

文章编号 1004-924X(2017)增-0259-07

脑磁共振图像中改进的模糊聚类分割方法研究

潘宏亮, 郭凯文, 刘丽伟, 李 阳, 侯阿临*
(长春工业大学 计算机科学与工程学院, 吉林 长春 130012)

摘要:模糊 C 均值聚类(Fuzzy C Mean clustering, FCM)算法因其聚类中心初始化敏感和陷入局部最优的缺点不能对非线性可分离数据进行聚类,而初始聚类中心的选择是随机选取,导致图像分割结果较差。为了克服这一缺点,提出了高斯核模糊 C 均值聚类算法。首先为保证传统 FCM 算法能够快速收敛到准确稳定的聚类中心,提出一种模拟退火遗传算法的模糊 C 均值聚类算法。该算法利用模拟退火较强的局部搜索能力和遗传算法较强的全局搜索能力,可以有效地选取初始聚类中心,提高收敛速度。为增强改进算法对聚类结果不理想的问题,借助高斯核函数进行非线性变换映射至高维空间,转换为高维空间的线性可分问题,改善脑 MRI 图像分割效果。在不同类型噪声和偏移场的影响下,本文算法分割时间比其它算法节约 1~3 倍,概率兰德指数(Probabilistic Rand Index, PRD)至少增加 0.01,可以发现本文算法在临床图像分割结果中,分割时间和脑白质与灰质分割精度都有明显优势。对临床脑 MRI 的图像分割,本文算法比传统 FCM 算法在速度和精度上具有更好的鲁棒性与有效性。

关键词:模糊 C 均值聚类;脑 MRI 图像;图像分割;模拟退火遗传算法;高斯核函数

中图分类号:TP391.4 **文献标识码:**A **doi:**10.3788/OPE.20172513.0259

Research on improved fuzzy clustering segmentation in magnetic resonance image for brain

PAN Hong-liang, GUO Kai-wen, LIU Li-wei, LI Yang, HOU A-lin*

(Changchun University of Technology, College of Computer Science and Technology, Changchun 130012, China)

* Corresponding author, E-mail: houalin@sohu.com

Abstract: The main defect for Fuzzy C mean clustering (FCM) was initialization sensitivity and trapping into the local optimum of clustering center, and nonlinear separable data could not be provided with clustering through this algorithm. Selection for initial clustering center was random selection, which causes relatively bad image segmentation results. In order to conquer this defect, Gaussian kernel Fuzzy C mean clustering was proposed. Fuzzy C mean clustering algorithm for stimulated annealing genetic algorithm was proposed to guarantee traditional algorithm can rapidly converge to accurate

收稿日期:2017-06-01; **修订日期:**2017-08-21.

基金项目:国家留学基金资助项目(No. 201308220163);吉林省高等教育学会 2017 年度高教科研课题资助项目(No. JGJX2017B13);吉林省教育科学“十三五”规划课题资助项目(No. GH170222);国家自然科学基金资助项目(No. 61303132);教育部国际合作科研项目资助项目(No. Z2011138);吉林省科技厅自然科学基金项目资助项目(No. 20101523);吉林省教育厅“十二五”科学技术研究资助项目(No. 2014142, No. 2016349);吉林省科技厅软科学研究资助项目(No. 20160418080FG)

and stable clustering center. Relatively strong local search ability for simulated annealing and relatively strong global search ability for genetic algorithm were utilized to effectively select initial clustering center and improved convergence speed. In order to strengthen and improve non-ideal problem of improved algorithm to clustering result, nonlinear transformation was conducted with the help of Gaussian kernel function to be mapped into higher dimensional space and be converted into linear separable problem of higher dimensional space, which could improve brain MRI image segmentation effect. Under the effect of noises and biased field with different types, segmentation time for algorithm in the thesis saved 1–3 times than other algorithms, and it was at least 0.01 larger than Probabilistic Rand Index (PRI). It can be found that time, white matter and grey matter segmentation accuracies of algorithm in the thesis both have obvious advantages in clinical image segmentation results. For brain MRI image segmentation, the algorithm has better robustness and effectiveness in speed and accuracy compared with traditional FCM algorithm.

Key words: fuzzy C means clustering; brain MRI image; image segmentation; simulated annealing genetic algorithm; gauss kernel function

1 引言

随着图像处理技术和医疗技术的不断发展,计算机辅助检测和诊断对于医学图像的研究有着重要实际应用价值,已经成为阿尔茨海默症、脑肿瘤、帕金森症等精神疾病检测的重要工具。准确地检测脑组织对治疗方案的制订有着十分重要的作用。现代医学影像技术可以采用非侵入的方式对人体内部组织结构进行数字成像,能够为临床医师提供更有价值的解剖学和功能学信息,可以帮助临床治疗方案的制定,避免手术中发生不必要的损伤,并且提高治愈率与患者治疗后的生存几率。对于医学图像识别而言,图像的快速精准分割是图像识别的关键环节。

图像分割是医学图像分析诊断中的一个重要环节。它将图像分为不同的部分,将同质像素分成簇。分割结果可以应用在医疗诊断组织分析、检测肿瘤、骨折和类似的许多问题。已经提出的分割技术有阈值法、区域增长法和聚类法。在各种聚类技术中,最广泛使用的技术包括 k 均值^[1]、模糊 C 均值^[2]及其演变。在 k 均值聚类中,数据被分成若干簇,其中的数据元素恰好属于一个簇。但由于医学图像被认为是模糊的区域边界,在非均匀强度变化方面存在不确定性。利用模糊理论,大大提高了医学成像领域;由 J C Bezdek 提出的模糊 C-均值算法 (Fuzzy C-Mean cluster,

FCM)是一种模糊聚类方法^[3],其对无噪声和偏移场的图像分割效果好,但对受噪声和偏移场污染的图像分割结果却不理想。传统的 FCM 鲁棒性差主要是由于初始聚类中心的随机性与不利用图像中的空间信息。遗传算法 (Genetic Algorithm, GA)是医学图像处理的有力方法^[4-5]。考虑到在优化问题上遗传算法所获得的巨大成功,本文利用模拟退火遗传算法优化初始聚类中心。但该算法采用的欧氏距离只对球类或凸类数据进行聚类效果理想,所以引入高斯核函数进行非线性变换映射至高维空间,将其转换为解决高维数据的线性可分不理想的问题。即通过改进 FCM 分割方法分割出脑灰质 (Gray Matter, GM)、脑白质 (White Matter, WM) 和脑脊液 (Cerebro-Spinal Fluid, CSF) 的图像,并最终与标准分割和目前脑图像分割方法中最常用的方法进行对比。

2 脑组织分割算法原理

2.1 模糊 C 均值聚类算法

FCM 聚类算法是将特征空间 $X = \{x_1, x_2, \dots, x_n\}$ 中的特征点分为 c 类 ($2 \leq c \leq n$), 第 i 类的聚类中心用 $v_i \in R$ 表示, 聚类中心构成矩阵 $V = \{v_1, v_2, \dots, v_n\}$, 其中任意特征点 $c_j \in R$, 属于第 i 类的隶属度为 u_{ij} ($0 \leq u_{ij} \leq 1$), U 是其模糊隶属度矩阵, 将其归一化后, 数据集的隶属度总和等于 1, u_{ij} 介于 $[0, 1]$ 间, 即 u_{ij} 满足如下条件:

$$\sum_{i=1}^c u_{ij} = 1, \forall j = 1, \dots, n.$$

FCM 的目标函数的一般化形式如下:

$$J_{\text{FCM}} = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2, \quad (1)$$

其中, $d_{ij} = \|c_i - x_j\|$ 为第 i 个聚类中心与第 j 个数据点间的欧式距离, $m \in [1, \infty)$ 是模糊指数。

在实际应用中模糊 C 均值聚类存在两个问题:一是由于对初始值敏感,初始聚类中心选择不当,它就会收敛到局部极小值点上,本文采用模拟退火遗传算法优化 FCM 初始聚类中心,使该算法更快速、更稳定地收敛到全局最优解。二是传统 FCM 算法大多采用欧式距离,对球状或凸状结构数据聚类有优势,但在样本高维特征空间线性不可分时会带来划分不准的缺点,算法结合高斯核函数用于改善图像分割效果。实验结果表明,此算法在提高图像分割效率的基础上,与传统算法相比具有更好的鲁棒性。

2.2 基于模拟退火遗传算法改进的 FCM 算法

模拟退火遗传算法 (Simulated Annealing Genetic Algorithm, SAGA)^[6] 是一种基于融合思想而产生的混合优化算法,是模拟退火算法与能够使搜索过程避免陷入局部最优解法的融合。本文采用的模拟退火遗传算法改进 FCM 算法基本流程如下:

步骤 1:初始化各种控制参数:设定 FCM 中参数,种群个数 sizepop、最大遗传代数 MAX-GEN、交叉概率 P_c 、变异概率 P_m 、退火初始温度 T_0 、终止温度 T_{end} 和温度冷却系数 q ;

步骤 2:设循环计数变量 $\text{gen} = 0$;

步骤 3:遗传算法进行选择、交叉、变异操作计算子代目标函数值,创建新种群和新目标函数值,模拟退火算法替换初始种群和目标函数值;

步骤 4:若 $\text{gen} < \text{MAXGEN}$, 则 $\text{gen} = \text{gen} + 1$, 转至步骤 3; 否则转至步骤 6;

步骤 5:若 $T_i < T_{\text{end}}$ 计算出最佳初始聚类中心和目标函数值; 否则, 执行降温操作 $T_i + 1 = KT_i$, 转至步骤 2;

步骤 6:最佳初始聚类中心在模糊 C 类算法中求出最佳聚类中心、隶属度矩阵和最佳目标函数值。

2.3 基于高斯核改进 FCM 算法

支持向量机理论中,原始空间中线性不可分

的样本经常需要利用核函数来加强对其特征的描述来划分数据。可以通过非线性变换映射到高维空间,转换为高维空间的线性可分问题。最为常见的核函数有线性核函数、多项式核函数和高斯核函数^[7],对于核函数的具体选用目前还没有一个统一的标准,但是其中的高斯核函数所对应的线性空间可以是任意维数的,因此用高斯核函数描述的映射问题一定是线性可解的,它的应用最为广泛。利用这一特点,本文采用高斯核函数,定义如下 $K(x, y) = \exp\left(-\frac{\|x-y\|^2}{\sigma^2}\right)$, 将图像的灰度特征通过非线性映射 Φ 变换到高维空间,则 FCM 算法的目标函数变为 $J_{\text{FCM}} = \sum_{i=1}^c \sum_{j=1}^N u_{ij}^m \|\Phi(x_j) - \Phi(v_i)\|^2$ 式中 $\|\Phi(x_j) - \Phi(v_i)\|^2 = \langle \Phi(x_j) - \Phi(v_i), \Phi(x_j) - \Phi(v_i) \rangle$ 可转换为内积, x_j ($j = 1, 2, \dots, C$) 表示图像像素灰度值的集合; c 是聚类数目; v_i ($i = 1, 2, \dots, N$) 是聚类中心集合。

为避免高维运算,将高维空间的内积运算通过低维空间的简单运算实现,无须考虑非线性映射 Φ 的形式,这个函数称作核函数,记为 $k(x, y)$, 满足:

$$k(x, y) = \langle \Phi(x), \Phi(y) \rangle. \quad (2)$$

则 FCM 算法目标函数改写为:

$$J_{\text{FCM}} = 2 \sum_{i=1}^c \sum_{j=1}^N u_{ij}^m (1 - K(x_j, v_i))^2. \quad (3)$$

为使 J_{FCM} 达到最小值,利用拉格朗日乘子优化函数,求得隶属度 u_{ij} 和聚类中心 v_i 的迭代更新表达式为:

$$u_{ij} = \left[\sum_{k=1}^c \left(\frac{d_{ji}(x_j, v_i)}{d_{jk}(x_j, v_k)} \right)^{\frac{1}{m-1}} \right]^{-1} = \left[\sum_{k=1}^c \left(\frac{1 - K(x_j, v_i)}{1 - K(x_j, v_k)} \right)^{\frac{1}{m-1}} \right]^{-1}, \quad (4)$$

$$v_i = \frac{\sum_{j=1}^m u_{ij}^m K(x_j, v_i) x_j}{\sum_{j=1}^n u_{ij}^m K(x_j, v_i)}. \quad (5)$$

3 实验结果及分析

3.1 仿真环境

为验证本文算法的优越性,在保证参数一致的情况下,分别将本文方法与传统 FCM 方法应

用于脑 MRI 图像分割,然后对比实验结果。本文实验环境为 Windows7 操作系统、Intel i5 处理器、4GB 内存,实验平台为 MATLAB R2014a。

3.2 算法评价方法

在本文中使用的概率兰德指数 (Probability Rand Index, PRI) 来定量地评价算法的分割性能。PRI 常用来评价实际分割结果与标准分割结果的接近程度, PRI 值的取值范围为 $[0, 1]$, PRI = 0 时表示实际分割结果与标准结果完全不同; PRI = 1 时表示实际分割结果与标准结果完全相同,因此, PRI 的值越大,则实际分割结果与标准分割结果越接近,被评价的算法越出色。

3.3 仿真实验

本文是以国际公认的模拟人脑 MRI 图像数据库 Brainweb [8] 提供的 T1 加权人脑 MRI 图像做为实验数据,由于可以自行设置不同噪声 (Noise) 和偏移场 (Biased Field, BF),方便测试脑图像分割的新算法,而真实的脑 MRI 图像不能做到这一点,所以模拟人脑 MRI 图像被广泛应用。数据库提供的图像维数为 $181 \times 217 \times 181$, 切

片厚度为 1 mm。

实验中,根据多次实验结果对算法参数进行设置以求达到最好效果,模拟退火算法参数设置为,冷却系数为 0.8,初始温度为 100,终止温度 0。遗传算法参数设置为,个体参数为 8,最大遗传代数为 10,变量的二进制位数为 10,交叉概率为 0.6,变异概率为 0.01。最大迭代次数设置为 100;迭代停止阈值为 0.001。高斯核函数为 3、聚类数为 4 时可以将脑白质、脑灰质、脑脊液和背景分离开。采用 toc 命令统计传统 FCM 算法、ARKFCM^[9]、FCMLSM^[10] 算法以及本文提出 SAGAKFCM 算法单幅图像的分割时间,所有算法均进行 5 次实验,平均分割时间如图 1 所示 (彩图见期刊电子版)。可知,4 种算法分割时间受噪声和偏移场影响较小,SAGAKFCM 分割时间与传统 FCM 算法相比节约 1.5 倍,这是由于 SAGAKFCM 算法是将模拟退火遗传算法 (SAGA) 核模糊 C 均值聚类相结合,用 SAGA 算法选取最佳初始聚类中心,极大减少了 KFCM 迭代次数,缩短了分割时间,可有效提高图像的后续处理效率。

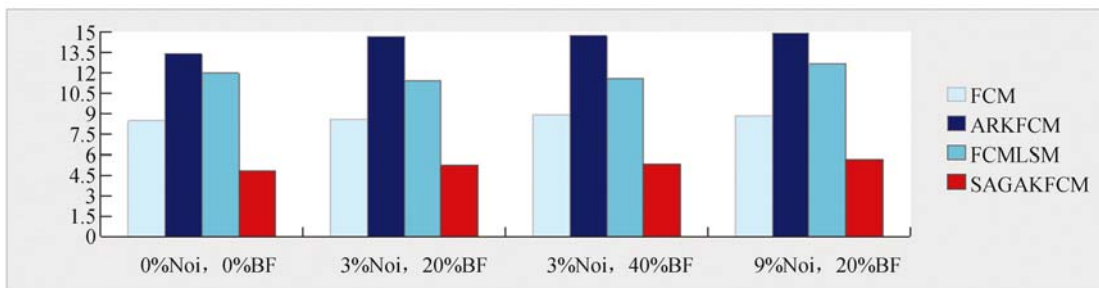
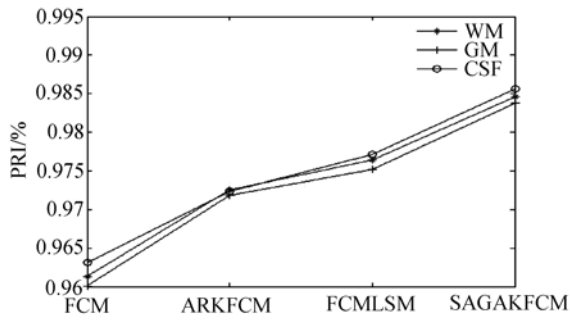


图 1 SAGAKFCM 算法与其他算法的分割时间对比

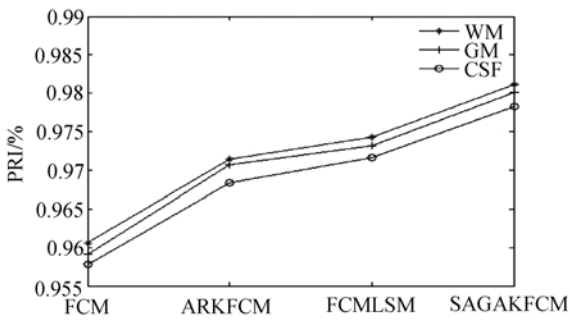
Fig. 1 Comparison of segmentation time of SAGAKFCM with those of other algorithms

图 2 为模拟人脑 MRI 图片在无噪声无偏移场、3%噪声 20%偏移场、3%噪声 40%偏移场、9%噪声 20%偏移场下的 PRI 数据图,可以发现本文分割算法在分割被不同噪声和偏移场污染的脑 MRI 图像时分割的优越性。从图 2 中可以发现 ARKFCM 算法和 FCMLSM 算法受噪声和偏移场影响导致稳定性不好;传统 FCM 算法虽然稳定性好,但分割效果差;本文算法在稳定性和分割效果都明显有优势。图 3(a) 为受 9%噪声 40%偏移场影响的模拟人脑 MRI 图像,(b)

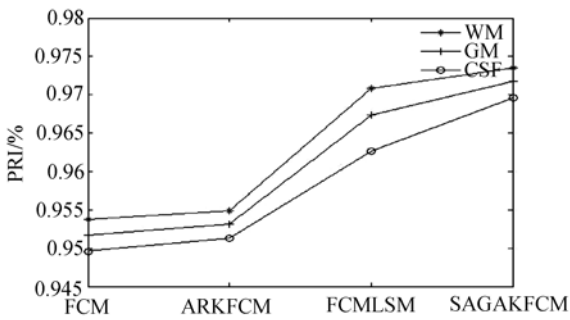
为标准分割结果,(c)为 FCM 算法分割结果,(d)为 ARKFCM 分割结果,(e)为 FCMLSM 分割结果,(f)为 SAGAKFCM 算法的分割结果。通过比较可以看出:FCM 算法的分割结果中仍然含有很强的噪声;ARKFCM 算法和 FCMLSM 算法分割结果中噪声相对于 FCM 算法有不同程度的减少;SAGAKFCM 算法在去除噪声以及细节保持方面都取得了较为理想的分割结果,解决 FCM 对于非线性可分问题效果不理想问题。



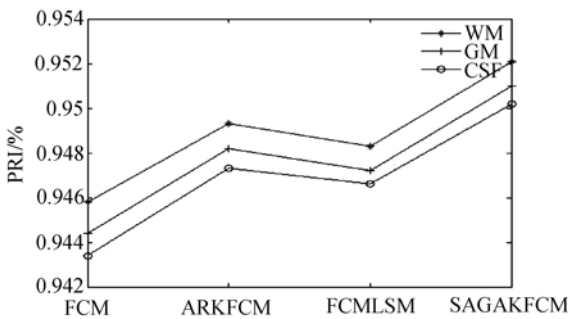
(a) 0% 噪声, 0% 偏移场
(a) 0% noise, 0% biased field



(b) 3% 噪声, 20% 偏移场
(b) 3% noise, 20% biased field



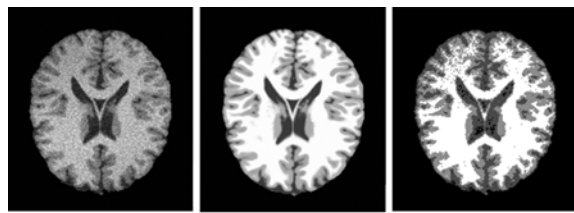
(c) 3% 噪声, 40% 偏移场
(c) 3% noise, 40% biased field



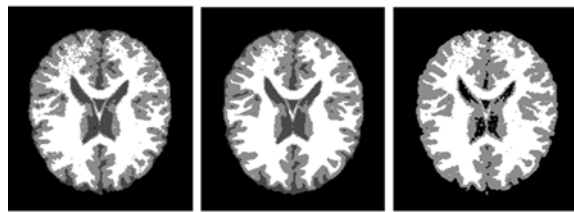
(d) 9% 噪声, 20% 偏移场
(d) 9% noise, 20% biased field

图 2 概率兰德指数数据

Fig. 2 Probability rand index data



(a) 9% 噪声, 40% 偏移场
(b) 标准
(c) FCM
(a) Image with 9% noise and 40% biased field



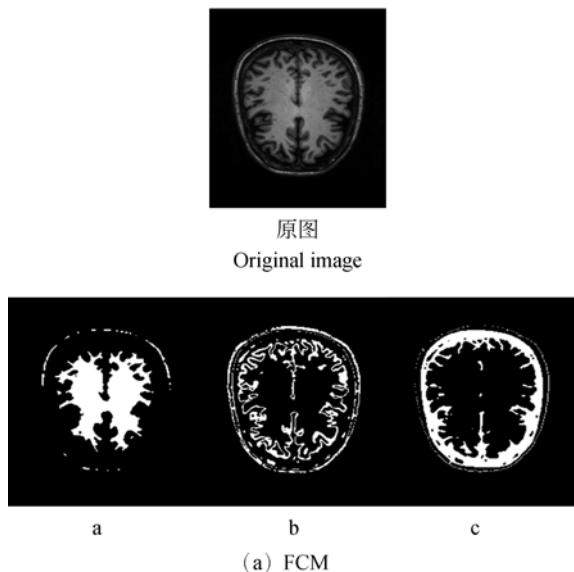
(d) ARKF (e) FCMLS (f) SAGAK

图 3 不同算法分割结果

Fig. 3 Segmentation results of different algorithms

3.4 临床实际脑 MRI 图像实验

临床实验数据是采用加利福尼亚大学神经影像实验室 (Laboratory of Neuro Imaging, LONI)^[1]提供的阿尔茨海默症患者 (ALZHEIMER'S DISEASE, AD) 的 MRI 脑图像。选用的原始图像是横向面第 169 层, 本文算法与其它算法对临床实际脑 MRI 图像分割结果如图 4 所示, (a) 组脑白质的分割有着明显的丢失, (b) 组和 (c) 组



(a) FCM

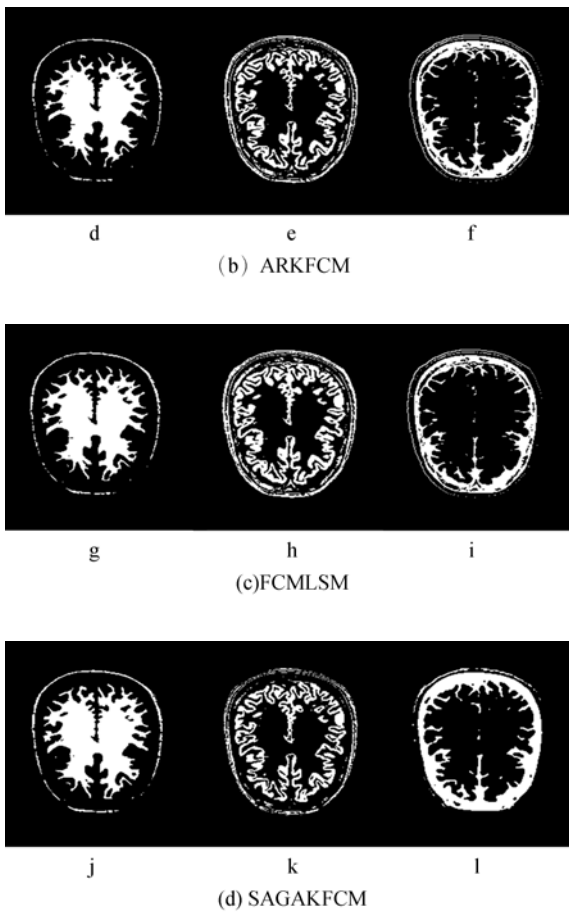


图 4 临床实际脑 MRI 图像分割结果

Fig. 4 Segmentation results of clinical brain MRI images

的脑积液分割结果与(d)组相比也有丢失,尤其在中间部位较为明显。图 5 为临床实际脑图像的分割时间,从中可以发现本文算法时间速度优于其它 3 种算法,本文算法在分割临床脑 MRI 图像在时间和分割精度都有着明显的优势。

参考文献:

- [1] RULANINGTYAS R, SUKSMONO A B, MENG-KO T, *et al.*. Multi patch approach in K-means clustering method for color image segmentation in pulmonary tuberculosis identification[C]. *International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering*. IEEE, 2015:75-78.
- [2] 王和春,宫迅凯. 用模糊聚类 Fuzzy C-Means 算法

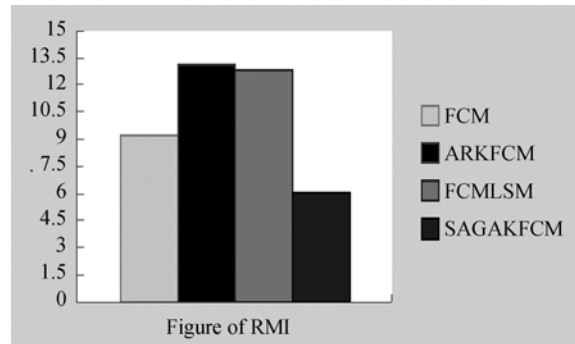


图 5 SAGAKFCM 算法与其他算法的分割时间对比
Fig. 5 Comparison of segmentation time of SAGAKFCM with those of other algorithm

4 结 论

本研究提出改进的 FCM 聚类方法-SAGAKFCM 算法,用于 MRI 图像脑组织的分割。SAGAKFCM 算法通过利用模拟退火较强的局部收搜索能力和遗传算法较强的全局搜索能力,有效地选取最佳初始聚类中心,减少模糊 C 均值算法的迭代次数,加快运算速度。算法结合高斯核函数提高 FCM 算法的鲁棒性。在不同类型噪声和偏移场的影响下,本文算法分割时间比传统 FCM 算法、ARKFCM 算法和 FCMLSM 算法节约 1~3 倍,概率兰德指数(PRI)比 ARKFCM、FCMLSM 至少增加 0.01。本文算法对噪声和偏移场具有较强的鲁棒性,可解决 FCM 对于非线性可分效果不理想的问题。在临床脑图像的分割上对比各种算法分割脑白质和灰质也能发现本文算法的优势。本文算法分割性能较为稳定,可以用于实际 MRI 脑组织分割。

实现图像分割[J]. *光学机械*,1991(4):65-70.

WANG H CH, GONG X K. By using the fuzzy clustering algorithm for image segmentation Fuzzy C-Means [J]. *Optical Machinery*, 1991(4): 65-70. (in Chinese)

- [3] BEZDEK J C, EHRLICH R, FULL W. FCM: The fuzzy c-means clustering algorithm[J]. *Computers & Geosciences*, 1984, 10(2):191-203.
- [4] 张怀柱,向长波,宋建中,等. 改进的遗传算法在实时图像分割中的应用[J]. *光学精密工程*,2008,

- (2):333-337.
- ZHANG H ZH, XIANG CH B, SONG J ZH, Qiao S. Application of improved genetic algorithm in real-time image segmentation [J]. *Opt. Precision Eng.*, 2008, (2): 333-337. (in Chinese)
- [5] KAVITHA A R, CHELLAMUTHU C. Brain tumour segmentation from MRI image using genetic algorithm with fuzzy initialisation and seeded modified region growing (GFSMRG) method [J]. *Imaging Science Journal*, 2016, 64(5):1-13.
- [6] 古良玲,王玉菡. 基于模拟退火遗传算法的图像增强[J]. *激光杂志*, 2015(2):19-22.
- GU L L, WANG Y N. Image enhancement based on simulated annealing genetic algorithm [J]. *Journal of Lasers*, 2015 (2): 19-22. (in Chinese)
- [7] 唐苦,王昕,王振雷,等. 一种基于高斯核化有效性指标的自适应优选聚类数的 FKCM[J]. *计算机与应用学*, 2012, 29(10):1199-1203.
- TANG K, WANG X, WANG ZH L, *et al.*. An adaptive optimization clustering number based on Gauss's nuclear validity index FKCM[J]. *Computer and Applied Chemistry*, 2012, 29 (10): 1199-1203. (in Chinese)
- [8] <http://brainweb.bic.mni.mcgill.ca/brainweb/>
- [9] AHMED E, WANG C, JIA F, *et al.*. Segmentation of brain tissues from magnetic resonance images using adaptively regularized kernel-based fuzzyC-means clustering[J]. *Computational & Mathematical Methods in Medicine*, 2015, 2015:485495.
- [10] SINGH P, BHADARIA H S, SINGH A. Automatic brain MRI image segmentation using FCM and LSM[C]. *International Conference on Reliability, INFOCOM Technologies and Optimization*, IEEE, 2015:1-6.
- [11] <http://www.loni.usc.edu/>

作者简介:



侯阿临(1972—),女,湖南临澧人,博士,长春工业大学教授,主要从事信号与信息处理、图像工程与光电信息处理方向研究,E-mail:houlin@sohu.com



潘宏亮(1987—),男,吉林人,硕士研究生,现就读于长春工业大学计算机科学与工程学院,主要从事图像处理与机器视觉方向研究。E-mail:252932727@qq.com