4.3. Video Compression

H. 261
H. 263
MPEG
Newer MPEG Standards

- *Reference:* Chapter 6 of Steinmetz and Nahrstedt

- Uncompressed video data are huge. In HDTV, the bit-rate could exceed 1 Gbps. --> big problems for storage and network communications.
- We will discuss both Spatial and Temporal Redundancy Removal -- *Intra-frame and Inter-frame coding*.

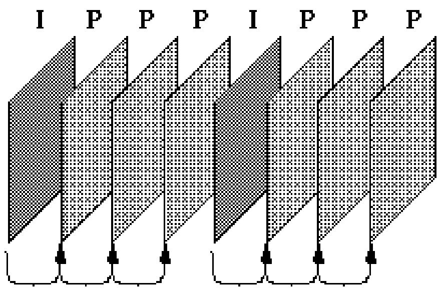
4.3.1. H. 261

- Developed by CCITT (Consultative Committee for International Telephone and Telegraph) in 1988-1990
- Designed for videoconferencing, video-telephone applications over ISDN telephone lines.

  Bit-rate is p x 64 Kb/sec, where p ranges from 1 to 30.

1. Overview of H. 261

- Frame Sequence

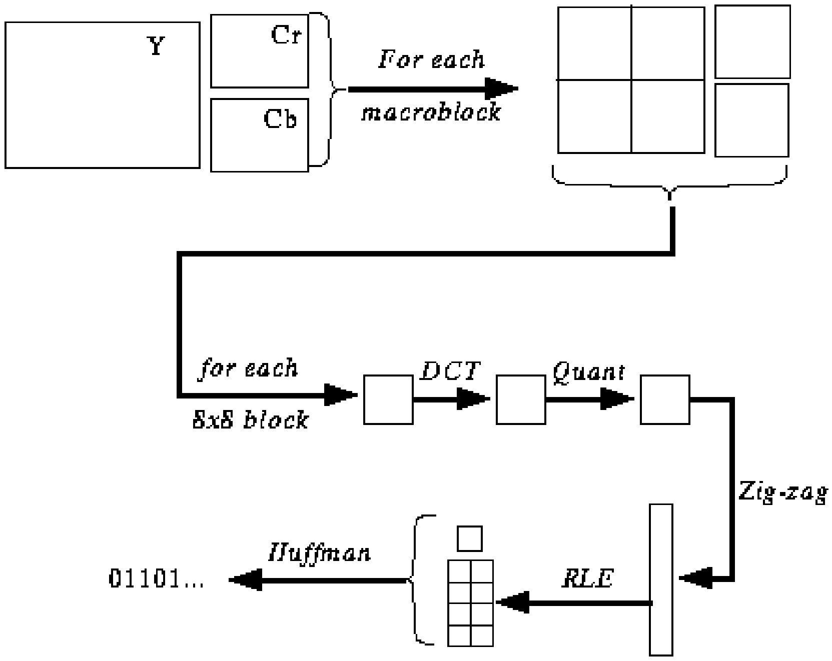


- Frame types are CCIR 601 CIF (352 x 288) and QCIF (176 x 144) images with 4:2:0 subsampling.
- Two frame types: Intra-frames (*I-frames*) and Inter-frames (*P-frames*):

I-frame provides an accessing point, it uses basically JPEG.

P-frames use **"pseudo-differences"** from previous frame ("predicted"), so frames depend on each other.

2. Intra-frame Coding



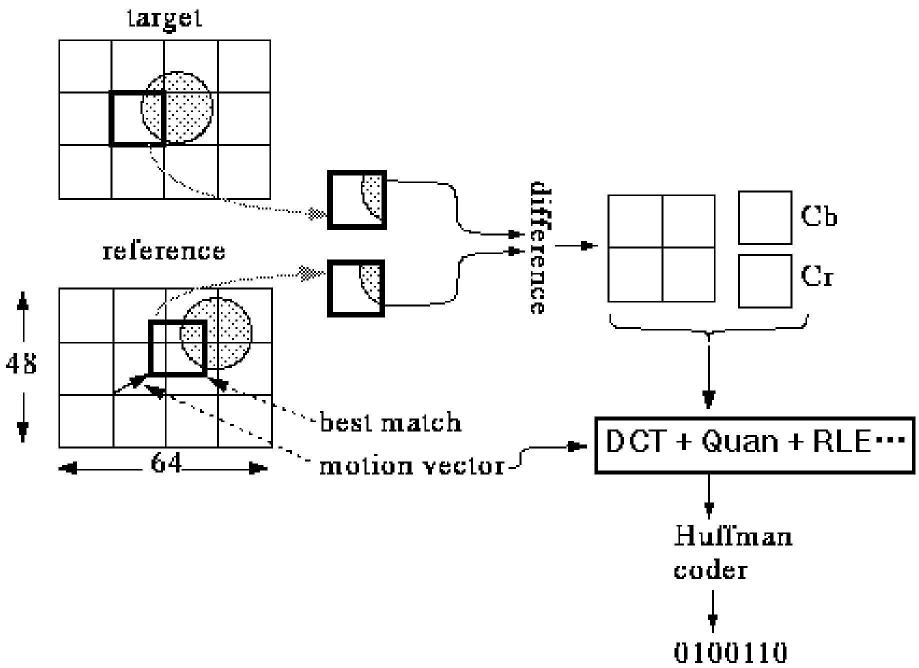
- Macroblocks are 16 x 16 pixel areas on Y plane of original image.

  A macroblock usually consists of 4 Y blocks, 1 Cr block, and 1 Cb block.

- Quantization is by constant value for all DCT coefficients (i.e., no quantization table as in JPEG).

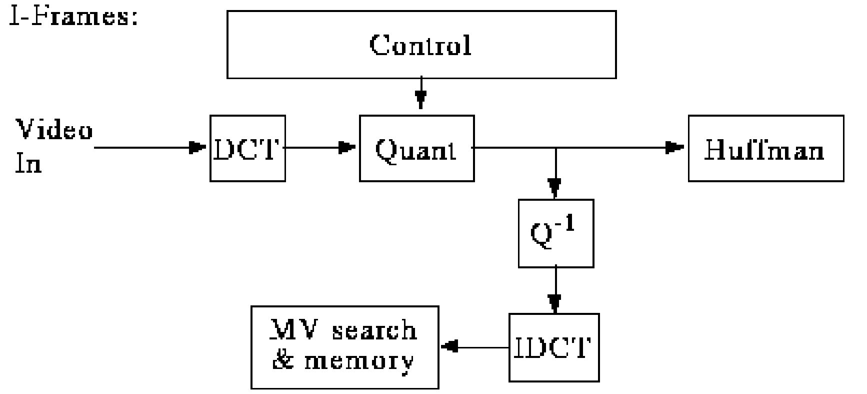
3. Inter-frame (P-frame) Coding

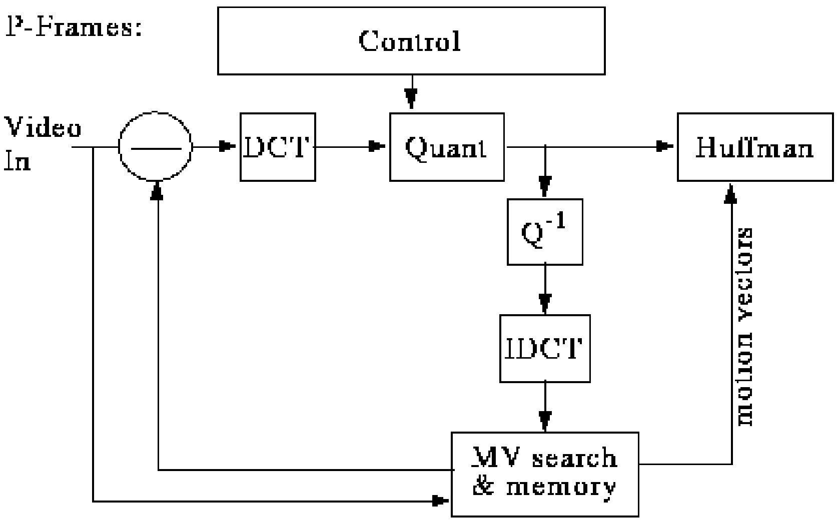
- An Coding Example (P-frame)



- Previous image is called *reference image*, the image to encode is called *target image*.
- **Points to emphasize:**
    1. The difference image (not the target image itself) is encoded.
    2. Need to use the decoded image as reference image, *not* the original.
    3. We're using "Mean Absolute Error" (MAE) to decide best block.
       Can also use "Mean Squared Error" (MSE) = sum(E*E)/N

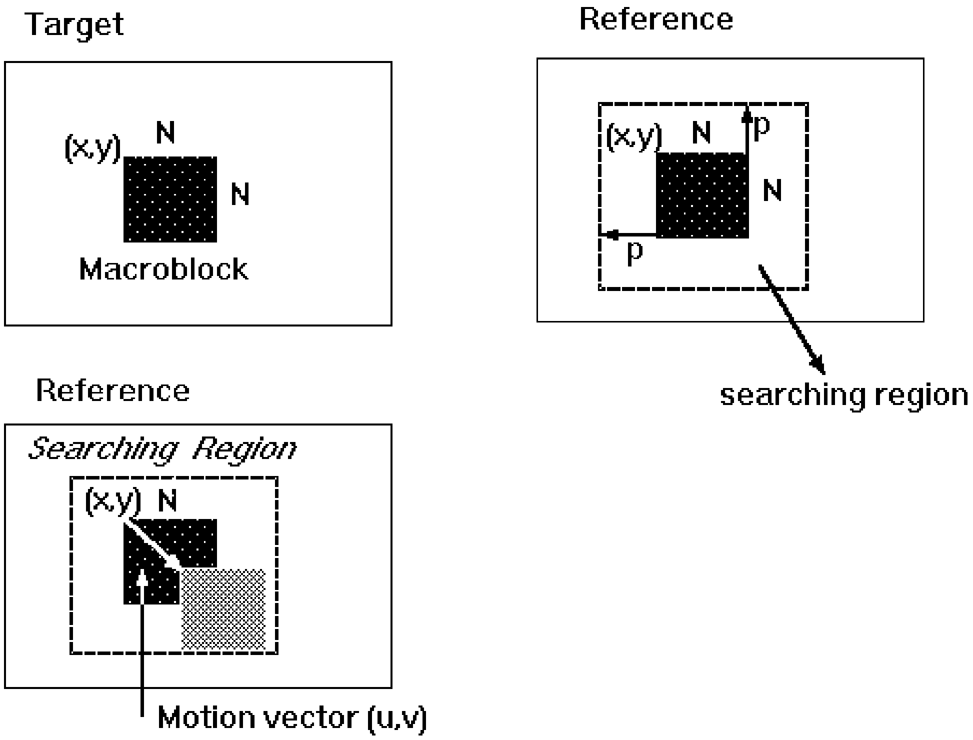
4. H. 261 Encoder





- "Control" -- controlling the bit-rate. If the transmission buffer is too full, then bit-rate will be reduced by changing the quantization factors.
- "memory" -- used to store the reconstructed image (blocks) for the purpose of motion vector search for the next P-frame.

5. Methods for Motion Vector Searches



- C(x + k, y + l) -- pixels in the macro block with upper left corner (x, y) in the Target frame.

  R(x + i + k, y + j + l) -- pixels in the macro block with upper left corner (x + i, y + j) in the Reference frame.

  Cost function is:

$$MAE(i,j) = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} |C(x+k, y+l) - R(x+i+k, y+j+l)|$$

  Where *MAE* stands for *Mean Absolute Error*.

- Goal is to find a vector (u, v) such that *MAE*(u, v) is minimum.

5.1 Full Search Method

Sequentially search the whole [-p, p] region --> very slow

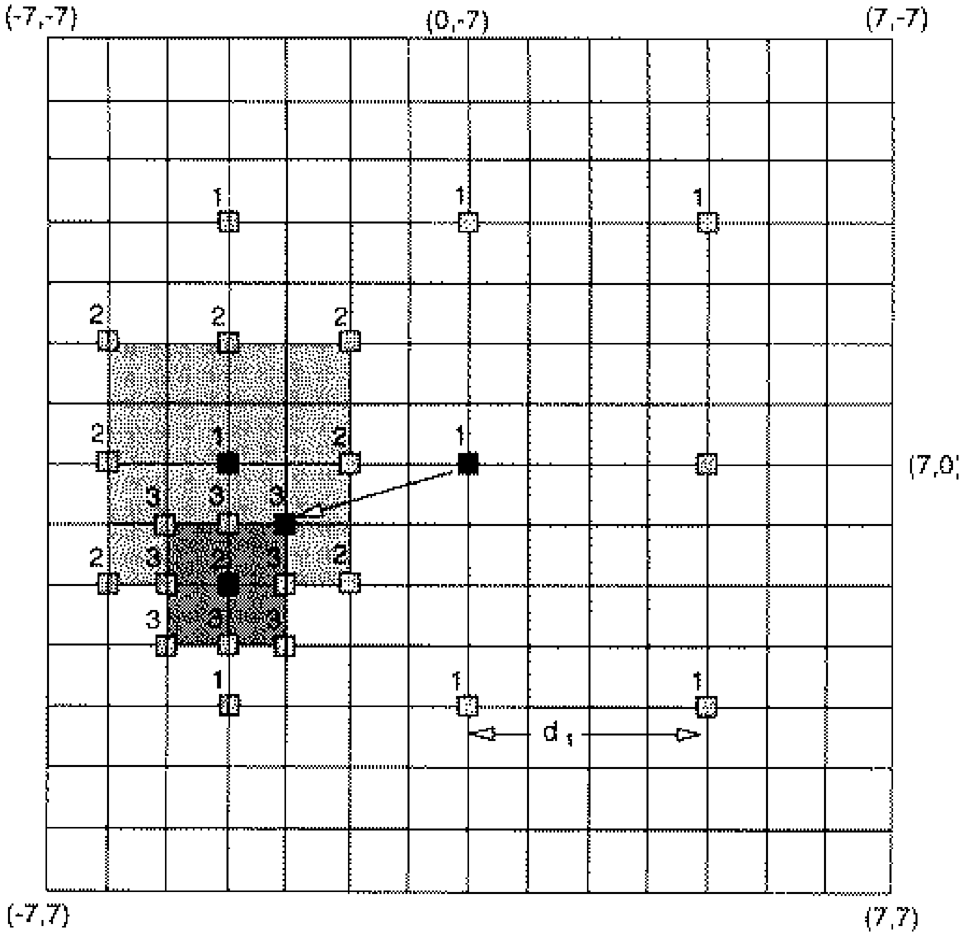
5.2 Two-Dimensional Logarithmic Search

Similar to binary search. MAE function is initially computed within a window of [-p/2, p/2] at nine locations as shown in the figure.

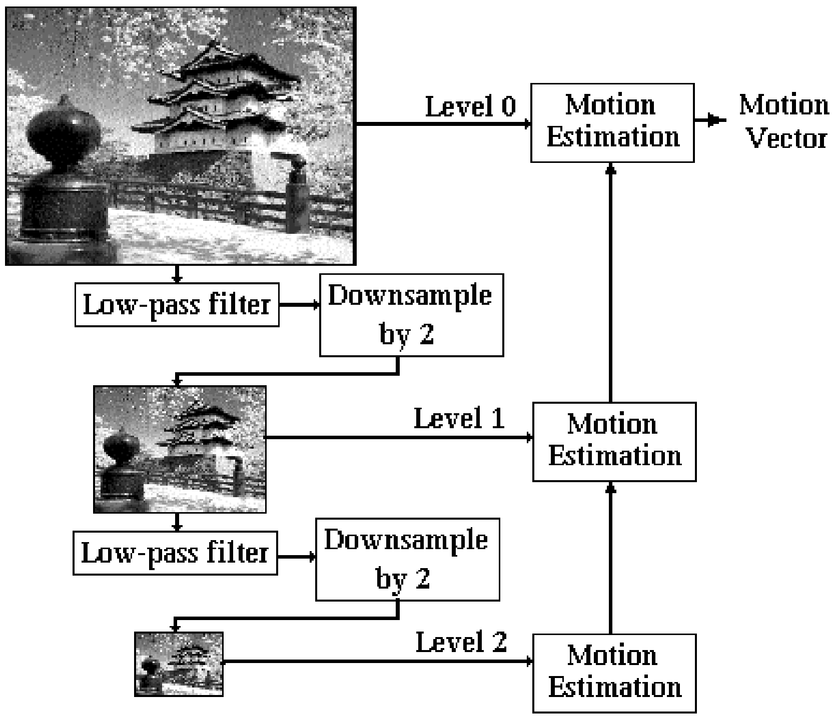
Repeat until the size of the search region is one pixel wide:

1. Find one of the nine locations that yields the minimum MAE

2. Form a new searching region with half of the previous size and centered at the location found in step 1.



5.3 Hierarchical Motion Estimation



1. Form several low resolution version of the target and reference pictures
2. Find the best match motion vector in the lowerest resolution version.
3. Modify the motion vector level by level when going up

6. Some Important Issues

- Avoiding propagation of errors
  1. Send an I-frame every once in a while
  2. Make sure you use decoded frame for comparison
- Bit-rate control
  - Simple feedback loop based on "buffer fullness"

    If buffer is too full, increase the quantization scale factor to reduce the data.

7. Details

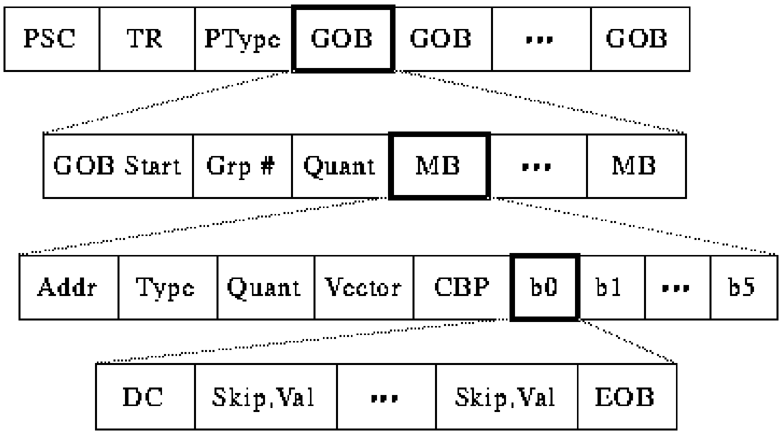
7.1 How the Macroblock is Coded ?

Addr	Type	Quant	Vector	CBP	b0	b1	...	b5

- Many macroblocks will be exact matches (or close enough). So send address of each block in image --> *Addr*

- Sometimes no good match can be found, so send INTRA block --> *Type*
- Will want to vary the quantization to fine tune compression, so send quantization value --> *Quant*
- Motion vector --> *vector*
- Some blocks in macroblock will match well, others match poorly. So send bitmask indicating which blocks are present (Coded Block Pattern, or *CBP*).
- Send the blocks (4 Y, 1 Cr, 1 Cb) as in JPEG.

7.2. H. 261 Bitstream Structure



- Need to delineate boundaries between pictures, so send Picture Start Code --> *PSC*
- Need timestamp for picture (used later for audio synchronization), so send Temporal Reference --> *TR*
- Is this a P-frame or an I-frame? Send Picture Type --> *PType*
- Picture is divided into regions of 11 x 3 macroblocks called Groups of Blocks --> *GOB*
- Might want to skip whole groups, so send Group Number (*Grp #*)
- Might want to use one quantization value for whole group, so send Group Quantization Value --> *GQuant*
- Overall, bitstream is designed so we can skip data whenever possible while still unambiguous.

4.3.2. H. 263

- H. 263 is a new improved standard for low bit-rate video, adopted in March 1996.
  As H. 261, it uses the transform coding for intra-frames and predictive coding for inter-frames.

- Advanced Options:

  - Half-pixel precision in motion compensation
  - Unrestricted motion vectors
  - Syntax-based arithmetic coding
  - Advanced prediction and PB-frames

- In addition to CIF and QCIF, H. 263 could also support SQCIF, 4CIF, and 16CIF.
  The following is a summary of video formats supported by H. 261 and H. 263:

Video Formats Supported

Video format	Luminance Image Resolution	Chrominance Image Resolution	H.261 support	H.263 support	Bit-rate (Mbit/s) (if uncompressed, 30 fps)		Max bits allowed per picture (BPPmaxKb)
					B / W	Color	
SQCIF	128 x 96	64 x 48	Optional	Required	3.0	4.4	64
QCIF	176 x 144	88 x 72	Required	Required	6.1	9.1	64
CIF	352 x 288	176 x 144	Optional	Optional	24.3	36.5	256
4CIF	704 x 576	352 x 288	n/a	Optional	97.3	146.0	512
16CIF	1408 x 1152	704 x 576	n/a	Optional	389.3	583.9	1024

4.3.3. MPEG

1. What is MPEG ?

- "Moving Picture Coding Experts Group", established in 1988 to create standard for delivery of video and audio.

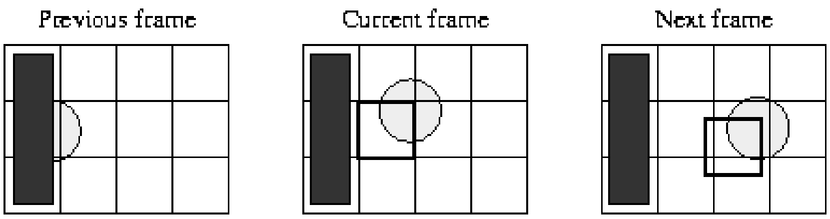
- MPEG-1 Target: VHS quality on a CD-ROM or Video CD (VCD) (352 x 240 + CD audio @ 1.5 Mbits/sec)

- Standard had three parts: Video, Audio, and System (control interleaving of streams)

2. MPEG Video

- Problem: some macroblocks need information not in the previous reference frame.

  Example: The darkened macroblock in Current frame does not have a good match from the Previous frame, but it will find a good match in the Next frame.

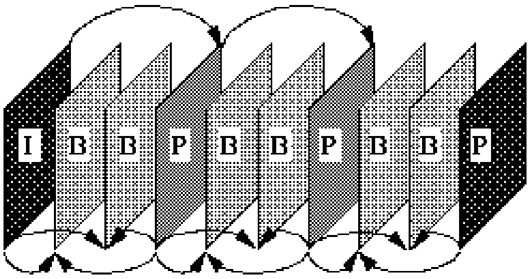


- MPEG solution: add third frame type: bidirectional frame, or *B-frame*

  In B-frames, search for matching macroblocks in both *past* and *future* frames.

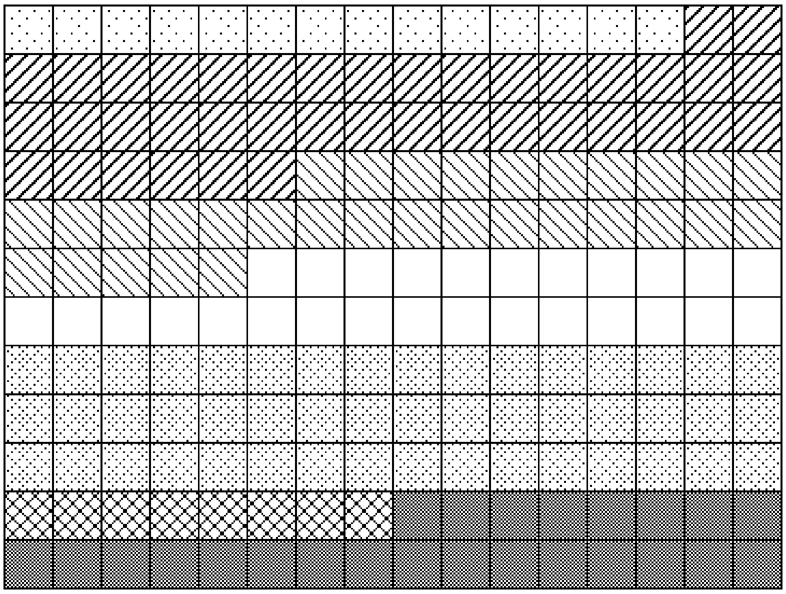
- Typical pattern is IBBPBBPBB IBBPBBPBB IBBPBBPBB

  Actual pattern is up to encoder, and need not be regular.

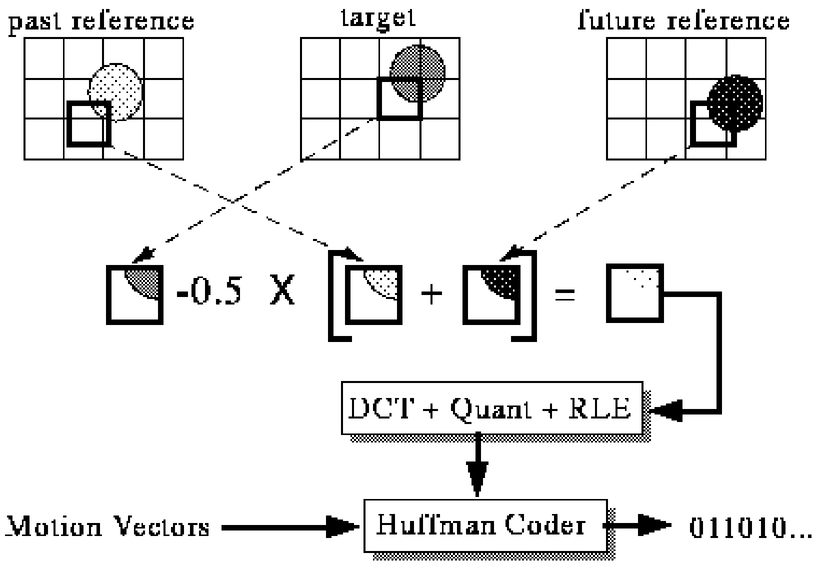


3. Differences from H. 261

- Larger gaps between I and P frames, so need to expand motion vector search range.
- To get better encoding, allow motion vectors to be specified to fraction of a pixel (1/2 pixel).
- Bitstream syntax must allow random access, forward/backward play, etc.
- Added notion of *slice* for synchronization after loss/corrupt data. Example: picture with 7 slices:



- B frame macroblocks can specify *two* motion vectors (one to past and one to future), indicating result is to be averaged.



- Compression performance of MPEG 1

Type	***Size***	***Compression***
I	18 KB	7:1
P	6 KB	20:1
B	2.5 KB	50:1
Avg	4.8 KB	27:1

4. MPEG Video Bitstream

- Click here for details
- Public domain tool **`mpeg_stat`** and **`mpeg_bits`** will analyze a bitstream.

5. Decoding MPEG Video in Software

- Software Decoder goals: portable, multiple display types
- Breakdown of time

Function	***% Time***
Parsing Bitstream	17.4%
IDCT	14.2%
Reconstruction	31.5%
Dithering	24.5%
Misc. Arith.	9.9%
Other	2.7%

4.3.4. Newer MPEG Standards

---

1. MPEG-2

Unlike MPEG-1 which is basically a standard for storing and playing video on a single computer at low bit-rates, MPEG-2 is a standard for digital TV. It meets the requirements for HDTV and DVD (Digital Video/Versatile Disc).

MPEG-2 Level Table:

Level	***size***	***Pixels/sec***	***bit-rate (Mbits)***	***Application***
Low	352 x 288 x 30	3 M	4	consumer tape equiv.
Main	720 x 576 x 30	12 M	15	studio TV
High 1440	1440 x 1152 x 60	96 M	60	consumer HDTV
High	1920 x 1152 x 60	128 M	80	film production

- Other Differences from MPEG-1:

  1. Support both field prediction and frame prediction.
  2. Besides 4:2:0, also allow 4:2:2 and 4:4:4 chroma subsampling
  3. Scalable Coding Extensions: (so the same set of signals works for both HDTV and standard TV)

        - SNR (quality) Scalability -- similar to JPEG DCT-based Progressive mode, adjusting the quantization steps of the DCT coefficients.

        - Spatial Scalability -- similar to hierarchical JPEG, multiple spatial resolutions.

        - Temporal Scalability -- different frame rates.

    4. Frame sizes could be as large as 16383 x 16383

    5. Non-linear macroblock quantization factor

    6. Many minor fixes (see MPEG FAQ for more details)

- MPEG-3: Originally planned for HDTV, got folded into MPEG-2

2. MPEG-4

- Version 1 approved Oct. 1998, Version 2 to be approved Dec. 1999.

- Originally targeted at very low bit-rate communication (4.8 to 64 Kb/sec), it now aims at the following ranges of bit-rates:
    - video -- 5 Kb to 5 Mb per second
    - audio -- 2 Kb to 64 Kb per second

- It emphasizes the concept of *Visual Objects* --> Video Object Plane (VOP)
    - objects can be of arbitrary shape, VOPs can be non-overlapped or overlapped
    - supports content-based scalability
    - supports object-based interactivity
    - individual audio channels can be associated with objects

- Good for video composition, segmentation, and compression; networked VRML, audiovisual communication systems (e.g., text-to-speech interface, facial animation), etc.

- Standards being developed for shape coding, motion coding, texture coding, etc.

3. MPEG-7 (Multimedia Content Description Interface)

- International Standard due by July 2001.

- MPEG-7 is a content representation standard for multimedia information search, filtering, management and processing.

    - *Descriptors* for multimedia objects and *Description Schemes* for the descriptors and their relationships.
    - A *Description Definition Language (DLL)* for specifying Description Schemes.
    - For visual contents, the lower level descriptions will be color, texture, shape, size, etc., the higher level could include a semantic description such as "this is a scene with cars on the

highway".

Further Exploration

The MPEG Home Page, A good MPEG FAQ, MPEG Resources on the Web.

Last Updated: ***3/15/00***