

基于改进模糊聚类分析的医学脑部 MRI 图像分割

周显国¹, 陈大可², 苑森淼²

(1. 吉林省人民医院, 长春 130022; 2. 吉林大学 通信工程学院, 长春 130022)

摘要:结合 MRI 图像的直方图统计信息, 提出了一种改进的快速 FCM(HF-KFCM)算法。算法首先利用多尺度窗口遍历的方法找到直方图的峰值点, 然后将其作为模糊聚类的初始化中心, 并使用基于统计信息的快速聚类方法进行遍历, 以减少每次迭代的运算量。仿真结果表明, 该算法相比于标准 FCM 算法和其他改进算法, 在聚类有效性和模糊性上的分割效果显著提高。

关键词:信息处理技术; 图像分割; 脑部磁共振图像; 模糊聚类; 直方图统计

中图分类号:TP391 **文献标识码:**A **文章编号:**1671-5497(2009)Sup. 2-0381-05

Medical brain MRI images segmentation by improved fuzzy C-Means clustering analysis

ZHOU Xian-guo¹, CHEN Da-ke², YUAN Sen-miao²

(1. People's Hospital of Jilin Province, Changchun 130022, China; 2. College of Communication Engineering, Jilin University, Changchun 130022, China)

Abstract: For the shortcomings of huge calculation in MRI images segmentation with standard Fuzzy C-Mean algorithm (FCM), a new algorithm combined with histogram statistical information of Improved Fast Fuzzy C-Mean algorithm (HF-KFCM) was proposed. Firstly, the method of multi-scale window traverse is used by this algorithm to find the histogram peaks. Then these peaks are defined as the fuzzy clustering initialization centre. Meanwhile, the fast FCM method based on histogram statistic is used as ergodicity to reduce each iteration calculation. The simulation results show that compared with standard FCM algorithm and other improved algorithms, the proposed algorithm can be improved significantly in fuzzy and clustering effectiveness.

Key words: information processing; image segmentation; brain MRI image; fuzzy C-Mean; histogram statistic

MRI 脑图像的分割问题已成为当今医学和计算机图像处理领域的一个研究热点^[1-2]。研究人员尝试了很多数学模型来改进 MRI 脑组织分割算法, 常用的有动态变形模型^[3]、人工神经网络^[4]、Gibbs 随机场^[5]、形态学算法^[6]等。其中应

用最为广泛的是基于模糊聚类模型(FCM)的图像分割算法。标准 FCM 算法^[7]在聚类过程中存在计算量大、迭代次数过多、容易收敛到局部极小值等缺点。其改进的快速 FCM 算法^[8](KFCM)虽然减少了计算量, 但在分割中仅利用了图像灰

收稿日期:2008-12-12.

基金项目:吉林省科学技术厅项目(20070323).

作者简介:周显国(-), 高级工程师. 研究方向: . E-mail: zxg2004zxg@126.com

度信息,没有考虑像素的空间统计信息,因此只适用于分割噪声含量很低的图像。

针对上述问题,本文提出了一种结合直方图统计信息的改进快速 FCM(HF-KFCM)算法。仿真结果表明,该算法在抗噪声性能和分割有效性上相对于 FCM 算法和其他改进算法有显著提高,同时迭代次数和运算量也有一定的减少。

1 基于 HF-KFCM 算法的图像分割

1.1 标准 FCM 算法

FCM 算法^[7]是由 Bezdek 在 Dunn 基础上提出的,其模糊 C-均值泛函的定义如下

$$J_m(U, V; X) = \sum_{k=1}^n \sum_{i=1}^c u_{ik}^m (d_{ik})_{\Lambda}^2 \quad (1)$$

式中: $m \in [1, \infty]$ 是在每一个模糊隶属度上的权重指数; $U \in M_{fc}$ 是一个 X 的约束模糊 C-划分,它是一个 $n \times c$ 的矩阵。约束条件为

$$\begin{cases} 0 \leq u_{ik} \leq 1, \forall k, i \\ \sum_{i=1}^c u_{ik} = 1, \forall k \end{cases} \quad (2)$$

$V = (v_1, v_2, \dots, v_c) \in R^n$ 为类中心矢量; $d_{ik} = \|x_k - v_i\|$ 为数据点 x_k 与第 i 类中心 v_i 之间的距离; Λ 是任意的 $P \times P$ 的正定矩阵。一般 d_{ik} 的计算可按欧氏范数或明氏范数。准则函数 $J_m(U, V; X)$ 表示样本 x_k 与各个聚类中心 v_i 的带权距离平方和。因此, J_m 是平方误差聚类准则,故样本集合 X 的最佳模糊 C-划分应该是 J_m 的最小方差稳定点 (U, V) , 由最小化 J_m 解得。

FCM 定理: 假设 $d_{ik} > 0, \forall i, k$, 对于 $m > 1$, 只有在满足约束条件(2)的情况下,且下式

$$u_{ik} = \left[\sum_{j=1}^c \left(\frac{d_{jk}}{d_{ik}} \right)^{2/(m-1)} \right]^{-1}, \text{ any } i, k \quad (3)$$

$$v_i = \frac{\sum_{k=1}^n (u_{ik})^m x_k}{\sum_{k=1}^n (u_{ik})^m}, \text{ any } i \quad (4)$$

成立时,数据对 (U, V) 可能使 J_m 达到最小。

具体实现步骤见参考文献[7]。

1.2 KFCM 算法

KFCM^[8](Kernelized fuzzy C-means)是用高斯核函数作为距离测度的模糊聚类算法。其核函数为

$$K(x, y) = \exp(-\beta \|x - y\|^2) \quad (5)$$

式中: $\beta = (\sum_{j=1}^n \|x_j - \bar{x}\|^2 / n)^{-1}$, $\bar{x} = \sum_{j=1}^n x_j / n$ 。

核函数的思想是将非线性的输入数据映射到一个潜在的线性的高维空间,从而使非线性问题可解,而且无论这个高维空间为几维,其内积的表示可以简化核函数的运算。核函数的内积表示为

$$K(x, y) = \langle \phi(x), \phi(y) \rangle \quad (6)$$

由于

$$\begin{aligned} & \| \phi(x_k) - \phi(v_i) \|^2 = \\ & (\phi(x_k) - \phi(v_i))^T (\phi(x_k) - \phi(v_i)) = \\ & K(x_k, x_k) + K(v_i, v_i) - 2K(x_k, v_i) \end{aligned} \quad (7)$$

对于高斯核函数

$$\| \phi(x_k) - \phi(v_i) \|^2 = 2(1 - K(x_k, v_i)) \quad (8)$$

由此得到的 KFCM 目标函数为

$$J_m = 2 \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m (1 - K(x_k, v_i)) \quad (9)$$

隶属度划分矩阵为

$$u_{ik} = \frac{(1 - K(x_k, v_i))^{-1/(m-1)}}{\sum_{j=1}^c (1 - K(x_k, v_j))^{-1/(m-1)}} \quad (10)$$

模糊聚类中心为

$$v_i = \sum_{k=1}^n u_{ik}^m K(x_k, v_i) x_k / \sum_{k=1}^n u_{ik}^m K(x_k, v_i) \quad (11)$$

1.3 HF-KFCM 算法

通过上述分析可以得出,标准 FCM 算法及其改进算法仅考虑像素的强度,而没有利用任何空间统计信息。脑部 MRI 图像可以看作是分片常数图像,相邻的像素具有相关性。因此,对脑 MRI 图像分割算法应充分考虑像素强度的空间统计信息。

灰度直方图是显示图像中每一个灰度级像素个数的函数。设图像的最大灰度级为 $L - 1$, $h(k), k = 0, 1, \dots, l - 1$ 为图像中灰度级为 k 的像素的个数,将 $h(k)$ 除以总像素个数即为灰度分布概率。

基于直方图统计的快速 FCM 算法(HF-KFCM)将针对像素灰度值的运算转换为针对灰度级的运算,模糊聚类中心的运算公式为

$$v_i = \sum_{k=0}^{l-1} h(k) (u_{ik})^m k / \sum_{k=0}^{l-1} h(k) (u_{ik})^m \quad (12)$$

使用多尺度窗口对直方图进行遍历的方法寻找峰值点。采用由小到大的窗口对直方图遍历。从宽度为 3 的窗口开始,当窗口内的直方图统计值的分布近似高斯分布时,保留;再用大小为 5 的窗口进行第二次遍历,判断窗口内的直方图统计

值是否近似高斯分布;依次遍历,直到寻找到的峰值的数量等于事先确定的峰值个数。

为了进一步提高算法的速度,可将隶属度矩阵的运算转化为查表运算。这样,对于一幅大小为 $M \times N$ 的 L 级灰度图像,进行一次迭代中,隶属度矩阵的计算量和存储量也由 $M \times N$ 减小到 L 级。

如图 1 所示,一幅已去除非脑组织的 MRI 图像可以被分为 4 类:脑白质、脑灰质、脑脊液和背景。脑部 MRI 图像每一类的像素灰度分布近似为高斯分布。使用该算法对脑组织 MRI 图像分割时,其聚类中心应该在直方图的峰值附近。用峰值作为模糊聚类中心的初始值时,由于初始聚类中心 $V^{(0)}$ 接近最优解 V^* ,因此可以减少迭代次数,加快收敛速度,同时减小分割结果陷入局部极小值的可能。

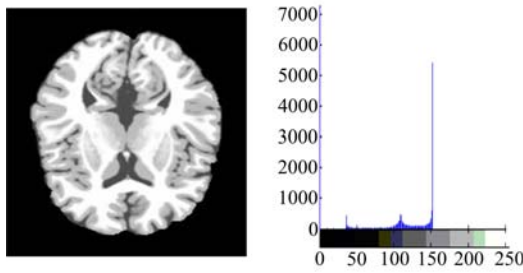


图 1 脑部 MRI 图像的直方图分析

Fig. 1 Histogram analysis on brain MRI image

2 算法执行步骤

该算法分为两部分:第一部分是完成对源图像的直方图统计和使用多尺度窗口对直方图遍历找到直方图的 4 个峰值点;第二部分是将从得到的直方图峰值点作为模糊聚类的初始化中心,采用高斯核函数距离测度来加强算法的分割有效性,并采用基于直方图统计的快速聚类算法以减少每一次迭代的运算量。具体步骤如下:

- (1) 初始化参数: $\epsilon, V^{(0)}$, 模糊指数 m ;
- (2) 对源图像进行直方图统计;
- (3) 使用多尺度窗口遍历的方法寻找 4 个峰值点:采用由小到大的窗口进行遍历,并判断窗口内的直方图统计值是否近似高斯分布,直到寻找到的峰值数量等于事先确定的峰值数;
- (4) 将得到的直方图峰值点作为模糊聚类的初始化中心并进行隶属度更新,然后依据公式 (12) 再对聚类中心更新;

(5) 重复步骤(4)至各聚类中心收敛;

(6) 根据距离测度最大原则对图像进行分割。

3 图像分割质量评价

目前,常用的磁共振图像分割质量评价方法包括像素数量误差和最终测量精度等。针对本文要实现的模糊聚类算法,采用的分割评价参数有聚类有效性函数,分割系数和分割熵^[9-10]。

(1) 聚类有效性函数(RII)

图像分割出来的每个区域内部特征应该是相似的,因此可以采用分割图中各区域内部特征均匀的程度来描述分割图像的质量。最小类间距离与最大类内距离的比值定义聚类有效性函数 RII 为

$$RII = \frac{\min(\text{dist}(s_i, s_j))}{\max(\text{diam}(s_k))} \quad (13)$$

式中: $\text{dist}(s_i, s_j)$ 表示类 s_i 与 s_j 间的距离; $\text{diam}(s_k)$ 为类 s_k 内的距离。

$$\text{diam}(s_i, s_j) = 1 - \frac{\sum_{k=1}^n \min[u_{ik}, u_{jk}]}{\sum_{k=1}^n \max[u_{ik}, u_{jk}]} \quad (14)$$

$$\text{diam}(s_i) = \max_{i=1, \dots, c} \left[\frac{\sum_{j=1}^n u_{ij} (x_j - v_i) (x_j - v_i)^T}{\sum_{j=1}^n u_{ij}} \right] \quad (15)$$

对于最优分类,由于类间距离可以达到最大,类内距离可以达到最小,因此聚类有效性测量函数 RII 的值趋于最大,表示该模糊划分是紧致的。

(2) 分割系数 V_{pc} 和分割熵 V_{pe}

针对模糊聚类算法,评估分割模糊性的代表函数是分割系数 V_{pc} 和分割熵 V_{pe} 。这两个有效性函数的意义在于具有较小模糊性(fuzziness)的分割是较好的分割。因此最好的分割是当 V_{pc} 最大或 V_{pe} 最小的时候得到的。其定义式如下

$$V_{pc} = \sum_j \sum_i u_{ij}^2 / N \quad (16)$$

$$V_{pe} = - \sum_j \sum_i [u_{ij} \log u_{ij}] / N \quad (17)$$

4 仿真结果及其性能分析

为了检验该算法的性能,本文采用加拿大蒙特利尔神经研究所的脑部 MRI 图像进行仿真,并

与标准 FCM 算法和 KFCM 算法的分割结果进行比较。参数设定为： $m=2, c=4, \epsilon=0.001$ ，初

始化模糊聚类中心使用直方图统计信息的峰值。各种算法的仿真结果如图 2 所示。

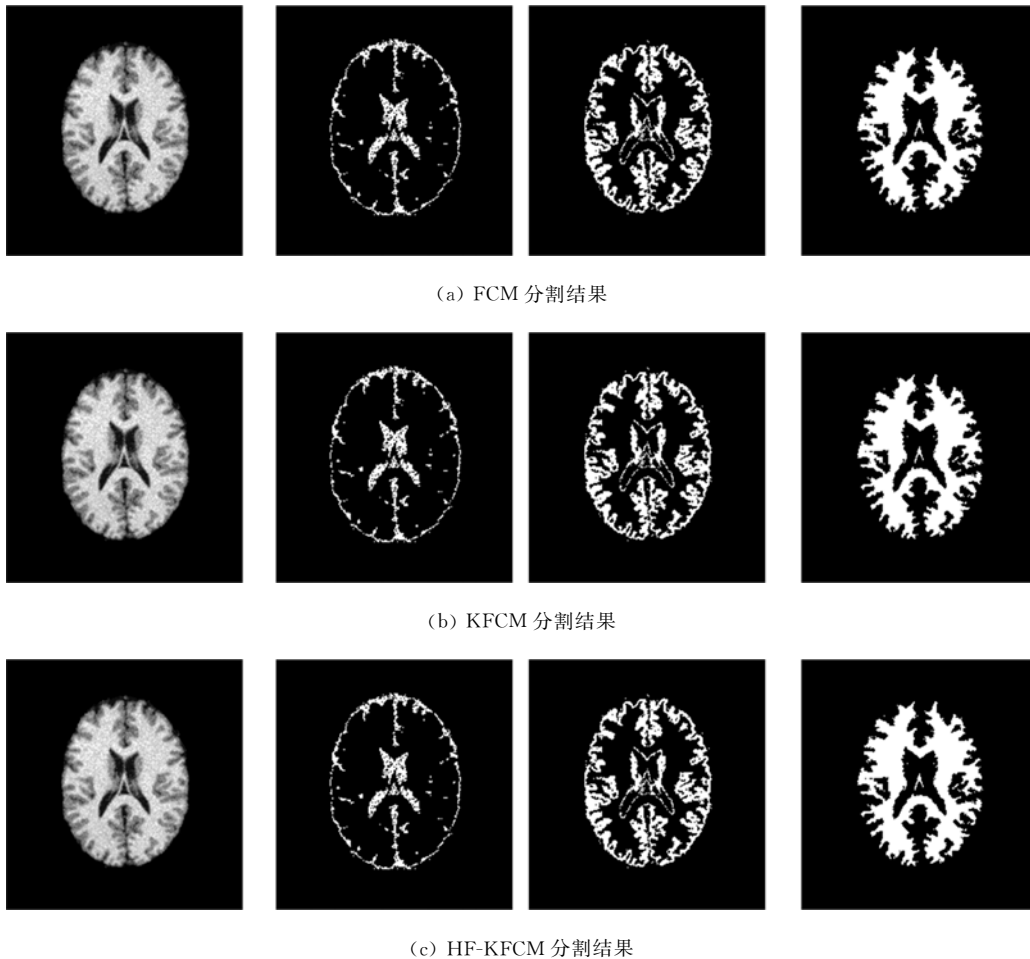


图 2 各种方法的分割结果图像

Fig. 2 Segmentation results of various methods

使用分割质量评估参数对算法的性能进行客观比较分析，得到表 1，图 3。

表 1 性能指标比较

Table 1 Comparison of performance index

算法	RII	V_{pc}	V_{pc}	误分率	迭代次数
FCM	79.3461	0.9444	0.1589	8.06	11
KFCM	58.0244	0.9253	0.1836	7.56	9
HF-KFCM	104.6736	0.9660	0.0841	5.64	3

从仿真结果图 2,3 及表 1 可以看出

(1)相比于标准 FCM 算法和 KFCM 算法，本文提出的 HF-KFCM 算法在聚类有效性上分别提高了 23.72%和 44.56%，分割系数 V_{pc} 分别增大了 2.24%和 4.21%，分割熵 V_{pc} 分别降低了 47%和 54%。

(2)HF-KFCM 算法只需要较少的迭代计算就会收敛到目标函数极小值。同时由于直方图的

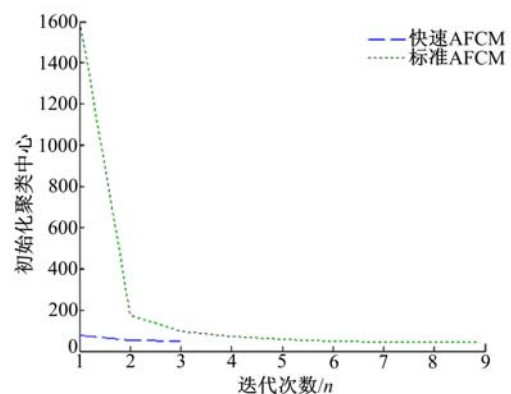


图 3 KFCM 算法与 HF-KFCM 算法收敛比较
Fig. 3 Convergence comparison between KFCM and HF-KFCM algorithm

峰值距离聚类中心较近，使得算法陷入局部极小值的可能性减少，算法的稳定性增强。

(3)HF-KFCM 算法运算量的减少是从两个

方面着手的。首先,初始化聚类中心的优化选择减少了迭代次数。仿真中, HF-KFCM 算法只需要 3 次迭代就可以收敛,而标准 KFCM 算法需要 9 次迭代。其次是单次迭代过程中,运算量的减少。HF-KFCM 算法单次迭代中使用了基于直方图统计信息的快速聚类算法,将运算量由 $N \times N$ 量级减小到了 L (L 代表灰度级)量级。

5 结束语

针对脑部 MRI 图像的分割提出了一种结合直方图统计信息和直方图峰值信息的模糊聚类算法,称为 HF-KFCM(Histogram Fast KFCM)算法。该算法较标准 FCM 算法和 KFCM 算法的改进之处有:利用图像的直方图峰值作为模糊聚类的初始聚类中心,使算法的迭代次数减少;在单次迭代中,使用基于直方图统计信息的快速聚类算法,以减少单次迭代中的计算量。仿真结果表明,该算法的迭代次数及运算量明显减少。同时,降低了算法陷入局部极小值的可能,增强了算法的鲁棒性。

参考文献:

- [1] Shan Shen, William Sandham. MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization[J]. IEEE Transactions on Information Technology, 2005, 9(3): 459-467.
- [2] Dzung L, Pham Jerry L. Adaptive fuzzy segmentation of magnetic resonance images[J]. IEEE Transactions on Medical Imaging, 1999, 18(9): 737-752.
- [3] 栾红霞,戚飞虎. 一种新的用于 MRI 脑图像分割的主动轮廓模型[J]. 仪器仪表学报, 2004, 25(4): 558-561.
- Luan Hong-xia, Qi Fei-hu. New active contour models for brain MRI segmentation [J]. Chinese Journal of Scientific Instrument, 2004, 25(4): 558-561.
- [4] Yoon Y, Kim J S. Adaptable fuzzy C-means for improved classification as a preprocessing procedure of brain parcellation[J]. Journal of Digital Imaging, 2001, 14(2): 238-240.
- [5] Clarke L. MRI segmentation methods and applications[J]. Magnetic Resonance Imaging, 1995, 13(3): 343-368.
- [6] 王华峰,陈武凡. 基于数学形态学的脑组织图像自动分割[J]. 第一军医大学学报, 2004, 24(7): 824-826.
- Wang Hua-feng, Chen Wu-fan. Automatic segmentation of brain tissue images based on mathematic morphology[J]. Journal of First Military Medical University, 2004, 24(7): 824-826.
- [7] Bezdek J C. Convergence theorem for the fuzzy ISO-DATA clustering algorithm[J]. IEEE Transactions on Pattern Analysis, 1980, 3(2): 1-8.
- [8] Zhang Dao-qiang, Chen Song-can. A novel kernelized fuzzy C-means algorithm with application in medical image segmentation [J]. Artificial Intelligence in Medicine, 2004, 32(1): 37-50.
- [9] 李彬,陈武凡. 基于模糊聚类空间模型的非均匀 MRI 图像分割[J]. 医疗卫生设备, 2006, 27(2): 3-4.
- Li Bin, Chen Wu-fan. Inhomogeneous magnetic resonance image segmentation by spatial model to fuzzy clustering[J]. Chinese Medical Equipment Journal, 2006, 27(2): 3-4.
- [10] He Ren-jie, Datta S, Sajja B R. Adaptive FCM with contextual constrains for segmentation of multi-spectral MRI[C]//Proceedings of the 26th Annual International Conference of the IEEE EMBS, USA, 2004: 1660-1663.