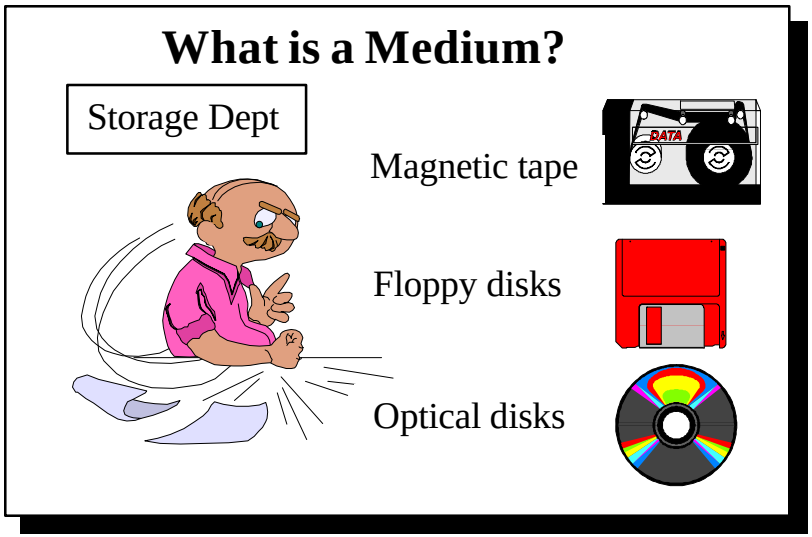
Multimedia: An Introduction



Raj Jain

Pro ces

Raj Jain is now at
Washington University in Saint Louis
Jain@cse.wustl.edu
http://www.cse.wustl.edu/~jain/

- Local Multimedia
- Video Fundamentals
- Compression methods
- Compression Standards: JPEG, MPEG,...

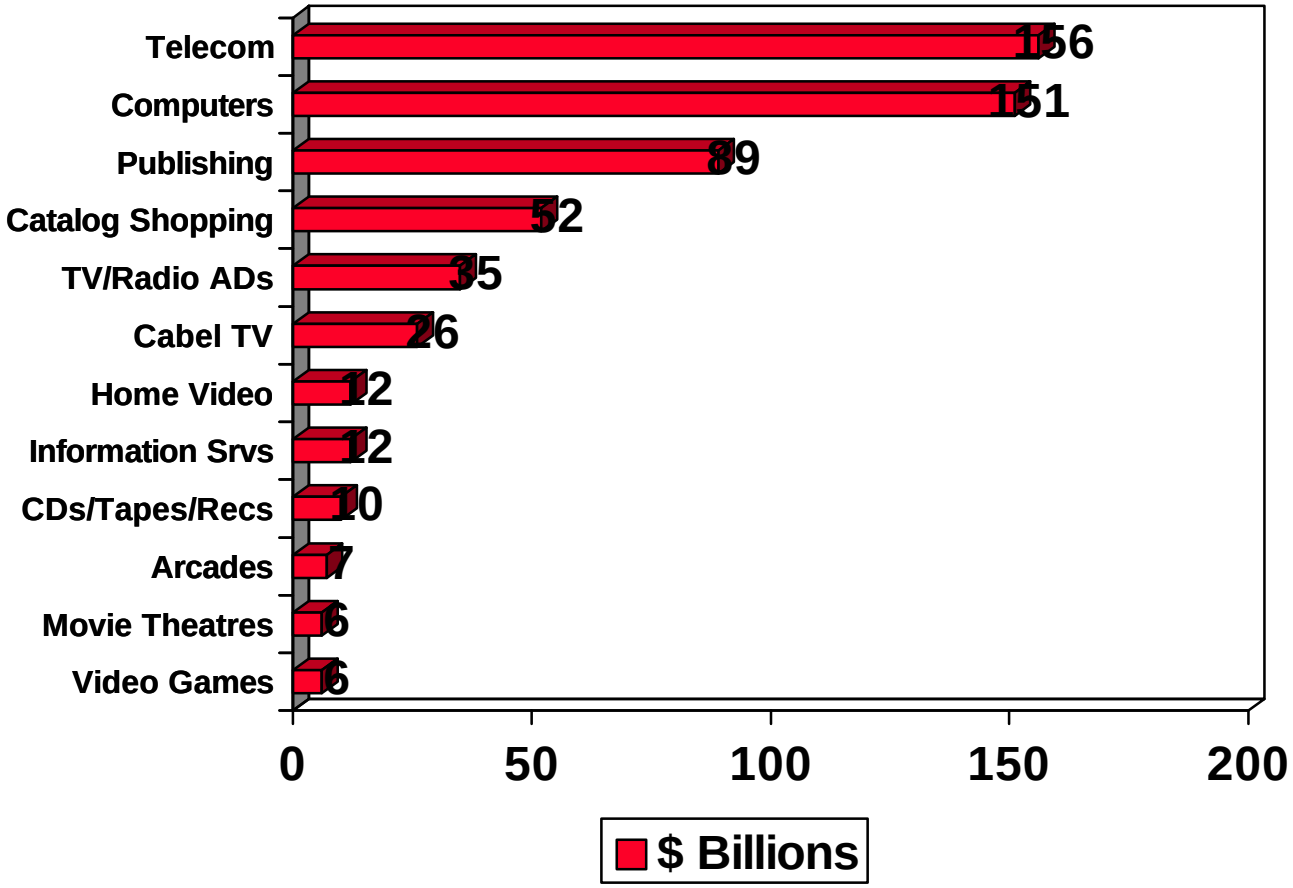
Multimedia: Key Ingredients

- Two or more media: Text, Image, Audio, Video
- Synchronization among the media
- Human users: senders, receivers or both
- Interaction: Search, fast forward, conversation

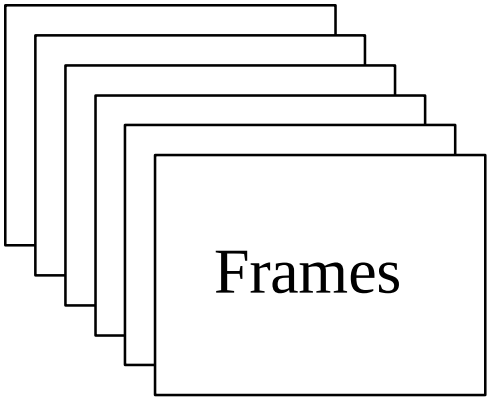
Multimedia Applications

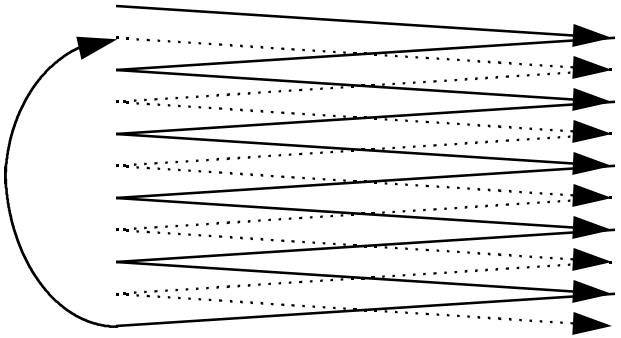
- Distance learning
- Healthcare: Telemedicine, Remote Diagnosis
  Remote access to health database
- Telepresence: Virtual proximity. Can control remote camera.
  - Real estate purchasers can drive down the virtual city
- Video Cruising: Virtually go through coworker's offices, coffee rooms
- Collaboration: Shared-screen systems

Multimedia Market



Television Formats

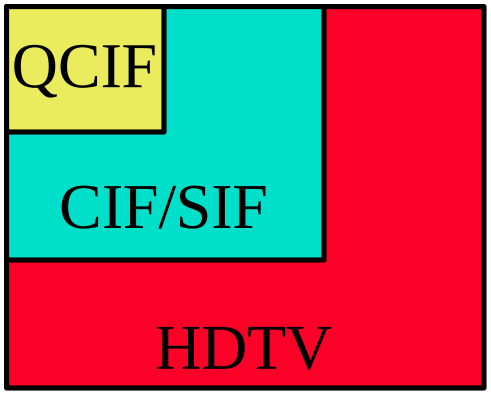




Lines

- US/Japan: 525/60 NTSC
  (National Television Standards Committee)
  Europe: 625/50 PAL (Phase Alternating Line),
  French/E. Europe: 625/50 SECAM (Seq. Color and Memory)
- 525 lines/frame (481 contain video, 22/vertical interval/field can be used for teletext, closed caption, timestamp, etc.)
- 60 Fields/sec. Fields = Odd lines, even lines (Interlaced)
  30 frames/sec

Video Formats



In 525 lpf countries:

- Video Conferencing
  $\Rightarrow$ Common Intermediate Format (CIF) = 352×240
- Source Input Format (SIF) by MPEG = CIF by CCITT
- Video Phone $\Rightarrow$ Quarter CIF = 176 ppl × 120 lpf
- Broadcast, Cable, VCR $\Rightarrow$ CCIR 601 = 704×480
- HDTV = 1280×720 (Progressive=NI) or 1440×960 (Interlace)
- 24 bits/pixel $\Rightarrow$ 704×480×30×24 = 242 Mbps, 109 GB/h

Video Compression Considerations

- High compression
  - 100-200 normal, 2500 possible with fractal methods
- Decoding must be simple $\Rightarrow$ Asymmetric
  - H.261, JPEG, AVI, QuickTime are Symmetric
  - DVI, MPEG are asymmetric
- Allow real time encoding/decoding
- Implementable in software, if possible
- Allow random-access, fast forward/reverse
- Transmission Error-tolerant
- Integratable with text, graphic files for storage/mailing
- Scalable: Allow a range of video quality

Compression Standards

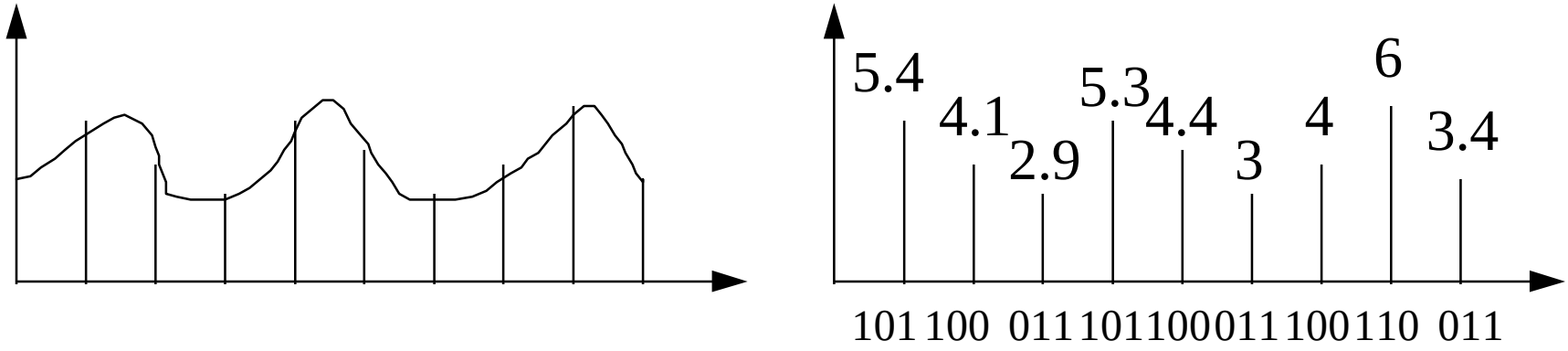
	Symmetric	Asymmetric
Hardware based	JPEG H.261	DVI/Intel MPEG CD-I
Software based	AVI QuickTime Ultimedia	

Video Compression Techniques

- Reducing the frame rate, lines/frame, pixels/line, bits/pixel
  Used for teleconferencing. Not acceptable for entertainment
- Redundancies: Spatial, Spectral, Temporal
- No loss entropy coding:
  - Run-length coding: $000011111111..100000=0^{4}1^{35}0^{5}$
  - Huffman coding: Frequent characters with fewer bits
- Discrete Cosine Transform
  - Only low frequency components are quantized
- Differential Pulse Code Modulation (DPCM)
  - Sections with large interframe differences are quantized
- Motion Compensation (Inter-frame)
  - Differences from predicted motion are quantized

Pulse Code Modulation (PCM)

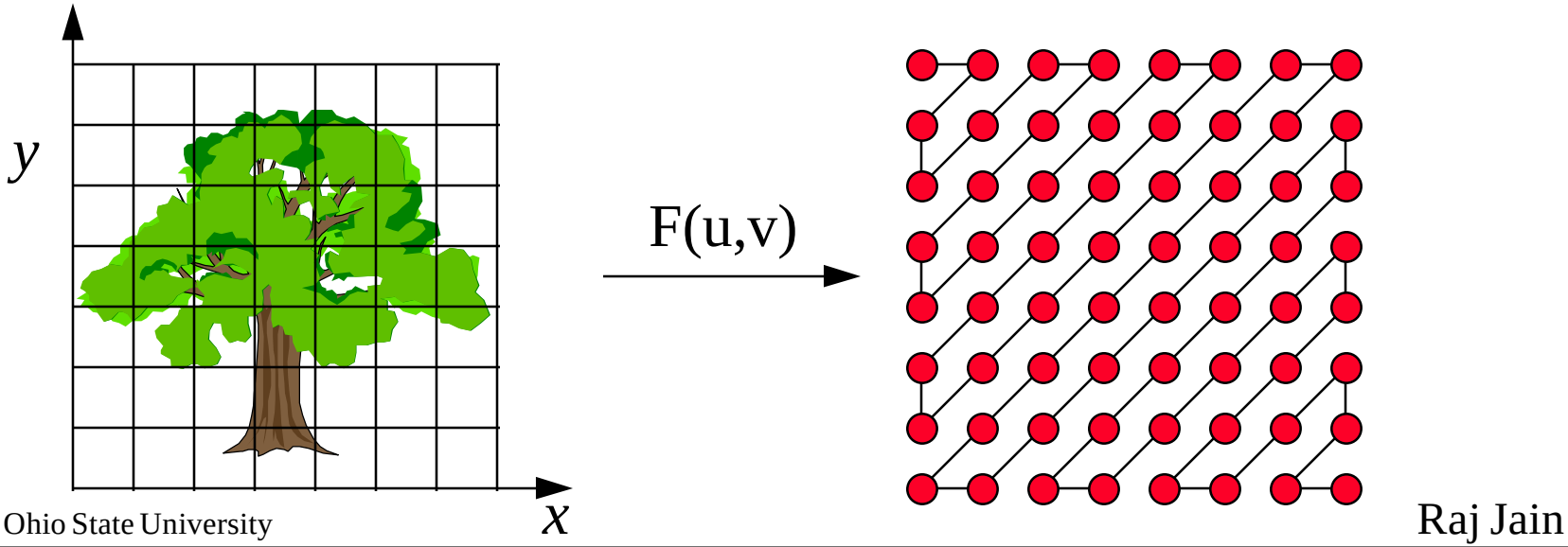
- Nyquist Sampling Theorem: Sample analog signal at twice the signal frequency



- Quantization Step = Unit of quantization
- Quantization error
  = Error due truncation or rounding
  Acceptable value depends upon noise and human perception

Discrete Cosine Transform

- Time domain to frequency domain transform in two dimensions
- With infinite precision DCT is lossless
- DC coefficient coded differentially from previous block
- Other coefficients entropy encoded using zig-zag sequence
- Quantization table specifies step size for each element $\Rightarrow$ Loss



Discrete Cosine Transform

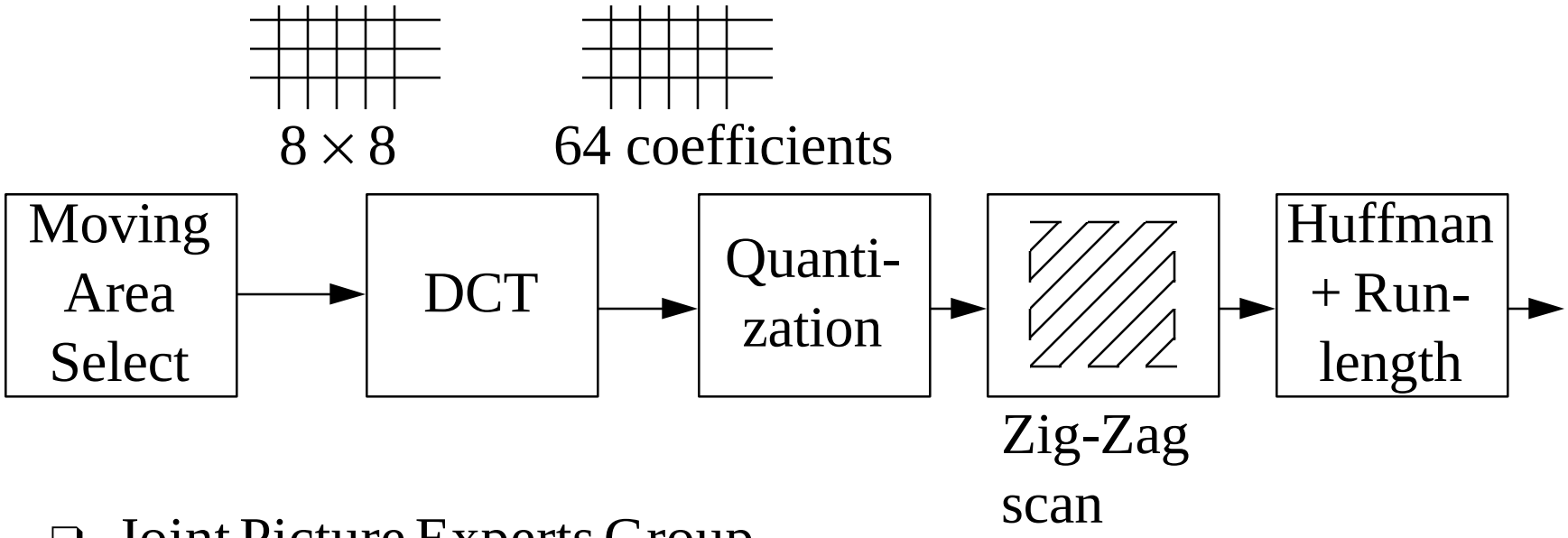
- Take 8×8 block:
- Compute 64 coefficients
- For u=0,1,...,7 and v=0,1,...,7

  $F(u,v)=0.25\, C(u)\, C(v)\, \{ \Sigma^{7}_{x=0}\Sigma^{7}_{y=0} f(x,y) \cos[(2x+1)u\, \pi/16] \cos[(2y+1)v\, \pi/16]\}$

  $f(u,v)=0.25\ \{\Sigma^{7}_{x=0}\Sigma^{7}_{y=0} C(u)\, C(v)\, F(x,y) \cos[(2x+1)u\, \pi/16] \cos[(2y+1)v\, \pi/16]\}$

  Where, C(i)=1/sqrt(2) for i=0
  =1 otherwise
- F(0,0) is the DC, others are AC
- F(u,v) is small for high values of u,v

JPEG



- Joint Picture Experts Group.
  Joint ⇒ ISO+CCITT
- DCT + Quantization + Entropy
- Entropy = Huffman or Arithmetic
- S/W coding possible: 10 s for 640X480X30 image on 68040

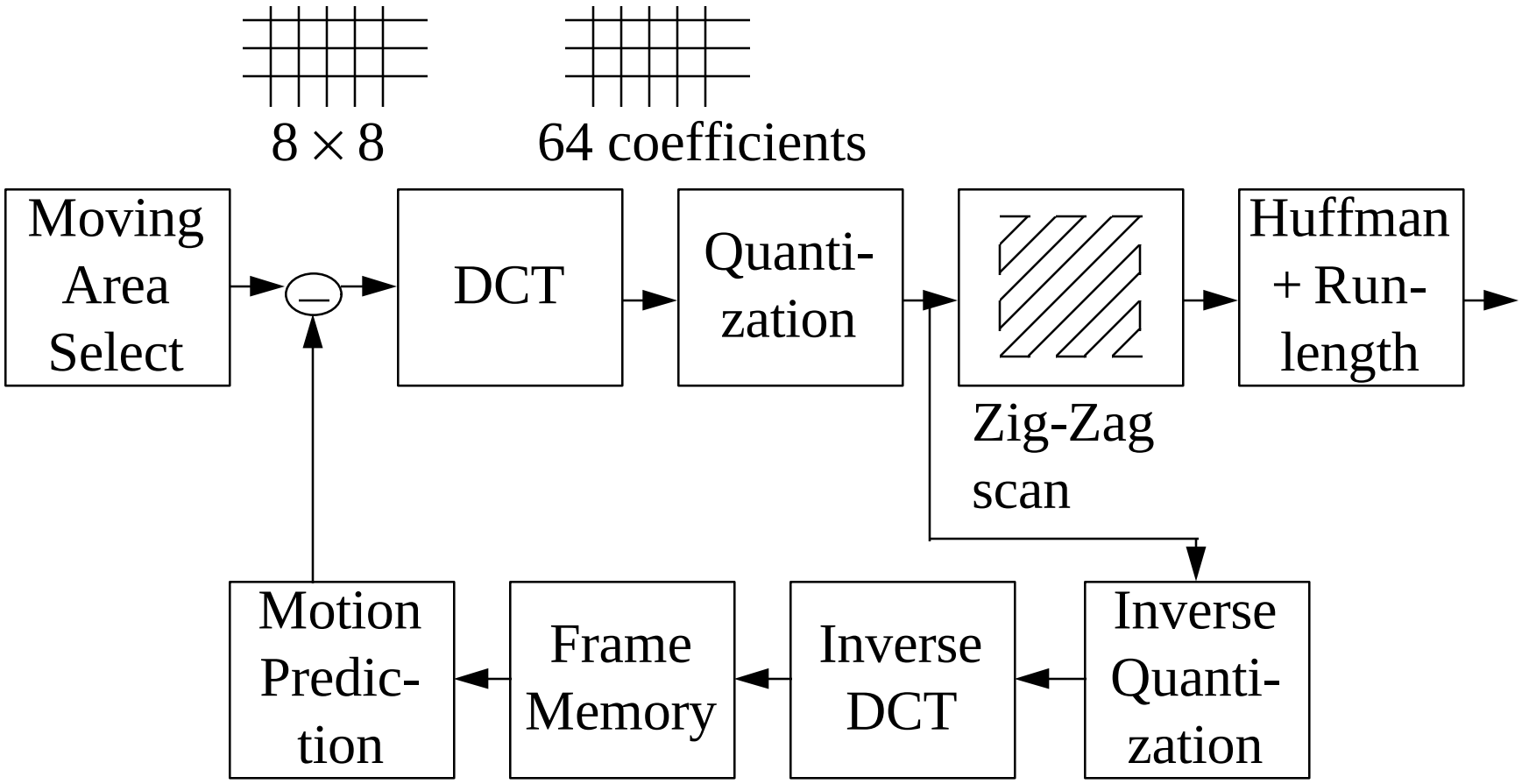
JPEG for Color

- Color Images = Multiple components
- The components may be interleaved
- Each component is divided into regions
- Each region consists of several 8X8 blocks
- All blocks of one region of one component followed by one region of second component and so on
- CODECs have up to 4 quantization tables and 4 entropy tables active

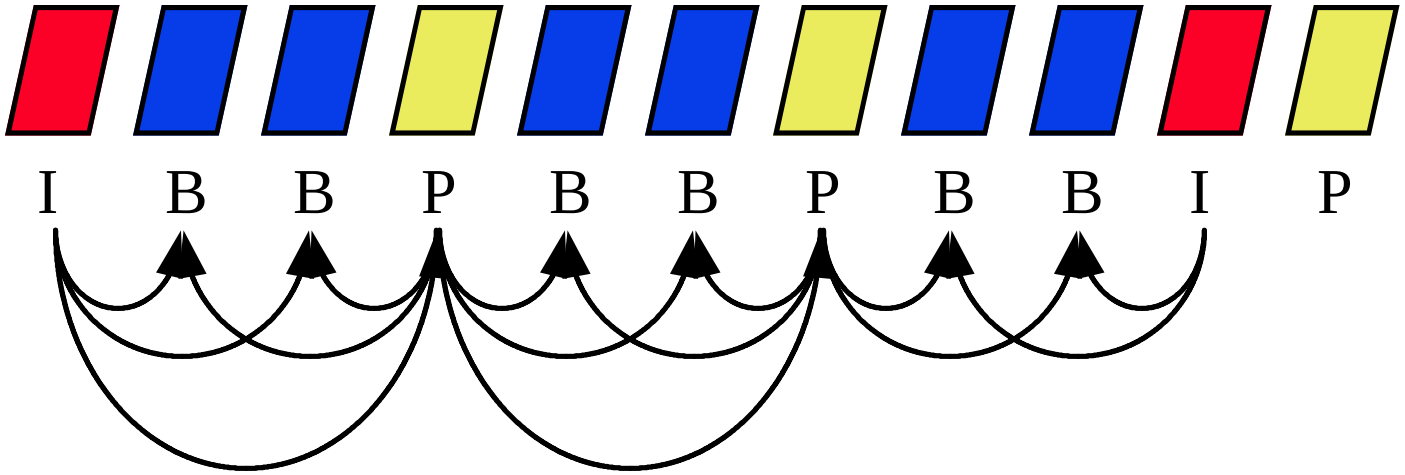
JPEG for Video

- Many vendors use JPEG for video
- Although designed only for images
- No interframe coding
  $\Rightarrow$ Fast random access
- 221.184 Mbps for 640X480X30X24
- 1:50 compression $\Rightarrow$ 4.4 Mbps
- Quarter window $\Rightarrow$ 1 Mbps

Motion Prediction



MPEG-1



- Inter-frame Coding
- I = Intraframe coded ⇒ Allows random access
- P = Predicted from previous P or I
- B = Bidirectional prediction
- Bandwidth Allocation: I:P:B::5:3:1
- Uses Motion prediction, DCT coding, quantization, entropy

MPEG-1 (Continued)

- ISO/IEC JTC1/SC29/WG11 CD 11172
  Motion Picture Experts Group
- VCR quality video for digital storage 1.5-2 Mbps
- Asymmetric: coding more complex than decoding
- Compression ratios of 200 $\Rightarrow$ VCR quality at 1.5-2 Mbps
- Ratio of 50 $\Rightarrow$ Broadcast quality at 6 Mbps
- Specifies rules for multiplexing audio/video streams
- 32 kbps to 384 kbps mono/stereo audio
- Synchronization using 33 bit timestamps 90 kHz clock

MPEG-2

- MPEG Phase 2: Broadcast quality or better
- 15 Mbps for NTSC, 60 Mbps for HDTV, 4-15 Mbps for VCR
- Compability: Backward/forward, Different picture formats
- Specifies 3 profiles: Simple, Main, Next
  Simple does not use bidirectional frames
  (No storage $\Rightarrow$ Low cost, Low delay)
  Main profile designed for most common applications
  Next profile supports hierarchical coding
- Each profile in 4 levels
  - High Level Type 1: HDTV, 1152 lines/frame, 1920 pixels/line, 60 fps $\Rightarrow$ 62.7 Mpps = 60 Mbps
  - High Level Type 2: HDTV, 1152 lpf, 1440 ppl, 60 fps $\Rightarrow$ 47 Mpps = 60 Mbps

MPEG-2 (Continued)

- Main Level: Full size (576 lpf, 720 ppl), 30 fps
  $\Rightarrow$ 10.4 Mpps = 15 Mbps
- Low Level: 1/4 Size (288 lpf, 352 ppl), 30 fps
  $\Rightarrow$ 2.5 Mpps = 4 Mbps

- Video industry will use Main profile/main level
- Cable industry may use Simple profile/main level
- US HDTV will use Main profile/High level
- European HDTV will use Next profile/High level
- MPEG-2 allows field or frame coding,
  MPEG-1 allows only frame coding
- MPEG-2 allows 3 chrominance formats: 4:4:4, 4:2:2, 4:2:0
  MPEG-1 allows only 4:2:0

ITU-T H.261 Standard

- Started in 1984 for m×384 kbps
- Later p×64 kbps p = 1, 2,...,30
- VCR quality video
- Resynchronization at receiver ⇒ Allows transmission over independent parallel channels
- DCT + Quantization + Motion-predicted compression
- Color represented by Y, Cb, Cr
- Y = Luminance, Cb = Chrominance = B-Y,
  Cr = Chrominance = R-Y
- Sampling rates 4:2:2

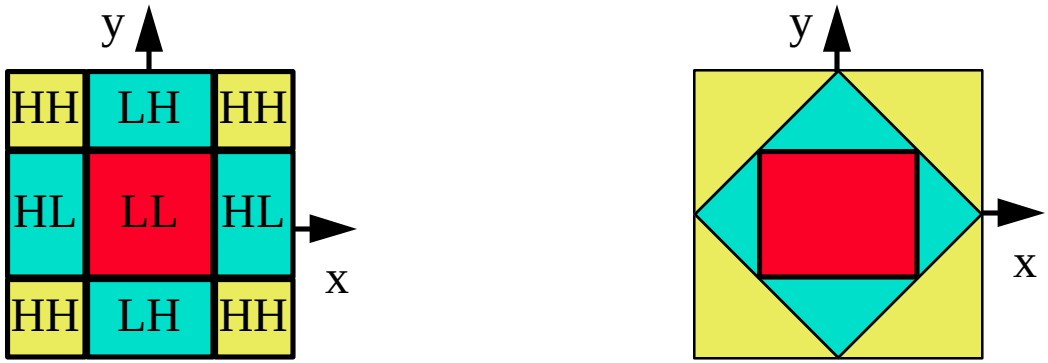
ITU-T H.261 Standard

- Common Intermediate Format (CIF) for Video conferencing
- Quarter CIF (QCIF) for desktop telephony

	QCIF	CIF
Luminance (y)	144 lpf, 180 ppl	288 lpf, 360 ppl
Chrominance (Cr,Cb)	72 lpf, 90 ppl	144 lpf, 180 ppl
Uncompressed rate	9.115 Mbps	29.97 Mbps

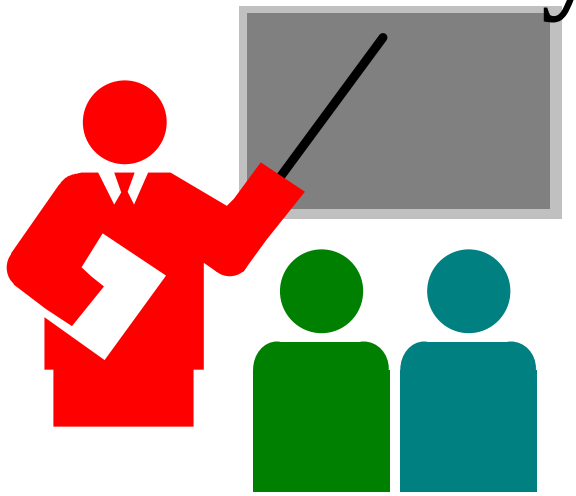
- Allows dropping frames $\Rightarrow$ reducing frame rate
- At 10 fps, QCIF requires compression ratio of 50 for 64 kbps
- p = 1 or 2 $\Rightarrow$ Face only (Video Phone)
- p = 6 for teleconferencing

Subband Hierarchical Coding



- Hierarchical coding
- Image signal is split into frequency bands
  Each band is processed differently
  Example: 4 bands LL, LH, HL, HH
- LL consists of low-frequency band in the horizontal scan and low-frequency in the vertical scan
- LL contains most useful information, coded using DCT
- Other bands coded using simple quantizers with a dead zone
- Dead zone results in long strings of zeros

Summary



- Video formats: Lines, pixels
- Compression techniques: Huffman, run-length, DCT, Motion prediction
- Compression Standards: JPEG, MPEG, H.261
- Hierarchical coding

References: Books

- D. Minoli and R. Keinath, "Distributed Multimedia Through Broadband Communications Services," Artech House, 1993, 312 pp.
- B. Lee, M. Kang, and J. Lee, "Broadband Telecommunications Technology," Artech House, Boston, 1993
- N. Ohta, "Packet Video," Artech House, Boston, 1994
- T. Trowt-Bayard, “Video Conferencing: The Whole Picture,” Flatiron Publishing, Inc., New York, 1994
- J. Keyes, Ed., “The McGraw-Hill Multimedia Handbook,” McGraw-Hill, New York, 1994, Chapter 32.
- J. Ozer, “Video Compression for Multimedia,” Academic Press, Cambridge, MA, 1995.

References: Papers

- IEEE Spectrum, Special issue on Digital Television, April 1995. [History and explanation of HDTV]
- Communications of ACM, Vol. 34, No. 4, Apr 1991.
- B. Furht, "Multimedia Systems: An Overview," IEEE Multimedia, Spring 1994, pp. 47-59. [JPEG, MPEG, pX64, Networking]
- T. Chiang and D. Anastassiou, "Hierarchical Coding of Digital Television," IEEE Communications Magazine, May 1994, pp. 38-45. [MPEG-2]
- E. Fox, "Advances in Interactive Digital Multimedia Systems," IEEE Computer, October 1991, pp. 9-21.

References: Standards

- [JPEG] JTC1/SC2/WG10, "Digital Compression and Coding of Continuous-Tone Still Images"
- [MPEG-1] JTC1/SC2/WG11, "Coding of Moving Pictures and Associated Audio for Digital Storage Media up to about 1.5 Mbit/s," ISO/IEC DIS 11172, 1992.
- [MPEG-2] "Information Technology- Generic Coding of Moving Pictures and Associated Audio Recommendations H.262," ISO/IEC 13818-2 CD, Nov 1993, Seol.
- [H.261] CCITT SG XV, "Video coder/decoder for audiovisual services at p × 64 kbps"

References: Online

- JPEG image compression: Frequently asked questions, http://www.cis.ohio-state.edu/hypertext/faq/usenet/jpeg-faq/faq.html
- Compresion: Frequently asked questions, ftp rtfm.mit.edu, files:/pub/usenet/news.answers/compression-faq/part[1-3]
- Distributed Multimedia Survey: Standards, http://cuiwww.unige.ch/OSG/MultimediaInfo/mmsurvey/standards.html#G.711
- MPEG: Frequently asked questions, http://www.crs4.it/HTML/LUIGI/MPEG/mpegfaq.html

Acronyms

- CD-I — Compact Disc Interactive
- DCT — Discrete Cosine Transform
- DVI — Digital Video Interactive
- HDTV — High Definition Television
- IVD — Integrated Voice Data
- JBIG — Joint Bilevel Image Experts Group
- JPEG — Joint Photographic Experts Group
- MHEG — Multimedia and Hypermedia Experts Group
- MIDI — Musical Instrument Digital Interface
- MPEG — Motion Picture Experts Group
- NTSC — National Television Standards Committee
- PAL — Phase Alternating Line
- SECAM — Sequential Color And Memory
- SMPTE — Society for Motion Picture and Television Engineers