Applied Soft Computing xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Soft fuzzy rough set-based MR brain image segmentation

Anupama Namburu[a,*], Srinivas kumar Samay[b], Srinivasa Reddy Edara[c]

[a] Department of CSE, Jawaharlal Nehru Technological University, Kakinada, India
[b] Department of ECE, Jawaharlal Nehru Technological University, Kakinada, India
[c] Department of CSE, Acharya Nagarjuna University, Guntur, India

ARTICLE INFO

Article history:
Received 31 October 2015
Received in revised form 11 July 2016
Accepted 8 August 2016
Available online xxx

Keywords:
Soft sets
Segmentation
K-means
Fuzzy c-means
Soft fuzzy rough c-means
Histogram

ABSTRACT

Fuzzy sets, rough sets are efficient tools to handle uncertainty and vagueness in the medical images and are widely used for medical image segmentation. Soft sets are a new mathematical approach to uncertainty and vagueness. In this paper, a hybrid segmentation algorithm based on soft sets namely soft fuzzy rough c-means is proposed to extract the white matter, gray matter and the cerebro spinal fluid from MR brain image with bias field correction. In this algorithm, soft fuzzy rough approximations are applied to obtain the rough regions of image. These approximations are free from defining thresholds, weight parameters and are less complex compared to the existing rough set based algorithms. Soft sets use similarity coefficients to find the similarity of the clusters formed in present and previous step. The proposed algorithm does not involve any negative region, hence all the pixels participate in clustering avoiding clustering mistakes. Also, the histogram based centroids choose the centroids close to the ground truth that in turn effect the definition of approximations, standardizing the clusters. The proposed algorithm evaluated through simulation, compared it with existing k-means, rough k-means, fuzzy c-means and other hybrid algorithms. The soft fuzzy rough c-means algorithm outperforms the considered algorithms in all analyzed scenarios even in extracting the tumor from the brain tissue.

1. Introduction

Image segmentation is the most critical function in image analysis and processing. The segmentation of magnetic resonance (MR) image has enormous application such as automating, delineating anatomical, diagnosis, for extracting tissue volumes, localization of pathology, treatment planning, partial volume correction of functional imaging data, and computer integrated surgery [1]. The MR brain image is often divided/segmented into white matter (WM), gray matter (GM) and the cerebro spinal fluid (CSF). The effective extraction of these regions is very important for quantitative analysis and hence, resulted in several methods for segmenting MR brain images.

MR imaging has advantages over other medical imaging models, including the high contrast among different soft tissues, relatively high spatial resolution across the entire field of view and multi spectral characteristics [2]. Hence, clustering methods are very well suited for segmenting the regions. Clustering methods are unsupervised with high reproducibility and mainly based on the information of image data itself and requires little or no assumption of the model/distribution of the image data [3] unlike classification.

The classical k-means (KM) clustering algorithm or the hard c-means is based on crisp set [4] where a pixel belongs to a single cluster at a time [5]. Its fuzzy version, fuzzy c-means (FCM) [6] has also gained tremendous attention, where there is a cluster overlap [7]. Several modifications and extensions on fuzzy c-means like possibilistic fuzzy c-means [8], fuzzy c-means with spatial information [9], fast c-means [10], robust c-means [11] and MICO fuzzy c-means for bias estimation [12] are reported in the literature. However, fuzzy c-means method does not take into account the spatial context and the boundary conditions, hence, is sensitive to noise and insufficient data.

In order to consider spatial context and boundary regions soft computing technique rough sets [13,14] are incorporated to KM and FCM frame work for image segmentation. Hybrid algorithms viz., rough k-means [15], rough fuzzy possibilistic c-means [16–18], shadowed c-means [19], etc., were presented in literature. There are some disadvantages with these hybrid algorithms. First disadvantage is, grouping the pixels into rough regions based on single threshold and defining negative region. The threshold is calculated using the least and greatest memberships of the pixels [15,16,19] which depends on the initial centroids of the cluster. Hence, improper selection of the initial centroids results in

* Corresponding author.
E-mail address: namburianupama@gmail.com (A. Namburu).

http://dx.doi.org/10.1016/j.asoc.2016.08.020
1568-4946/

inaccurate results. The generalized fuzzy c-means [20], uses a rough set algorithm that has two thresholds and group the pixels into three regions positive, boundary and negative. This definition of rough approximation of the universe makes some basic properties of rough sets to evaporate and unfortunately eliminates pixels in the negative region from clustering, resulting in inaccurate clusters. Lingras et al. [21], presented a rough set algorithm to find the lower and upper approximations using a threshold with out using negative region. Here, the pixels definitely belong to lower or upper approximation or both. Hence, there is no elimination of pixels while clustering. However, this technique uses the random initialization of cluster prototype which yields inaccurate results when chosen wrongly. Shabir et al. [22] has proposed soft rough set to define the lower and upper approximations that does not include the negative region. However these definitions are for crisp values which cannot handle the boundary pixels.

Second disadvantage of rough set is, defining the rough regions is very complex. All the rough set based algorithms discussed before involve complex instructions to calculate the approximations. Moreover, hybrid algorithms that combines intuitionistic fuzzy sets and rough sets were proposed in literature [23–25]. These algorithms used triple memberships indicating belongingness to a cluster, non-belongingness to a cluster and hesitancy to a cluster. The lower and upper approximations and the regions are calculated using the triple memberships. Hence, the calculations involved in these algorithms are more complex than rough sets.

Besides these disadvantages in rough set, setting the weight parameters (lower weight and the upper weight) for cluster region preference remains one of the great challenges since no well accepted guidelines have been proposed [26]. These drawbacks of rough set theory are due to lack of parameterization tools [27].

The absence of any restrictions on the approximate description in soft sets makes it very convenient and applicable to many applications [28]. Aktas and Cagman [29] compared soft sets with the related concepts of fuzzy sets, rough set and proved that every fuzzy set, rough set is a soft set. Molodtsov applied this theory to several directions [28] and formulated the notions of soft number, soft derivative and soft integral, etc. The soft set theory has been applied to many different fields with great success. Maji et al. [30], worked on the theoretical study of soft sets in detail, and presented an application of soft set in the decision making problem using the reduction of rough sets. Feng et al. [31], proposed the concept of soft rough sets, in which instead of equivalence classes, parameterized soft set defines the lower and upper approximations. However, Feng's method uses the negative region and hence, eliminates the elements present in negative region [22]. The shabir's method by proposing modified soft sets granulated the soft sets with crisp values. Handling the boundary pixels (imprecise environment) is not possible with crisp soft sets. Hence, with fuzzy soft sets the imprecise environment can be parameterized. Meng et al. [32] proposed a method to calculate the lower and upper approximations for fuzzy soft sets that effectively define the boundary.

Fuzzy sets and rough set were applied to MR brain image segmentation and proved to handle the rough regions of the image with better segmentation results. However, considering the disadvantages of rough sets, the image segmentation using rough sets is very complex and depends on various parameters. Hence, in this paper, MR brain image segmentation is proposed based on soft fuzzy rough sets using Meng's method to reduce the complexity and time involved in segmentation. This algorithm proposes the soft sets representation of image and define the lower and upper approximation of rough regions of the image with soft fuzzy rough approximations [32]. This algorithm for clustering avoids the usage of threshold, weight parameters and complex upper and lower approximations unlike rough set algorithms. The proposed algorithm does not use negative region, hence, all the pixels can participate in clustering. Soft set makes use of similarity measure to quantify the uncertainty pertaining to which extent the previous cluster and the present clusters obtained are alike. Also, a histogram based initialization of cluster prototype is used to choose the appropriate centroids that help in calculation of upper and lower approximations. The performance of the proposed hybrid algorithm is compared to KM, rough k-means (RKM), rough fuzzy c-means (RFCM), and generalized rough c-means (GFCM).

The paper is organized as follows: the background of the proposed algorithm for image segmentation is discussed in Section 2. The proposed algorithm based on soft set is presented in Section 3. Implementation and experimental results of the proposed hybrid algorithm is presented in Section 4. The discussion related to the proposed algorithm is discussed in Section 5. Conclusions and the future scope of the proposed algorithms are presented in Section 6.

2. Background

2.1. Bias field model

The MR brain image I with the bias field can be modeled as

$$I = b \cdot J + n \tag{1}$$

where I is the observed image, J is the true image, b is the bias field to be estimated from the observed image and n is the additive Gaussian noise with zero mean. The true image J is characterized by the physical properties of the tissue, which takes a specific value for the voxel with in the same tissue. Hence, it is assumed that J takes a constant value X_i for all the X voxels in the ith tissue. The bias field is assumed to vary smoothly and can be estimated using linear combination of bases functions.

2.2. Rough set

Rough sets are used to approximate the uncertain data using lower and upper approximation. Let U denote the universe and R is an equivalence relation and U/R is a set of n equivalence classes $\{x_1, x_2, \ldots, x_n\}$ which form partitions in X. The pair (U, R) is the approximation space. The lower and upper approximations for a subset $X \subseteq U$ are denoted by:

$$\underline{R}(x_i) = \bigcup_{x_i \in X} x_i \quad \bar{R}(x_i) = \bigcup_{x_i \cap X \neq \phi} x_i \tag{2}$$

$\underline{R}(x_i)$ indicates the lower approximation space of X where an object x_i belongs to X. $\underline{R}(x)$ indicates the upper approximation space of X where an object x_i possibly belongs to X. The approximation space of X is classified into three distinct region.

$$\begin{aligned} positive(x_i) &= \underline{R}(x_i) \\ boundry(x_i) &= \bar{R}(x) - \underline{R}(x_i) \\ negative(x_i) &= U - \bar{R}(x_i) \end{aligned} \tag{3}$$

Using rough set theory hybrid algorithms like rough k-means and rough fuzzy c-means algorithms are proposed. In order to understand rough set application to image segmentation, consider that a pixel belongs to all the clusters with very less membership then including that pixel in cluster adds noise and removing the pixel from respective clusters will generate insufficient data. These incomplete, uncertainty and vagueness in the clusters can be defined using lower and upper approximations of rough sets. The lower approximation of a cluster is the set of the pixels that definitely belong to the cluster and the upper approximation is the set of pixels that possibly belong to the clusters. The pixels in the boundary region (upper–lower approximations) are given a second chance, such that in next iteration they can move to the appropriate cluster. Hence, with rough sets clustering mistakes can be reduced.