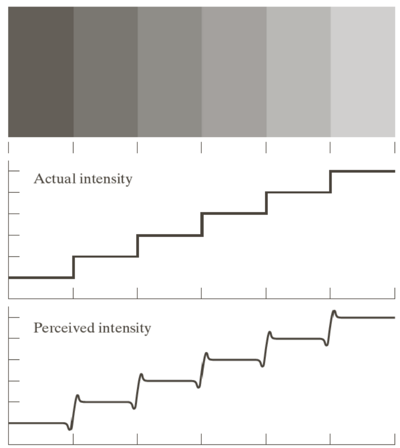
Digital Image Fundamentals

SUMMARY:

- Elements of Visual Perception
- Digitization
- Thresholding
- Connectivity and distance

Intensity perception



Mach bands: Perceived intensity is not equal to physical intensity.

Simultaneous contrast

Optical illusions

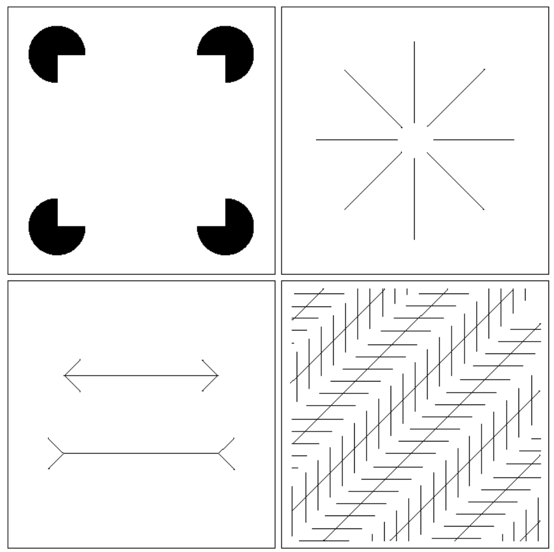
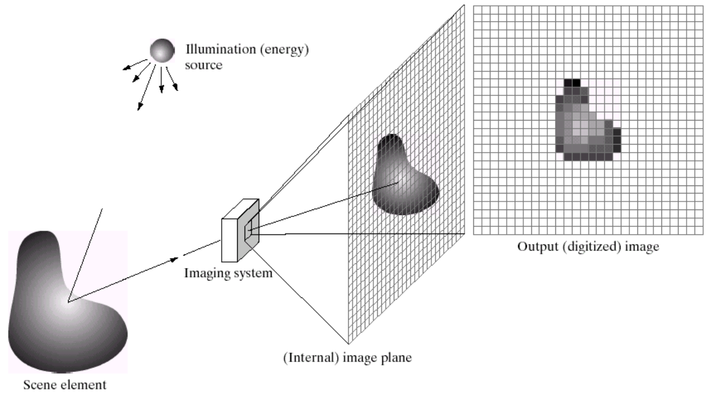
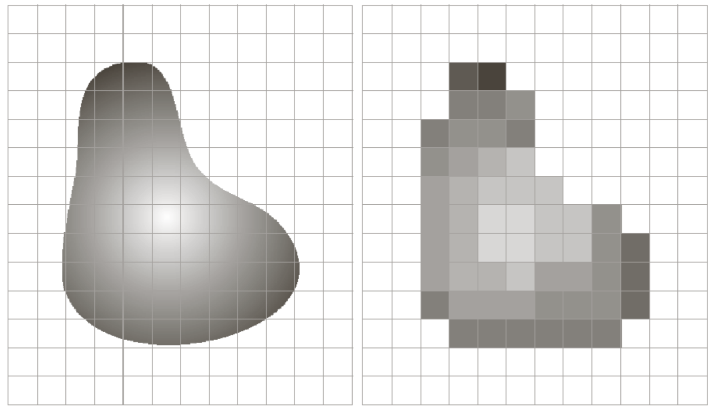


Image Acquisition



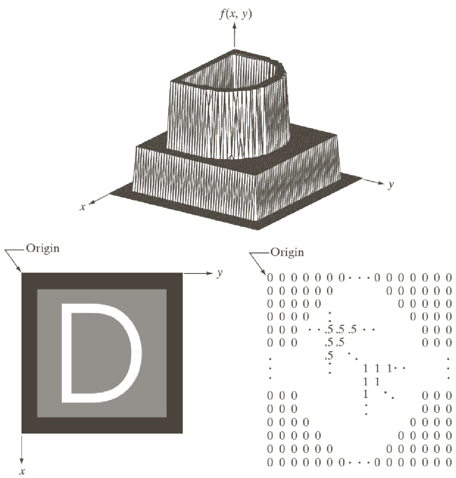
(a) Illumination source. (b) Element of scene. (c) Imaging system. (d) projection on image plane. (d) Digitized image.

Image sampling & quantization



(a) Continuous image projected on sensor array. (b) Result of sampling & quantization.

Image representations



(a) Surface. (b) Visual intensity array. (c) Numerical 2D array.

Image representations

Image with M rows and N columns:

$$\mathbf{f} = \begin{pmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & \vdots & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1,N-1) \end{pmatrix}$$

Element $f(i,j)$ is value of pixel on row i and column j.

Image representations

1	4	7
2	5	8
3	6	9

3×3 image

$$\mathbf{f} = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$$

matrix repr.

$$\vec{f} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{pmatrix}$$

vector repr.

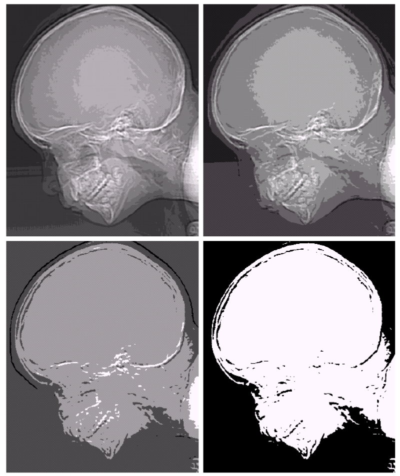
Resolution

- Spatial resolution: number of pixels per unit distance (dpi: dots per inch)
- Intensity resolution: number of bits used to quantize the intensity range

Reducing spatial resolution

(a) 1250 dpi. (b) 300 dpi. (c) 150 dpi. (d) 72 dpi.

Reducing intensity resolution



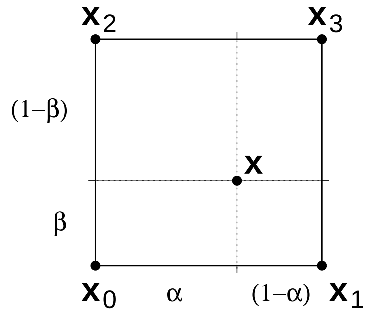
(a) 16 levels. (b) 8 levels. (c) 4 levels. (d) 2 levels.

Image interpolation

What is the image value S_X in an arbitrary point X?

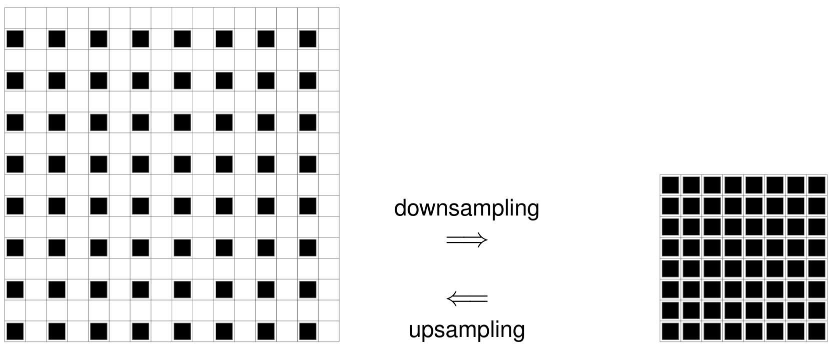
- nearest neighbour interpolation:
  value of the closest pixel
- bilinear interpolation:
  weighted average of 4 neighbouring pixels
- higher-order interpolation:
  more accurate, more computation because more neighbouring pixels are used

Bilinear interpolation



$$\begin{aligned} S_X &= (1-\alpha)(1-\beta)S_{X_0} + \alpha(1-\beta)S_{X_1} \\ &\quad +(1-\alpha)\beta S_{X_2} + \alpha\beta S_{X_3} \end{aligned}$$

Downsampling and upsampling



Downsampling and upsampling (zooming) by a factor of 2 in each dimension

Zooming by pixel replication

a	b	c	d
e	f	g	f
i	j	k	l
m	n	o	p

Input image

a	a	b	b	c	c	d	d
a	a	b	b	c	c	d	d
e	e	f	f	g	g	f	f
e	e	f	f	g	g	f	f
i	i	j	j	k	k	l	l
i	i	j	j	k	k	l	l
m	m	n	n	o	o	p	p
m	m	n	n	o	o	p	p

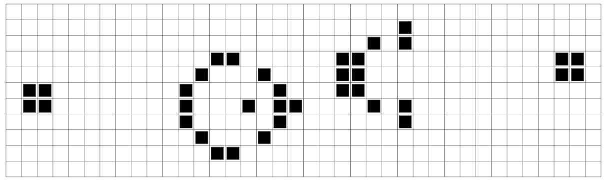
Zoomed image

Shrink and zoom

Image (3692×2812 pixels) shrunk to 72 dpi (row 1) or 150 dpi (row 2) and zoomed back to original size by nearest neighbour interpolation (column 1), bilinear interpolation (column 2), bicubic interpolation (column 3).

Binary images

A binary image consists of 1-pixels (pixels with value 1, foreground) and 0-pixels (pixels with value 0, background).



Thresholding

Let $I(r,c)$ be a grey value image. Let t be a fixed grey value called the threshold.
Define a binary image I_t, called the thresholded image, by:

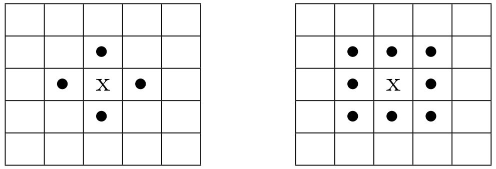
$$I_t(r,c) := 1 \quad if \quad I(r,c) \geq t$$
$$I_t(r,c) := 0 \quad if \quad I(r,c) < t$$

Thresholding

(a) (b)

(a): input image. (b): threshold image of (a).

Neighbours of a pixel



Left: the pixels • form the set $N_4(\mathrm{x})$ (the 4-neighbours of x).
Right: the pixels • form the set $N_8(\mathrm{x})$ (the 8-neighbours of x).
The diagonal neighbours form the set $N_D(\mathrm{x}) = N_8(\mathrm{x}) \setminus N_4(\mathrm{x})$.

Adjacency

Let V be a subset of the set of pixel values. For binary images, $V = \{0\}$ or $V = \{1\}$.
A V-pixel is a pixel whose value is in V.

- 4-adjacency: Two V-pixels p and q are 4-adjacent if q is a 4-neighbour of p ($q \in N_4(p)$).
- 8-adjacency: Two V-pixels p and q are 8-adjacent if q is an 8-neighbour of p ($q \in N_8(p)$).

Paths on a grid

A (digital) path from p to q of length ℓ is an $\ell + 1$-tuple $(p_0, p_1, \ldots, p_\ell)$, with $p_0 = p$, $p_\ell = q$, such that for all $i = 0, 1, \ldots, \ell - 1$ p_i and p_{i+1} are neighbours.

One may define 4-, 8-, or m-paths, depending on the chosen adjacency.

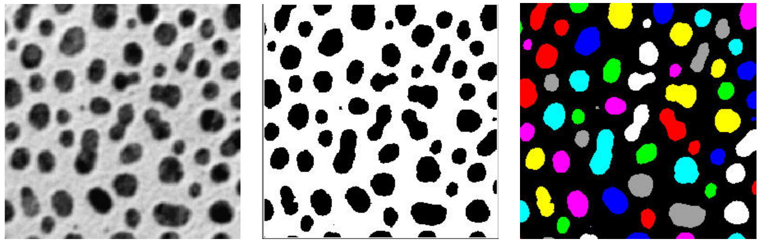
1	1	1	0	1
1	1	1	0	1
0	1	0	0	1
0	1	1	0	1
0	0	1	1	1

A 4-adjacent path of 1-pixels.

Connected components

- Let S be a subset of V-pixels of an image. Two V-pixels p and q are said to be connected in S if there is a path from p to q within S.
- For any p in S, the set of pixels connected to p in S is called a connected component of S.
- If S has only one connected component, S is called connected.
- A connected set of V-pixels is also called a region.

Component labeling



(a) (b) (c)

(a): input image. (b): threshold image of (a). (c): connected components of (b) labeled by colour.

Connected components

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

a	a	a	a	a
a	a	1	a	a
a	2	b	3	a
a	a	4	a	a
a	a	a	a	a

a	a	a	a	a
a	a	1	a	a
a	1	a	1	a
a	a	1	a	a
a	a	a	a	a

a	a	a	a	a
a	a	1	a	a
a	1	b	1	a
a	a	1	a	a
a	a	a	a	a

Upper left: binary image. Upper right: 4-adjacency for 0- and 1-pixels. Lower left: 8-adjacency for 0- and 1-pixels. Lower right: 8-adjacency for 1-pixels and 4-adjacency for 0-pixels.

Distance function

Let p, q and r be pixels on the grid $E = \mathbb{Z}^2$. The function $D : E \to \mathbb{R}$ is called a distance function if for all p, q and r:

1. $D(p,q) \geq 0$, with $D(p,q) = 0$ iff $p = q$
2. $D(p,q) = D(q,p)$ symmetry
3. $D(p,r) \leq D(p,q) + D(q,r)$ triangle inequality

Distance measures

- Euclidean distance:

$$D_e(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2}$$

- City-block distance:

$$D_4(p,q) = |p_1 - q_1| + |p_2 - q_2|$$

- Chessboard distance:

$$D_8(p,q) = \max(|p_1 - q_1|, |p_2 - q_2|)$$

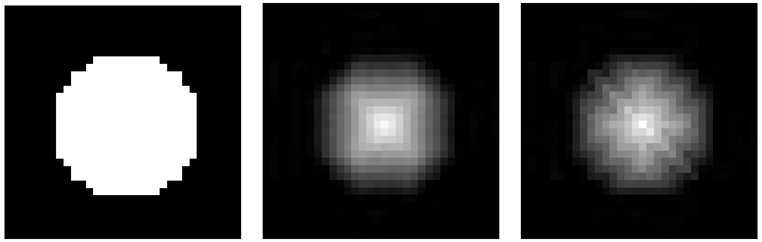
Distance transform: definition

Let $A \subseteq E$. The distance transform of A is the function $dt : E \to \mathbb{R}$ which associates to point x the D_α (D_e, D_4, D_8) distance of x to the complement of A:

$$dt_\alpha(x) = D_\alpha(x, A^c)$$

Notice that $dt_\alpha(x) = 0$ for points $x \in A^c$.

Distance transform



binary image $\quad dt_4 \quad dt_8$