GNR602
Advanced Methods in Satellite Image Processing

Instructor: Prof. B. Krishna Mohan

CSRE, IIT Bombay

bkmohan@csre.iitb.ac.in

Slot 13

Lectures 09-11 Fuzzy Sets and Applications

Contents of the Lectures

- Introduction to Fuzzy Sets
- Operations on Fuzzy Sets
- Fuzzification and Defuzzification
- Measures of Fuzziness
- Applications to Image Processing

Resources

- Any standard text book for fundamentals
- See the chapter by Haubecker and Tizhoosh “Fuzzy Image Processing” in the book **COMPUTER VISION AND APPLICATIONS** edited by Bernd Jähne, Horst Haussecker IITB Central Library accession no. 211301
- Chapter on Fuzzy Sets in Tso and Mather’s book on ***Classification Methods for Remotely Sensed Data, CRC Press, 2009 (2nd edition)***

DEFINITIONS

A. Definitions

1. Sets

Classical sets (also referred to as crisp sets) – either an element belongs to the set or it does not. For example, for the set of integers, either an integer is even or it is not (it is odd). Likewise, either you are allotted hostel 11/12 or you are not.

You can relate the element under consideration to be 100% belonging to the set, or 0%.

$a \in A => m_A(a) = 1$

$a \notin A => m_A(a) = 0$

Classical sets

Classical sets are also called *crisp* (sets).

Lists: A = {apple, banana, grape, papaya}

$A = \{a_1, a_2, a_3\}$

$A = \{2, 4, 6, 8, \ldots\}$

Formulas: A = {x | x is an even natural number}

A = {x | x = 2n, n is a natural number}

Membership or characteristic function

$$X_A(x) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \notin A \end{cases}$$

Fuzziness

- You are approaching a red light and must advise a driving student when to apply the brakes. What would you say:
- “Start slowing down *100 metres* from the zebra-crossing”?
- “Apply the brakes *soon*”?

Everyday language is one example of the ways vagueness is used and propagated.

Imprecision in data and information

statistical (e.g., the outcome of a coin toss is a matter of chance) or *non-statistical* (e.g., “apply the brakes soon”).

This latter type of uncertainty is called ***fuzziness*****.**

Fuzzy Sets and Membership Fuzzy Sets and Membership Functions

We all assimilate and use fuzzy data, vague rules, and imprecise information.

Statistical models deal with random events and outcomes; fuzzy models attempt to capture and quantify nonrandom imprecision.

Fuzzy Sets and Membership Functions

Fuzzy sets contain objects that satisfy *imprecise properties* to varying degrees, for example, the "set" of numbers F that are "close to 7."

In the case of fuzzy sets, the membership function, $m_F(r)$, maps numbers into the entire unit interval [0,1]. The value $m_F(r)$ is called the *grade of membership* of r in F.

Fuzzy sets correspond to continuously-valued logic:

- all shades of gray between black (= 1) and white (= 0)

Fuzzy Sets and Membership Functions

Because the property "close to 7" is fuzzy, there is not a unique membership function for *F*. Rather, it is left to the modeler to decide, based on the potential application and properties desired for *F*, what $m_F(r)$ should be like.

The *membership function* is the basic idea in fuzzy set theory; its values measure degrees to which objects satisfy imprecisely defined properties.

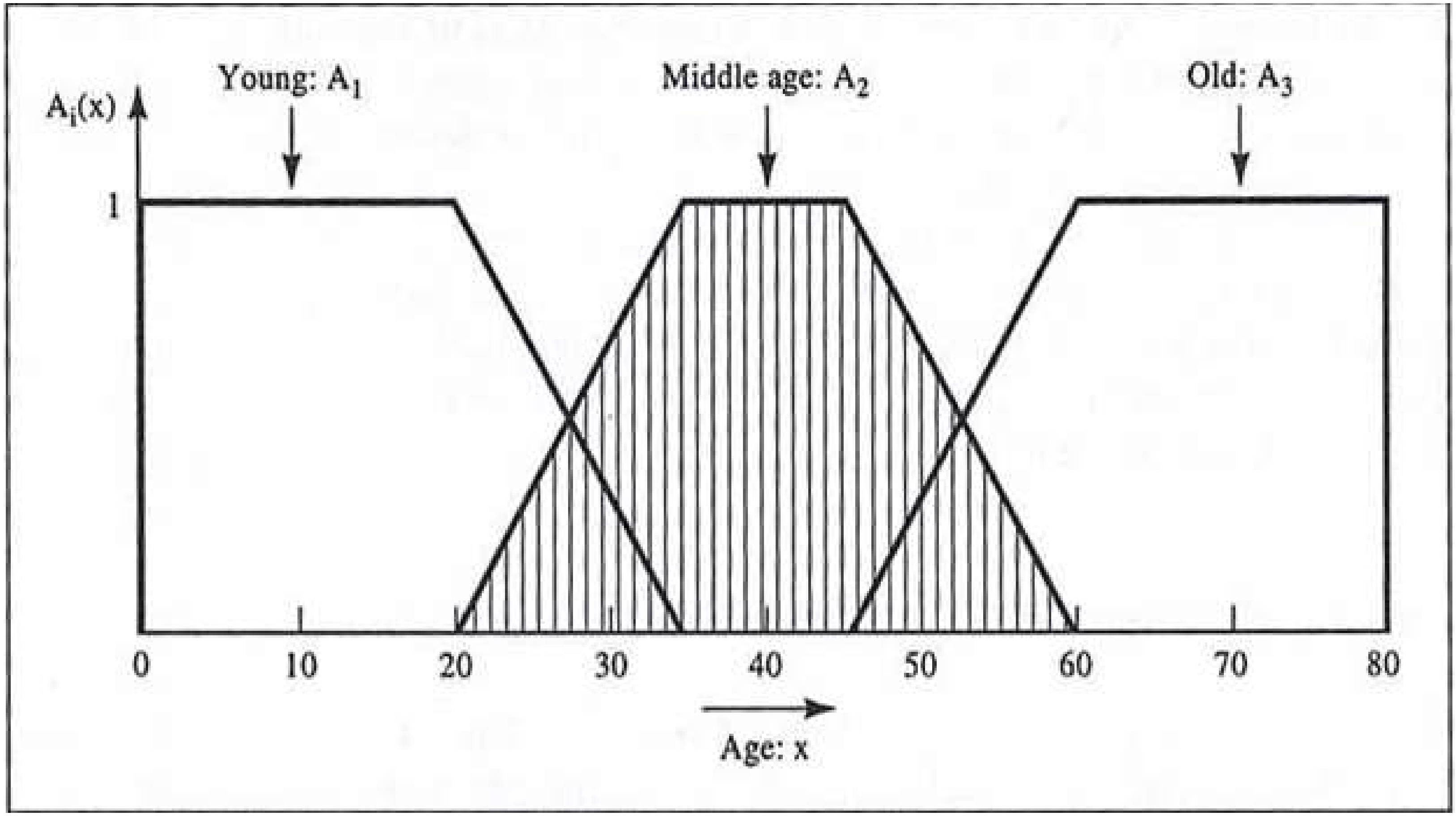
Fuzzy memberships represent similarities of objects to imprecisely defined properties.

Membership values determine how much fuzziness a fuzzy set contains.

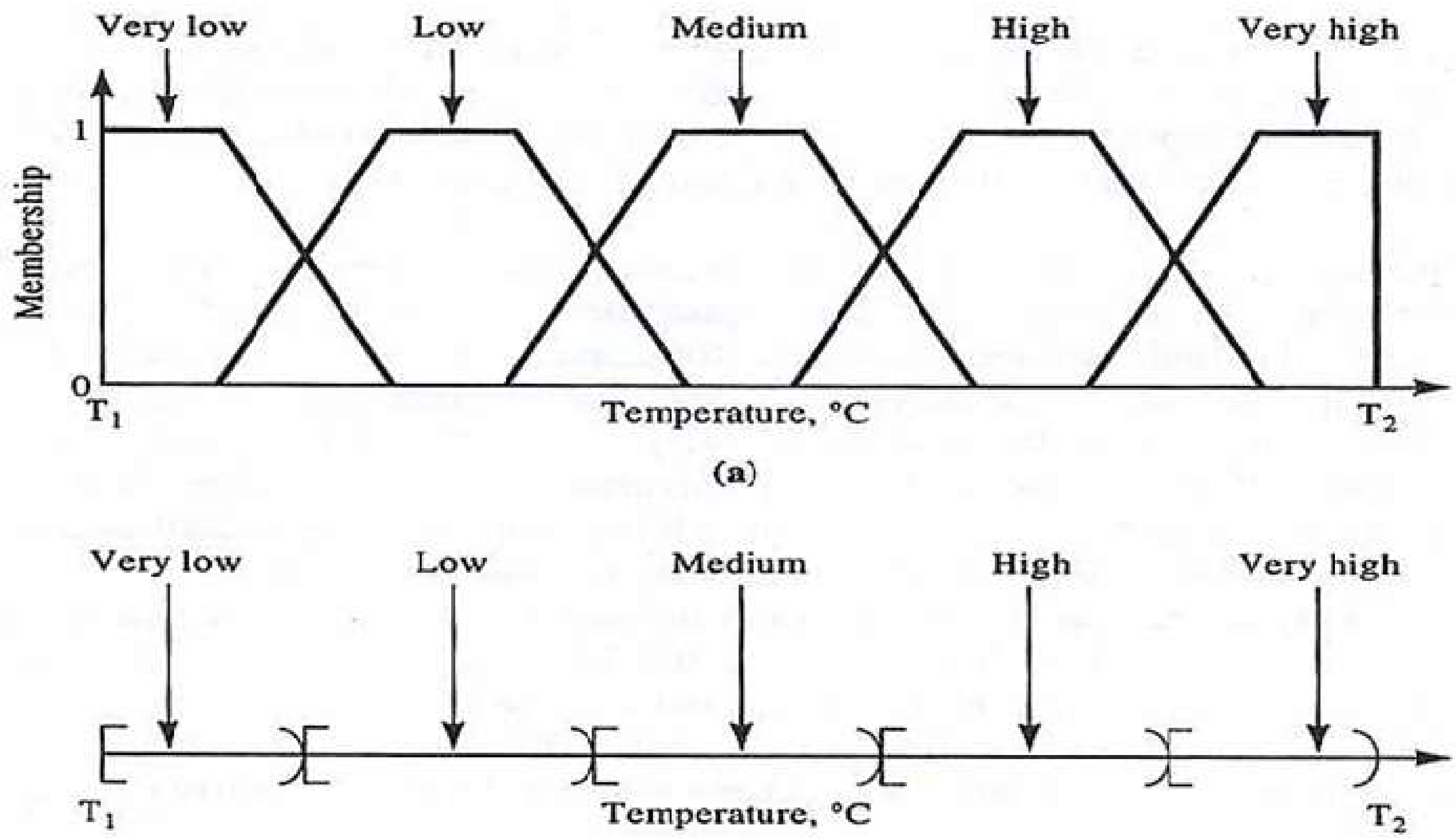
Notion of Membership

Fuzzy sets – admit gradation such as all tones between black and white. A fuzzy set has a graphical description that expresses how the transition from one to another takes place. This graphical description is called a membership function.

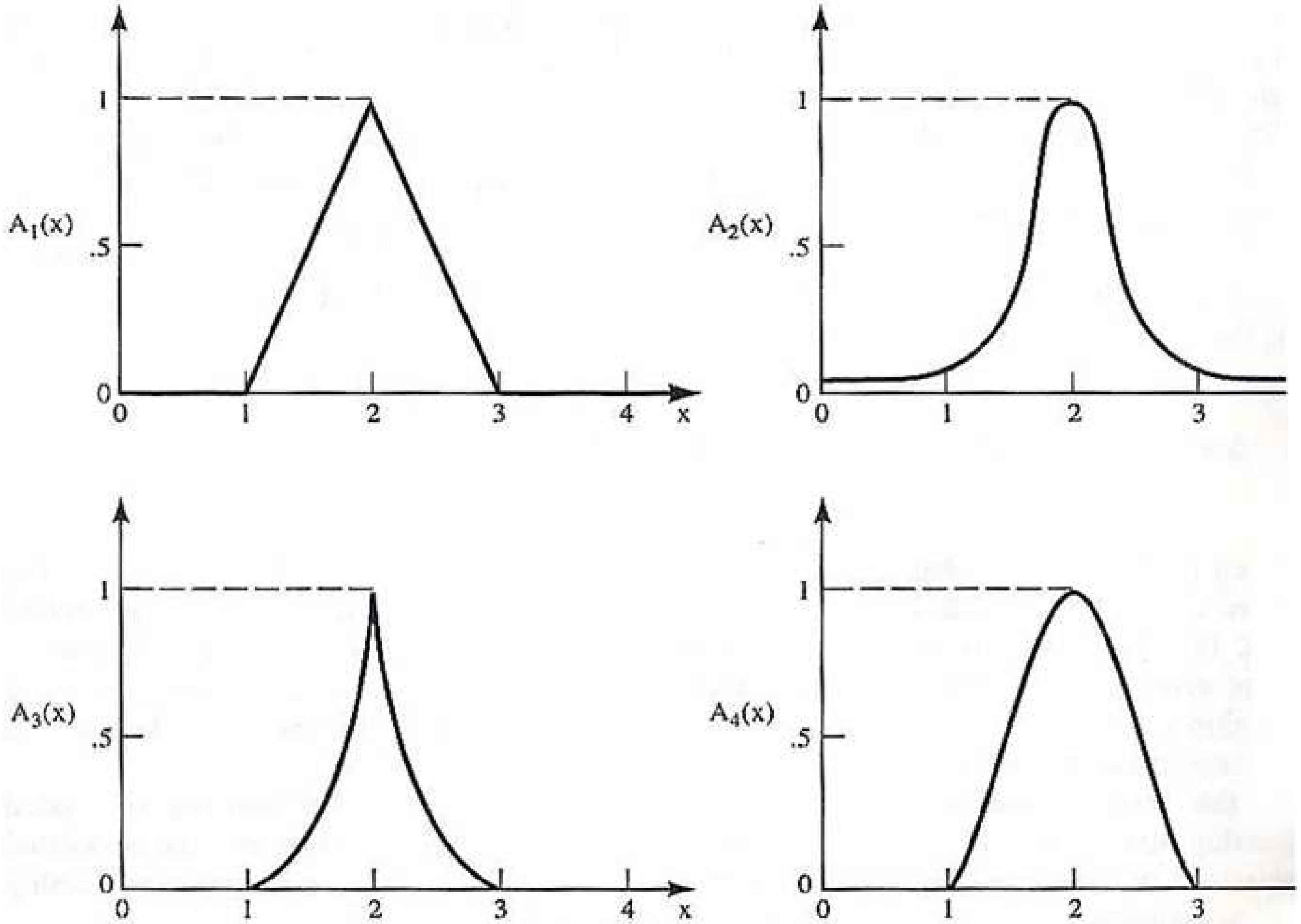
Sample Membership Function



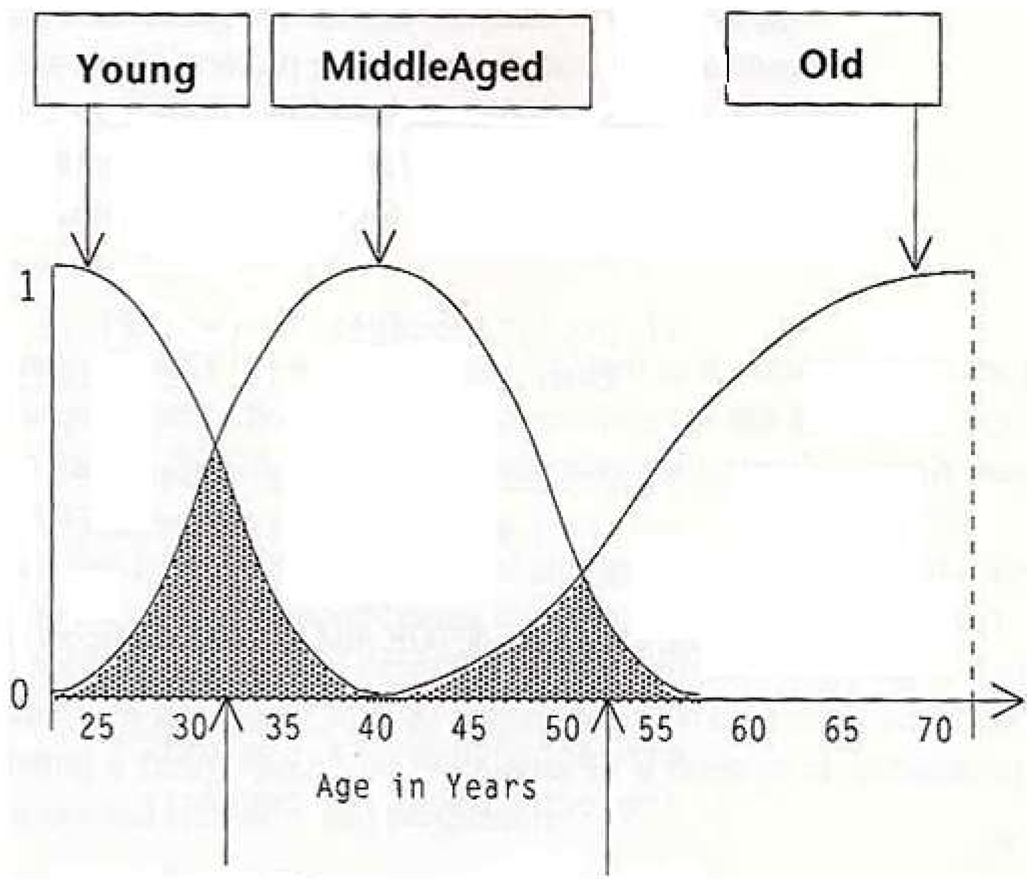
More Gradations



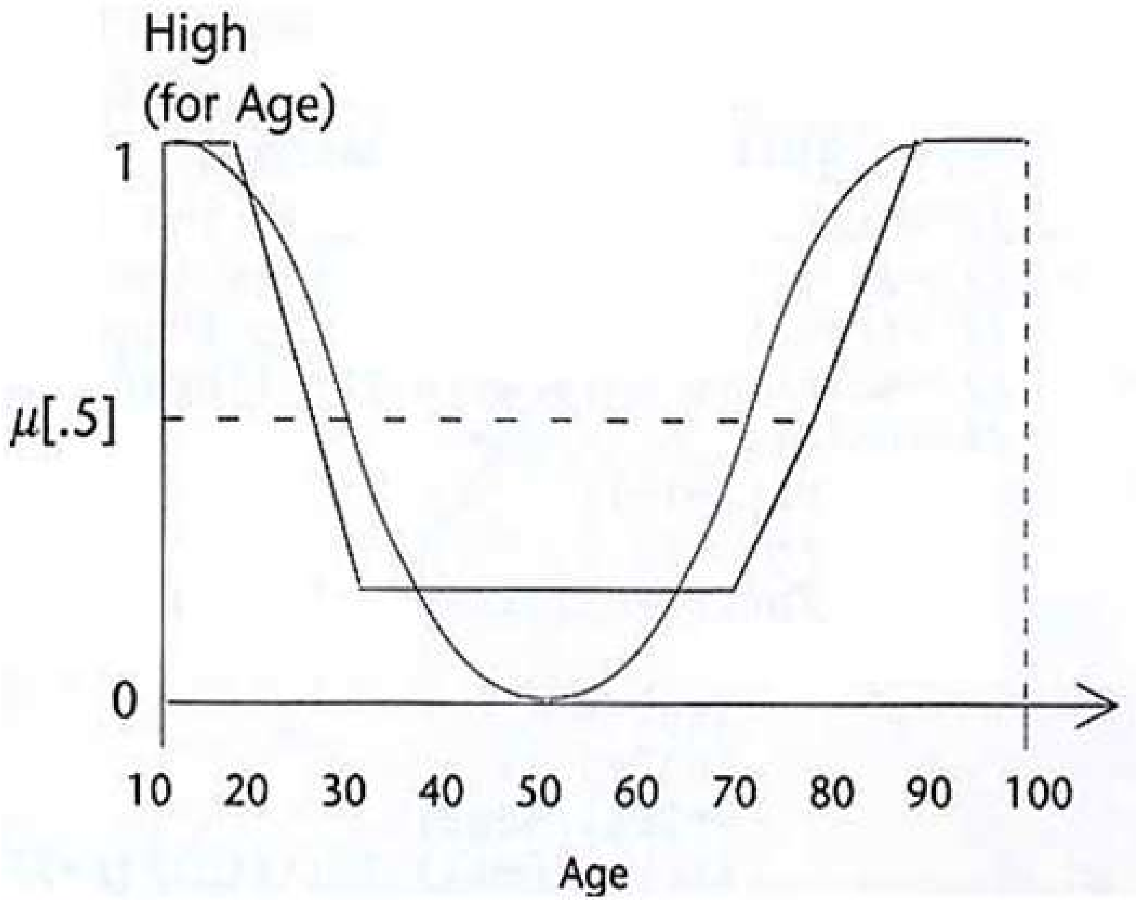
More Examples



Ambiguous Regions

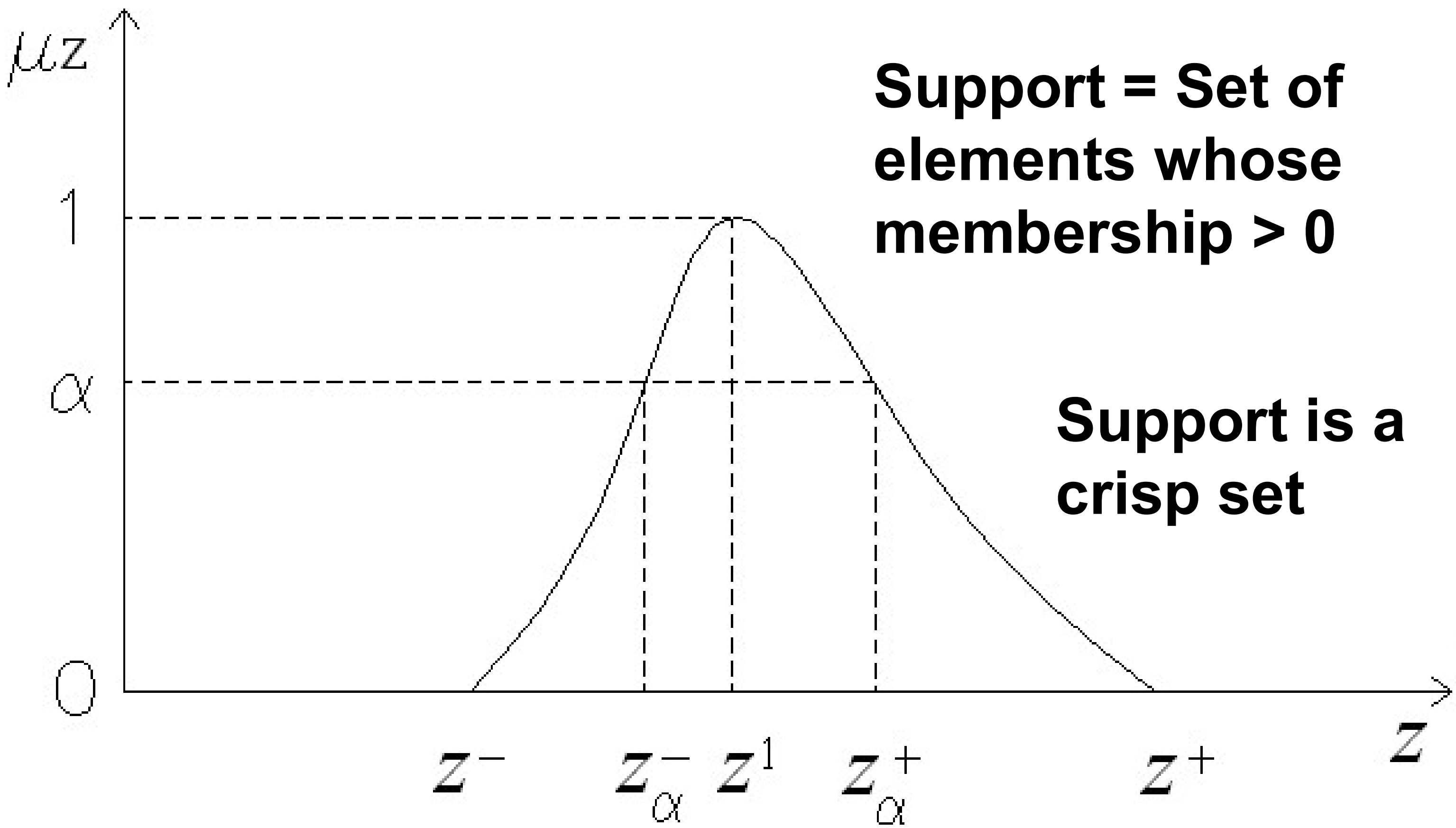


Variations in Membership



Operations on Fuzzy Sets

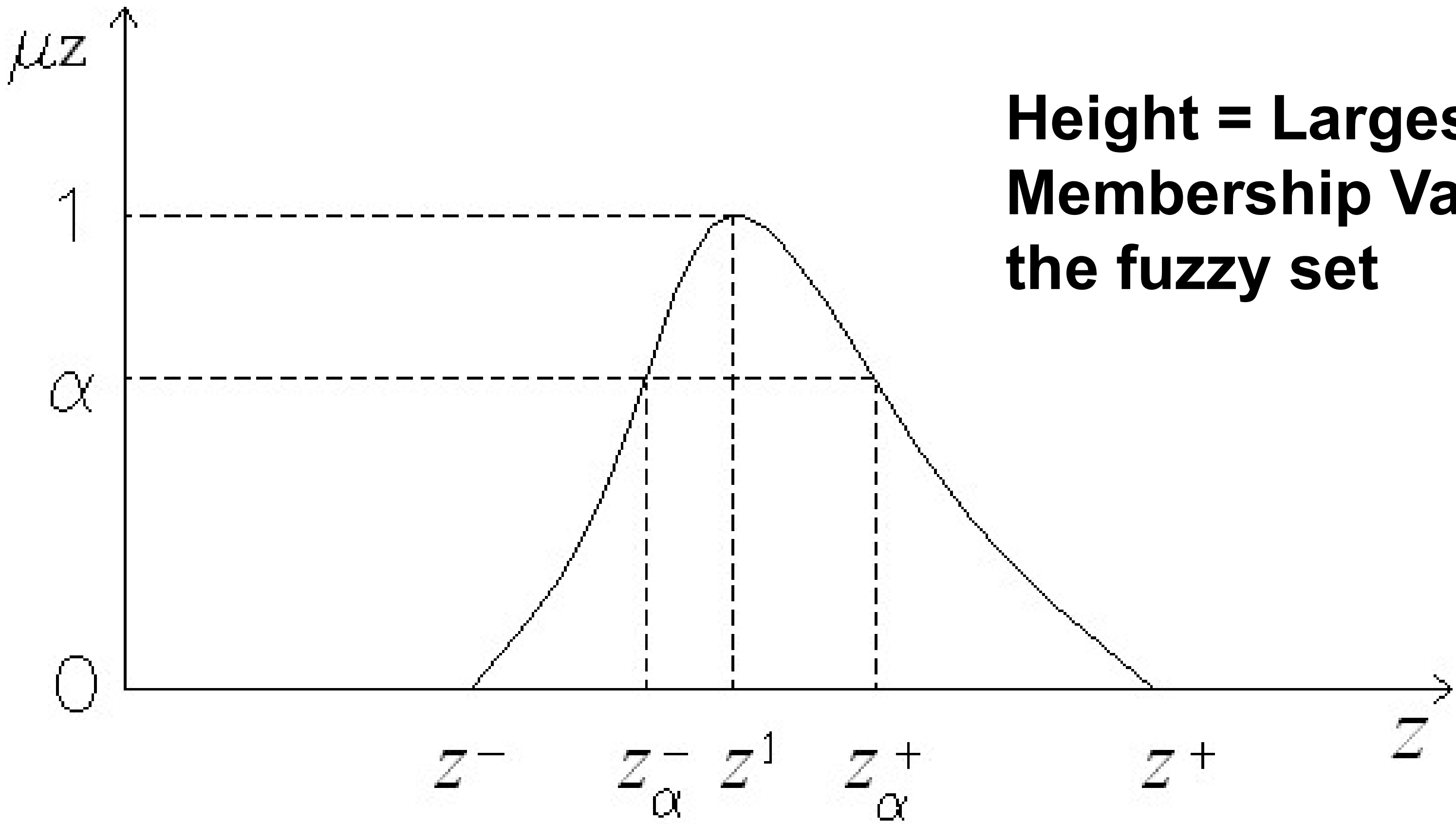
Support of a Fuzzy Set



Support = Set of elements whose membership > 0

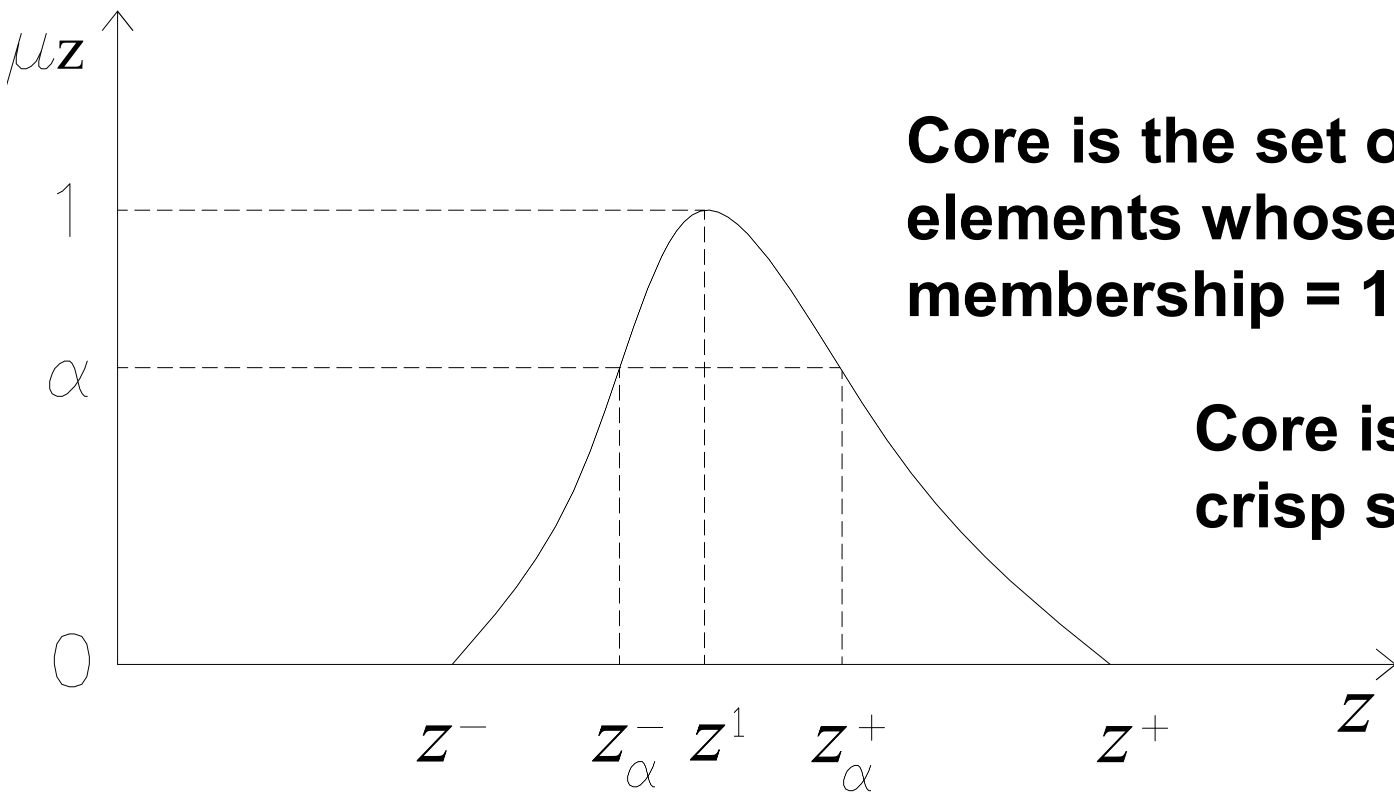
Support is a crisp set

Height of a Fuzzy Set



Height = Largest Membership Value in the fuzzy set

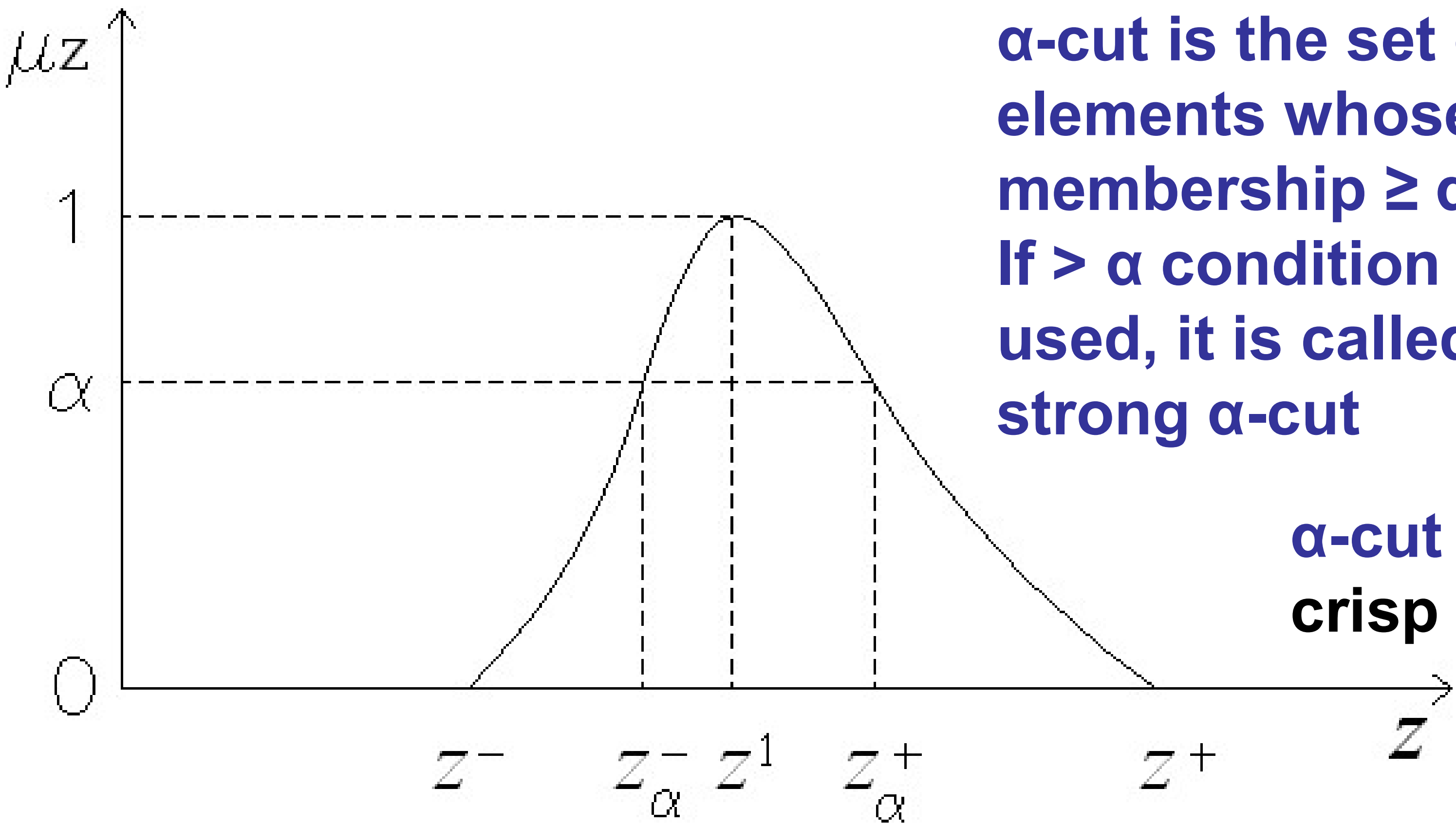
Core of a Fuzzy Set



Core is the set of all elements whose membership = 1

Core is a crisp set

α-cut



α-cut is the set of all elements whose membership ≥ α
If > α condition is used, it is called strong α-cut

α-cut is a crisp set

Set Theoretic Operations on fuzzy sets

- Since fuzzy sets are introduced in 1965, the properties of the operations on fuzzy sets are to be defined carefully so that they are meaningful
- The desirable properties lead to a unique definition of the operations such as union, intersection and complement

Universal Set

- Let X be a universal set, consisting of elements x_1, $x_{2,}$, …
- All these elements x_i are members of X
- Any fuzzy set A is a subset of X, since the membership of any x_i in A is partial, in the interval [0 1].

Operations on Fuzzy Sets

- Complement
- If A is a complement of B ($A = \bar{B}$), $\mu_A(x) = 1 - \mu_B(x)$
- Properties
  - depends only on $\mu_B(x)$
  - should satisfy the rules of ordinary complementation: if $\mu_A(x) = 1$ then its complement=0 and vice versa
  - monotonous and strictly decreasing with increasing membership of a function
  - Involutive – complement(complement(membership))=membership

Properties of Complement

- Important condition

$$\mu_B(x_1) - \mu_B(x_2) = \mu_{\overline{B}}(x_2) - \mu_{\overline{B}}(x_1)$$

- This results in the definition

$$Complement(\mu) = 1 - \mu$$

Containment

- A fuzzy set A is a subset of B if
- $\mu_A(x) \leq \mu_B(x)$ for all x in X
- A special case is B = X, the universal set, which is stated earlier.
- Obviously, A = B when $\mu_A(x) \leq \mu_B(x)$ and
- $\mu_B(x) \leq \mu_A(x)$, i.e.,
- $\mu_A(x) = \mu_B(x)$ for all x in X

Union

- The union of two fuzzy sets A and B is given by
- C = A U_f B such that
- $\mu_c(x) = V(\mu_A(x), \mu_B(x))$
- V is the symbol for supremum
- The membership value of x in a compound fuzzy set depends only on its constituent values and not on anything else
- Like conventional sets, the individual sets are contained in their union

Intersection

- The intersection of two fuzzy sets A and B is given by
- $C = A \cap_f B$ such that
- $\mu_c(x) = \wedge(\mu_A(x), \mu_B(x))$
- $\wedge$ is the symbol for infimum (or minimum)
- The membership value of x in a compound fuzzy set depends only on its constituent values and not on anything else
- Like conventional sets, the intersection is contained in the individual sets

Properties of union and intersection

- V and $\wedge$ are monotonically increasing
- For instance, if
- $\mu_A(x_1) = \mu_B(x_1) > \mu_A(x_2) = \mu_B(x_2)$, then
- $\mu_{AUB}(x_1) > \mu_{AUB}(x_2)$
- $\mu_{A\cap B}(x_1) > \mu_{A\cap B}(x_2)$
- Complete membership in A and B implies complete membership in A∩B
- Complete absence of membership in A and B implies complete absence of membership in AUB.

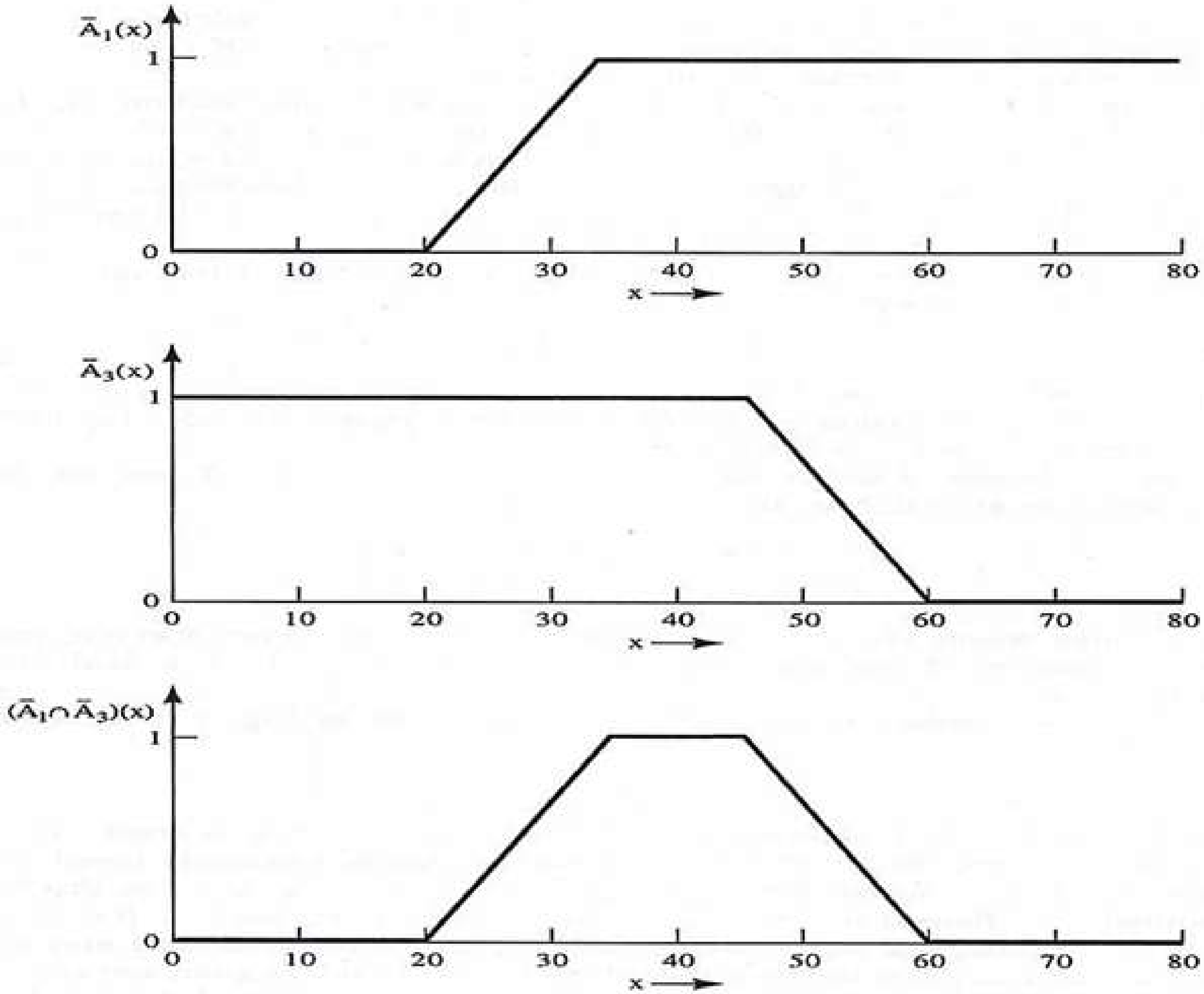
Departure from Ordinary Set Theory

- Law of Contradiction and Excluded Middle
- In ordinary set theory,
- $A \cap A^c = \varnothing$, $A \cup A^c = X$
- In fuzzy set theory, this is no longer true:
- A fuzzy set theory has no definite boundary

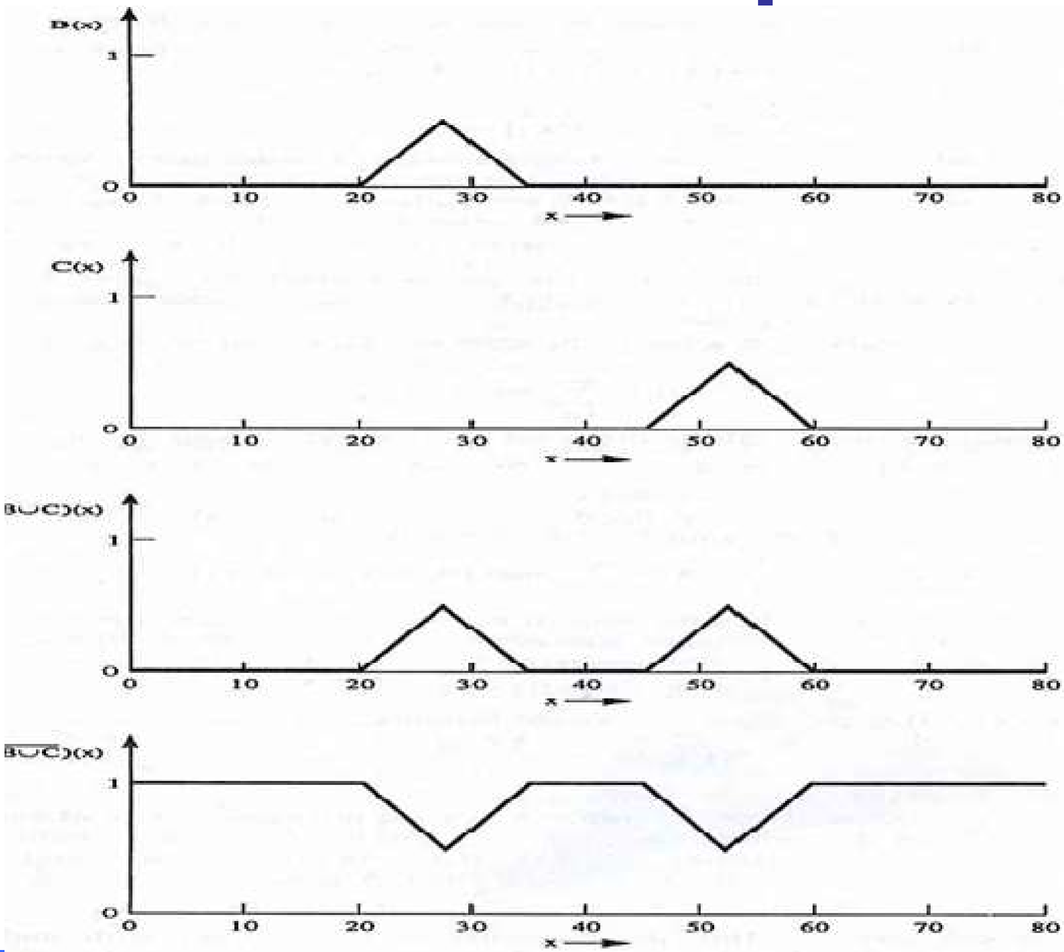
$$\min\left(\mu_A(x), \mu_{A^c}(x)\right) \leq 0.5 \text{ i.e., } \mu_{A \cap A^c} \leq 0.5$$

$$\max\left(\mu_A(x), \mu_{A^c}(x)\right) \geq 0.5 \text{ i.e., } \mu_{A \cup A^c} \geq 0.5$$

Fuzzy Intersection



Union and Complement



Other Properties

- **Product of A and B**
- $\mu_{AB}(x) = \mu_A(x).\ \mu_B(x)$
- **α^{th} power of A**
- $\mu_A^{\alpha}(x) = (\mu_A(x))^{\alpha}$
- **Probabilistic sum of A and B**

$$\mu_{A_P+B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x)$$

- **Bounded sum of A and B**

$$\mu_{A_B+B}(x) = \Lambda(1, (\mu_A(x) + \mu_B(x))$$

- **Bounded product of A and B**

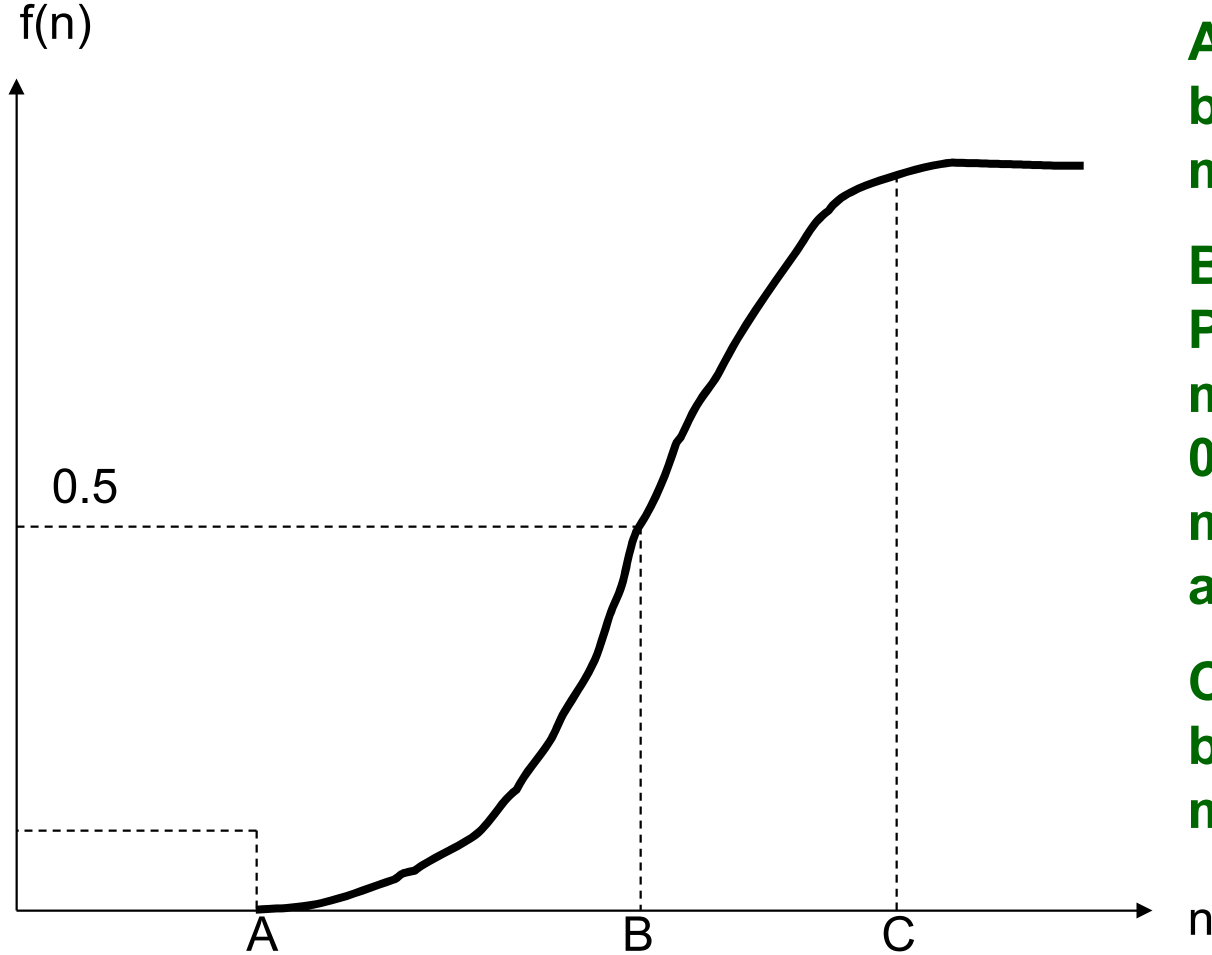
$$\mu_{A_{BP}+B}(x) = \mathrm{V}(0, (\mu_A(x) + \mu_B(x) - 1)$$

Creating Fuzzy Sets

Fuzzification Procedures

- A dataset can be represented in a fuzzy set theoretic environment by defining a fuzzy set that is a subset of the universal set, and the mapping of various elements with varying degrees of membership into the fuzzy set

S-function



A – lower cutoff below which membership is 0

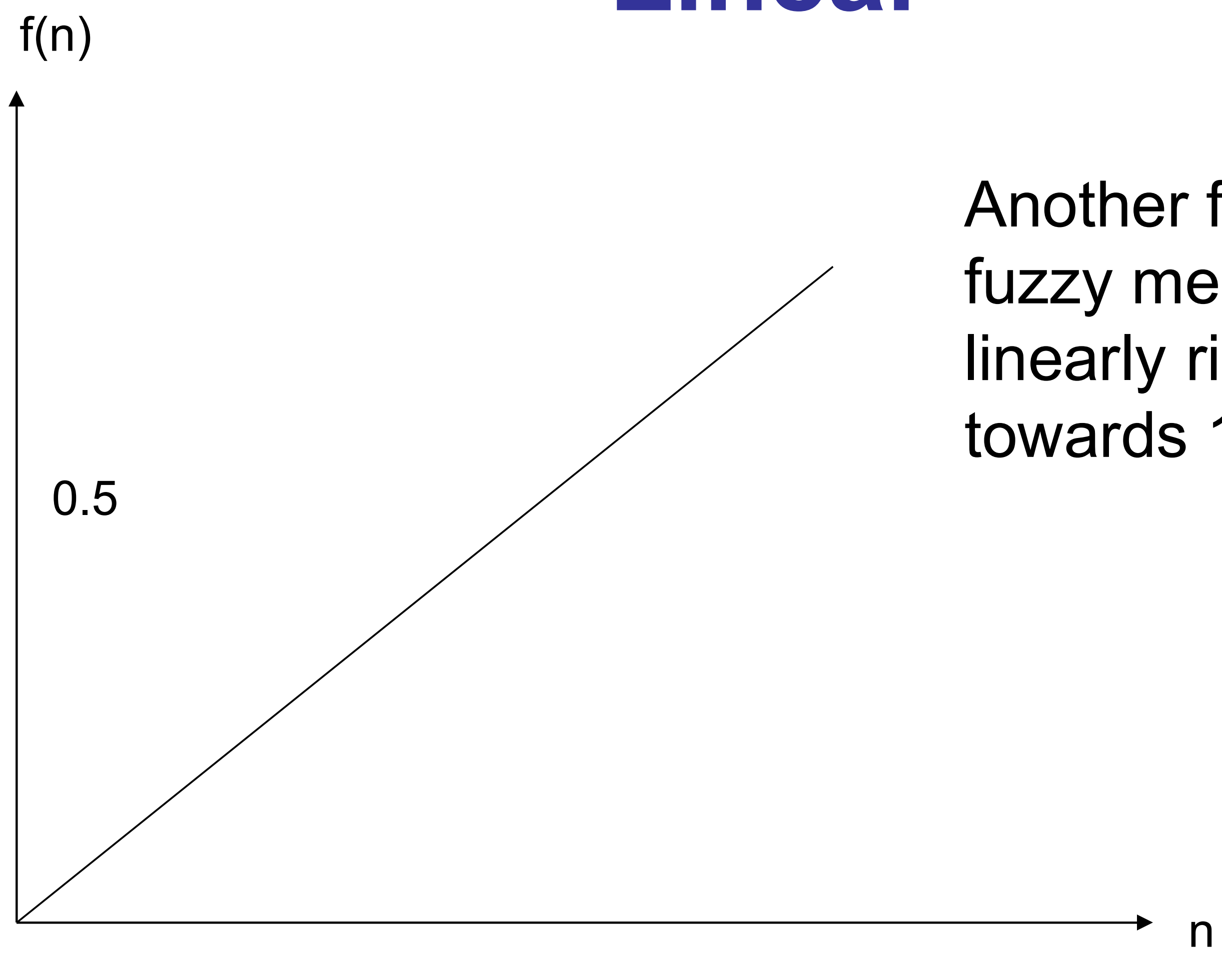
B – Crossover Point where membership is 0.5 (point of maximum ambiguity)

C – upper cutoff beyond which membership is 1

Properties of S-Function

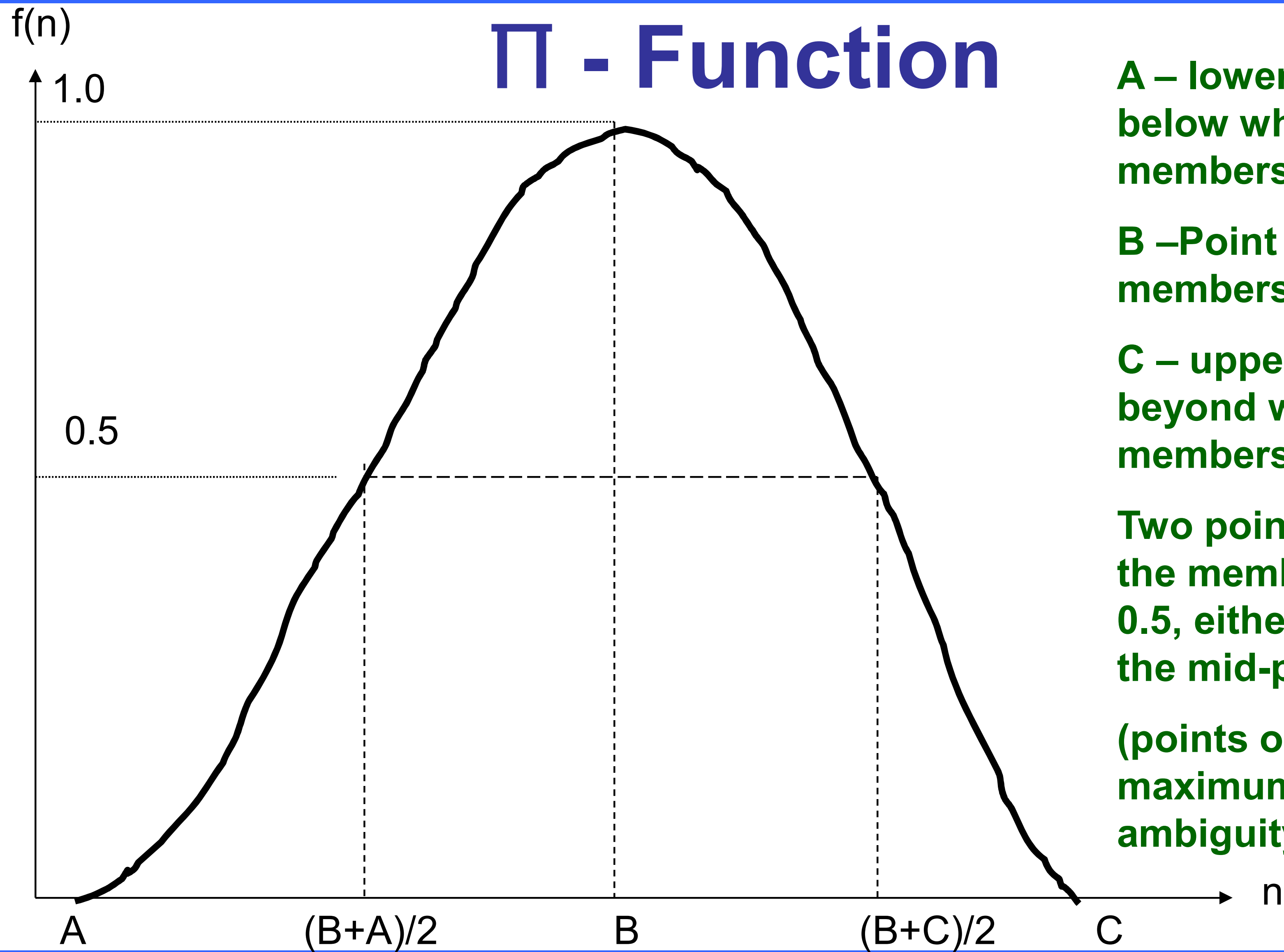
- S-functions are well suited to represent the gray level distribution.
- Membership assignment μ_n = S(n;A,B,C)
- **$\mu_n = 0$ if $n \leq A$**
- **$\mu_n = 2\{(n-A)^2 / (C-A)^2\}$ if $A \leq n \leq B$**
- **$\mu_n = 1 - 2\{(C-n)^2/(C-A)^2\}$ if $B \leq n \leq C$**
- **$\mu_n = 1$ if $n \geq C$**
- **B = (C+A)/2**
- It is possible to have asymmetric S-functions, such that the membership increases at different rates between 0 and 0.5, and between 0.5 and 1.

Linear



Another function for fuzzy memberships, linearly rising from 0 towards 1.

Π - Function



A – lower cutoff below which membership is 0

B –Point where membership is 1.0

C – upper cutoff beyond which membership is 0

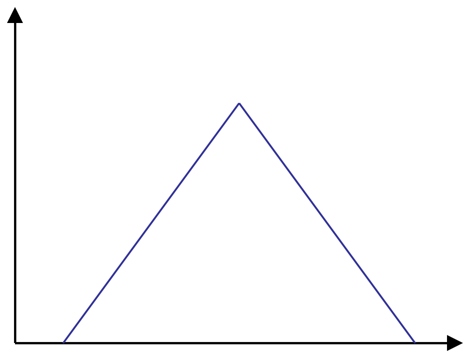
Two points where the membership is 0.5, either side of the mid-point B

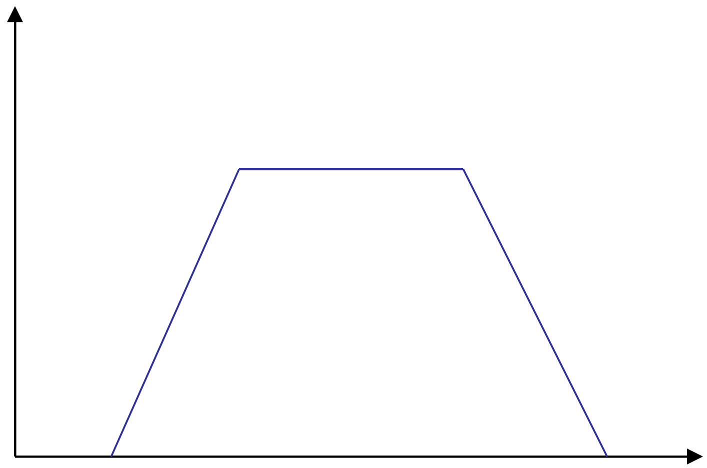
(points of maximum ambiguity)

Properties of Π - Function

- $\mu_n = 1$ if $n = B$
- $\mu_n = 0$ for $n \le A$, or $n \ge C$
- Π – Function can also be represented as a mixture of two S-functions
- $\Pi(n;A,B,C) = S(n;A,(B-A)/2;B)$ if $B \ge n$
- $\Pi(n;A,B,C) = 1-S(n;B,(B+C)/2;C)$ if $n \ge B$
- Π – Function is used in classification applications. The point of membership = 1 corresponds to the feature mean, and around the mean the membership tends to 0 on either side

Other Functions





A library of membership functions allows the user to choose a membership function that suits the application.

Comparing Fuzzy Sets

- Given two fuzzy sets created out of the same input data, by different mappings, how can the two fuzzy sets be compared?
- One way is to see which mapping leaves more elements getting mapped close to 0.5 membership and which one maps them closer to 0 or 1.
- When a fuzzy set has more elements with membership close to 0.5, there is more ambiguity regarding their belongingness to the fuzzy set.

Measures of Fuzziness - How fuzzy is a fuzzy set?

- By Hamming distance:
- The membership of an element into a set A is maximally uncertain if its membership is 0.5. Hence the difference of fuzzy membership and its complement for such an element is 0.
- For unambiguous membership, (membership = 0 or 1), the Hamming distance $f(\mu_A) = 0$
- $f(\mu_A) = \Sigma_x \{1 - | \mu_A(x) - (1 - \mu_A(x))|\}$
- f(.) is 0 for crisp sets, and maximum for a fuzzy set with all memberships = 0.5

How fuzzy is a fuzzy set? – Shannon Entropy

- Shannon Entropy
- $f_1(\mu_A) = -\Sigma_x \mu_A(x) . \ln(\mu_A(x))$
- $f_1(.)$ is maximum when all memberships are equal to 0.5
- $f_1(.)$ is minimum when each membership is equal to 0 or 1

How fuzzy is a fuzzy set? – De Luca and Termini Entropy

- De Luca and Termini's entropy
- based on Shannon function
- $f_2(\mu_A) = -\Sigma_x \mu_A(x) . \ln(\mu_A(x)) -$
- $-\Sigma_x (1-\mu_A(x)) . \ln((1-\mu_A(x)))$
- $f_2(.)$ is based on the sum of ambiguities associated with the memberships and their complements

How fuzzy is a fuzzy set? – Implications for data classification

- Issues in a classification situation
- When a pixel is classified, a classifier may be assessed in terms of the amount of fuzziness associated with the pixel’s memberships into various classes
- A pixel’s membership should be high ONLY for one class and small for all other classes

Identification of mixed pixels

- This issue is relevant when classifying pixels to various classes, where a pixel may be containing two or more classes
- In such a case, according to neighbors, a mixed pixel may have sizeable membership in more than one class

Fuzzification of Gray level Images

- In a crisp set representation of an image, we can identify two levels – white and black
- A gray level that is neither black nor white is black to some extent, and white to some extent
- In other words, a gray level has a fuzzy membership into the set of white images, 0 if the gray level is 0, and 1 if the gray level is 255. Intermediate levels have progressively increasing membership

Gray levels

- It is also possible to define a black image as a fuzzy set, with gray level 0 having a membership of 1, and gray level 255 a membership of 0. Intermediate levels decrease progressively from 1 towards 0 as the gray levels increase

Membership into set of bright pixels

- The image can be treated as a set of bright pixels
- Membership of each pixel will be based on its gray level
- Pixels with low gray levels have very small or no membership into this set
- Membership increases with gray level, according to, say, the S-function

Fuzzy Image Processing

Operations for Image Enhancement

- Contrast Intensification – generates a new fuzzy set A' from a given fuzzy set A.
- A' = T(A)
- $T_1(\mu_n) = 2.\ \mu_n^2,\ 0 \leq \mu_n \leq 0.5$
- $T_2(\mu_n) = 1 - 2.[1 - \mu_n]^2,\ 0.5 \leq \mu_n \leq 1$
- Successive application of T_1 and T_2 tend to push the gray levels towards 0 and 1, leading to a 2-level image from a multilevel image

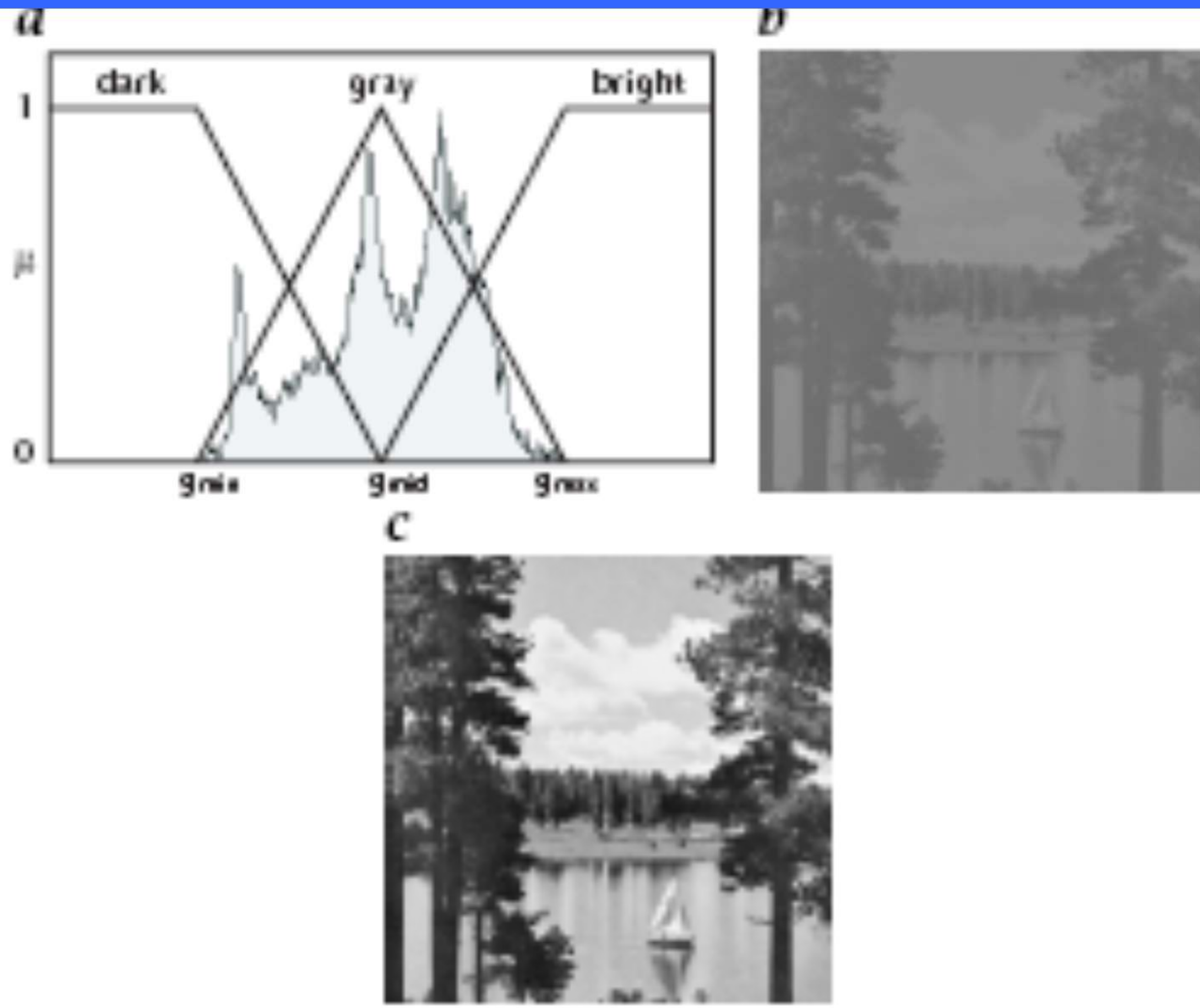
Example

Used with permission of Prof. H. Tizhoosh

Fuzzy Rules

- If pixel is dark, make it darker
- If the pixel is bright, make it brighter
- If the pixel is gray, reduce ambiguity by pushing it towards bright or dark

Example



Reducing Contrast

- The reverse of contrast intensification is contrast dilution, where the membership difference between pixels is reduced.
- DIL (μ_n)= Sqrt(μ_n)
- Consider two memberships 0.8 and 0.4, where the difference is 0.4
- Intensification increases the difference between the membership: $0.8^2 = 0.64$, $0.4^2 = 0.16$
- Dilution decreases the difference between the membership: sqrt(0.8) ~ 0.9, sqrt(0.4)~0.64
- Successive application of Dilution effectively diminishes the contrast in the image, but without considering the neighboring pixels

Membership into set of edge pixels

- The membership is often computed over a set of neighboring pixels; high differences in pixel gray levels leads to high membership into set of edge pixels. One formula suggested:

$$\mu_e = 1 - \left[1 + \frac{1}{g_{\max}} \sum_{i=0}^{8} \| g_0 - g_i \| \right]^{-1}$$

- g_0 is the central pixel intensity; g_{max} is the local maximum intensity in the neighborhood

Fuzzy Edge Detection

- One approach considers computing the intensity difference between the central pixel of a 3x3 neighborhood and two neighbors at a time.
- Using the membership assignment mentioned before, we compute the edge membership of each triplet μ_i, i=1,…,4
- Overall edge membership $\mu_e = \max \{\mu_i\}$ i=1,…,4
- No edge is marked by $\mu_n = 1 - \mu_e$

Example

a

INPUT IMAGE

b

FUZZY EDGE MAP

Used with permission of Prof. H.Tizhoosh

Clustering

- Given a set of N elements, the K-means algorithm distributes them into K clusters such that

$$\sum_{i=1}^{K} N_i = N$$

- Each cluster is represented by its mean m_i, i=1,2, …, K

Algorithm

- The K-means algorithm assumes that the value of K is correct, and that each cluster is represented by its mean m_i.
- In order to determine the means, we need to have *a priori* information pertaining to the clusters, or starting from a random assignment of means that are refined by some criteria.

Algorithm

- Suppose that the K clusters are initialized by randomly assigned means m_i. In general, each mean is a vector in case of mutidimensional data.
- Each data element x_i, i=1,…,N is now compared with the randomly initialized means. The method is based on some user chosen distance measure, such as the Euclidean distance between the feature vector associated with the data element and the mean vector.

Algorithm contd.

- Let us define

$$d_k = \sum_{j=1}^{L} (\mathbf{x}_{\mathbf{ij}} - \mathbf{m}_{\mathrm{kj}})^2, \ \mathrm{k}=1,2, \ldots , \mathrm{K}, \ \mathrm{i}=1,2, \ldots, \mathrm{N}$$

- Then the data element $\mathbf{x_i}$ is assigned to class P if $d_p \leq d_j$, j=1,2,…,K

Algorithm contd.

- All the N elements are assigned to the K clusters.
- In order to update the cluster means, we need some measure to minimize an error or maximize a gain function.
- Defining the error term

$$E_i = \sum_{j=1}^{N_i} \sum_{k=1}^{L} (\mathbf{x}_{\mathbf{jk}} - \mathbf{m}_{\mathbf{ik}})^2$$

- Minimize E_i to find the correct $\mathbf{m}_\mathbf{i}$**'s.**

Algorithm contd.

- Differentiating E w.r.t. $\mathbf{m_j}$,
- $$\frac{\partial E_i}{\partial \mathbf{m_j}} = 0 \rightarrow \quad \mathbf{m_j} = \frac{1}{N_j} \sum_{k \in \{N_j\}} \mathbf{x}_{jk}$$

This means that in a mean-squared error sense, the new mean of a cluster is computed as the average of the elements that are assigned to that particular cluster.

K-means Algorithm contd.

- Let the means be estimated at two consecutive iterations m_j^n and m_j^{n+1}, j=1,2,…,K.
- If these means are very close then we have the converged estimate of the K-mean vectors

Fuzzy C-Means Clustering

- Input: Unlabeled data set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$

$\mathbf{x}_k \in \Re^p$

***n* is the number of data points in X**

p is the number of features in each vector

- Output

A c-partition of X, which is *c x n* matrix U

$$\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_c\} \subset \Re^p$$

Cluster centres

Fuzzy C-Means Clustering

- Optimization of an "objective function" or "performance index"

$$\min_{(\mathbf{U},\mathbf{V})}\left\{J_m(\mathbf{U},\mathbf{V}) = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m} D_{ik}^{2}\right\}$$

Constraint

$$\sum_{i=1}^{c} u_{ik} = 1 \;, \forall k$$

Distance

$$D_{ik}^{2} = \left\| \mathbf{x}_k - \mathbf{v}_i \right\|_{\mathbf{A}}^{2}$$

A-norm

$$\|\mathbf{x}\|_{\mathbf{A}} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_{\mathbf{A}}} = \sqrt{\mathbf{x}^T \mathbf{A} \mathbf{x}}$$

Degree of Fuzzification $m \geq 1$

Minimizing Objective Function

- Zeroing the gradient of J_m with respect to V

$$u_{ik} = \left[\sum_{j=1}^{c} \left(\frac{D_{ik}}{D_{jk}} \right)^{\frac{2}{m-1}} \right]^{-1}, \forall\, i, k$$

$$\mathbf{U}_t = F_\partial(\mathbf{V}_{t-1})$$

- Zeroing the gradient of J_m with respect to U

$$\mathbf{v}_i = \left(\sum_{k=1}^{n} u_{ik}^{m} \mathbf{x}_k \Big/ \sum_{k=1}^{n} u_{ik}^{m} \right), \forall i$$

$$\mathbf{V}_t = G_\partial(\mathbf{U}_{t-1})$$

Initial Choices

- Initial Choices
  - Number of clusters $1 < c \leq n$
  - Maximum number of iterations (e.g. 100) *T*
  - Weighting exponent (Fuzziness degree)
    - m=1: crisp
    - m=2: Typical
  - Termination measure $E_t = || V_t - V_{t-1}||$
  - Termination threshold (e.g. 1.0)

Algorithm for Iterations

- Guess Initial Cluster Centers
- Alternating Optimization (AO)

$t \leftarrow 0$

REPEAT (

$t \leftarrow t+1$

$U_t = F_\theta(V_{t-1})$

$V_t = G_\theta(U_{t-1})$

UNTIL ($t = T$) or ($\|V_t - V_{t-1}\| \leq \varepsilon$)

)

$(U,V) \leftarrow (U_t, V_t)$

Alternative

- Process could be shifted one half cycle
  - Initialization is done on U_n
  - Iterates become $U_{t-1} \rightarrow V_t \rightarrow U_t$
  - Termination criterion $|| U_t – U_{t-1}|| \leq \varepsilon$
- The convergence theory is the same in either case
- Initializing and terminating on V is advantageous
  - Convenience
  - Speed
  - Storage

Convergence of FCM Algorithm

- Advantages
  - Unsupervised
  - Always converges
- Disadvantages
  - Long computational time
  - Sensitivity to the initial guess (speed, local minima)
  - Sensitivity to noise
    - One expects low (or even no) membership degree for outliers (noisy points)

Optimal Number of Clusters

- **Performance Index** $\min_{(c)}\left\{P(c)=\sum_{i=1}^{c}\sum_{k=1}^{n}u_{ik}^{m}\left(\left\|\mathbf{x}_k-\mathbf{v}_i\right\|^2-\left\|\mathbf{v}_i-\overline{\mathbf{x}}\right\|^2\right)\right\}$

Mean of all feature vectors $\overline{\mathbf{x}}=\frac{1}{n}\sum_{k=1}^{n}\mathbf{x}_k$

$$\sum_{i=1}^{c}\sum_{k=1}^{n}u_{ik}^{m}\left(\left\|\mathbf{x}_k-\mathbf{v}_i\right\|^2\right)$$

Sum of the within fuzzy cluster fluctuations (small value for optimal c)

$$\sum_{i=1}^{c}\sum_{k=1}^{n}u_{ik}^{m}\left(\left\|\mathbf{v}_i-\overline{\mathbf{x}}\right\|^2\right)$$

Sum of the between fuzzy cluster fluctuations (large value for optimal c)

FCM example

Input Image

FCM example

K-Means Result
7-clusters

FCM example

FCM Result
7-clusters

Fuzzy Supervised Classification (An approach)

- An approach presented in an IEEE conference combined two fuzzy supervised classifier assigned memberships and iteratively updated the memberships till the values stopped changing. (**DOI:** 10.1109/ICICISYS.2009.5357646)
- The two classifiers are Fuzzy Mahalanobis Classifier and Fuzzy Maximum Likelihood Classifier.

Fuzzy Supervised Classification contd…

- Fuzzy memberships
- Fuzzy Mean Vector

$$\begin{cases} 0 \le f_i(x) \le 1 \\ \sum_{i=1}^{m} f_i(x) = 1 \end{cases}$$

$$\mu_k^* = \frac{\sum_{i=1}^{n} f_k(x_i) x_i}{\sum_{i=1}^{n} f_k(x_i)}$$

- Fuzzy Covariance Matrix

$$V_k^* = \frac{\sum_{i=1}^{n} f_k(x_i)(x_i - \mu_k^*)(x_i - \mu_k^*)^T}{\sum_{i=1}^{n} f_k(x_i)}$$

Fuzzy Supervised Classification contd…

- Mahalanobis distance based classifier
- $$f_k^{*}(x_i) = \frac{\left(\frac{1}{d_k^{*2}(x_i)}\right)^t}{\sum_{k=1}^{m}\left(\frac{1}{d_k^{*2}(x_i)}\right)^t}$$

$$d_k^{*2}(x_i) = (x_i - \mu_k^{*})^T (V_k^{*})^{-1} (x_i - \mu_k^{*})$$

Fuzzy Supervised Classification contd…

- Traditional ML classifier

$$f_k^{m}(x_i) = \frac{P_k^*(x_i)}{\sum_{k=1}^{m} P_k^*(x_i)}$$

- Class likelihood $P^*_k(\mathbf{x}_i)$

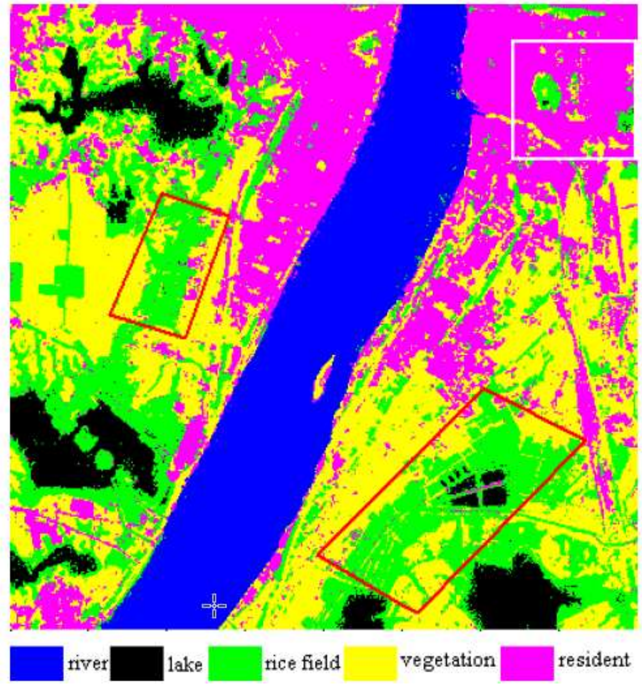
$$P_k^*(x_i) = \frac{1}{(2\pi)^{n/2}\left|V_i^*\right|^{1/2}} \exp\left[-\frac{1}{2}(x_i - \mu_i^*)^T (V_i^*)^{-1}(x_i - \mu_i^*)\right]$$

Fuzzy Supervised Classification contd…

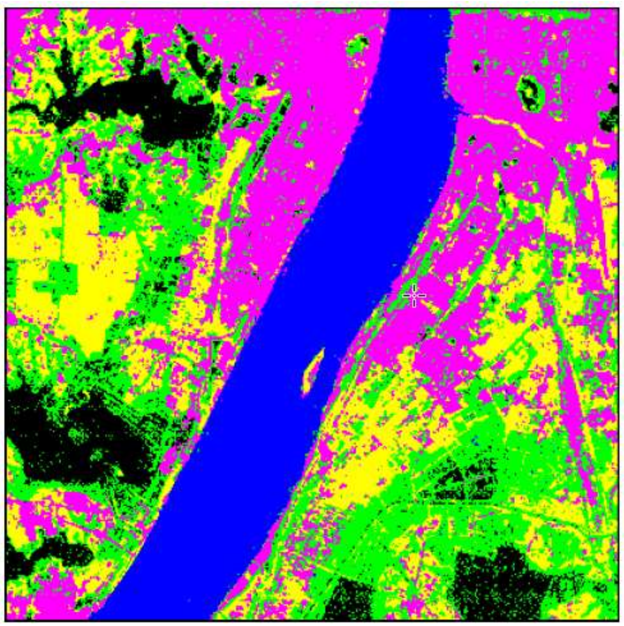
- Algorithm: Likely in case of FCM, iteratively update the fuzzy memberships for the classes for both the Mahalanobis distance classifier and the Maximum Likelihood Classifier. Then, final membership:

$$f_k(x_i) = \frac{(f_k^{'}(x_i))^2 + (f_k^{''}(x_i))^2}{\sum_{k=1}^{m}[(f_k^{'}(x_i))^2 + (f_k^{''}(x_i))^2]}$$

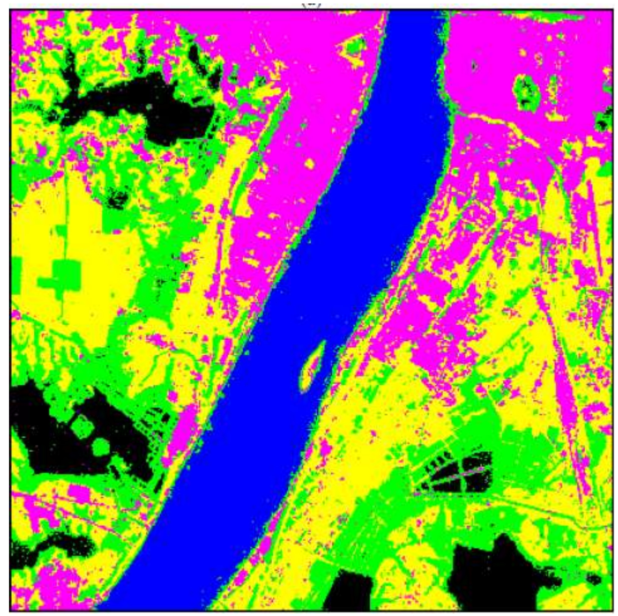
Input Image



Combined Classification



Mahalanobis Distance based



Max. Likelihood based

Overall Assessment

- Accuracy of classification of major classes improved, while some small classes were affected.
- The problem with small classes is that the number of pixels would be very few, so computation of covariance matrix will not be reliable.

Fusion of Fuzzy Memberships

Fusion of Fuzzy Memberships

- It is possible to obtain, for a given area, data from different sources, and thereby results through different algorithms
- How to put together the multiple results into one result, by reinforcing / weakening them based on different individuals?

Combining Fuzzy Memberships

- Assume that we are given different classification results produced by one or more fuzzy classifiers applied to different sources of data
- Fusion operations are applied to each pixel in the images, i.e., for any pixel k, selecting all the class memberships for all datasets

Fuzzy AND

- $\mu_{combination} = \mathrm{MIN}\{\mu_A, \mu_B, \ldots , \}$
- This implies the minimum confidence in the evidences obtained from different individual results
- This is a conservative estimate of the membership

Fuzzy OR

- $\mu_{combination} = MAX\{\mu_A, \mu_B, \ldots ,\}$
- This implies the maximum confidence in the evidences obtained from different individual results
- This operation is applicable when the number of choices is limited, and any positive evidence is acceptable

Fuzzy Algebraic Product

- $\mu_{combination} = \Pi_i(\mu_i)$
- The result of the fuzzy algebraic product is smaller since several numbers, each less than 1, are multiplied. (Decreasing result)
- Each membership has a role in the output, unlike the previous options, where the minimum or maximum alone were responsible for the final result

Fuzzy Algebraic Sum

- $\mu_{combination} = 1 - \Pi_i(1 - \mu_i)$
- This result is always greater than or equal to the largest result (Increasing output)
- Fuzzy Algebraic product is a product of fuzzy memberships, while fuzzy algebraic sum is not.

Gamma Operator

- Geometric combination of the algebraic sum and product operators
- $\mu_{combination} = [\Pi_i(\mu_i)]^{\gamma}\ [1 - \Pi_i(1 - \mu_i)]^{1-\gamma}$
- If $\gamma = 1$, this reduces to algebraic product
- If $\gamma = 0$, it reduces to algebraic sum
- In geological applications, gamma operator was found better than algebraic sum or product operators
- See Krishna Mohan et al. IJRS 2000 for an application of these operators for fusion classification results of multidate IRS image data

To be continued …