Multimedia Systems and Applications

Video Compression

James Wang

Some notes are adapted from Prof. Lawrence A. Rowe's original slides at http://www.BMRC.Berkeley.EDU/~larry

Representations

- **Composite**
  - NTSC - 6MHz (4.2MHz video), 29.97 frames/second
  - PAL - 6-8MHz (4.2-6MHz video), 50 frames/second
- **Component**
  - Separation video (luma, chroma) - svhs, Hi8mm
  - RGB, YUV, YIQ, …
  - YC_BC_R - used for most compressed representations
- **Separation video called "s-video"**

2

Analog Video Representations

- **NTSC**
  - Y = 0.299R + 0.587G + 0.114B
  - I = 0.596R - 0.275G - 0.321B
  - Q = 0.212R - 0.523G + 0.311B
  - composite = Y + I*cos*(Fsc t) + Q*sin*(Fsc t)
- **PAL**
  - Y = 0.299R + 0.587G + 0.114B
  - U = 0.492(B-Y)
  - V = 0.877(R-Y)
  - composite = Y + U*sin*(Fsc t) + V*cos*(Fsc t)

3

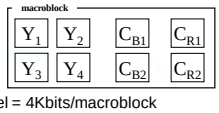
Digitizing

- **Analog TV is a continuous signal**
- **Digital TV uses discrete numeric values**
  - Signal is sampled
  - Samples are quantized
  - Small, discrete regions are digitized
- **Image represented by pixel array**

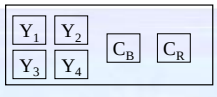
4

Digital Video Block Structure

- **4:2:2 YC_BC_R**
  - 16x16 macroblock
  - 8x8 pixel blocks
  - 8 bits/sample = 16 bits/pixel = 4Kbits/macroblock



- **4:1:1 YC_BC_R**
  - 3Kbits/macroblock
  - 12 bits/pixel



5

What is Video Data Rate?

- **Digital**
  - 720x483 = 347,760 pixels/frame
  - 4:2:2 sampling gives 695,520 bytes/frame
  - 21 MB/sec (167 Mbs)
  - 4:4:4 sampling gives 250 Mbs
- **ATV (MPEG MP@ML)**
  - 1280x720 = 921,600 pixels/frame
  - 4:2:0 sampling gives 1,382,400 bytes/frame
  - 41 MB/sec (328 Mbs)

(Note: MPEG coded streams are 1.5-80 Mbs)

6

What is Video Data Rate (cont.)?

- **ATSC (720P)**
  - 720x1280 = 921,600 pixels per frame
  - 4:2:2 sampling = 1,843,200 bytes per frame
  - 24 fps = 44,236,800 bytes per second
  - ➔ 44 MB/s = 354 Mbs
- **ATSC (studio 1080I)**
  - 1080x1920 = 2,073,600 pixels per frame
  - 4:4:4 sampling = 6,220,800 bytes per frame
  - 30 fps = 186,624,000 bytes per second
  - ➔187MB/s = 1.5 Gbs

7

Human Perception

- **What is smooth motion**
  - Depends on source material
  - Most action is perceived as smooth at 24 fps
- **Human most sensitive**
  - Low frequencies
  - Changes in luminance and blue-orange axis
- **Vision emphasizes edge detection**
  - Strong bias to horizontal and vertical lines
- **Visual masking by large luminance changes**

8

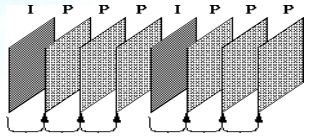
H. 261

- **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.**

9

Overview of H. 261

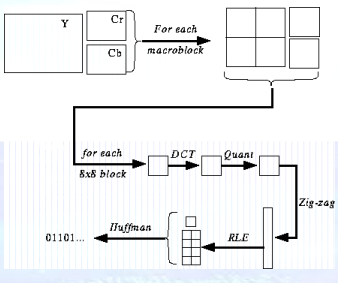
- **Frame Sequence**



- **Frame types are CCIR 601 CIF (352 x 288) and QCIF (176 x 144) images with 4:2:0 sub-sampling.**
- **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.**

10

Intra-frame Coding



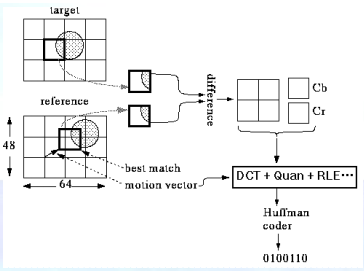
11

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).**

12

Inter-frame (P-frame) Coding



A Coding Example (P-frame)

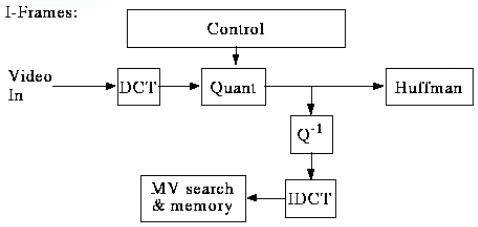
13

Inter-frame (P-frame) Coding

- **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

14

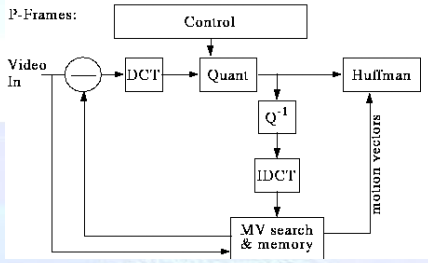
H. 261 Encoder



- 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.

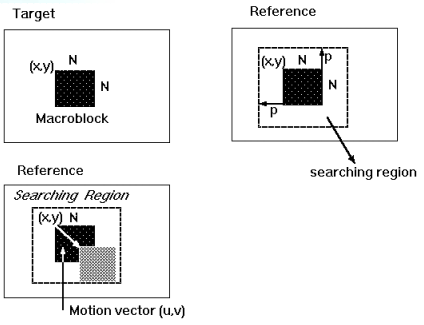
15

H. 261 Encoder



16

Methods for Motion Vector Searches



17

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.

18

Methods for Motion Vector Searches

Full Search Method

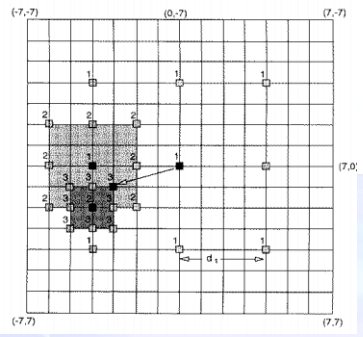
- Sequentially search the whole [-p, p] region --> very slow

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 centred at the location found in step 1.

19

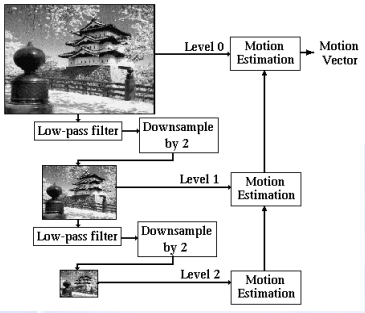
Methods for Motion Vector Searches



20

Methods for Motion Vector Searches

Hierarchical Motion Estimation



21

Hierarchical Motion Estimation

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

22

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.

23

Details

How the Macroblock is Coded ?

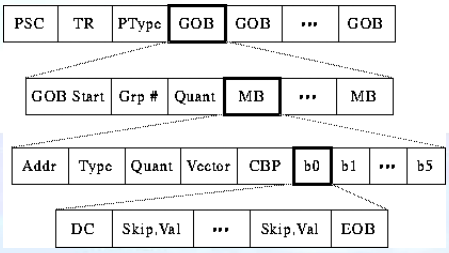
Addr	Type	Quant	Vector	CBP	b0	b1	...	b5

- **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.***

24

Details

H. 261 Bitstream Structure



25

Details

- 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
- Generally, bitstream is designed so that we can skip data whenever possible and still remain unambiguous.

26

H. 263

- H. 263 is an 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 page is a summary of video formats supported by H. 261 and H. 263:

27

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) B / W	Bit-rate (Mbit/s) (if uncompressed, 30 fps) Color	Max bits allowed per picture (BPPmax, Kb)
SQCIF	128 x 96	64 x 48	n/a	Required	3	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	512
16CIF	1408 x 1152	704 x 576	n/a	Optional	389.3	583.9	1024

28

MPEG

What is MPEG ?

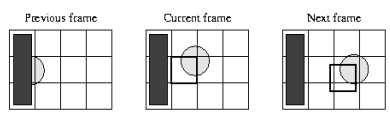
- "Motion Picture Coding Experts Group", established in 1988 to create standard for delivery of video and audio.
- MPEG-1 Target: VHS quality on a CD-ROM or Video CD (VCD) (352 x 240 + CD audio @ 1.5 Mbits/sec)
- Standard had three parts: Video, Audio, and System (control interleaving of streams)

29

MPEG-1 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.

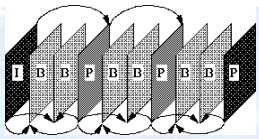
30

MPEG-1 Video

- Typical pattern is

  IBBPBBPBB IBBPBBPBB IBBPBBPBB

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



31

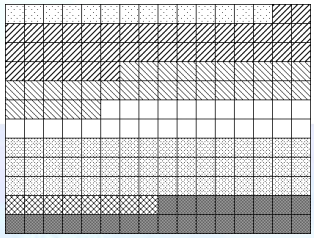
Differences from H. 261

- 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:

32

Differences from H. 261

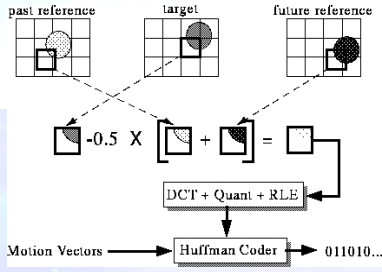
- Example of 7 slices:



33

B Frame

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



34

MPEG-1 Compression

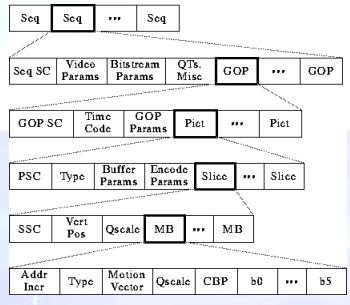
Compression performance of MPEG 1

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

35

MPEG Video Bitstream

- Public domain tool such as mpeg_stat and mpeg_bits is able to analyze an MPEG bitstream.



36

MPEG Video Bitstream

- **Sequence Information**
  1. *Video Params* include width, height, aspect ratio of pixels, picture rate.
  2. *Bitstream Params* are bit rate, buffer size, and constrained parameters flag (means bitstream can be decoded by most hardware)
  3. Two types of QTs: one for intra-coded blocks (I-frames) and one for inter-coded blocks (P-frames).
- **Group of Pictures (GOP) information**
  1. *Time code*: bit field with SMPTE time code (hours, minutes, seconds, frame).
  2. *GOP Params* are bits describing structure of GOP. Is GOP *closed*? Does it have a dangling pointer *broken*?
- **Picture Information**
  1. *Type*: I, P, or B-frame?
  2. *Buffer Params* indicate how full decoder's buffer should be before starting decode.
  3. *Encode Params* indicate whether half pixel motion vectors are used.

37

MPEG Video Bitstream

- **Slice Information**
  1. *Vert Pos*: what line does this slice start on?
  2. *QScale*: How is the quantization table scaled in this slice?
- **Marcoblock (MB) Information**
  1. *Addr Incr*: number of MBs to skip.
  2. *Type*: Does this MB use a motion vector? What type?
  3. *QScale*: How is the quantization table scaled in this MB?
  4. *Coded Block Pattern (CBP)*: bitmap indicating which blocks are coded.

38

Decoding MPEG Video in Software

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

Function	% Time
Parsing Bitstream	17.40%
IDCT	14.20%
Reconstruction	31.50%
Dithering	24.50%
Misc. Arith.	9.90%
Other	2.70%

39