

· 生物电子学 ·

## 遗传模糊C-均值聚类算法应用于MRI分割

曾翎<sup>1</sup>, 王美玲<sup>2,3</sup>, 陈华富<sup>1</sup>

(1. 电子科技大学生命与技术学院 成都 610054; 2. 电子科技大学应用数学学院 成都 610054;  
3. 解放军信息工程大学电子技术学院 郑州 450004)

**【摘要】**通过分析经典的将图像分割成C类的常用的模糊C-均值聚类算法(FCMA)依赖于初始聚类中心的选择, 通常得到的是局部最优解而非全局最优解, 又由于遗传算法能搜索到全局最优解, 因此将遗传算法(GA)与FCMA相结合, 对MRI直接进行聚类, 利用遗传算法搜索全局最优解, 从而有效地避免了模糊C-均值聚类算法收敛到局部最优的问题, 并在此基础上实现了对MRI的分割, 得到了比较满意的效果。

**关键词** 模糊聚类; 遗传算法; MRI分割  
**中图分类号** R318.04 **文献标识码** A

## Genetic Fuzzy C-Means Clustering Algorithm for Magnetic Resonance Images Segmentation

ZENG Ling<sup>1</sup>, WANG Mei-ling<sup>2,3</sup>, and CHEN Hua-fu<sup>1</sup>

(1. School of Life Science and Technology, University of Electronic Science and Technology of China Chengdu 610054  
2. School of Applied Mathematics, University of Electronic Science and Technology of China Chengdu 610054;  
3. Institute of Electronic Technology, PLA Information Engineering University Zhengzhou 450004)

**Abstract** Based on the class fuzzy C-means clustering algorithm (FCMA) is a well-known clustering method to partition an image into homogeneous region. We know FCMA is dependent on the choice of the initial distribution of cluster center, and consequently the algorithm ends up in a local optimum. Because of the genetic algorithm which can achieve the global optimum, we directly unified them in the magnetic resonance images (MRI) segmentation. By applying genetic algorithm, we can achieve the global optimum in MRI segmentation application.

**Key words** fuzzy cluster; genetic algorithm; MRI segmentation

磁共振图像(magnetic resonance images, MRI)是一种有广泛应用的神科学、心理学、临床医学图像, 它能以一种非扩散的方式在2维、3维解剖过程中提供大量丰富的量化的信息。随着临床医学的发展, 脑部结构信息系统得到了广泛的应用。一般来说, 白质、灰质和脑机体是脑中3种最基本的物质。MRI的脑组织分割<sup>[1-3]</sup>是基于MRI和先验知识获得的信息, 把2维或3维的每个像素或者体素划分为不同的组织类。经典的C-均值聚类算法(fuzzy C-means clustering algorithm, FCMA)<sup>[4]</sup>是一种常用的聚类方法。但是, 该算法依赖于初始聚类中心的选择, 最终结果经常会收敛到局部最优解而不是全局最优解<sup>[5-6]</sup>。本文将遗传算法与C-均值聚类算法相结合<sup>[7-8]</sup>, 将

C-均值聚类算法推广到高维空间, 对MRI直接进行聚类, 用遗传算法搜索全局最优的聚类中心, 有效地避免了直接使用C-均值聚类算法所带来的收敛到局部最优的问题, 并在此基础上实现了对MRI的分割。对真实MRI图像进行的提取分析结果表明, 该方法是可行的。

### 1 模糊C-均值聚类算法

经典的FCMA聚类算法是将一幅图像分为C类区域的常用方法, FCMA聚类算法的目标函数是:

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

收稿日期: 2006-12-15; 修回日期: 2007-05-13

资助项目: 国家自然科学基金(30370507)

作者简介: 曾翎(1965-), 男, 副教授, 主要从事脑功能成像方面的研究。

式中  $N$  为处理的样本总量;  $C$  为分类数目;  $m > 1$  为控制聚类结果的模糊程度的常数;  $\mathbf{x}_i$  为  $P$  维向量;  $\mathbf{v}_k$  为第  $i$  类聚类中心;  $u_k(\mathbf{x}_i)$  是第  $k$  个样本对第  $i$  类的隶属度函数。

具体的FCMA聚类算法步骤如下:

- 1) 设定聚类数  $C$ 。
- 2) 任选  $C$  个初始聚类中心  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_C$ 。
- 3) 重复计算隶属度函数和各聚类中心的运算, 直到各个样本的隶属度函数稳定:

- (1) 首先用当前的聚类中心计算隶属度函数:

$$u_k(\mathbf{x}_i) = \frac{(1/\|\mathbf{x}_i - \mathbf{v}_k\|^2)^{1/(m-1)}}{\sum_{l=1}^C (1/\|\mathbf{x}_i - \mathbf{v}_l\|^2)^{1/(m-1)}} \quad (2)$$

- (2) 然后用当前的隶属度函数更新计算各聚类中心:

$$\mathbf{v}_k = \frac{\sum_{i=1}^N [u_k(\mathbf{x}_i)]^m \mathbf{x}_i}{\sum_{i=1}^N [u_k(\mathbf{x}_i)]^m} \quad (3)$$

当算法收敛时, 就得到各类的聚类中心和各个样本对于各类的隶属度值, 从而完成了模糊聚类的划分。这种迭代过程的收敛速度比较快。但是, 该算法依赖于初始聚类中心的选择, 因此经常会收敛于局部最优, 而不是通常所要的全局最优。

## 2 遗传模糊C-均值聚类算法

遗传算法是一类借鉴生物界自然选择和自然遗传机制的随机并行搜索方法, 是一种全局优化搜索方法<sup>[9]</sup>。将遗传算法应用于FCMA聚类算法, 可有效避免FCMA聚类算法对初始聚类中心的依赖, 获取全局最优。遗传算法(GA)的主要操作有: 初始化群体、选择、交叉、变异。遗传算法用于FCMA聚类算法的具体过程如下:

- (1) 确定聚类数目  $C$ 。

(2) 初始化群体: 随机生成含有  $P$  个个体的群体, 每个个体代表  $C$  类中每个类的中心  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_C\}$ ;  $\mathbf{v}_j$  为  $P$  维向量。

(3) 交叉: 设交叉概率为  $p_c$ , 对  $P$  中每一个个体随机生成一个  $0 \sim 1$  之间的随机数  $r$ , 若  $r < p_c$ , 则进行交叉操作。对  $P$  随机选中一个个体, 随机生成一个在  $1 \sim C$  之间的整数  $i$ , 按  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_C\} - \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j, \mathbf{v}'_{j+1}, \dots, \mathbf{v}'_C\}$  和  $\{\mathbf{v}'_1, \mathbf{v}'_2, \dots, \mathbf{v}'_C\} - \{\mathbf{v}'_1, \mathbf{v}'_2, \dots, \mathbf{v}'_j, \mathbf{v}_{j+1}, \dots, \mathbf{v}_C\}$  方式对位置  $j$  处进行交叉操作, 也就是第  $j$  类的中心交换。

(4) 变异: 设变异概率为  $p_m$ , 对选中的每一个个体  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_C\}$  随机选定某一位置  $\mathbf{v}_j$ , 对其进行变异操作; 随机产生一个新的  $\mathbf{v}_j$ 。

- (5) 个体的适应度为:

$$F = \frac{1}{1 + J(u, v)} \quad (4)$$

对交叉、变异产生的新个体与群体中的个体进行比较, 适应度  $F$  较大的被选入群体, 而适应度最小的群体被淘汰。

## 3 结果

将遗传模糊C-均值聚类算法GFCMA (genetic fuzzy c-means clustering algorithm)应用于MRI分割, 图像大小为  $109 \times 91$ 、切片个数为91、图像为大脑切片原图像1a的第37张、1d的第51张和1g的第70张图片。所得结果如图1所示。图中, 图1a、1d、1g为原始图像; 图1b、1e、1h为相应的FCMA分割后的图像; 图1c、1f、1i为相应的GFCMA分割后的图像。

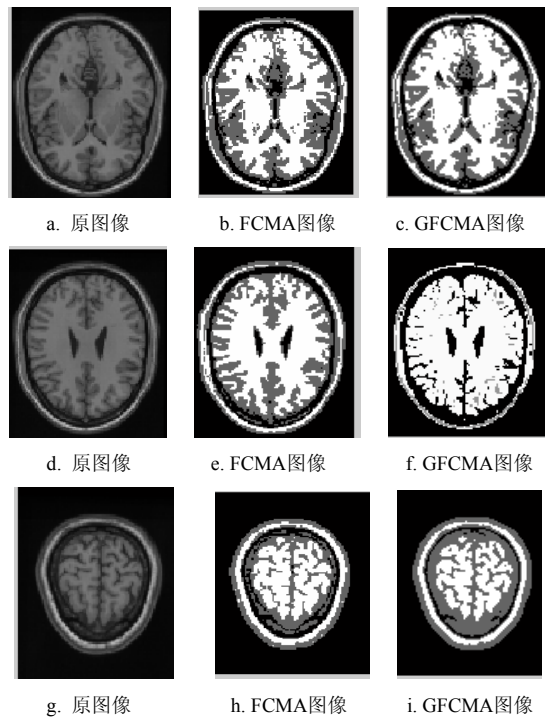


图1 原图像、FCMA图像与GFCMA图像

在GFCMA中, 群体规模为80、 $C=3$ (灰质, 白质, 脑机体)、循环次数为100、交叉概率  $p_c=0.7$ 、变异概率  $p_m=0.09$ 。

将原始的灰度值与分割后的图像灰度值做均方误差比较, FCMA 和GFCMA在脑图像分割中的算法计算结果如表1所示。从表1中可以看出, GFCMA算法的均方误差基本比FCMA的均方误差要小, 表

明GFCMA算法要优于FCMA算法。

表1 3幅图片的FCMA和GFCMA算法计算结果

图片	类别数	FCMA均方误差	GFCMA均方误差
a	3	12.287 6	12.238 5
b	3	11.816 1	10.472 4
g	3	9.652 3	10.171 8

### 4 结 束 语

从本文的算法过程可以看出, 当需要处理的问题的样本或样本维数很大时, 适应度函数的计算量变得非常巨大, 导致算法需要花费大量的时间才能收敛到最优值。为减少计算量, 本文中实际计算时, 将遗传算法与模糊C-均值聚类方法结合起来。具体就是初始化群体后, 先对群体中的每个个体应用FCMA迫使其收敛到局部最优, 然后再进行交叉变异操作; 最后再对遗传算法中每一代交叉变异所产生的新个体在可能被加入群体之前也先应用FCMA使之收敛到局部最优。然而尽管作如此的改进, GFCMA算法仍是相当耗时的一项工作。

#### 参 考 文 献

[1] VELTHUIZEN R P, HALL L O, CLARKE L P. Feature extraction for MRI segmentation[J]. Neuro Imaging, 1999, 9(2): 85-90.  
 [2] BARRA V, BOIRE J V. Tissue segmentation on MR images of the brain by possibilistic clustering on a 3D wavelet representation[J]. Magn Reson Imag, 2000, (11): 267-278.

[3] SCHROETER P, VESIN J M, LANGENBERGER T, et al. Robust parameter estimation of intensity distribution for brain magnetic resonance images[J]. IEEE Trans Med Imaging, 1998, 17(2): 172-186.  
 [4] PHAM D L, PRINCE J L. Adaptive fuzzy segmentation of magnetic resonance images[J]. IEEE Trans Med Imaging, 1999, 18: 737-752.  
 [5] 章 晋. 图像分割[M]. 北京: 科学出版社, 2001. ZHANG Jin. Images Segmentation[M]. Beijing: Science Press, 2001.  
 [6] 雷英杰, 张善文, 李续武, 等. 遗传算法工具箱及应用[M]. 西安: 西安电子科技大学出版社, 2005. LEI Ying-jie, ZHANG Shan-wen, LI Xu-wu, et al. Genetic algorithm tools & applications[M]. Xi'an: Xidian University Press, 2005.  
 [7] WANSCHURE T, COLEY D A, VENNART W, et al. Automat realignment of time-separated MR images by genetic algorithm[J]. Magnersson Imaging, 1999, 17(2): 313-317.  
 [8] 周叙国, 邢汉承. 基于遗传模糊C-均值聚类算法的地图分割[J]. 计算机与数字工程, 2005, 33(6): 77-78. ZHOU Xu-guo, XING Han-cheng. Topographic map segmentation based on genetic fuzzy C-means clustering algorithm[J]. Computer & Digital Engineering, 2005, 33(6): 77-78.  
 [9] 边肇祺, 张学工. 模式识别[M]. 北京: 清华大学出版社, 2004. BIAN Zhao-qi, ZHANG Xue-gong. Pattern recognition[M]. Beijing: Tsinghua University Press, 2004.

编 辑 熊思亮

(上接第609页)

#### 参 考 文 献

[1] DEMING R, PERLOVSKY L. Concurrent multi-target localization, data association, and navigation for a swarm of flying sensors[J]. Information Fusion, 2007, 8(3): 316-330.  
 [2] ZHU H. The algorithm and simulations for the asynchronous track-to-track association[C]//Proceedings of the Sixth International Conference on Information Fusion. Washington: IEEE, 2003: 1400-1404.  
 [3] OUSSALAH M, SCHUTTER J. Hybrid fuzzy probabilistic data association filter and joint probabilistic data association filter[J]. Information Sciences, 2002, 142(1): 195-226.  
 [4] MEGALOOIKONOMOU V, YESHA Y. Space efficient quantization for distributed estimation by a multi-sensor fusion system[J]. Information Fusion, 2004, 5(4): 299-308.  
 [5] DRUMMOND E. Performance Metrics for Multiple-Sensor, Multiple -Target Tracking[C]//Signal and Data Processing of Small Targets 2000, Proceedings of SPIE. Atlanta: IEEE, 2000.  
 [6] CHANG K, TIAN Z MAP track fusion performance evaluation[C]//Proc Fusion 2002. Washington DC: [s.n.],

2002.  
 [7] REID D. Optimising the flow of information within a C3I network[J]. Mathematical and Computer Modelling, 1999, 30(9): 225-243.  
 [8] SONG H, Compton K. Rounds SPIN: a model checker for reconfigurable hybrid systems based on SPIN[J]. Electronic Notes in Theoretical Computer Science, 2006, 145(14): 167-183.  
 [9] VENEMA Y. Automata and fixed point logic: a coalgebraic perspective[J]. Information and Computation, 2006, 204(4): 637 -678.  
 [10] 高 军, 杨冬青. 基于树自动机的XPath在XML数据流上的高效执行[J]. 软件学报, 2005, 16(2): 223-232. GAO Jun, YANG Dong-qing. Tree automata based efficient XPath evaluation over XML data stream[J]. Journal of Software, 2005, 16(2): 223-232.  
 [11] DIERKS H. PLC-automata: a new class of implemenTab real-time automata[J]. Theoretical Computer Science, 2001, 253: 61-93.

编 辑 税 红