18-796

Multimedia Communications:
Coding, Systems, and Networking

Prof. Tsuhan Chen
tsuhan@ece.cmu.edu

Introduction

What is Multimedia?

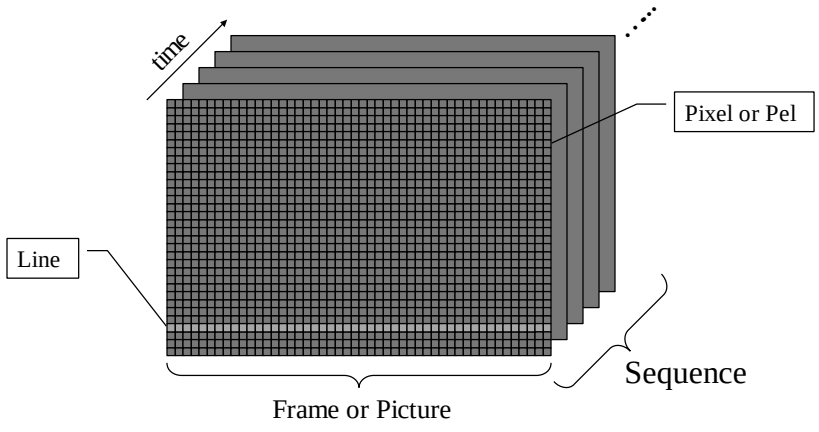
- Multimedia
  - Text, speech, music, audio, image, graphics, video, and many more...
- Multimedia research
  - Compression/Coding
  - Standards: H.series, MPEG, DAVIC, VRML...
  - Networking: streaming, QoS, VBR…
  - Implementation: architectures, low-power, MMX...
  - Databases: retrieval and indexing
  - Human-machine interface

Multimedia Communications...

- Coding
  - Compression algorithms for audio, images, and video
- Systems
  - Integrating audio, video, and other components
- Networking
  - Transmission of multimedia over networks

Coding

Images and Video



Why Compression?

- Still images
  - 512 × 512 × 3 bytes/pel = 6.29 Mbits
  - Needs 112 sec at 56 kbits/s
- Video

	Pels/line	Lines	Frames/s	Bytes/pel	Bit rate
Video Telephony (CIF)	352	288	10	1.5	12.2 Mbits/s
Broadcast TV (ITU-R 601 4:2:2)	720	480	30	2	166 Mbits/s
HDTV	~1280	~720	60	2	885 Mbits/s

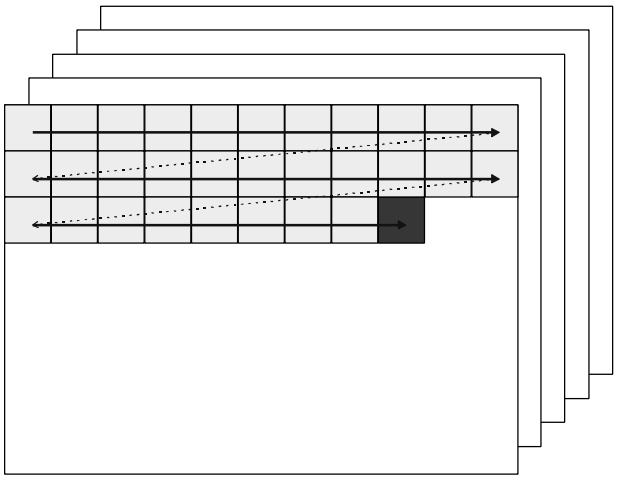
How to Compress?

- Removal of statistical redundancy
  - Spatial redundancy: intra coding
  - Temporal redundancy: inter coding
  - Non-stationary statistics of images/video
- Human visual system
  - Spatial masking
    - Flat vs. texture areas
  - Temporal masking
    - Scene cuts
- Lossless compression vs. lossy compression

Spatial Redundancy: Intra Coding

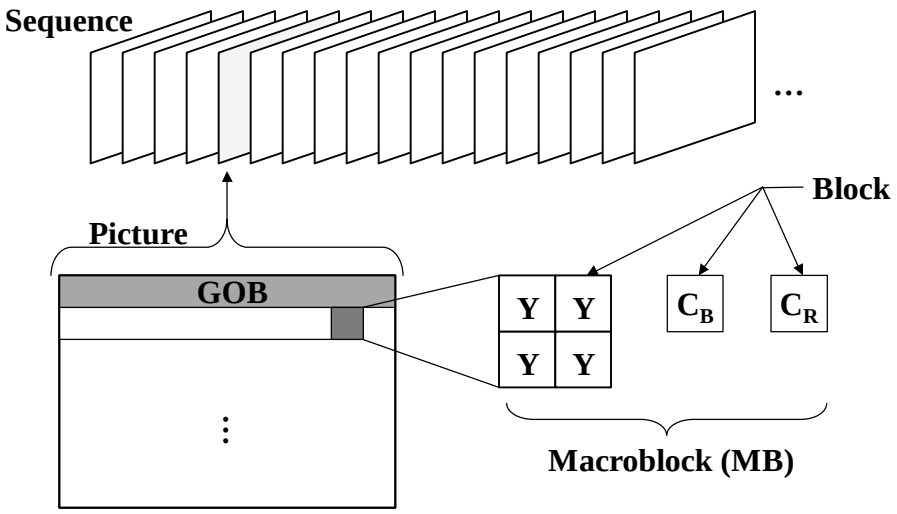
- Block-based schemes
  - Transform coding
  - Vector quantization (VQ)
- Non block-based schemes
  - Subband/Wavelet coding
  - Pyramid coding

Block-Based Coding



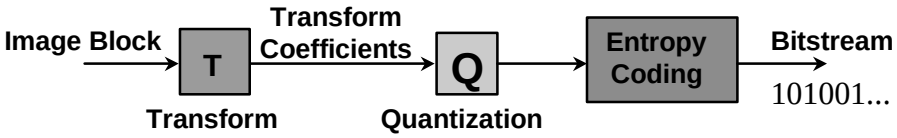
Typical block size: 8×8 or 16×16

Block-Based Coding

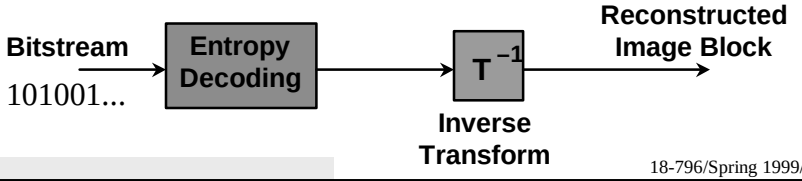


Transform Coding

Encoder



Decoder



Selection of Transform

- Decorrelation of transform coefficients
  - To remove redundancy
- Energy concentration
  - To allow selection of coefficients
  - Easy for entropy coding (cf. run-length coding)
- Discrete Cosine Transform (DCT)
  - Close to optimal for typical images
  - Well-known algorithm
  - Used in JPEG, H.26x, MPEG

2D Discrete Cosine Transform

$$\begin{bmatrix} Y_{mn} \end{bmatrix} = \begin{bmatrix} C_{mn} \end{bmatrix}^T \begin{bmatrix} X_{mn} \end{bmatrix} \begin{bmatrix} C_{mn} \end{bmatrix}$$

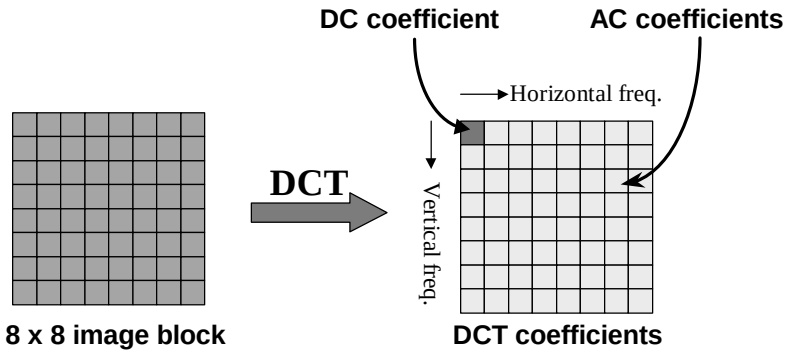
Transform Coefficients — Image Block

- For 8×8 blocks

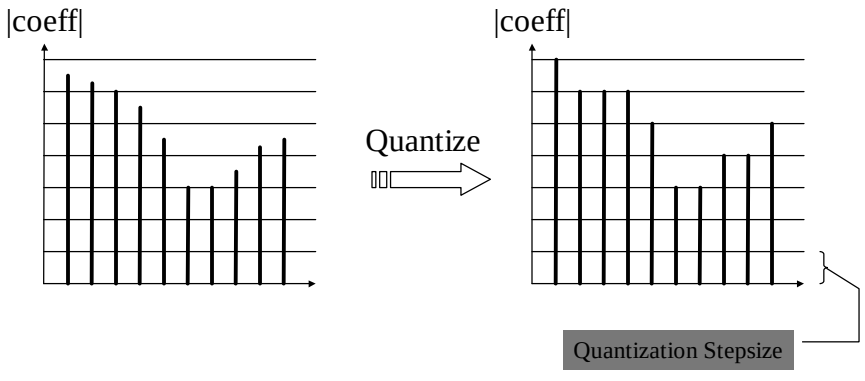
$$C_{mn} = k_n \cos\left[\frac{(2m+1)np}{16}\right] \text{ where } k_n = \begin{cases} 1/(2\sqrt{2}) & \text{when } n = 0 \\ 1/2 & \text{otherwise} \end{cases}$$

- Question: Inverse DCT?

DC and AC Coefficients

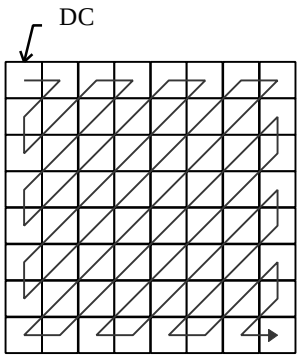


Quantization



Zigzag Scan

- Convert 2-D coefficients block to 1-D coefficients
- To generate long runs of zeros



Entropy Coding

- DC coefficients
  - Differential coding
- AC coefficients
  - **run-level** symbols
    - **run**: length of the zero run
    - **level**: amplitude of the nonzero coefficient
  - Huffman coding
    - Short codes for frequent symbols (Question: Why?)
    - Variable length codes (VLC)

An Example VLC...

Run	Level	Code
EOB		10
0	1	1s If first coefficient in block
0	1	11s Not first coefficient in block
0	2	0100 s
0	3	0010 1s
0	4	0000 110s
0	5	0010 0110 s
0	6	0010 0001 s
0	7	0000 0010 10s
0	8	0000 0001 1101 s
0	9	0000 0001 1000 s
0	10	0000 0001 0011 s
0	11	0000 0001 0000 s
0	12	0000 0000 1101 0s
0	13	0000 0000 1100 1s
0	14	0000 0000 1100 0s
0	15	0000 0000 1011 1s
1	1	011s
1	2	0001 10s
1	3	0010 0101 s
1	4	0000 0011 00s
1	5	0000 0001 1011 s
1	6	0000 0000 1011 0s
1	7	0000 0000 1010 1s
2	1	0101 s
2	2	0000 100s
2	3	0000 0010 11s
2	4	0000 0001 0100 s
2	5	0000 0000 1010 0s
3	1	0011 1s
…	…	…

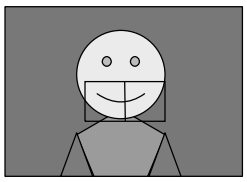
0 0 0 -1 6 0 3 EOB

001111 001000010 001001010 10

Temporal Redundancy: Inter Coding

- Conditional replenishment
  - Transmit only the changing blocks

Previous Frame



Current Frame

Inter Coding (cont.)

- Motion

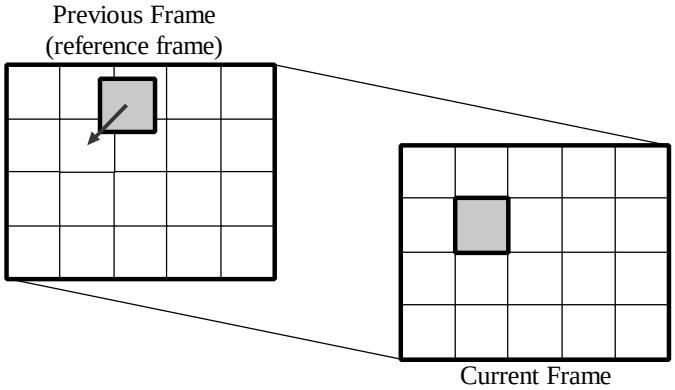
Previous Frame

Current Frame

- Motion compensation
  - Block-based motion
  - Object-based motion
  - Pel-based motion

Block-based Motion Compensation

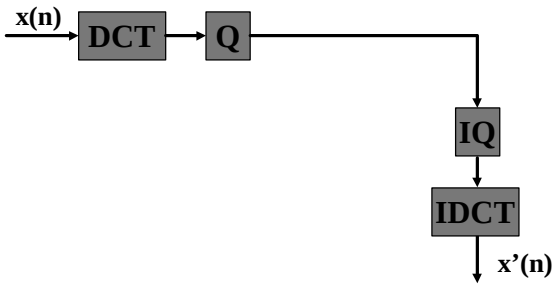
- Block matching for motion estimation (ME)



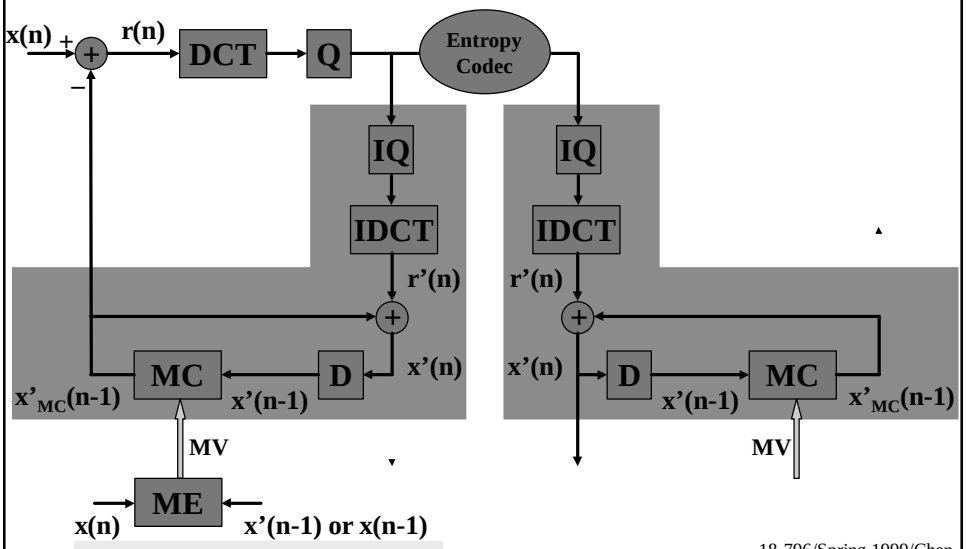
Block-based Motion Compensation (cont.)

- Offset: motion vector
  - Differential coding in x
  - Differential coding in y
- Residue: prediction error
  - Coded as in intra coding

Codec (intra mode)



Codec (inter mode)



International Standards

Why Standards?

- Important for communication
- Customers prefer standards to proprietary schemes: Freedom to choose
- Adoption of standards increases volume and brings down cost of
  - service providers
  - manufacturers
- Reduce the risk of deploying new technology
- Major players often participate
- Research opportunities

Types of Standards

- Industrial/Commercial standards
  - Mutual agreement among companies
  - May become de facto standards
- Voluntary standards
  - By volunteers in open committees
  - Based on consensus
  - Market driven
  - Need to stay ahead of technology

Global Standards Arena

- International
  - **ITU**: International Telecommunication Union
    - ITU-T: ITU Telecommunication Standardization Sector (CCITT)
    - ITU-R: ITU Radio Communication Sector (CCIR)
  - **ISO**: International Standards Organization
  - IEC: International Electrotechnical Commission
  - JTC1: Joint Technical Committee on Information Technology
- Regional
  - CEN/CENELEC: Committee for European Normalization
  - PASC: Pacific Area Standards Congress
- National
  - ANSI: American National Standards Institute

Principles of Coding Standards

- Specify only the decoder
- Standardize the minimum

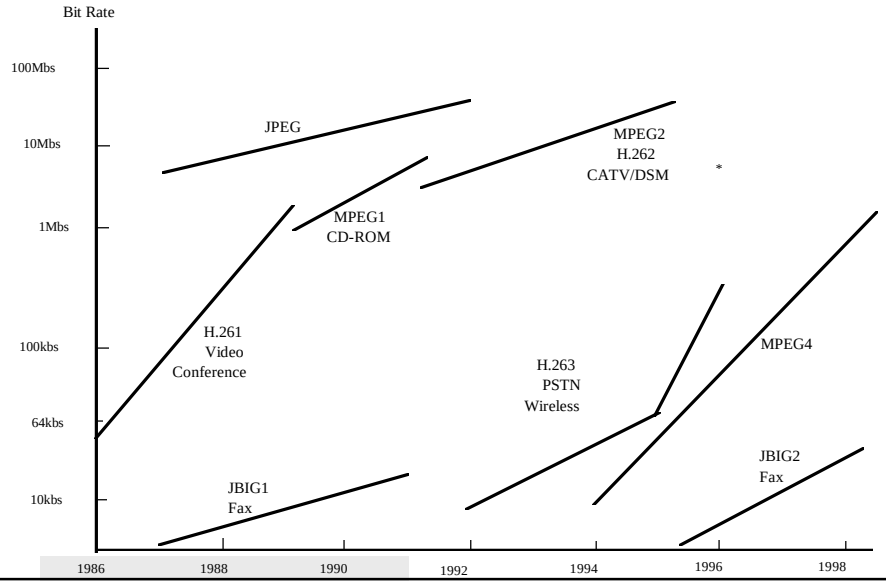
"ISO/IEC JTC1 SC29 WG11"?

- Subcommittee (SC) 29
  - Working Group (WG) 1
    - Joint Bi-Level Image Group (JBIG)
      - Still pictures (1-bit to 4-5 bits)
    - Joint Photographic Expert Group (JPEG)
      - Still pictures (8-bit to 24-bit)
  - WG 11: Moving Picture Experts Group (MPEG)
      - Full-motion video and associated audio
  - WG 12: Multimedia-Hypermedia Experts Group (MHEG)
      - Data a related to multimedia and hypermedia applications

Video Coding Standards

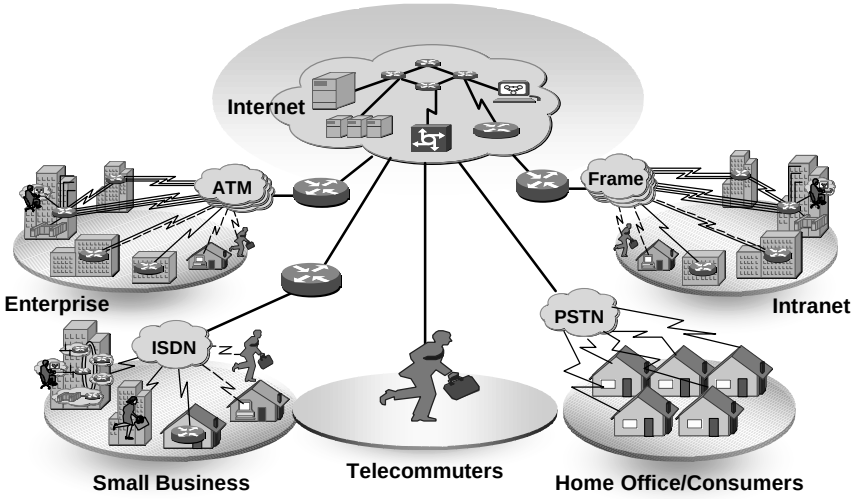
Standards Organization	Video Coding Standard	Typical Range of Bit Rates	Typical Applications
ITU-T	H.261	p×64 kbits/s, p=1…30	ISDN Video Phone
ISO	IS 11172-2 MPEG-1 Video	1.2 Mbits/s	CD-ROM
ISO	IS 13818-2 MPEG-2 Video	4-80 Mbits/s	SDTV, HDTV
ITU-T	H.262		
ITU-T	H.263	64 kbits/s or below	PSTN Video Phone
ISO	CD 14496-2 MPEG-4 Video	24-1024 kbits/s	
ITU-T	H.263 Version 2	< 64 kbits/s or above	PSTN Video Phone
ITU-T	H.26L	< 64 kbits/s	-

Time Line and Bit Rate for Coding Standards



Systems and Networking Issues

The Big Picture...



Issues in Networked Multimedia

- Real-time constraints: delay, delay jitter
- Bandwidth requirement, VBR or CBR, symmetrical or asymmetrical
- Quality of Service (QoS): delay, delay jitter, packet loss, bit-error-rate, burst-error-rate, burst error length...
- Synchronization of video, audio, data, applications...
- Error robustness: error resilience, error concealment
- Cost

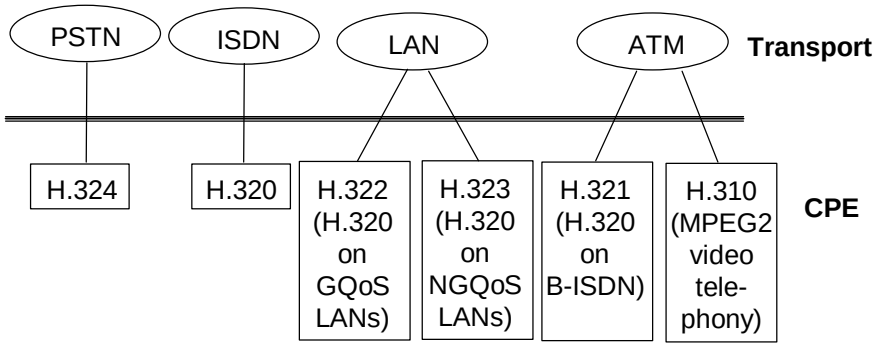
Network Characteristics

- PSTN: up to 33.6 kbits/s, ubiquitous, low cost
- N-ISDN: 128 kbits/s, widely available, low cost
- ATM (B-ISDN): broadband cell-switched network, guaranteed QoS, variable bit-rate, priority, not widely available
- Ethernet: packet-switched network, non-guaranteed QoS, delay, delay variation, packet loss, congestion, widely available, low cost
- IsoEthernet: guaranteed QoS, not widely available, higher cost
- Mobile: low-bit-rate, fading, bit errors
- xDSL, cable, satellite, etc.

Purposes of System Standards

- Media multiplexing
  - Video, audio, data, and control streams
- Capability negotiation
  - Coding algorithms, bit rate, frame rate, data capability, network capability, encryption, etc.
- System control

ITU-T System Standards



PSTN: Public Switched Telephone Network
ISDN: Integrated Switched Digital Network
LAN: Local Area Network
ATM: Asynchronous Transfer Mode

GQoS: Guaranteed Quality of Service
NGQoS: Non-Guaranteed QoS
CPE: Customer Premises Equipment

ITU-T Audiovisual Recommendations

Network		Overall	Video	Audio	Mux	Control/ Signaling	Comm. Interface
WAN	PSTN, Mobile	H.324	H.261, H.263	G.723.1	H.223	H.245	V.34
	N-ISDN	H.320	H.261	G.7xx*	H.221	H.242	I.400
	B-ISDN	H.321	H.261	G.7xx*	H.221	Q.2931	I.361/363 I.400
		H.310	H.261/ H.262	G.7xx* 11172-3	H.222	H.245	I.361/363 I.432
LAN	ISO Ethernet	H.322	H.261	G.7xx*	H.221	H.242	TCP/UDP IP
	Ethernet	H.323	H.261, H.263	G.7xx* G.723.1	H.225.0	H.245	

G.7xx*: G.711, G.722, G.728
11172-3: ISO/IEC 11172-3 MPEG-1 Audio

Topics to be Covered...

- VQ and Subband Coding
- JPEG
- H.261, H.263, H.263 Version 2
- MPEG-1,2,4,7
- MPEG Audio
- Networking Issues
  - Error resilience and network characteristics
- Multimedia over IP
  - RTP, RTCP, RTSP, RSVP
- Multimedia over ATM