CSEP 590
Data Compression
Autumn 2007

Video Compression

Human Perception of Video

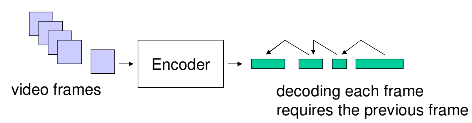
- 30 frames per second seems to allow the visual system to integrate the discrete frames into continuous perception. Even 10 frames per second is acceptable.
- If distorted, nearby frames in the same scene should have only small details wrong.
  - A difference in average intensity is noticeable
- Compression choice when reducing bit rate
  - skipped frames cause stop action
  - lower fidelity frames may be better

Applications of Digital Video

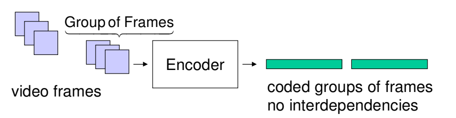
- Teleconference or video phone
  - Real-time video
  - Very low delay (1/10 second is a standard)
- Live Broadcast Video
  - Modest delay is tolerable (seconds is normal)
  - Error tolerance is needed.
- Video-in-a-can (DVD, Video-on-Demand)
  - Random access to compressed data is desired
  - Encoding can take a lot of time
- Decoding must always be at at least the frame rate.

Video Encoding

Frame-by-Frame coding



Group-of-Frames coding



Coding Techniques

- Frame-by-frame coding with prediction
  - Very low bit rates
  - low delay
  - Not error resilient
- Group-of-frames coding
  - Higher bit rates – within a group prediction is used
  - Error resilient
  - Random Access
  - Higher delay

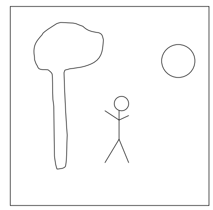
Digital Video Data

- CCIR 601 (4,2,2 scheme)
  - 13.5 MHz sample rate for luminance channel
  - 6.75 MHz sample rate for each of two chrominance channels
  - 8 bits per sample is a bit rate of 27 x 8 = 216 Mb per second
  - MPEG-SIF – ½ sample rate for luminance and ¼ for chrominance – 81 Mb per second
- CIF (Common Interchange Format)
  - 288 x 352 pixels per frame for luminance channel
  - 144 x 176 pixels per frame for each of two chrominance
  - 8 bits per pixel and 30 frames per second gives 48.7 Mb per second
  - QCIF (Quarter - CIF) is ¼ the data or 12.2 Mb per second.

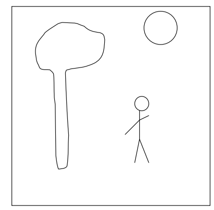
High Compression Ratios Possible

- Nearby frames are highly correlated. Use the previous frame to predict the current one.
- Need to take advantage of the fact that usually objects move very little in 1/30 th of a second.
  - Video coders use motion compensation as part of prediction

Motion Compensation

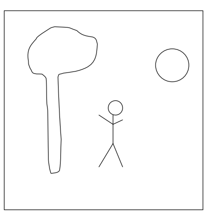


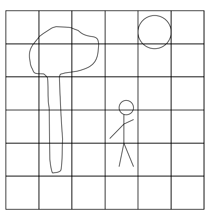
Previous Frame



Frame

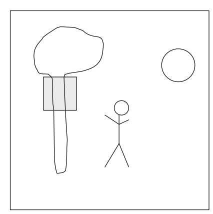
Block Based Motion Compensation

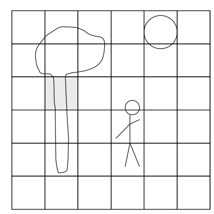




motion compensation blocks

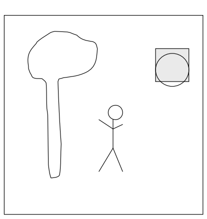
Motion Vectors

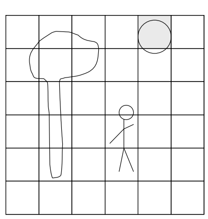




motion vector = (0,0)

Motion Vectors





motion vector = (20,5)
20 down and 5 to right

Motion Compensation

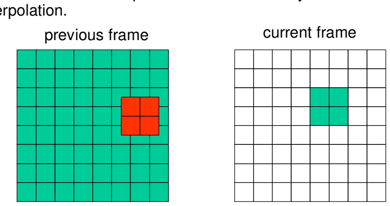
- For each motion compensation block
  - Find the block in the previous decoded frame that gives the least distortion.
  - If the distortion is too high then code the block independently. (intra block)
  - Otherwise code the difference (inter block)
- The previous decoded frame is used because both the encoder and decoder have access to it.

Issues

- Distortion measured in squared error or absolute error
  - Absolute error is quicker to calculate
- Block size
  - Too small then too many motion vectors
  - Too large then there may be no good match
- Searching range to find best block
  - Too large a search range is time consuming
  - Too small then may be better matches
  - Prediction can help.
- Prediction resolution
  - Full pixel, half-pixel, quarter-pixel resolution
  - Higher resolution takes longer, but better prediction results

Fractional Motion Compensation

- Half or quarter pixel motion compensation may achieve better predictions.
- Fractional motion compensation is achieved by linear interpolation.

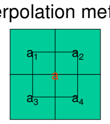


Half pixel motion compensation

Linear Interpolation

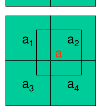
- Calculate an interpolated pixel as the average of overlapping pixels.
- Better interpolation methods exist.

Half



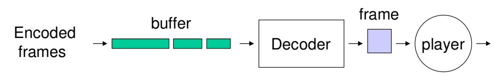
$$a = \frac{a_1}{4} + \frac{a_2}{4} + \frac{a_3}{4} + \frac{a_4}{4}$$

Quarter



$$a = \frac{3a_1}{16} + \frac{9a_2}{16} + \frac{a_3}{16} + \frac{3a_4}{16}$$

Rate Control



- Buffer is filled at a constant rate (almost).
- Buffer is emptied at a variable rate.

Underflow Problem

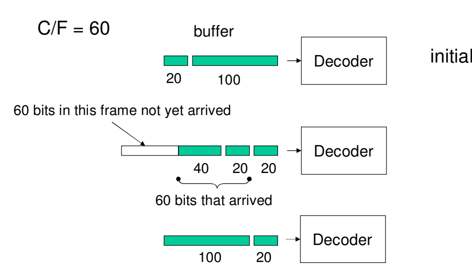
- Set up
  - Constant rate channel at C bits per second
  - Frame rate F in frames per second.
  - b_i is the number of bits in compressed frame i
  - Initial occupancy of buffer B
- B_i is the number of bits in the buffer at frame i

$$B_0 = B$$

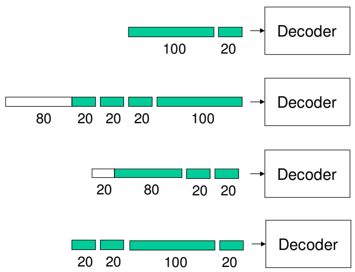
$$B_{i+1} = B_i + C/F - b_i$$

- Buffer should never empty $B_i \geq 0$ for all i

Example



Example



Rate Control

- The rate control buffer is seeded to allow for variability in number of bits per frame and variability in channel rate.
- Causes of variability in bits per frame.
  - Encoder does not predict well how many bits will be used in a frame - scalar quantization.
  - Encoder allocates more bits in frames that are hard to encode because they are not predicted well - scene changes.
- Causes of channel rate variability
  - Congestion on the internet

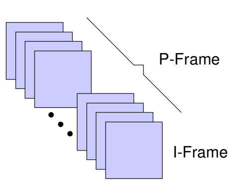
Rate Control Algorithms

- On-line solution
  - Send a few frames to seed the buffer
  - Encoder simulates the buffer, should the buffer threaten to empty start sending more frames at lower fidelity or skip frames (decoder will interpolate skipped frames).
- Off-line solution
  - Attempt to allocate bits to frames to assure even fidelity.
  - Seed the buffer with enough frames to prevent underflow.

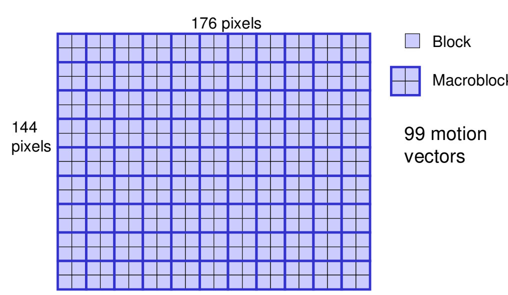
H.261

- Application – low bit rate streaming video
- Frame-by-frame encoder
- DCT based with 8x8 coding block
  - Uses JPEG style coding
- Motion compensation based on 16x16 macroblocks.
- Half pixel motion compensation

H.261

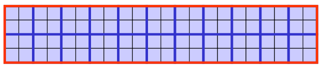


H.261(QCIF)



H.261

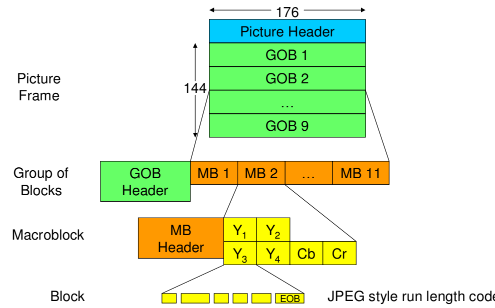
- Within a group of blocks (GOB) prediction is used with motion vectors for coding.



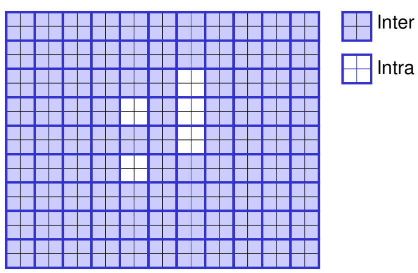
GOB

Also called a slice

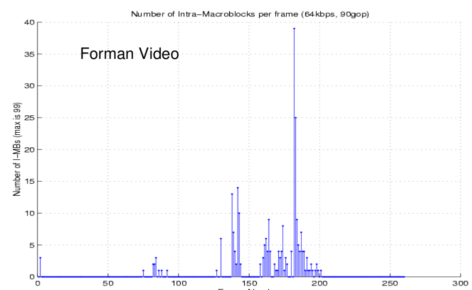
H.261 Organization



P-Frame



Intra-Macroblock Distribution

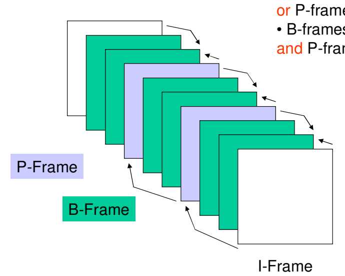


MPEG-1

- Application – Video coding for random access
- Group-of-frames encoder
- DCT based with 8x8 coding block
  - Uses JPEG style coding
- Motion compensation based on 16x16 macroblocks.
- Forward and Backward Prediction within a group of frames

MPEG-1

- P-frames predicted by I-frames or P-frames
- B-frames predicted by I-frames and P-frames



Orders

Display Order

10	9	8	7	6	5	4	3	2	1

Coding/Decoding Order

9	8	10	6	5	7	3	2	4	1

Added delay is one frame time

MPEG-1 Notes

- Random access unit = Group-of-Frames
  - Called GOP for group-of-pictures
- Error resilient
  - B-frames can be damaged without propagation
- Added delay
  - Coding order different than display order
- Encoding time consuming
  - Suitable for non-interactive applications

Beyond MPEG-1

- MPEG-2
  - Application independent standard
- MPEG-4
  - Multimedia applications
  - Model based coding
- H.263
  - More error resilience

Newest Trends

- H.264
  - Just out in 2003, many new features
  - Quarter pixel motion compensation
  - Variable size motion blocks
  - Multiple frame prediction
- 3-D Wavelet Coding
  - Third dimension is time
  - 3-D SPIHT has been implemented
  - Delay is large because GOP is large
- GTV
  - Group testing for video
  - Bits per frame can be controlled enabling off-line rate control to succeed.