Digital Image Fundamentals

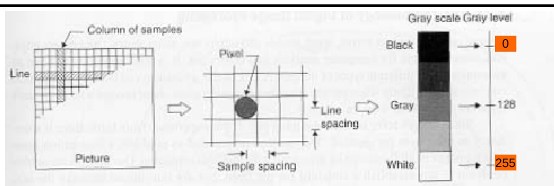
Xiaojun Qi
-- REU Site Program in CVMA
(2010 Summer)

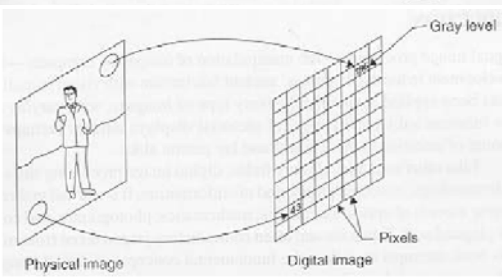
1

Outline

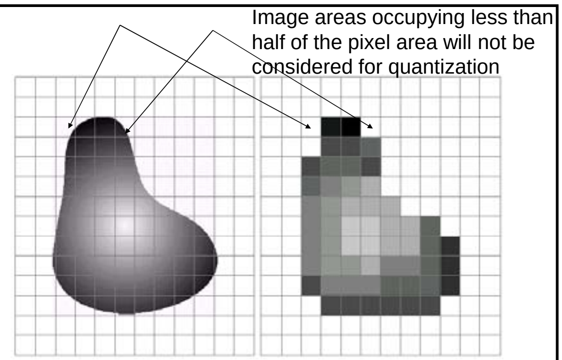
- Some basic concepts
  - Resolution
  - Pixel
  - Gray level intensity
  - Histogram
  - Color Space
- Color Features
- Texture Features
- Shape Features

2



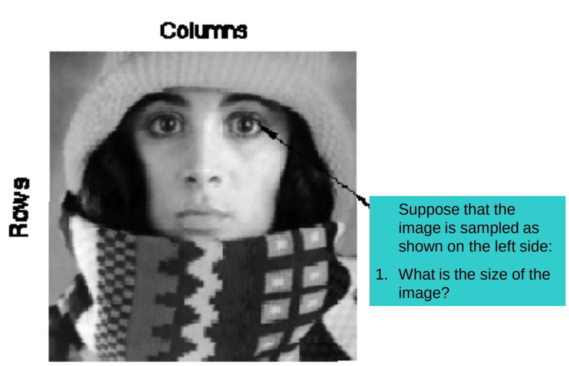


3



a b
FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.

4



5

Digital Images

- An image is a mapping of spatial coordinates (x, y) into values identified with a set of gray levels.
- Pixel or picture element is a triple
  (x, y, g(x, y)) = (x, y, L) where L is a gray level. The number of gray levels typically is **an integer power of 2** due to processing, storage, and sampling hardware considerations.
- The gray level 256 is fairly common because the data can be stored in one byte.

6

DIP Concepts

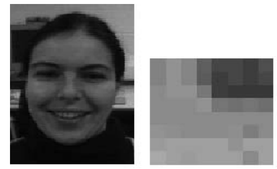
- Image:

  An image may be defined as a 2-D function, *f(x,y)*, where *x* and *y* are spatial coordinates, and the amplitude of *f* at any pair of coordinates *(x, y)* is called the **intensity** or **gray level** of the image at that point.

- Digital Image:

  x, *y*, and the amplitude values of *f* are all **finite, discrete quantities.**

7



(a) (b)

130	146	133	95	71	71	62	78
130	146	133	92	62	71	62	71
139	146	146	120	62	55	55	55
139	139	139	116	117	112	117	110
139	139	139	139	139	139	139	139
116	112	139	139	139	143	124	139
156	159	159	159	159	146	159	159
168	159	156	159	159	159	139	159

(c) A 8*8 array

(a) Image of a face (b) subimage from the right eye region
(c) Matrix notation of the subimage

8

DIP Concepts (Cont.)

- Pixels (Picture Elements, Image Elements, Pels):

  The elements in a digital image with a particular location and value.

- Digital Image Processing:

  Manipulation and analysis of digital images (pictorial information) by computer.

9

An M*N digital image can be represented in the following compact matrix form:

$$f(x,y)=\begin{bmatrix} f(1,1) & f(1,2) & \ldots\ldots & f(1,N) \\ f(2,1) & f(2,2) & \ldots\ldots & f(2,N) \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ f(M,1) & f(M,2) & \ldots\ldots & f(M,N) \end{bmatrix}$$

- Common choices for *M*, *N* and *L*

$M = 2^n$, $N = 2^k$, and $L = 2^m$

examples for *M*, *N*: 256 x 256 or 512 x 512

examples for *L*: 256 (m=8 bits/pixel), 4096 (m=12 bits/pixel)

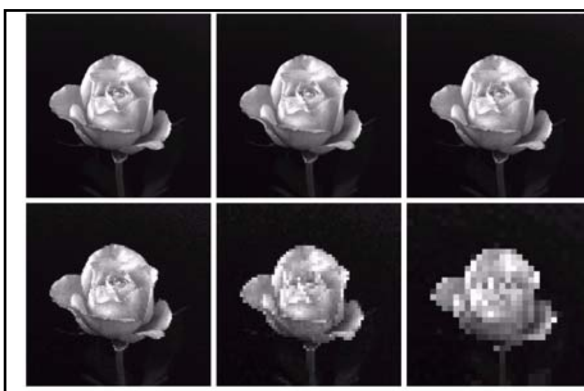
- Storage requirements

N x *M* x *m* bits (m=8, N=M=1024, 1MB image)

10

- Pixel is the smallest unit in an image. It depends on the sampling.
- Resolution is the smallest number of discernible line pairs per unit distance.
  - Dense sampling produces a high resolution image
  - Coarse sampling produces a low resolution image
- Spatial resolution of an image is the physical size of a pixel in that image. That is, spatial resolution is the smallest discernible detail in an image.
  - A digital image of size M by N has a spatial resolution of M * N pixels.
- Gray-level resolution of an image is the number of gray-levels in that image. It depends on the number of bits used for quantization. It also determines the gray-level (intensity) range.
  - An L-level digital image has a gray-level resolution of L level.

11



a b c
d e f

Ex1: Effects of varying the number of samples in a digital image while keeping the number of gray levels constant

FIGURE 2.20 (a) 1024 × 1024, 8-bit image. (b) 512 × 512 image resampled into 1024 × 1024 pixels by row and column duplication. (c) through (f) 256 × 256, 128 × 128, 64 × 64, and 32 × 32 images resampled into 1024 × 1024 pixels.

One Possible Implementation (a→b→c→d→e→f)

Step 1: Subsampling: Delete every other row and column from the original image.

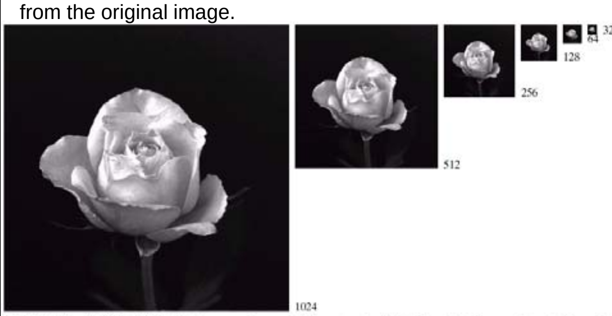


FIGURE 2.19 A 1024 × 1024, 8-bit image subsampled down to size 32 × 32 pixels. The number of allowable gray levels was kept at 256.

Step 2: Scale up: Bring all the subsampled images up to the size of original image by **row and column pixel replication**.

13

Ex2: Effects of varying the number of gray levels in a digital image while keeping the spatial resolution constant

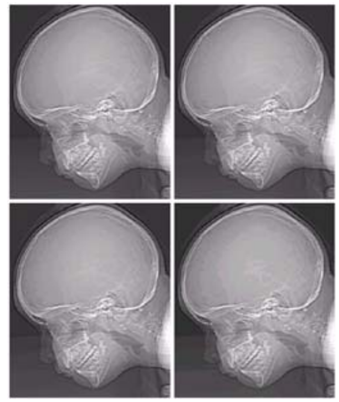
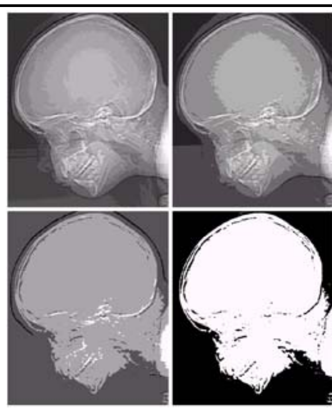


FIGURE 2.21 (a) 452 × 374, 256-level image. (b)–(d) Image displayed in 128, 64, and 32 gray levels, while keeping the spatial resolution constant.

14

FIGURE 2.21 (*Continued*) (e)–(h) Image displayed in 16, 8, 4, and 2 gray levels. (Original courtesy of Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)



It is accomplished by using: Original Intensity * (2^k-1) / 255 where k is the decreased bits for the gray level.

15

Three Coordinate Systems

1. Cartesian Coordinate: Origin is at lower left corner and usually starts with (0, 0).
2. Array (Matrix) Coordinate: Origin is at upper left corner and usually starts with (1, 1). Also refers to row-column (i.e., x-y) coordinate systems.
3. Pixel Coordinate: Origin is at upper left corner and usually starts with (1, 1). The first component x increases to the right, while the second component y increases downward.

16

Digital Image Notations

- (x, y): spatial coordinates [or p=(x, y): a point]
- g(x, y), or g(p): image gray-level at (x, y)
- (x, y, g(x,y)): Pixel
- G: Image Grid
- Nx: Number of x's on the image grid (Height)
- Ny: Number of y's on the image grid (Width)
- Nt=Nx * Ny: Total number of pixels on the image grid.
- L: The number of gray-levels. The gray-levels range from 0 to L-1.

17

Neighbors and Neighborhood

- Np: The neighbors of a pixel p
  - $N_4(p)$: 4 horizontal and vertical neighbors of p
  - $N_D(p)$: 4 diagonal neighbors of p
  - $N_8(p)$: 4 horizontal and vertical neighbors of p plus 4 diagonal neighbors of p; that is $N_4(p)$ and $N_D(p)$
- Wp: The window of a pixel p

  It is the neighborhood of a pixel p with certain gray-level values
  - Wp = {Np, g(w)} where g(w) = g(p') for all p' in Np

18

Sample Matlab Codes

Calling Matlab Function bwlabel(a, 4) returns:

1	1	1	0	0	0	0	0
1	1	1	0	1	1	0	0
1	1	1	0	1	1	0	0
1	1	1	0	0	0	1	0
1	1	1	0	0	0	1	0
1	1	1	0	0	0	1	0
1	1	1	0	0	1	1	0
1	1	1	0	0	0	0	0

→

1	1	1	0	0	0	0	0
1	1	1	0	2	2	0	0
1	1	1	0	2	2	0	0
1	1	1	0	0	0	3	0
1	1	1	0	0	0	3	0
1	1	1	0	0	0	3	0
1	1	1	0	0	3	3	0
1	1	1	0	0	0	0	0

Adjacency:
4-Adjacency: p and q with values from a set V are 4-adjacent if q is in the set N4(p).
8-Adjacency: p and q with values from a set V are 8-adjacent if q is in the set N8(p).

19

Multi-channel Images and Color

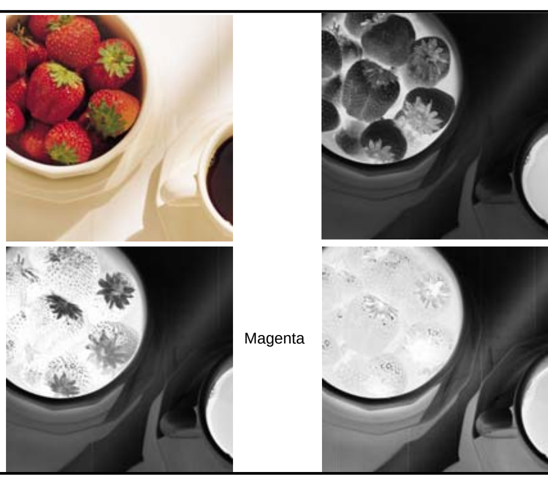
- N-channel or N-band Image:
  - On a common referenced coordinate systems, there is a sequence of gray-level images [g1(x,y), g2(x,y), …, gn(x, y)]
- Color Model: Represent colors and their relationship to each other.

20

- RGB (red-green-blue): It is an additive model.
  - F(x,y) = (r, g, b)
- CMY (cyan-magenta-yellow): It is a subtractive model since all three primaries are subtracted from white light to produce the required color.
- YIQ: Recoding of RGB for transmission efficiency and for maintaining compatibility with monochrome television standards.
- HIS (hue-intensity-saturation): Decouple the intensity component from the color information. H is the dominant color, S is the degree of non-dilution in white, and I is the relative brightness.

21

22



23

24

Color Models
-- HSI Color Model

- Given an image in RGB color format, the H component of each RGB pixel is obtained by using the equation:

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases}$$

with

$$\theta = \cos^{-1}\left\{ \frac{\frac{1}{2}\left[(R-G)+(R-B)\right]}{\left[(R-G)^2+(R-B)(G-B)\right]^{1/2}} \right\}$$

How about the denominator is zero?

25

Color Models
-- HSI Color Model

- The saturation component is given by

$$S = 1 - \frac{3}{(R+G+B)}\left[\min(R,G,B)\right]$$

- The intensity component is given by

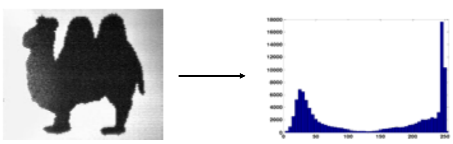
$$I = \frac{1}{3}(R+G+B)$$

- Here the RGB values have been normalized to the range [0, 1] and that the angle theta is measured with respect to the red axis of the HSI space.

26

Color Features: Color Histogram

- **Histogram** (gray-scale images)



- –Invariant to translation, rotation, and small variations
- –Normalized histogram is invariant to scale
- –Not very sensitive to noise
- –But: removes a lot of information!

27

Histogram

- The **histogram** of a digital image with gray levels in the range [0, L-1] is a discrete function $h(r_k) = n_k$, where r_k is the kth gray level and n_k is the number of pixels in the image having gray level r_k.
- A **normalized histogram** is given by $p(r_k) = n_k/n$ for k = 0, 1, …, L-1 and n is the total number of pixels in the image. That is, $p(r_k)$ gives an estimate of the probability of occurrence of gray level r_k.

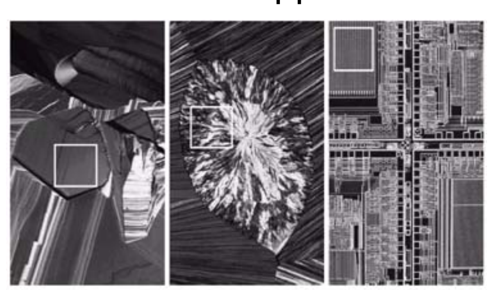
28

Texture Features

- The three principal approaches used in image processing to describe the texture of a region are statistical, structural, and spectral.
  - Statistical approaches yield characterizations of textures as smooth, coarse, grainy, and so on.
  - Structural techniques deal with the arrangement of image primitives, such as the description of texture based on regularly spaced parallel lines.
  - Spectral techniques are based on properties of the Fourier spectrum and are used primarily to detect global periodicity in an image by identifying high-energy, narrow peaks in the spectrum.

29

Texture Features: Statistical Approach



a b c

FIGURE 11.22 The white squares mark, from left to right, smooth, coarse, and regular textures. These are optical microscope images of a superconductor, human cholesterol, and a microprocessor. (Courtesy of Dr. Michael W. Davidson, Florida State University.)

30

TABLE 11.2 Texture measures for the subimages shown in Fig. 11.22.

Texture	Mean	Standard deviation	R (normalized)	Third moment	Uniformity	Entropy
Smooth	82.64	11.79	0.002	−0.105	0.026	5.434
Coarse	143.56	74.63	0.079	−0.151	0.005	7.783
Regular	99.72	33.73	0.017	0.750	0.013	6.674

$$\text{Mean}: m = \sum_{i=0}^{L-1} z_i p(z_i)$$

$$\text{Standard Deviation}: \sigma = \sqrt{\sum_{i=0}^{L-1} (z_i - m)^2 p(z_i)}$$

$$\text{R}: R = 1 - \frac{1}{1+\sigma^2(z)}$$

$$\text{Third Moment}: \mu_3(z) = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i)$$

$$\text{Uniformity}: U = \sum_{i=0}^{L-1} p^2(z_i)$$

$$\text{Entropy}: e = -\sum_{i=0}^{L-1} p(z_i)\log_2 p(z_i)$$

Standard Deviation is a measure of gray-level contrast that can be used to establish descriptors of relative smoothness.

R: Normalized variance in the range of [0, 1]

Third moment is a measure of the skewness of the histogram.

Uniformity: Histogram based measure.

Average Entropy is a measure of variability and is 0 for a constant image.

31

Shape Features: Fourier Descriptors

- Fourier Transform of the Signature s(t):

$$u_n = \frac{1}{N}\sum_{t=0}^{N-1} s(t) e^{-j2\pi nt/N}$$

$$\text{for } n = 0, 1, \ldots, N-1$$

The complex coefficients u_n are called the Fourier descriptors of the boundary, and are denoted as FDn.

32

Shape Signatures: Complex Coordinates

- The boundary can be represented as the sequence of coordinates s(t) = [x(t), y(t)] for t = 0, 1, 2, …, N-1, where x(t) = x_t and y(t) = y_t; (x_t, y_t)'s are encountered in traversing the boundary in the counterclockwise direction and N is the total number of points on the boundary.

Z(t) = x(t) + iy(t)

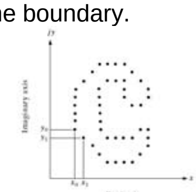


FIGURE 11.13 A digital boundary and its representation as a complex sequence. The points (x_0, y_0) and (x_1, y_1) shown are (arbitrarily) the first two points in the sequence.

33

Shape Feature: Fourier Descriptors (Cont.)

- The inverse Fourier transform of these coefficients restores s(t). That is:

$$s(t) = \sum_{n=0}^{N-1} u_n e^{j2\pi nt/N} \quad \text{for t} = 0, 1, \ldots, \text{N-1}$$

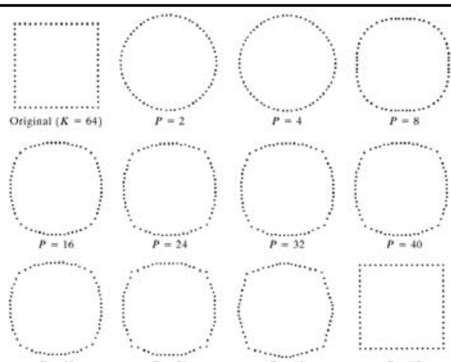
- Suppose that only the first P coefficients are used (that is, setting u_n = 0 for n > P-1). The result is the following approximation to s(k):

$$\hat{s}(t) = \sum_{n=0}^{P-1} u_n e^{j2\pi nt/N}$$

for t = 0, 1, 2, …, N-1.

34

FIGURE 11.14 Examples of reconstruction from Fourier descriptors. *P* is the number of Fourier coefficients used in the reconstruction of the boundary.



The goal is to use a few Fourier descriptors to capture the gross essence of a boundary. These coefficients carry shape information and can be used as the basis for differentiating between distinct boundary shapes.

35

Transformation	Boundary	Fourier Descriptor
Identity	$s(k)$	$a(u)$
Rotation	$s_r(k) = s(k)e^{j\theta}$	$a_r(u) = a(u)e^{j\theta}$
Translation	$s_t(k) = s(k) + \Delta_{xy}$	$a_t(u) = a(u) + \Delta_{xy}\delta(u)$
Scaling	$s_s(k) = \alpha s(k)$	$a_s(u) = \alpha a(u)$
Starting point	$s_p(k) = s(k - k_0)$	$a_p(u) = a(u)e^{-j2\pi k_0 u/K}$

$$s_t(k) = [x(k) + \Delta x] + j[y(k) + \Delta y]$$

$$s_p(k) = x(k - k_0) + jy(k - k_0)$$

1) Magnitude |FDn| is translation and rotation invariant
2) |FD0| carries scale-information
3) "Low-frequency" terms (*t* small): smooth behavior
4) "High-frequency" terms (*t* large): jaggy, bumpy behavior

36

Shape Feature: -- Normalized Fourier Descriptor

$$f = [\frac{|FD_1|}{|FD_0|}, \frac{|FD_2|}{|FD_0|}, ..., \frac{|FD_m|}{|FD_0|}]$$

Why?

When two shapes are compared, m=N/2 coefficients are used for central distance, curvature and angular function.

m=N coefficients are used for complex coordinates.

$$d = \sqrt{\sum_{i=1}^{m} (f_i^q - f_i^t)^2}$$

$$where \quad f_q = (f_q^1, f_q^2, ..., f_q^m) \text{ and } f_t = (f_t^1, f_t^2, ..., f_t^m)$$

are the feature vectors of the two shapes respectively.

37

Matlab

- Arguments to functions are passed by reference (i.e., pointers) in most cases to conserve memory. If an argument is modified by the function, then the argument is passed by value, which means a copy of the argument is made.
- For example the file Change.m contains the following:

```
function B = Change(A)
A = A + 10;
B=A;
```

38

Useful Matlab Commands

- imread, imwrite, rgb2gray, rgb2hsv
- imhist, hist
- size, length, ones, zeros
- repmat, reshape, find
- rand, randn, randperm,
- whos, help
- for, while, if, elseif, else, end
- display, pause
- cputime, tic, toc
- :, ;, .*, '

39