Application of fuzzy set theory in image analysis

Nataša Sladoje

Centre for Image Analysis

Our topics for today

- Crisp vs fuzzy
- Fuzzy sets and fuzzy membership functions
- Fuzzy set operators
- Approximate reasoning
- Defuzzification
- Fuzziness and images
- Fuzzy segmentation methods
  - Fuzzy tresholding
  - Fuzzy c-means
  - Fuzzy connectedness

Fuzzy systems –why

- Fuzzy systems and models are capable of representing diverse, inexact, and inaccurate information.
- The qualifiers they can deal with are like those used by humans describe knowledge, i.e., they are linguistic variables.
- Examples: a rotten apple, a bright image, a medium dark wall, a dark sky.

Fuzzy systems and knowledge reprezentation

- Two forms of knowledge:
  - Objective knowledge – mathematical knowledge, used in engineering problems.
  - Subjective knowledge – exists in linguistic form, often not possible to quantify.

- Fuzzy systems can coordinate these two forms of knowledge.
- Fuzzy systems can handle numerical data and linguistic knowledge simultaneously.

What is a fuzzy set?

What is a set? ”... to be an element...”

Let us observe a set X, $X = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$

Let us form a subset C of X, $C = \{x \mid 3 < x < 8\}$.

$C = \{4, 5, 6, 7\}$

Easy! ”Yes or no.”

C is a **crisp** set.

What is a fuzzy set?

What is a set? ”... to be an element...”

Let us observe a set X, $X = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$

Let us form a subset C of X, $C = \{x \mid 3 < x < 8\}$.

$C = \{4, 5, 6, 7\}$

Easy! ”Yes or no.”

C is a **crisp** set.

Let us form a subset F of big numbers in X

$F = \{10, 9, 8, 7, 6, 5, 4, 3, 2, 1\}$

”Yes or no”? More like graded.

F is a **fuzzy** set.

Crisp vs. Fuzzy

Crisp

Accept, or reject.

A characteristic function of a set $A(x) = \begin{cases} 1, & if \;\; x \in A \\ 0, & if \;\; x \notin A \end{cases}$

An example: Set of apples.

Fuzzy

Admit intermediate values of memberships to a set.

An example: Set of ripe apples.

Example – Fuzzy set of *tall men*

Name	Height, cm	Degree of Membership	
		Crisp	*Fuzzy*
Chris	208	1	1.00
Mark	205	1	1.00
John	198	1	0.98
Tom	181	1	0.82
David	179	0	0.78
Mike	172	0	0.24
Bob	167	0	0.15
Steven	158	0	0.06
Bill	155	0	0.01
Peter	152	0	0.00

Crisp: A man is either tall, or not.

Fuzzy: The degree of membership to a set of tall men depends on the height.

Fuzzy sets and fuzzy membership functions

Each element of a reference set is assigned its degree of belongingness to a fuzzy set.

Define a fuzzy set ↔ Define a membership function

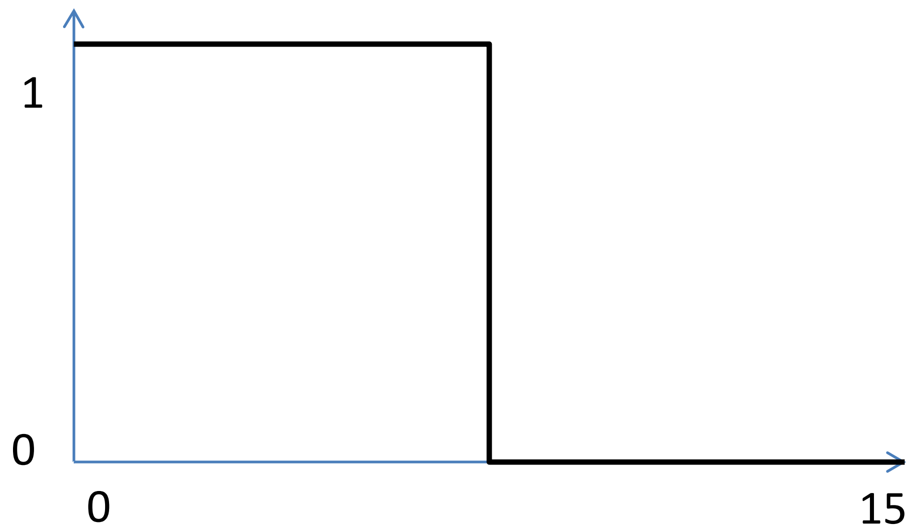
A fuzzy subset S of a reference set X is a set of ordered pairs $S = \{(x, \mu(x)) \mid x \in X\}$ where the membership function $\mu(x) \in [0,1]$ represents the grade of membership of *x in S.*

Example – Small numbers

x	$\mu_S(x)$
0	1
2	1
3	1
5	1
7	1
7	1
8	0
9	0
10	0
11	0
13	0
15	0

Reference set

$X=\{0,15,13,2,11,7,8,9,3,7,5,10\}$

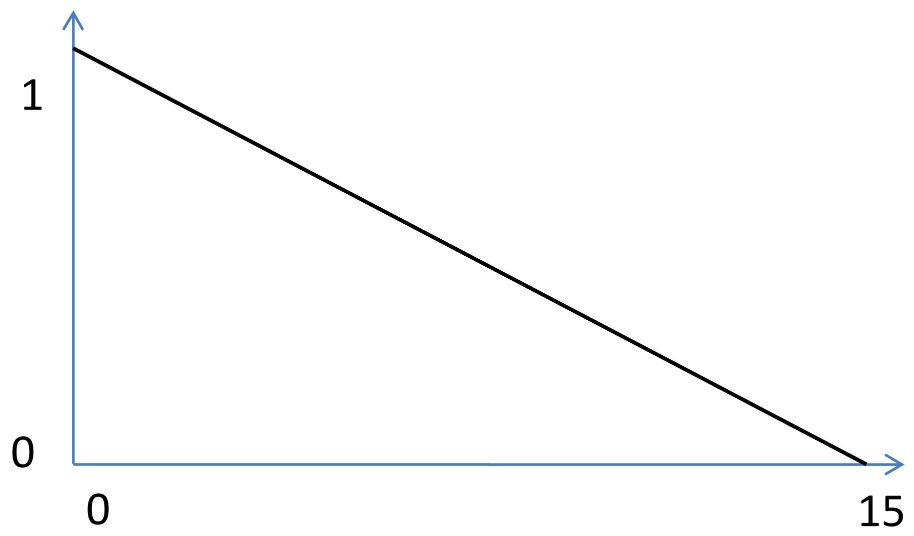


Example – Small numbers

x	$\mu_S(x)$
0	1
2	13/15
3	12/15
5	10/15
7	8/15
7	8/15
8	7/15
9	6/15
10	5/15
11	4/15
13	2/15
15	0

Reference set

X={0,15,13,2,11,7,8,9,3,7,5,10}

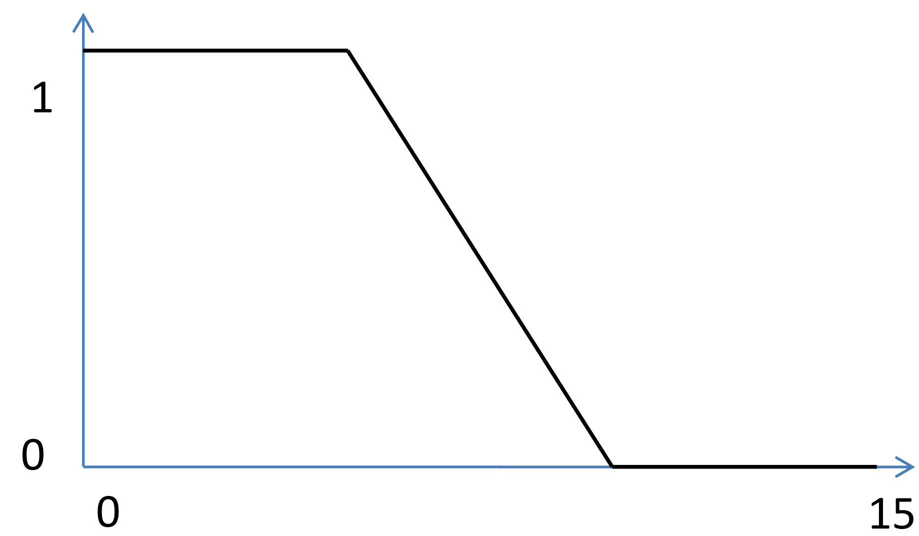


Example – Small numbers

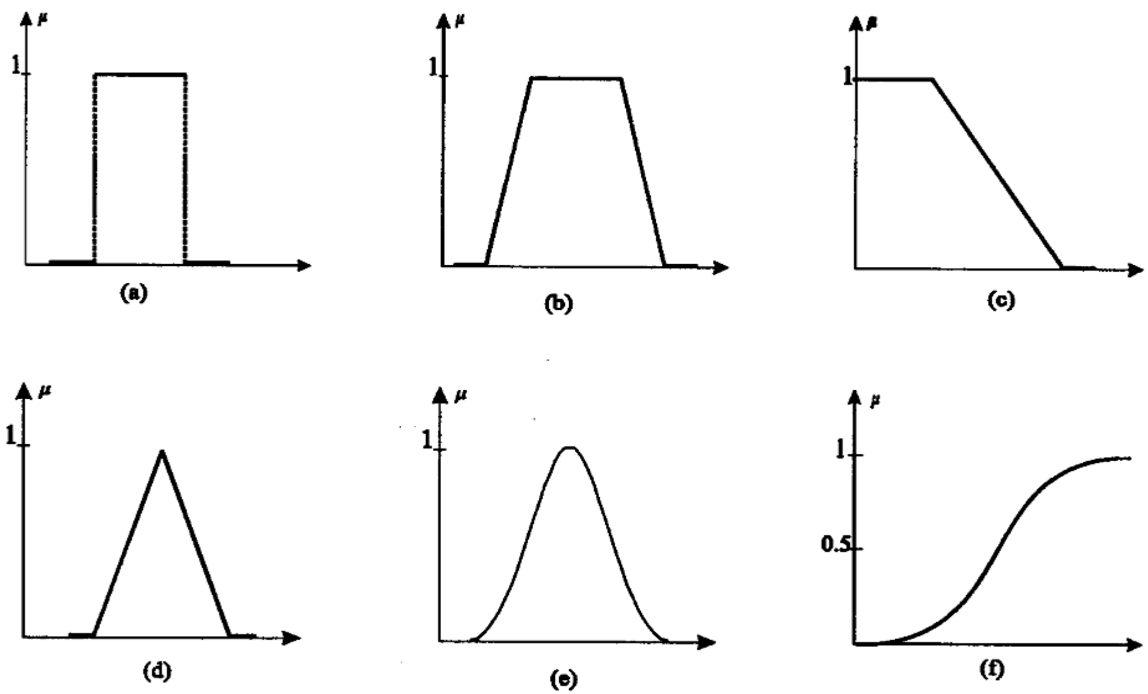
x	$\mu_S(x)$
0	1
2	1
3	1
5	1
7	2/4
7	2/4
8	1/4
9	0
10	0
11	0
13	0
15	0

Reference set

X={0,15,13,2,11,7,8,9,3,7,5,10}



Membership functions -examples



Membership function

(a) π-function (crisp)

(b) Trapezoidal function

(c) Semi-trapezoidal

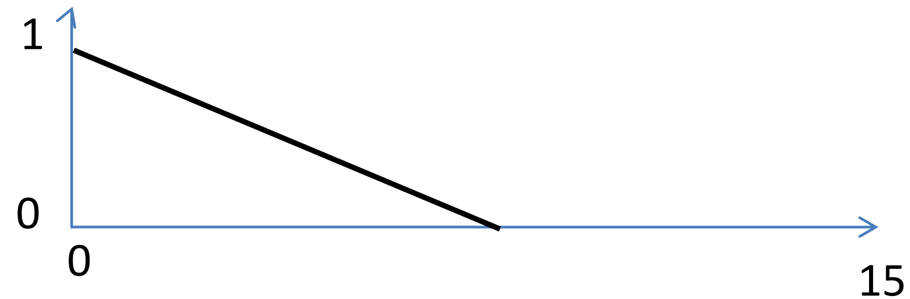
(d) Triangular function

(e) Gaussian

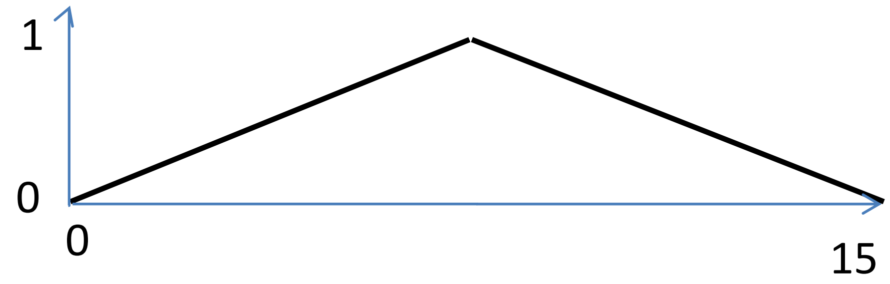
(f) S-function

Membership functions - examples

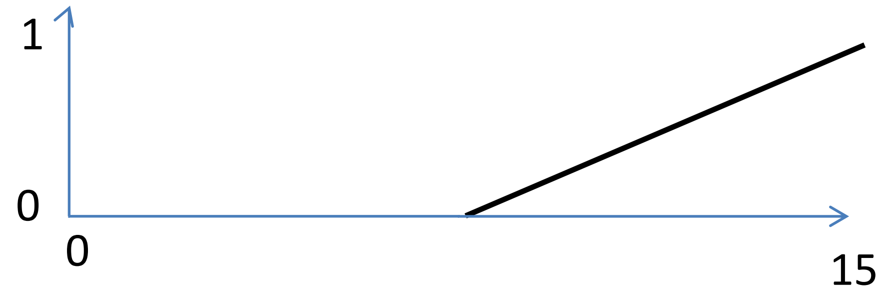
$$\mu_{SMALL}(x)$$



$$\mu_{MEDIUM}(x)$$



$$\mu_{LARGE}(x)$$



Fuzziness vs. Probability

Number 10 is not probably big!

...and number 2 is not probably not big.

Uncertainty is a concequence of

non-sharp boundaries

between the notions/objects,

and not because of lack of information.

Terminology:
Support , core, α-cut of a fuzzy set

- The **support** of a fuzzy set A is the (crisp) set of all elements of X with non-zero membership to A:

$$Supp(A) = \{x \in X \mid \mu_A(x) > 0\}$$

- The **core** of a fuzzy set A is the (crisp) set of all elements of X with membership to A equal one:

$$Core(A) = \{x \in X \mid \mu_A(x) = 1\}$$

- An **α-cut** of a fuzzy set A is a crisp set of all the elements in X with membership to A not smaller than α:

$${}^{\alpha}A = \{x \in X \mid \mu_A(x) \geq \alpha\}$$

Fuzzy set operations

As for crisp sets, we can define set operations for fuzzy sets…in infinitely many ways.

Three best known and most often applied are:

Intersection A *and* B

$$\mu_{A\cap B}(x)=\min(\mu_A(x),\mu_B(x))$$

Union A *or* B

$$\mu_{A\cup B}(x)=\max(\mu_A(x),\mu_B(x))$$

Complement *not* A

$$\mu_{A^C}(x)=1-\mu_A(x)$$

Approximate reasoning

General schema is of the form:

Rule 1:	If $\mathcal{X}$ is A_1, then $\mathcal{Y}$ is B_1
Rule 2:	If $\mathcal{X}$ is A_2, then $\mathcal{Y}$ is B_2
	...
Rule n:	If $\mathcal{X}$ is A_n, then $\mathcal{Y}$ is B_n
Fact:	$\mathcal{X}$ is A'
Conclusion:	$\mathcal{Y}$ is B'

A', A_j are fuzzy sets on X,

B', B_j are fuzzy sets on Y, for all j.

Approximate reasoning

Most common way to determine B' is by using

method of interpolation.

Step 1. Calculate the degree of consistency between the given fact and the antecedent of each rule.
Use height of intersection of the associated sets:

$$r_j(A') = h(A' \wedge A_j) = \sup_{x \in X} \min[A'(x), A_j(x)].$$

Step 2. Truncate each B_j by the value $r_j(A')$ and determine B' as the union of truncated sets:

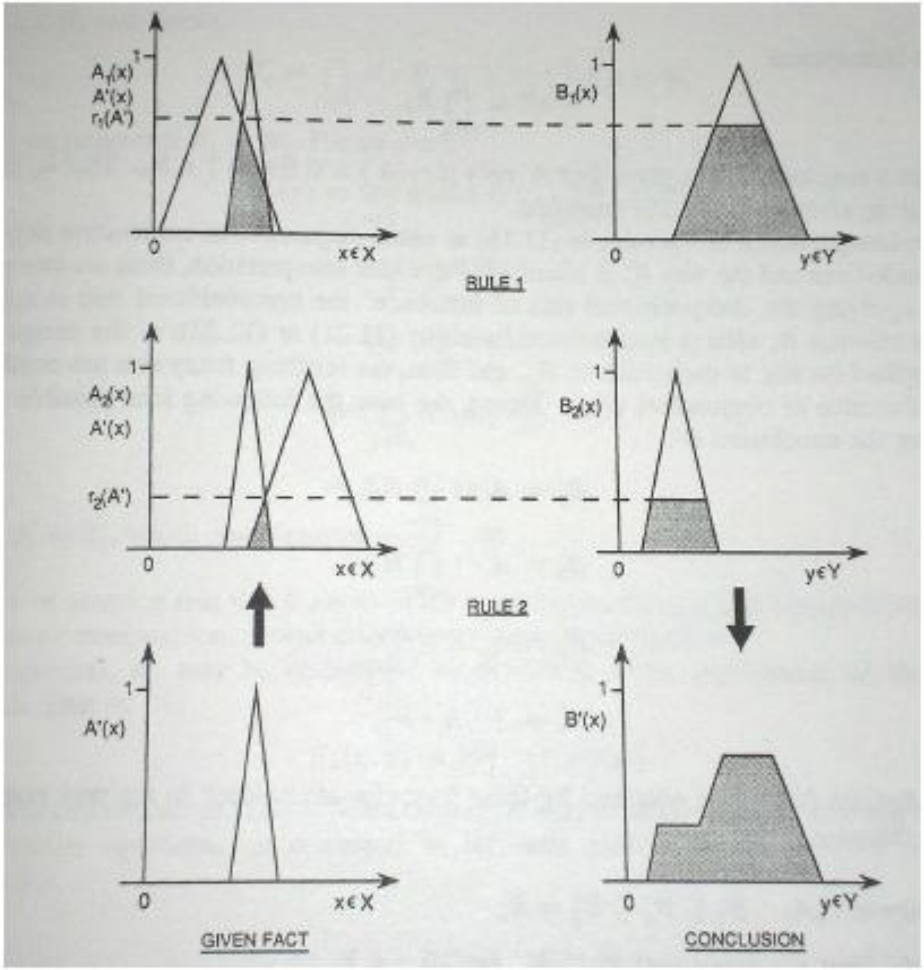
$$B'(y) = \sup_{j \in \mathbb{N}_n} \min[r_j(A'), B_j(y)], \quad \text{for all } y \in Y.$$

Note that interpolation method is a special case of the composition rule of inference, with

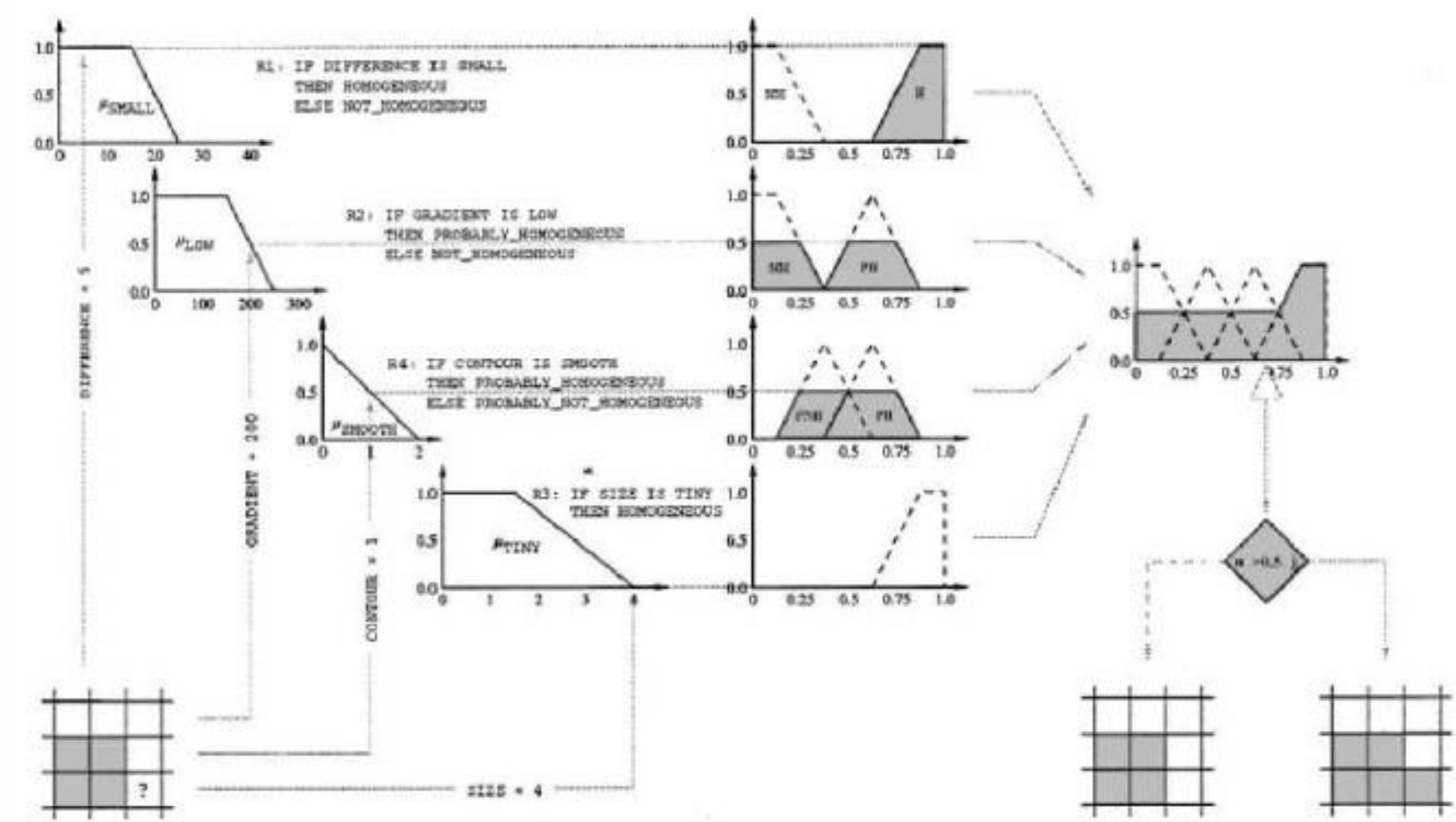
$$R(x, y) = \sup_{j \in \mathbb{N}_n} \min[A_j(x), B_j(y)]$$

where then $B'(y) = \sup_{x \in X} \min[A'(x), R(x, y)] = (A' \circ R)(y)$.

Approximate reasoning



An application
Region growing using fuzzy rule based system



Defuzzification

To find a crisp solution, we need **defuzzification**.

We want to select a good crisp representative of a fuzzy set.

Defuzzification to a point:

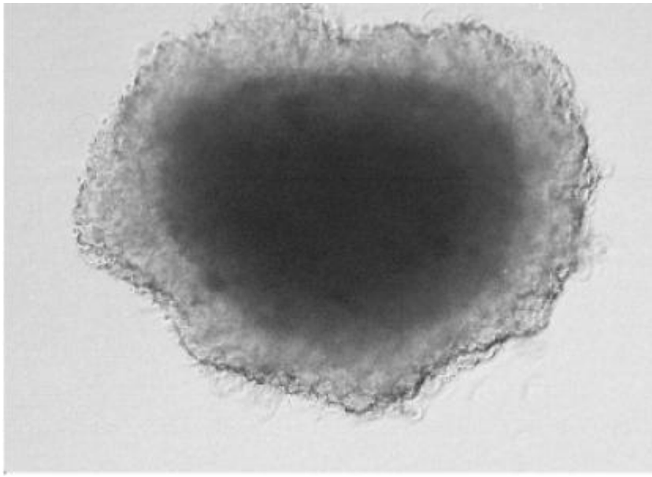
- Composite moments: Select the centroid of the fuzzy set.
- Composite maximum: Select a point from the core of the fuzzy set.

Defuzzification to a set:

- Most often an appropriate α-cut is selected.

Fuzzy sets in image processing

- Image data are rarely of perfect quality.
- Fuzziness is intrinsic property of images.
- Fuzzy sets and fuzzy techniques are capable of representing diverse, non-exact, uncertain, and inaccurate knowledge or information.



Discrete spatial fuzzy sets

- Object of interest is represented as a (discrete) spatial fuzzy subset of a grid.
- The mapping $\mu: X \to [0,1]$ becomes

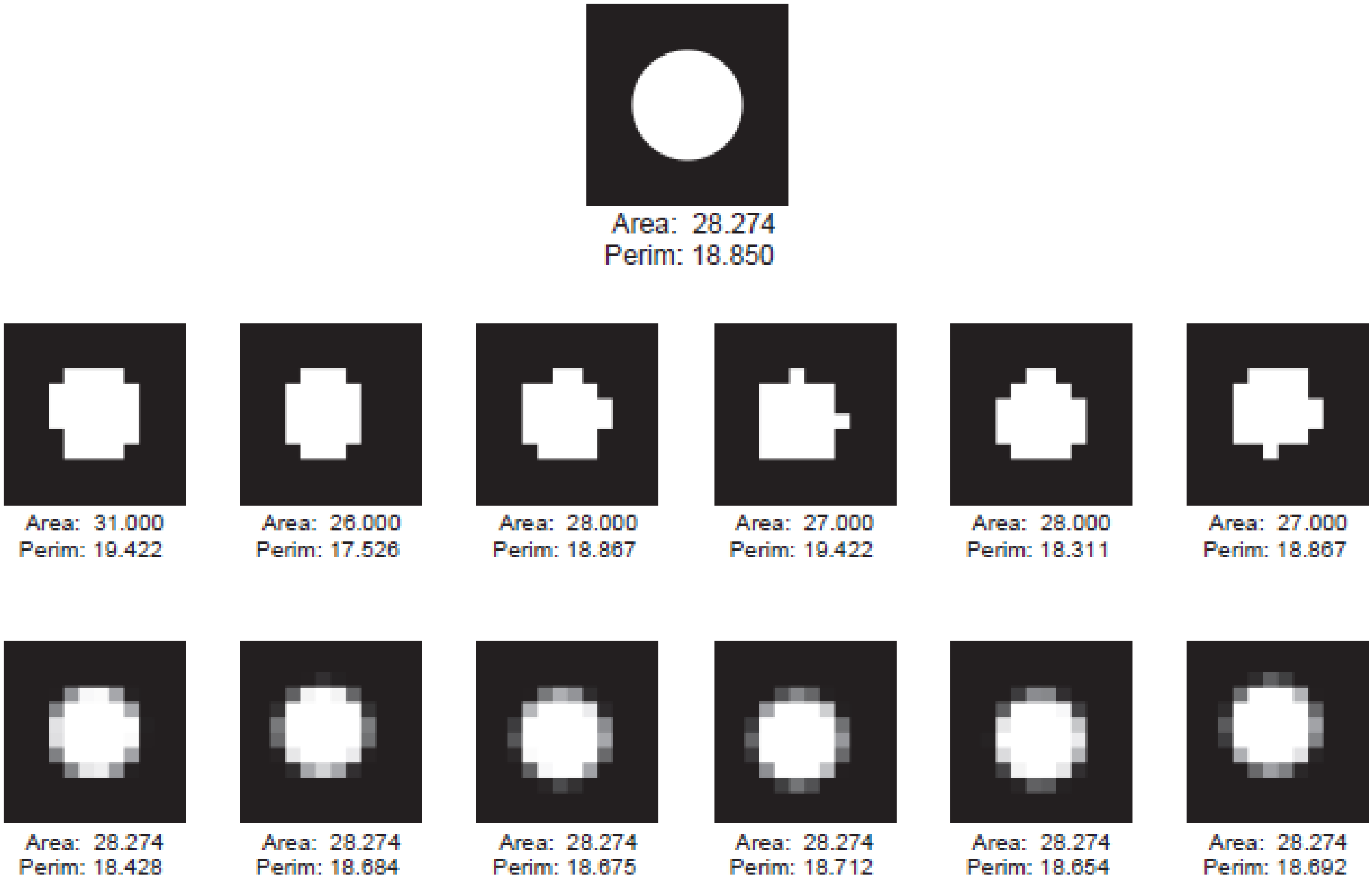
$$\mu: Z \times Z \to \{0, 1, 2, \dots, m\} \quad \text{(in 2D)}$$

m – maximal number of grey-levels available

(e.g., m=255 for 8-bit pixel representation)

- Negative effects of discretization (loss of data) can be signifficantly decreased by utilizing fuzzy object representations and corresponding analysis tools.

Objects with fuzzy borders – an example



Objects with fuzzy borders

- Most of the pixels in images are easily classified as object pixels, or as background pixels.
- Pixels close to the border of the object are more difficult to classify. They can, e.g., partly belong to several objects.
- We assign to them a fuzzy membership value according to the extent of their belongingness to the object.
- An intuitive aproach is the pixel/voxel coverage approach. Pixel/voxel value is determined as its relative size (area/volume) covered by the observed object.

Fuzzy thresholding

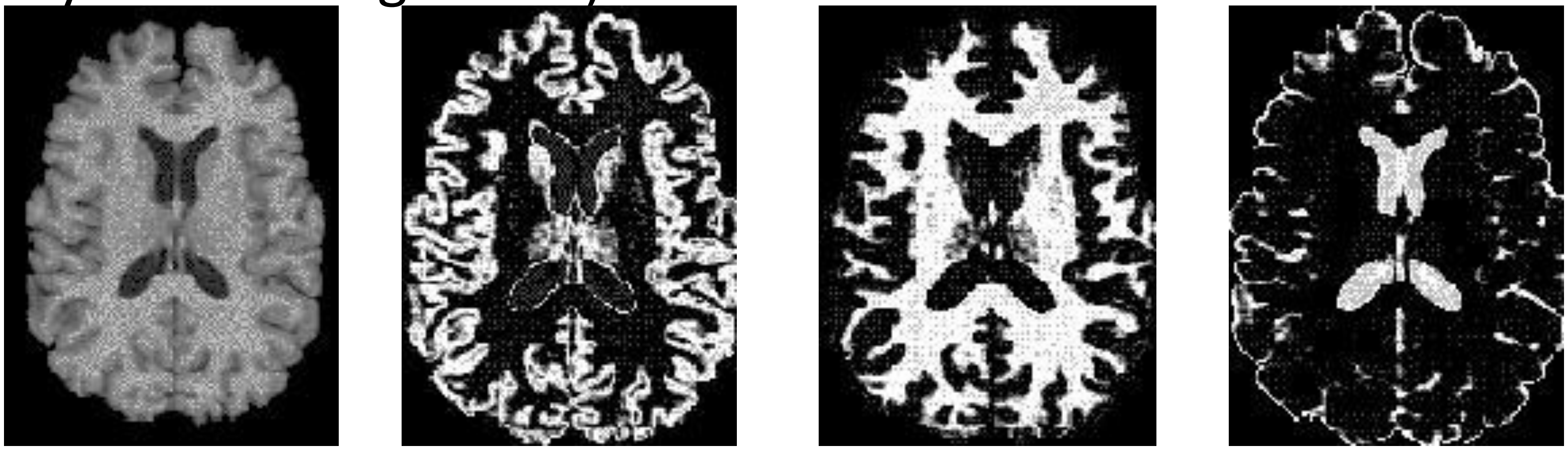
In general, instead of setting a hard threshold, we can apply fuzzy thresholding and obtain soft transitions between “in” and “out” regions. Fuzzy thresholding functions can be defined in many ways. Their general form is:

$$g(x,y) = \begin{cases} 0 & \text{if } f(x,y) < T_1 \\ \mu_{g(x,y)} & \text{if } T_1 \leq f(x,y) < T_2 \\ 1 & \text{if } T_2 \leq f(x,y) < T_3 \\ \mu_{g(x,y)} & \text{if } T_3 \leq f(x,y) < T_4 \\ 0 & \text{if } T_4 \leq f(x,y) \end{cases}$$

Fuzzy image is a grey-level image…

- (a) A sample slice from acquired MRI data set.
- Membership functions: (b) gray matter (GM),
  (c) white matter (WM),
  (d) cerebrospinal fluid (CSF).

(fuzzy c-means algorithm)



(a) (b) (c) (d)

Fuzzy c-means clustering

The fuzzy c-means algorithm (FCM) iteratively optimizes an objective function in order to detect its minima, starting from a reasonable initialization.

Its objective is to partition a collection of numerical data into a series of overlapping clusters. The degrees of belongingness are interpreted as fuzzy membership values.

Fuzzy c-means clustering
Extends K-means, ch. 9.2.5

- K-means algorithm is a popular, simple, non-parametric, non-hierarchical approach to cluster analysis.
- K-means algorithm minimizes the sum of within-cluster variances for the K observed clusters:

$$E_K^2 = \sum_{i=1}^{K} \sum_{k=1}^{n} I_{k,i} (d_{ki})^2$$

where $I_{k,i}$ is an element of a *n*x*K* matrix *I* which represents a *K*-partition of the data set $X = \{x_1, x_2, \ldots, x_n\}$, v_i is the cluster center of the class $i, (1 \le i \le K)$ and $d_{ki}^2 = \|x_k - v_i\|^2$, for an inner product norm metric $\|\cdot\|$.

K-means clustering

Example: Clustering of *n*=4 points into K=3 clusters.

Partition matrix contains the (crisp) membership of each point to each cluster.

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Point four (row) does not belong to cluster two (column).

Distance matrix contains the distances between each point and each cluster center.

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \\ d_{41} & d_{42} & d_{43} \end{bmatrix}$$

The distance between point four and center of cluster two.

K-means clustering

- Number of clusters, *K*, is decided in advance, preferably by some a priori knowledge. Naturally, $2 \leq K \leq n$.
- Distance measure determines the shape of the cluster; Euclidean will produce hyper-spherical clusters, Mahalanobis distance will lead to hyper-elliptical clusters.
- The algorithm iteratively updates cluster centers as the means of the clusters created in a previous iteration, re-computes the distances of the points to the new cluster centers and re-partitions the data.
- Partition of the data is crisp!
  - The matrix *I* contains only elements 0 and 1.
  - Each row (corresponding to an element) contains exactly one element equal to 1 (the element is assigned to exactly one cluster)

Fuzzy c-means clustering

- A partition of the observed set is represented by a nxc matrix $U=[u_{ki}]$
- u_{ki} corresponds to the membership value of the k^{th} element (out of n), to the i^{th} cluster (out of c).
- Boundaries between the subgroups are not crisp.
- Each element may belong to more than one cluster – its ”overall” membership equals one.
- Objective function includes parameter controlling degree of fuzziness.

Fuzzy c-means clustering

- The fuzzy *c*-means algorithm iteratively optimizes the objective function in order to detect its minima, starting from a reasonable initialization.
- The objective function belongs to the family of fuzzy *c*-means functionals, using a particular inner product norm metric as a similarity measure:

$$J_m = \sum_{i=1}^{c} \sum_{k=1}^{n} (u_{ki})^m (d_{ki})^2$$

where u_{ki} is an element of a *nxc* matrix *U* which represents a fuzzy *c*-partition of the data set $X = \{x_1, x_2, ..., x_n\}$, v_i is the cluster center of the class $i, (1 \le i \le c)$, $m, (1 \le m < \infty)$ is the parameter controlling fuzzines of the partition, and $d_{ki}^2 = \|x_k - v_i\|^2$, for an inner product norm metric $\|\cdot\|$.

Fuzzy c-means clustering

Example: Clustering of n=4 points into c=3 clusters.

Partition matrix contains the fuzzy membership of each point to each cluster.

$$U = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.3 & 0.3 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$$

Point four (row) has membership 0.1 to cluster two (column).

Distance matrix contains the distances between each point and each cluster center.

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \\ d_{41} & d_{42} & d_{43} \end{bmatrix}$$

The distance between point four and center of cluster two.

Fuzzy c-means algorithm: the basic steps

Choose

c – the number of clusters,

m - the weighting exponent (between 1.5 and 2.5),

the inner product induced norm metric (e.g., Euclidean norm),

the matrix norm (e.g., *sup* norm),

the terminating criterion.

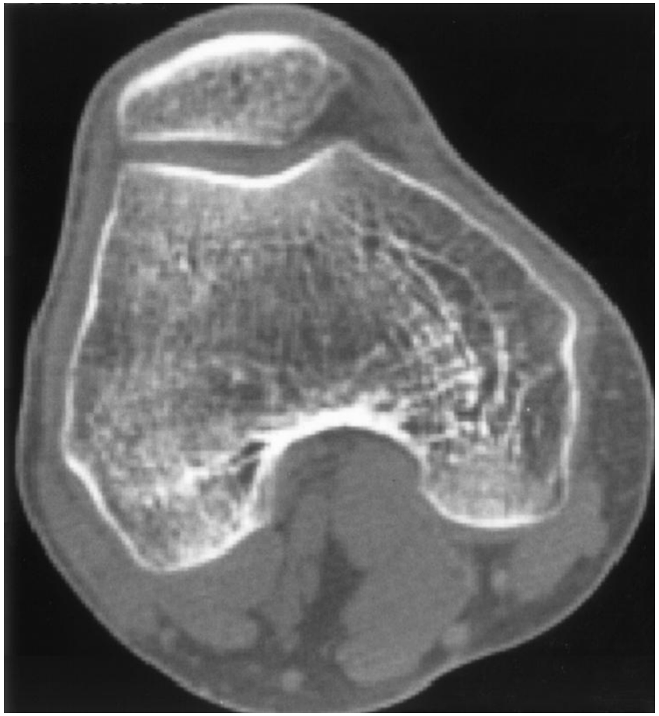
Initialize (randomly) the fuzzy c-partition or cluster centres vector.

Calculate iteratively the next partition, using formulae for updating membership values and cluster centres, starting from initial cluster centres and the initial partition.

Stop when two successive iterations produce partitions, or centres, that are close enough according to the given termination criterion when compared in the chosen matrix, or vector, norm.

Fuzzy connectedness (Ch. 7.4)

Images are intrinsically fuzzy



Graded composition

Heterogeneity of intensity in the object region due to heterogeneity of object material and blurring caused by the imaging device.

Hanging-togetherness

In spite of intensity heterogeneity, a human viewer readily sees natural grouping of voxels constituting an object in a display of the scene.

Hanging togetherness

If two regions have about the same grey-level and if they are relatively close to each other, then they likely belong to the same object (hang together).

To group pixels that seem to hang together

- Observe local hanging-togetherness based on
  similarity in spatial location
  similarity in intensity(-derived features)
- Determine relationship between each pair of pixels in the entire image.
- Derive global hanging-togetherness (connectedness).

Fuzzy connectedness

Fuzzy connectedness combines

- fuzzy adjacency (closeness in space)
- fuzzy affinity (closeness in terms of intensities or other properties)

and assigns a strength of connectedness to each pair of image points determined as the strength of the weakest link of the strongest path between the points.

Fuzzy connectedness

- A path $P_{c,d}$ between the points *c* and *d* is any sequence of points $\langle c = c_1, c_2, \ldots, c_n = d \rangle$.
- The strength of connectedness of a path $P_{c,d}$ is

$$\mu_{P_{c,d}} = \min_{j=1,\ldots,n-1} \mu_K(c_j, c_{j+1}).$$

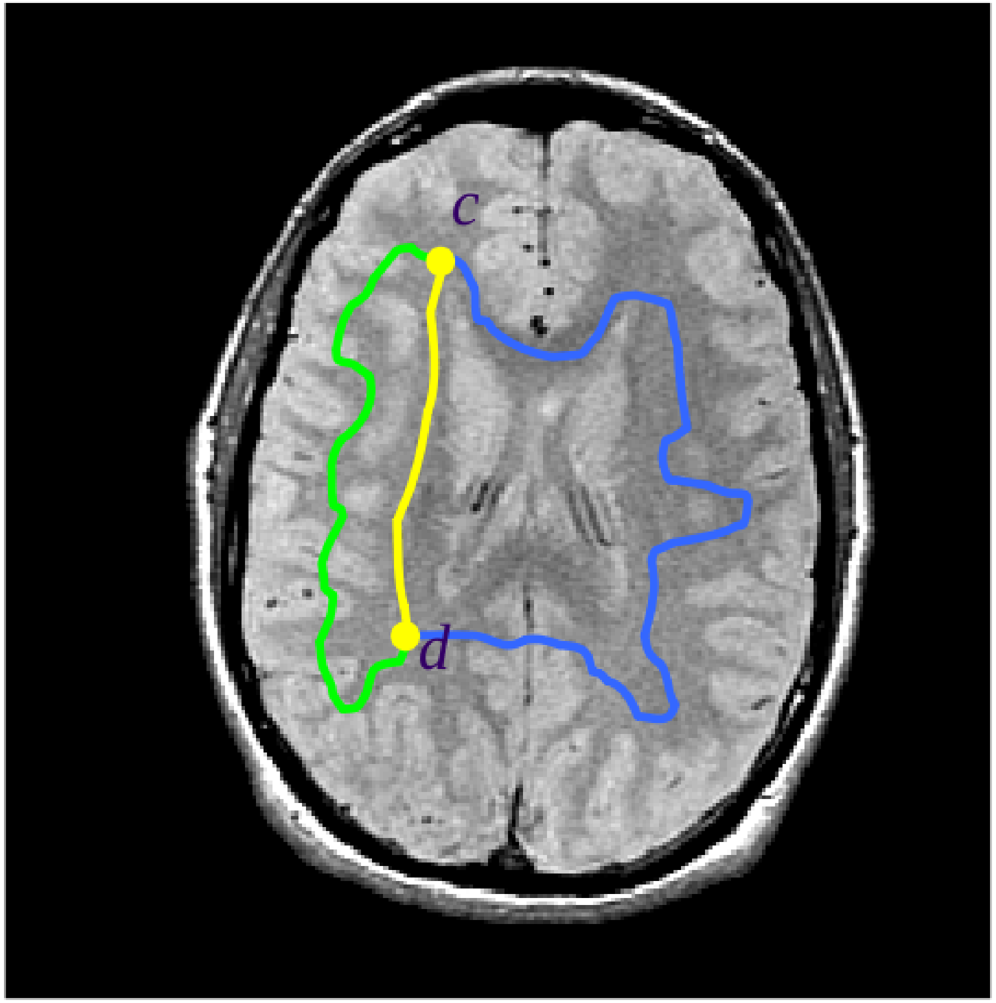
- Let $\mathbf{P}_{c,d}$ denote the set of all paths between *c* and *d.*

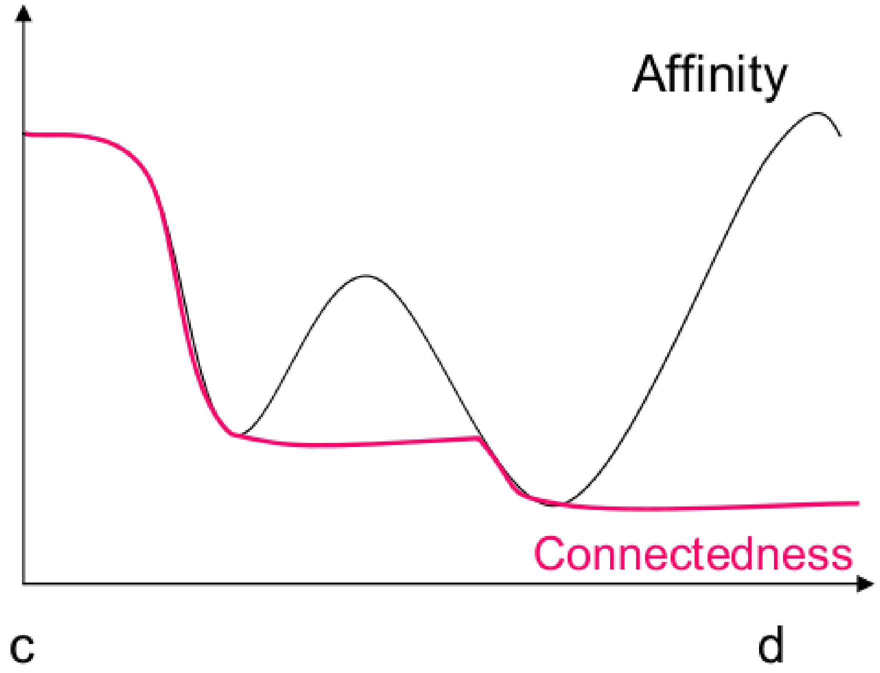
  The fuzzy connectedness between *c* and *d* is defined as

$$\mu(c,d) = \max_{P_{c,d} \in \mathbf{P}_{c,d}} \mu_{P_{c,d}}.$$

Fuzzy connectedness

The strength of the weakest link of the strongest path





Fuzzy affinity – local hanging- togetherness

$$\mu_K(c,d) = \mu_\omega(c,d) \cdot g(\mu_\varphi(c,d), \mu_\delta(c,d))$$

Fuzzy adjacency

Fuzzy affinity- homogeneity based component

Fuzzy affinity- object-feature based component

Fuzzy adjacency

Fuzzy adjacency function determines spatial closeness of the image elements.

Can be hard, when only the elemens with common face/edge (or vertex) have non-zero adjacency
(e.g., 4- or 8-adjacency in binary 2D images).

An example of fuzzy adjacency is

$$\mu_\omega(c,d) = \begin{cases} \dfrac{1}{1 + k_1\left(\sqrt{\sum_{i=1}^{n}(c_i - d_i)^2}\right)}, & \text{if } \sum_{i=1}^{n} |c_i - d_i| \leq n \\ 0, & \text{otherwise,} \end{cases}$$

Fuzzy affinity

$$\mu_K(c,d) = \mu_\omega(c,d) \cdot g(\mu_\varphi(c,d), \mu_\delta(c,d))$$

Expected properties of g :

Range within [0,1];

Monotonically non-decreasing in both arguments.

Examples:

$$\mu_K(c,d) = \frac{1}{2}\mu_\omega(c,d)\big(\mu_\varphi(c,d) + \mu_\delta(c,d)\big)$$

$$\mu_K(c,d) = \mu_\omega(c,d)\sqrt{\mu_\varphi(c,d) \cdot \mu_\delta(c,d)}$$

Fuzzy affinity
Homogeneity based component

The degree of local hanging-togetherness due to the similarity in intensity.

$$\mu_\varphi(c,d) = W_\varphi(|f(c) - f(d)|)$$

Expected properties of W_φ :

Range within [0,1] and $W_\varphi(0) = 1$;

Monotonically non-increasing.

Examples:

The right-hand side of an appropriately scaled box, trapezoid, or Gaussian function.

Fuzzy affinity
Object-feature-based component

The degree of local hanging-togetherness with respect to some given feature, e.g., intensity distribution.

$$\mu_\delta(c,d) = \begin{cases} 1 & \text{if } c = d \\ \dfrac{W_o(c,d)}{W_b(c,d) + W_o(c,d)} & \text{otherwise} \end{cases}$$

$$W_o(c,d) = \min\left[W_o\big(f(c)\big), W_o\big(f(d)\big)\right]$$
$$W_b(c,d) = \max\left[W_b\big(f(c)\big), W_b\big(f(d)\big)\right]$$

Expected properties of W_o and W_b

Range within [0,1].

Examples:

An appropriately scaled and shifted box, trapezoidal, or Gaussian function.

Fuzzy affinity

A concrete example

In the computer exercise (and in the book):

$$\mu_\varphi(c,d) = \frac{1}{1 + k_2|f(c) - f(d)|}$$

$$\mu_K(c,d) = \mu_\omega(c,d)\mu_\varphi(c,d)$$

$$\mu_K(c,d) = \frac{\mu_\omega(c,d)}{1 + k_2|f(c) - f(d)|}$$

An object as a fuzzy connected component

Given one or several seeds:

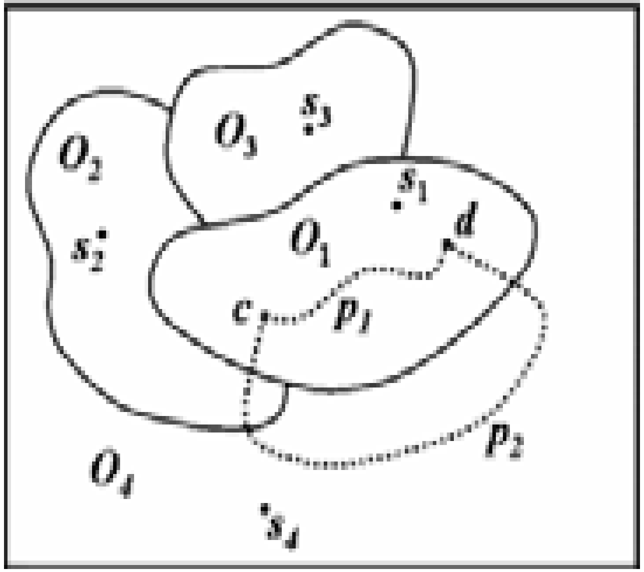
Compute connectedness map for all possible paths.

Threshold the connectedness map.

An object is a fuzzy connected component of a given strength.

Variations are proposed to improve the performance.

E.g., if the threshold is known in advance, computation can be more efficient.



An object as a fuzzy connected component

- How to set a threshold is not an easy question.
- The answers led to improvements of the initial idea of the (absolute) fuzzy connectivity algorithm.

Relative fuzzy connectivity (for two, as well as multiple objects)

Instead of thresholding the connectivity map, two (or more) objects are competing for points.

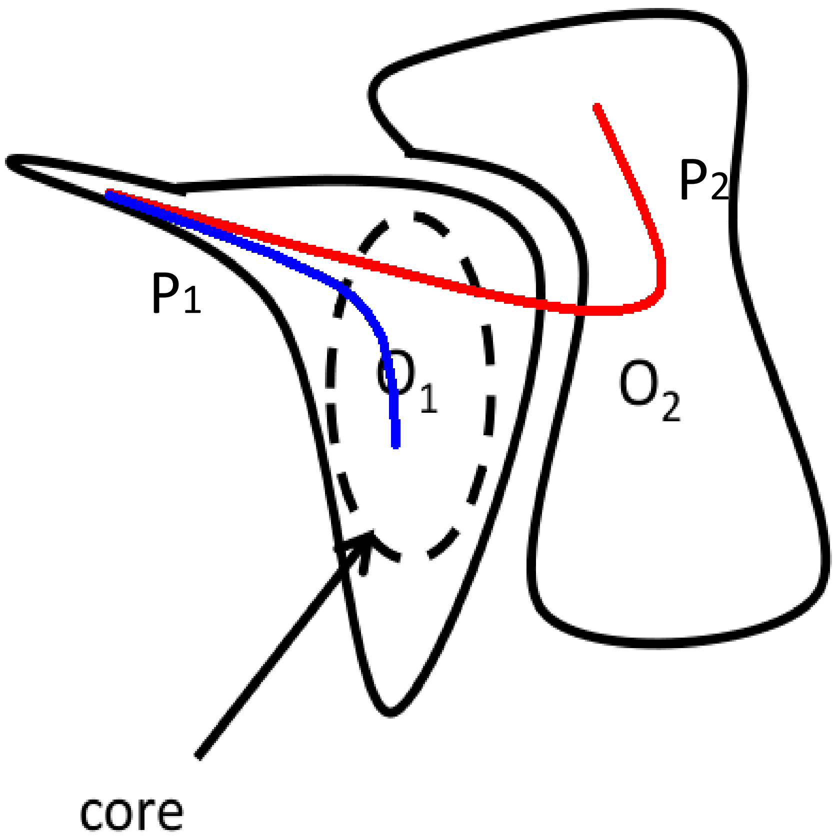
Iterative fuzzy connectivity

Repeated steps in fuzzy connectedness computation to overcome problems with weak object borders.

Scale-based fuzzy connectivity

Affinity is computed w.r.t. scale, and the scale is adapted to locations. Improved performance, at a considerable computational cost.

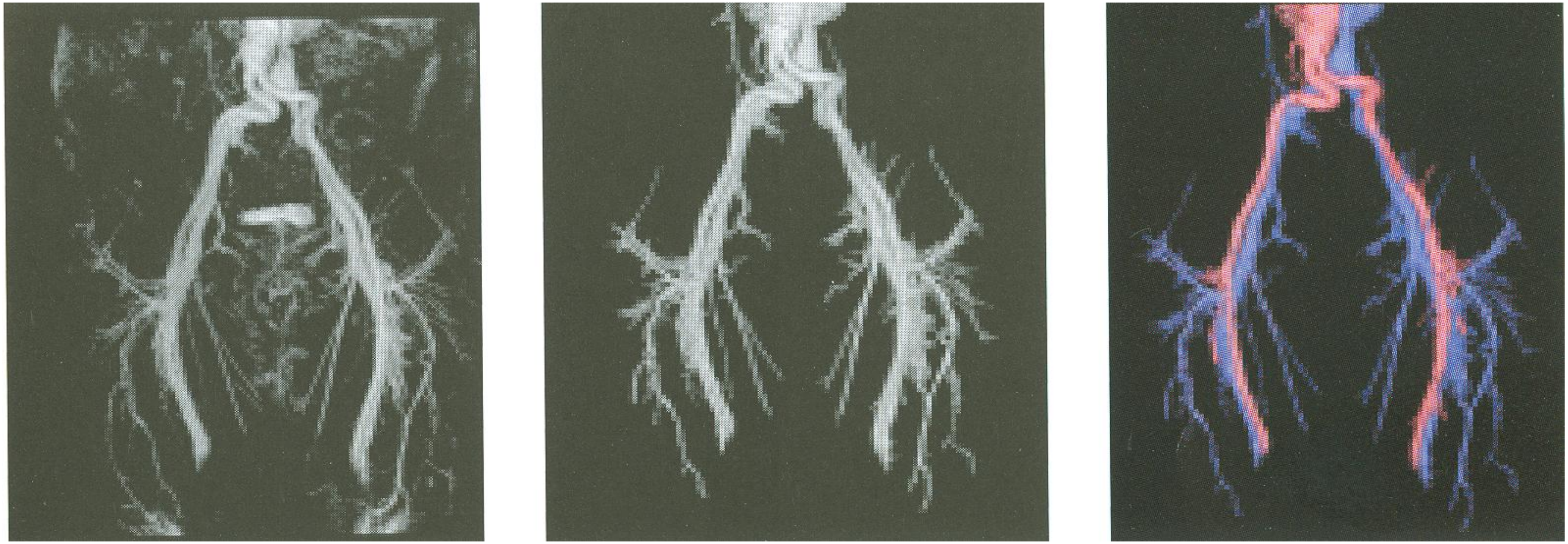
Iterative relative fuzzy connectedness



Due to weak boundary between objects O1 and O2, the costs of paths P1 and P2 may be too similar, or even equal.

It is suggested to first determine the core of each object (by relative fuzzy connectedness) and then, in repeated computation of connectedness, not allow paths from a seed to pass through the core of another seed/object.

An example



Segmentation of vascular trees. (a) MIP.
(b) Segmentation of the entire vascular tree by absolute fuzzy connectedness.
(c) Artery-vein separation using relative fuzzy connectedness. Multiple seeds are determined in an interactive way.

Algorithm for Computing Fuzzy Connectedness (Dijkstra-like)

Set all elements of FC to 0 except s which is set to 1 ;
Push s to Q ;
While Q is not empty do
    Remove a spel c from Q for which $FC(c)$ is maximal ;
    For each spel e such that $\mu_K(c,e) > 0$ do
        Set $fc = \min(FC(c), \mu_K(c,e))$;
        If $fc > FC(e)$ then
            Set $FC(e) = fc$;
            If e is already in Q then
                Update e in Q ;
            Else
                Push e to Q ;

Summary

Today, we talked about:

- Fuzzy sets (definition, properties, operations)
- Fuzzy reasoning
- Applications of fuzzy sets in image processing
  - Image segmentation
    - fuzzy thresholding
    - fuzzy region growing
    - fuzzy connectedness
  - Cluster analysis based on fuzzy c-means