Digital Image Processing
CS390S

Feng Jiang
2018 Spring

METROPOLITAN STATE UNIVERSITY℠
OF DENVER

Human visual system & visual saliency

Visual attention

Innovative R&D by NTT

Visual attention is a built-in mechanism of the human visual system for scene understanding.

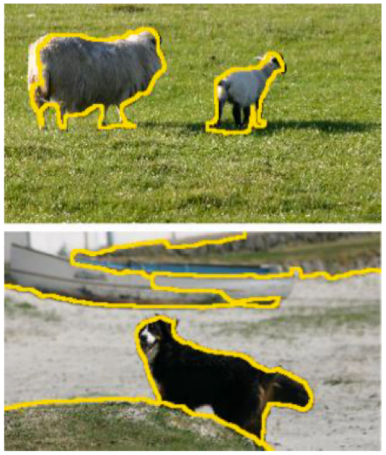
http://www.tobii.com/eye-tracking-research/global/library/white-papers/tobii-eye-tracking-white-paper/

Human visual system & visual saliency

Simulating visual attention is essential

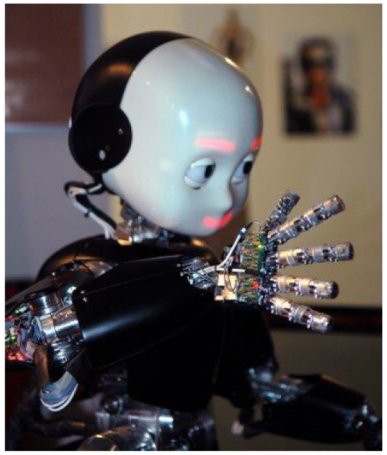
Innovative R&D by NTT

Such a pre-selection mechanism would be essential in enabling computers to undertake



[Donoser *et al*. 09]

Object detection



[http://www.icub.org]

HCI

[https://www.google.com/glass]

Visual assistance

Human visual system & visual saliency

Saliency as a measure of attention

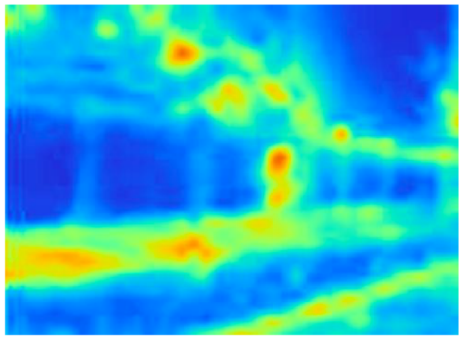
Innovative R&D by NTT

Saliency = attractiveness of visual attention

- Simple, easy to implement, reasonable outputs

Estimating human visual focus of attention

Input image



Saliency map [Itti *et al.* 98]

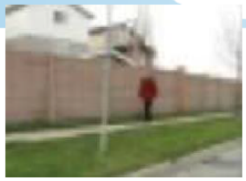
Human visual system & visual saliency

Related work

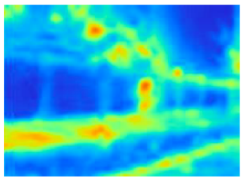
Innovative R&D by NTT

Visual saliency

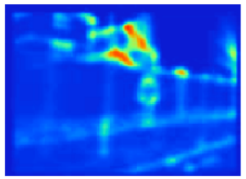
- Saliency map model [Itti 1998]
- Shannon self-information [Bruce 2005]
- Incorporating temporal dynamics [Itti 2009]



Input image



[Itti *et al*. 98]



[Bruce *et al*. 05]

Visual saliency

- Required paper reading (GBVS): “Graph-Based Visual Saliency” by Jonathan Harel, Christof Koch , Pietro Perona

- Papers for reading
- Two papers in Blackboard: saliency1 saliency2

- Code (GBVS)

Visual saliency

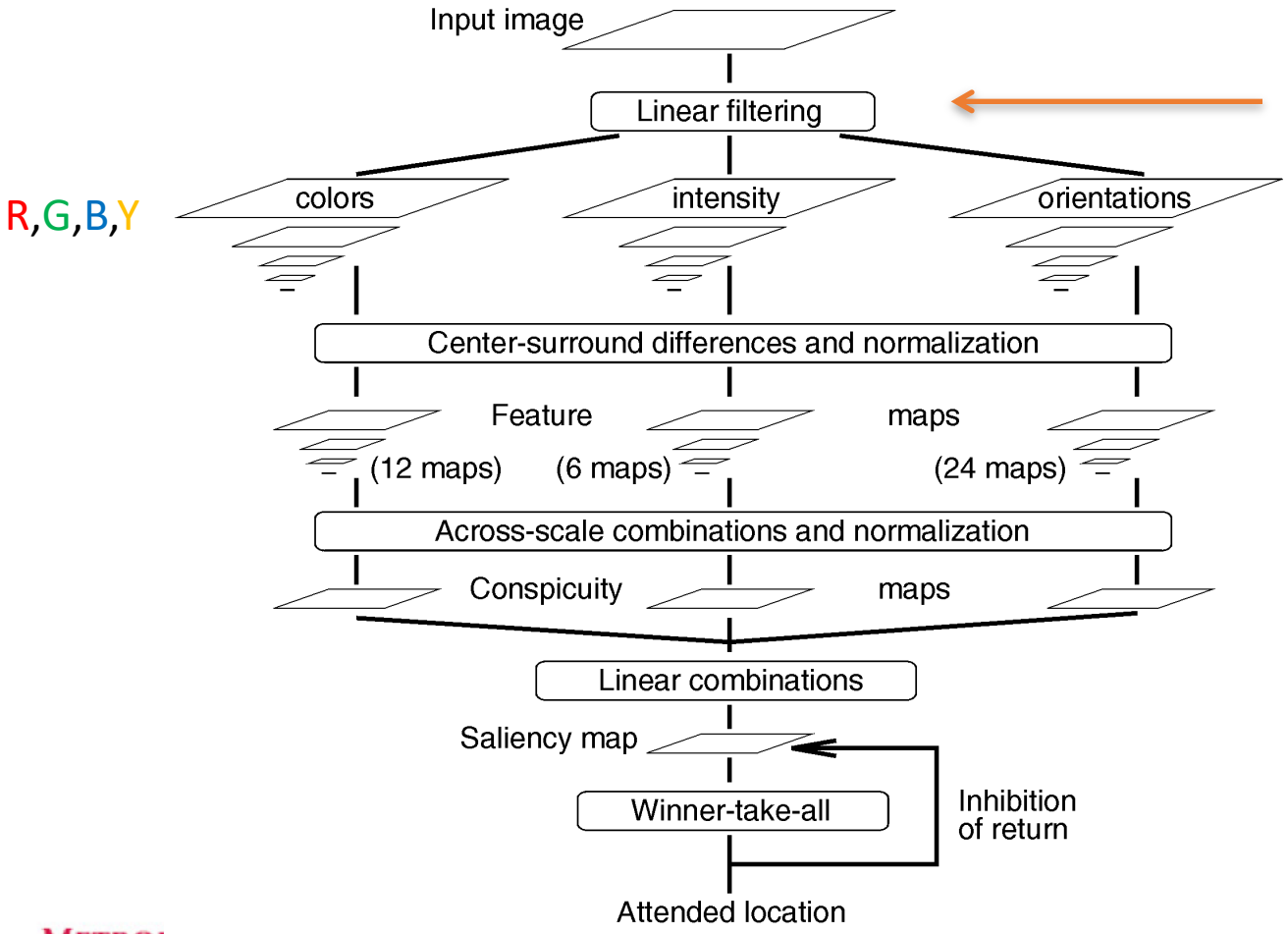
- Applications “SalApplications.pdf”

Visual saliency

- http://www.vision.caltech.edu/~harel/share/gbvs.php

- Code install
- Download GBVS from blackboard
- Install GBVS:
- http://www.vision.caltech.edu/~harel/share/gbvs/readme.txt

L. Itti's approach



Gaussian Pyramids

Gabor pyramids for $\theta = \{0º, 45º, 90º, 135º\}$

Question

Fig. 1. Reference Image I18.

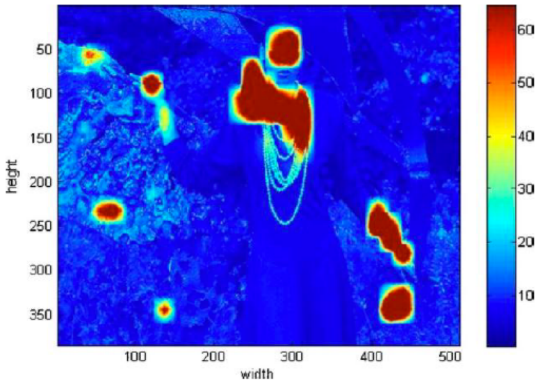


Fig. 2. Saliency map of I18 with face detection.

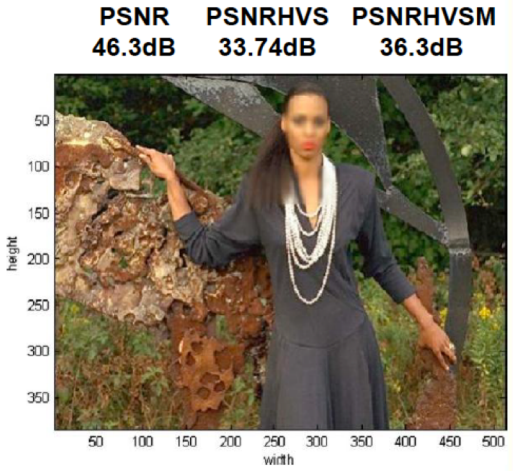


Fig. 3. I18 with noise in one salient region

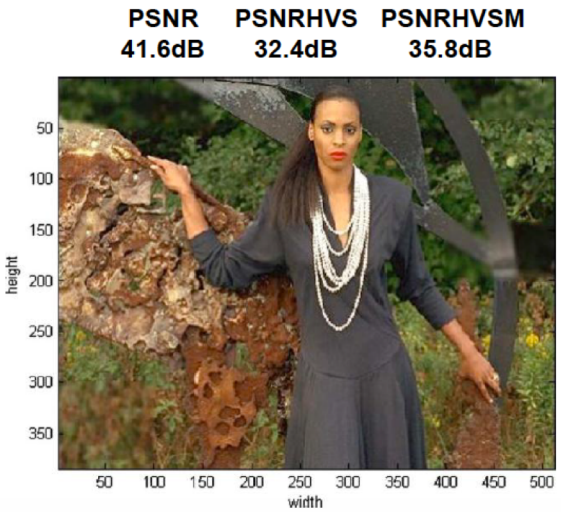


Fig. 4. I18 with noise in nonsalient region

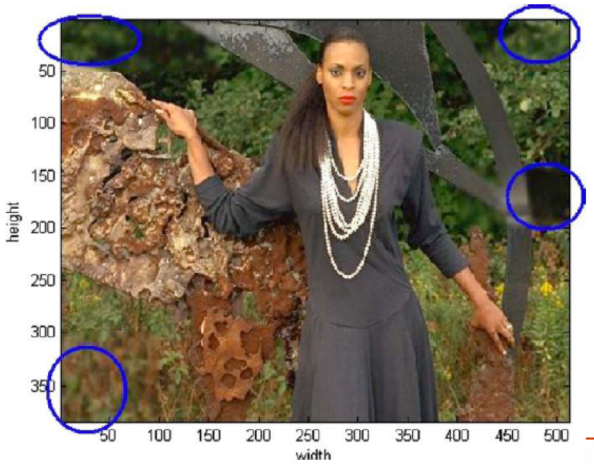


Fig. 5. I18 with distortion in four nonsalient region

Assignment

- Image with object
- Distortion at salient area & distortion at non-salient area
- Quality measurement

Review

- PSNR
- MSE
- SSIM ssim.m
- Your metric

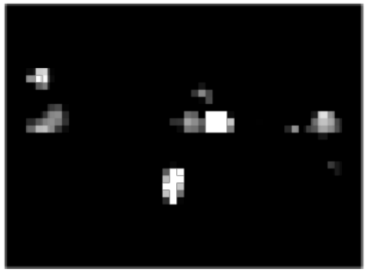
$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (x_{ij} - y_{ij})^2$$

$$\mathrm{PSNR} = 10 \log_{10} \frac{L^2}{\mathrm{MSE}}$$

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

Assignment

- Your quality metric based on saliency!



Data Compression

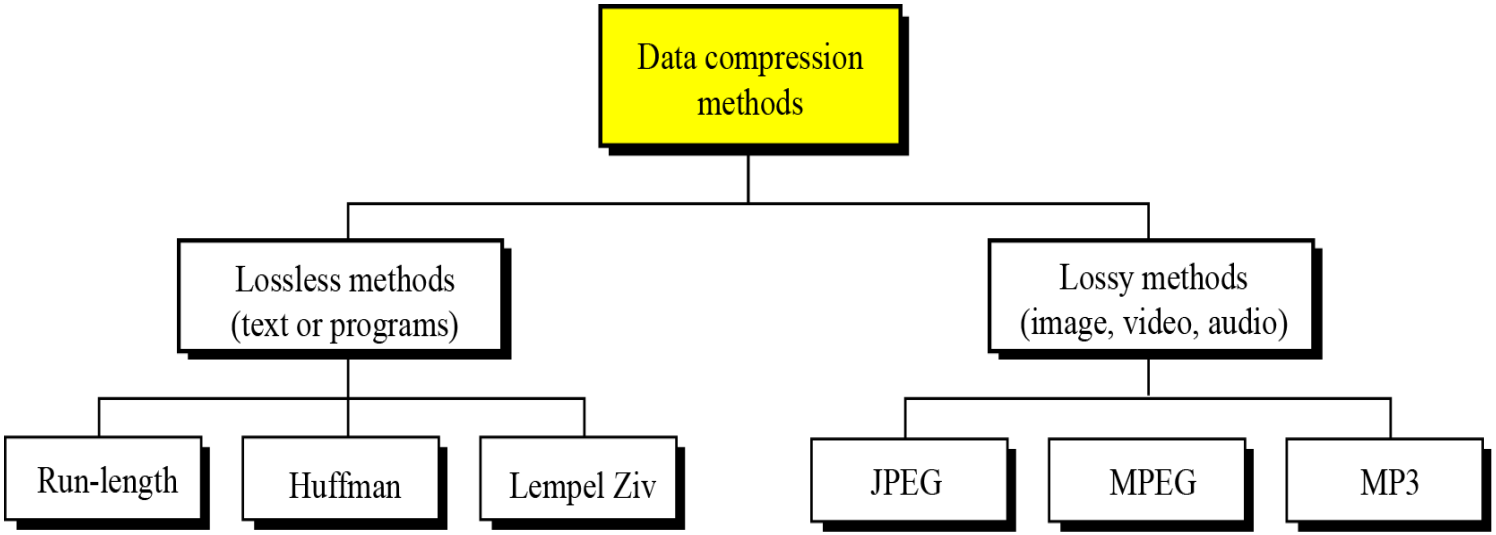
- Reason of compression /Motivation

- True reason of compression

  Data Redundancy

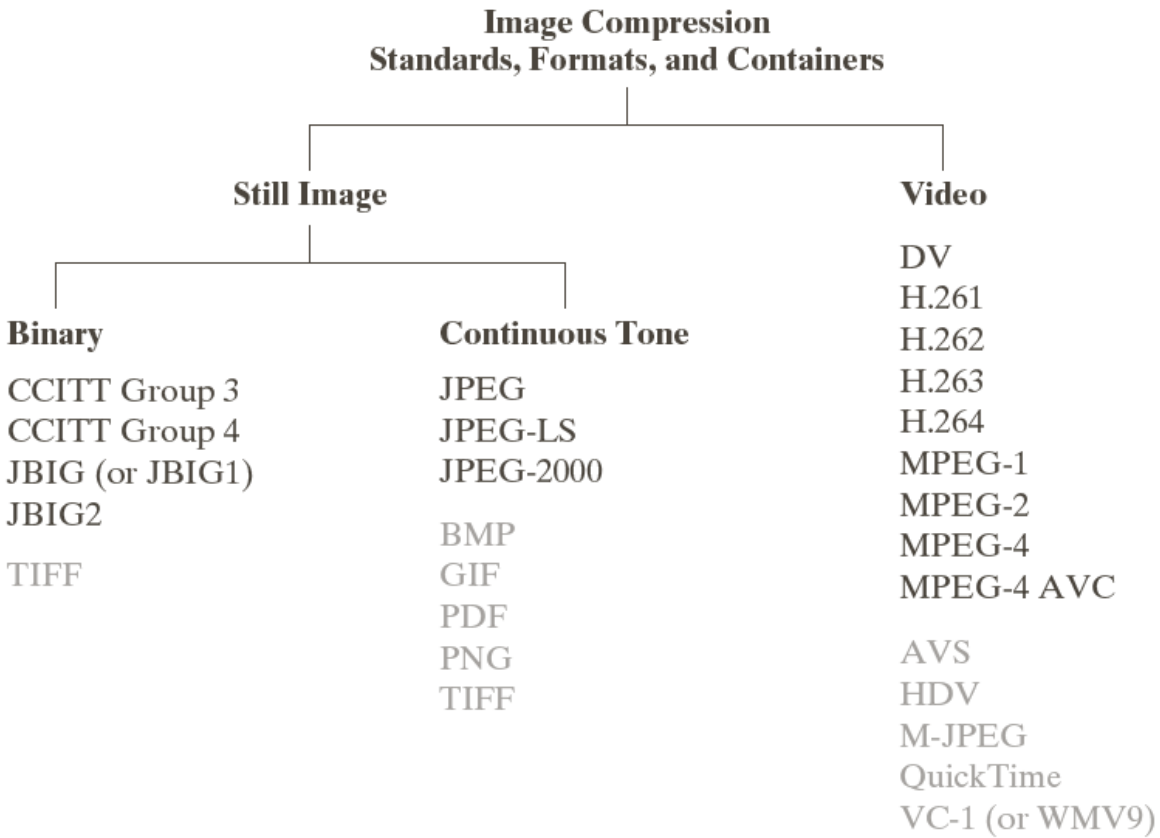
Data Compression

- Data compression is about storing and sending a smaller number of bits.
- There're two major categories for methods to compress data: lossless and lossy methods



Image/Video Compression

- Standards



Image/Video Compression

- Standards
- https://homepages.cae.wisc.edu/~ece533/project/f06/aguilera_rpt.pdf

- BMP (bitmap) uncompressed
- PNG (Portable Network Graphics) (1996) lossless
- (unique color, or with small variations of color)
- TIFF is a complicated format that incorporates an extremely wide range of options.
- JPEG
- JPEG2000
- JPEG-LS

Image Compression

- Data redundancy

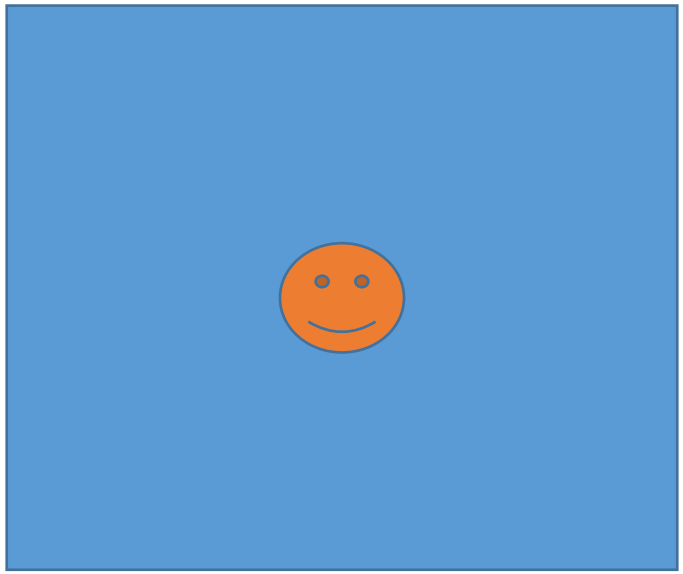


Image Compression

a b c

FIGURE 8.1 Computer generated 256 × 256 × 8 bit images with (a) coding redundancy, (b) spatial redundancy, and (c) irrelevant information. (Each was designed to demonstrate one principal redundancy but may exhibit others as well.)

Image Compression (chapter 8)

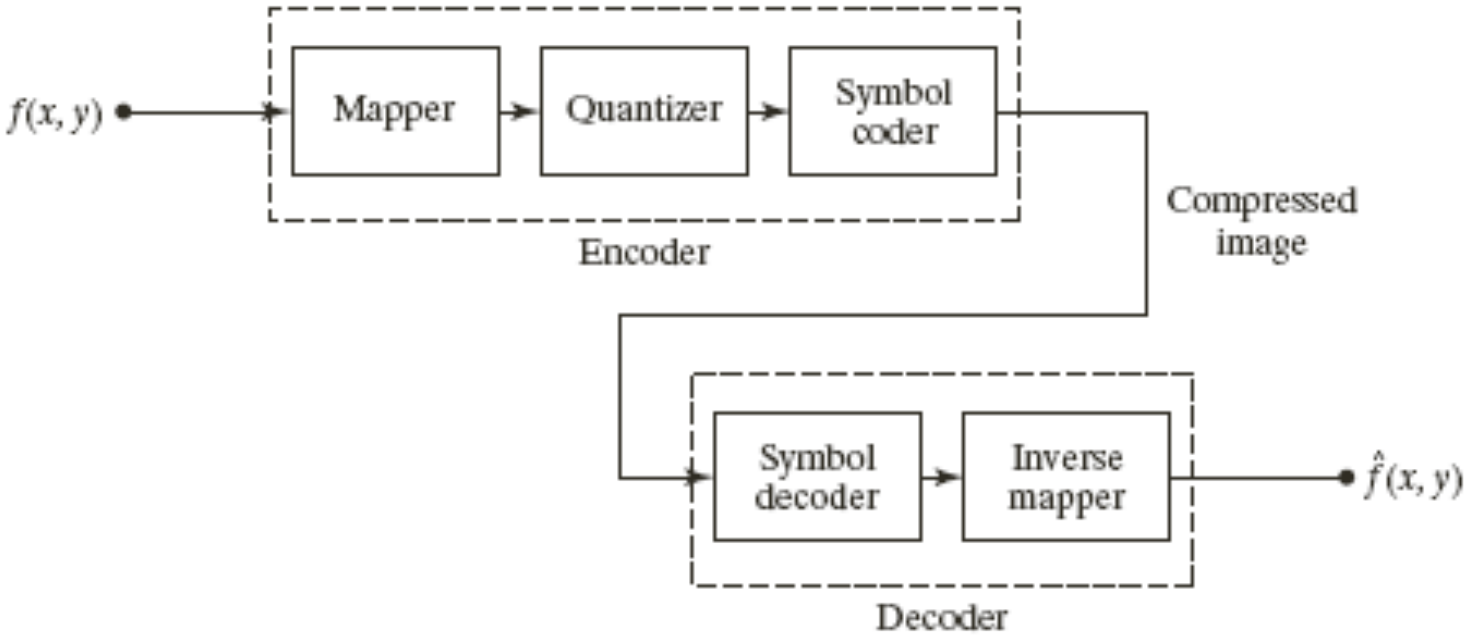
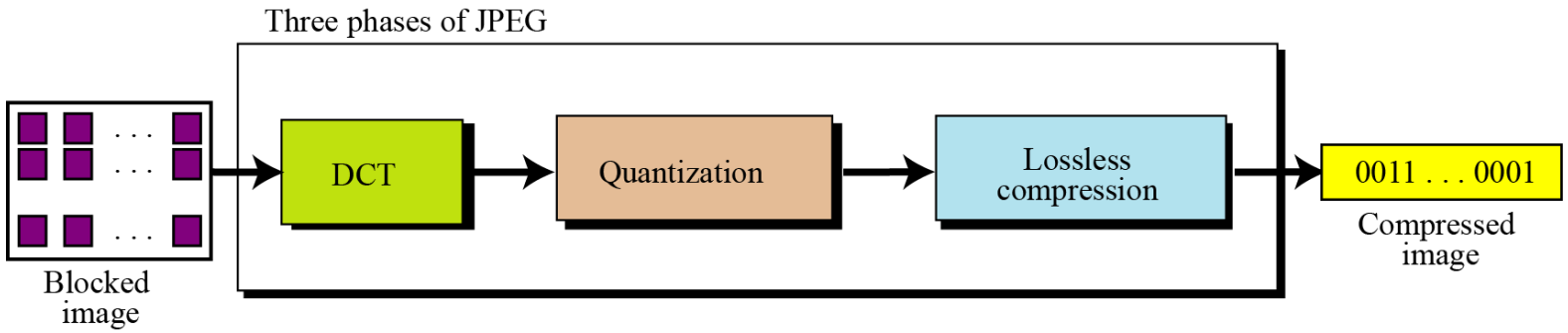


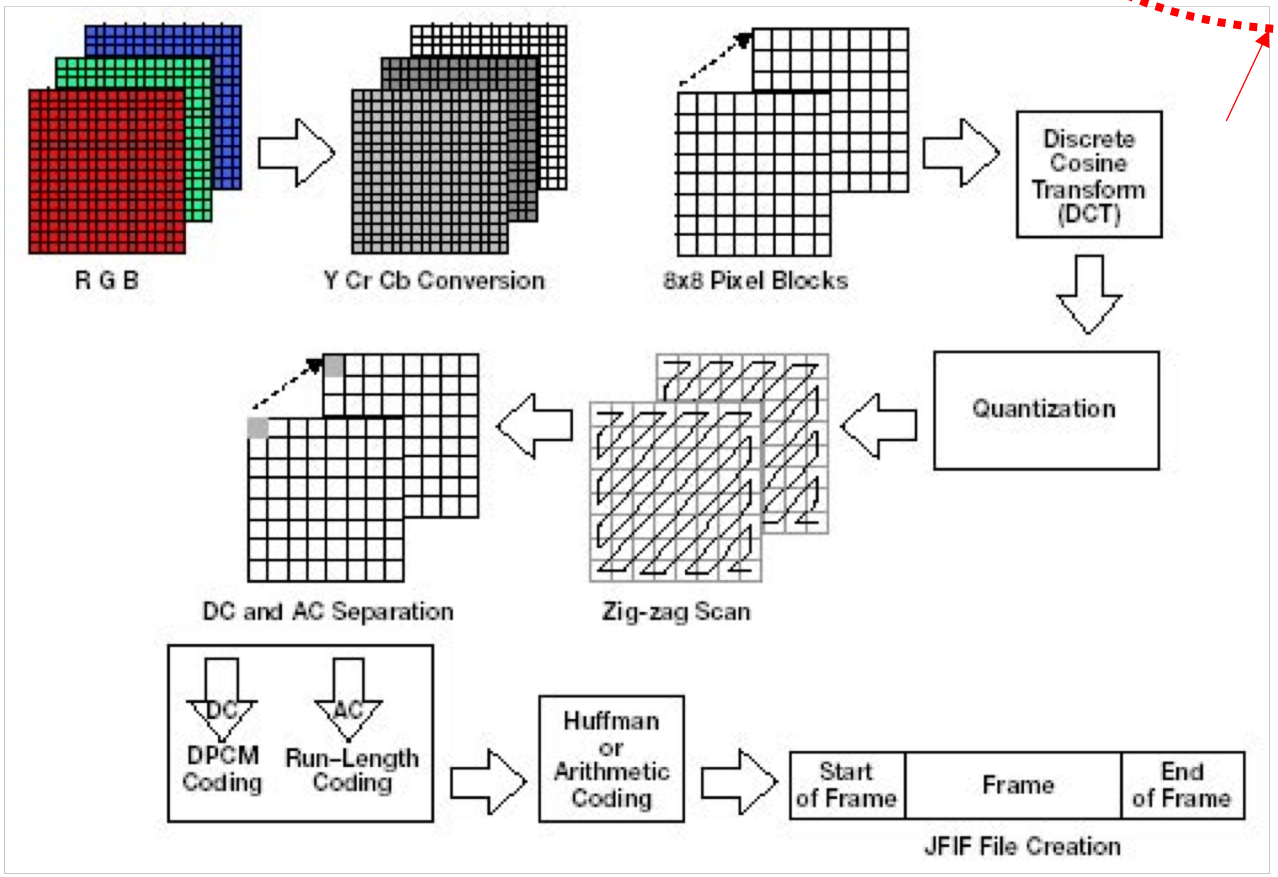
FIGURE 8.1
A general image compression system block diagram.

JPEG Encoding

- Used to compress pictures and graphics.
- In JPEG (Joint Photographic Experts Group), a grayscale picture is divided into 8x8 pixel blocks to decrease the number of calculations.
- Basic idea:
  - Change the picture into a linear (vector) sets of numbers that reveals the redundancies.
  - The redundancies is then removed by one of lossless compression methods.



JPEG Encoding



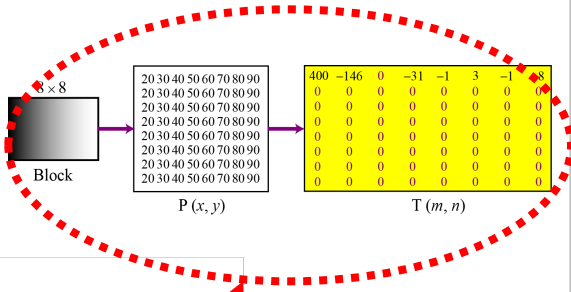


Image Compression

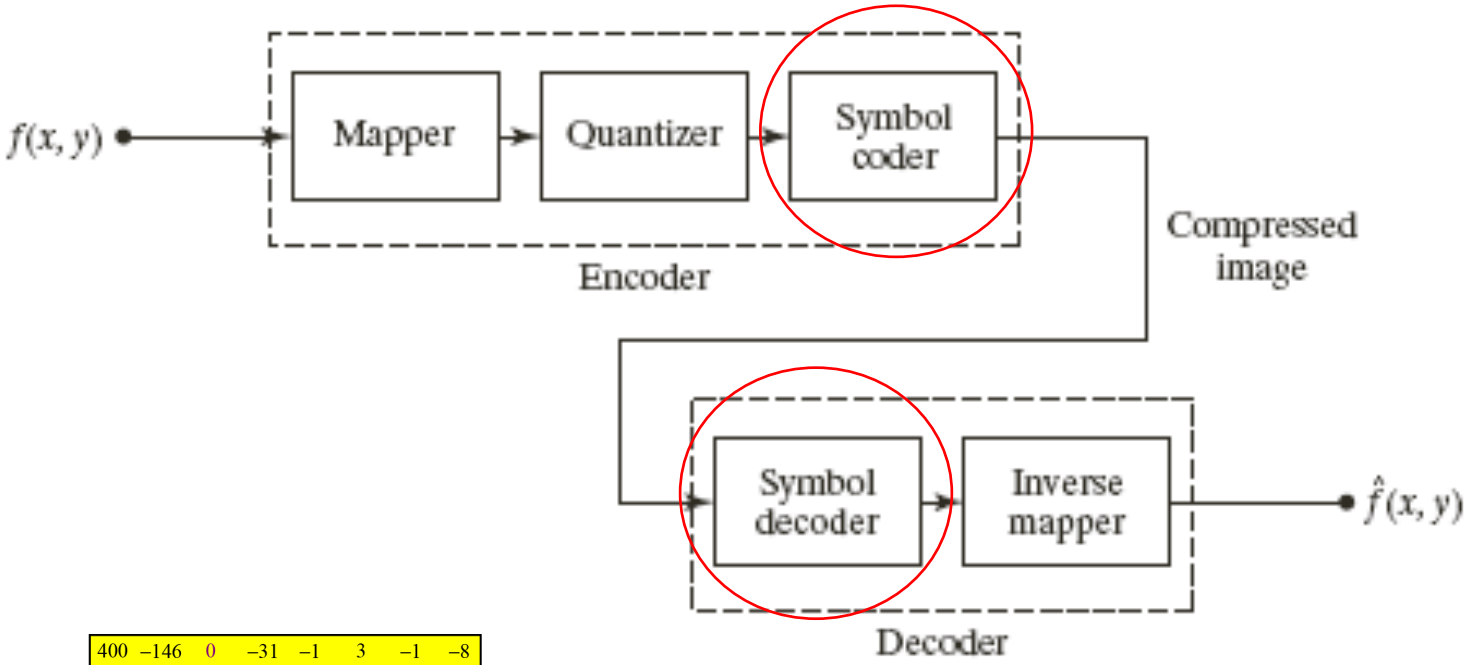


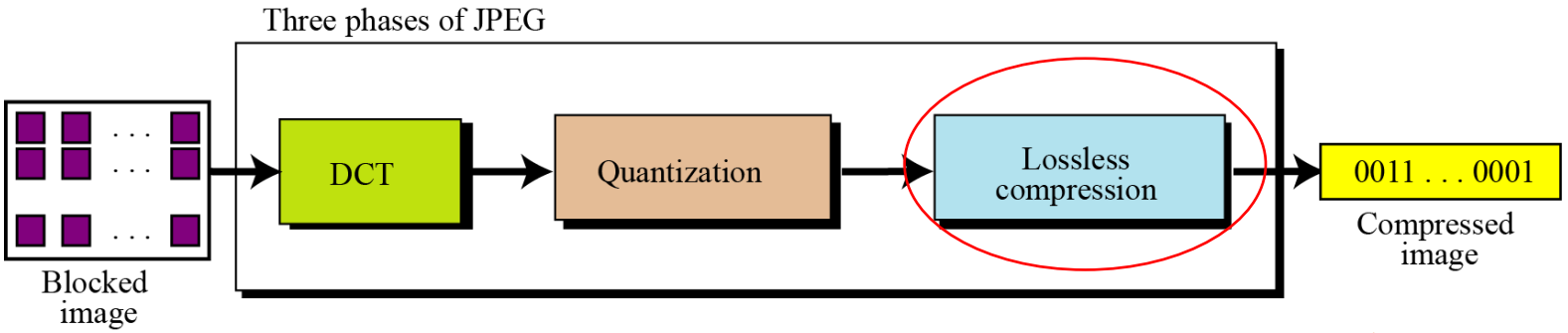
FIGURE 8.1 A general image compression system block diagram.

400	−146	0	−31	−1	3	−1	−8
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

T (m, n)

Image Compression

- Entropy encoding



400	−146	0	−31	−1	3	−1	−8
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

T (m, n)

fixed-length encoding

- count number of different symbols.
- ⌈lg M⌉ bits suffice to support M different symbols

Ex. genomic sequences
- 4 different codons
- 2 bits suffice

char	code
a	00
c	01
t	10
g	11

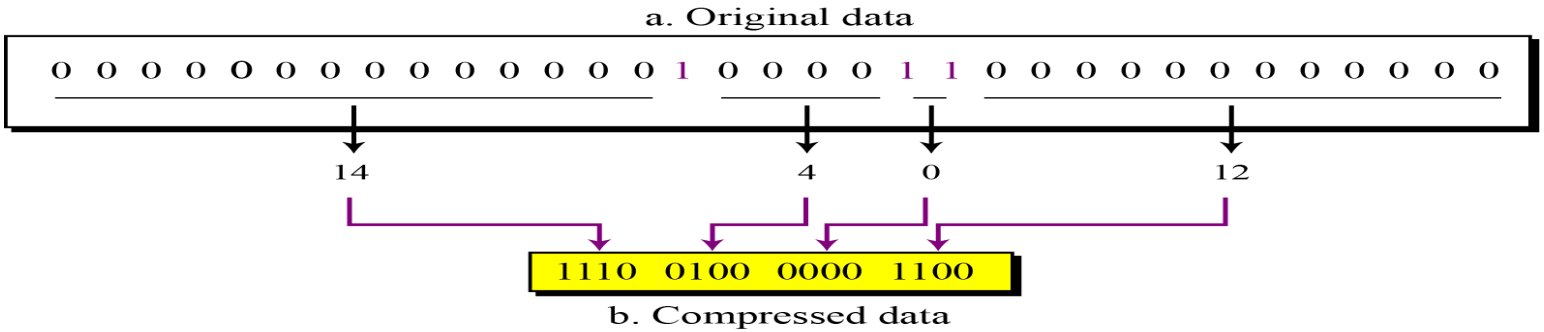
2N bits to encode genome with N codons

a	c	t	a	c	a	g	a	t	g	a
00	01	10	00	01	00	11	00	10	11	00

https://www.cs.princeton.edu/~rs/AlgsDS07/20Compression.pdf

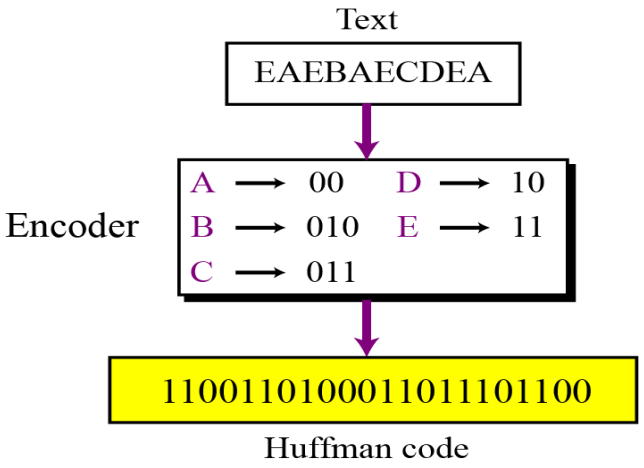
Run-length encoding

- Simplest method of compression.
- If the data uses only 2 symbols (0s and 1s) in bit patterns and 1 symbol is more frequent than another.



Huffman Coding

- Encoding



Prefix free

- Decoding

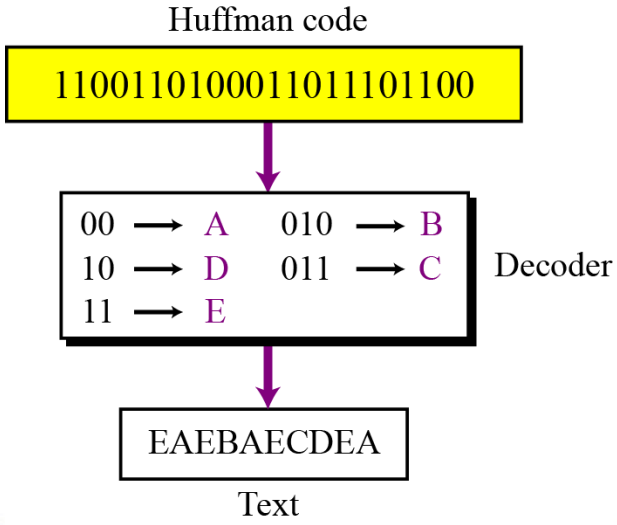


Image Compression

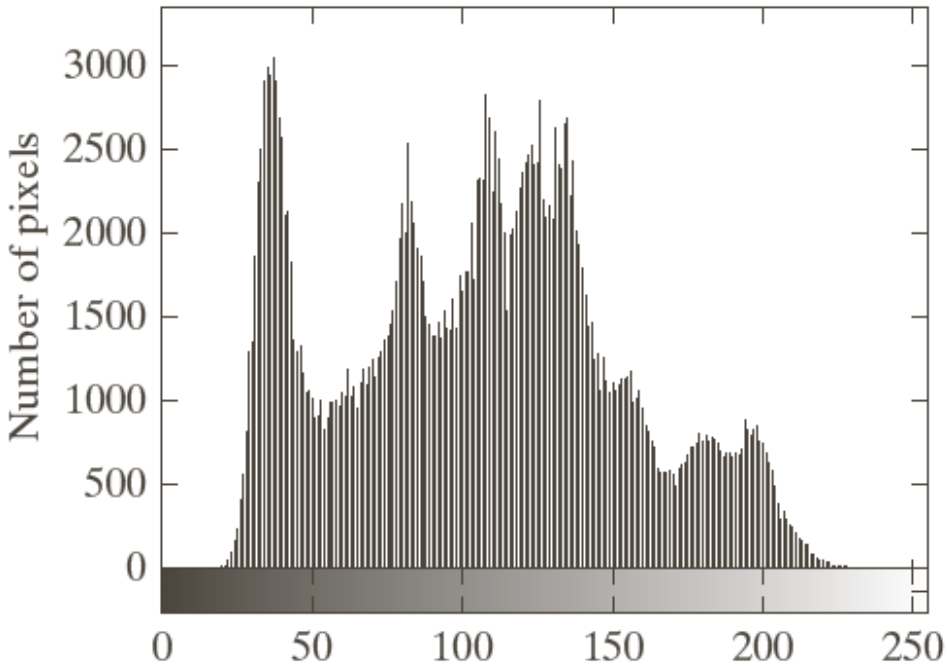


Image Compression

r_k	$p_r(r_k)$	Code 1	$l_1(r_k)$	Code 2	$l_2(r_k)$
$r_{87} = 87$	0.25	01010111	8	01	2
$r_{128} = 128$	0.47	10000000	8	1	1
$r_{186} = 186$	0.25	11000100	8	000	3
$r_{255} = 255$	0.03	11111111	8	001	3
r_k for $k \neq 87, 128, 186, 255$	0	—	8	—	0

Image Compression

r_k	$p_r(r_k)$	Code 1	$l_1(r_k)$	Code 2	$l_2(r_k)$
r_1	0.1875	00	2	011	3
r_2	0.5000	01	2	1	1
r_3	0.1250	10	2	010	3
r_4	0.1875	11	2	00	2

TABLE 8.1
Illustration of coding redundancy: $L_{avg} = 2$ for Code 1; $L_{avg} = 1.81$ for Code 2.

Image Compression

a
b

FIGURE 8.2
Huffman (a) source reduction and (b) code assignment procedures.

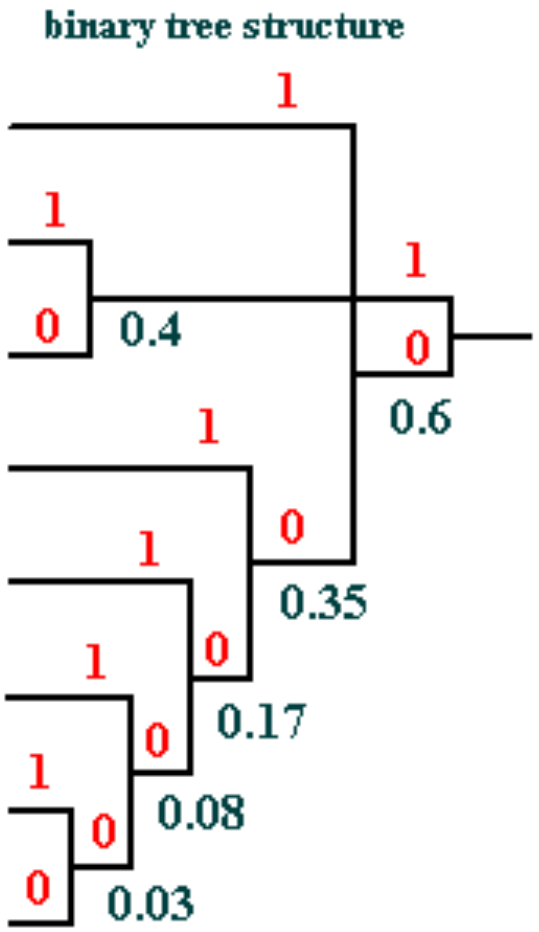
Original Source		Source Reduction	
Symbol	Probability	1	2
a_2	0.5	0.5	0.5
a_4	0.1875	0.3125	0.5
a_1	0.1875	0.1875	
a_3	0.125		

1
0
1
0
1
0

Original Source			Source Reduction			
Symbol	Probability	Code	1		2	
a_2	0.5	1	0.5	1	0.5	1
a_4	0.1875	00	0.3125	01	0.5	0
a_1	0.1875	011	0.1875	00		
a_3	0.125	010				

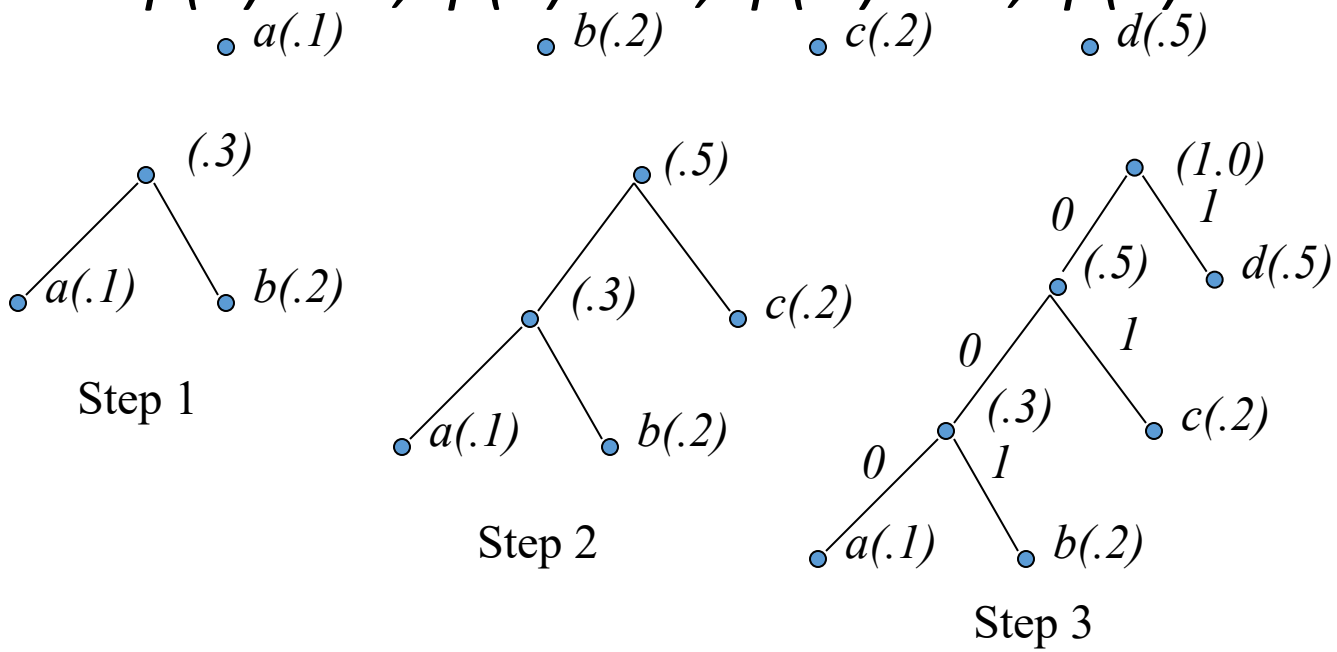
Huffman Coding

Original codewords	Huffman codewords	Symbol probabilities
111	01	P(h)=0.25
110	11	P(g)=0.2
101	10	P(f)=0.2
100	001	P(e)=0.18
011	0001	P(d)=0.09
010	00001	P(c)=0.05
001	000001	P(b)=0.02
000	000000	P(a)=0.01



Example

$p(a) = .1,\ p(b) = .2,\ p(c) = .2,\ p(d) = .5$



$a=000,\ b=001,\ c=01,\ d=1$

Compressing Digital Video

- Exploit *spatial redundancy* within frames (like JPEG: transforming, quantizing, variable length coding)
- Exploit *temporal redundancy* between frames
  - Only the sun has changed position between these 2 frames

Previous Frame

Current Frame

Image Compression

- http://www.ux.uis.no/~karlsk/ICTools/ictools.html

https://www.mathworks.com/help/comm/ref/huffmanenco.html