

# An Improved C-Means Model for MRI Segmentation

Ying Shen, Weihua Zhu

**Abstract**—Medical images are important to help identifying different diseases, for example, Magnetic resonance imaging (MRI) can be used to investigate the brain, spinal cord, bones, joints, breasts, blood vessels, and heart. Image segmentation, in medical image analysis, is usually the first step to find out some characteristics with similar color, intensity or texture so that the diagnosis could be further carried out based on these features. This paper introduces an improved C-means model to segment the MRI images. The model is based on information entropy to evaluate the segmentation results by achieving global optimization. Several contributions are significant. Firstly, Genetic Algorithm (GA) is used for achieving global optimization in this model where fuzzy C-means clustering algorithm (FCMA) is not capable of doing that. Secondly, the information entropy after segmentation is used for measuring the effectiveness of MRI image processing. Experimental results show the outperformance of the proposed model by comparing with traditional approaches.

**Keywords**—Magnetic Resonance Image, C-means model, image segmentation, information entropy.

## I. INTRODUCTION

MRI refers to a medical photographing technology which uses radiology techniques to form images of the anatomy and physiology processes [1]. The images have the both health and disease bodies through making full use of strong magnetic fields, radio waves, and field gradients [2]. For taking an MRI image, a strong magnetic field will be used and radio waves are applied for a short time in different directions. This sudden shift results in certain atoms in the patient's body to make special signals which could be detected by the MRI scanners. After that, a computer creates an image of the inner body by using the received signal information so that doctors can diagnose disorders of the body [3]. Therefore, MRI can be used to investigate the brain, spinal cord, bones, joints, breasts, blood vessels, heart, etc. [4]. However, some metallic objects such as hip, shoulder and knee replacements will distort the images.

Image segmentation, in medical image analysis, is usually the first step which plays an important role in many diagnosis applications with the aims of dividing an image into several different areas/regions where similar color or intensity could be observed. Thus, image segmentation has been widely used for analyzing and visualizing the anatomical structures delineating pathological areas in brain MRI diagnosis which is significant for treating some patients with serious diseases on their brains [5], [6]. In the previous decades, various

segmentation techniques have been developed and reported in the literature according to the development of ad-hoc technologies such as Internet of Things and Big Data Analytics [7]-[10]. Brain MRI segmentation is a critical task in many applications because it influences the outcome of the entire diagnosis. Therefore, there are many different processing steps which are based on accurate segmentation of anatomical regions [11], [12]. Take MRI image segmentation for example, this process is widely used for evaluating and diagnosing different brain structures. Thus, design and development of the segmentation approaches should be based on different applications [13]. Some considerations such as complexities, accuracies, and effectiveness are concerned during this process. In realizing brain injury and exploring brain anatomy using MRI images, some progresses are achieved such as using advanced models and algorithms to help making decisions. The advances in brain MR imaging have also provided vast number of data with an increasingly high level of quality. Analysis of these large and complex MRI data becomes a complex task for clinicians or doctors, who have to manually extract important information. This manual operation is often time-consuming and prone to errors due to various inter- or intra-operator variability studies [14]. These challenges in brain MRI data analysis open up an urgent requirement invention in computerized methods to improve disease diagnosis and testing. Currently, computerized methods for MR image segmentation, registration, and visualization have been used to assist doctors in qualitative diagnosis [15].

Unfortunately, there are still several difficulties in doing the MRI segmentation. Firstly, most of the research uses GA to achieve the global optimum in MRI segmentation. However, the computational time is high so that it is difficult to implement in practical cases. Secondly, C-means clustering algorithm has been widely used in literature. But, it heavily relies on the initial clustered center so that global optimum is hardly achieved. Finally, the information degree which is used to measure the results is scarcely reported.

In order to address the above difficulties, this paper introduces an improved C-means model using information entropy for MRI segmentation. This model extends the classic C-means into a higher space so as to cluster the MRI images by making full use of GA approach which achieves the global optimum. Additionally, based on the extension, MRI segmentation is carried out and the information entropy is proposed to measure the results so that the quality of segmentation could be ensured.

The rest of this paper follows the routes: Section II reports on the improved C-means Model for MRI segmentation. It highlights the information entropy measurement after the segmentation. Section III demonstrates the experiments and

Ying Shen is an associate professor in Department of Engineering, Xinyu College, Xinyu, Jiangxi, China (e-mail: Erinshen2010@163.com).

Weihua Zhu is an associate professor in Department of Computer Science, Xinyu College, Xinyu, Jiangxi, China (corresponding author, phone: +86-0790-6666117; e-mail: 969257518@qq.com).

discussions. Section IV concludes this paper by giving our contributions and future research directions.

## II. AN IMPROVED C-MEANS MODEL FOR MRI SEGMENTATION

### A. Classic C-Means

In fuzzy clustering which is usually talked in fuzzy theory, clustering involves allocating data points to different clusters so that they may have the similarity or same characteristics as possible [16]. When we use this approach in image processing, FCMA is widely adopted. Assume that any point  $x$  has a set of coefficients which give the degree of being in the  $k^{\text{th}}$  cluster  $w_k(x)$ , the centroid of a cluster  $c_k$  is the mean of all points:

$$c_k = \frac{\sum_x w_k(x)^m x}{\sum_x w_k(x)^m} \quad (1)$$

FCMA aims to minimize an objective function which divides an image into clusters [17]:

$$\arg \min_c \sum_{i=1}^n \sum_{j=1}^c w_{ij}^m \|x_i - c_j\|^2 \quad (2)$$

$$w_{ij} = \frac{1}{\sum_{k=1}^c \left( \frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \quad (3)$$

From (2), for a specific MRI image, we can get:

$$J(u, v) = \sum_{k=1}^C \sum_{i=1}^N [u_k(x_i)]^m \|x_i - v_k\|^2 \quad (4)$$

where  $N$  is the total number of samples.  $C$  is the cluster number.  $m > 1$  is a constant number of fuzzy degree which is used for controlling the cluster results.  $x_i$  is  $P$  dimensional vector.  $v_k$  is the  $i$ th cluster center.  $u_k(x_i)$  is the membership function of  $k$ th sample in  $i$ th cluster.

### B. Improved Model

In the previous FCMA method, the membership grade of each sample in individual cluster could be obtained when the algorithm is converging. That aims to achieve the fuzzy clustering for the images. However, the convergence speed is fast due to the iterations. That heavily relies on the selection of initial cluster center [18]-[20]. Therefore, the local optimization is achieved rather than global optimization. An improved model is necessary for achieving global optimization which will make better segmentation results.

To this end, this paper uses GA approach twining with

FCMA. GA is able to make global optimization through initializing operators, selection, crossover, mutation, accepting, replacement, and testing [21]. By using the GA in the FCMA, the following steps are used for the improved model.

- Confirm the Cluster quantity  $C$ ;
  - Initialization: A population of  $P$  individual is randomly generated. Each individual represents the center  $\{v_1, v_2, \dots, v_c\}$  of each  $C$ .  $v_i$  is a  $P$  vector.
  - Crossover: Assume that the crossover probability is  $p_c$ , for each population in  $P$ , a random value  $r$  is created,  $0 \leq r \leq 1$ . If  $r \leq p_c$ , the crossover operations will be conducted. For  $P$ , we can randomly select a population, an integer  $i$  is generated,  $1 \leq i \leq C$ . The location  $j$  will be conducted the crossover operation given  $\{v_1, v_2, \dots, v_c\} - \{v_1, v_2, \dots, v_j, v'_{j+1}, \dots, v'_c\}$  and  $\{v'_1, v'_2, \dots, v'_c\} - \{v'_1, v'_2, \dots, v'_j, v_{j+1}, \dots, v_c\}$ .
- This operation is going to switch the center of  $j$ . It means that the central pixel of  $j$  in the image will be segmented.
- Mutation: Assume that the mutation probability is  $p_m$ , for each selected individual  $\{v_1, v_2, \dots, v_c\}$ , a location  $v_j$  will be randomly chosen for mutation operation. Thus, a new  $v_j$  will be produced as  $v'_j$ .
  - The fitness degree of an individual is:

$$Fit = \frac{1}{1 + J(u, v)} \quad (5)$$

- Compare the new individual after mutation with each individual population, the individual with larger  $Fit$  will be selected and the ones with smaller  $Fit$  will be omitted. A threshold will be followed in applications.

### C. Entropy-Based Measurement

Entropy refers to the expected average value of the information in any data flow so that the message could be examined and measured [22]. The measurement of the message is based on the mean information which shows how much useful data/message is carried out by the information so that the unpredictability of the information/message could be evaluated [23]-[25]. The logarithm of the probability distribution could be used for measuring the entropy. Thus, in this paper, the conditional entropy of two elements  $v_j$  and  $v'_j$  from the improved model is considered to validate the mutation results.

The entropy  $H$  of a discrete random variable  $V$  with possible values  $\{v_1, v_2, \dots, v_c\}$  and the probability mass function  $P(V)$  is:

$$H(V) = E[I(V)] = E[-\ln(P(V))] \quad (6)$$

$E$  is the expected value operator and  $I$  is the information contents of  $V$ . In this paper, the information contents come from the pixels of MRI images. Thus, the entropy could explicitly be written as:

$$H(V) = \sum_{j=1}^C P(v_j) I(v_j) = - \sum_{j=1}^C P(v_j) \log_b P(v_j) \quad (7)$$

$b$  is the base of the logarithm used. In this paper,  $b$  is set as 2 with Shannon entropy. Therefore, we can get the conditional entropy for segmenting purpose using:

$$H(V | V') = \sum_{i,j} p(v_i, v'_j) \log \frac{p(v_i, v'_j)}{p(v'_j)} \quad (8)$$

### III. EXPERIMENT AND DISCUSSION

This section reports on the experiments used for validating the proposed model for MRI segmentation. The images are from National Alliance for Medical Image Computing which is a free and open medical image source, including large number of images for researchers and practitioners. The images are 136×152px. Fig. 1 shows the experiment results.

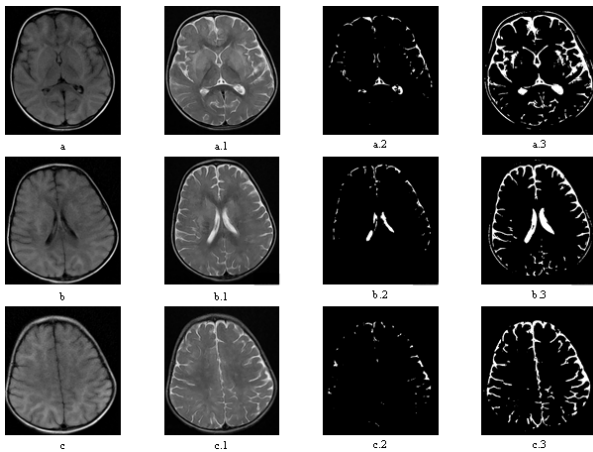


Fig. 1 Experiment Results

Fig. 1 presents several images where a, b, and c are the original images from MRI. The second column a.1, b.1, and c.1 are the results from using noise cleansing algorithm in [21] which aims to remove most of the noises and to smoothen the original images. It could be observed that, after using the cleansing method, the images are much smoother and some features are much clearer such as the alba, ectocinerea, and neurolymph. For example, from a and a.1, some tissues are easily identified and the characteristics such as veins could be observed so that doctors may easily carry out the diagnosis. The third column images are the results using traditional FCMA approach and the last column are the results using the

proposed model in this paper.

In the experiments, the proposed model has the following parameters: the population is 80. According to the main tissues in MRI images like alba, ectocinerea, and neurolymph,  $C$  is set as 3. The iteration loop is 100 times. Crossover possibility and mutation possibility are  $p_c = 0.7$  and

$p_m = 0.1$ .

TABLE I  
QUANTITATIVE ANALYSIS

Group	Experimental Results				
	Images	MSE 1	MSE 2	Time	Entropy
a	a.2	9.18	9.05	26 vs 36	0.29
	a.3	9.26	9.11	28 vs 32	0.68
b	b.2	10.15	10.02	36 vs 44	0.36
	b.3	10.23	10.10	38 vs 50	0.85
c	c.2	12.28	11.88	44 vs 62	0.26
	c.3	11.81	10.84	48 vs 72	0.82

The segmentation results from Fig. 1 show that the proposed model is able to segment the MRI images effectively compared with the traditional FCMA approach. For example, most of the key features could not be observed by traditional FCMA because local optimization is used. However, using GA approach to achieve a global optimization, more features are identified with more significant information provided for decision-makings such as disease diagnosis. Fig. 1 presents the qualitative analysis by comparing the experimental segmentation results. Then, Table I shows the quantitative analysis of the experiments.

Table I presents the quantitative analysis which compares the grayscale values before and after the segmentation using difference approaches. The first column is the group which uses three different MRI images for the experiments. Six individual MRI images are presented. MSE 1 and 2 are the mean square error (MSE) values which are obtained from traditional FCMA and the proposed improved model. Time indicates the computational time/cost that each approach uses for achieving the results. The last column presents the entropy value from the definition that is used to evaluate the information carrying after the segmentation processing.

From Table I, it could be observed that values of MSE 2 are much smaller than those of MSE 1. That implies the outperformance of the proposed model comparing with traditional FCMA approach. The proposed model spends more computational cost for achieving better performance. Averagely speaking, it takes 32.9% of more time to conduct the segmentation comparing with traditional FCMA method. That is because, in the proposed model, GA approach is used for achieving global optimization which requires calculation time for crossover and mutation, which attributes to the more time cost. Despite the fact that the more time will be required, the information carried after the segmentation is much more significant compared with that from the traditional FCMA. That is why the proposed model is able to identify more features after segmentation which could be observed in Fig. 1.

## IV. CONCLUSION

This paper introduces an improved C-means model for MRI segmentation. This model uses information entropy to measure the segmentation results. GA method is integrated to the FCMA approach so as to achieve global optimization. Experiments from this study show the outperformance of the proposed model by comparing traditional FCMA method. This paper has several contributions. Firstly, GA approach is used to get the global optimization so as to avoid the heavy dependence on initial cluster center in FCMA. Thus, better performance could be achieved in the proposed model. Secondly, information entropy is adopted to evaluate the effectiveness of segmentation. The information contents come from the pixels of MRI images where the key features could be observed by using the entropy.

Future research directions will be carried out to improve this study. First of all, how to improve the efficiency of GA so that less computational cost could be applied? In order to address this question, parallel approaches will be studied in the future to divide the GA calculation by using different parameters. Secondly, how to validate the proposed model by comparing with other approaches, such as ACO? In the future, more experimental scenarios will be conducted to comparing this model with different approaches so that more effective and efficient segmentation methodologies could be figured out. Thus, practical applications with suitable methodologies will be applied.

## ACKNOWLEDGMENT

Authors would like to acknowledge 2016 Department of Education Jiangxi Science and Technology Fund (GJJ161183).

## REFERENCES

- [1] R. D. Nowak, "Wavelet-based Rician noise removal for magnetic resonance imaging," *IEEE Transactions on Image Processing*, vol. 8, pp. 1408-1419, 1999.
- [2] N. Weiskopf, K. Mathiak, S. W. Bock, F. Scharnowski, R. Veit, W. Grodd, et al., "Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI)," *IEEE transactions on biomedical engineering*, vol. 51, pp. 966-970, 2004.
- [3] M. Lysaker, A. Lundervold, and X.-C. Tai, "Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time," *IEEE Transactions on image processing*, vol. 12, pp. 1579-1590, 2003.
- [4] K. Ugurbil, "Magnetic resonance imaging at ultrahigh fields," *IEEE Transactions on Biomedical Engineering*, vol. 61, pp. 1364-1379, 2014.
- [5] I. Despotović, B. Goossens, and W. Philips, "MRI segmentation of the human brain: challenges, methods, and applications," *Computational and mathematical methods in medicine*, vol. 2015, 2015.
- [6] R. Y. Zhong, Q. Dai, T. Qu, G. Hu, and G. Q. Huang, "RFID-enabled real-time manufacturing execution system for mass-customization production," *Robotics and Computer-Integrated Manufacturing*, vol. 29, pp. 283-292, 2013.
- [7] R. Y. Zhong, G. Q. Huang, S. L. Lan, Q. Y. Dai, C. Xu, and T. Zhang, "A Big Data Approach for Logistics Trajectory Discovery from RFID-enabled Production Data," *International Journal of Production Economics*, vol. 165, pp. 260-272, 2015.
- [8] M. Evertsson, P. Kjellman, M. Cinthio, S. Fredriksson, R. in't Zandt, H. Persson, et al., "Multimodal detection of iron oxide nanoparticles in rat lymph nodes using magnetomotive ultrasound imaging and magnetic resonance imaging," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 61, pp. 1276-1283, 2014.
- [9] E. S. H. Ibrahim, R. A. Pooley, M. D. Bridges, J. G. Cernigliaro, and W. E. Haley, "Kidney stone imaging with 3D ultra-short echo time (UTE) magnetic resonance imaging. A phantom study," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 2014, pp. 2356-2359.
- [10] R. Y. Zhong, S. T. Newman, G. Q. Huang, and S. L. Lan, "Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives," *Computers & Industrial Engineering*, vol. 101, pp. 572-591, 2016.
- [11] W. H. Zhu and Y. Shen, "A segmentation approach for tissue images using non-dominated sorting GA," in *Anti-counterfeiting, Security, and Identification (ASID), 2016 10th IEEE International Conference on*, 2016, pp. 1-5.
- [12] T. Chitiboi and L. Axel, "Magnetic resonance imaging of myocardial strain: A review of current approaches," *Journal of Magnetic Resonance Imaging*, 2017.
- [13] A. Makropoulos, I. S. Gousias, C. Ledig, P. Aljabar, A. Serag, J. V. Hajnal, et al., "Automatic whole brain MRI segmentation of the developing neonatal brain," *IEEE transactions on medical imaging*, vol. 33, pp. 1818-1831, 2014.
- [14] R. Y. Zhong, S. Lan, C. Xu, Q. Dai, and G. Q. Huang, "Visualization of RFID-enabled shopfloor logistics Big Data in Cloud Manufacturing," *The International Journal of Advanced Manufacturing Technology*, vol. 84, pp. 5-16, April 2016.
- [15] S. Roy, Q. He, E. Sweeney, A. Carass, D. S. Reich, J. L. Prince, et al., "Subject-specific sparse dictionary learning for atlas-based brain MRI segmentation," *IEEE journal of biomedical and health informatics*, vol. 19, pp. 1598-1609, 2015.
- [16] M. Gong, Y. Liang, J. Shi, W. Ma, and J. Ma, "Fuzzy c-means clustering with local information and kernel metric for image segmentation," *IEEE Transactions on Image Processing*, vol. 22, pp. 573-584, 2013.
- [17] S. L. Lan, H. Zhang, R. Y. Zhong, and G. Q. Huang, "A customer satisfaction evaluation model for logistics services using fuzzy analytic hierarchy process," *Industrial Management & Data Systems*, vol. 116, pp. 1024-1042, 2016.
- [18] E. Özceylan, M. Kabak, and M. Dağdeviren, "A fuzzy-based decision-making procedure for machine selection problem," *Journal of Intelligent & Fuzzy Systems*, vol. 30, pp. 1841-1856, 2016.
- [19] M. L. Wang, R. Y. Zhong, Q. Y. Dai, and G. Q. Huang, "A MPN-based scheduling model for IoT-enabled hybrid flow shop manufacturing," *Advanced Engineering Informatics*, vol. 30, pp. 728-736, 2016.
- [20] R. Y. Zhong, G. Q. Huang, S. L. Lan, Q. Y. Dai, T. Zhang, and C. Xu, "A two-level advanced production planning and scheduling model for RFID-enabled ubiquitous manufacturing," *Advanced Engineering Informatics*, vol. 29, pp. 799-812, 2015.
- [21] Y. F. Zhang, G. Zhang, T. Qu, Y. Liu, and R. Y. Zhong, "Analytical target cascading for optimal configuration of cloud manufacturing services," *Journal of Cleaner Production*, vol. 151, pp. 330-343, 2017.
- [22] S. Bruers, "A discussion on maximum entropy production and information theory," *Journal of Physics A: Mathematical and Theoretical*, vol. 40, p. 7441, 2007.
- [23] C. H. Liu, R. Y. Zhong, Y. E. Yan, and X. Hu, "CEP-Based Massive Data Processing Approach for RFID Data," *Advanced Materials Research*, vol. 317, pp. 350-353, 2011.
- [24] R. Y. Zhong, G. Q. Huang, and Q. Y. Dai, "A Big Data Cleansing Approach for n-dimensional RFID-Cuboids," *Proceeding of the 2014 IEEE 18th International Conference on Computer Supported Cooperative Work in Design (CSCWD 2014)*, 21-23 May, Taiwan, pp. 289-294, 2014.
- [25] S. P. Lu, C. Xu, and R. Y. Zhong, "An Active RFID Tag-Enabled Locating Approach with Multipath Effect Elimination in AGV," *IEEE Transactions on Automation Science and Engineering*, vol. 13, pp. 1333-1342, 2016.

**Ying Shen** is an associate professor from department of engineering, Xinyu College, got the master degree of communication and information system in 2003 from Nan Chang University. She has been published a rich set of related papers in international journals and conferences. Her main research area is digital image processing.

**Weihua Zhu** is a senior lecture from department of Computer Science, got the master degree of communication and information system from Nan Chang University. He has been published a large number of articles in international

journals and conferences. His main research areas are algorithm design and embedded system.