames***, or ***intraframes***, are compressed independently, as if they were isolated still images, using JPEG compression.
  - I frames serve as reference points for the ***P frames*** (***interframes***, also called ***forward prediction frames***) and ***B frames*** (***bidirectional frames***), which are compressed both spatially and temporally.



Main steps of MPEG compression

- **Step 2.** Divide each frame into macroblocks.
  - A ***macroblock*** is a 16 × 16 pixel area.
  - A 16 × 16 macroblock can be divided into 8 × 8 blocks. The way that macroblocks are divided depends on the particular compression standard, which can apply different types of chrominance subsampling.

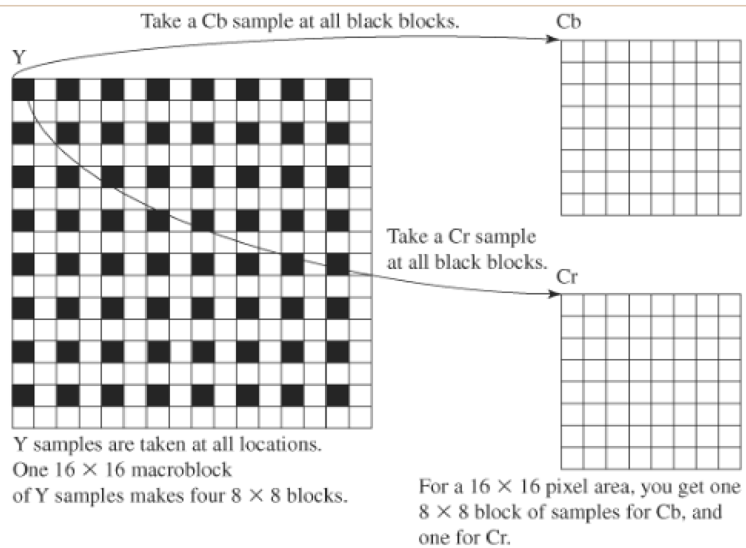


Figure 7.22 Macroblock for 4: 2: 0 subsampling

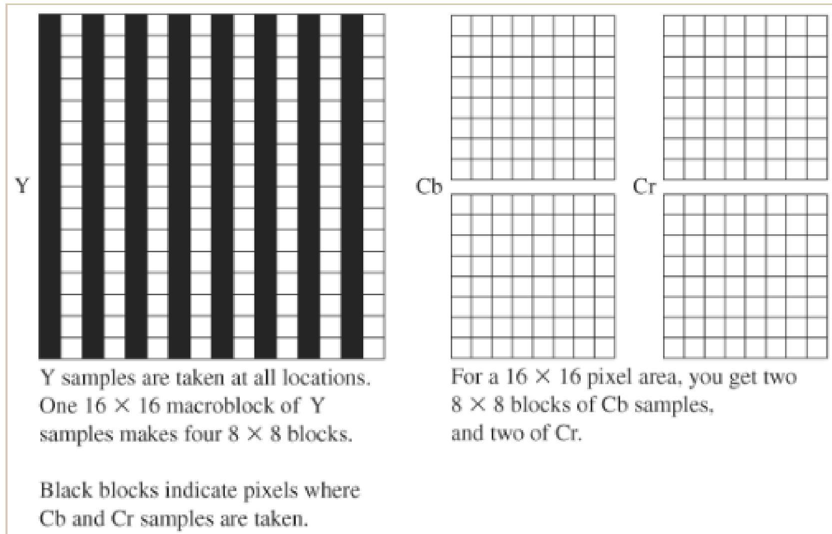
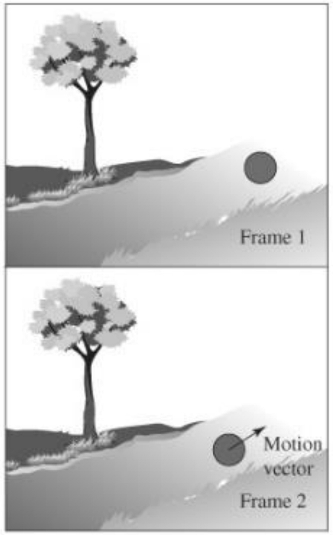


Figure 7.23 Macroblock for 4: 2: 2 subsampling

Main steps of MPEG compression

- **Steps 3 and 4.** For each P and B frame, compare the frame to the related I frame to determine a motion vector. Record differential values for P and B frames.
  - This step is called ***motion estimation.***
  - It's more economical to convey the difference between one frame and the next, a method called ***differential encoding.***
  - Motion estimation determines how much a frame has "moved" since the previous frame.
  - The difference between the macroblock in frame 2 and the matching macroblock in frame 1 is called the **prediction error**.

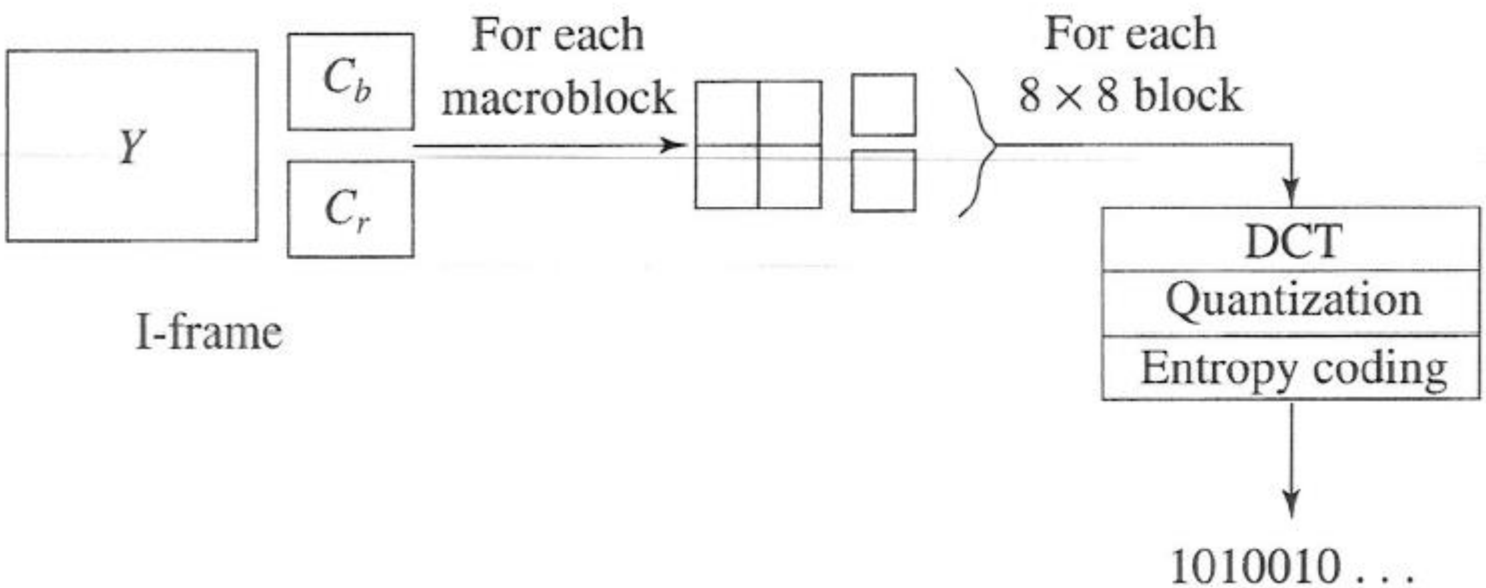


Main steps of MPEG compression

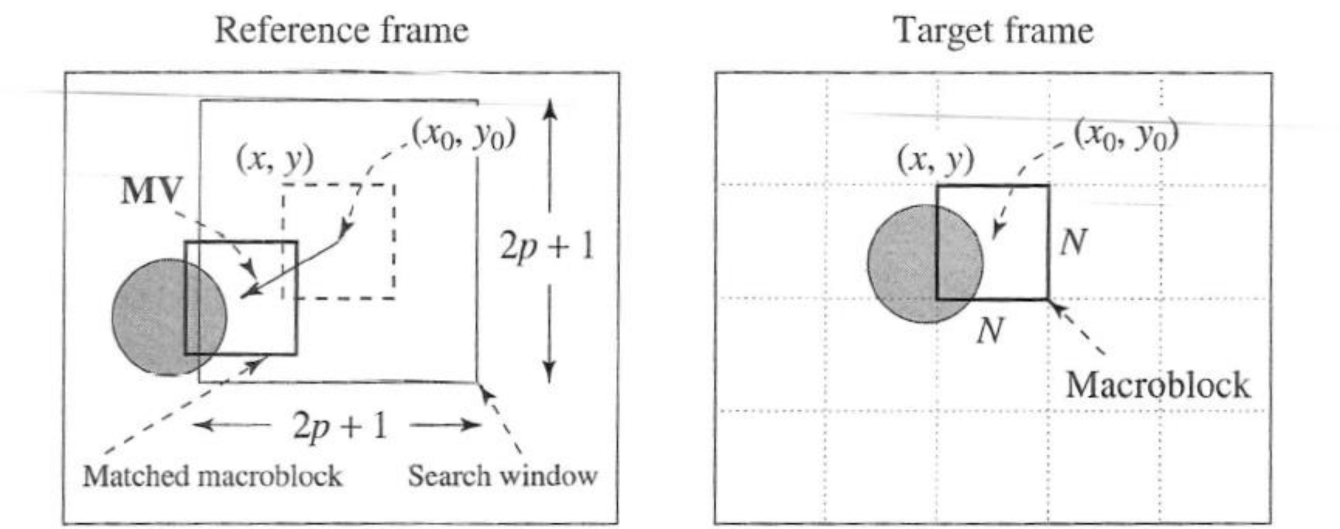
- The P or B frame being compressed is called the ***target frame***, and the I frame is named ***reference frame***.
- The reference frame to which a P frame is compared is called its ***forward prediction frame***. The reference frame to which a B frame is compared is called its ***backward prediction frame.***
- Assume we have a macroblock in the target frame T. We will search for a matching macroblock in reference frame R. We want to look in the vicinity of $\boldsymbol{R}_{x,y}$ for the macroblock that most closely matches $\boldsymbol{T}_{x,y}$.

Intra-frame coding

- Various lossless and lossy compression techniques use – like JPEG

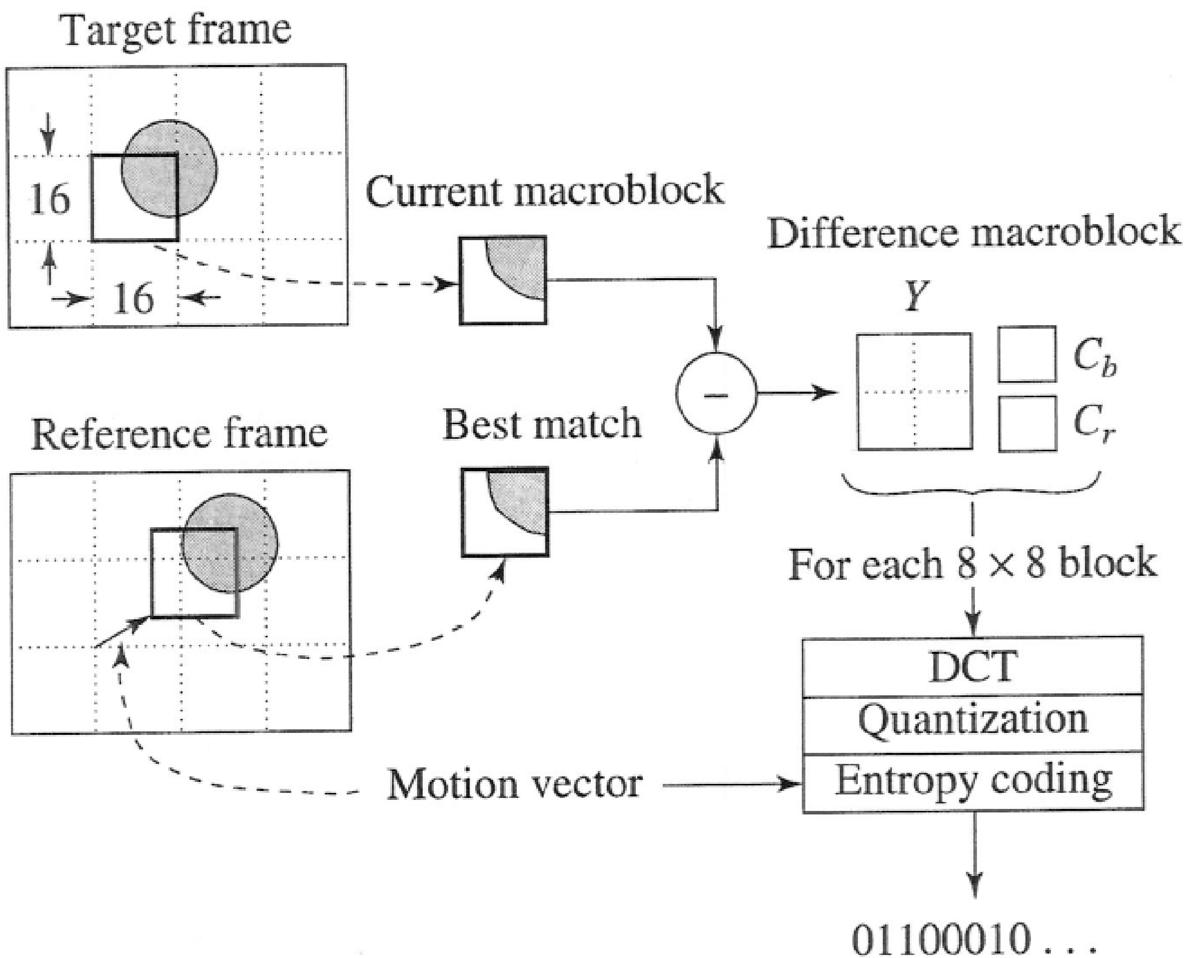


Motion Compensation



In H.261, motion vectors are in the range [-15,15]x[-15,15], e.g, p = 15.

MPEG Video for P-Frames



Estimate motion vectors

- Basic idea to find motion vectors is to search macroblocks
  - Within a $\pm p \times p$ pixel search window
  - Calculate **Sum of Absolute Difference** (SAD) of each macroblock(or **Mean Absolute Error** (MAE))
  - Choose macroblock which SAD/MAE is a minimum
- If the encode decides that no acceptable match exists then
  - Coding that macroblock as an intra macroblock
  - In this manner, high quality video is maintained at a slight cost to coding efficiency

Sum of Absolute Difference (SAD)

- SAD is computed by

$$SAD(\boldsymbol{i},\boldsymbol{j}) = \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} |T(x+k,y+l) - R(x+k+\boldsymbol{i},y+l+\boldsymbol{j})|$$

- N is size of macroblock window typically (16 or 32 pixels)
- (x, y) the position of the **target** macroblock T, and R is the **reference** region to compute the SAD.
- $T(x+k, y+l)$ — pixels in the macroblock with upper left corner (x, y) in the target
- $R(x+k+i, y+l+j)$ — pixels in the macroblock with upper left corner $(x+i, y+j)$ in the reference

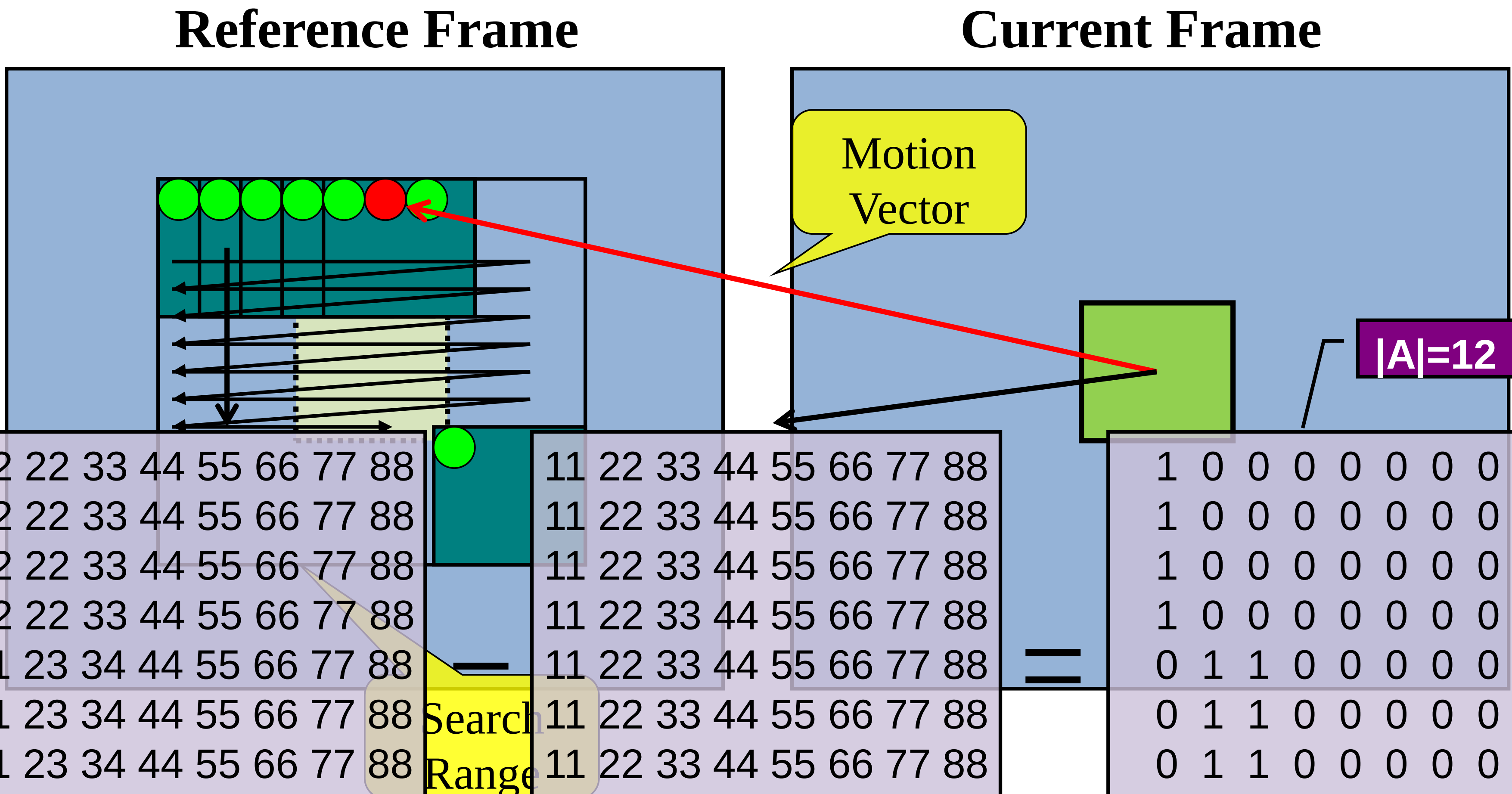
Sum Square Differences (SSD)

- Alternatively: sum of squared differences

$$SSD(\boldsymbol{i},\boldsymbol{j}) = \sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\Big(\big(T(x+k,y+l) - R(x+k+\boldsymbol{i},y+l+\boldsymbol{j})\big)\Big)^2$$

- Goal is to find a vector (i,j) such that SAD/SSD (i,j) is minimum

Exhaustive Block-Matching Algorithm



12	22	33	44	55	66	77	88
12	22	33	44	55	66	77	88
12	22	33	44	55	66	77	88
12	22	33	44	55	66	77	88
11	23	34	44	55	66	77	88
11	23	34	44	55	66	77	88
11	23	34	44	55	66	77	88
11	23	34	44	55	66	77	88

−

11	22	33	44	55	66	77	88
11	22	33	44	55	66	77	88
11	22	33	44	55	66	77	88
11	22	33	44	55	66	77	88
11	22	33	44	55	66	77	88
11	22	33	44	55	66	77	88
11	22	33	44	55	66	77	88
11	22	33	44	55	66	77	88

=

1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	1	1	0	0	0	0	0
0	1	1	0	0	0	0	0
0	1	1	0	0	0	0	0
0	1	1	0	0	0	0	0

Fast Block-Matching Algorithms

- **The characteristics of fast algorithm**
  - Not as accurate as exhaustive search algorithm
  - Save large amount of computation
- **fast motion estimation algorithm**
  - 2-D logarithm search method

2-D logarithm Search Method

- The search is accomplished by successively **reducing the area of search**.
- Each step consists of searching five locations which contain the center of the area, and the midpoints between the center and the four boundaries of the area along the axes passing through the center.
- In the final step all the nine locations are searched and the location corresponding to the minimum is the *direction of minimum distortion*.

J.R. Jain and A.K. Jain, “Displacement measurement and its application in interframe image coding,” *IEEE Trans. Commun.*, Vol. COM-29, pp. 1799–1808, Dec. 1981

The algorithm is given below.

For any integer $m > 0$, we define

$$N(m) = \{(i, j); \quad -m \leqslant i, j \leqslant m\}$$

$$M(m) = \{(0, 0), (m, 0), (0, m), (-m, 0), (0, -m)\}.$$

A 2-D Logarithmic Search Procedure for DMD:

Step 1: (initialization)

Error function

$$D(i, j) = \infty \qquad (i, j) \notin N(p)$$

$$n' = \lfloor \log_2 p \rfloor$$

P is search range

$$n = \max \cdot \{2, 2^{n'-1}\}$$

$$q = l = 0 \text{ (or an initial guess for DMD)}$$

where $\lfloor \cdot \rfloor$ is a lower integer truncation function.

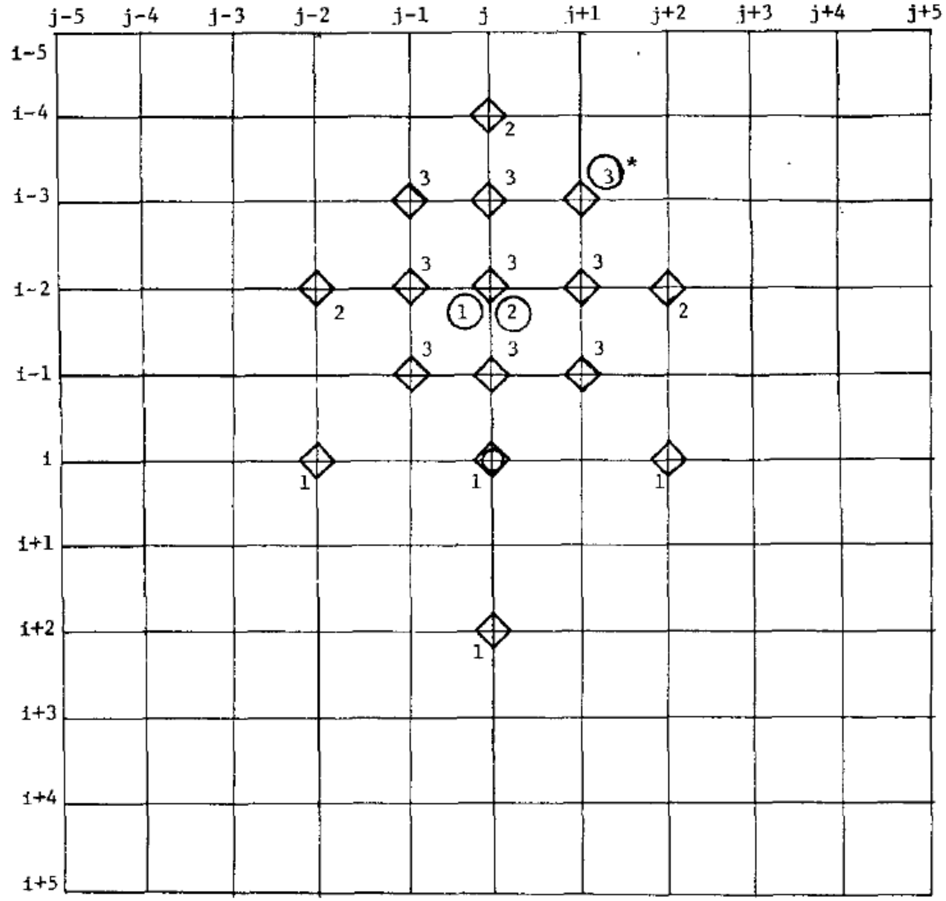
Step 2: $M'(n) \leftarrow M(n)$.

Step 3: Find $(i, j) \in M'(n)$ such that $D(i + q, j + l)$ is minimum. If $i = 0$ and $j = 0$, go to Step 5; otherwise go to Step 4.

Step 4: $q \leftarrow q + i, l \leftarrow l + j$; $M'(n) \leftarrow M'(n) - (-i, -j)$; go to Step 3.

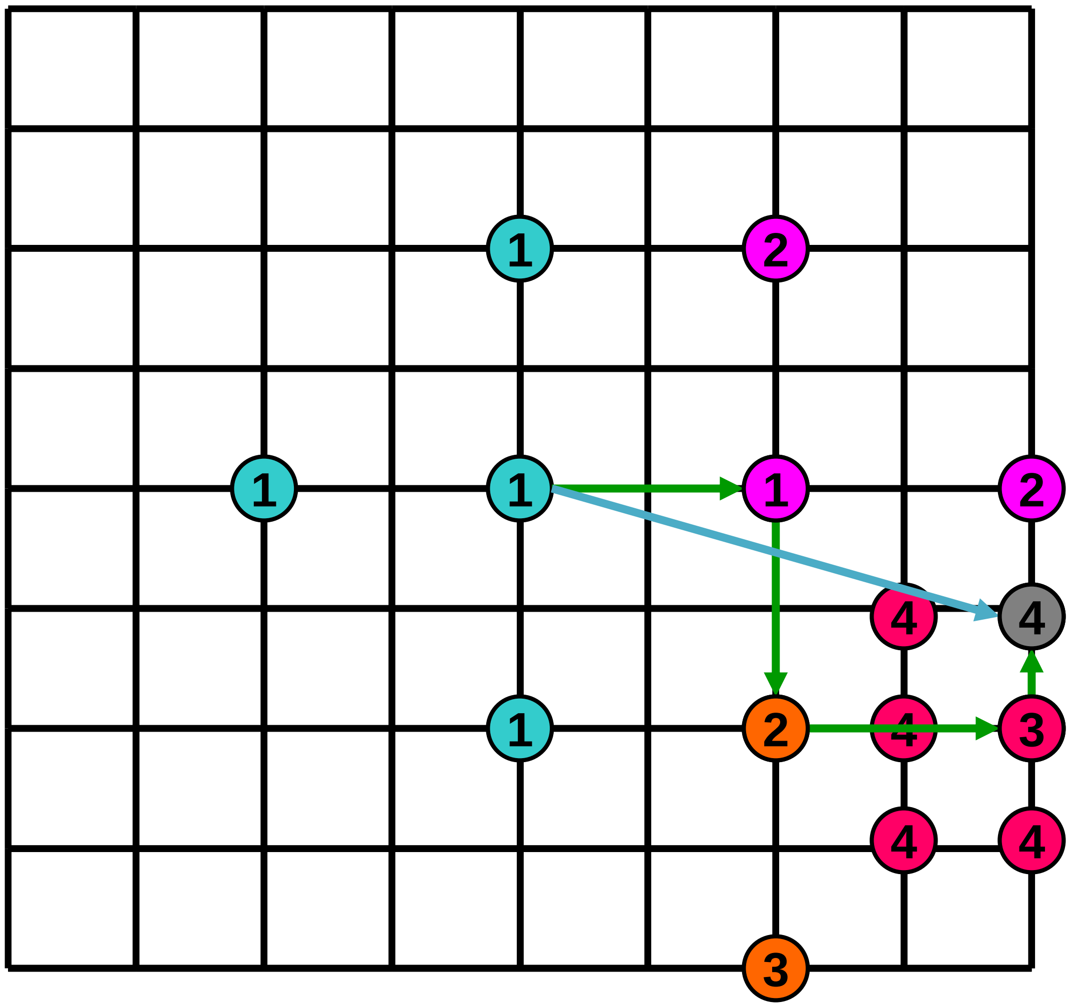
Step 5: $n \leftarrow n/2$. If $n = 1$, go to Step 6; otherwise, go to Step 2.

Step 6: Find $(i, j) \in N(1)$ such that $D(i + q, j + l)$ is minimum: $q \leftarrow i + q, l \leftarrow l + j$. (q, l) then gives the DMD.



P=5

2-D logarithm Search Method



Main steps of MPEG compression

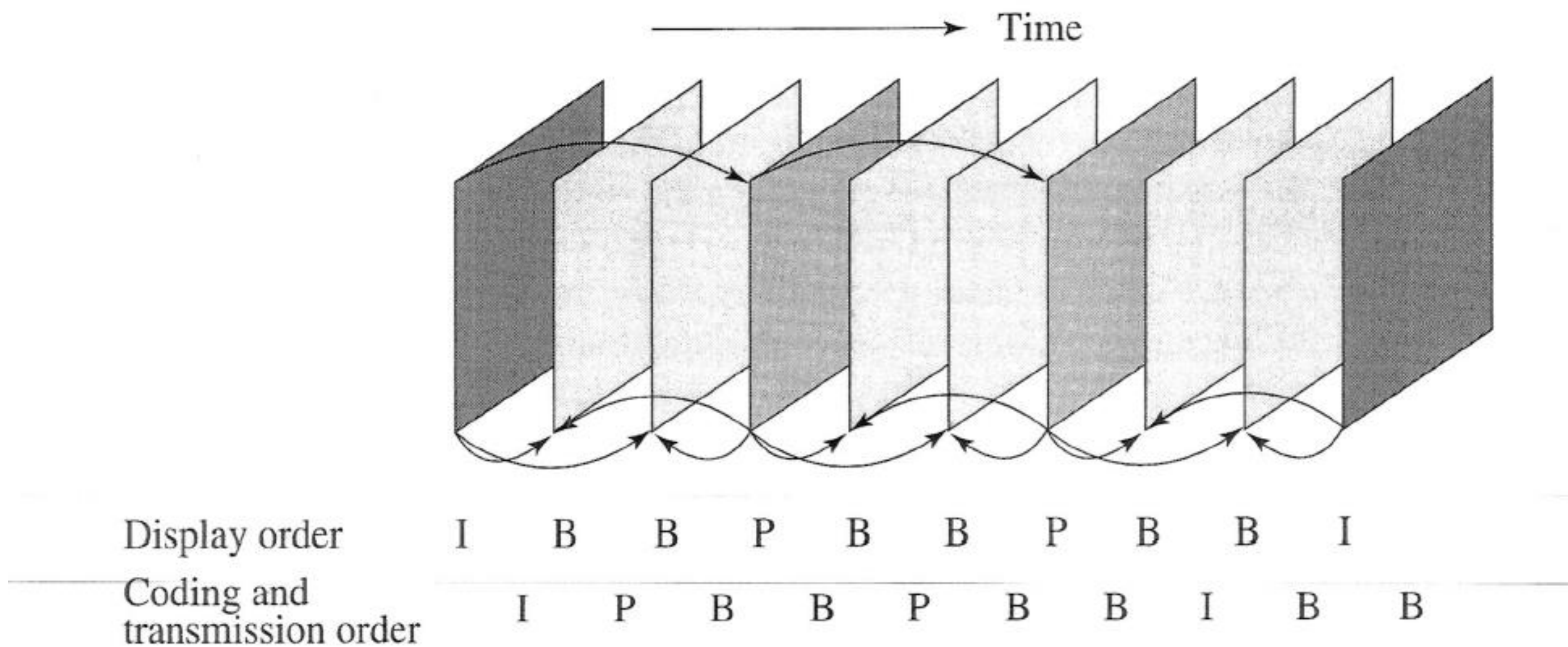
- **Step 5:** For all frames, compress with JPEG compression.
  - Compressing a frame of video is just like compressing a still image, and thus JPEG compression can be applied
  - **I frames** undergo intraframe compression only, without reference to any other image.
  - **P and B frames** first undergo motion prediction. Then the difference between the expected value of a pixel and its actual value is encoded.

MPEG-1 Video

- MPEG-1 was approved by ISO and IEC in 1991 for "Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5Mbps".
- MPEG-1 standard is composed of
  - System
  - Video
  - Audio
  - Conformance
  - And Software
- MPEG-1's video format is called SIF(Source Input Format)
  - 352x240 for NTSC at 30f/s
  - 352x288 for PAL at 25f/s

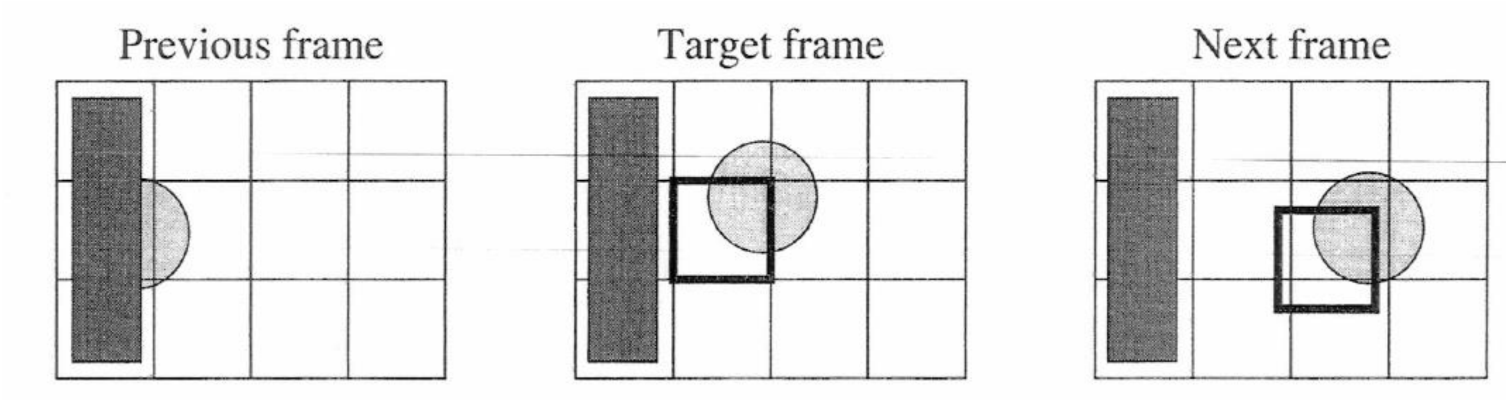
MPEG-1 Motion Compensation

- MPEG-1 introduces a new type of compressed frame: the B-frame.



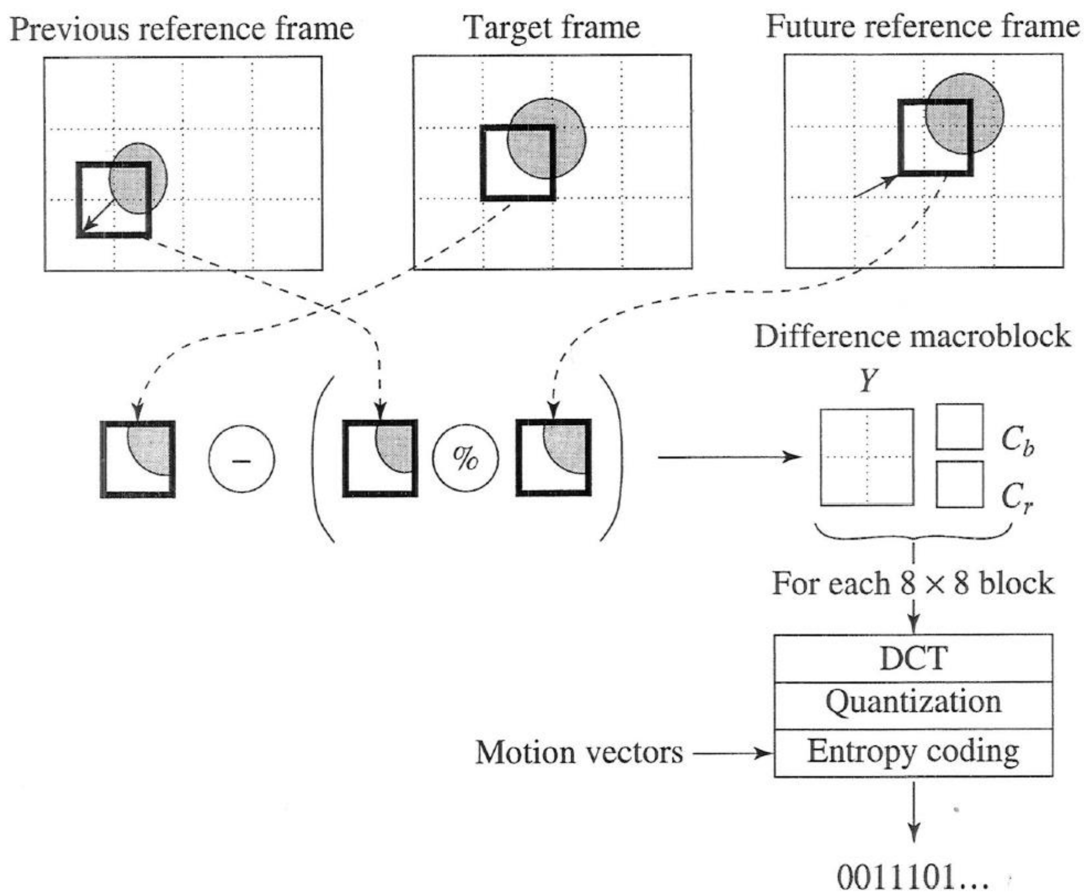
Why do we need B-frames?

- Bi-directional prediction works better than only using previous frames when occlusion occurs.

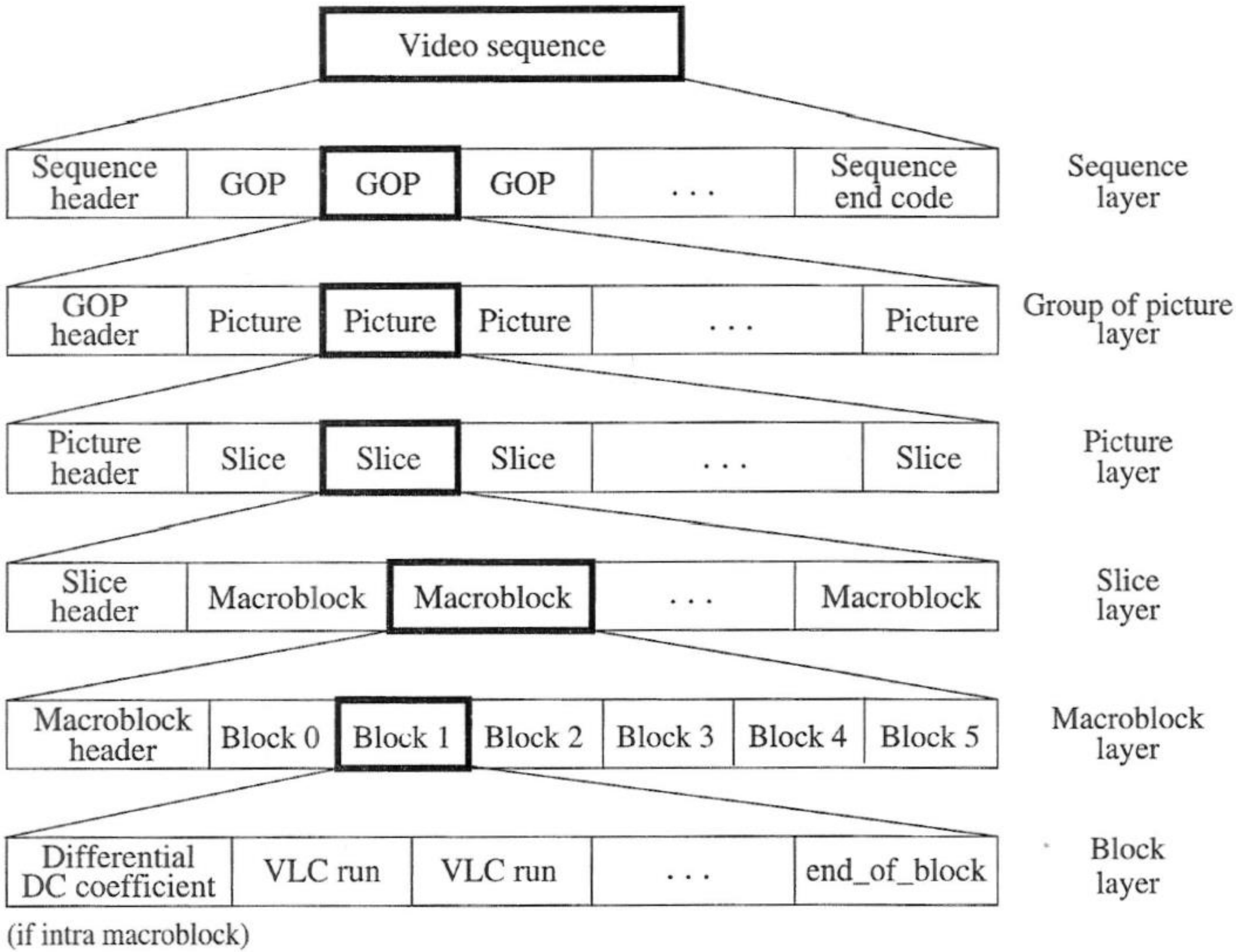


For this example, the prediction from next frame is used and the prediction from previous frame is not considered.

Compression of B-frames



MPEG-1 Video Stream



MPEG-2

- **MPEG-2** (aka H.222/H.262 as defined by the ITU) is a standard for "the generic coding of moving pictures and associated audio information“
- Backwards compatibility with existing hardware and software means it is still widely used, for example in the DVD-Video standard
- MPEG-2 evolved out of the shortcomings of MPEG-1

MPEG-2

- MPEG-2 profiles and levels:

Level	Simple profile	Main profile	SNR scalable profile	Spatially scalable profile	High profile	4:2:2 profile	Multiview profile
High		*			*		
High 1440		*		*	*		
Main	*	*	*		*	*	*
Low		*	*				

Four levels in the main profile of MPEG-2.

Level	Maximum resolution	Maximum fps	Maximum pixels/sec	Maximum 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 equivalent

Scalability

- SNR scalability
  - Base layer uses rough quantization, while enhancement layers encode the residue errors.
- Spatial scalability
  - Base layer encodes a small resolution video; enhancement layers encode the difference of bigger resolution video with the “un-sampled” lower resolution one.
- Temporal scalability
  - Base layer down-samples the video in time; enhancement layers include the rest of the frames.
- Hybrid scalability
- Data partitioning

MPEG-4

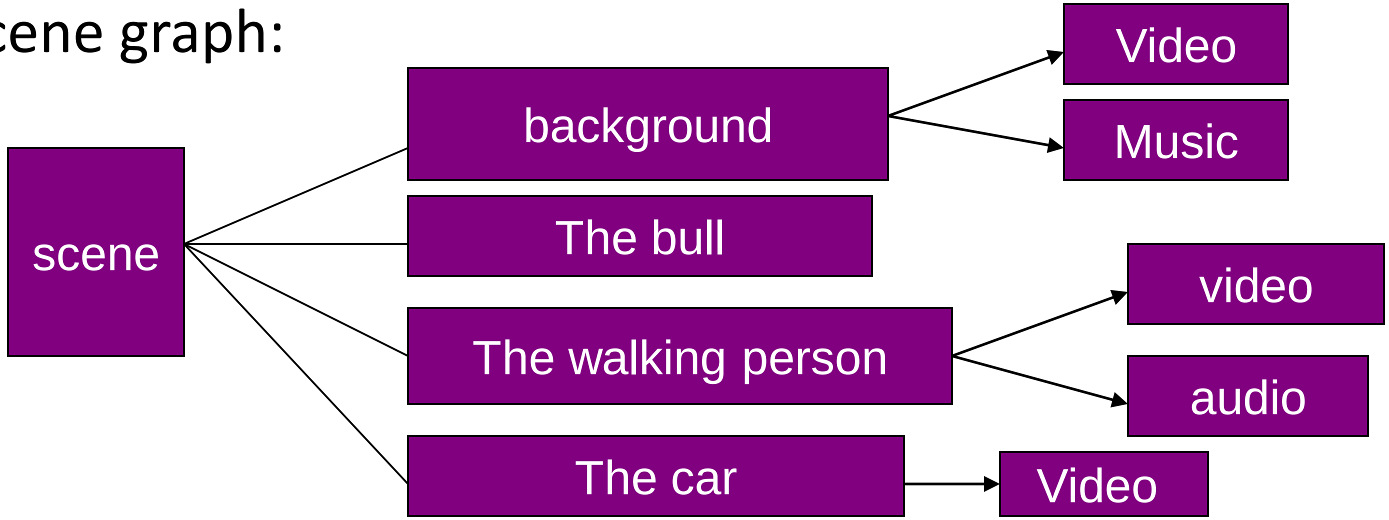
- Officially up to 10 Mbits/sec.
- Improved encoding efficiency.
- Content-based interactivity.
- Content-based and temporal random access.
- Integration of both natural and synthetic objects.
- Temporal, spatial, quality and object-based scalability.
- Improved error resilience.
- Support object-based features for content
- Enable dynamic rendering of content
  - defer composition until decoding
- Support convergence among digital video, synthetic environments, and the Internet

Audio-Video Object

- MPEG4 is based on the concept of media objects.

Audio Video Objects

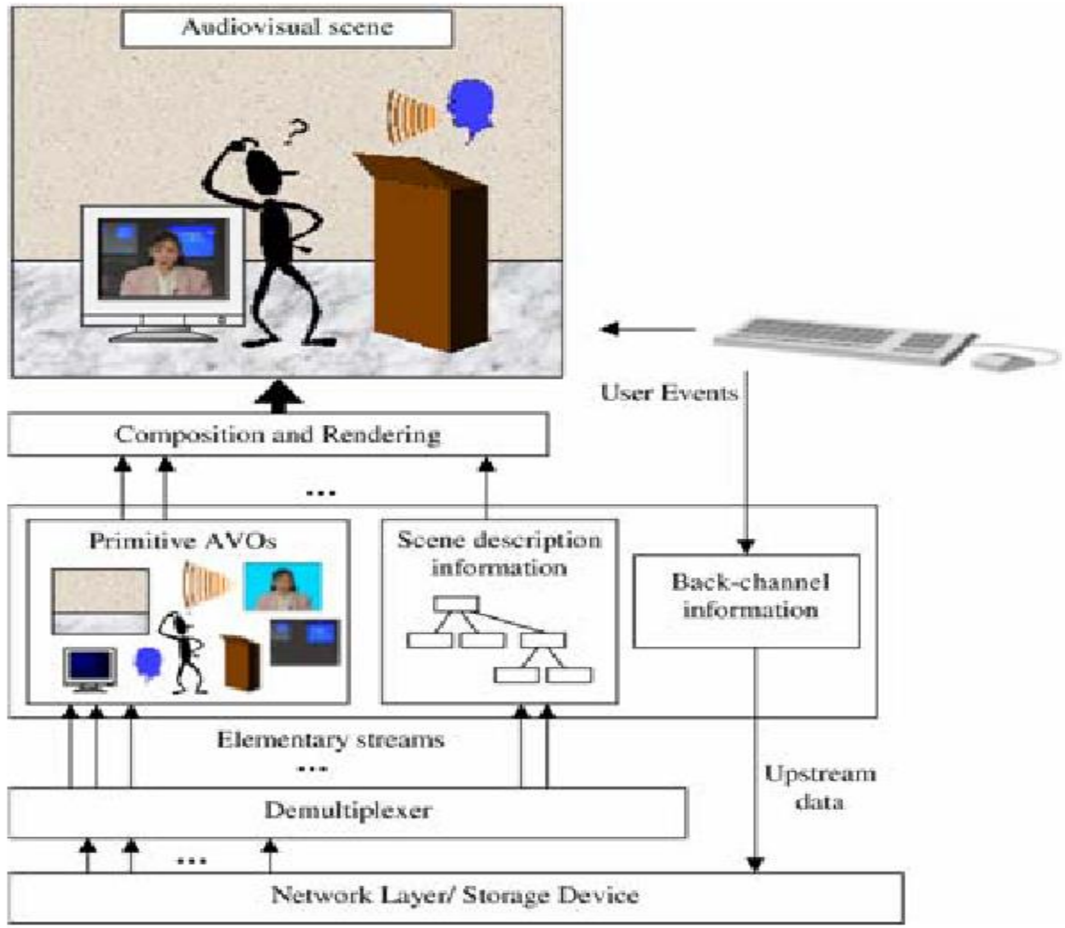
- A media object in MPEG4 could be
  - A video of an object with “shape”.
  - The speech of a person.
  - A piece of music.
  - A static picture.
  - A synthetic 3D cartoon figure.
- In MPEG4, a scene is composed of media objects based on a scene graph:



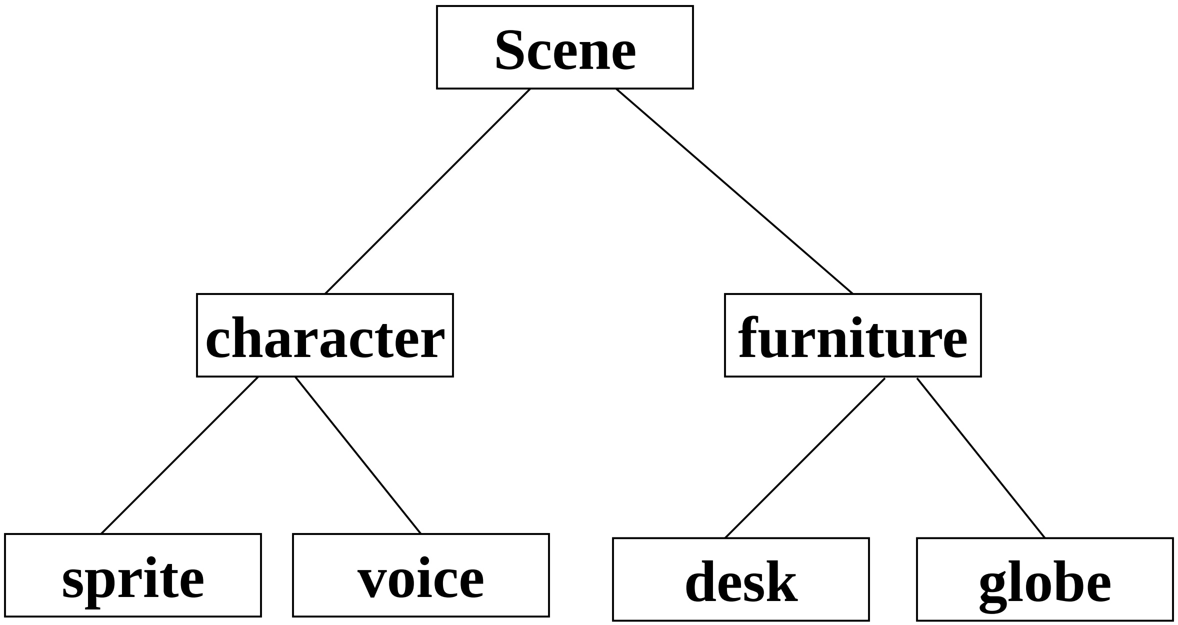
MPEG-4 Standard

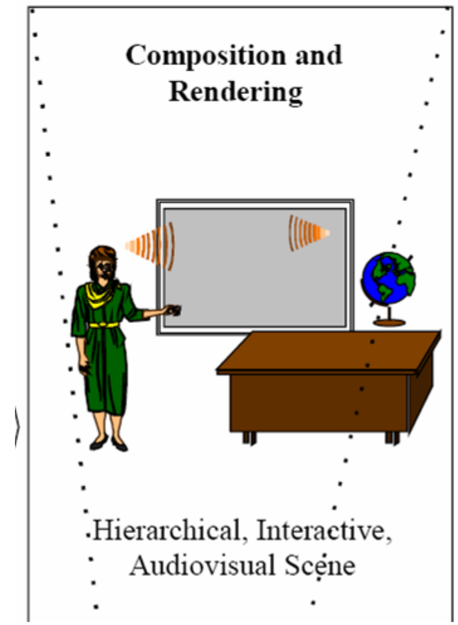
- Defines the scheme of encoding audio and video objects
  - Encoding of shaped video objects.
  - Sprite encoding.
  - Encoding of synthesized 2D and 3D objects.
- Defines the scheme of decoding media objects.
- Defines the composition and synchronization scheme.
- Defines how media objects interact with users.

Composition and Interaction



Composition



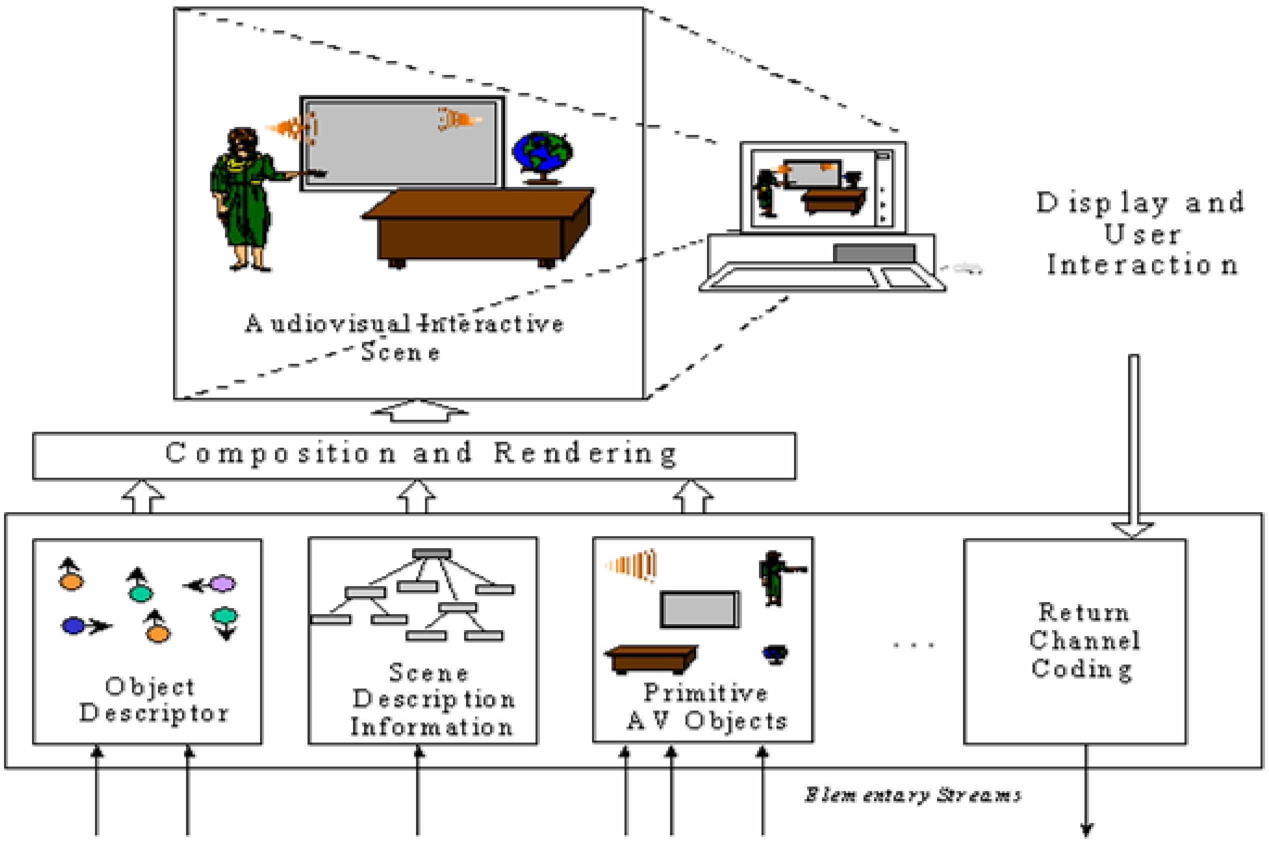


ISO N3536 MPEG4

Composition (cont.)

- Encode objects in separate channels
  - encode using most efficient mechanism
  - transmit each object in a separate *stream*
- Composition takes place at the decoder, rather than at the encoder
  - requires a binary scene description (BIFS)
- BIFS is low-level language for describing:
  - hierarchical, spatial, and temporal relations

MPEG-4 Rendering



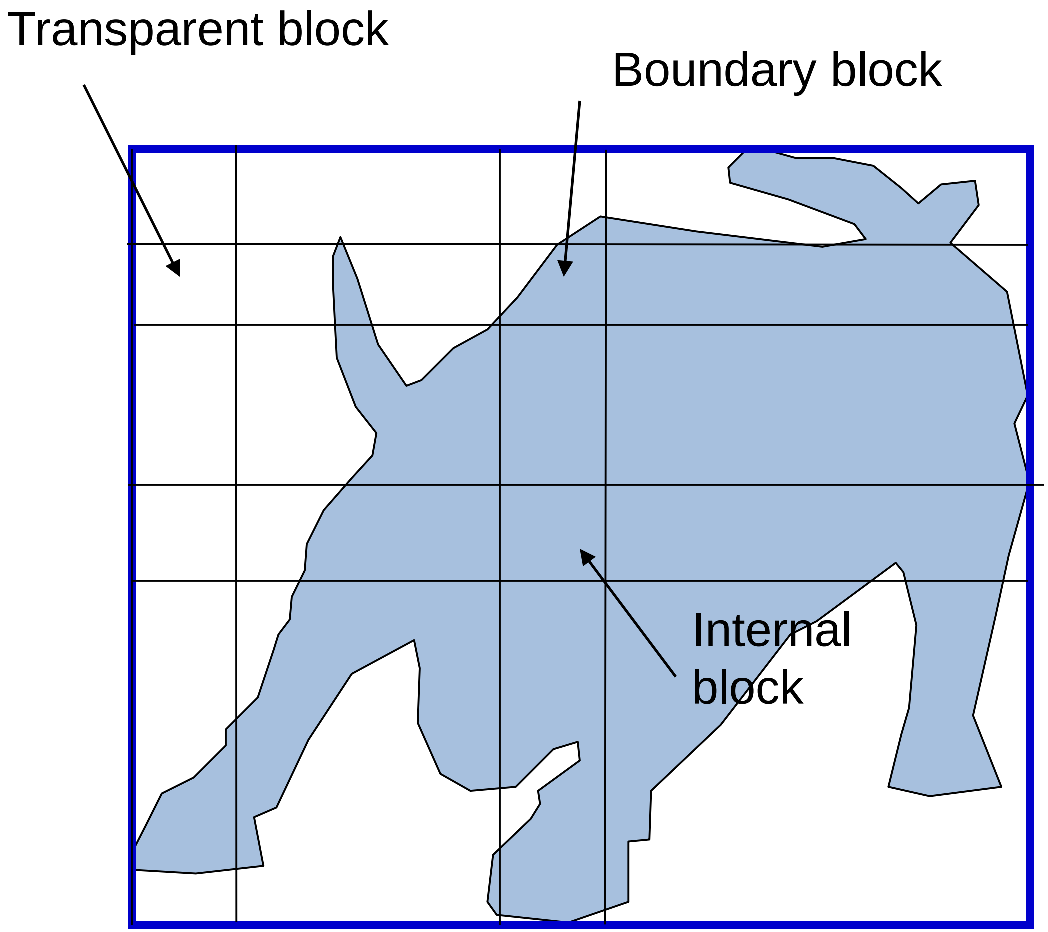
Interaction as Objects

- Change colors of objects
- Toggle visibility of objects
- Navigate to different content sections
- Select from multiple camera views
  - change current camera angle
- Standardizes content and interaction
  - e.g., broadcast HDTV and stored DVD

Video Coding in MPEG-4

- Support for 4 types of video coding:
  - Video Object Coding
    - For coding of natural and /or synthetic originated, rectangular or arbitrary shaped video objects.
  - Mesh Object Coding
    - For visual objects represented with a mesh structure.
  - Model-based Coding
    - For coding of a synthetic representation and animation of a human face and body.
  - Still Texture Coding
    - For wavelet coding of still textures.

Shape Coding

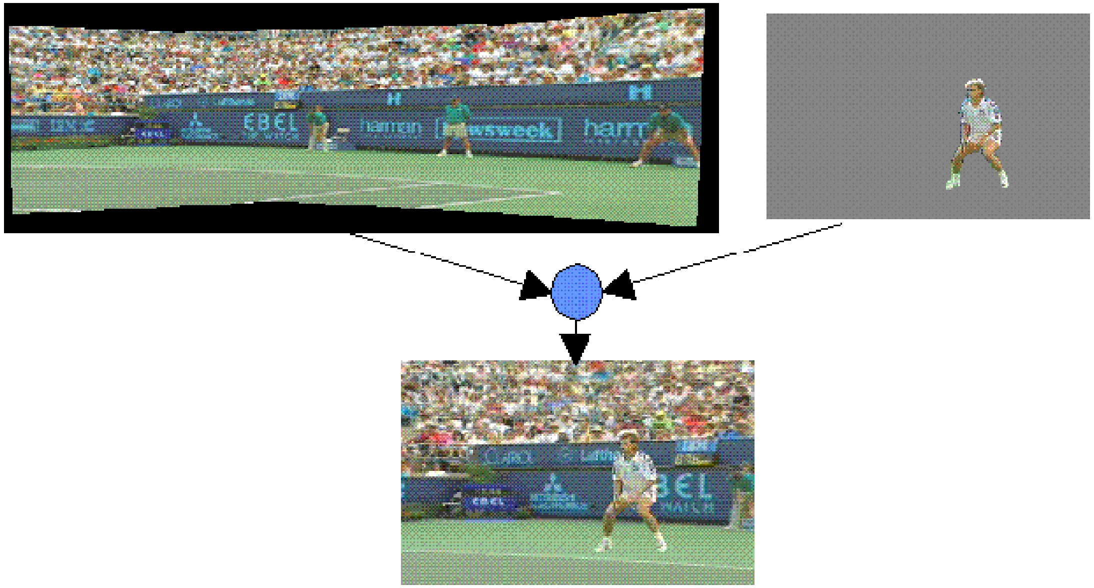


Shape coding:

- Bitmap image of a shape – alpha plane
  - Binary alpha plane.
  - Grayscale alpha plane.
- Binary alpha plane – shape information only.
- Grayscale alpha plane – shape and transparency information.
- Inter and Intra coding for the binary shapes.

Sprite Coding

- Sprite coding is used for encoding a scene with large static background with small foreground objects.
- Background is coded only once at the beginning of the sequence as an Intra-VOP.
- It uses global motion parameters to manipulate the background.



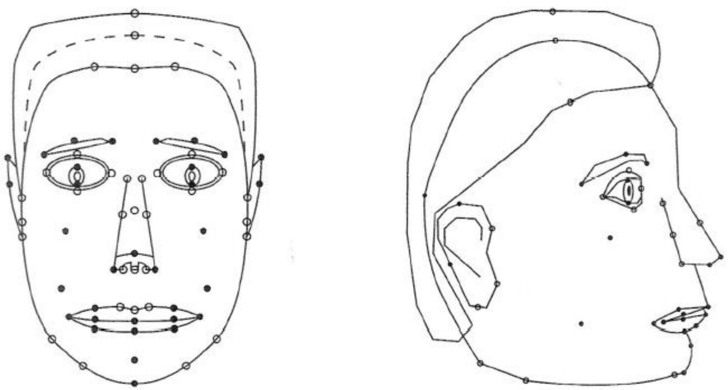
Mesh Coding

- Mesh
  - Partitioning of an image into polygonal patches.
- MPEG-4 supports 2D meshes

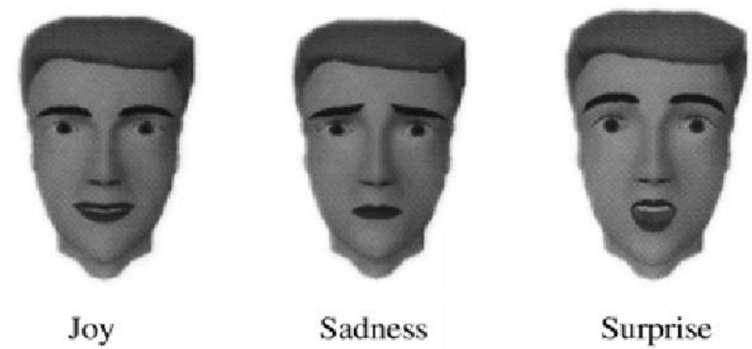
  with triangular patches.
- Benefits of using mesh coding
  - Easy to manipulate an object.
  - Easy to track the motion of a video object after it has been encoded.
- Superior compression

Model Based Coding

- MPEG-4 supports 2 types of models
  - Face object model
    - Synthetic representation of the human face with 3D polygon meshes that can be animated.



The face model



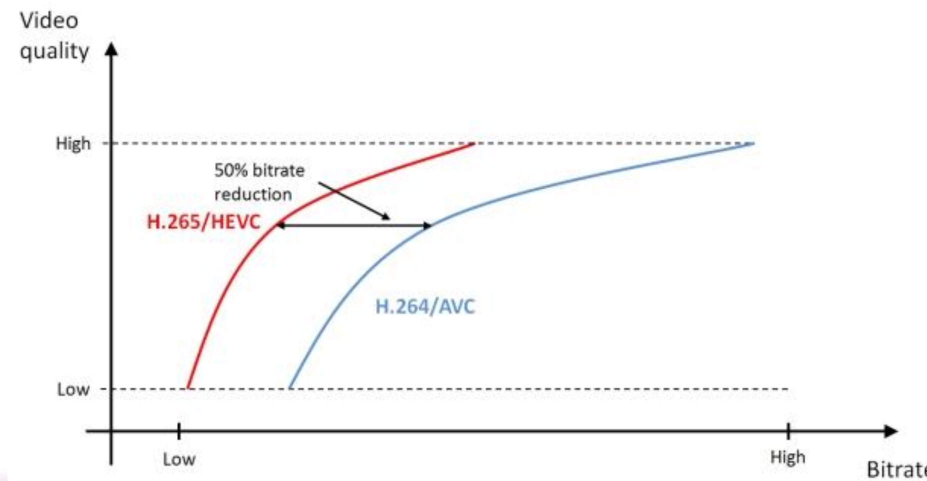
  - Body object model
    - Synthetic representation of a human body with 3D polygon meshes that can be rendered to simulate body movement.

MPEG-4 Summary

- A lot of MPEG-4 examples with interactive capabilities
- Content-based Interactivity
  - Scalability
    - Spatial Scalability
    - Temporal Scalability
  - Sprite Coding
- Improved Compression Efficiency
- Universal Accessibility
  - re-synchronization
  - data recovery
  - error concealment

High Efficiency Video Coding (HEVC)

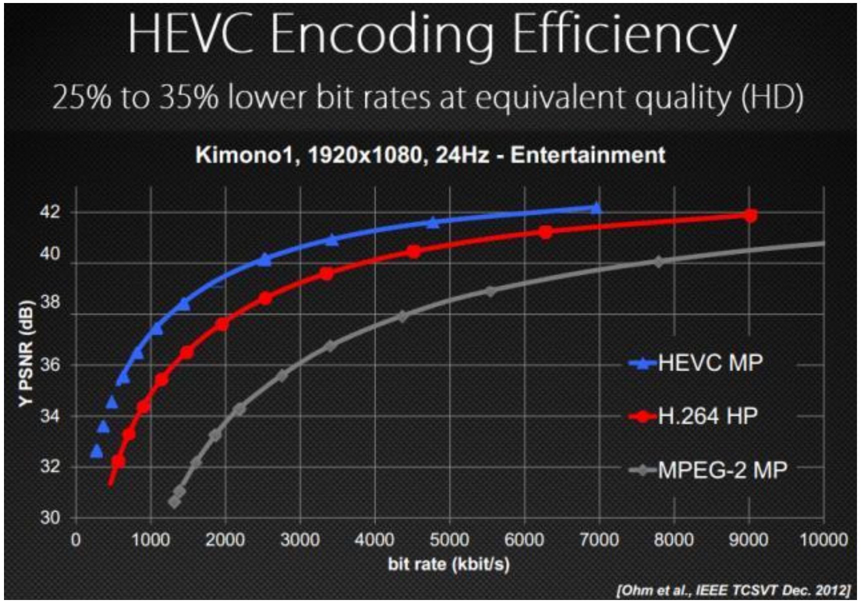
- High Efficiency Video Coding (HEVC) is the latest generation video compression standard.
- This standard was developed by the **ISO/IEC** MPEG and **ITU-T** VCEG, through their JCT-VC
- HEVC is also known as ISO/IEC 23008-2 MPEG-H Part 2 and ITU-T **H.265**



- Have a bit rate reduction of 50% at the same subjective image quality compared to the H.264/MPEG-4 AVC High profile

High Efficiency Video Coding (HEVC)

- It can support 8K UHD and resolutions up to 8192×4320
- HEVC is said to double the data compression ratio compared to H.264/MPEG-4 AVC at the same level of video quality



Subjective video performance comparison[61]

Video coding standard	Average bit rate reduction compared to H.264/MPEG-4 AVC HP			
	480p	720p	1080p	4K UHD
HEVC	52%	56%	62%	64%

H.265/HEVC/MPEG-H Part 2

- Main drivers
  - Get **low bitrate target – target 2:1** over H.264
  - Cheat your eyes – how much can you cut bits and still see the same quality
  - Improve **resolutions (8K by 4K and 4K by 2K) and frame rates**
  - Launch **1080p50/60 services** to compete against BluRay
  - Expect <10x more computational complexity and 2x-3x (decode)

H.265

- Derived from H.264
  - More modes, tools and more interdependencies
  - More efficient search algorithms
  - More complex intra-prediction
  - Macroblocks vs Partitions

AVC	HEVC
16x16 macro-blocks	Coding unit size 64x64 to 8x8
8x8 and 4x4 transform sizes	32x32, 16x16, 8x8 and 4x4 transform sizes

Summary

- Standard formats for video
- Main properties of video
- Deinterlacing
- Different kinds of codecs
- MPEG compression
- Motion estimation
- Main features of MPEG-4
- H.265