* Digital image processing

UNIT - 4

Object Recognition

Knowledge Representation

- **What to Represent?**

Let us first consider what kinds of knowledge might need to be represented in AI systems:

- **Objects**
  - -- Facts about objects in our world domain. *e.g.* Guitars have strings, trumpets are brass instruments.
- **Events**
  - -- Actions that occur in our world. *e.g.* Steve Vai played the guitar in Frank Zappa's Band.

Thus in solving problems in AI we must represent knowledge and there are two entities to deal with:

- **Facts**
  - -- truths about the real world and what we represent. This can be regarded as the *knowledge level*
- **Representation of the facts**
  - which we manipulate. This can be regarded as the *symbol level* since we usually define the representation in terms of symbols that can be manipulated by programs.

- We can structure these entities at two levels
- **the knowledge level**
  - -- at which facts are described
- **the symbol level**
  - -- at which representations of objects are defined in terms of symbols that can be manipulated in programs.

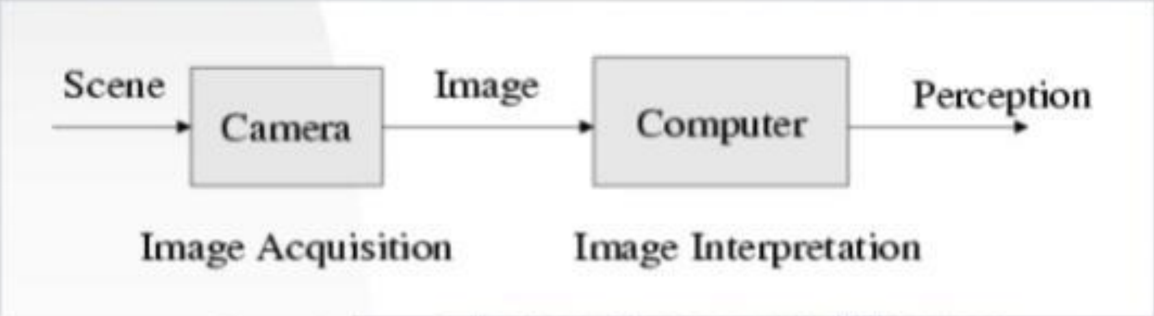
Figure: Simple Relational Knowledge

Musician	Style	Instrument	Age
Miles Davis	Jazz	Trumpet	deceased
John Zorn	Avant Garde	Saxophone	35
Frank Zappa	Rock	Guitar	deceased
John Mclaughlin	Jazz	Guitar	47

- We can structure these entities at two levels
- **the knowledge level**
  - -- at which facts are described
- **the symbol level**
  - -- at which representations of objects are defined in terms of symbols that can be manipulated in programs.

Computer Vision

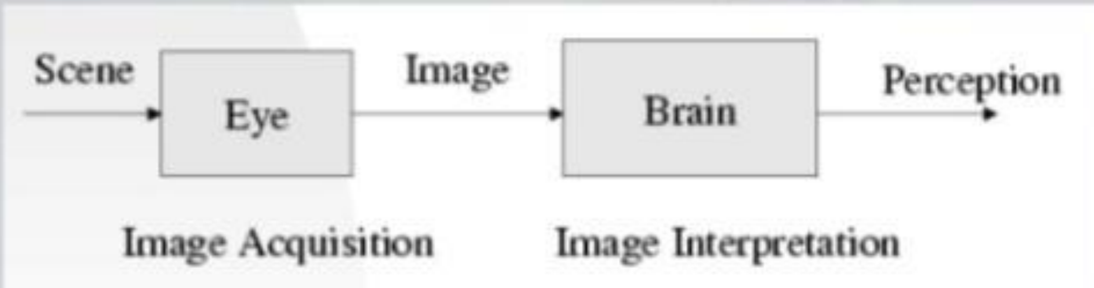
- Computer Vision is the study of analysis of pictures and videos in order to achieve results similar to those as by humans

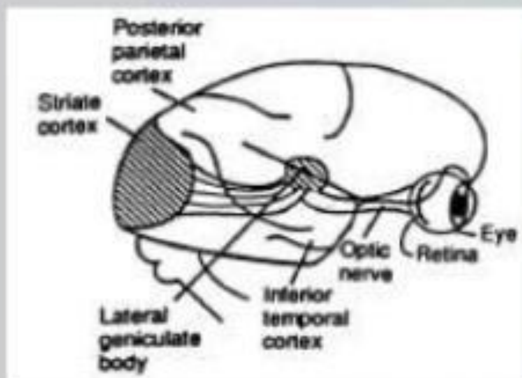


- Analogously, given a set of TV camera
  - What computer architectures, data structures & algorithms should use to create a machine that can "see" as we do?

Human Vision

- Vision is the process of discovering what is present in the world and where it is by looking





- Computational algo implemented in this massive network of neurons; they obtain their inputs from retina, & produce as output an "understanding" of the scene in view
  - But what does it mean to "understand" the scene? What algos & data representation are used by brain?

Human Vision VS Computer Vision

What we see

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

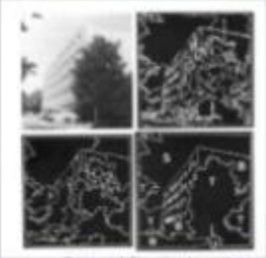
Outline

- Fingerprint Recognition
  - Definition
  - Fingerprint Matching Using Ridge-End and Burification
  - Fingerprint Image
    - Binarization
    - Thinning
  - Minutiae Extraction
- Car Number Plate Recognition
  - What is an ALPR System?
  - ALPR Procedure
    - Characters Recognition
    - Characters Segmentation
    - Normalization of Characters
- New Innovations in Object Recognition
- References

Brief History of Computer Vision

- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts
- 2030's: robot uprising?

Guzman '68



Ohta Kanade '78

Turk and Pentland '91

Statistical Pattern Recognition

- The Problem: Given a set of measurements **x** obtained through observation, Assign the pattern to one of C possible classes w_i, $i=1,2,\ldots C$

- A decision rule partitions the measurement space into C regions $\mathbf{W}_i$, $i=1,\ldots,C$
- If a pattern vector falls in the region $\mathbf{W}_i$, then it is assumed to belong to class w_i
- If it falls on the boundary between regions $\mathbf{W}_i$, we may reject the pattern or withhold a decision until further information is available

Statistical Approach

- Each pattern is represented in D features in d dimensional space as a point.
- Objective to establish decision boundaries in the feature space which separate pattern of different classes.
- Discriminate analysis based approach for classification
- Using mean squared error criteria
- Construct the decision boundaries of the specified form

Statistical Pattern Recognition

- Many sources of variability in speech signal
- Much more than known deterministic factors
- Powerful mathematical foundation
- More general way of handling discrimination

Statistical Pattern Recognition (cont..)

- If all of the class conditional densities is known then **Bayes decision rule** can be used to design a classifier.
- If the form of class conditional densities is known **(multivariate gaussian)** but parameter like an mean vectors and covariance matrix) not known then we have a **parametric decision problem.** Replace the unknown paramters with estimated value.
- If form of class conditional density not known that we are in **non parametric mode**. In such cases we used **Parzen window (estimate the density function)** or directly construct boundry by using **KNN rule**.
- Optimizing the classifier to maximize its performance on training data will **NOT** give such result on test data.

Statistical Pattern Recognition

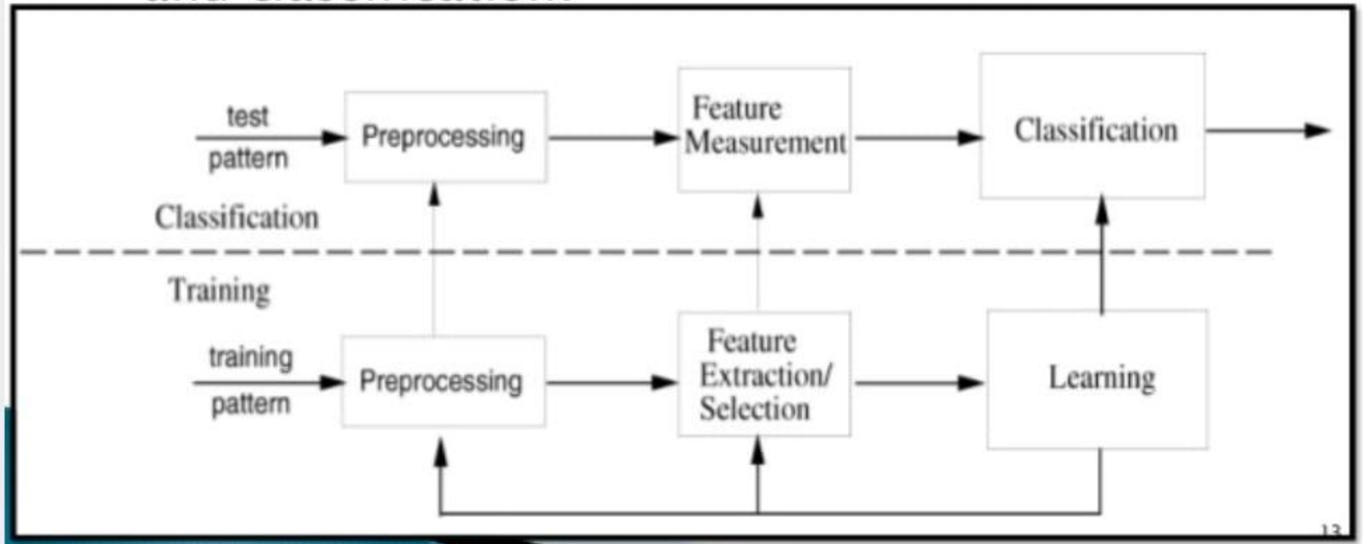
Decision Making Process

- Pattern assign to one of the C categories/Class W1,W2,...,Wc based on a vector of d features values x=(x1,x2,...,xd)
- Class conditional Probability = $P(x|wi)$
- Conditional Risk = $R(wi|x)=\Sigma L(wi,wj).P(wj|X)$
  where L(wi,wj) is loss in curred in deciding wi when true class is wj.
- Posterior Probability = $P(Wj|X)$
- For 0/1 loss function = $L(wi,wj)=\{0, i=j$
  $\{1, i \neq j$
- Assign input pattern x to class wi if

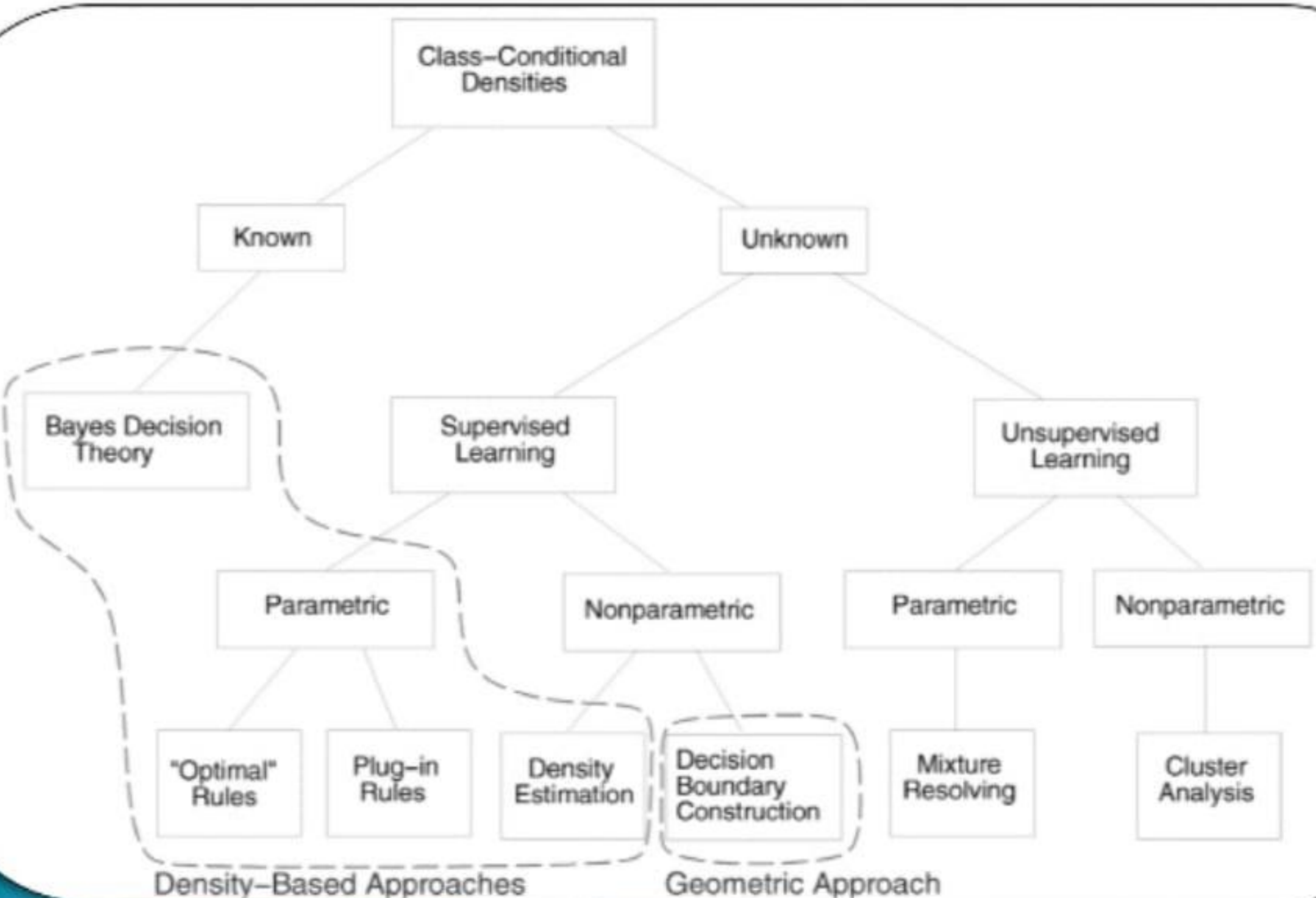
  $P(Wi|X) > P(Wj|X)$ for all $j \neq i$

Statistical Pattern Recognition

- Pattern is represented by set of d features/attributes viewed as D-dimensional feature space.
- System is operating in two modes i.e Training and classification.



Various Approaches in Statistical Pattern Recognition



Morphological operations

Morphology is a technique of image processing based on shapes.

The value of each pixel in the output image is based on a comparison of the corresponding pixel in the input image with its neighbors.

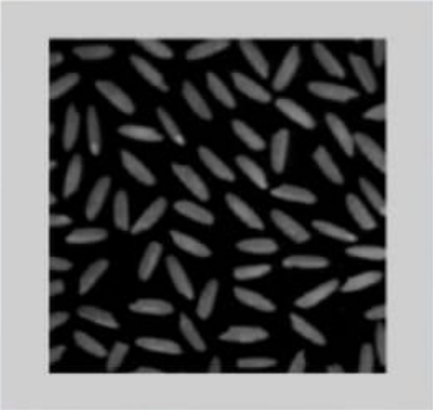
By choosing the size and shape of the neighborhood, a morphological operation can be applied that is sensitive to specific shapes in the input image.

Image Processing Toolbox morphological functions in Matlab can be used to perform common image processing tasks, such as contrast enhancement, noise removal, thinning, skeletonization, filling, and segmentation.

Morphological functions

Adjust the Image Contrast

I3 = imadjust(I2, stretchlim(I2), [0 1]);

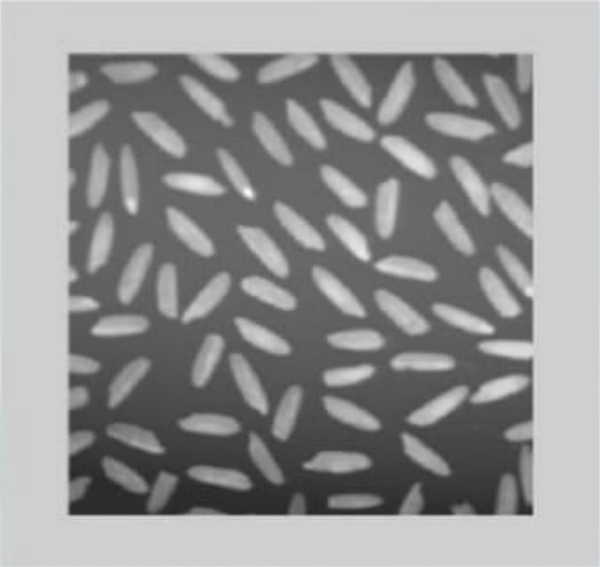


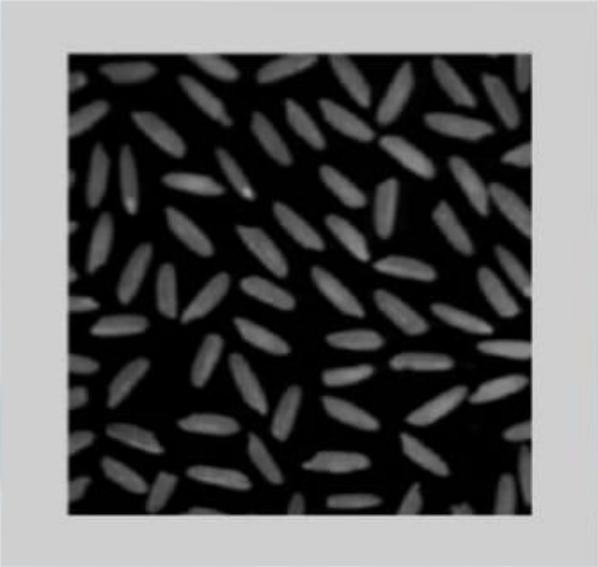
imadjust command to increase the contrast in the image. The imadjust function takes an input image and can also take two vectors: [low high] and [bottom top]. The output image is created by mapping the value low in the input image to the value bottom in the output image, mapping the value high in the input image to the value top in the output image, and linearly scaling the values in between.

Morphological functions

Now subtract the background image, background, from the original image, I, to create a more uniform background.

```
I2 = imsubtract(I,background);
```





Morphological functions

```
I = imread('rice.tif');
background = imopen(I,strel('disk',15));
imshow(background)
```

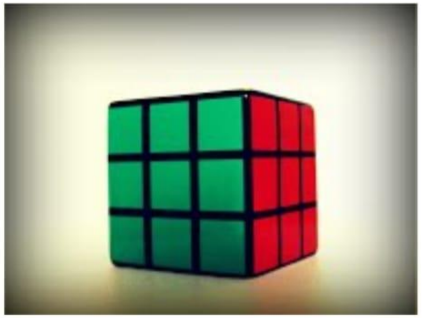
Morphological opening operation by calling imopen with the input image, I, and a disk-shaped structuring element with a radius of 15.

The structuring element was created by the strel function.

The morphological opening has the effect of removing objects that cannot completely contain a disk of radius 15.

The purpose of this chapter is to introduce the two basic concepts of mathematical morphology, namely *dilations* and *erosions*. Three possible uses are given. Firstly, dilation and erosion can be combined to produce two other morphological concepts *(openings* and *closings)* which have rich structural content. Secondly, the Hausdorff distance between objects has a simple morphological interpretation. Finally, the chance of detecting an object by regular sampling can be simply expressed in terms of dilations.

Morphology – the internal structure of words



Morphology is the study of the internal structure of words and forms a core part of linguistic study today.

- The term morphology is Greek and is a makeup of morph- meaning 'shape, form', and -ology which means 'the study of something'.
- Morphology as a sub-discipline of linguistics was named for the first time in 1859 by the German linguist **August Schleicher** who used the term for the study of the form of words.[1]

- An approach for processing digital image based on its **shape**
- A mathematical tool for investigating **geometric structure** in image
- The language of morphology is **set theory**

* What is the mathematical morphology ?

* Goal of morphological operations

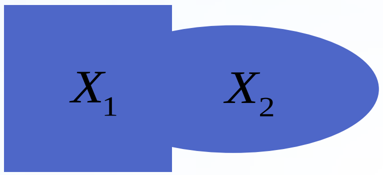
- ➢Simplify image data, preserve essential shape characteristics and eliminate noise
- ➢Permits the underlying shape to be identified and optimally reconstructed from their distorted, noisy forms

- Identification of objects, object features and assembly defects correlate directly with **shape**
- Shape is a prime carrier of information in machine vision

* Shape Processing and Analysis

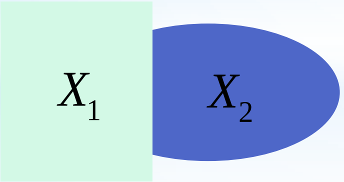
Shape Operators

- Shapes are usually combined by means of :
  - **Set Union** (overlapping objects):



$$X_1 \cup X_2$$

  - **Set Intersection** (occluded objects):



$$X_2 \setminus X_1 = X_1^c \cap X_2$$

- The primary morphological operations are **dilation** and **erosion**
- More complicated morphological operators can be designed by means of combining erosions and dilations

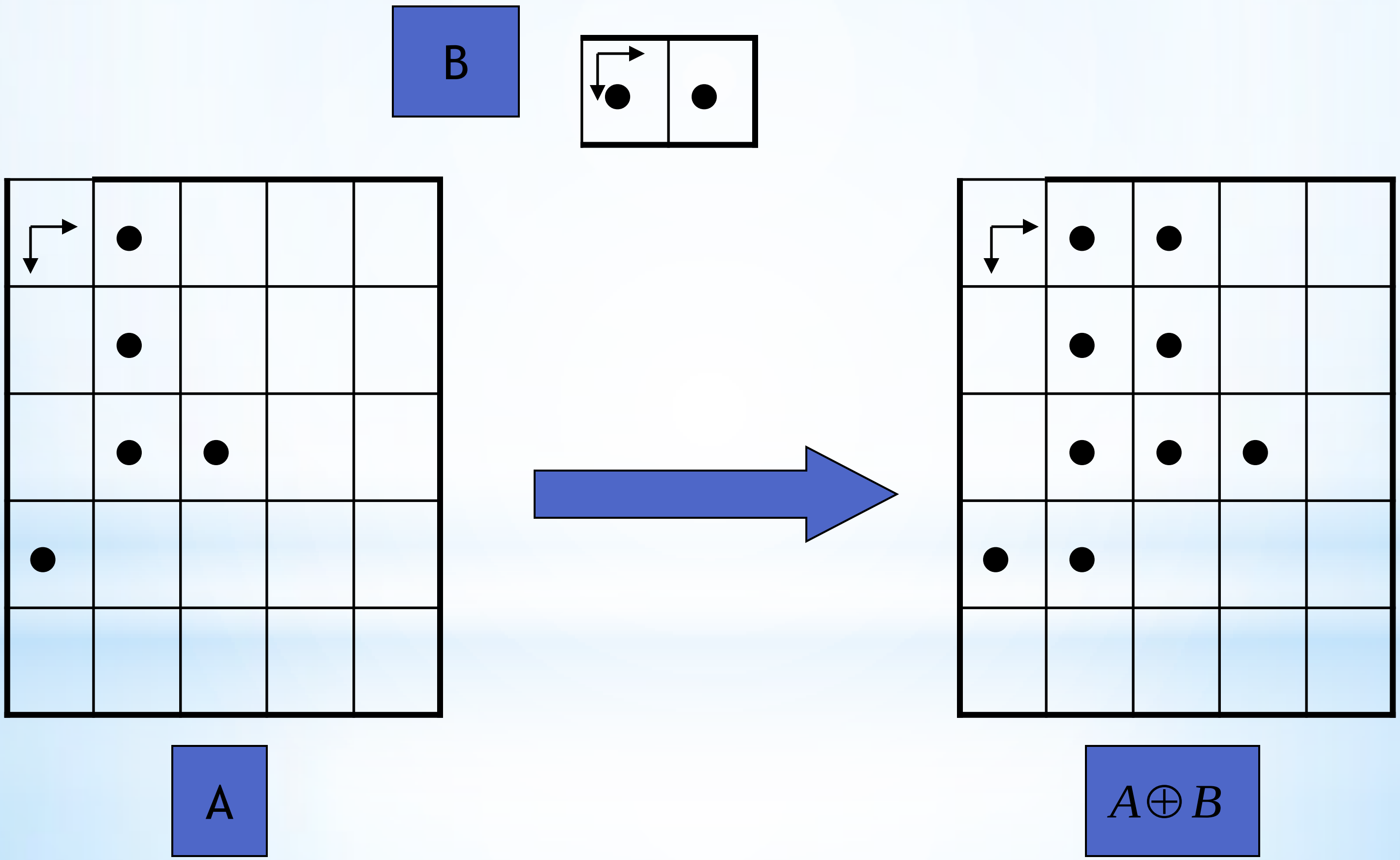
* Morphological Operations

Dilation

- Dilation is the operation that combines two sets using vector addition of set elements.
- Let A and B are subsets in 2-D space. A: image undergoing analysis, B: Structuring element, $\oplus$ denotes dilation

$$A \oplus B = \{c \in Z^2 \mid c = a + b \quad for \quad some \quad a \in A, b \in B\}$$

*Dilation



* Dilation

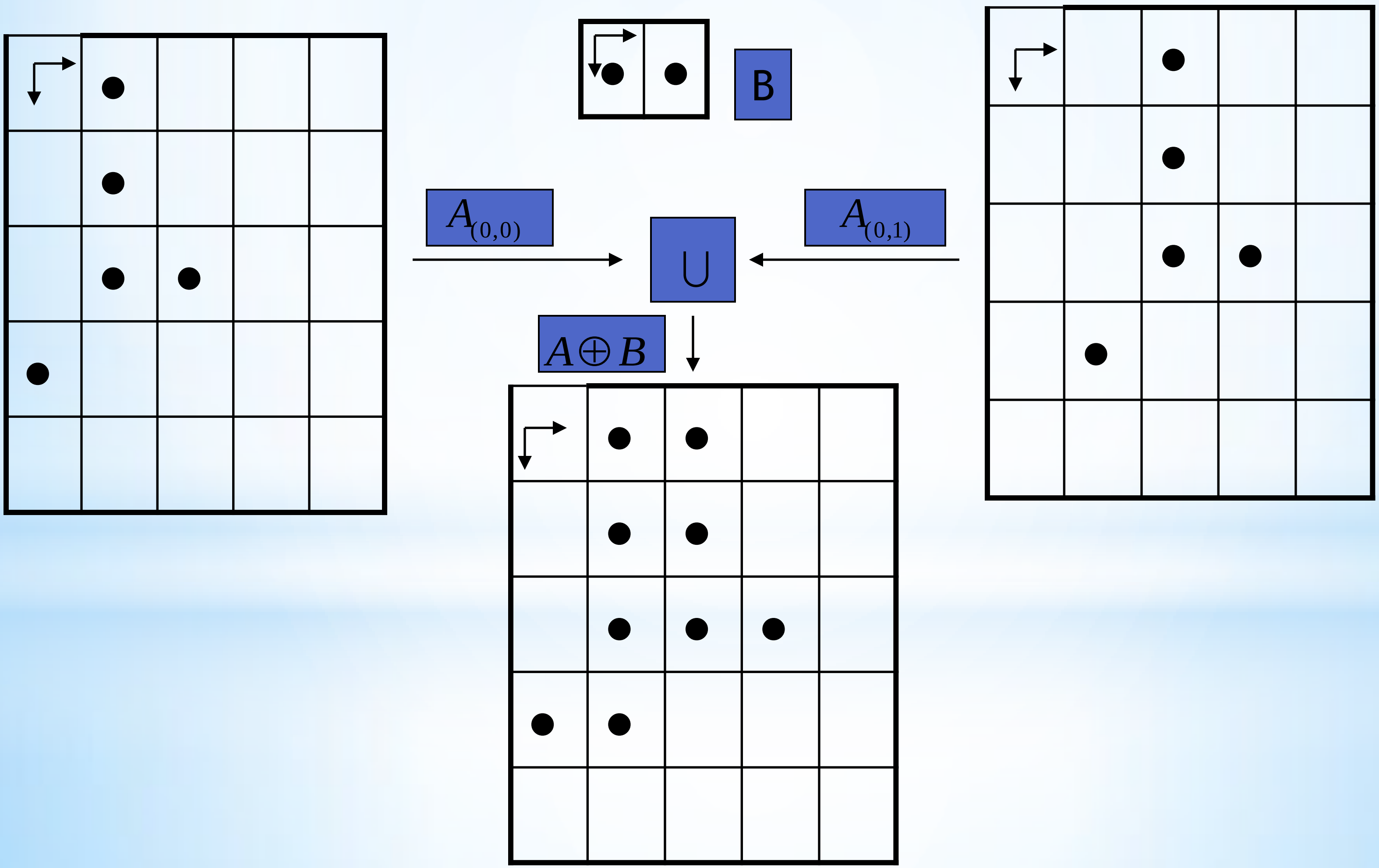
- Let A be a Subset of Z^2 and $x \in Z^2$. The translation of A by x is defined as

$$(A)_x = \{c \in Z^2 \mid c = a + x, \quad for \quad some \quad a \in A\}$$

- The dilation of A by B can be computed as the u elem

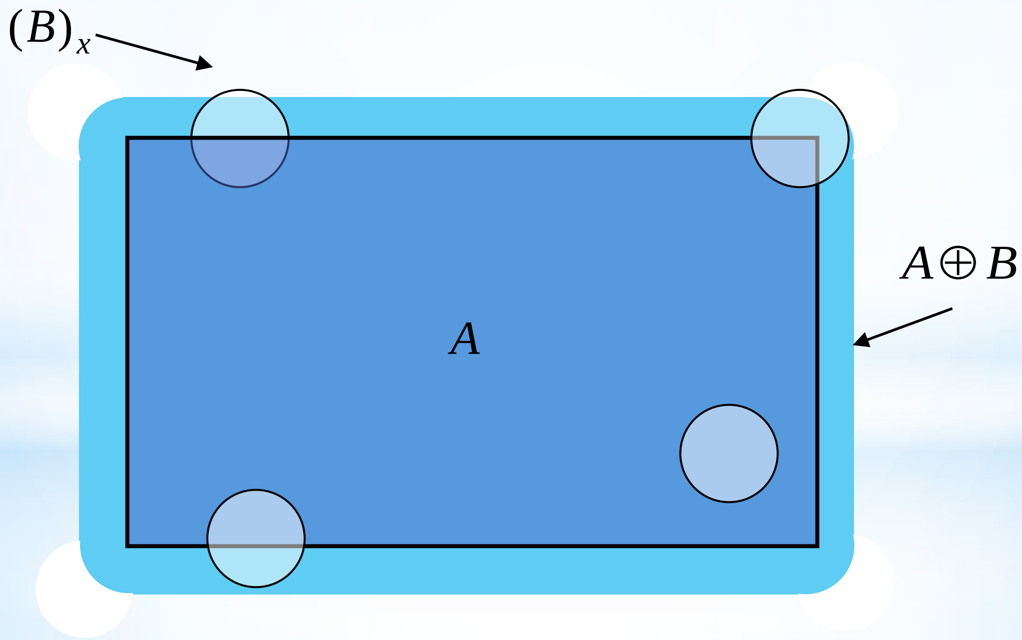
$$A \oplus B = \bigcup_{b \in B} (A)_b = \bigcup_{a \in A} (B)_a$$

*Dilation



*Dilation

$$A \oplus B = \bigcup_{b \in B} (A)_b = \bigcup_{a \in A} (B)_a$$



Properties of Dilation

- Commutative
- Associative
- Extensivity
- Dilation is increasing

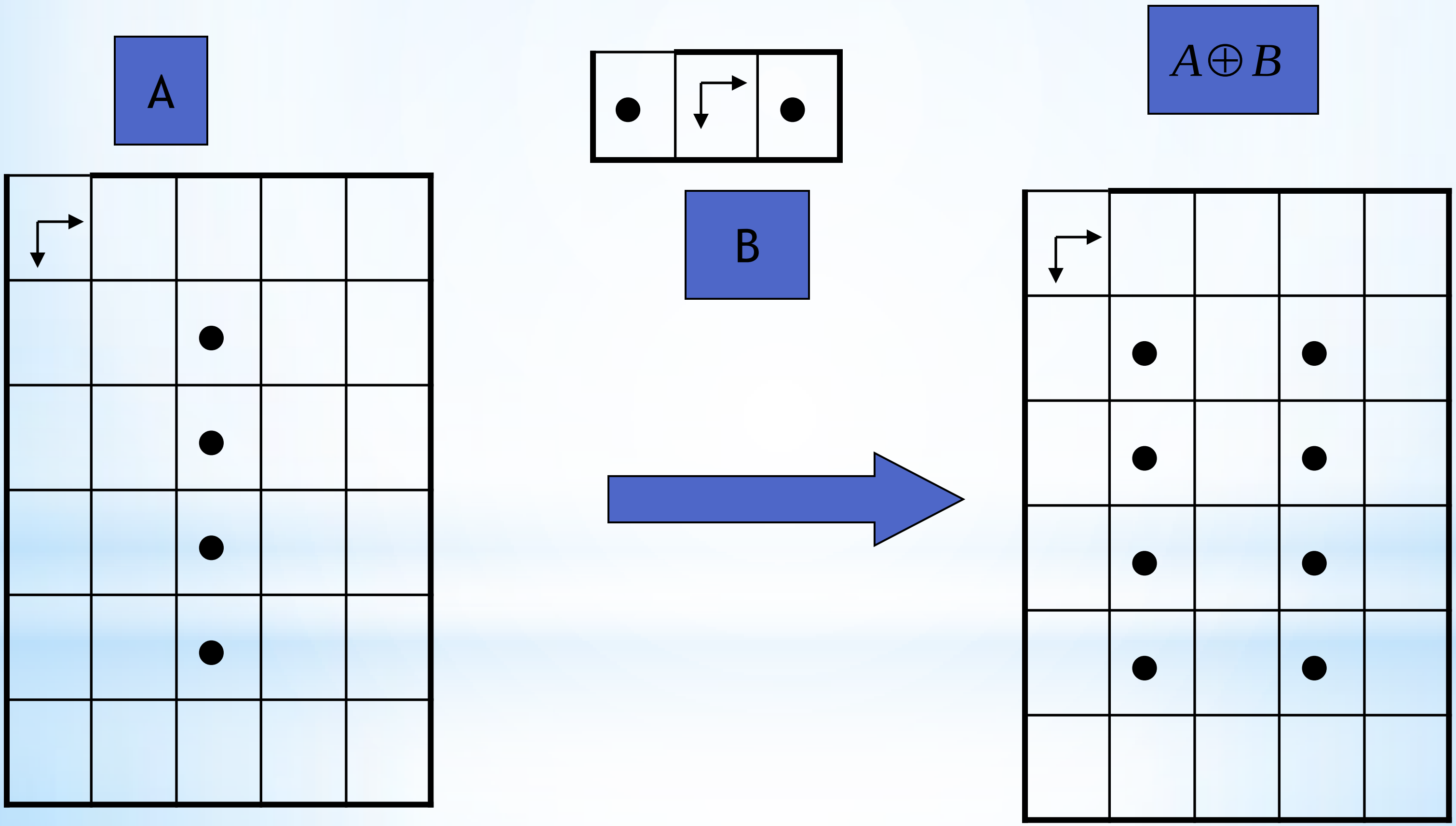
$$A \oplus B = B \oplus A$$

$$A \oplus (B \oplus C) = (A \oplus B) \oplus C$$

$$if \;\; 0 \in B, A \subseteq A \oplus B$$

$$A \subseteq B \; implies \;\; A \oplus D \subseteq B \oplus D$$

*Extensitivity



* Properties of Dilation

- Translation Invariance
- Linearity
- Conta
- Decor

$$(A)_x \oplus B = (A \oplus B)_x$$

$$(A \cup B) \oplus C = (A \oplus C) \cup (B \oplus C)$$

$$(A \cap B) \oplus C \subseteq (A \oplus C) \cap (B \oplus C)$$

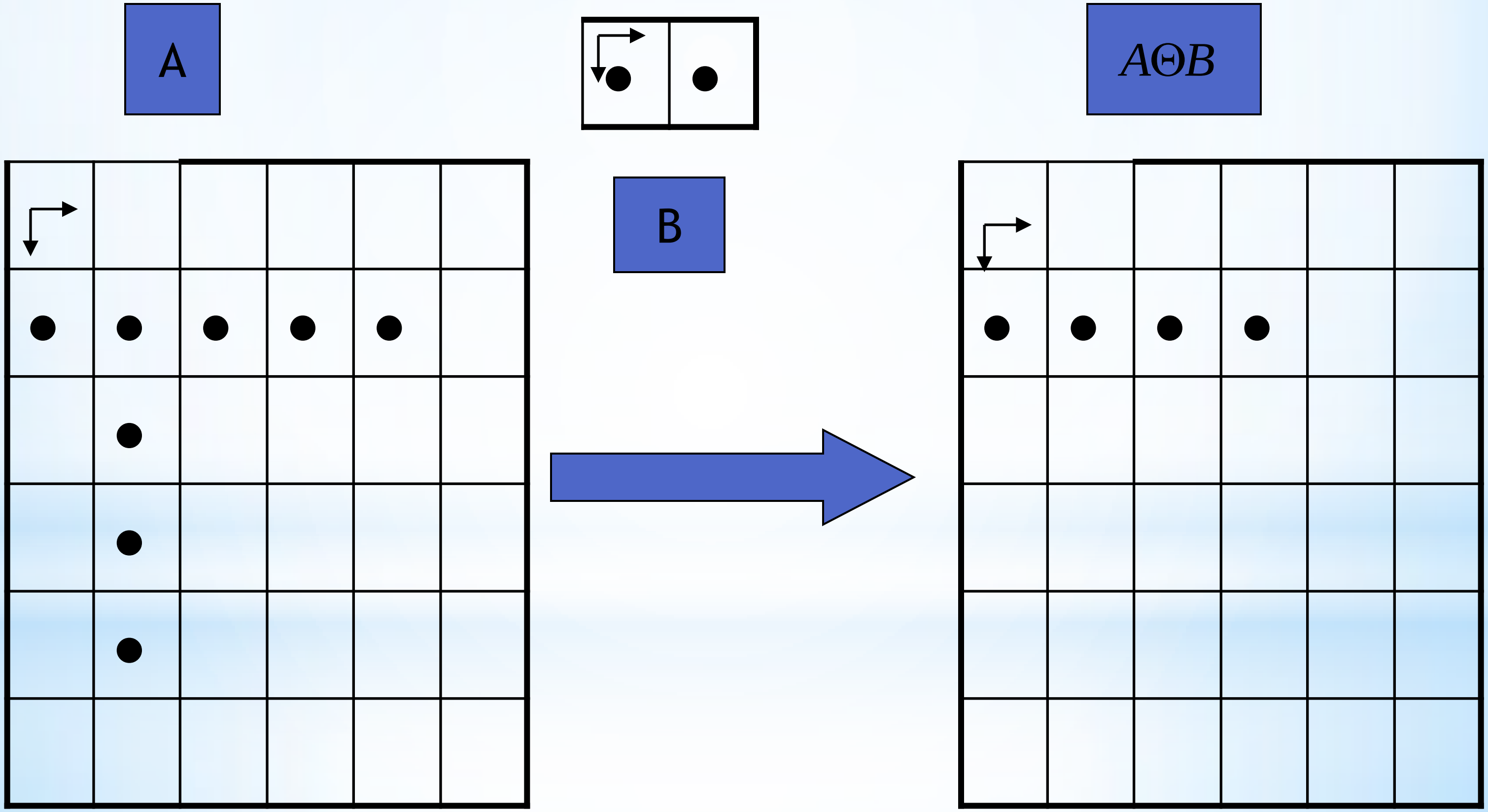
$$A \oplus (B \cup C) = (A \oplus B) \cup (A \oplus C)$$

Erosion

- Erosion is the morphological dual to dilation. It combines two sets using the vector subtraction of set elements.
- Let $A \Theta B$ denotes the erosion of A by B

$$A \Theta B = \{x \in Z^2 \mid \; for \;\; every \; b \in B, exist \; an \; a \in A \; s.t. x = a - b\}$$
$$= \{x \in Z^2 \mid x + b \in A \;\; for \;\; every \;\; b \in B)$$

Erosion



Erosion

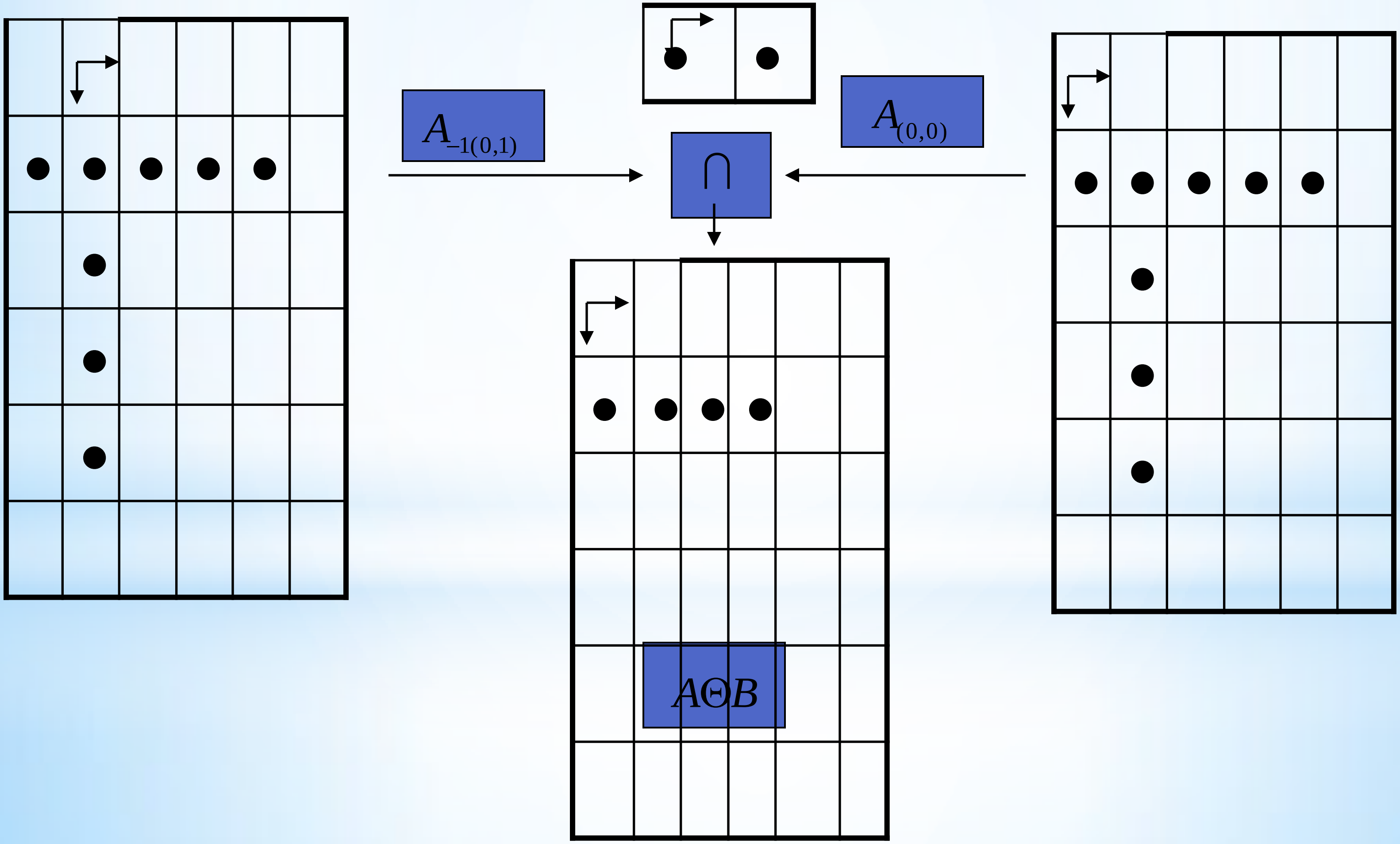
- Erosion can also be defined in terms of translation

$$A \Theta B = \{x \in Z^2 \,|\, (B)_x \subseteq A)$$

- In terr

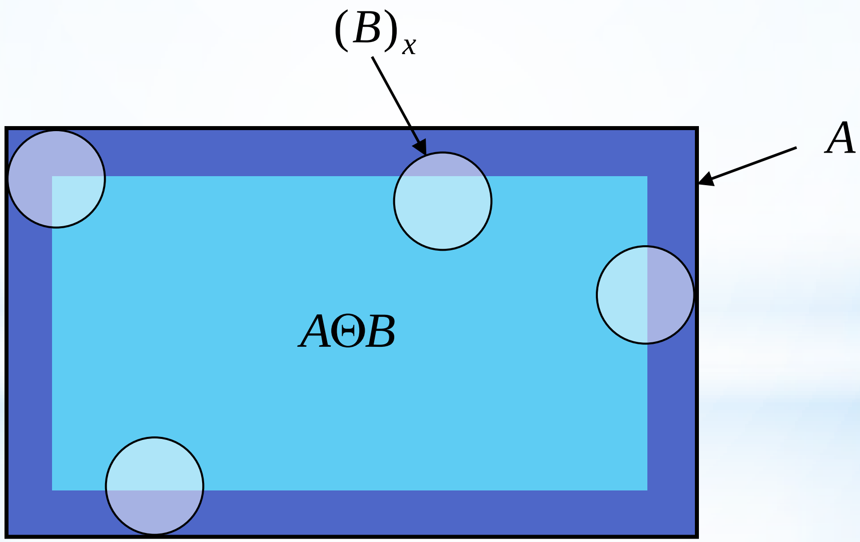
$$A \Theta B = \bigcap_{b \in B} (A)_{-b}$$

*Erosion



*Erosion

$$A \Theta B = \{x \in Z^2 \mid (B)_x \subseteq A\}$$



- Erosion is not commutative!

$$A \Theta B \neq B \Theta A$$

- Extensivity

$$if \;\; 0 \in B, A \Theta B \subseteq A$$

- Dilat

$$A \subseteq C \text{ implies } A \Theta B \subseteq C \Theta B, \; B \supseteq C \text{ implies } A \Theta B \subseteq A \Theta C$$

- Chai

$$A \Theta (B_1 \oplus \ldots \oplus B_k) = (\ldots(A \Theta B_1) \Theta \ldots \Theta B_k)$$

* Properties of Erosion

- Translation Invariance
- Linearity
- Conta
- Decom

$$A_x \Theta B = (A \Theta B)_x,\ \ A \Theta B_x = (A \Theta B)_{-x}$$

$$(A \cap B) \Theta C = (A \Theta C) \cap (B \Theta C)$$

$$(A \cup B) \Theta C \supseteq (A \Theta C) \cup (B \Theta C)$$

$$A \Theta (B \cup C) = (A \Theta B) \cap (A \Theta C)$$

Duality Relationship

- Dilation and Erosion transformation bear a marked similarity, in that what one does to image foreground and the other does for the image background.
- $B \in Z^2$, the reflection of B, $\breve{B}$ is defined as

$$\breve{B} = \{x \mid for\ some\ b \in B,\ x = -b\}$$

- Erosion and Dilation Duality Theorem

$$(A \Theta B)^c = A^c \oplus \breve{B}$$

Opening and Closing

- Opening and closing are iteratively applied dilation and erosion

Opening

Closing

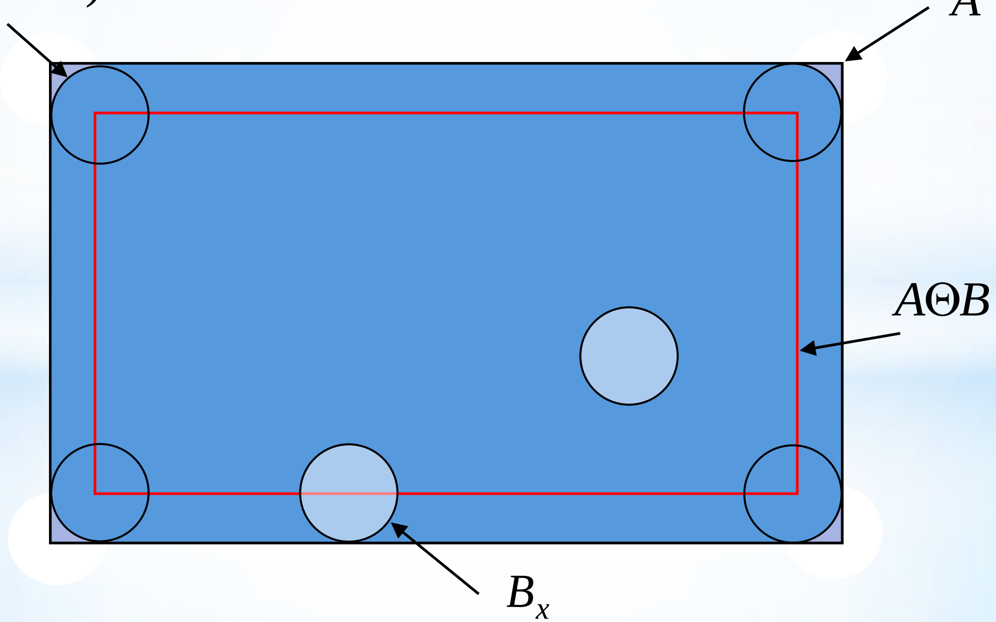
$$A \circ B = (A \Theta B) \oplus B$$

$$A \bullet B = (A \oplus B) \Theta B$$

*Opening and Closing

$$A \circ B = \bigcup_{\{x | B_x \subseteq A\}} B_x$$

$$A \circ B = (A \Theta B) \oplus B$$



- They are idempotent. Their reapplication has not further effects to the previously transformed result

$$A \bullet B = (A \bullet B) \bullet B$$

*Op[illegible] Closing

$$A \circ B = (A \circ B) \circ B$$

Opening and Closing

- Translation invariance

$$A \circ (B)_x = A \circ B \qquad A \bullet (B)_x = A \bullet B$$

- Antiextensivity of opening

$$A \circ B \subseteq A$$

- Extensivity of closing

$$A \subseteq A \bullet B$$

- Duality

$$(A \bullet B)^c = A^c \circ \breve{B}$$

- Main idea
  - Examine the **geometrical** structure of an image by matching it with small patterns called structuring elements at various locations
  - By varying the **size** and **shape** of the matching patterns, we can extract useful information about the shape of the different parts of the image and their interrelations.

*Morphological Filtering

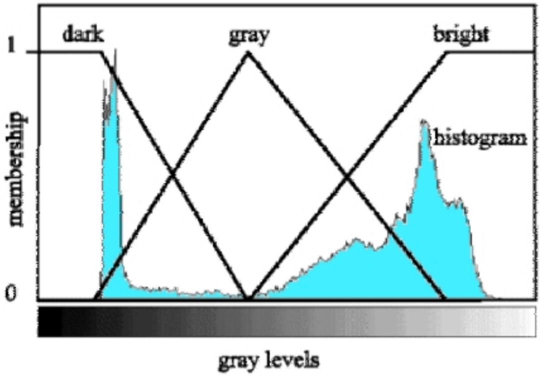
Summary

- Mathematical morphology is an approach for processing digital image based on its **shape**
- The language of morphology is **set theory**
- The basic morphological operations are **erosion** and **dilation**
- Morphological filtering can be developed to extract useful shape information

Fuzzy Rule-Based

- Step 1:
  Setting the parameters of inference system (input features, membership functions,..)

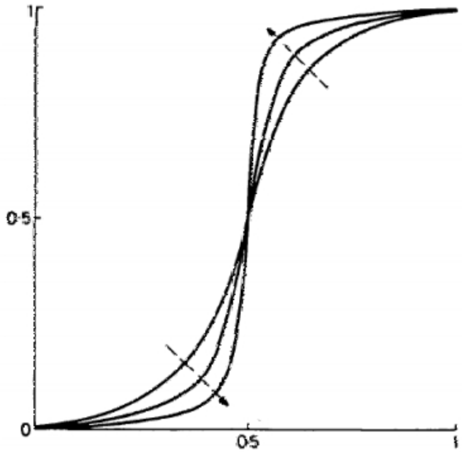
- Step 2:
  Fuzzification of the actual pixel (memberships to the dark, gray and bright sets of pixels)



INT-Operator (Contd…)

- Step 3: Generate new gray-levels

$$g'_{mn} = G^{-1}\left(\mu'_{mn}\right) = g_{max} - F_d\left(\left(\mu'_{mn}\right)^{\frac{-1}{F_e}} - 1\right)$$



INT-Operator

- Step 1: Define the membership function

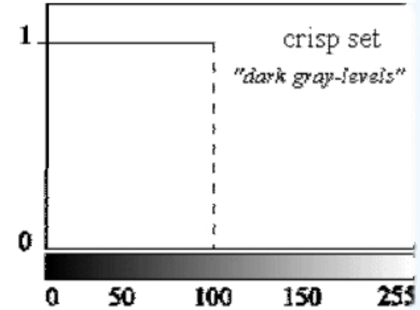
$$\mu_{mn} = G(g_{mn}) = \left[1 + \frac{g_{max} - g_{mn}}{F_d}\right]^{-F_e}$$

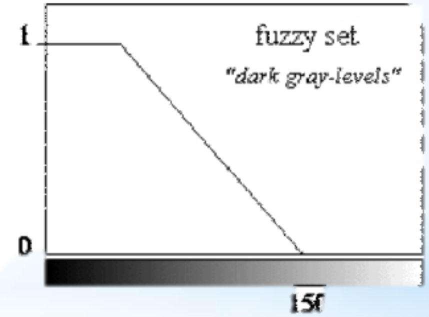
- Step 2: Modify the membership values

$$\mu'_{mn} = \begin{cases} 2 \cdot [\mu_{mn}]^2 & 0 \le \mu_{mn} \le 0.5 \\ 1 - 2 \cdot [1 - \mu_{mn}]^2 & 0.5 \le \mu_{mn} \le 1 \end{cases}$$

Fuzzy Sets

- Fuzzy set theory is the extension of conventional (crisp) set theory
- It handles the concept of partial truth using a membership function
- Instead of just black and white, the color belonging to a set has degree of whiteness & blackness

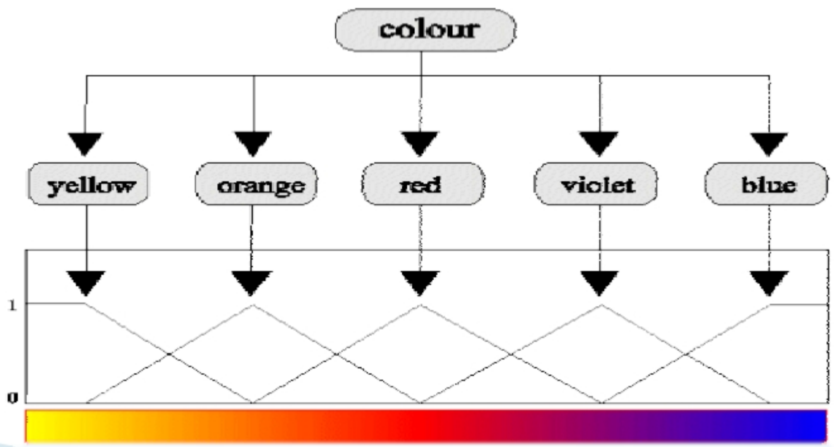




Contd

- As an example, we can regard the variable color as a fuzzy set

color = {yellow, orange, red, violet, blue}



The Learning Process

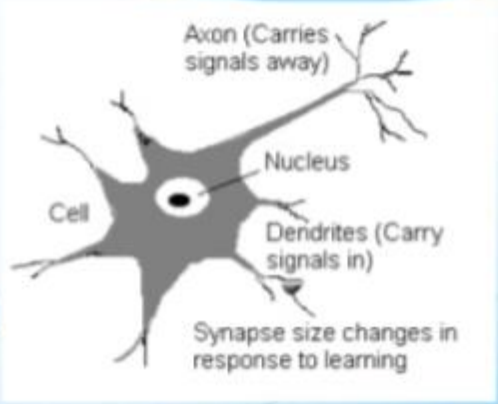
- **Associative mapping** in which the network learns to produce a particular pattern on the set of input units whenever another particular pattern is applied on the set of input units. The associative mapping can generally be broken down into two mechanisms:

Why using neural networks?

* Neural networks enable us to find solution where algorithmic methods are computationally intensive or do not exist.
* There is no need to program neural networks they learn with examples.
* Neural networks offer significant speed advantage over conventional techniques.

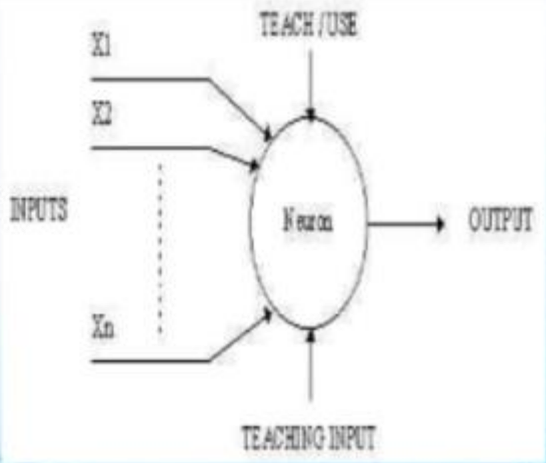
Inspiration from Neurobiology

* A neuron: many-inputs / one-output unit
* output can be *excited* or *not excited*
* incoming signals from other neurons determine if the neuron shall excite ("fire")
* Output subject to attenuation in the *synapses*, which are junction parts of the neuron



A simple neuron

* Takes the Inputs .
* Calculate the summation of the Inputs .
* Compare it with the threshold being set during the learning stage.



Features of finger prints

Finger prints are the unique pattern of ridges and valleys in every person's fingers.

- Their patterns are permanent and unchangeable for whole life of a person.
- They are unique and the probability that two fingerprints are alike is only 1 in 1.9x10^15.
- Their uniqueness is used for identification of a person.

Preview

- "***Morphology*** " – a branch in biology that deals with the **form and structure** of animals and plants.
- "***Mathematical Morphology***" – as a tool for **extracting image components**, that are useful in the **representation and description of region shape**
- **What are the applications of Morphological Image Filtering?**
  - boundaries extraction
  - skeletons
  - convex hull
  - morphological filtering
  - thinning
  - Pruning
- The language of mathematical morphology is – **Set theory**.
- Unified and powerful approach to numerous image processing problems

Sets in mathematical morphology represent objects in an image:

- **binary image** (0 = black, 1 = white) :

the element of the set is the coordinates (x,y) of pixel belong to the object Z2

- **gray-scaled image** :

the element of the set is the coordinates (x,y) of pixel belong to the object and the gray levels Z3

9.1 Basic Concepts in Set Theory

- Subset

  $A \subseteq B$
- Union

  $A \cup B$
- Intersection

  $A \cap B$

  disjoint / mutually exclusive $A \cap B = \emptyset$
- Complement $A^c \equiv \{w \mid w \notin A\}$
- Difference $A - B \equiv \{w \mid w \in A, w \notin B\} = A \cap B^c$
- Reflection $\mathcal{B} \equiv \{w \mid w = -b, \quad \forall b \in B\}$
- Translation $(A)z = \{c \mid c = a + z, \quad \forall a \in A\}$

9.2.1 Erosion

- Erosion is used for shrinking of element A by using element B
- Erosion for Sets A and B in Z^2, is defined by the following equation:

$$A \ominus B = \{z \mid (B)_z \subseteq A\} \qquad (9.2\text{-}1)$$

$$A \ominus B = \{z \mid (B)_z \cap A^c = \varnothing\} \qquad (9.2\text{-}2)$$

- This equation indicates that **the erosion of A by B is the set of all points z such that B, translated by z, is contained in A.**
- Erosion can be used to
  - **Shrinks** or **thins** objects in binary images
  - **Remove image components**(**how?**)
    - Erosion is a morphological filtering operation in which image details smaller than the structuring elements are filtered(removed)

9.2.1 Erosion – Example 1

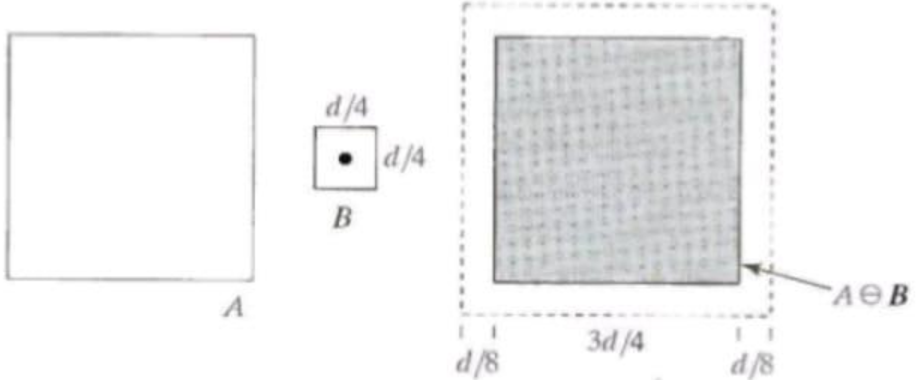


FIGURE 9.6 (a) Set A. (b) Square structuring element. (c) Erosion of A by B shown shaded

9.2.2 Dilation

- Dilation is used for expanding an element A by using structuring element B
- Dilation of A by B and is defined by the following equation:

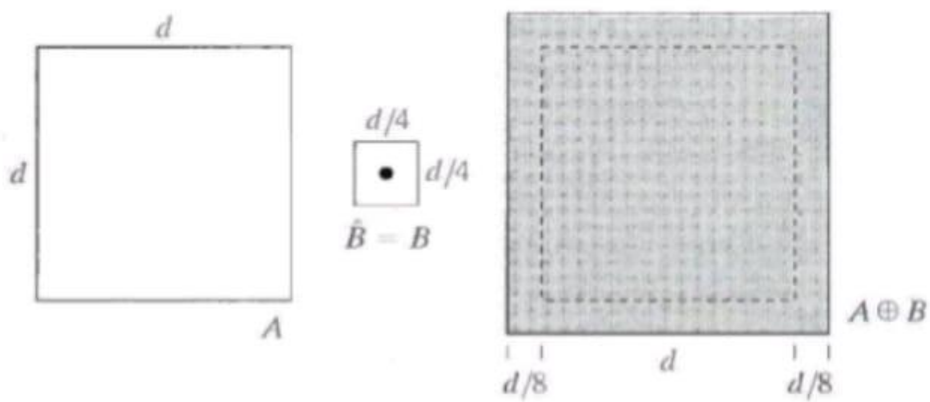
$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \varnothing\} \qquad (9.2\text{-}3)$$

- This equation is based on **obtaining the reflection of B about its origin and shifting this reflection by z**.
- The dilation of A by B **is the set of all displacements z, such that $\hat{B}$ and A overlap by at least one element**. Based On this interpretation the equation of (9.2-1) can be rewritten as:

$$A \oplus B = \{z | [(\hat{B})_z \cap A] \subseteq A\} \qquad (9.2\text{-}4)$$

- Relation to Convolution mask:
  - Flipping
  - Overlapping

9.2.2 Dilation – Example 1



a b c

FIGURE 9.4
(a) Set A.
(b) Square structuring element (dot is the center).
(c) Dilation of A by B, shown shaded.