Image Segmentation: Type–2 Fuzzy Possibilistic C-Mean Clustering Approach

M.H. Fazel Zarandi* & M. Zarinbal

Mohammad Hossein Fazel Zarandi, Professor of Department of Industrial Engineering, Amirkabir University of Technology, Tehran, Iran
Mohammad Zarinbal, Department of Industrial Engineering, Amirkabir University of Technology, Tehran, Iran

KEYWORDS
Brain Tumors Diagnosis, Image segmentation, Type-2 Fuzzy Logic, Type-2 PCM

ABSTRACT
Image segmentation is an essential issue in image description and classification. Currently, in many real applications, segmentation is still mainly manual or strongly supervised by a human expert, which makes it irreproducible and deteriorating. Moreover, there are many uncertainties and vagueness in images, which crisp clustering and even Type-1 fuzzy clustering could not handle. Hence, Type-2 fuzzy clustering is the most preferred method. In recent years, neurology and neuroscience have been significantly advanced by imaging tools, which typically involve vast amount of data and many uncertainties. Therefore, Type-2 fuzzy clustering methods could process these images more efficient and could provide better performance. The focus of this paper is to segment the brain Magnetic Resonance Imaging (MRI) in to essential clusters based on Type-2 Possibilistic C-Mean (PCM) method. The results show that using Type-2 PCM method provides better results.

© 2012 IUST Publication, IIIEPR, Vol. 23, No. 4, All Rights Reserved.

1. Introduction
Human medicine is one of the most important parts of natural science and tries to understand and control the human body, its structure and function under all conditions of health, illness, and injury. In recent years, neurology and basic neuroscience have been significantly advanced by imaging tools that enable in vivo monitoring of the Central Nervous System. Among the various existing imaging tools, Magnetic Resonance Imaging (MRI) provides rich information about the human soft tissue anatomy, which made it an indispensable tool for human medicine and medical diagnosis. Applying this information frequently requires segmenting the MRI in to different classes. This task has proven to be problematic, due to vast amount of data, uncertainties and different characteristics of the brain tissues.

Various clustering methods especially fuzzy clustering methods are used to overcome these problems such as [2-5]. In this paper, we applied the Type–2 Possibilistic C-Means clustering (PCM) method proposed by [1] to segment the MRI and to detect the abnormalities in these images.

This detection could help the physicians to diagnose the abnormalities or tumors more efficient.

The rest of this paper is organized as follows: Section II explains the image segmentation methods shortly. Clustering methods, the most applied segmentation method, are described in Section III. Fundamentals of Type-2 fuzzy Possibilistic C-Means Method are described in Section IV. Section V addresses the brain tumors and the diagnosing methods. Section VI is dedicated to the experimental results. Finally, the conclusions are presented in Section VII.

2. Image Segmentation
Segmentation is an essential issue in image description and classification. It is based on a definition of uniformity, which usually depends on the
particular task and its context. Currently, in many real applications, segmentation is still mainly manual or strongly supervised by a human expert. The level of supervision affects the performance of the segmentation method in terms of time consumption and makes it irreproducible and deteriorating. Hence, there is a real need for automated segmentation methods, which could be classified into four main categories [6]: Tresholding [7], Edge detecting [8], Clustering [6], and Region Extracting [6].

- Tresholding: it is one of the simplest methods for obtaining a crisp segmentation. Tresholding generates a binary image in which the pixels belonging to the objects have the value 1 and pixels belonging to the background have the value 0 [7].
- Edge Detection: Edge detection is an important topic in computer vision and image processing, and applied in many areas [8].
- Clustering: Clustering is the most important methods in modern data mining, which is used in processing large databases. The general philosophy of clustering is to divide the initial set into homogenous groups and to reduce unnecessary data [6].
- Region Extraction: It contains three categories, Region Growing, Region-Based Segmentation and Contour Models [6].

The Image Segmentation is usually done by using clustering method. In clustering, the partitioning structure is usually obtained as exclusive clusters (groups) of objects in the observation space with respect to attributes (or variables). However, such a partitioning structure is inadequate to explain the related features of attributes (or variables). Fuzzy clustering can solve this problem and obtain the degrees of belongingness of objects to the fuzzy clusters.

That is, in a fuzzy clustering result, there exist objects that belong to several fuzzy clusters with certain degrees. Fuzzy clustering has been shown to be advantageous over crisp clustering in that total commitment of a vector to a given class is not required in each iteration [6]. Therefore, the automated image segmentation is usually done by using clustering method.

3. Clustering Methods

Clustering is an unsupervised learning task that aims at decomposing a given set of objects into subgroups or clusters based on similarity. The goal is to divide the data set in such a way that objects belonging to the same cluster are as similar as possible, whereas objects belonging to different clusters are as dissimilar as possible [9]. Clustering is useful in several exploratory decision-making, machine learning, data mining, image segmentation, and pattern classification [10].

Most of the clustering methods can be classified into two types: Crisp clustering and Fuzzy clustering. Crisp clustering assigns each data to a single cluster and ignores the possibility of belonging to other clusters. Since the boundaries between clusters could not be defined precisely, some of the data could belong to more than one cluster with different positive degrees of memberships. Fuzzy clustering method considers each cluster as a fuzzy set and the membership function measures the degree of belonging of each feature to a cluster [1].

In the last 30 years, many fuzzy clustering models for crisp data have been suggested. Fuzzy K-Means clustering model introduced by Dunn (1974) and Bezdek (1974) and extended by Bezdek (1981). Fuzzy C-Mean (FCM) is the best-known and applied model in the body of literature [9]. However, the FCM’s memberships do not always correspond well to the degrees of belonging and it may be inaccurate in a noisy environment. To relieve these weaknesses, Krishnapuram and Keller (1993) created a Possibilistic approach for the clustering purpose, which used a Possibilistic type of membership function to describe the degree of belonging [1]. These two important models of fuzzy clustering are described as follows:

**Fuzzy C-Mean (FCM):** Fuzzy C-Means clustering model can be defined as follows [1]:

$$\min \left\{ J(x, \mu, c) = \sum_{i=1}^{N} \sum_{j=1}^{c} \mu_{ij}^{m} d_{ij}^{2} \right\}$$

ST:

$$0 < \sum_{j=1}^{c} \mu_{ij} < N \quad \forall i \in (1, 2, ..., c)$$

$$\sum_{i=1}^{N} \mu_{ij} = 1 \quad \forall j \in (1, 2, ..., N)$$

where, $\mu_{ij}$ is the degree of belonging of the $j^{th}$ data to the $i^{th}$ cluster, $d_{ij}$ is the distance between the $j^{th}$ data and the $i^{th}$ cluster center, $m$ is the degree of fuzziness, $c$ is the number of clusters, and $N$ is the number of the data. However, when the data set is noisy the performance of FCM might be inadequate. In addition, when the norm method is different to the Euclidean, introducing restrictions is necessary, such as fuzzy covariance matrices proposed by Gustafson and Kessel [11].

**Possibilistic C-Mean (PCM):** To improve the FCM weaknesses, Krishnapuram and Keller (1993) created a possibilistic approach, which uses a possibilistic type of membership function to describe the degree of belonging. It is desirable that the memberships of the
representative feature points be as high as possible, while the memberships of unrepresentative points be as low as possible. The objective function, which satisfies this requirement, is formulated as follows [12]:

$$\min \left\{ J_m(x, \mu, c) = \sum_{i=1}^{n} \sum_{j=1}^{c} \mu_i^c d_{ij}^2 + \sum_{i=1}^{n} \sum_{j=1}^{c} (1 - \mu_i^c)^2 \right\}$$

(4)

where, $d_{ij}$ is the distance between the $j^{th}$ data and the $i^{th}$ cluster center, $\mu_i^c$ is the degree of belonging of the $j^{th}$ data to the $i^{th}$ cluster, $m$ is the degree of fuzziness, $c$ is the number of clusters, and $N$ is number of the data. $\eta_i$ can be obtained by using Eq.(5) [12]:

$$\eta_i = \frac{1}{1 + \left( \frac{d_{ij}^2}{\eta_i} \right)^{\frac{1}{m}}}$$

(5)

The value of $\eta_i$ determines the distance at which the membership value of a point in a cluster becomes 0.5. In practice, Eq.(6) is used to obtained $\eta_i$ values. The value of $\eta_i$ can be fixed or changed in each iteration by changing the values of $\mu_i^c$ and $d_{ij}$ [12]:

$$\eta_i = \frac{\sum_{j=1}^{c} \mu_i^c d_{ij}^2}{\sum_{j=1}^{c} \mu_i^c}$$

(6)

Updating the membership values depends on the distance measurements. The Euclidean and Mahalanobis distance are two common ones. The Euclidean distance works well when a data set is compact or isolated and Mahalanobis distance takes into account the correlation of data by using the inverse of the variance-covariance matrix of data set [13]. Gustafson and Kessel [11] proposed a new approach based on Mahalanobis distance, which enables the detection of ellipsoidal clusters [1]. Satisfying the underlying assumptions, such as shape and number of clusters, is another important issue in clustering methods, which could be obtained by validation indices such as Kwon validity index. This index is shown in Eq.(7) [14]:

$$V_k(c) = \frac{\sum_{j=1}^{c} \sum_{j=1}^{N} \mu_i^c \|v_j - v_i\|^2 + \frac{1}{c} \sum_{c}^c \|v_i - \bar{v}\|^2}{\min_{i \in j} \|v_i - v_j\|^2}$$

(7)

where, $\mu_i^c$ is the degree of belonging of the $j^{th}$ data to the $i^{th}$ cluster, $v_i$ is center of the $i^{th}$ cluster, $\bar{v}$ is the mean of the cluster centers, $m$ is the degree of fuzziness, $c$ is the number of clusters, and $N$ is number of the data. To assess the effectiveness of clustering method, the smaller the $V_k$, the better the performance [14].

In sum, the PCM is more robust in the presence of noise, in finding valid clusters, and in giving a robust estimate of the clusters centers [1]. In sum, the PCM is more robust in the presence of noise, in finding valid clusters, and in giving a robust estimate of the clusters centers [1].

However, as mentioned before, there exist many uncertainties in real world, which Type-1 fuzzy clustering could not handle. Hence, Type-2 fuzzy clustering is the most preferred method and with more accurate and better results. This is explained shortly in the next section.

4. Type-2 Fuzzy Possibilistic C-Means Method

Fuzzy logic is an area of soft computing that enables a computer system to reason with uncertainty [15]. Type-II fuzzy set (introduced by Zadeh [16]) is characterized by fuzzy membership function, and applied in many clustering methods such as [17-20]. Type-2 fuzzy logic is increasingly being advocated as methodology for reasoning in situations where large uncertainties are present [21]. The uncertainties may arise from the following sources [22]:

- The words used in the antecedents and the consequents of rules can mean different things to different people.
- Consequents obtained by polling a group of experts may differ.
- The training data are noisy.
- The measurements are noisy.

However, Type-2 fuzzy sets are difficult to understand and use because:

- The three-dimensional nature of Type-2 fuzzy sets makes them very difficult to draw.
- There is no simple collection of well-defined terms, which makes us effectivly communicate about Type-2 fuzzy sets, and make us be mathematically precise about them.
- Derivations of the formulas for the union, intersection, and complement of Type-2 fuzzy sets all rely on using Zadeh’s Extension Principle, which in itself is a difficult concept.
- Using Type-2 fuzzy sets is computationally more complicated than using Type-1 fuzzy sets [18].

A Type-2 fuzzy set is characterized by fuzzy membership function, i.e., the membership grade for each element is a fuzzy set in interval [0,1]. Such sets
can be used in situations where there are uncertainties about the membership values [23]. There are essentially two types of Type-2 fuzziness: Interval-Valued Type-2 fuzzy and Generalized Type-2 fuzzy. The membership grade of a generalized Type-2 fuzzy set is a Type-1 fuzzy set, with a support bounded in [0,1] and is known as a secondary membership function (Fig. 1.b). A simplified version of generalized Type-2 fuzzy sets is Interval Type-2 fuzzy sets, which are predominate in the literature. However, the interval Type-2 fuzzy sets have less expressive power than generalized Type-2 fuzzy sets. Interval Type-2 fuzzy sets have membership grades that are crisp intervals and its processing is simplified by only ever working with the end points of these intervals (Fig. 1.a) [21].

Ref. [1] proposed a new PCM method in which the membership functions are Type-2 Fuzzy, the distance is assumed Mahalanobis and Type-2 Kwon validity index is used to find the optimal degree of fuzziness \(m\) and number of clusters \(c\). This method is as follows:

\[
\min \left\{ J_n(x, \bar{\mu}, c) = \sum_{i=1}^{c} \sum_{j=1}^{N} \tilde{\mu}_j D_{ij} + \sum_{i=1}^{c} \eta_i \sum_{j=1}^{N} (1 - \tilde{\mu}_j)^m \right\} \tag{8}
\]

S.T.:

\[
0 < \sum_{j=1}^{N} \tilde{\mu}_j < N \quad \forall i \in \{1,2,...,c\} \tag{9}
\]

\[
\tilde{\mu}_i \in [0,1] \quad \forall i,j \tag{10}
\]

\[
\text{max} \tilde{\mu}_j > 0 \quad \forall j \tag{11}
\]

where, \(\tilde{\mu}_i\) is Type-2 membership for the \(i^{th}\) data in the \(j^{th}\) cluster (obtained by Eq.(14)), \(D_{ij}\) is the Mahalanobis distance of the \(i^{th}\) data to the \(j^{th}\) cluster’s center, \(\eta_i\) is positive numbers (obtained by 13), \(c\) is the number of the clusters, and \(N\) is the number of input data. The first term in the objective function make the distance to the cluster’s center be as low as possible and the second term make the membership values in a cluster to be as large as possible. The membership values for data in each cluster must lie in the interval [0,1], and their sum are restricted to be smaller than the number of input data, as shown in Eq. (8-11) [1].

\[
\bar{\mu}_i = \frac{1}{1 + \frac{D_j}{\eta}} \tag{12}
\]

\[
\eta_i = \frac{\sum_{j=1}^{N} \tilde{\mu}_j D_{ij}}{\sum_{j=1}^{N} \tilde{\mu}_j^m} \tag{13}
\]

Since the parameter \(\eta_i\) is independent of the relative location of the clusters, the membership value \(\bar{\mu}_i\) depends only on the distance of a point to the cluster center. Hence, the membership of a point in a cluster is determined solely by how far a point is from the center and is not coupled with its location with respect to other clusters [1]. The clustering method needs a validation index to define number of clusters \(c\) and degree of fuzziness \(m\) used in Eq.(12-13). A Type-2 Kwon Index Eq.(14) is used for this purpose:
tumor diagnosis greatly by histologic type and age at diagnosis.

Astrocytoma is the most common type of brain tumor, contains more than 30% of all brain tumors and is a usually malignant one. Astrocytoma can be subdivided into 4 grades: Grade I, Grade II, Grade III, and Grade IV [26]. These grading systems have been shown to correlate well with survival, so that survival ranges are more than 5 years for grade II, between 2-5 years for grade III, and less than 1 year for Grade IV. However, location of the tumor, extent of surgical resection, and response to therapy also significantly affect the survival.

This information is shown in Table 1 [27]. Therefore, diagnosing the brain tumors in an appropriate time is very essential for further treatments.

<table>
<thead>
<tr>
<th>Tumors</th>
<th>% of epidemiology</th>
<th>Mean age at diagnosis</th>
<th>% of survival after 2 year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade I</td>
<td>1.8</td>
<td>17</td>
<td>91</td>
</tr>
<tr>
<td>Grade II</td>
<td>1.3</td>
<td>47</td>
<td>67</td>
</tr>
<tr>
<td>Grade III</td>
<td>4.3</td>
<td>50</td>
<td>46</td>
</tr>
<tr>
<td>Grade IV</td>
<td>22.6</td>
<td>62</td>
<td>9</td>
</tr>
</tbody>
</table>

Different characteristics of brain tissues (such as intensity of each tissue) and Astrocytoma tumors (such as regions of tumor and infiltration) make it very difficult to assign an exact degree of belonging to one cluster.

Therefore, using Type-1 fuzzy logic and defining Type-1 fuzzy membership function is inadequate. On the other hand, Type-2 fuzzy membership values are themselves fuzzy and have the potential to provide better performance in diagnosing brain tumors based on image processing approaches [6]. Therefore, we use the Type-2 PCM method proposed by [1] to diagnosis the abnormality in brain MRI.
6. Experimental Results

Applications that use the morphologic contents of MRI such as drug treatment and abnormality detection frequently require segmentation of the image volume into tissue types. Accurate segmentation of brain MRI is of interest in the study of many brain disorders. Volumetric analysis of Gray Matter (GM), White Matter (WM) and Cerebrospinal Fluid (CSF) is important to characterize morphological differences between subjects [31]. Such studies typically involve vast amount of data and many uncertainties in MRI such as uncertainties related to boundaries, non-homogeneous regions and different characteristics of the brain tissues.

Clustering methods could process large amount of data and according to [1], Type-2 fuzzy logic can provide better performance than Type-1 fuzzy in order to process the existing uncertainties. Therefore, to detect the abnormalities in MRI we use the Type-2 PCM method proposed by [1], which is described shortly in Section IV.

In this paper, in order to segment the brain MRI, we assumed that the number of clusters, c, is assumed to be 4 (four main tissues in MRI CSF, WM, GM, and Abnormality). For the sake of simplicity the degree of fuzziness, m, is assumed to be 2.5. For evaluating the performance of the clustering method proposed by [1] 50 patient's brain T1-weighted MRIs (such as Fig. 2) are used.

The type-2 membership functions for these 4 clusters are shown in Fig. 3. Based on these membership functions and the afore mentioned Type-2 fuzzy clustering method, we generate a MATLAB code, segment the MRI and extract the abnormality cluster by using Center of Gravity method. The output image for the Abnormality cluster is shown in Fig. 4. Using Type-II PCM for MRI provides more reliable results. This method could also provide better tools for physicians to diagnose the abnormality in a suitable time.

7. Conclusions

In this paper, we applied the Type-2 Possibilistic C-Means clustering (PCM) method proposed by [1] to segment the MRI and to detect the abnormalities in these images. By using this automated segmenting method, good results are obtained, which does not suffer the weaknesses of manual segmenting procedures. The results help the physicians to diagnose the abnormalities or tumors more efficient in a suitable time.

References


