

# Application of Rough Set Based Dynamic Parameter Optimization to MRI Segmentation

S. Widz

Department of Robotics and Multi-Agent  
Systems  
Polish-Japanese Institute of Information  
Technology  
Warsaw, Poland  
sebastian.widz@pjwstk.edu.pl

D. Slezak\*

Department of Computer Science  
University of Regina  
Regina, Canada  
slezak@uregina.ca

K. Revett

Department of Computer Science  
University of Luton  
Luton, United Kingdom  
biomodelling@aol.com

**Abstract – We introduce a multi-spectral MRI segmentation technique based on approximate reducts derived from the data mining paradigm of the theory of rough sets. We use genetic algorithms to tune the parameterized attributes and search for the best segmentation models based on approximate reducts.**

## I. INTRODUCTION

Magnetic Resonance Imaging (MRI) data segmentation is a task of assigning proper labels to pixels (called voxels in 3D data sets), where labels for brain images correspond to tissue types: white matter (*WM*), grey matter (*GM*), cerebrospinal fluid (*CSF*), bone, and background. Relative distributions and changes at the level of brain tissues allow for diagnosing specific diseases and pathologies, like Alzheimer's disease, Parkinson's disease, Multiple Sclerosis, stroke, etc. [3, 4].

Segmentation accuracy is a similarity measure between the algorithm's result and the expert's evaluation. We need an objective and verifiable data set with a properly segmented brain image (phantom). We use Simulated Brain Database (SBD)<sup>1</sup> containing 3D volumetric multi-spectral MRI images (*T1*, *T2*, *PD*) with axial orientation. Every set contains 181 slices (1mm thickness), every slice is 181x217 voxels. Various data sets are available with varying slice thickness, noise ratio and field inhomogeneity (INU) levels. We test the algorithm's accuracy across all those image parameters.

Earlier MRI segmentation methods have been performed using cluster analysis, histogram extraction, neural networks, etc. [1, 5, 7, 17]. We present an approach based on the concept of approximate reducts derived from the theory of rough sets [6, 9, 10, 11, 12, 13]. We use methods described in our earlier work [15]. Furthermore, we propose modifications based on adaptive tuning of parameterized attributes, resulting in a potentially higher accuracy level.

The article is organized as follows: Section II describes data preparation. Section III describes algorithms employed to find approximate reducts, rules used for testing segmentation, as well as possible extensions. Section IV presents our results. A brief discussion follows in Section V.

## II. DATA PREPARATION

In the rough set theory the sample of data takes the form of an *information system*  $A=(U,A)$ , where each attribute  $a \in A$  is a function  $a:U \rightarrow V_a$  from the universe  $U$  into the value set  $V_a$ . In our case the elements of  $U$  are voxels taken from MRI images. We restrict our classification algorithm to *WM*, *GM*, and *CSF*. Below we characterize the method we employ to extract the attributes in  $A$  from the MRI images.

### A. Magnitude

*Magnitude attributes*, denoted by  $mag_{T1}$ ,  $mag_{T2}$ ,  $mag_{PD}$ , have are derived from frequency histograms for *T1*, *T2*, and *PD* modalities. Figure 1 displays the total voxel content of a single SBD *T1*-weighted slice. There are several peaks which correspond to certain tissue types. In [15] we describe how to extract the peaks using standard polynomial interpolation. Because it is hard to choose proper degree for a polynomial, we propose a genetic algorithm ( $GA_{mag}$ ) that searches for the peak values. Every gene in a chromosome encodes a Gaussian distribution. The sum of densities encoded by distributions is compared to the histogram's shape and *error* is calculated.

Genetic algorithms simulate evolution of *individuals* within a population [2, 16]. Evolution's result is an increase in the average fitness of members of a population, which strives towards a global optimum. Algorithm's behavior depends on the specification of the *fitness function*, which evaluates individuals and determines which of them are likely to survive. In our case, fitness is specified by (1). By changing the  $\gamma$  parameter we control the importance of chromosome length which has a direct influence on number of the peaks found by the algorithm on the histogram, where  $\gamma \in [0,1]$ :

$$F(x) = 1/(error+1) + \gamma * 1/(number\ of\ peaks) \quad (1)$$

For crossover we select two parents, and randomly choose a cut for each of them separately. Then we exchange their tails, which results in creation of two new individuals in next population. To minimize the number of peaks (and therefore – genes), we use the  $\gamma$  values at the order of 0.01 – 0.1.

\* The second author is supported by the grant awarded from the Faculty of Science at the University of Regina

<sup>1</sup> The SBD data sets were provided by the Brain Imaging Center, Montreal Neurological Institute (<http://www.bic.mni.mcgill.ca/brainweb>)

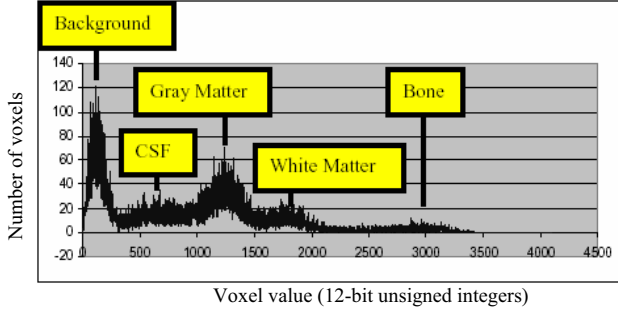


Fig. 1 A single bin frequency histogram from a  $T1$  SBD slice #81 (1mm slice thickness, 3% noise and 20% INU).

Mutation is a random change performed at the gene level, either on the mean or standard deviation of distribution. Each gene has two hidden normal distributions attached to two parameters defining genes' performance. During crossover those distributions are exchanged as well. It can be compared to the classical evolutionary strategy approaches [2].

In [15] the extraction of the Full-Width Half-Maximum (FWHM) interval for each peak is applied. Here we propose a slightly different method. We label objects-voxels with the corresponding peak numbers. Namely, we label voxel  $v$  with number  $i$  if  $v$ 's closeness to the  $i$ -th peak (measured on the frequency histogram) is at least  $\mu$  times larger than to the other peaks. Otherwise, if  $v$  is in between the  $i$ -th and  $(i+1)$ -th peak and we cannot decide between them at the level of accuracy specified by parameter  $\mu > 0$ , we label  $v$  with  $i+1/2$ . In this way, by tuning  $\mu$  we can build a discretized attribute based on the previously extracted histogram's peaks. The proposed way of operating with the closeness to the peaks can be referred to the *Rough Bayesian* model developed in [11].

#### B. Discrete Laplacian and Neighbor

*Discrete Laplacian attributes*, denoted by  $lap_{T1}$ ,  $lap_{T2}$ ,  $lap_{PD}$ , have the values derived by a general non-directional gradient operator, which is used in this context to determine whether the neighboring voxels have enough homogenous values. For instance,  $lap_{T1}$  takes the value 0 for a given voxel, if its neighborhood for  $T1$  is homogeneous, and 1 otherwise. We use parameter  $\beta > 0$  to specify the difference between voxel values, which is still considered as a homogenous region. Similarly to  $\mu$ , the value of  $\beta$  can be tuned. The values of discrete Laplacian attributes are calculated as follows:

$$\begin{aligned} & \text{if} ( |n_0 - \sum_i n/\delta| \leq \beta ) \\ & \quad lap = 0 \\ & \text{else} \\ & \quad lap = 1 \end{aligned}$$

The associated *neighbor attributes*, denoted by  $nbr_{T1}$ ,  $nbr_{T2}$ ,  $nbr_{PD}$ , replace original values of  $mag_{T1}$ ,  $mag_{T2}$ ,  $mag_{PD}$  on the basis of the region's homogeneity. Although they are derivable from the magnitude and the discrete Laplacian attributes, they turn out to be helpful in the segmentation process. They are calculated using the following procedure:

```

if (lap == 0)
    nbr = magnitude value from the decision table
else // (we want to adjust the middle voxel to its neighbors)
    if (there is unique most frequent value in neighborhood)
        nbr = the most frequent value in neighborhood
    else
        nbr = the value of a randomly chosen neighbor

```

#### C. Mask

It can happen that two different tissues have similar voxel value and appear on the histogram in one peak. Knowing the relative position of a voxel may help remove the resulting classification ambiguity. It provides us with extra information about the voxel location (e.g. *CSF* tissue is located in the middle of the brain and bone on the most outside parts).

The *mask attribute*, denoted by  $msk$ , is a rough estimation of the position of a voxel within the brain. Mask is created as follows: First we isolate the brain region by creating a binary mask. On a histogram we find a frequency value, below which every magnitude value corresponds to background (first peak on a histogram). After the artifact removal and holes filling, we are left with a single solid masked brain region. Then a central point is calculated, which is the central position  $(x,y)$  of the masked voxels. We divide the mask into four parts by drawing two orthogonal lines that cross at the center. Then three translations are made of all four parts. The radii are controlled by parameter  $rad$ , which can be tuned again, like in case of  $\mu$  and  $\beta$  before.

The above procedure yields concentric circles defining the approximations of bone, *GM*, *WM*, and *CSF*. Values of  $msk \in A$  in  $A = (U, A \cup \{d\})$  are defined by membership of voxels to particular regions.

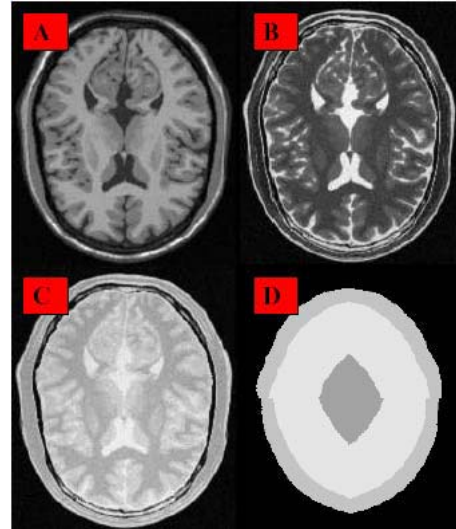


Fig. 2 Modalities  $T1$  (Picture A),  $T2$  (B), and  $PD$  (C) from the SBD data set, generated for slice thickness 1mm, 3% noise, and 20% field inhomogeneity (INU). Picture D presents the mask obtained for these modalities. Please note that the tissue approximation induced by mask is not very accurate due to the way of its heuristic specification. Still, it contains important complementary information with respect to the other attributes.

### III. METHODS

#### A. Approximate Reduction

When modeling complex phenomena, one must strike a balance between accuracy and complexity. In the current context, this balance is achieved through the use of a *decision reduct*: an irreducible subset  $B \subseteq A$  determining  $d$  in decision table  $A=(U, A \cup \{d\})$ . The obtained decision reducts are used to produce the decision rules from the training data.

For smaller reducts we generate shorter and more general rules, better applicable to new objects. Therefore, it is worth searching for reducts with a minimal number of attributes. Sometimes it is better to remove more attributes to get even shorter rules at the cost of slight inconsistencies. One can specify a measure  $M(d/\cdot):P(A) \rightarrow R$  which evaluates the degree of influence  $M(d/B)$  of subsets  $B \subseteq A$  on  $d$ . Then one can decide which attributes may be removed from  $A$  without a significant loss of accuracy. Given  $A=(U, A \cup \{d\})$ , accuracy measure  $M(d/\cdot)$ , and approximation threshold  $\varepsilon \in [0,1)$ , let us say that  $B \subseteq A$  is an  $(M, \varepsilon)$ -approximate decision reduct, if it satisfies inequality  $M(d/B) \geq (1-\varepsilon)M(d/A)$  and none of its proper subsets does it. For a detailed study on such reducts we refer the reader to [10]. In this article, we consider the multi-decision relative gain measure [13, 15]:

Definition 1. Let  $A=(U, A \cup \{d\})$  and  $\varepsilon \in [0,1)$  be given. We say that  $B \subseteq A$  is an  $(R, \varepsilon)$ -approximate decision reduct, if and only if it is irreducible set of attributes that satisfies inequality

$$R(d/B) \geq (1-\varepsilon)R(d/A) \quad (3)$$

where

$$R(d/B) = \sum_{\text{rules } r \text{ induced by } B} \left( \frac{\text{number of objects recognizable by } r}{\text{number of objects in } U} * \max_i \frac{\text{probability of the } i\text{-th decision class induced by } r}{\text{prior probability of the } i\text{-th decision class}} \right) \quad (4)$$

Measure (4) expresses the average gain in determining decision classes under the evidence provided by the rules generated by  $B \subseteq A$  [13]. In particular, it can be applied to evaluate the potential influence of single attributes on the decision. For instance, we can use it while searching for the peaks in the frequency histograms. The quantity of *1/error* in equation (1) can be replaced by  $R(d/\{mag_\mu\})$ , where attribute  $mag_\mu$  is obtained by closeness-based discretization described in Section II, for appropriately tuned parameter  $\mu > 0$ .

#### B. Order Based Genetic Algorithms

The quantities of  $R(d/\{a\})$ ,  $a \in A$ , reflect information gain from one-attribute rules. They are, however, not sufficient to select the *subsets* of relevant attributes. Several attributes with low values of  $R$  can create a subset  $B \subseteq A$  with high  $R(d/B)$  – they may represent complementary knowledge about decision which should be put together while constructing decision rules. The problems of finding approximate reducts are in general NP-hard [10]. Therefore, even for the case of  $A=(U, A \cup \{d\})$  with ten attributes  $A = \{mag_{T1}, lap_{T1}, nbr_{T1},$

$msk, mag_{T2}, lap_{T2}, nbr_{T2}, mag_{PD}, lap_{PD}, nbr_{PD}\}$  described in the previous section, it means that one would prefer to consider the use of a heuristic rather than an exhaustive search. We extend *order based genetic algorithm* (o-GA), developed for searching for minimal reducts in [16], to find (sub)optimal reducts specified by Definition 1. We follow the way of extension proposed in [11] for finding reducts approximately preserving the measure of information entropy. As a *hybrid algorithm* [2], our o-GA consists of two parts:

- *Genetic part*, where each chromosome encodes a permutation of attributes
- *Heuristic part*, where permutations  $\tau$  are put into the following algorithm:

$(R, \varepsilon)$ -REDORD algorithm [12, 16]:

1. For  $A=(U, A \cup \{d\})$  and  $\tau: \{1, \dots, |A|\} \rightarrow \{1, \dots, |A|\}$ , let  $B_\tau = A$ ;
2. For  $i = 1$  to  $|A|$  repeat steps 3 and 4;
3. Let  $B_\tau \leftarrow B_\tau \setminus \{a_{\tau(i)}\}$ ;
4. If  $B_\tau$  does not satisfy condition (3), undo step 3.

We define fitness of a given permutation-individual  $\tau$  due to the quality of  $B_\tau$  resulting from  $(R, \varepsilon)$ -REDORD. The reduct quality is usually based on its length (cf. [6, 16]). Therefore, we use the following measure for a reduct [16]:

$$\text{Fitness}(\tau) = 1 - \text{card}(B_\tau) / \text{card}(A) \quad (5)$$

To work on permutations, we use the order cross-over (OX) and standard mutation switching randomly selected genes [2, 8]. Results are *always*  $(R, \varepsilon)$ -approximate decision reducts, i.e. satisfy (4) and cannot be further reduced without its failure.

#### C. Dynamic Parameter Optimization

As an alternative to tuning parameters for each attribute separately, we develop Genetic Algorithm  $GA_{\text{param}}$  analyzing the space of parameter configurations and choosing the one corresponding to the best-found accuracy.  $GA_{\text{param}}$  is in some sense a hybridization of evolutionary-like strategy for finding peaks of frequency histograms with o-GA searching for  $(R, \varepsilon)$ -approximate decision reducts. The main idea is to be able to search for the reducts in parallel to the parameter tuning.

A chromosome encodes permutation of attributes and uses it similarly as in case of  $(R, \varepsilon)$ -REDORD algorithm. Every gene represents, besides an attribute, some parameters influencing its values' specification. They can refer, e.g., to  $\mu$  in case of magnitude attributes,  $\beta$  for Laplacian-related attributes, and  $rad$  for the mask.

The fitness of an individual corresponds to segmentation accuracy induced by an  $(R, \varepsilon)$ -approximate decision reduct obtained from the encoded permutation, over the training set. Chromosomes are initialized as permutations with randomly picked values taken from the value sets allowed for particular parameters. Permutation-related genetic operators designed for o-GA act in parallel to those concerning parameters, referring to the basic idea of evolutionary strategies.

Although the next section presents results of order-based genetic algorithms working over constant attributes, tuned manually before starting the evolution process, our initial experiments show that the above extension can improve accuracy and further optimize the MRI-related parameters.

### III. RESULTS

Experimental results are obtained for 300 segmentation test cases. For each test a training set was generated using 10 random brain slices chosen from the slice range (61-130) in SBD database. For each thickness 1/3/5/7/9mm and the noise levels (noise/INU) 3/20, 3/40, 9/20, and 9/40, we performed the classification tests based on the following procedure:

1. Generate  $(R, \epsilon)$ -approximate decision reducts using o-GA based on  $(R, \epsilon)$ -REDORD for a given  $\epsilon \in [0, 1]$ ;
2. For each obtained  $(R, \epsilon)$ -approximate decision reduct  $B \subseteq A$  generate rules with conditions induced by B and its values in the universe;
3. Sort decision rules according to their support, in order to choose the most significant rules which recognize each given object;
4. For a new object choose  $\alpha$  most significant applicable rules (where  $\alpha$  is yet another algorithm's parameter);
5. Choose a decision, which is overall best supported within the set of the most significant rules.

The above procedure is more challenging when applied to testing the slices with a higher thickness/noise/INU level, as illustrated by Figure 3. On the other hand, the approximate rough set classifier often performs better when trained over slices from such more difficult levels, as shown in Table I.

More specifically, for each of thickness/noise/INU levels, we considered 15 tests parameterized by the choice of  $\epsilon$  and  $\alpha$  as illustrated by Table II. The results were obtained by testing 20 random slices (range 61-130) from the same image volumes (across all 3 imaging modalities). For the slice thicknesses higher than 1mm we tested on all given slices within the specified range above because of decreasing number of slices when thickness increases.

For each level of  $(\epsilon, \alpha) \in ([0, 1], [0, 1])$  we calculated the average length of  $(R, \epsilon)$ -approximate reducts obtained for various data sets. By varying from  $\epsilon = 0.002$  to  $0.010$ , this length decreases from 3.82 to 2.35, without dramatic changes in the average accuracy (between 80.70% and 81.12% on average for various settings of  $\alpha$ ). It shows that using o-GA with  $(R, \epsilon)$ -REDORD algorithm enables us to reduce thoroughly the number of attributes necessary for reasonable classification.

The results also show that the reduction procedure can yield better classification model than in case of selection of attributes, which seem to provide the highest information gain separately. Attributes with high values of measure (4) are not necessarily those frequently occurring in the reducts and, therefore, in rules used for classification. For  $\epsilon = 0.008$  which has the best average accuracy level, attributes *msk*, *nbrT2*, *magPD* lose their importance, although their relative gain

values are higher than, e.g., that of *lapT1*, which starts to be crucial in decision rules. Moreover, for  $\epsilon \geq 0.008$ , the most frequently occurring  $(R, \epsilon)$ -approximate reducts were always  $\{magT1\}$  and  $\{lapT1, nbrT1, magT2\}$ . This suggests that *T1* and *T2* images provide sufficient attributes for MRI segmentation.

TABLE I  
RESULTS COMPARISON

Thickness	ANN Accuracy <sup>a</sup>	Our Accuracy (o-GA) <sup>b</sup>	
1	90.0%	90.0%	91.9%
3	84.0%	87.1%	89.3%
5	74.0%	82.8%	86.2%

Noise	ANN Accuracy <sup>a</sup>	Our Accuracy (o-GA) <sup>b</sup>	
0	89.0%	90.0%	91.9%
3	91.0%	95.5%	95.7%
5	84.0%	92.8%	92.0%
7	84.0%	88.0%	85.3%
9	81.0%	82.4%	74.2%

INU	ANN Accuracy <sup>a</sup>	Our Accuracy (o-GA) <sup>b</sup>	
0	90.0%	90.0%	91.9%
20	91.0%	94.0%	95.2%
40	86.0%	93.3%	92.4%

<sup>a</sup> Back-propagation Artificial Neural Network (ANN) Results taken from SBD (<http://www.bic.mni.mcgill.ca/users/rkwan/vasco.html>).

<sup>b</sup> "Our Accuracy" for 3% noise and 20% INU used as a training set (1st sub-column), and 0% noise and 0% INU used as a training set (2nd sub-column).

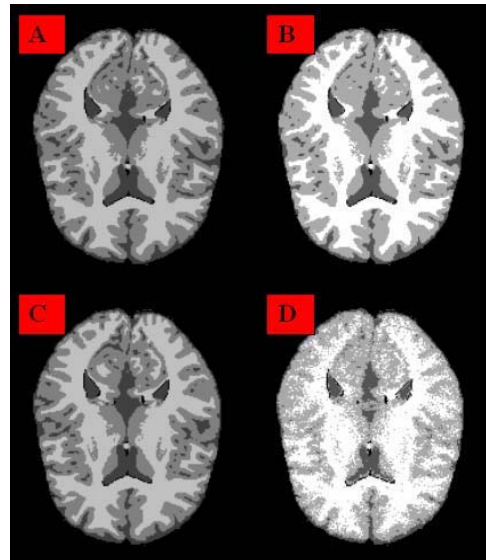


Fig. 3 Phantom image for 1mm slice thickness (Picture A), our result for a 1 mm thickness (B), phantom image for a 9mm slice (C), our result for 9 mm slice thickness (D).

Table II  
RESULTS ACCORDING TO  $\epsilon$  AND  $\alpha$  PARAMETERS

Thick- ness (mm)	$\epsilon$	3% noise / 20% INU (test accuracy in %)			3% noise / 40% INU (test accuracy in %)		
		$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
1		$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
	0.002	94.65	94.67	94.42	91.26	<b>92.61</b>	91.90
	0.004	94.83	94.73	95.05	91.23	92.38	91.70
	0.006	94.46	94.79	94.73	91.53	91.25	91.93
	0.008	94.53	94.82	<b>95.10</b>	91.56	91.95	91.78
	0.010	94.45	94.71	94.90	91.72	91.93	92.50
3		$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
	0.002	87.76	87.80	<b>88.54</b>	89.69	89.76	89.72
	0.004	87.80	87.81	87.86	89.67	89.68	89.71
	0.006	87.80	88.35	88.35	89.64	<b>90.00</b>	<b>90.00</b>
	0.008	87.80	88.24	88.34	89.64	<b>90.00</b>	<b>90.00</b>
	0.010	87.80	88.35	88.35	89.64	89.88	89.89
5		$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
	0.002	75.19	83.12	81.94	80.74	81.34	<b>81.39</b>
	0.004	83.23	83.72	83.72	80.84	81.33	81.33
	0.006	83.23	83.60	83.23	80.84	81.33	81.33
	0.008	83.23	83.72	83.72	80.84	81.33	81.33
	0.010	83.23	83.60	<b>83.72</b>	80.84	81.32	80.98
7		$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
	0.002	78.77	78.77	78.77	64.66	67.89	67.90
	0.004	78.84	79.25	79.25	71.89	71.52	70.29
	0.006	78.84	79.25	79.25	72.27	<b>72.78</b>	<b>72.78</b>
	0.008	78.84	79.25	79.20	72.27	72.73	72.78
	0.010	78.84	79.25	<b>79.25</b>	72.27	72.78	72.42
9		$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
	0.002	77.21	77.22	<b>79.19</b>	71.03	71.06	71.06
	0.004	77.20	77.21	77.31	71.27	<b>71.67</b>	<b>71.67</b>
	0.006	77.22	77.54	77.54	71.27	71.55	71.55
	0.008	77.22	77.54	77.54	71.27	71.67	71.67
	0.010	77.21	77.33	77.22	71.27	71.55	71.65

Thick- ness (mm)	$\epsilon$	9% noise / 20% INU (test accuracy in %)			9% noise / 40% INU (test accuracy in %)		
		$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
1		$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
	0.002	83.54	82.54	83.04	76.72	76.44	77.81
	0.004	82.92	82.31	82.20	78.08	77.09	75.63
	0.006	82.30	81.84	81.36	<b>78.72</b>	76.39	76.46
	0.008	82.68	82.46	82.95	77.47	77.72	74.37
	0.010	82.09	<b>84.35</b>	81.42	76.89	77.19	78.34
3		$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
	0.002	83.77	84.01	84.01	80.96	81.12	<b>81.61</b>
	0.004	84.06	84.19	84.10	79.04	79.08	79.08
	0.006	84.06	84.03	84.03	79.04	79.08	79.08
	0.008	84.07	84.04	<b>84.19</b>	79.05	78.89	79.06
	0.010	84.06	84.03	84.03	79.01	76.40	78.92
5		$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
	0.002	<b>82.73</b>	81.91	81.91	80.31	80.31	80.31
	0.004	81.66	81.62	81.62	80.33	80.37	80.37
	0.006	81.66	81.85	81.85	80.36	80.26	80.23
	0.008	81.66	81.69	81.85	80.33	80.37	80.37
	0.010	81.66	80.54	80.54	80.32	79.90	<b>80.56</b>
7		$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
	0.002	78.75	78.69	78.46	77.52	77.52	77.52
	0.004	78.85	78.79	78.85	77.81	78.03	78.03
	0.006	78.85	<b>79.06</b>	78.89	77.81	78.03	78.03
	0.008	78.85	78.79	78.85	77.52	77.52	77.52
	0.010	78.86	78.01	78.01	77.81	77.34	<b>78.04</b>
9		$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
	0.002	75.29	75.66	75.40	74.50	74.67	74.74
	0.004	76.49	76.65	76.56	74.82	74.90	74.90
	0.006	76.49	76.65	76.65	74.82	74.82	74.82
	0.008	76.49	76.54	<b>76.67</b>	74.82	75.02	<b>75.18</b>
	0.010	76.36	76.08	76.55	74.54	74.51	74.95

#### IV. CONCLUSIONS AND DISCUSSION

Automated image segmentation must have a reliable level of accuracy across all tissue classes. Our results indicate that with a reasonable amount of noise (3-5%), typical field inhomogeneities (20% or less), and a reasonable slice thickness (3-5mm), our approximate reduct algorithm is capable of yielding segmentation accuracy on the order of 90%+, consistent – if not more accurate than other approaches. Not only does our algorithm achieve high segmentation accuracy, but it works across all three major imaging modalities ( $T1$ ,  $T2$ ,  $PD$ ). This is a contrast to most other segmentation algorithms which classify on a single modality [5, 7, 17].

The results are obtained without any image preprocessing such as median filtering and other smoothing operations. We are pursuing investigating the segmentation accuracy when various standardized filtering and averaging processes are embedded into the algorithm. The results indicate that the major impediment to accurate segmentation is slice thickness and noise. They display a reduction of just 18% from the best to the worst case. This problem may not be solved with technological advances, but may instead, require a fresh computational perspective, one such as that provided by the rough set theory.

There are numerous issues that should be additionally taken into account. First of all, we considered only healthy brain images with constant number of histogram peaks. In future, we will focus on pathologies and extend the system to read most of known MRI data standards. Secondly, our algorithm is a supervised method. In a real live data, however, we would not have a known segmented phantom. This could be resolved by using more training data sets generated from, e.g., some brain atlas data sets.

Finally, we provided results only for manually specified parameters, which – although justified by expert knowledge – do not need to induce optimal attributes. For instance, images with high INU have a general “speckled” appearance. This may be due in part to the variance threshold  $\beta$  applied to the discrete Laplacian function. Therefore, we are going to follow the idea presented in the last part of Section III and tune the attribute parameters dynamically, in parallel to optimization of the rough set based segmentation model.

#### ACKNOWLEDGMENT

We would like to thank Mr. Lukasz Budkiewicz for help in preprocessing the MRI data, as well as Dr. Jakub Wroblewski for valuable comments on implementation and usage of the order based genetic algorithms

#### REFERENCES

- [1] Cocosco, C.A., Zijdenbos, A.P., Evans, A.C.: Automatic Generation of Training Data for Brain Tissue Classification from MRI. In: Proc. of ICCAI'2002 (2002).
- [2] Davis, L. (ed.): Handbook of Genetic Algorithms. Van Nostrand Reinhold (1991).
- [3] Kamber, M., Shinghal, R., Collins, L.: Model-based 3D Segmentation of Multiple Sclerosis Lesions in Magnetic Resonance Brain Images. IEEE Trans Med Imaging 14(3) (1995) pp. 442–453.
- [4] Kaus, M., Warfield, S.K., Nabavi, A., Black, P.M., Jolesz, F.A., Kikinis, R.: Automated Segmentation of MRI of Brain Tumors. Radiology 218 (2001) pp. 586–591.
- [5] Kollokian, V.: Performance Analysis of Automatic Techniques for Tissue Classification in Magnetic Resonance Images of the Human Brain. Master's thesis, Concordia University, Montreal, Canada (1996).
- [6] Komorowski, J., Pawlak, Z., Polkowski, L., Skowron, A.: Rough sets: A tutorial. In: S.K. Pal, A. Skowron (eds): Rough Fuzzy Hybridization – A New Trend in Decision Making. Springer Verlag (1999) pp. 3–98.
- [7] Kovacevic N., Lobaugh N.J., Bronskill M.J., Levine B., Feinstein A. and Black, S.E.: A Robust Extraction and Automatic Segmentation of Brain Images. NeuroImage 17 (2002) pp. 1087–1100.
- [8] Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. Springer-Verlag (1994).
- [9] Pawlak, Z.: Rough sets – Theoretical aspects of reasoning about data. Kluwer (1991).
- [10] Slezak, D.: Approximate Entropy Reducts. Fundamenta Informaticae (2002).
- [11] Slezak, D.: The Rough Bayesian Model For Distributed Decision Systems. In: Proc. of RSCTC'2004.
- [12] Slezak, D., Wroblewski, J.: Order-based genetic algorithms for the search of approximate entropy reducts. In: Proc. of RSFDGrC'2003. Chongqing, China (2003).
- [13] Slezak, D., Ziarko, W.: Attribute Reduction in Bayesian Version of Variable Precision Rough Set Model. In: Proc. of RSKD'2003. Elsevier, ENTCS 82(4) (2003).
- [14] Vannier, M.W.: Validation of Magnetic Resonance Imaging (MRI) Multispectral Tissue Classification. Computerized Medical Imaging and Graphics 15(4) (1991).
- [15] Widz, S., Slezak, D., Revett K.: An Automated Multi-spectral MRI Segmentation Algorithm Using Approximate Reducts. In: Proc. of RSCTC'2004.
- [16] Wroblewski, J.: Theoretical Foundations of Order-Based Genetic Algorithms. Fundamenta Informaticae 28(3-4) (1996) pp. 423–430.
- [17] Xue J.H., Pizurica A., Philips W., Kerre E., Van de Walle R., Lemahieu, I.: An Integrated Method of Adaptive Enhancement for Unsupervised Segmentation of MRI Brain Images. Pattern Recognition Letters, Vol 24(15) (2003) pp. 2549–2560.
- [18] Zijdenbos, A.P., Dawant, B.M., Margolin, R.A., Palmer, A.C.: Morphometric Analysis of White Matter Lesions in MR Images: Method and Validation. IEEE Trans. Med. Imaging 13(4) (1994) pp. 716–724.