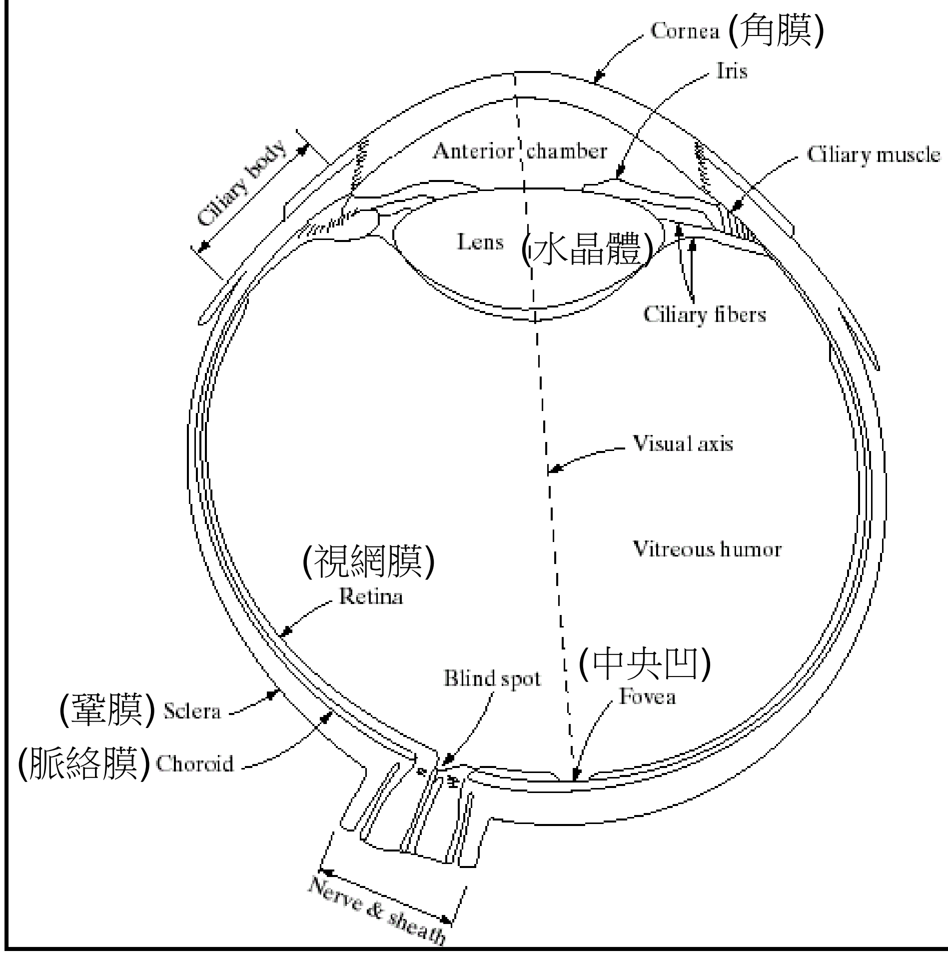
Chap 2 Digital Image Fundamentals

2.1 Elements of Visual Perception

- Although the digital image processing built on a foundation of mathematical and probabilistic formulations, human intuition and analysis generally play a central role in the choice of one technique versus another, and this choice often is made based on subjective, visual judgements.
- Developing a basic understanding of human visual perception as a first step is appropriate. The interest is in the mechanics and parameters related to how images are formed and perceived by humans.
- Physical limitations of human vision are also important.

2.1 Elements of Visual Perception

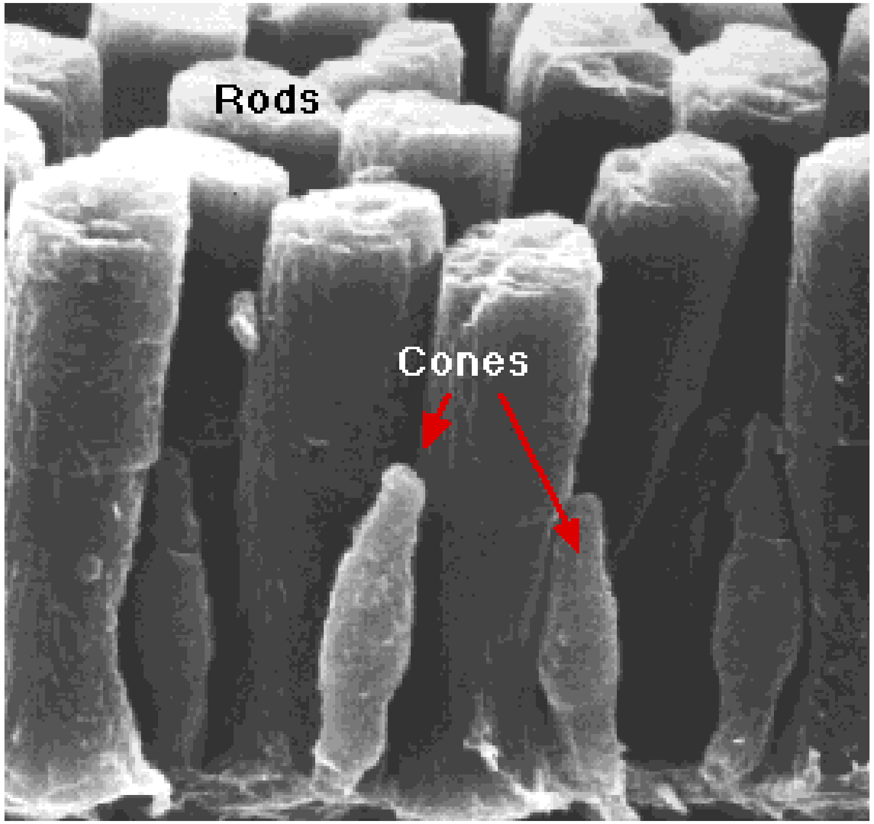


- A simplified diagram of a cross section of the human eye.
- The eye is nearly a sphere, with an average diameter of approximately 20 mm.

2.1 Elements of Visual Perception

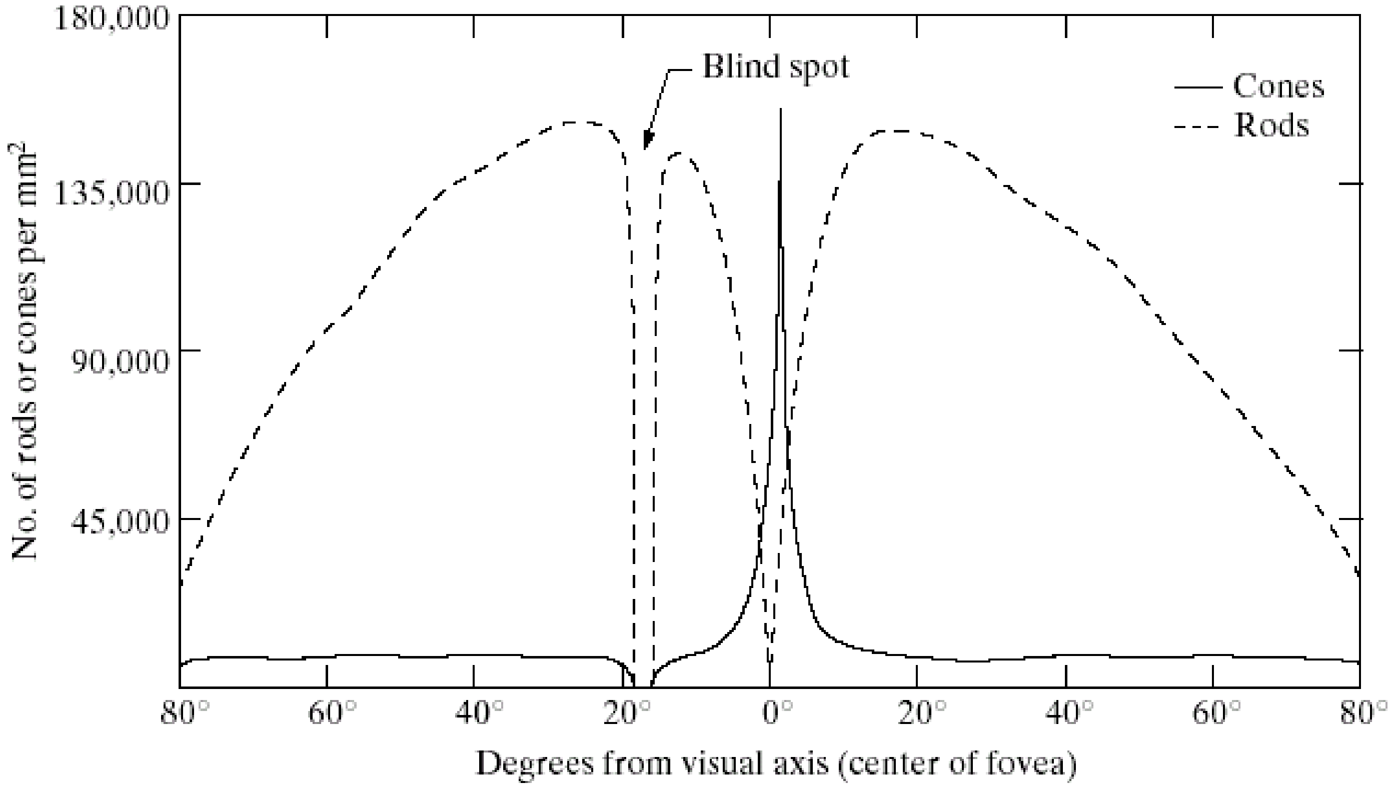
- Three membranes enclose the eye: cornea and sclera outer over; choroid ; retina
- There are two classes of receptors in retina: cones and rods
- The cones are located primarily in the central portion of the retina, called the fovea, and are highly sensitive to color. Human can resolve fine details with these cones largely.
- The rods are distributed over the retinal surface. Rods serve to give a general, overall picture of the field of view.

2.1 Elements of Visual Perception



(Ref: http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/V/Vision.html)

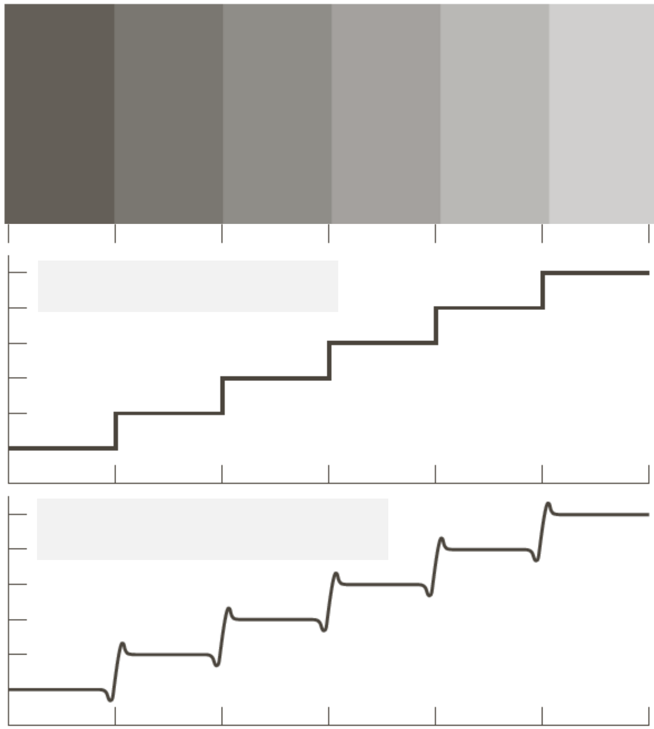
2.1 Elements of Visual Perception



6

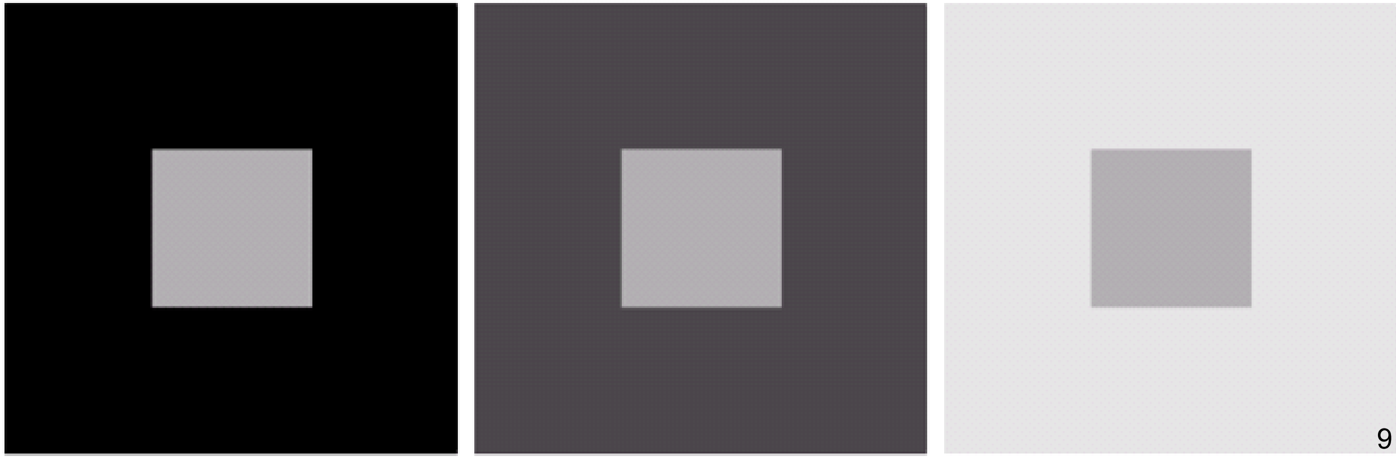
2.1.3 Brightness Adaptation and Discrimination

- Two phenomena demonstrate that **human perceived brightness is not a simple function of intensity**.
- The first one is based on the fact that the visual system tends to undershoot or overshoot around the boundary of regions of different intensities.
- The figure shows the Mach band effect, in which the perceived intensity is not simple function of actual intensity.

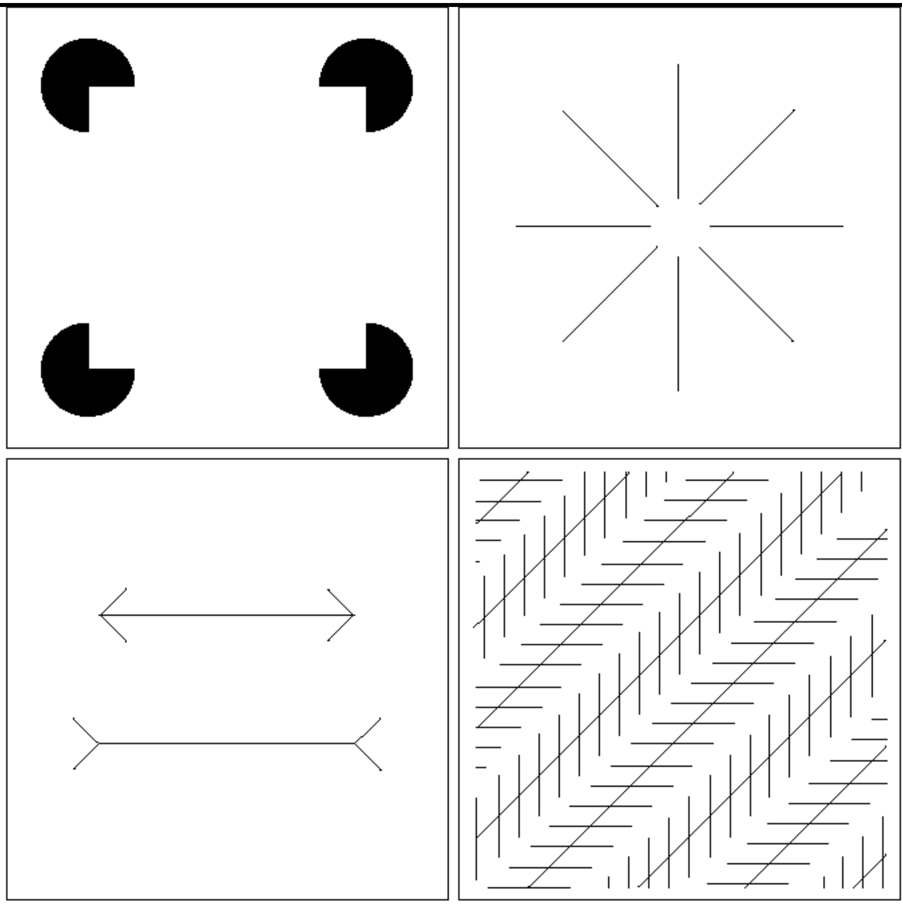


2.1.3 Brightness Adaptation and Discrimination

- The second phenomenon, called simultaneous contrast (聯合對比度)
  - A region's perceived brightness does not depend simply on its intensity.
  - Ex., in this figure, all the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

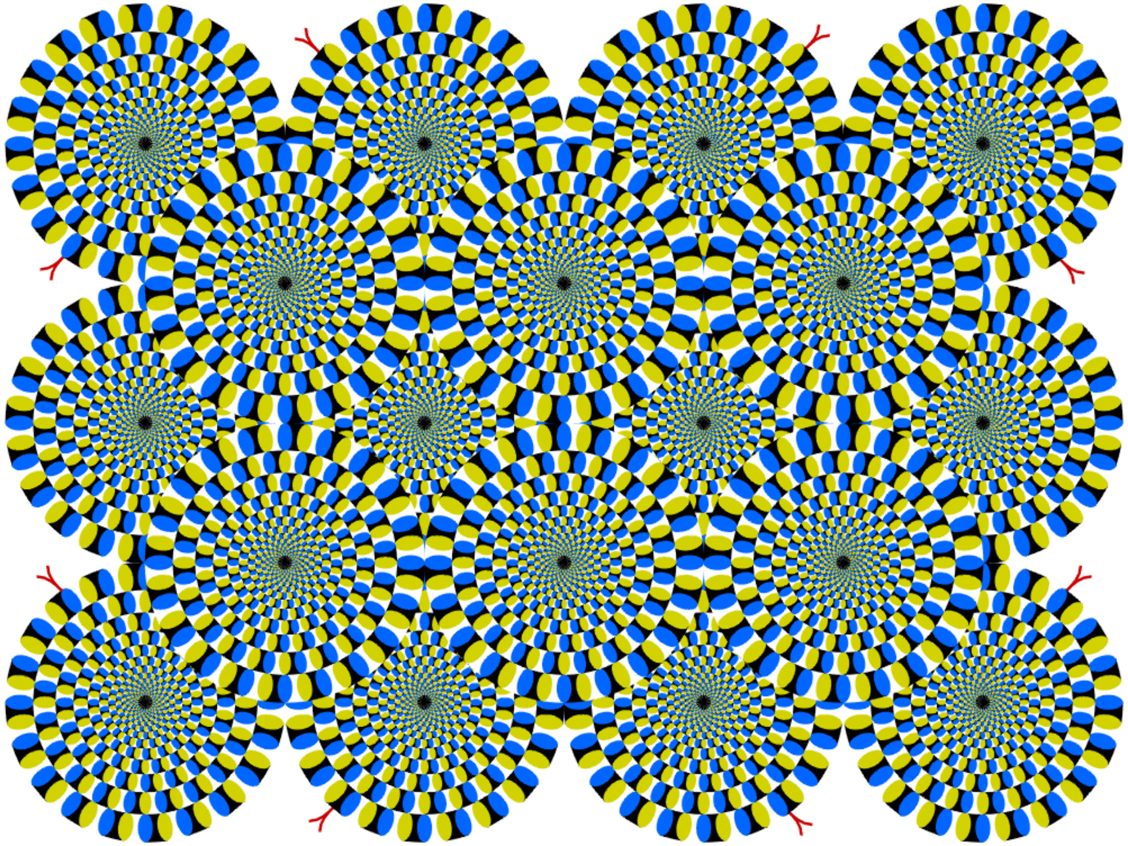


9



- Other examples of human perception phenomena are **optical illusions**: the eye fills in non-existing information or wrongly perceives geometrical properties of objects.

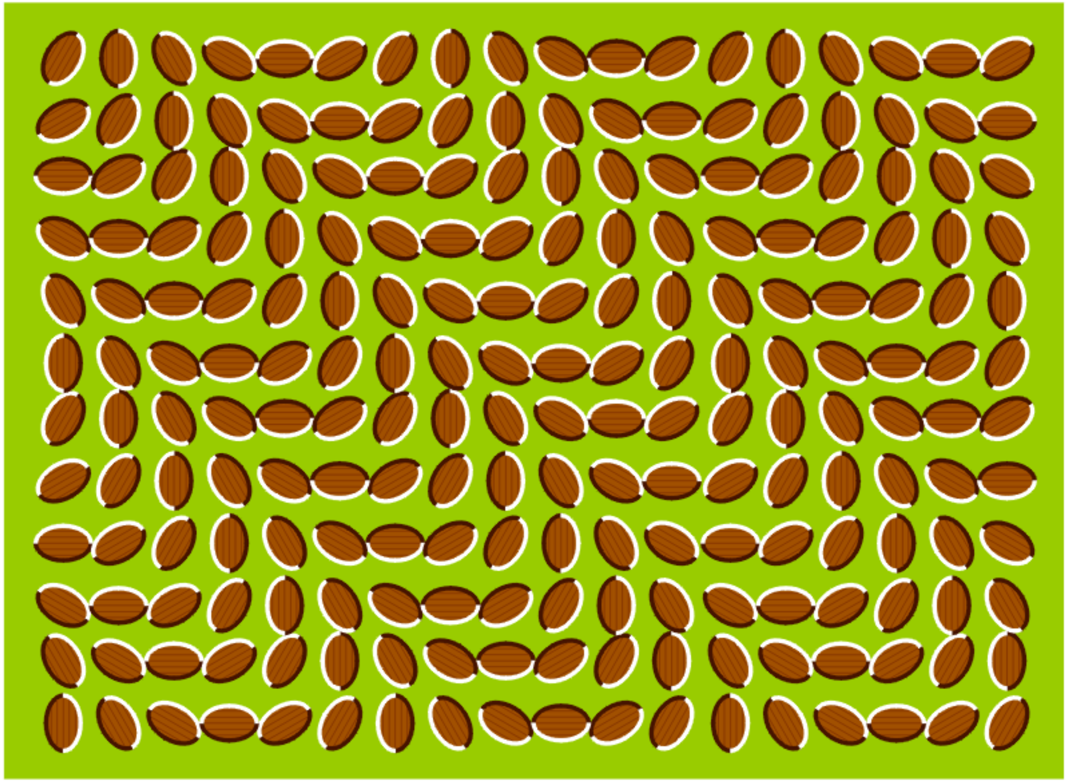
Illusion



(Ref: http://www.ritsumei.ac.jp/~akitaoka/index-e.html)

11

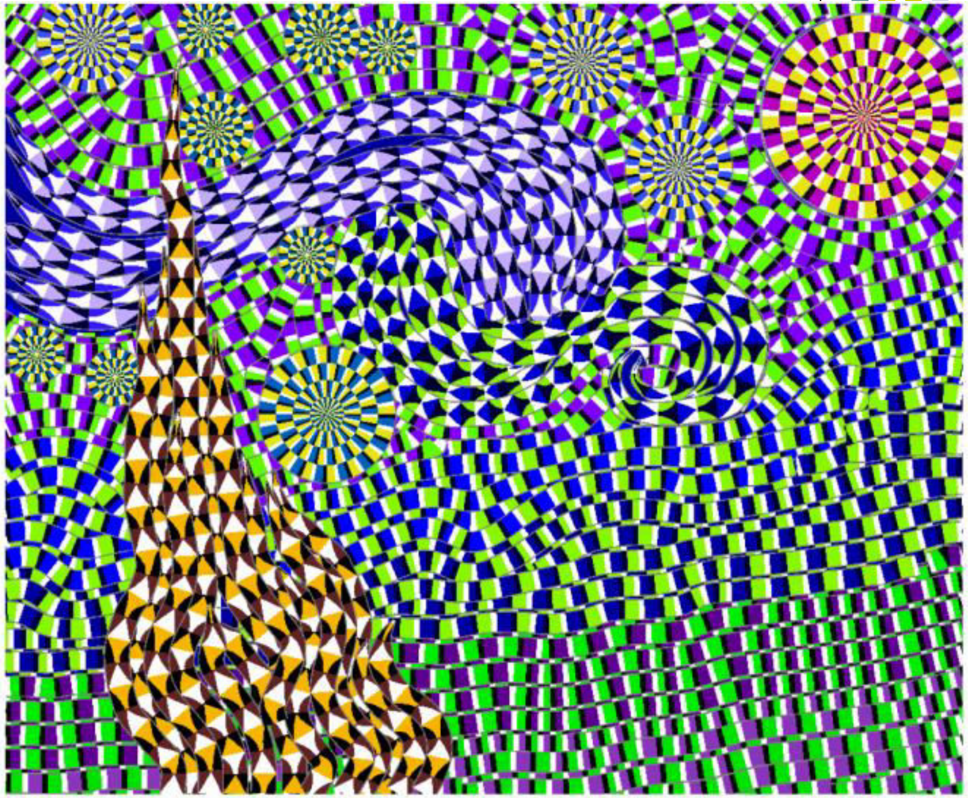
Illusion



(**Refs:** 1. T.N. Cornsweet, *Visual Perception,* Harcourt Brace Jovanovich, 1970
2. D.H. Hubel, *Eye, Brain, and Vision,* Scitific American, 1988)

Illusion

The Starry Night
by van Gogh

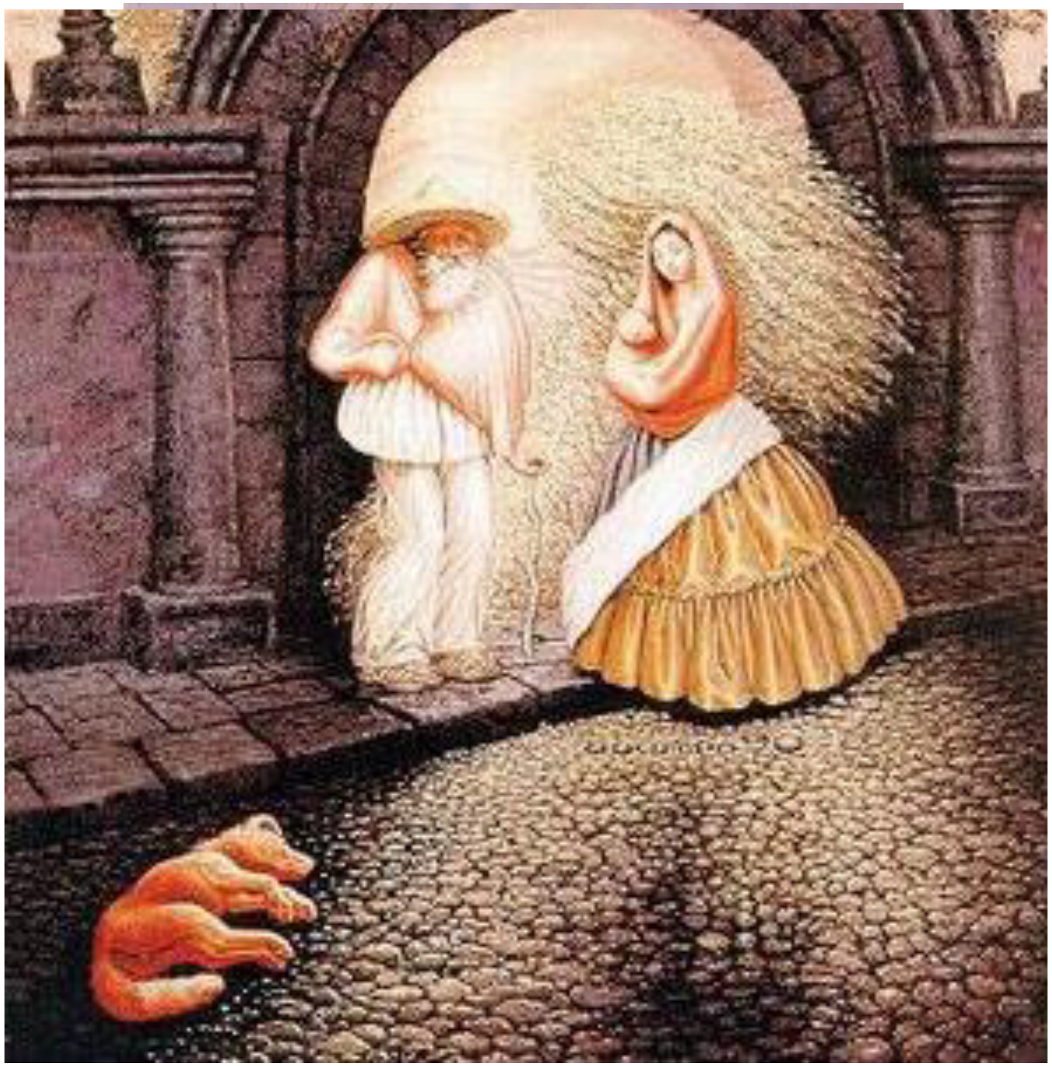


(**References:** Self-Animating Images: Illusory Motion Using Repeated Asymmetric Patterns, SIGGRAPH 2008)

13

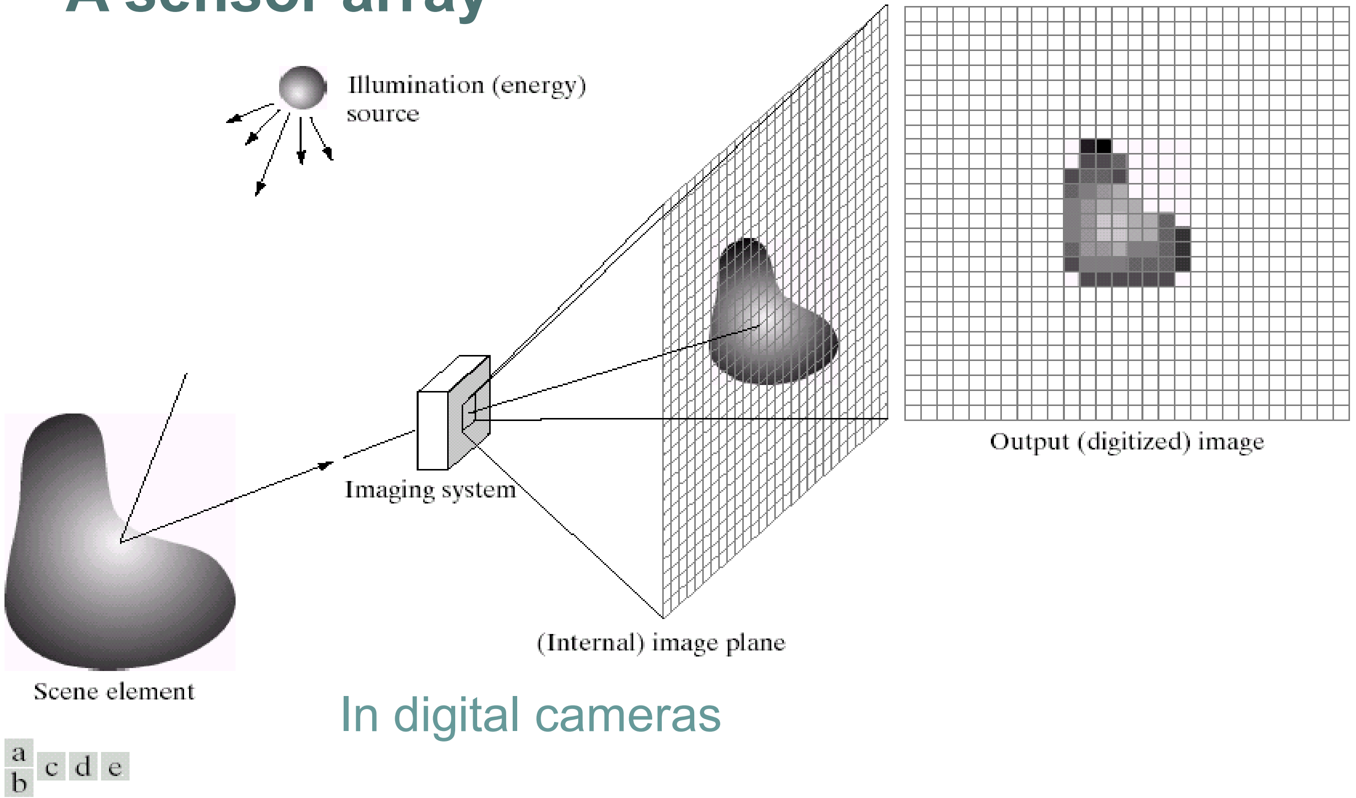

„Stillvergnügt"
„Délice silencieux"

15



2.3 Image Sensing and Acquisition

A sensor array



In digital cameras

FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

2.3 Image Sensing and Acquisition

- Images are denoted by two-dimensional functions of the form $f(x, y)$. The value of f at spatial coordinates (x, y) is a positive scalar quantity whose physical meaning is determined by the source of the image.
- **A simple image formation model** is given by

$$\boldsymbol{f(x,y) = i(x,y)r(x,y) + n(x,y)}$$

  - $f(x, y)$: Image function
  - $i(x, y)$: Illumination function
  - $r(x, y)$: Reflectance function $(0.0 < r(x, y) < 1.0)$
  - $n(x, y)$: Random noise function

  $L_{min} \leq$gray level $f(x, y) \leq L_{max}$

  $0 \leq$gray level $f(x, y) \leq L - 1$, **gray level normalization**

- $i(x, y)$ is determined by illumination source and $r(x, y)$ is determined by the characteristics of the imaged objects.

2.4 Image Sampling and Quantization

- To acquire digital images from the continuous sensed data $f(x,y)$:
  - Digitization in coordinate values: **Sampling**
  - Digitization in amplitude values: **Quantization**
- The resulting image has M rows and N columns as

$$f(x,y)=\begin{bmatrix} f(0,0) & f(0,1) & f(0,N-1) \\ f(1,0) & f(1,1) & f(1,N-1) \\ & & \\ f(M-1,0) & f(M-1,1) & f(M-1,N-1) \end{bmatrix}$$

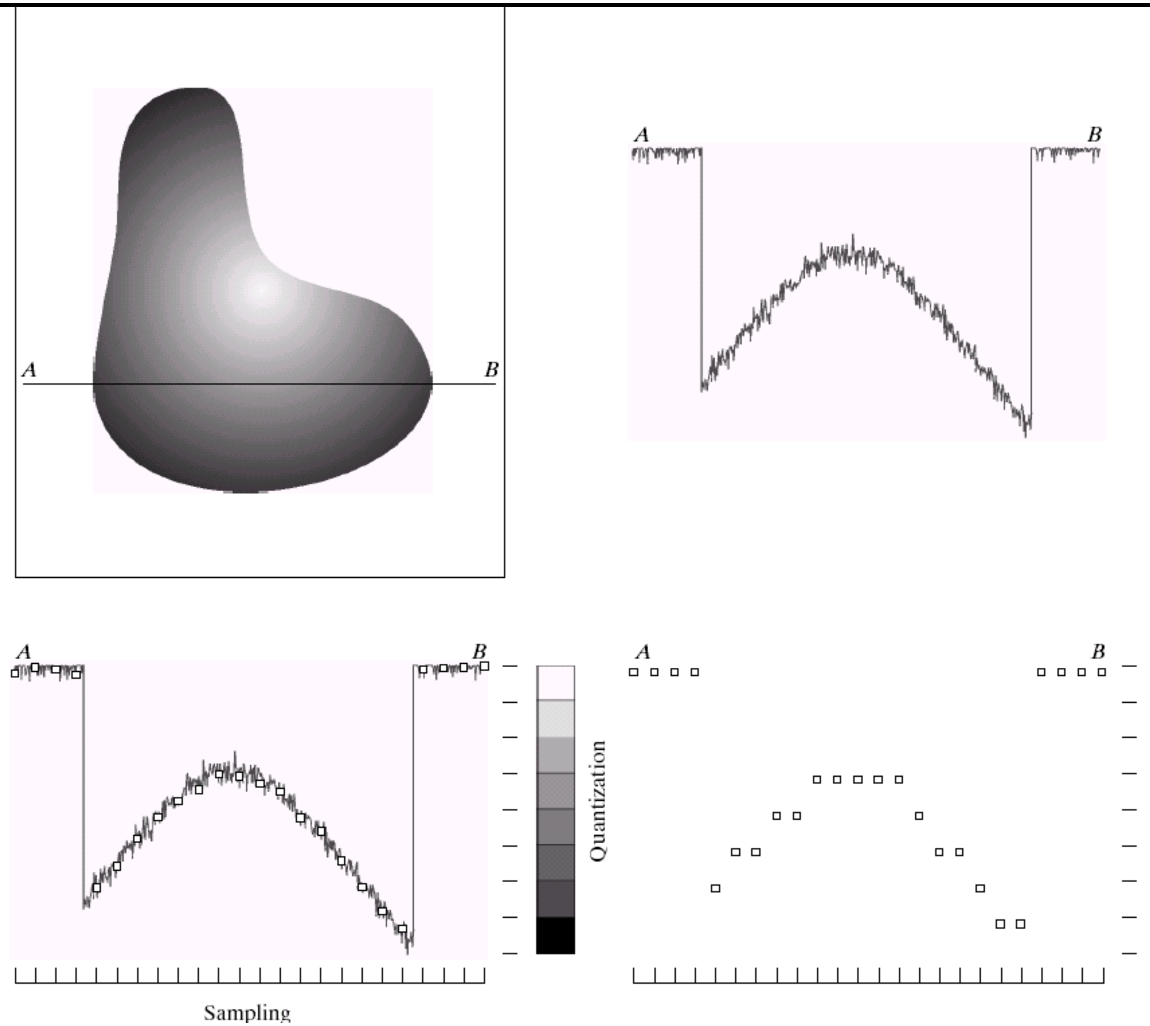
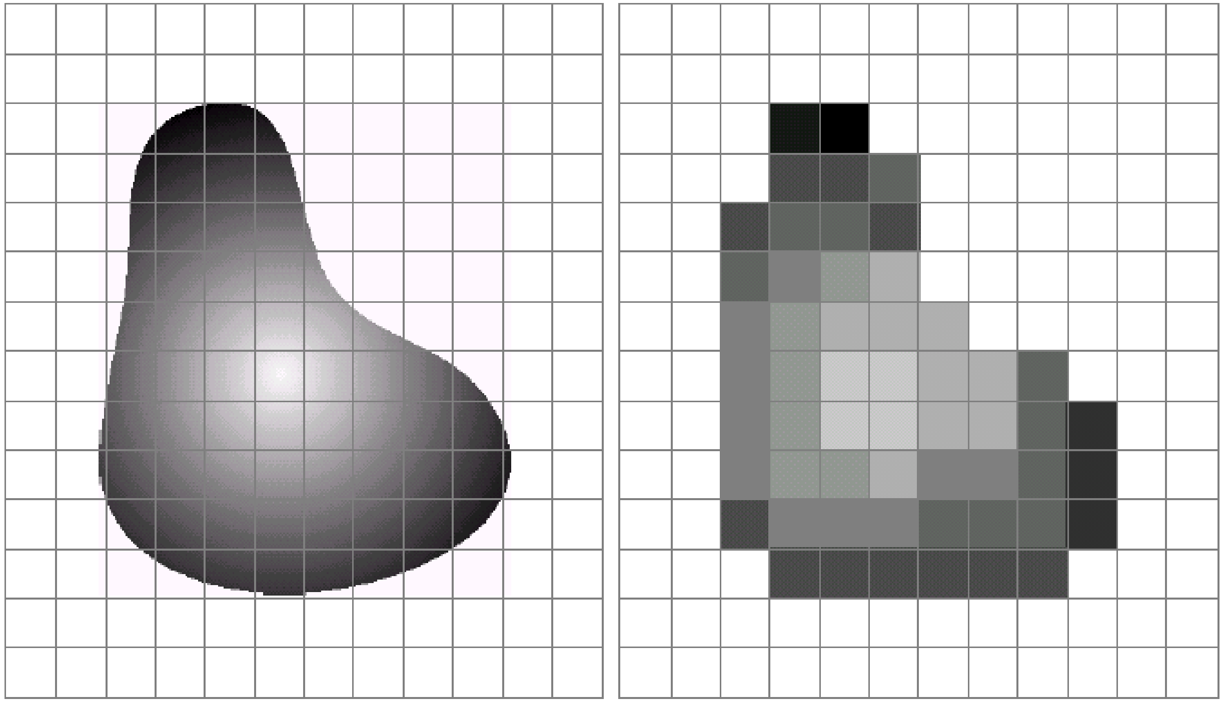
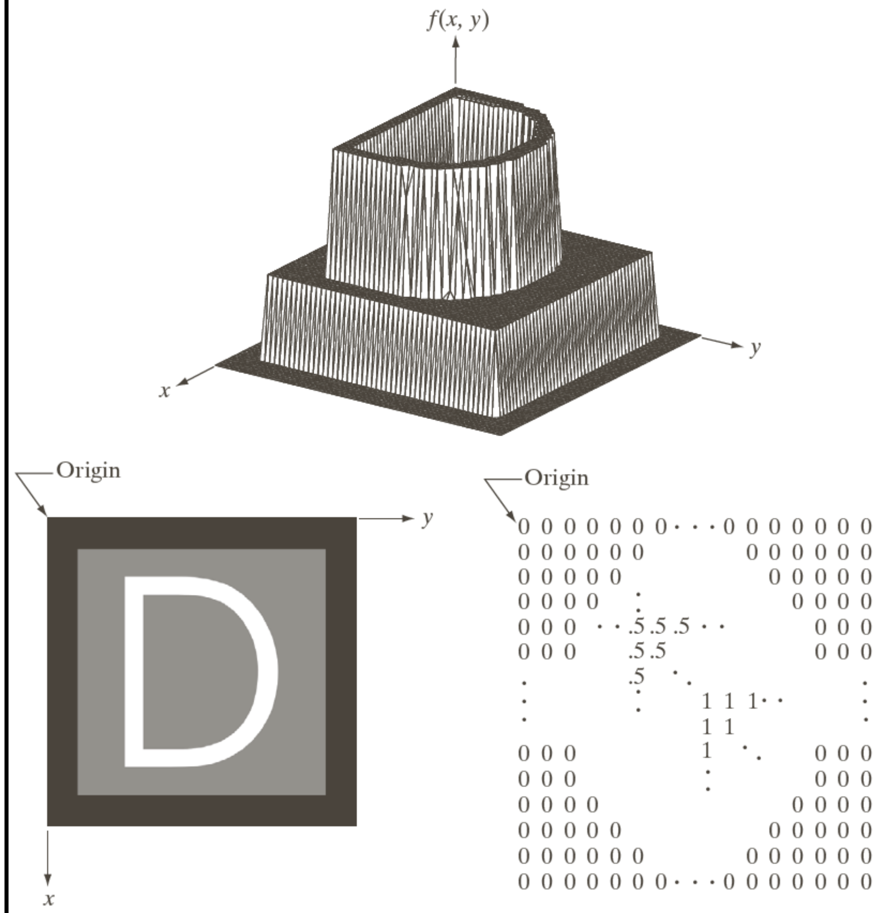


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.



- Left: continuous image projected onto a sensor array.
- Right: result of image sampling and quantization.

2.4.2 Representing Digital Images



Three basic ways to represent $f(x, y)$

- A 2.5D plot of the function. Look like a topography (useful for some specific algorithms).
- An image appear on a monitor (for visual display).
- Print the numerical values in an array (for implementation).

2.4 Image Sampling and Quantization

- The digitization process requires to determine the $M, N,$ and L.
  - M and N: **image size**
  - L: **gray-level resolution** (**radiometric resolution**)

    $L = 2^k$, L = gray-level
- **Dynamic range**: the range of values spanned by the **gray scale**, $[L_{min}, L_{max}]$.
  - **High dynamic range = high contrast image**
- The number of bits required to store the image

  b (bit)= $M \times N \times k$ or

  b (bit)= $N^2 \times k$

2.4 Image Sampling and Quantization

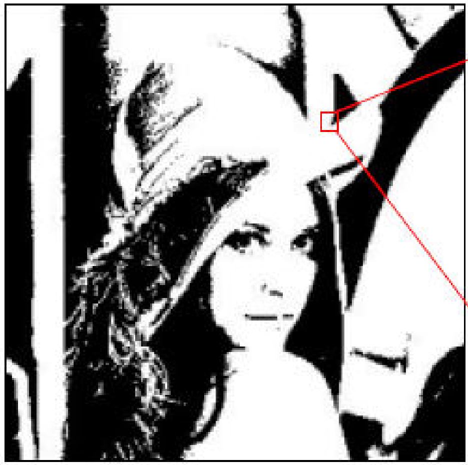
TABLE 2.1
Number of storage bits for various values of N and k.

N/k	1 (L = 2)	2 (L = 4)	3 (L = 8)	4 (L = 16)	5 (L = 32)	6 (L = 64)	7 (L = 128)	8 (L = 256)
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	393,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608
2048	4,194,304	8,388,608	12,582,912	16,777,216	20,971,520	25,165,824	29,369,128	33,554,432
4096	16,777,216	33,554,432	50,331,648	67,108,864	83,886,080	100,663,296	117,440,512	134,217,728
8192	67,108,864	134,217,728	201,326,592	268,435,456	335,544,320	402,653,184	469,762,048	536,870,912

N=256, k=8: 65536 bytes
N=2048, k=8: 12 M bytes
N=8192, k=8: 192 M bytes

Image representation

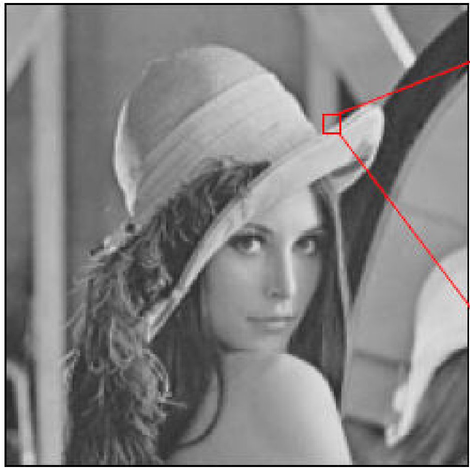
- Binary image: $f(x, y) = a \text{ or } b$
  - Only two values for a pixel: represented by 1 bit (0 or 1)
  - Size of a 256×256 image file: 65,536 bits = 8,192 bytes



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

Image representation

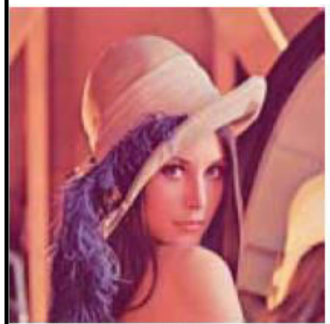
- **Gray-level image**: $\mathbf{0 \leq f(x,y) \leq 255}$
  - A pixel value is represented by **1 byte** (2^8); **0~255: from black to white** (256 gray levels)
  - Size of a 256×256 image file: 65,536 bytes



147	135	116	104	72	54
141	129	113	90	66	56
138	126	104	82	74	121
137	119	98	99	166	210
131	118	150	197	219	206
157	192	214	210	198	186

Image representation

- **Color models** of color images
  - **RGB** model: Red, Green, and Blue primaries
  - Number of representable colors: $(2^8)^3 = 2^{24} = 16{,}777{,}216$ (**true color**)
  - E.g., (0, 255, 0): green, (255, 0, 0): red,
    (0, 120, 0): light green, (100, 100, 0): yellow


Original image

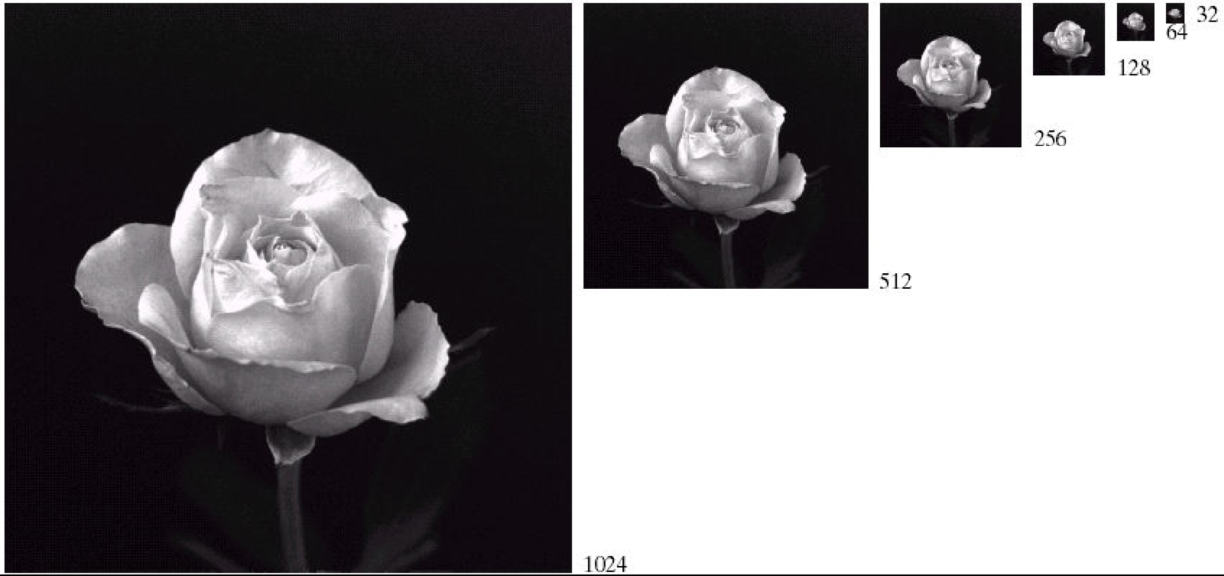

Red component


Green component


Blue component

2.4.3 Spatial and Intensity Resolution

- Spatial resolution is the smallest discernible (可識別的) detail in an image. Quantitatively, spatial resolution can be stated as dots (pixels) per unit distance.
- Spatial resolution is highly related to image size, but they have different meaning.



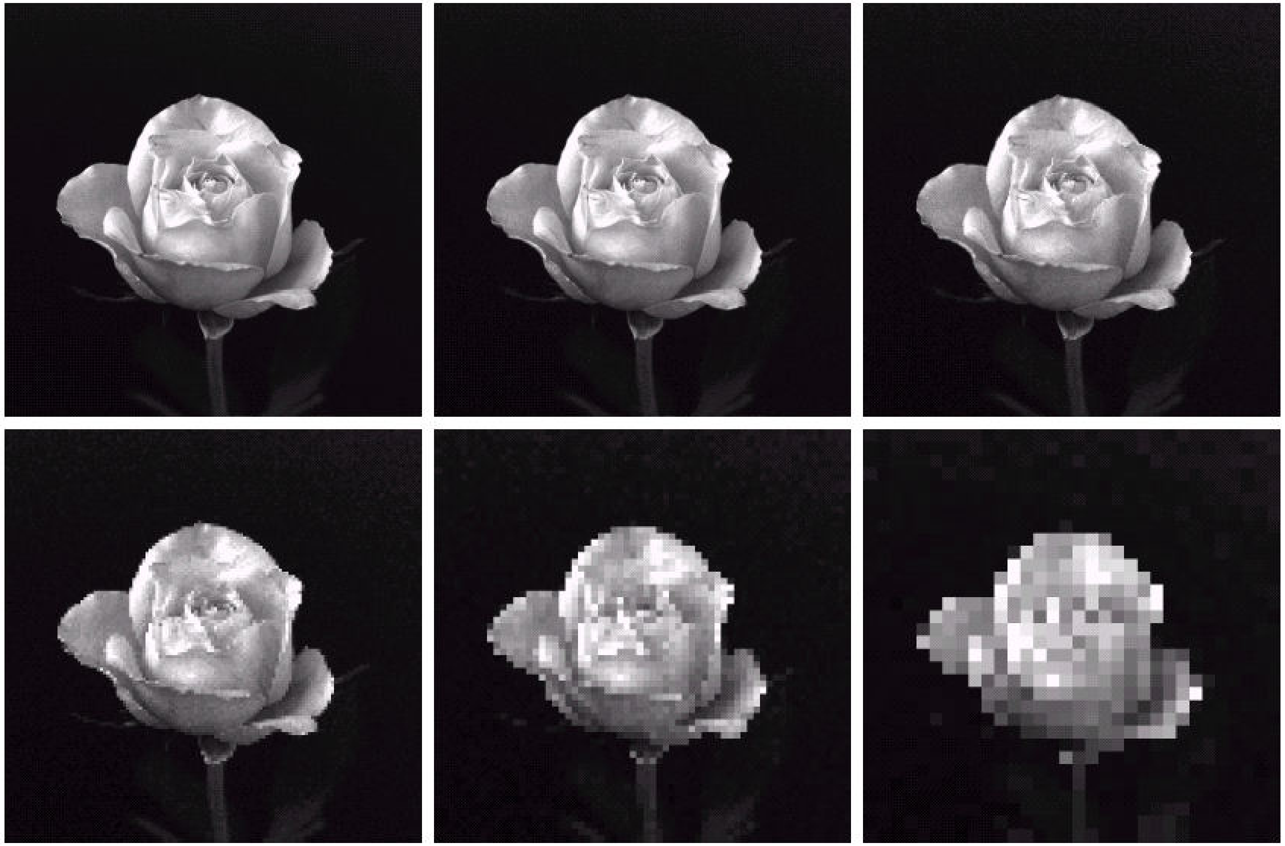
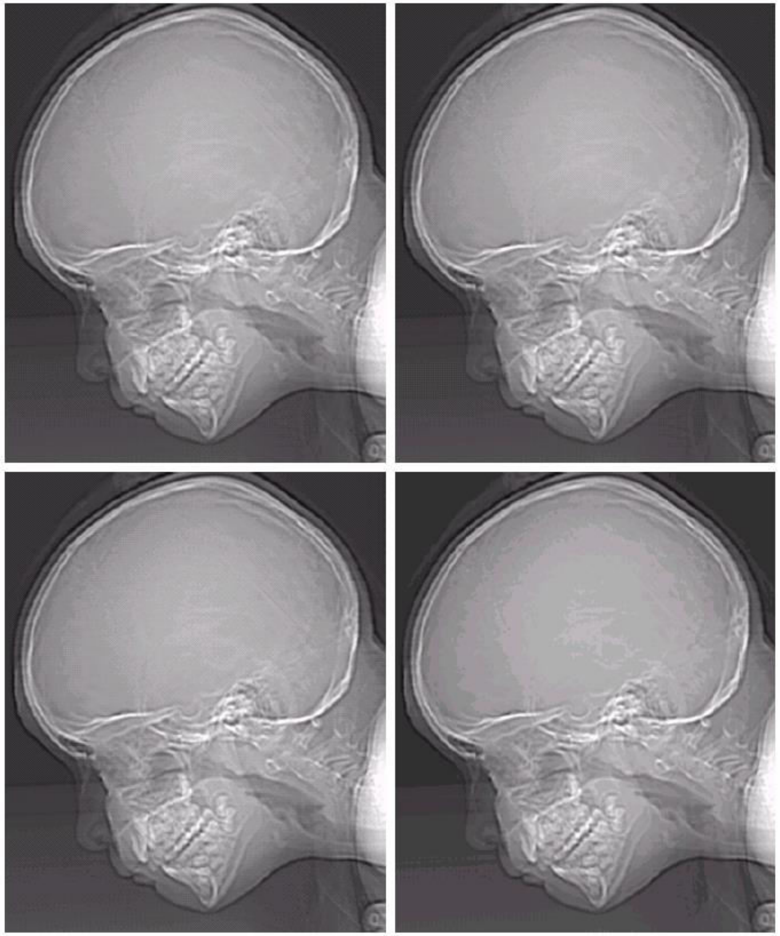


FIGURE 2.20 (a) 1024×1024, 8-bit image. (b) 512×512 image resampled into 1024×1024 pixels by row and column duplication. (c) through (f) 256×256, 128×128, 64×64, and 32×32 images resampled into 1024×1024 pixels.

The image sizes are the same, but spatial resolutions are different.

- **Gray-level resolution**: similarly refers to the smallest discernible change in intensity level.



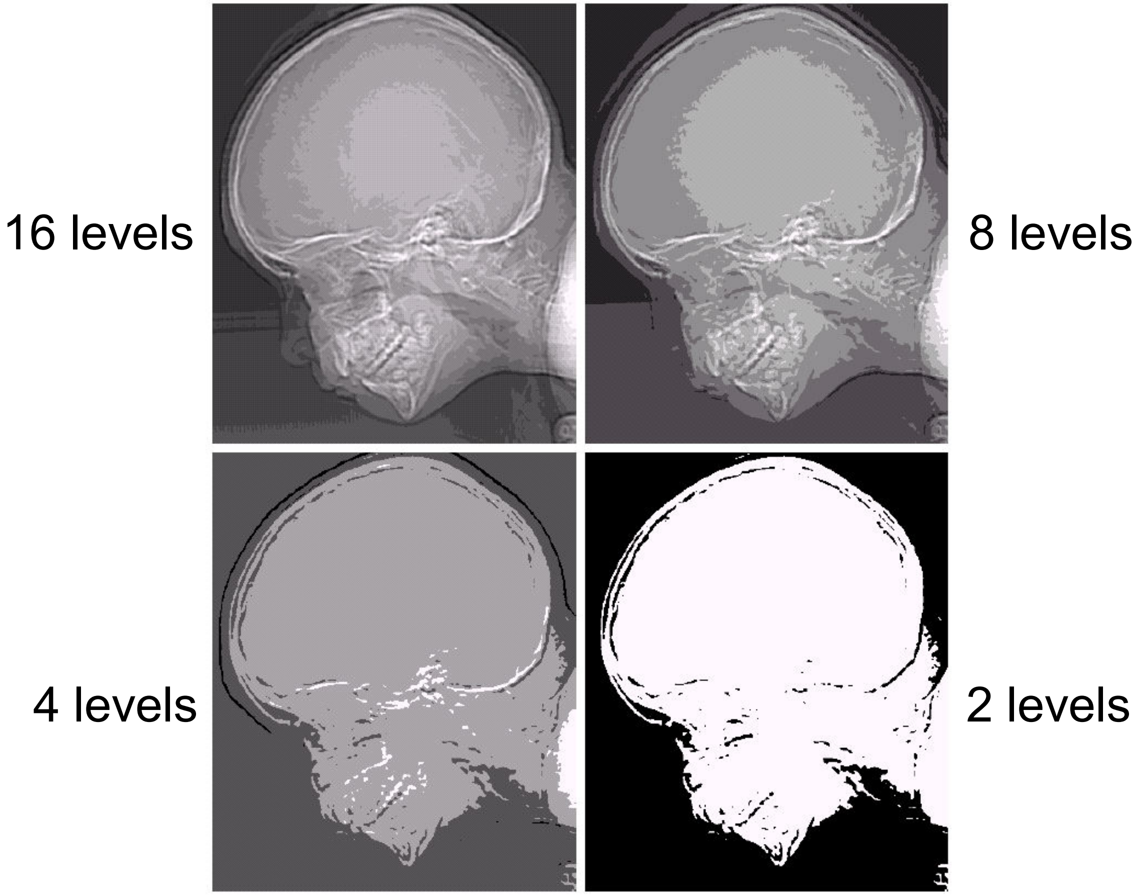
256 levels

128 levels

64 levels

32 levels

- **Gray-level resolution**: similarly refers to the smallest discernible change in intensity level.



Empirical Study of Resolutions

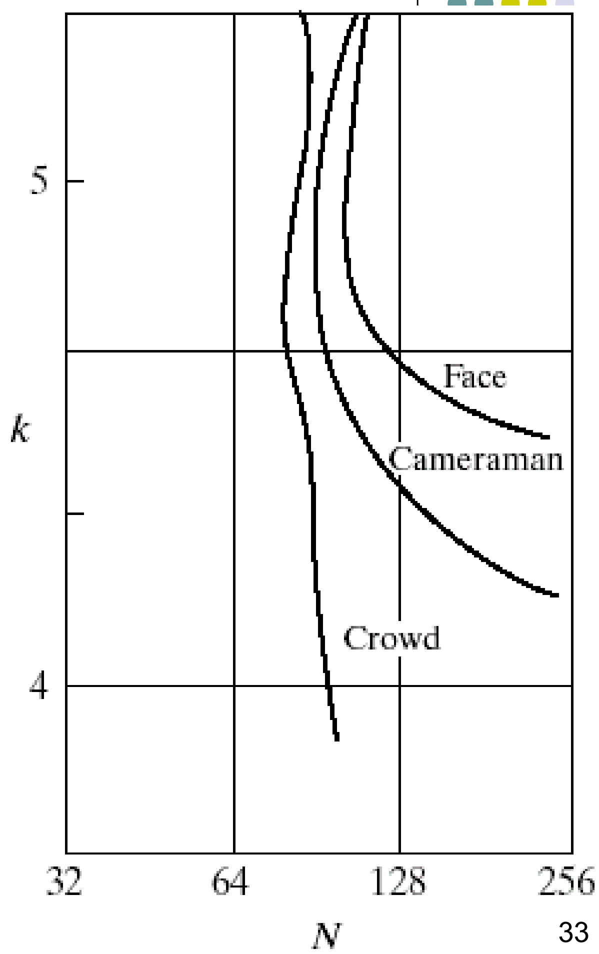
- Goal: **How k and N affect the image quality**
- 2^k-level digital image of size $N \times N$

Details (frequency)

Empirical Study of Resolutions

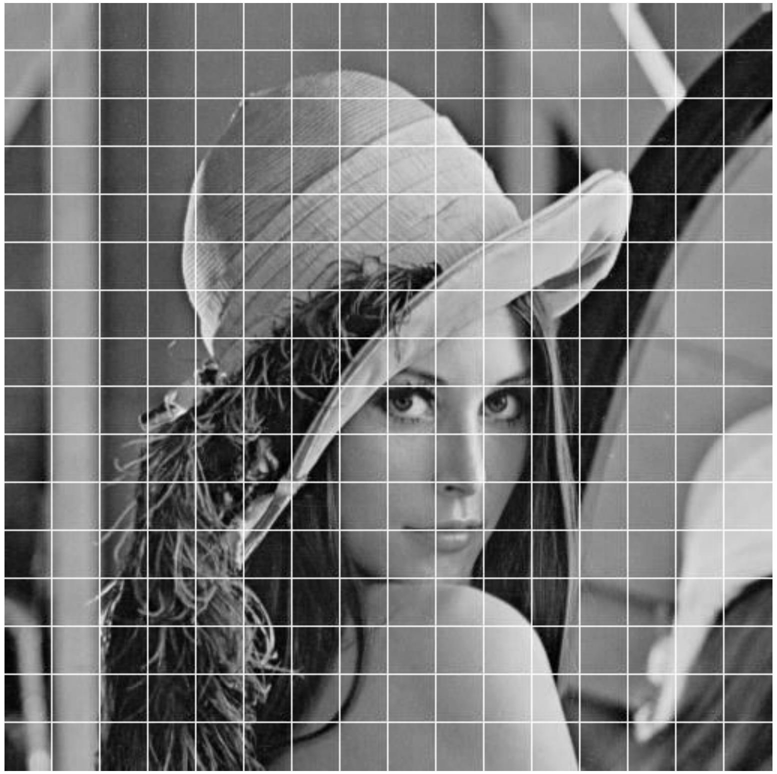
Iso-preference Curves

- Curves tends to shift right and upward. It simply means larger values for N and k implies better picture quality.
- Curve tends to become more vertical as the detail in the image increases.
- This result suggests that for images with a large amount of detail only a few intensity levels may be needed, and vice versa.



33

2.4.5 Zooming and Shrinking Digital Images

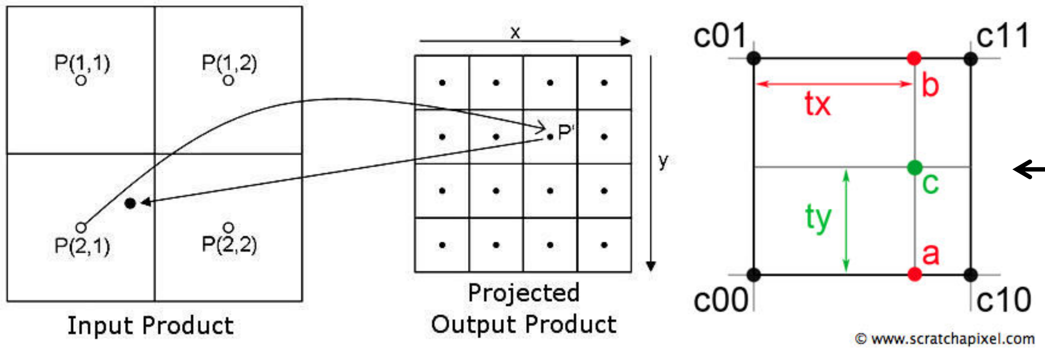


Idea: adjust the **grid size** over the original image

2.4.5 Zooming and Shrinking Digital Images

Zoming:

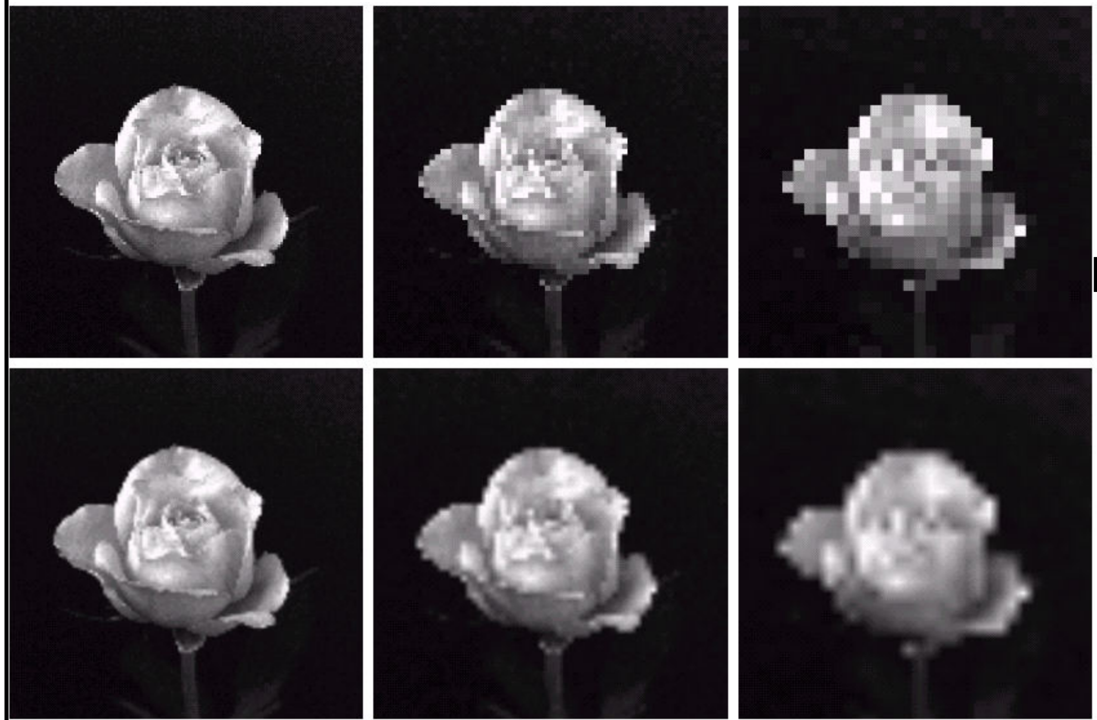
- Create several new pixel locations.
- Assign a gray-level to each of those new locations
  - Nearest neighbor interpolation
    - Pixel replication: a chessboard effect
  - Bilinear interpolation: using four nearest neighbors



  - Higher-order non-linear interpolation: using more neighbors for interpolation

35

Zooming Example



Nearest neighbor interpolation

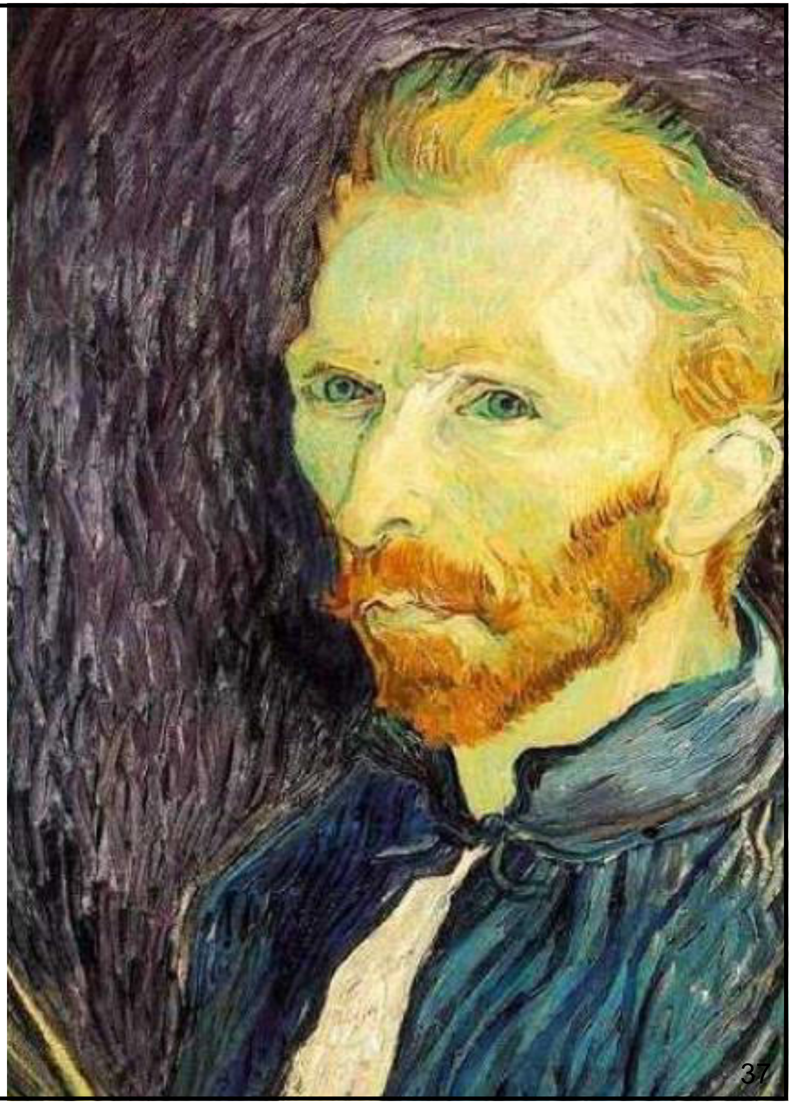
Bilinear interpolation

a b c
d e f

FIGURE 2.25 Top row: images zoomed from 128 × 128, 64 × 64, and 32 × 32 pixels to 1024 × 1024 pixels, using nearest neighbor gray-level interpolation. Bottom row: same sequence, but using bilinear interpolation.

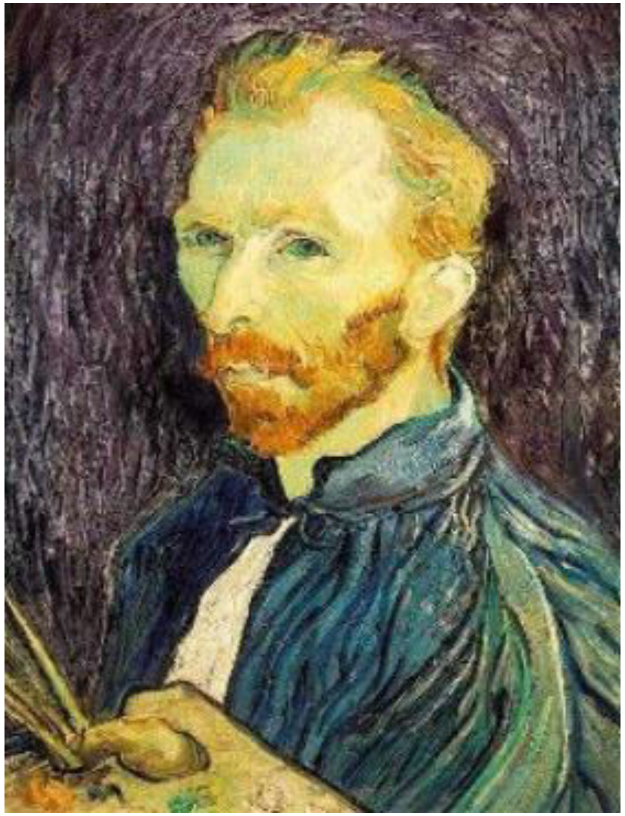
Shrinking Digital Images

- An image is too big to fit on the screen. How to reduce it? How to generate a half-sized version?
- Shrinking:
  - Direct shrinking (remove some rows and columns) causes **aliasing**
  - **Blur the image before shrinking it, which can reduce aliasing**



37

Image Shrinking
-- Naïve method

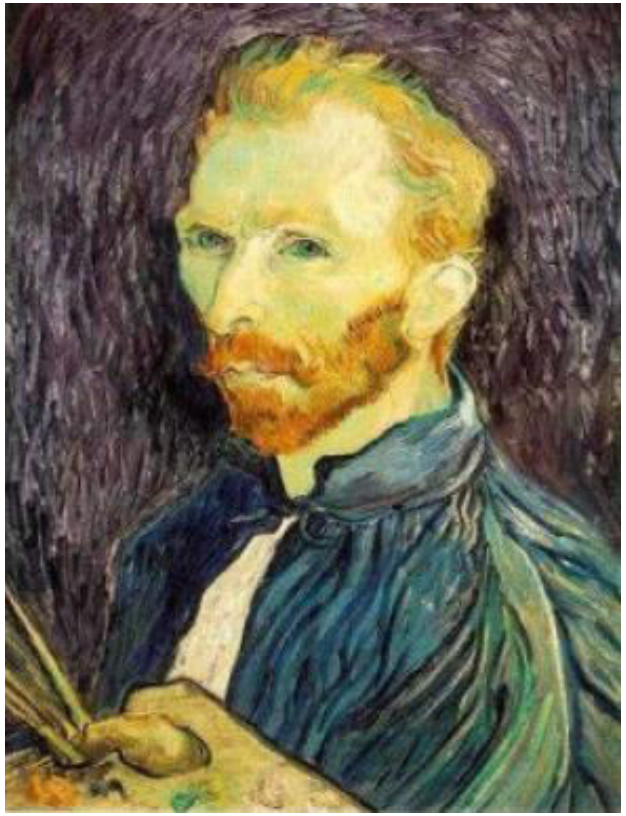


1/4

1/8

Throw away every some rows and columns to create a half-sized image

Image Shrinking
-- The common method



Gaussian 1/2

G 1/4

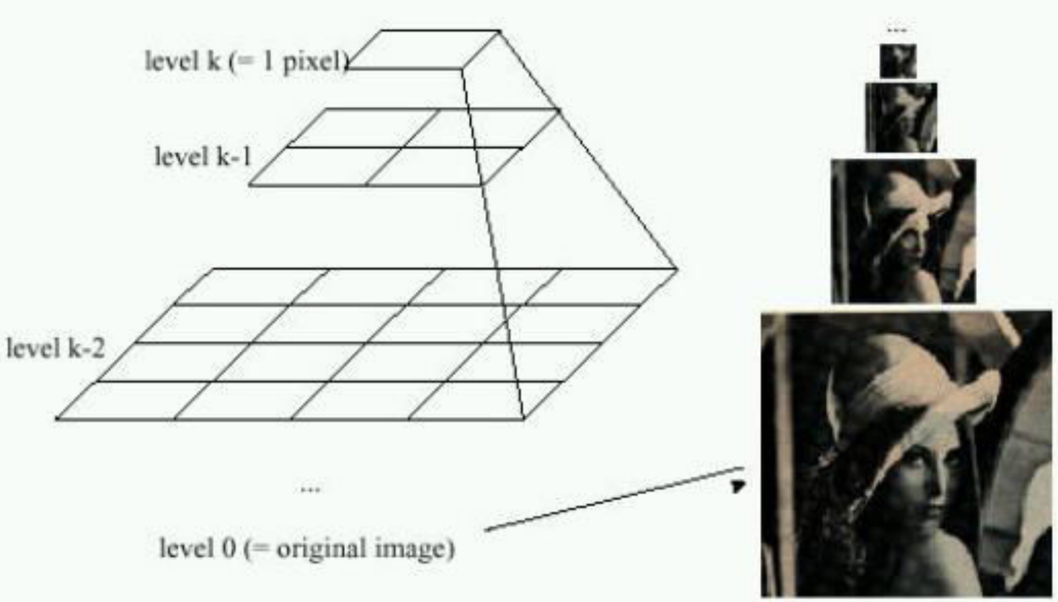
G 1/8

- Solution: subsampling with Gaussian pre-filtering
- Filter the image using Gaussian filter, then subsample

39

Hierarchical Structure

Idea: Represent NxN image as a "pyramid" of 1×1, 2×2, 4×4, ..., $2^k \times 2^k$ images (assuming $N=2^k$)



- Known as a **Gaussian Pyramid** [Burt and Adelson, 1983]
  - In computer graphics, a *mip map* [Williams, 1983]
  - A precursor to *wavelet transform*

Newest Method – Image Retargeting

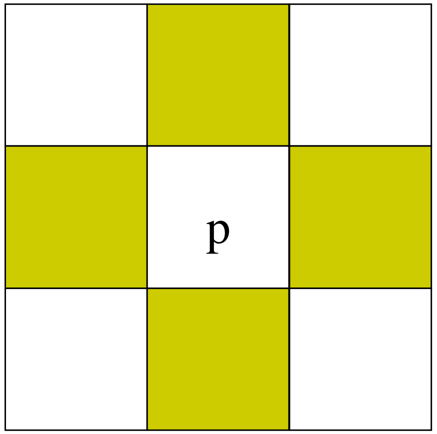
41

2.5 Basic Relations between Pixels

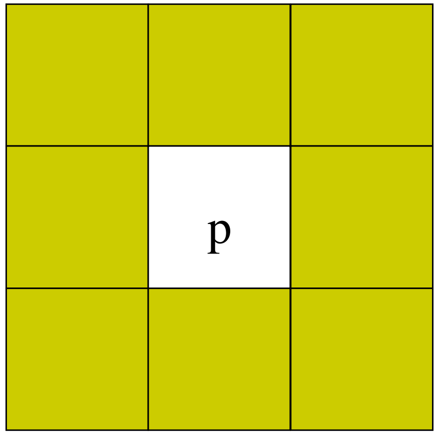
- **Neighbors** of a pixel (x, y)
  - Horizontal neighbors
    - $(\boldsymbol{x+1}, y)$, $(\boldsymbol{x-1}, y)$
  - Vertical neighbors
    - $(x, \boldsymbol{y+1})$, $(x, \boldsymbol{y-1})$
  - Four diagonal neighbors: $N_D(p)$
    - $(x+1, y+1)$, $(x+1, y-1)$, $(x-1, y+1)$, $(x-1, y-1)$
  - **4-neighbors of *p***: $N_4(p)$ (including horizontal and vertical neighbors).
  - **8-neighbors of *p***: $N_8(p)$.
    - $N_8(p) = N_4(p) \cup N_D(p)$
- **Adjacency**, **Connectivity**, **Regions**, **Boundary**

2.5 Basic Relations between Pixels

- Adjacency



4-adjacency: $N_4(p)$



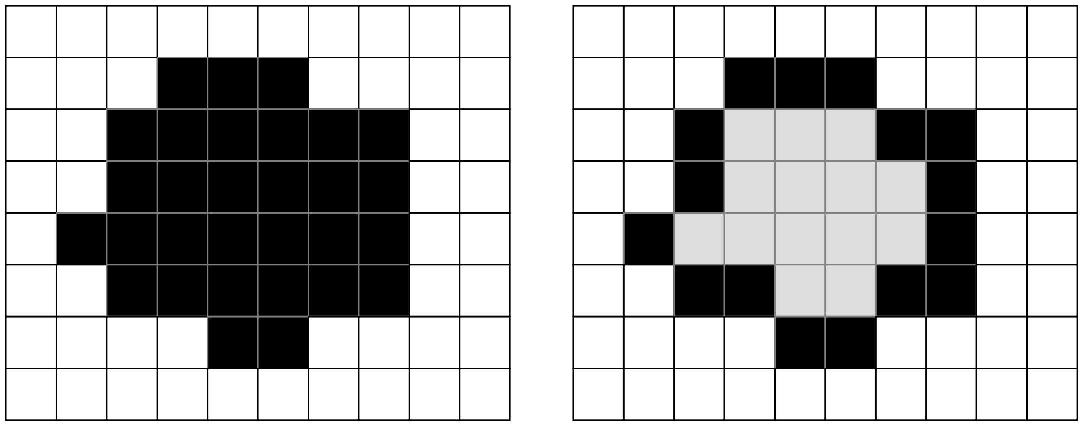
8-adjacency: $N_4(p) \cup N_D(p)$

Connectivity

- **Path**:
  - $(x_0, y_0), (x_1, y_1), \cdots, (x_n, y_n)$ where (x_i, y_i) and (x_{i+1}, y_{i+1}) are adjacent.
  - Closed path: $(x_n, y_n) = (x_0, y_0)$
- **Connectivity**:
  - Two pixels are said connected if they have **the same value** and **there is a path** between them.
  - If a **S** is a set of pixels,
    - For any pixel *p* in **S**, the set of pixels that are connected to it is called a **connected component** of **S**.
    - If **S** has only one connected component, **S** is called a connected set.

Regions

- R is a region if R is a connected set.
- The pixel in the boundary (contour) has at least one 4-adjacent neighbor whose value is 0.



45

- Distance measures
  - Euclidean distance
  - City-block distance or D_4 distance.
    $D_4(p,q) = |x-s| + |y-t|$
  - $\boldsymbol{D_8}$ distance or chessboard distance.
    $D_8(p,q) = max(|x-s|, |y-t|)$

$$\begin{array}{ccccc} & & 2 & & \\ & 2 & 1 & 2 & \\ 2 & 1 & 0 & 1 & 2 \\ & 2 & 1 & 2 & \\ & & 2 & & \end{array} \quad D_4$$

$$\begin{array}{ccccc} 2 & 2 & 2 & 2 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{array} \quad D_8$$