An Improved K-means Algorithm for Brain MRI Image Segmentation

Jianwei Liu[1, a], Lei Guo[1, b *]

[1]School of Automation, Northwestern Polytechnical University, Xi'an, Shanxi, 710072, China

[a]xatu07@163.com, [b]lguo@nwpu.edu.cn

Keywords: Magnetic Resonance Imaging (MRI)， Segmentation， K-means

Abstract. For the problem of low accuracy by the traditional K-means clustering algorithm to segment noised brain magnetic resonance imaging (MRI) images. This paper proposed an improved K-means algorithm. The traditional K-means algorithm only considers the brain image gray value itself, ignoring the relationship between pixels. Due to the characteristics of brain MRI image adjacent pixels most likely belonging to the same class, this paper adopts average value of small neighborhood of each image pixel and image gray value to compose a new sample point, in order to reduce the impact of noise on the clustering accuracy. Experimental results show that the improved K-means algorithm can effectively improve the segmentation accuracy of the noised brain MRI image.

Introduction

Brain image segmentation is one of important contents of medical image processing, especially accurate segmentation will benefit clinicians and patients, which has important meaning for clinical diagnostic. Most of research in medical image segmentation uses magnetic resonance imaging (MRI) images because of high contrast. Accurate brain tissue segmentation can improve reliability of brain disease diagnosis and effectiveness of treatment. Brain MRI images mostly contain noise, intensity non-uniformity, etc., therefore, accurate segmentation of brain MRI images is a very difficult task. Due to clustering methods do not need a large number of training samples and prior knowledge, using the distance and distribution of sample to category, clustering algorithms becomes a popular brain MRI image segmentation tools. Clustering methods usually adopt the optimization iteration method to obtain the best image segmentation results to maintain the maximum distance between different classes and within the maximum similarity in the same class. The K-means algorithm as a kind of classic unsupervised learning clustering algorithm has been widely used because of its advantages, such as simple, efficient, easy to implement and so on[1, 2, 3].

Brain MRI image usually is partitioned into white matter (WM), gray matter (GM), cerebrospinal fluid (CSF) and the background by the traditional K-means algorithm, but K-means clustering algorithm only considers brain image gray value itself, ignoring the relationship between pixels, and obtains low brain segmentation accuracy especially for low signal to noise ratio (SNR) data. Due to the characteristics of brain MRI image adjacent pixels most likely belonging to the same class, this paper adopts average value of small neighborhood of each image pixel and image gray value to compose a new sample point, in order to reduce the impact of noise on the clustering accuracy. Experiment results show that the improved K-means algorithm effectively improves segmentation accuracy of the noised brain MRI image.

The rest of the paper is organized as follows. Improved K-means clustering algorithm is given in section 2. Section3 presents the segmentation results and discussion. Finally, the conclusion and acknowledgement follows.

Improved K-means algorithm

K-means algorithm is a classical algorithm to solve the clustering problem proposed by MacQueen in 1967, and is also a simple and efficient clustering algorithm. It can make the minimum square distance between all point in clustering domain and the cluster center[4,5].

K-means algorithm classifies a sample set $X(x_1, x_2, \cdots, x_n)$ into k clusters with the aim at minimizing an objective function, where n is the number of samples and C_i represents the clustering center, $C_i = \frac{1}{N_i}\sum_{x \in x_i} x$, $i = 1,2,\cdots, k$, N_i is the sample number of the ith cluster x_i .The objective function J is given as:

$$J = \sum_{i=1}^{k}\sum_{j=1}^{n}\left\|x_j^{(i)} - C_i\right\|^2 \tag{1}$$

Where, $\left\|x_j^{(i)} - C_i\right\|^2$ is a distance measure between a data point x_j and the cluster center C_i , the Euclidean distance is adopted as the distance measure in this paper. Therefore, the K-means algorithm is an iterative clustering algorithm that finds a correct classification result which minimizes the sum squared error. The K-means algorithm is described as follows:

(1) Initialize cluster centroids C_i with k random samples;

(2) Assign each sample point x_j to the nearest cluster center.

(3) Recalculating each clustering center: $C_i = \frac{1}{N_i}\sum_{x \in x_i} x$, $i = 1,2,\cdots, k$;

(4) Repeat steps (2) and (3) until C_i no longer changes.

The traditional K-means algorithm makes the brain image gray value as the clustering samples, ignoring the relationship between pixels. Due to the characteristics of brain MRI image adjacent pixels most likely belonging to the same class, this paper adopts average value of small neighborhood of each image pixel to form a new sample point, in order to reduce the influence of noise on the clustering segmentation accuracy. The neighborhood of each pixel can be 3×3 , 5×5 , 7×7 . 3×3 is a good choice through experimental verification. Let $f(i, j)$ represents source image, whose size is $\mathrm{M}\times N$, $g(i, j)$ is the new sample image, then,

$$g(i, j) = \frac{1}{N\times N}\sum_{k=i-1}^{i+1}\sum_{l=j-1}^{j+1} f(i, j)\text{, } i = 1,2,\cdots,M\text{, } j = 1,2,\cdots,N \tag{2}$$

Where, $f(i, j)$ needs to be padded with zeros and $g(i, j)$ needs to trancate zeros in the image boundary. This paper uses the gray value of $f(i, j)$ and $g(i, j)$ to compose new sample points. Differences between the improved algorithm and the traditional K-means algorithm in this paper lies in the rational use gray similar information of the image space as a new sample point. Samples includes not only the gray information of the image, but also the spatial information, effectively compensating for the shortcomings of the K-means algorithm that is sensitive to noise. Therefore, the classification reliability of samples is greatly improved . In this paper, the samples are taken from the normalized column vector $f(:), g(:)$, and normalized formula is as follows:

$$f(:) = \frac{f(:)}{\max(f(:))} \tag{3}$$

$$g(:) = \frac{g(:)}{\max(g(:))} \tag{4}$$

Here, $\max(\bullet)$ represents maximum of column vector.

Results and discussion

In order to evaluate the effectiveness of the improved K-means algorithm, simulated brain MRI images are segmented by improved K-means algorithm and the traditional k-means respectively. In order to quantify the evaluation segmentation effect, this paper adopts the Jaccard similarity (JS) coefficient analyze and compare the segmentation result quantitatively [6], which is defined as follows:

$$J^k(S,T) = \left| \frac{V_S^k \cap V_T^k}{V_S^k \cup V_T^k} \right| \tag{5}$$

Where, V_T^k is the ground truth of class k, and V_S^k is the segmental results, k represents brain tissue classification. The ratio is between 0 and 1, the greater the value is, the better segmentation results get.

The experimental data comes from brain MRI data of McGill Montreal University Neurology Institute (http://www.bic.mni.mcgill.ca/brainweb). The website provides a lot of MRI simulated brain data with a variety of slice thicknesses, noise levels, and levels of intensity non-uniformity. The simulated brain MRI image database is often used as the gold standard for image segmentation. The size of experiment data is 181 x 217 x 181, and its spatial resolution is 1 mm isotropic. The experiment adopts T1-weighted data with 40% intensity non-uniformity and noise level 5%, 7%, 9%. Here, 40% intensity non-uniformity and noise level 9% T1-weighted simulated brain MRI image data as an example shows the segmentation results. Figure 1 is the input image data, Fig.2 is reference image data, Fig.3 is segmentation result by traditional K-means algorithm, and Fig.4 segmentation results by improved K-means algorithm

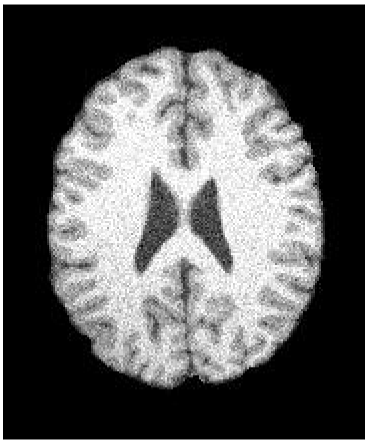

Fig.1 Simulation brain MRI image

Fig. 2 reference image

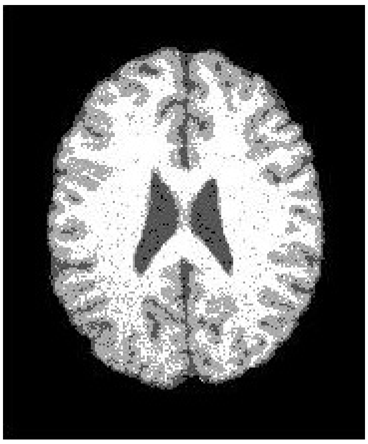

Fig.3 segmentation results by traditional K-means algorithm

Fig. 4 segmentation results by improved K-means algorithm

Table 1 gives evaluation of brain tissue segmentation results of WM, GM, and CSF in terms of Jaccard Similarity Coefficients for the simulated brain MRI images with different noise. As can be seen from Table 1, improved K-means algorithm obtains better segmentation accuracy even at low SNR brain images.

Table 1 Evaluation of Brain Tissue Segmentation accuracy

noise level	Tissue	traditional K-means algorithm	improved K-means algorithm
5%	WM	0.8996	0.9159
	GM	0.8074	0.8116
	CSF	0.7653	0.7508
7%	WM	0.8624	0.8984
	GM	0.7491	0.7908
	CSF	0.7228	0.7504
9%	WM	0.8122	0.8902
	GM	0.6721	0.7818
	CSF	0.6671	0.7549

Conclusion

This paper proposed an improved K-means algorithm, using gray value and the mean of pixel neighborhood of brain MRI image to compose new sample points, effectively compensating for the shortcomings of the K-means algorithm that is sensitive to noise. The experimental results showed that the segmentation accuracy of improved K-means algorithm was superior to the traditional K-means algorithm for noised brain MRI images.

Acknowledgments

This work is fully supported by a grant from the National Natural Science Foundation of China (Project No.61273362).

References

[1] A. Norouzi, M. S. M. Rahim , A. Altameem , et al, Medical image segmentation methods, algorithms, and applications, IETE Technical Review. 31(2014) 199-213.

[2] M.A. Balafar, A.R. Ramli, M.I. Saripan, et al, Review of brain MRI image segmentation methods, Artificial Intelligence Review. 33(2010) 261-274.

[3] A. Ortiz, J.M. Górriz, J.Ramírez, et al, Two fully-unsupervised methods for MR brain image segmentation using SOM-based strategies, Applied Soft Computing. 13(2013) 2668-2682.

[4] K. Somasundaram1, T. Genish, Binarization of MRI with Intensity Inhomogeneity using K-Means Clustering for Segmenting Hippocampus, International Journal of Multimedia and its Applications. 5(2013) 11-18.

[5] M.K. Date, M.S.P. Akarte, Brain image segmentation algorithm using K-means clustering, International Journal of Computer Science and Applications. 6(2013) 285-289.

[6] S.Wang, Z.Geng, J.Zhang, et al. A Fuzzy C-means Model Based on the Spatial Structural Information for Brain MRI Segmentation, International Journal of Signal Processing, Image Processing and Pattern Recognition.7(2014) 313-322.