Image and Video Compression

EE398

Bernd Girod
Information Systems Laboratory
Department of Electrical Engineering
Stanford University

Winter 2005/06

Image and Video Compression Everywhere …

- Fax machines
- Digital still cameras
- Digital camcorders
- Digital television broadcasting
- Digital video/versatile disk (DVD)
- Personal video recorder (PVR, aka TiVo)
- World Wide Web
- Internet video streaming
- Video conferencing
- . . .

Motivating Image Compression

- Binary image (facsimile Group 3)
  - 8.5 x 11 in document scanned at 7.7 lines/mm (“fine mode”), 1664 pixels/line, with 1 bit/pixel
  - 3.255 Mbits for 1 page = 5.65 minutes over 9600 baud connection
- Photos on 35 mm film
  - Scanned at 12μ resolution (3656x2664 pixels) with 8 bits per color and 3 colors
  - 233 Mbits for 1 photo, 2/3 of 48 Mbyte compact flash card

Motivating Video Compression

- Digital video studio standard ITU-R Rec. 601

	Y	Cb	Cr
Sampling rate	13.5 MHz	6.75 MHz	6.75 MHz
Quantization	8 bit	8 bit	8 bit
Raw bit rate		**216 Mbps**	
W/o blanking intervals		**166 Mbps**	

- Some interesting bit-rates
  - Terrestial TV broadcasting channel ~20 Mbps
  - DVD (max. 17 GB/length of movie) 10...20 Mbps
  - Ethernet/Fast Ethernet <10/100 Mbps
  - DSL downlink 384...2048 kbps
  - Wireless cellular data 9.6...384 kbps

Personal Video Recorder

MPEG-2 Qualities

Best	7.7 Mbps
High	5.4 Mbps
Medium	3.6 Mbps
Basic	2.2 Mbps

Introductory Lecture

- Goal: provide a first introduction of some key concepts of image compression, without rigorous treatment
- Lossless vs lossy compression
- Measuring distortion and compression
- Statistical and visual redundancy
- Need for structured compression schemes

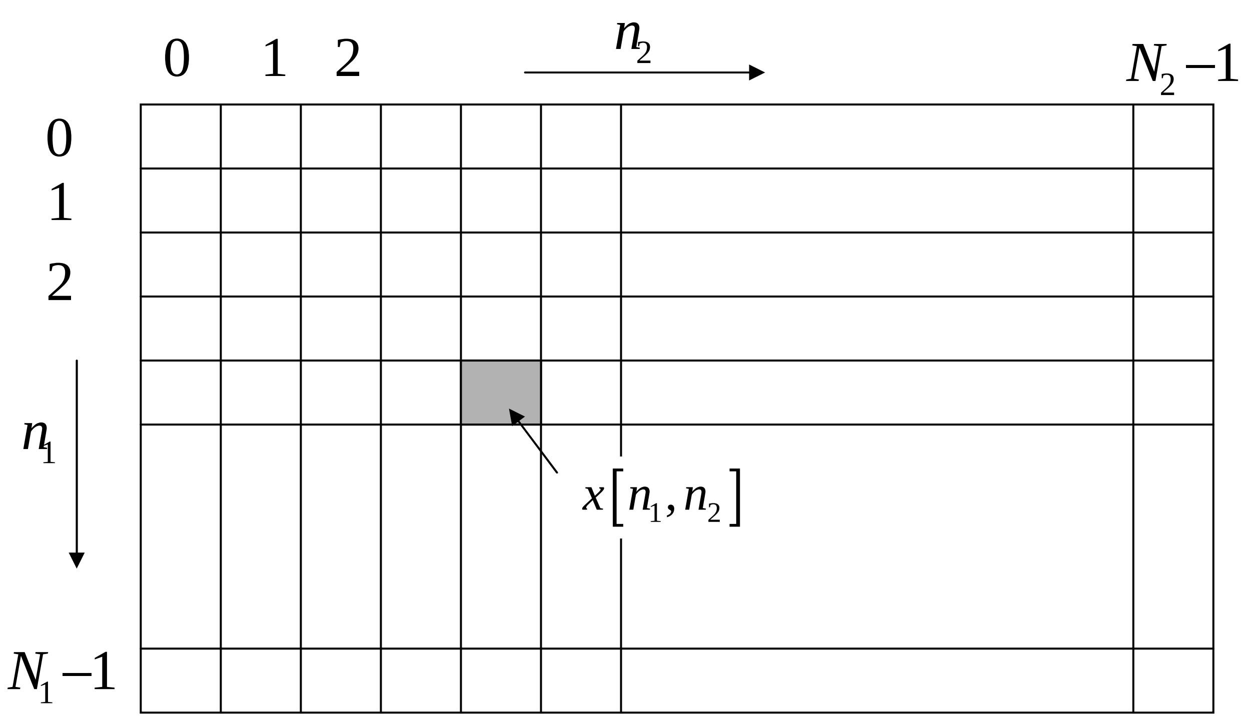
- EE398 Organisation

Digital Image: Quantized Array of Samples

Real-valued intensities, range 0…1 or –1/2…+1/2

Rounding to nearest integer

$$x[n_1,n_2] = \left\langle 2^B x'[n_1,n_2] \right\rangle \qquad 0 \le n_1 < N_1, \quad 0 \le n_2 < N_2$$



Multiple Image Components

- Color images typically represented by three values per sample location, for red, green and blue primary components

$$x_R[n_1, n_2], \quad x_G[n_1, n_2], \quad x_B[n_1, n_2]$$

- General multi-component image

$$x_C[n_1, n_2], \qquad c = 1, 2, \ldots, C$$

- Examples:
  - Color printing: cyan, magenta, yellow, black dyes, sometimes more
  - Hyperspectral satellite imaging: 100s of channels

Lossless Compression

- Minimize number of bits required to represent original digital image samples w/o any loss of information.
- All B bits of each sample must be reconstructed perfectly.
- Achievable compression usually rather limited.
- Applications
  - Binary images (facsimile)
  - Medical images
  - Master copy before editing
  - Palettized color images

Lossy Compression

- Some deviation of decompressed image from original ("distortion") is often acceptable:
  - Human visual system might not perceive loss, or tolerate it.
  - Digital input to compression algorithm is imperfect representation of real-world scene
- Much higher compression than with lossless.
- Lossy compression used widely for natural images (e.g. JPEG) and motion video (e.g. MPEG).

Lossy Compression: Measuring Distortion

- Most commonly employed: Mean Squared Error

$$\mathrm{MSE} = \frac{1}{N_1 N_2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \left(x[n_1, n_2] - \hat{x}[n_1, n_2] \right)^2$$

. . . or, equivalently, Peak Signal to Noise Ratio

$$\mathrm{PSNR} = 10 \log_{10} \frac{\left(2^B - 1\right)^2}{\mathrm{MSE}} \ \mathrm{dB}$$

- Advantages
  - Easy calculation
  - Mathematical tractability in optimization problems
- Disadvantage
  - Neglects properties of human vision

Measures of Compression

- Image represented by "bit-stream" $\mathbf{c}$ of length $\|\mathbf{c}\|$.
- Compare no. of bits w/ and w/o compression

$$\text{compression ratio} = \frac{N_1 N_2 B}{\|\mathbf{c}\|}$$

- Alternatively

$$\text{bit-rate} = \frac{\|\mathbf{c}\|}{N_1 N_2} \ \text{bits/pixel}$$

- For lossy compression, bit-rate more meaningful than compression ratio, as B is somewhat arbitrary.

Typical Bit-rates after Compression

- Dependent on image content: consider typical natural images
- Lossless compression: $(B\text{-}3)$ bpp (bits per pixel)
- Lossy compression,
  - Perceived distortion depends on sampling density and contrast
  - Assume viewing on computer monitor, 90 pixels/inch.
  - high quality: 1 bpp
  - moderate quality: 0.5 bpp
  - usable quality: 0.25 bpp

How Does Compression Work?

- Exploit statistical redundancy.
  - Take advantage of patterns in the signal.
  - Describe frequently occuring events efficiently.
  - <u>Lossless coding:</u> only statistical redundancy
- Introduce acceptable deviations.
  - Omit “irrelevant” detail that humans cannot perceive.
  - Match the signal resolution (in space, time, amplitude) to the application
  - <u>Lossy coding:</u> exploit statistical <u>and</u> visual redundancy

Statistical Redundancy

- Trivial example: Given two B-bit integers (e.g., representing two adjacent pixels)

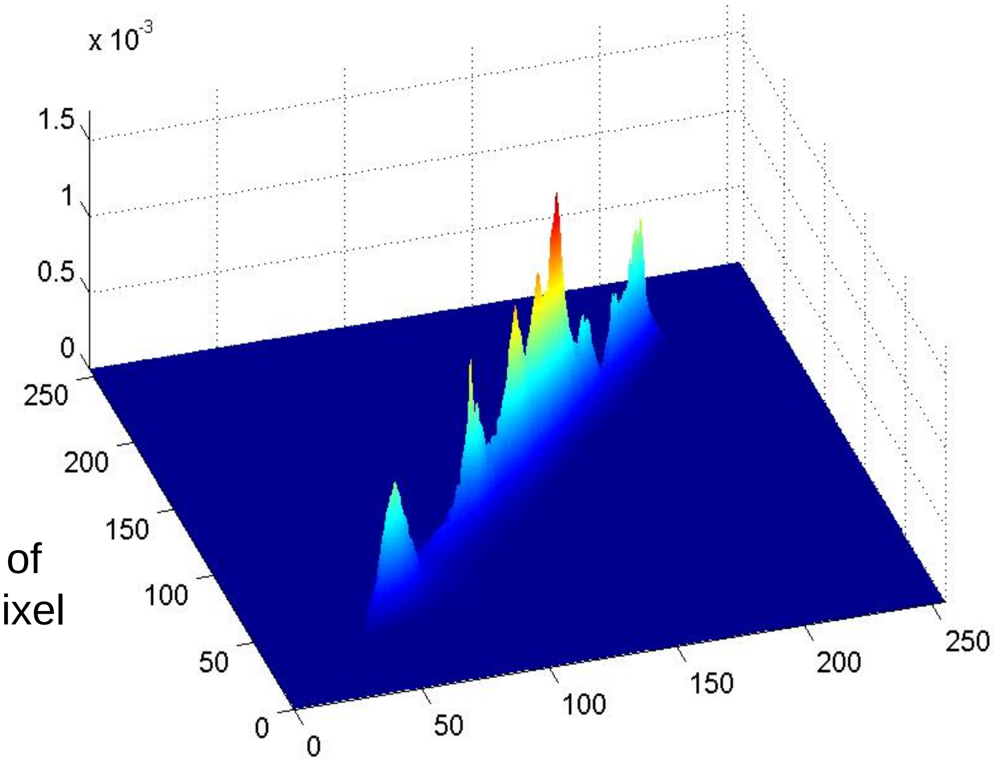
$$x_1, x_2 \in \left\{0, 1, \ldots, 2^B - 1\right\}$$

- Assume that x_1, x_2 only takes on values $\{0,1\}$
  - Compression to 1 bpp
- Further assume, that $x_1 = x_2$
  - Compression to 0.5 bpp
- Hope: bit-rate increases only slightly, as long as the above assumptions hold with high probability

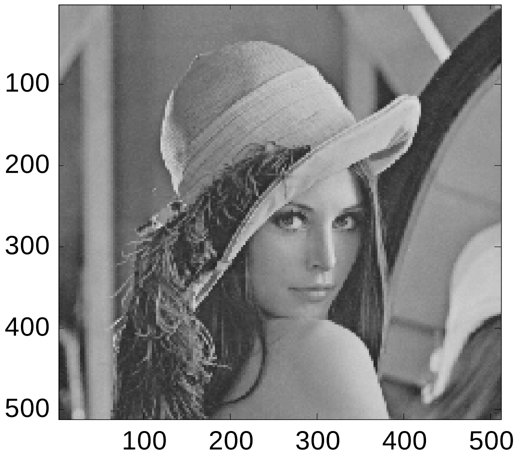
Joint Histogram of two Horizontally Adjacent Pixels

PMF

Amplitude of adjacent pixel



Amplitude of current pixel



('Lena', 512 x 512 pixels, 8 bpp)

Visual Redundancy

- For images to be viewed by humans, no need to represent more than the visible resolution in
  - space
  - time
  - brightness
  - color
- Required resolution might depend on image content ("masking")
- For some applications, only a specific region of the image might be relevant, e.g., in
  - medical imaging
  - military imaging

Exploiting Limitations of Color Vision

- Human visual system has much lower acuity for color hue and saturation than for brightness
- Use color transform to facilitate exploiting that property

$$\begin{pmatrix} x_Y \\ x_{Cb} \\ x_{Cr} \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.5 \\ 0.5 & -0.419 & -0.0813 \end{pmatrix} \cdot \begin{pmatrix} x_R \\ x_G \\ x_B \end{pmatrix}$$

Luminance component (x_Y)

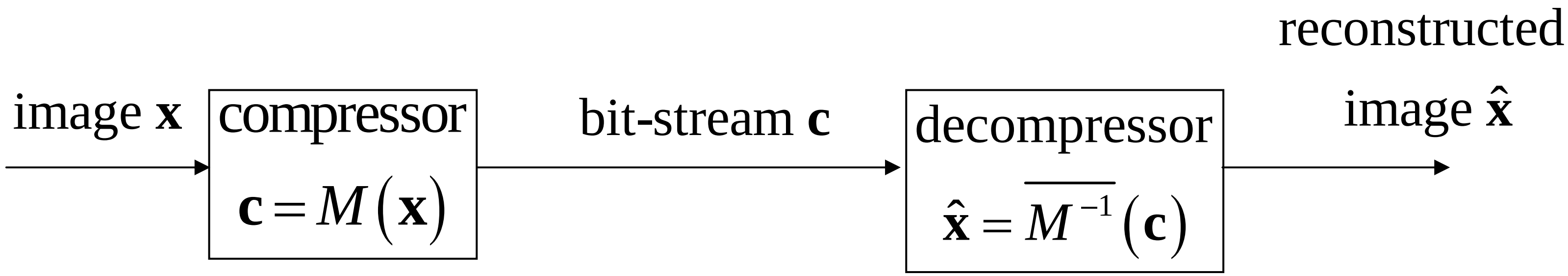
Chrominance components (x_{Cb}, x_{Cr})

RGB components (γ-predistorted)

- Note $x_{Cb} = 0.564\left(x_B - x_Y\right)$ $\quad$ $x_{Cr} = 0.713\left(x_R - x_Y\right)$

- Cb and Cr often sub-sampled 2:1 relative to Y.

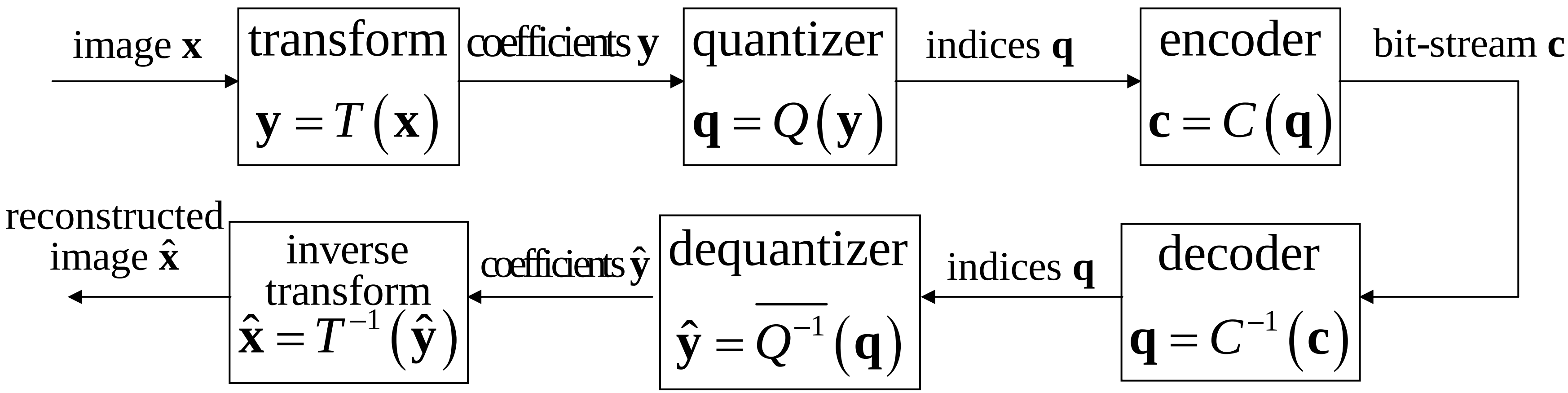
Compression as a Global Mapping



Lookup table interpretation	lookup table $2^{N_1 N_2 B}$ entries	fixed length $\lVert\mathbf{c}\rVert$	lookup table $2^{\lVert\mathbf{c}\rVert}$ entries

- Lossless compression $\overline{M^{-1}} = M^{-1}$
- Lossy compression $\mathbf{c} = M(\mathbf{x}) = \arg\min_{\mathbf{c}'} D\left(\mathbf{x}, \overline{M^{-1}}(\mathbf{c}')\right)$

Typical Structured Compression System



- Transform $T(\mathbf{x})$ usually invertible
- Quantization $Q(\mathbf{y})$ not invertible, introduces distortion
- Combination of encoder $C(\mathbf{q})$ and decoder $C^{-1}(\mathbf{c})$ lossless

Outline EE398

- Entropy and lossless coding techniques
- Run-length coding, fax standards
- Arithmetic coding
- Rate-distortion limits and quantization
- Lossless and lossy predictive coding
- Transform coding, JPEG standard
- Subband coding, wavelets, JPEG-2000
- Motion compensated coding, MPEG standards

EE398 Organisation

- Regularly check class home page:

  http://www.stanford.edu/class/ee398
- Co-instructor: Dr. Markus Flierl
- Assistants
  - General TA: David Varodayan
  - ISE lab TA: Shantanu Rane
  - Course assistant: Kelly Yilmaz

EE398 Organisation (cont.)

- Homeworks
  - 7 problem sets, require computer + Matlab (Image Proc. Toolbox)
  - Handed out Tuesdays, due one week later, 2:45 p.m.
- Grading
  - Homeworks: 50%
  - Term project: 50%
  - No mid-term, no final

EE398 Term Projects - General

- Work in groups of 2-3 students, 40-50 hours per person
- Project proposal required, deadline: February 9
- Class-room presentations of projects: March 14/16
- Project report due: March 16
- Project grade based on
  - Technical quality, significance, and originality of results 50%
  - Project report 25%
  - Class-room presentation 25%

ISE laboratory

- Created by equipment grants from Hewlett-Packard, Xerox, and Intel
- Exclusively a teaching laboratory
- Location: Packard room 021
- 20 Linux PCs, 2 Windows PCs, scanners, printers etc.
- Access:
  - door combination for lab entry will be provided by TA
  - Account on ise machine will be provided to all enrolled in class

Further reading

- Slides available as hand-outs and as pdf files on the web
- Recommended text books
  - *D. S. Taubman, M. W. Marcellin, „JPEG2000 – Image Compression Fundamentals, Standards, and Practice,“ Kluwer Academic Publishers, 2002.*
  - *Y. Wang, J. Ostermann, Y.-Q. Zhang, „Video Processing and Communications,“ Prentice-Hall, 2002.*
- Selected papers will be posted on Web site

Reading for this chapter

- Girod, Gray, Kovacevic, Vetterli, "Image and Video Coding," SP Magazine. March 1998.
- Taubman+Marcellin, Chapters 2.1 and 2.2