18-899 Special Topics in Signal Processing

Multimedia Communications:
Coding, Systems, and Networking

Prof. Tsuhan Chen
tsuhan@ece.cmu.edu

Lecture 1

Fundamentals

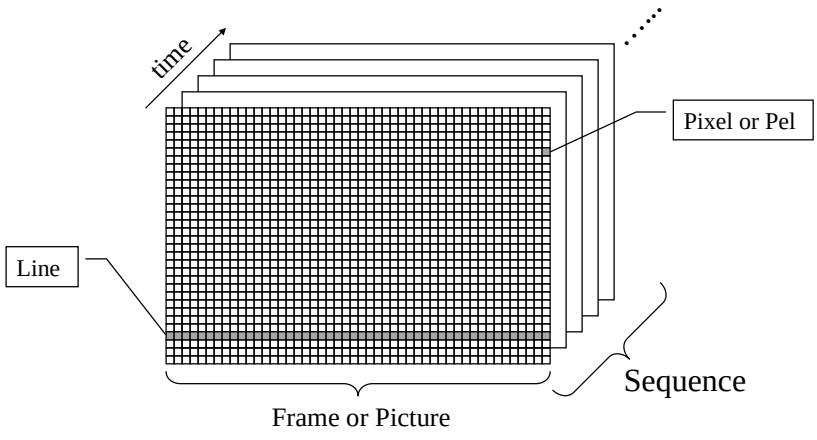
What is Multimedia?

- 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

Images and Video



Why Compression?

- Still images
  - $512 \times 512 \times 3$ 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

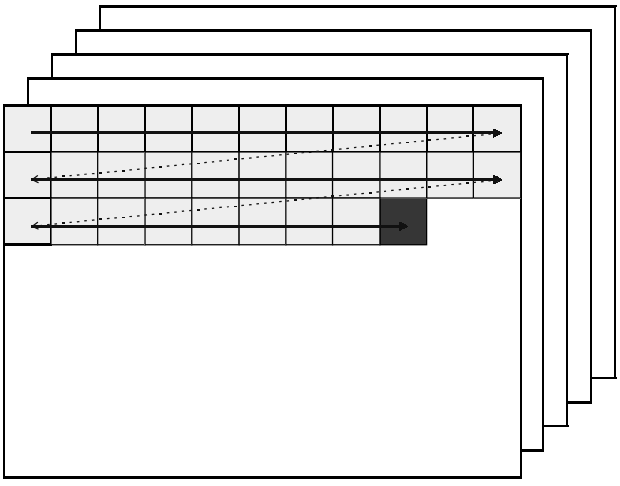
How to compress?

- 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

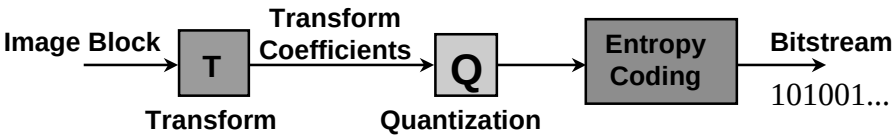
Block-Based Coding



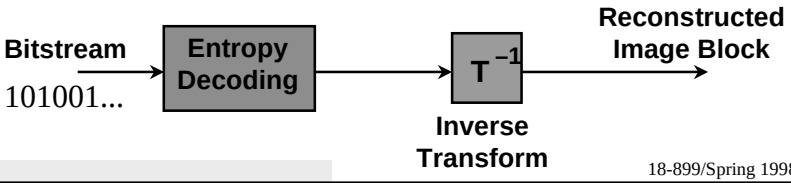
Typical block size: 8×8 or 16×16

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

$$[Y_{mn}] = [C_{mn}]^T [X_{mn}] [C_{mn}]$$

Transform Coefficients | Image Block

- For 8×8 blocks

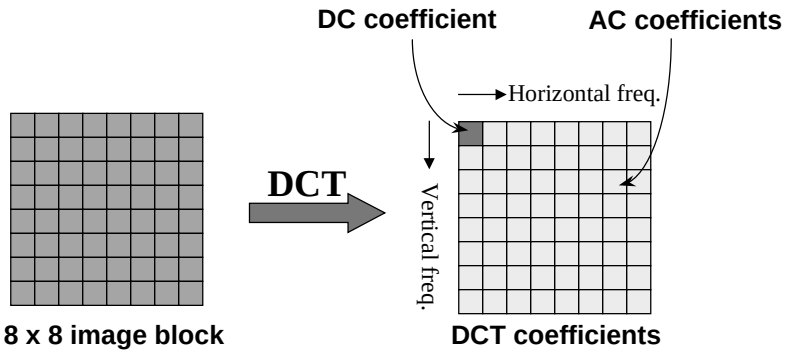
$$C_{mn} = k_n \cos\left[\frac{(2m+1)n\pi}{16}\right] \text{ where } k_n = \begin{cases} 1/(2\sqrt{2}) & \text{when } n = 0 \\ 1/2 & \text{otherwise} \end{cases}$$

- Question: Inverse DCT?

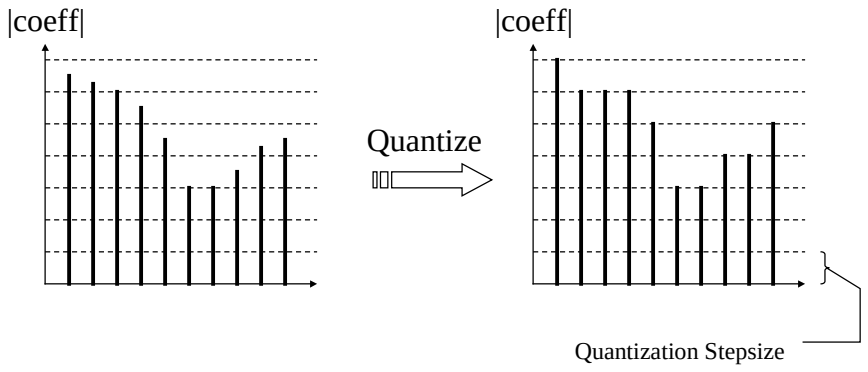
Basis Vectors

0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	→
0.49	0.42	0.28	0.10	-0.10	-0.28	-0.42	-0.49	→
0.46	0.19	-0.19	-0.46	-0.46	-0.19	0.19	0.46	→
0.42	-0.10	-0.49	-0.28	0.28	0.49	0.10	-0.42	→
0.35	-0.35	-0.35	0.35	0.35	-0.35	-0.35	0.35	→
0.28	-0.49	0.10	0.42	-0.42	-0.10	0.49	-0.28	→
0.19	-0.46	0.46	-0.19	-0.19	0.46	-0.46	0.19	→
0.10	-0.28	0.42	-0.49	0.49	-0.42	0.28	-0.10	→

DC and AC Coefficients

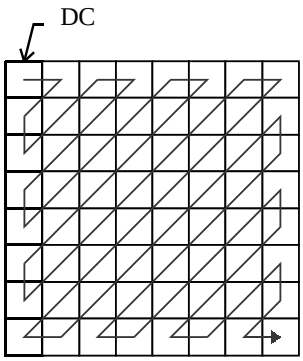


Quantization



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...

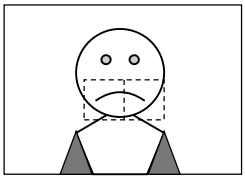
Run	Level	Code
EOB		10
0	1	1s If first coefficient in block
0	1	11s Not first coefficient in block
0	2	0100 s
0	3	0010 1s
0	4	0000 110s
0	5	0010 0110 s
0	6	0010 0001 s
0	7	0000 0010 10s
0	8	0000 0001 1101 s
0	9	0000 0001 1000 s
0	10	0000 0001 0011 s
0	11	0000 0001 0000 s
0	12	0000 0000 1101 0s
0	13	0000 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

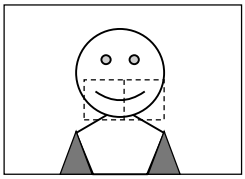
001111 001000010 001001010 10

Temporal Redundancy: Inter Coding

- Conditional replenishment
  - Transmit only the changing blocks



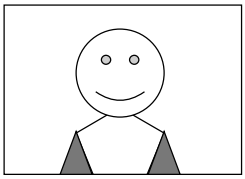
Previous Frame



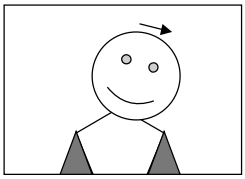
Current Frame

Inter Coding (cont.)

- Motion



Previous Frame

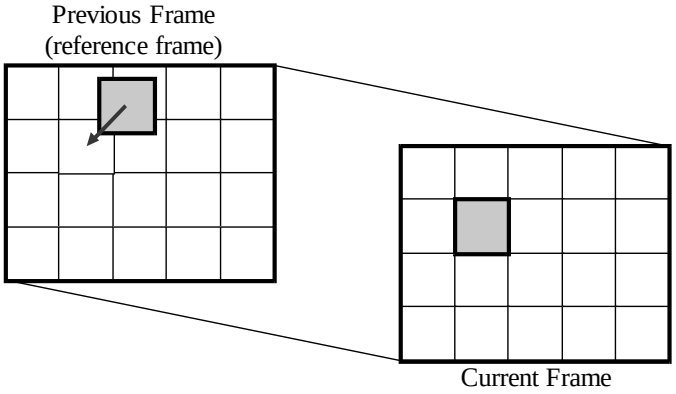


Current Frame

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

Block-based Motion Compensation

- Block matching for motion estimation (ME)

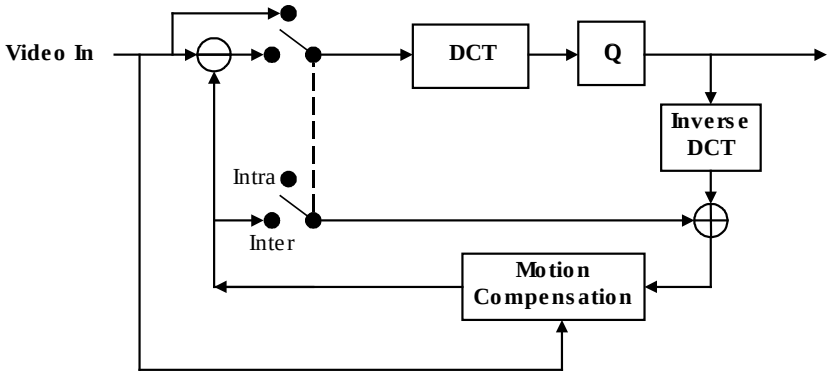


Block-based Motion Compensation (cont.)

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

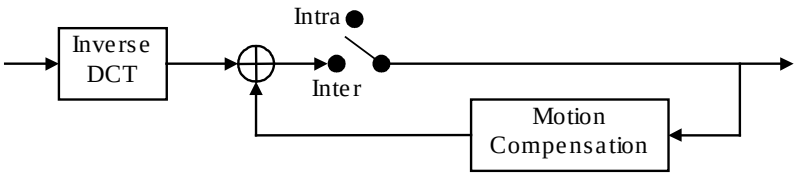
Codec

- Encoder



Codec (cont.)

- Decoder



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

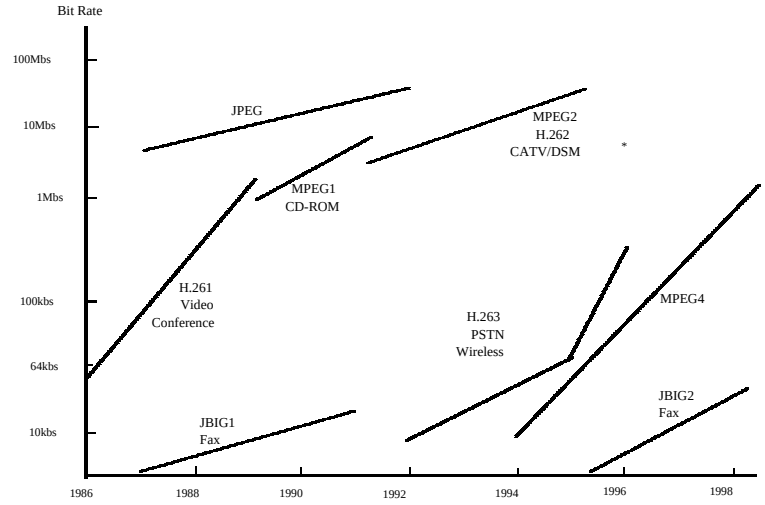
"ISO/IEC JTC1 SC29 WG11"?

- 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	PSTN Video Phone
ITU-T	H.263L	< 64 kbits/s	-

Time Line and Bit Rate for Coding Standards



Levels of Modeling

MODELS	CODED INFORMATION	EXAMPLES
Pixels	**Color of pixels**	**PCM**
Statistically dependent pixels	**Prediction error or transform coeffs**	**Predictive Coding Transform Coding**
Moving blocks	**Motion vectors and prediction error**	**Block-based coding H.261/263, MPEG-1/2**
Moving regions	**Shapes, motion, and colors of regions**	**Region-based coding H.263+, MPEG-4**
Moving objects	**Shapes, motion, and colors of objects**	**Model-based coding MPEG-4**
Facial models	**Action units**	**MPEG-4**
A/V objects	**Descriptive languages**	**MPEG-7**

Levels of Modeling (cont.)

- Better modeling implies
  - Higher compression
  - More content accessibility
  - More complexity
  - Less error resilience
- Currently
  - Block-based: H.261, H.263
  - 2D region-based: H.263+, MPEG-4 Video
  - Model-based: MPEG-4 SNHC
  - High-level descriptive language: MPEG-7