Application of fuzzy set theory in image analysis

Computer Assisted Image Analysis II

natasa.sladoje@it.uu.se
Fuzzy systems – why

- Fuzzy systems and models are capable of representing inexact and inaccurate information.
- The qualifiers they can deal with are **linguistic variables**.
  
  Examples: a rotten apple, a bright image, a medium dark wall, a dark sky.
- Fuzzy systems can handle numerical data and linguistic variables simultaneously.
Crisp vs. fuzzy
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

\[ A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases} \]

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*

<table>
<thead>
<tr>
<th>Name</th>
<th>Height, cm</th>
<th>Crisp</th>
<th>Fuzzy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chris</td>
<td>208</td>
<td>1</td>
<td>1.00</td>
</tr>
<tr>
<td>Mark</td>
<td>205</td>
<td>1</td>
<td>1.00</td>
</tr>
<tr>
<td>John</td>
<td>198</td>
<td>1</td>
<td>0.98</td>
</tr>
<tr>
<td>Tom</td>
<td>181</td>
<td>1</td>
<td>0.82</td>
</tr>
<tr>
<td>David</td>
<td>179</td>
<td>0</td>
<td>0.78</td>
</tr>
<tr>
<td>Mike</td>
<td>172</td>
<td>0</td>
<td>0.24</td>
</tr>
<tr>
<td>Bob</td>
<td>167</td>
<td>0</td>
<td>0.15</td>
</tr>
<tr>
<td>Steven</td>
<td>158</td>
<td>0</td>
<td>0.06</td>
</tr>
<tr>
<td>Bill</td>
<td>155</td>
<td>0</td>
<td>0.01</td>
</tr>
<tr>
<td>Peter</td>
<td>152</td>
<td>0</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**Crisp:** A man is either tall, or not.

**Fuzzy:** The degree of membership to a set of tall men depends on the height.
Fuzzy membership function
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 $\leftrightarrow$ Define a membership function

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

- (a) $\pi$-function (crisp)
- (b) Trapezoidal function
- (c) Semi-trapezoidal
- (d) Triangular function
- (e) Gaussian
- (f) S-function
Membership functions - examples

\[ \mu_{\text{SMALL}}(x) \]

\[ \mu_{\text{MEDIUM}}(x) \]

\[ \mu_{\text{LARGE}}(x) \]
Fuzziness vs. Probability

Number 10 is not probably big!
...and number 2 is not probably not big.

Uncertainty is a consequence of non-sharp boundaries between the notions/objects, and not because of lack of information.
Terminology:
Support, core, $\alpha$-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 $A$:
  \[
  \text{Supp}(A) = \{x \in X \mid \mu_A(x) > 0\}
  \]

- The **core** of a fuzzy set $A$ is the (crisp) set of all elements of $X$ with membership to $A$ equal one:
  \[
  \text{Core}(A) = \{x \in X \mid \mu_A(x) = 1\}
  \]

- An **$\alpha$-cut** of a fuzzy set $A$ is a crisp set of all the elements in $X$ with membership to $A$ not smaller than $\alpha$:
  \[
  ^\alpha A = \{x \in X \mid \mu_A(x) \geq \alpha\}
  \]
Fuzzy set operations

We can define set operations for fuzzy set (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**  \( \text{not } A \)

\[
\mu_{A^c}(x) = 1 - \mu_A(x)
\]
An application of approximate reasoning
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
Fuzzy sets in image processing
Objects with fuzzy borders

- Image data are rarely of perfect quality.
- Fuzziness is intrinsic property of images.

- 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.
Fuzzy thresholding

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.

Apply a sequence of S-functions to the image intensities. Threshold with the one providing minimal fuzziness. This may also result in hard (crisp) thresholding.
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)
Fuzzy c-means clustering
Fuzzy c-means clustering

Objective: to partition a collection of numerical data into a series of overlapping clusters. The degrees of belongingness are interpreted as fuzzy membership values of the data to the clusters.

The fuzzy c-means algorithm (FCM) iteratively minimizes an objective function, starting from a reasonable initialization.
Fuzzy c-means clustering
Extends K-means

- **K-means** algorithm minimizes the sum of within-cluster variances for the K observed clusters:

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

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

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}_{n \times K}$$

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.
- The algorithm iteratively updates cluster centers as the means of the clusters created in a previous iteration, recomputes 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 $n \times c$ 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**

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

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

\[
U = \begin{bmatrix}
0.8 & 0.1 & 0.1 \\
0.2 & 0.5 & 0.3 \\
0.4 & 0.3 & 0.3 \\
0.1 & 0.1 & 0.8 \\
\end{bmatrix}_{nxc}
\]

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 clustering

• The fuzzy c-means algorithm iteratively minimizes the objective function, starting from a reasonable initialization.

• The objective function is of the form:

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

where \( u_{ki} \) is an element of a \( n \times c \) matrix \( U \) which represents a fuzzy c-partition of the data set \( X = \{x_1, x_2, \ldots, x_n\} \), \( v_i \) is the cluster center of the class \( i \), \( 1 \leq i \leq c \), \( m \), \( 1 \leq m < \infty \) is the parameter controlling fuzziness of the partition, and \( d_{ki}^2 = \|x_k - v_i\|^2 \) for an inner product norm metric \( \| \| \).
Fuzzy c-means algorithm: the basic steps

Choose

- \( c \) – the number of clusters,
- \( m \) - the weighting exponent (between 1.5 and 2.5),
- the 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, starting from initial cluster centres and the initial partition.

- Cluster center is the mean of all cluster points, weighted by their membership to the cluster.
- Membership of a point to a cluster is inversely related to the distance of the point to the cluster center, computed in the previous iteration.

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
Fuzzy connectedness

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 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.

\[
\mu(c, d) = \max_{P_{c,d} \in P_{cd}} \mu_{P_{c,d}}.
\]

\[
\mu_{P_{c,d}} = \min_{j=1, \ldots, n-1} \mu_K(c_j, c_{j+1}).
\]
Fuzzy connectedness
The strength of the weakest link of the strongest path
Fuzzy connectedness

Fuzzy adjacency determines spatial closeness of the image elements.

- **hard** - e.g., 4- or 8-adjacency in binary 2D images,
- fuzzy adjacency usually decreasing with decreasing distance between the pixels.

Fuzzy affinity

Homogeneity based component: The degree of local hanging-togetherness due to the similarity in intensity.
Object-feature based component: The degree of local hanging-togetherness with respect to some given feature, e.g., intensity distribution.
Fuzzy connectedness

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

- Fuzzy adjacency
- Fuzzy affinity-homogeneity based component
- Fuzzy affinity-object-feature based component

Spatial intensity

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

Examples:

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

Breast density as measured from the volume of dense tissue in the breast is considered to indicate a risk factor for breast cancer.

A digitized X-ray mammogram, the fuzzy connectivity scene of a dense (fibro glandular) region (as opposed to fatty regions), and the segmented binary region.
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.
An example - MR Angiography

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.
Summary

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