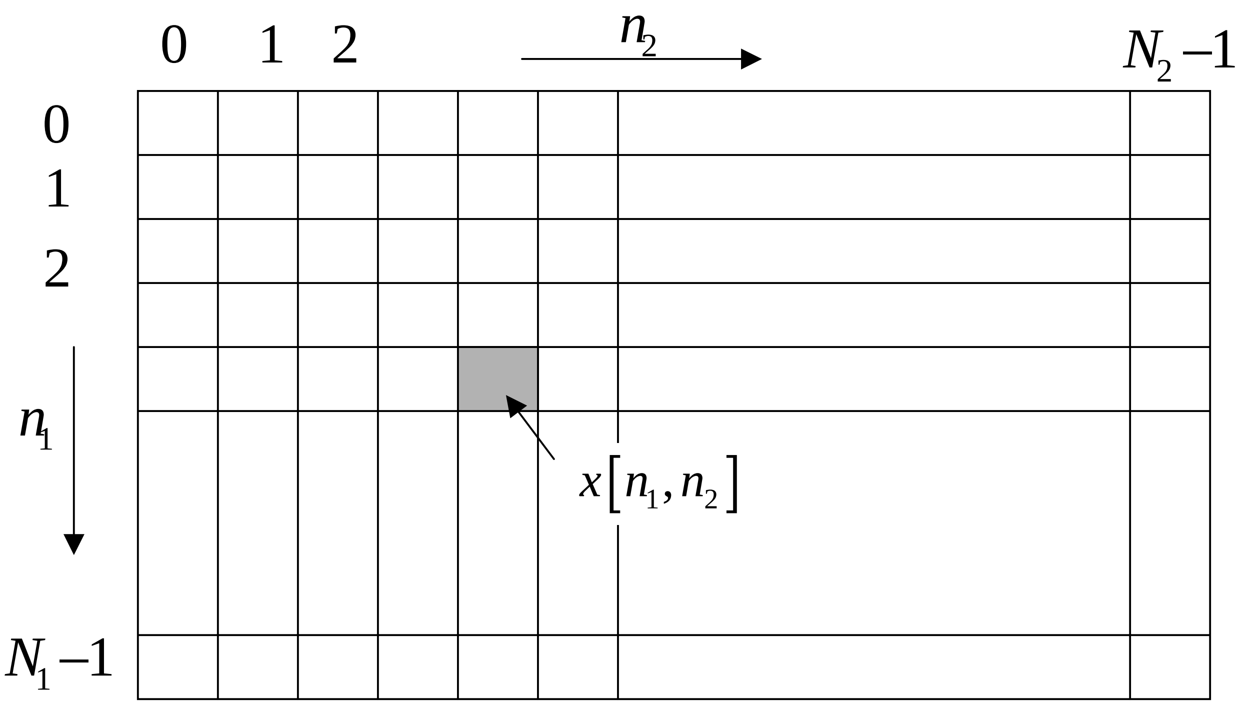
Image Compression Overview

- Goal of this chapter: provide a first introduction of some key terms and ideas of image compression, without rigorous treatment
- Lossless vs lossy compression
- Measuring distortion and compression
- Statistical redundancy and entropy
- Compression as a vector quantization problem
- Quantization
- Transforms
- Compression with predictive feedback

Digital images

- A digital image is a two-dimensional sequence of samples

$$x[n_1, n_2], \qquad 0 \le n_1 < N_1, \quad 0 \le n_2 < N_2$$



Discrete image intensities

- Unsigned B-bit imagery $x[n_1, n_2] \in \{0, 1, \ldots, 2^B - 1\}$
- Signed B-bit imagery $x[n_1, n_2] \in \{-2^{B-1}, -2^{B-1} + 1, \ldots, 2^{B-1} - 1\}$
- Most common: $B=8$, but larger B are used in medical, military, or scientific applications.
- Useful interpretation

$$x[n_1, n_2] = \langle 2^B x'[n_1, n_2] \rangle$$

Real-valued intensities, range 0…1 or –1/2…+1/2 (for $x'[n_1, n_2]$)

Rounding to nearest integer (for $\langle \cdot \rangle$)

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

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

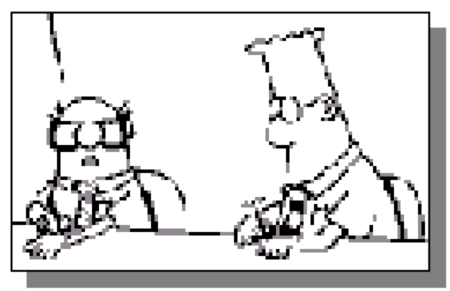
- Examples:
  - Color printing: cyan, magenta, yellow, black dyes, sometimes more
  - Hyperspectral satellite imaging: 100s of channels

Some classes of imagery

- “Natural” image

- Text image

- Graphics



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

Gamma correction

- Display devices (e.g. CRTs) highly nonlinear

$$L \sim x^{\gamma}, \qquad \gamma = 1.8 \ldots 2.8$$

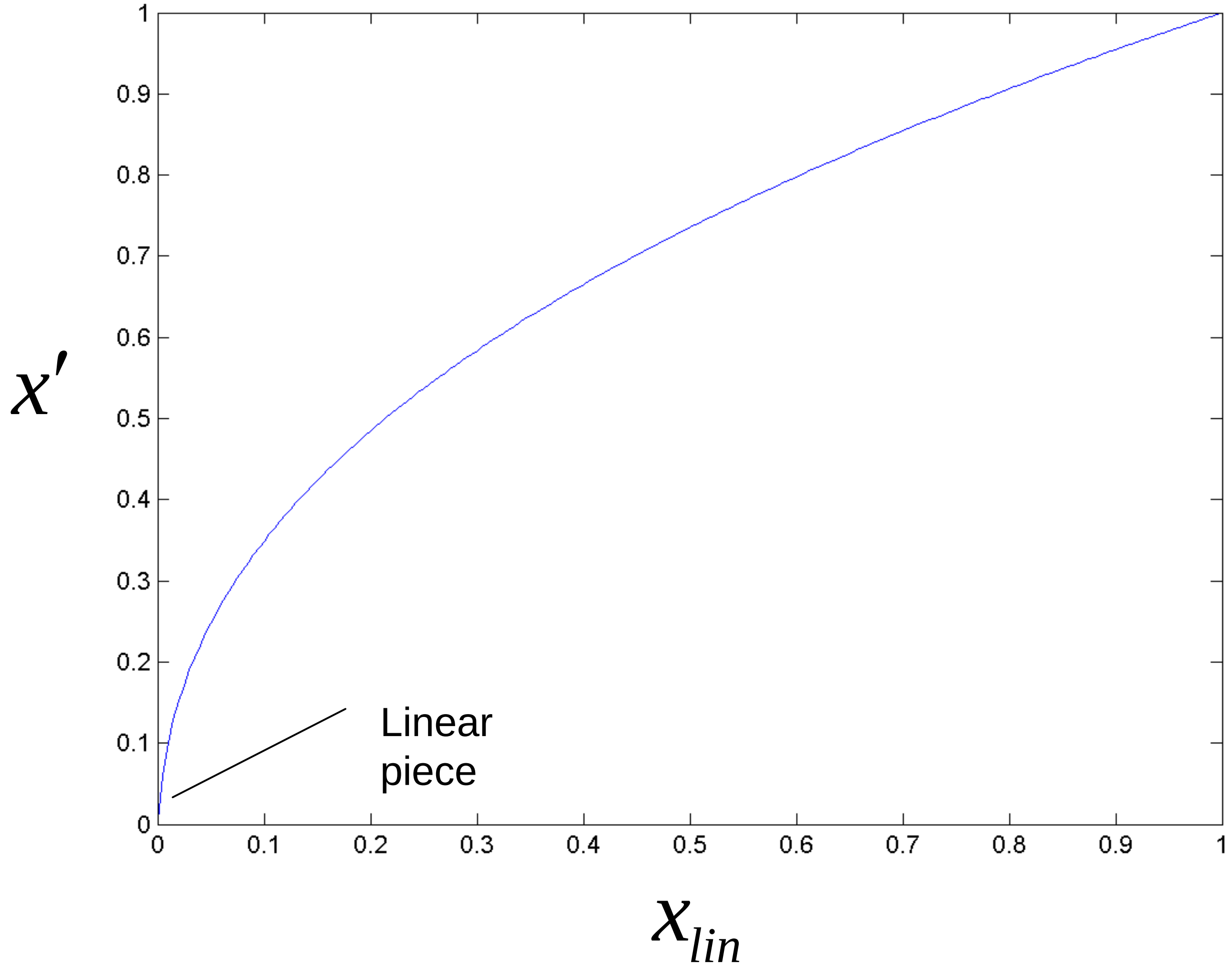
Screen luminance

- Cameras compensate by γ-predistortion circuitry
- IEC 61966-2-1 standard for γ-predistorted color space sRGB (“standard RGB”), each normalized linear RGB component mapped by

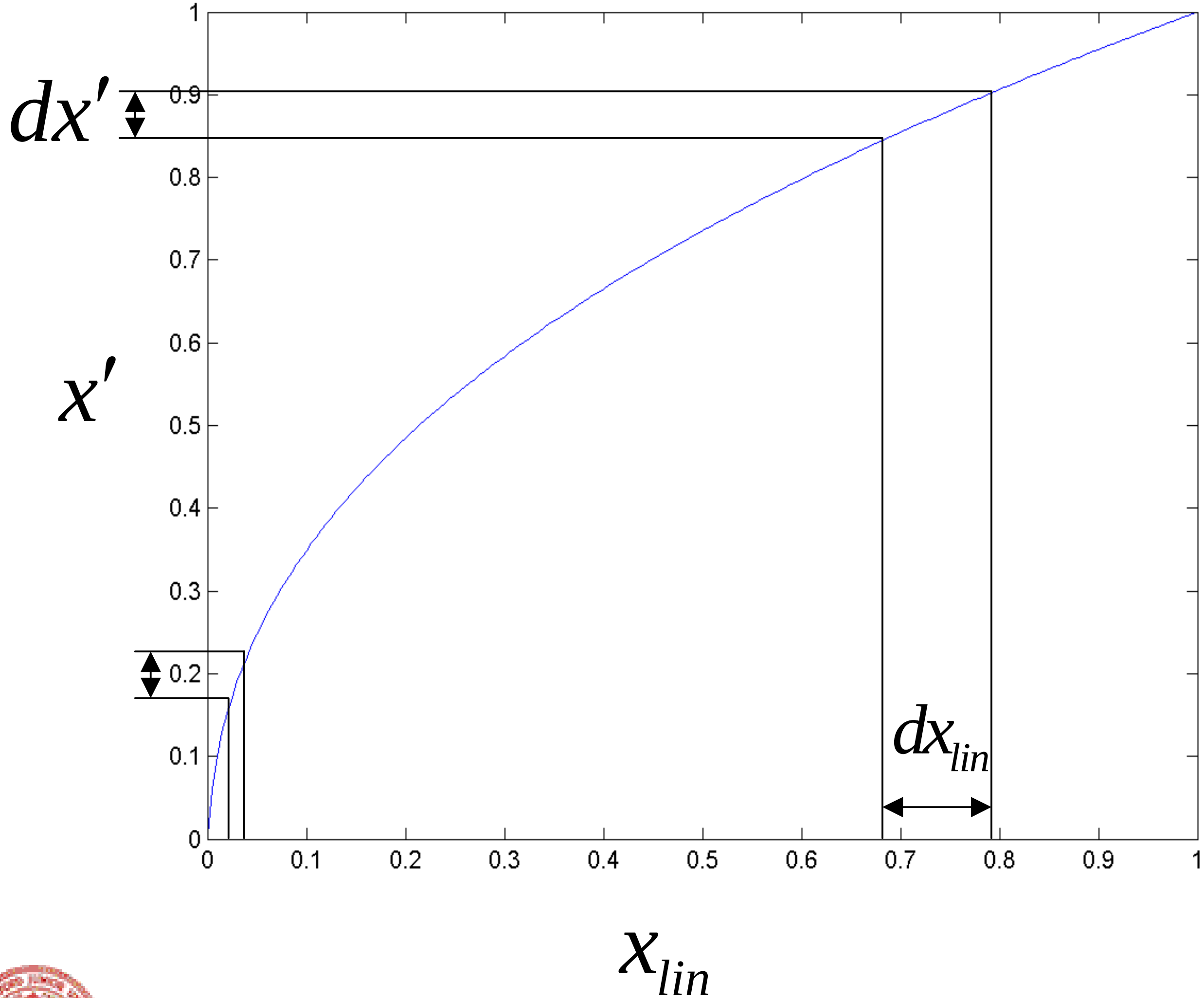
$$x' = \begin{cases} g x_{lin} & \text{if } 0 \le x_{lin} \le \varepsilon \\ (1+\beta)\, x_{lin}^{1/\gamma} - \beta & \text{if } \varepsilon \le x_{lin} \le 1 \end{cases}$$

$$\gamma = 2.4, \quad \beta = 0.055, \qquad \varepsilon = \left(\frac{\beta}{(1-\beta)\left(1 - \frac{1}{\gamma}\right)} \right)^{\gamma}, \quad g = \frac{\beta}{\varepsilon(\gamma - 1)}$$

sRGB γ-characteristic



Effect of a small error



$$dx_{lin} = \gamma \left(x' \right)^{\gamma - 1} dx'$$
$$= \gamma \left(x_{lin} \right)^{1 - 1/\gamma} dx'$$

Roughly matches Weber's Law.

PSNR distortion measure only useful in the γ-corrected domain.

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

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

How does compression work?

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

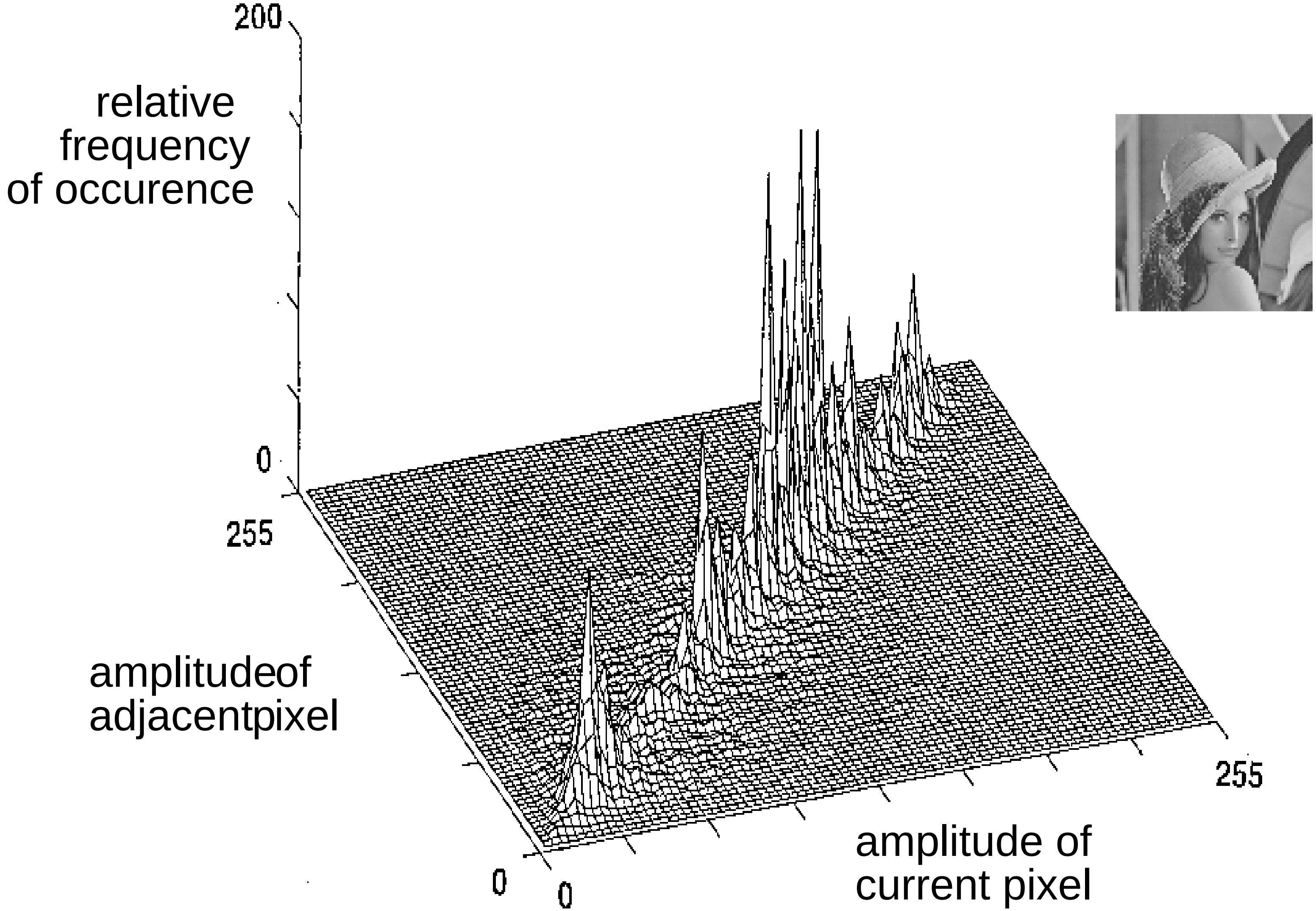
Statistical redundancy

- 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



Entropy

- Statistical redundancy is formalized by the concept of entropy of Shannon's information theory
- Minimum bit-rate for encoding x_1 or x_2 separately

$$H(x_1) \qquad H(x_2)$$

- Minimum bit-rate for encoding x_1 and x_2 jointly

$$H(x_1, x_2) \leq H(x_1) + H(x_2)$$

- Entropies $H(x_1),\ H(x_2), H(x_1, x_2)$ are determined entirely by the statistical distribution of values x_1 and x_2

Visual Irrelevance

- 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

Irrelevance in Color Imagery

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

Luminance component

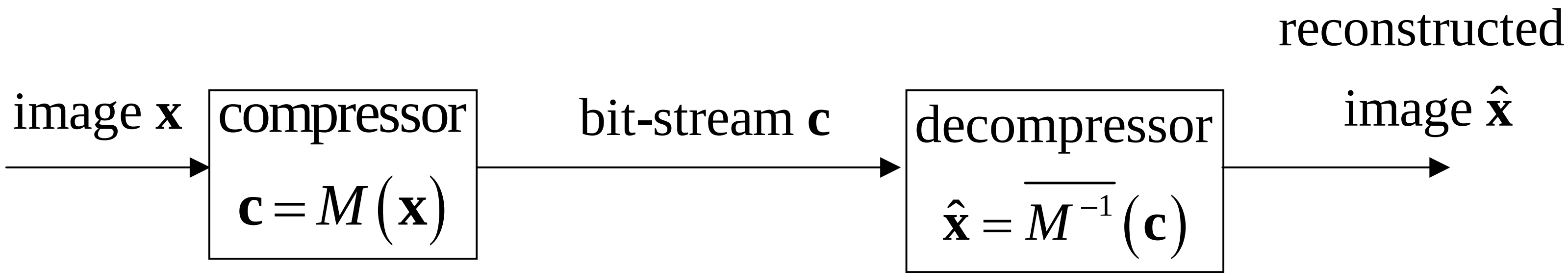
$$\begin{pmatrix} x_Y \\ x_{Cb} \\ x_{Cr} \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.5 \\ 0.5 & -0.419 & -0.0813 \end{pmatrix} \cdot \begin{pmatrix} x_R \\ x_G \\ x_B \end{pmatrix}$$

Chrominance components

RGB components (γ-predistorted)

- Note $x_{Cb} = 0.564(x_B - x_Y)$ $\qquad$ $x_{Cr} = 0.713(x_R - x_Y)$
- 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 $\|\mathbf{c}\|$	lookup table $2^{\|\mathbf{c}\|}$ entries

- Lossless compression $\overline{M^{-1}} = M^{-1}$
- Lossy compression

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

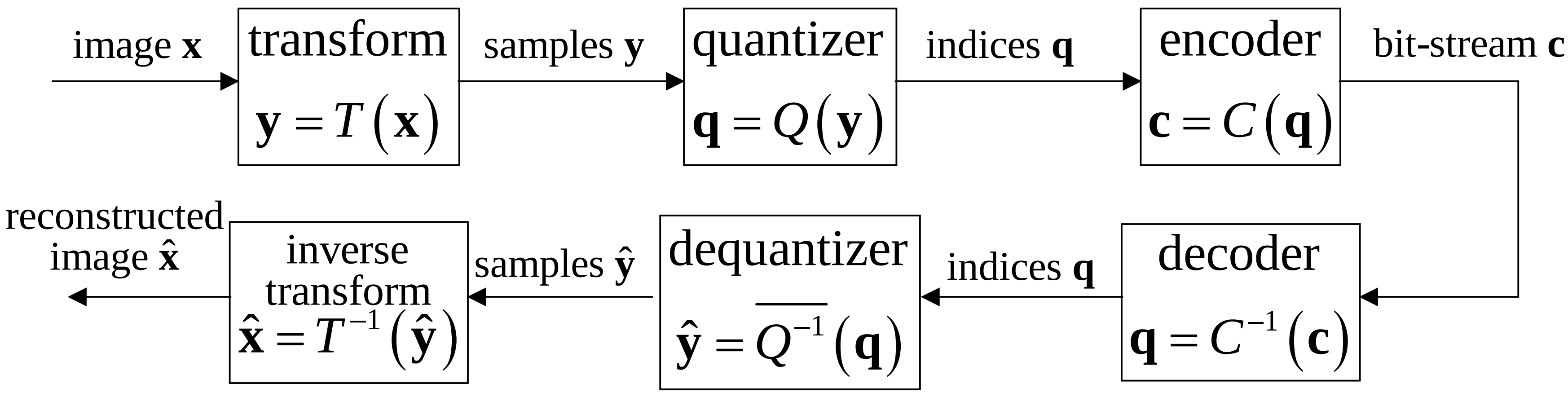
- Leads to idempotent compression system

$$M\left(\overline{M^{-1}}(\mathbf{c})\right) = \mathbf{c}$$

Compression as a global mapping (cont.)

- Variable length encoding: need formulation of trade-off between rate and distortion
- Lookup table formulation corresponds to approach taken by fixed-length vector quantization
- Size of lookup tables for entire image is impractical (both for encoder and decoder)
- Impose structure on mappings M and $\overline{M^{-1}}$
  - Avoid exponential increase of complexity with image size and size of bit-stream
  - Maintain achievable performance close to unconstrained system

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

Variable length coding

- Goal: exploit non-uniformity of probability distribution of quantization indices
- Simple example: Probabilities Code words

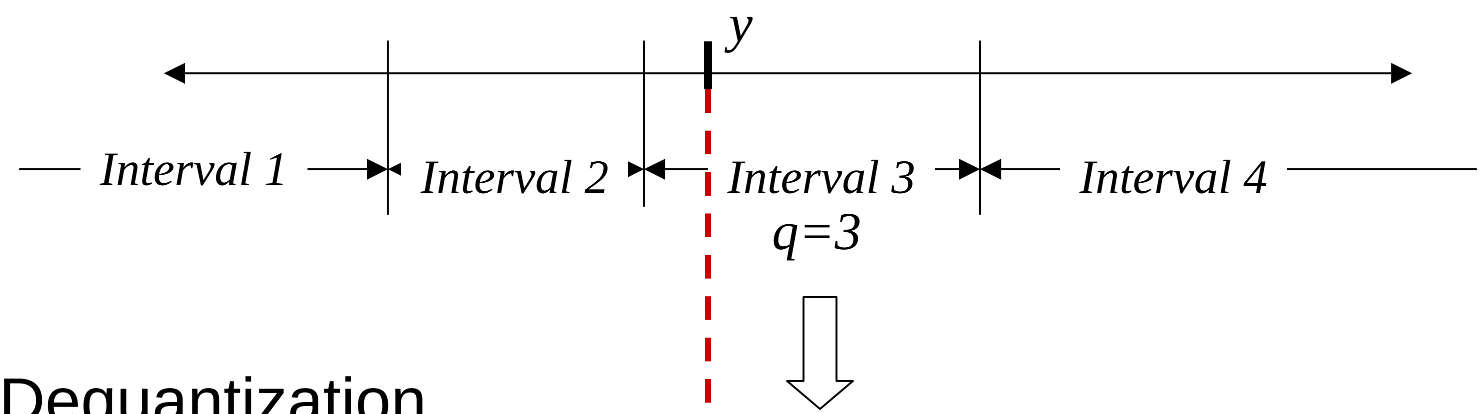
$$P_q(q) = \begin{cases} 1/2 & \text{if } q = 0 \\ 1/4 & \text{if } q = 1 \\ 1/8 & \text{if } q = 2 \\ 1/8 & \text{if } q = 3 \end{cases} \qquad \begin{matrix} 1 \\ 01 \\ 001 \\ 000 \end{matrix}$$

Average code word length $\frac{1}{2}\cdot 1\text{ bits} + \frac{1}{4}\cdot 2\text{bits} + 2\cdot\frac{1}{8}\cdot 3\text{bits} = 1.75\text{ bits}$

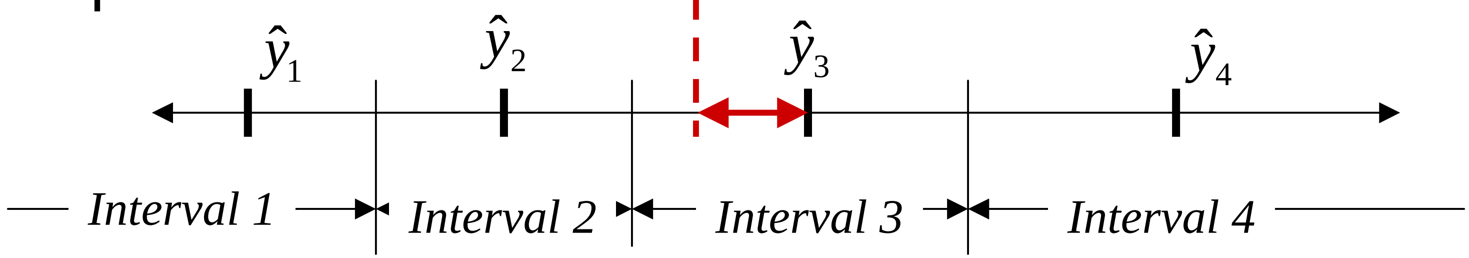
- Often much bigger gains for image compression

Quantization

- Goal: reduce the number of possible amplitude values for coding
- Simple scalar quantizer with four output indices

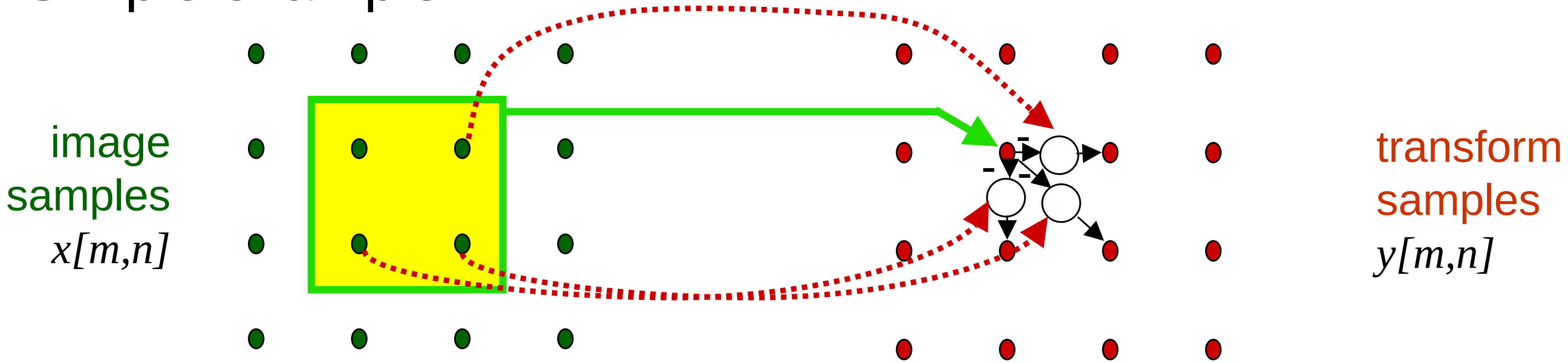


- Dequantization



Transforms

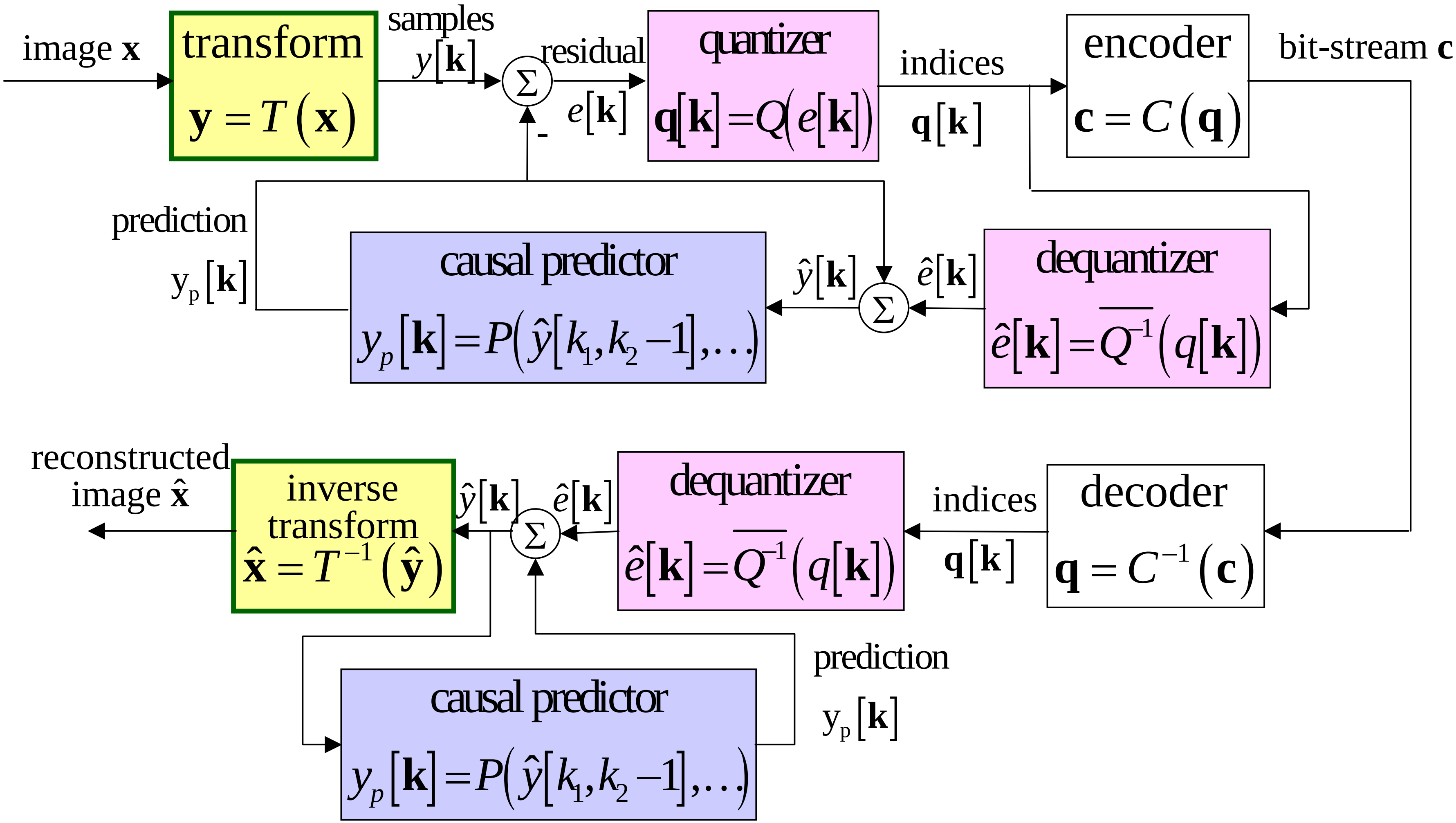
- Goal: represent the image samples in a different form, such that statistical dependencies are greatly reduced
- Simple example:



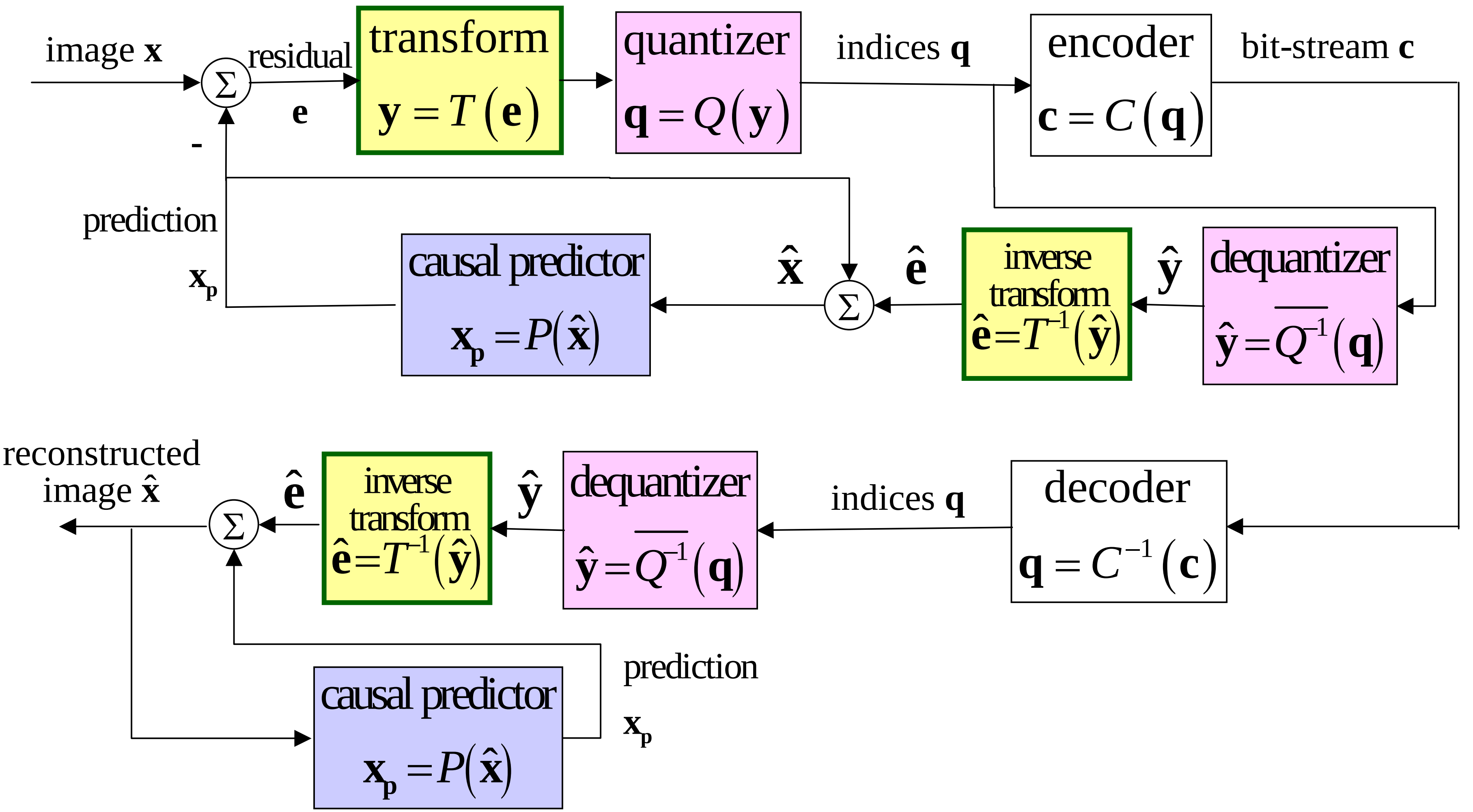
$$y[m,n]=\begin{cases}\frac{1}{4}\sum_{i,j\in\{0,1\}} x[m+i,n+j] & \text{if } m,n \text{ both even}\\ x[m,n]-y\left[2\left\lfloor\frac{m}{2}\right\rfloor,2\left\lfloor\frac{n}{2}\right\rfloor\right] & \text{otherwise}\end{cases}$$

$$\hat{x}[m,n]=\begin{cases}2\hat{y}[m,n]-\sum_{i,j\in\{0,1\}} \hat{y}[m+i,n+j] & \text{if } m,n \text{ both even}\\ \hat{y}[m,n]-\hat{y}\left[2\left\lfloor\frac{m}{2}\right\rfloor,2\left\lfloor\frac{n}{2}\right\rfloor\right] & \text{otherwise}\end{cases}$$

Compression with predictive feedback in the transform domain



Compression with predictive feedback in the image domain



Reading Assignment

Taubman+Marcellin, Chapters 2.1 and 2.2