

基于多目标规划的模糊 C 均值聚类算法

王丹丹 李彬 陈武凡

(南方医科大学生物医学工程学院, 广州 510515)

摘要 模糊 C 均值聚类算法(FCM)是一种非常经典的非监督聚类技术,已被广泛地应用到医学图像分割。由于传统的 FCM 聚类算法在分割图像时仅利用了图像的灰度信息,未利用图像的空间信息,在分割叠加了噪声的磁共振(MR)图像时分割效果不理想。考虑到脑部 MR 图像真实的灰度值具有分片为常数的特性,按照合理利用图像空间信息的原则,对传统的 FCM 聚类算法进行了改进,引入多目标规划的概念,提出了一种新的,更加合理的应用图像空间信息的聚类算法。实验结果表明,应用该算法可以有效地分割含有噪声的图像。

关键词 图像分割 模糊 C 均值聚类算法 多目标规划 图像的空间信息

中图分类号:TP391 文献标识码:A 文章编号:1006-8961(2008)08-1492-04

An Improved FCM Algorithm Based on Multiple Objective Programming

WANG Dan-dan, LI Bin, CHEN Wu-fan

(School of Biomedical Engineering, Southern Medical University, Guangzhou 510515)

Abstract Fuzzy C-means(FCM) clustering algorithm is one of well-known unsupervised clustering techniques, which has been widely used in medical image segmentation. However, when the conventional FCM algorithm is used for image segmentation, no spatial information is taken into account. This causes the FCM algorithm to generate unexpected results of segmentation when dealing with noise of magnetic resonance images(MRI). Considering the intensities of ideal MRI are piecewise constant, we present an improved model to FCM algorithms using multiple objective programming. The algorithm can reasonably use the spatial information of the image and improve the accuracy of segmentation. The new algorithm is applied to both synthetic images and magnetic resonance images and the results show better robustness of our algorithm to noise and other artifacts than the conventional fuzzy image segmentation algorithms.

Keywords image segmentation, fuzzy C-means algorithm, multiple objective programming, spatial information of image

1 引言

图像分割是医学图像处理的重要技术,是由图像处理到图像分割的关键技术。医学图像分割就是按照图像的解剖学特性将图像分割成不同的组织类别。常用的分割方法有阈值法、区域生长法和聚类法等。属于聚类法的模糊 C 均值算法(FCM)是一种非常经

典的非监督聚类算法,其依据最小二乘原理,采用迭代方法优化目标函数,最终获得对数据集的模糊划分,已被广泛应用于脑部图像的分割研究^[1,2]。

模糊 C 均值聚类算法是一种基于灰度的聚类算法,在聚类过程中各像素相互独立,未考虑相邻像素之间的影响,未能利用图像的空间信息,在分割叠加了噪声的图像时,分割效果很不理想。许多文献提出了利用图像空间信息的改进算法,提高了对低信噪比

基金项目:国家 973 计划项目(2003CB71610)

收稿日期:2006-11-20;改回日期:2007-04-09

第一作者简介:王丹丹(1981~),女。南方医科大学生物医学工程专业硕士研究生。主要研究方向为医学图像处理。E-mail:jingzou.net@163.com

通讯作者:陈武凡。E-mail:chenwf@fimmu.com

图像的分割精度^[3-5]。但是,这些算法一方面会增加计算复杂度,降低分割速度,另一方面由于未能充分利用医学图像的性质,难以达到较好的分割效果。

传统的模糊 C 均值只涉及一个目标,即求一组变量在满足约束条件下,使目标函数达到极大值或极小值。但是对于 MR 图像,不仅要考虑图像的灰度信息,还要考虑图像的空间信息,才能做出合理的分割。考虑到图像的空间信息,本文改进了传统的模糊 C 均值聚类算法的目标函数,把单目标问题转化为多目标问题,增加了关于隶属度函数的二级目标,得到了一种新的模糊 C 均值聚类算法。即基于多目标规划的模糊 C 均值聚类算法(MOP-FCM)。

2 传统的模糊 C 均值聚类算法

传统的 FCM 是由 Bezdek 等人提出的^[6],其基本思想就是通过迭代寻找聚类中心 v_i 和隶属度函数 u_{ik} ,使得目标函数:

$$J_{FCM} = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p \|y_k - v_i\|^2 \quad (1)$$

达到最小,以实现图像的优化分割。

式中, $\{y_k, k=1, 2, \dots, N\}$ 表示图像灰度值的集合。 c 为预定的类别数目, $v_i, i=1, 2, \dots, c$ 为每个聚类的中心, u_{ik} 是第 k 个样本对于第 i 类的隶属度函数。 $p > 1$ 是一个可以控制聚类结果的模糊程度的常数。

在满足 $\sum_{i=1}^c u_{ik} = 1$ 的条件下,根据 Lagrange 乘数法,可得目标函数式(1)取得极小值的必要条件为

$$u_{ik} = \frac{\left(\frac{1}{\|y_k - v_i\|^2}\right)^{\frac{1}{(p-1)}}}{\sum_{j=1}^c \left(\frac{1}{\|y_k - v_j\|^2}\right)^{\frac{1}{(p-1)}}} \quad (2)$$

$$v_i = \frac{\sum_{k=1}^N u_{ik}^p y_k}{\sum_{k=1}^N u_{ik}^p} \quad (3)$$

3 MOP-FCM 聚类算法

脑部 MR 图像真实的灰度值具有分片为常数的特性,相邻像素之间的关联关系比较紧密。由于传统的 FCM 聚类算法在聚类过程中仅考虑了每个像素的灰度值分别与各聚类中心的距离,未考虑相邻

像素之间的影响,未能利用图像的空间信息,因此在分割叠加了噪声的脑部 MR 图像时,会产生较大的偏差。针对脑部 MR 图像的特性,从运筹学的多目标规划得到启发^[7],提出了一种对传统的 FCM 聚类算法的目标函数进行改进的新模型。即增加了隶属度函数的二级目标函数,使得目标函数变为

$$J = \alpha \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p \|y_k - v_i\|^2 + (1 - \alpha) \sum_{i=1}^c \sum_{k=1}^N \frac{1}{N_r} \sum_{r,y_r \in R_k} \|u_{ir} - u_{ik}\|^2$$

换一种目标函数的形式:

$$J = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p \|y_k - v_i\|^2 + \beta \sum_{i=1}^c \sum_{k=1}^N \frac{1}{N_r} \sum_{r,y_r \in R_k} \|u_{ir} - u_{ik}\|^2 \quad (4)$$

式中, R_k 是第 k 个像素的邻域。 N_r 是邻域范围常数。 $\beta = \frac{1-\alpha}{\alpha}$ 是控制隶属度函数二级目标作用权重的常数。选取合适的隶属度和均值使得目标函数最小。

为使 J_{FCM} 最小,用 Lagrange 乘数法

$$F = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p \|y_k - v_i\|^2 + \beta \sum_{i=1}^c \sum_{k=1}^N \frac{1}{N_r} \sum_{r,y_r \in R_k} \|u_{ir} - u_{ik}\|^2 + \lambda \left(1 - \sum_{i=1}^c u_{ik}\right)$$

取 $p = 2$, 分别求 u_{ik}, v_i 的偏导数,并令其偏导数为 0, 解得

$$u_{ik} = \frac{-\lambda/2 + \beta \frac{1}{N_r} \sum_{r \in R_k} u_{ir}}{(y_k - v_i)^2 + \beta} \quad (5)$$

其中, $\lambda = 2 \times \frac{\beta \frac{1}{N_r} \sum_{r \in R_k} u_{ir}}{\sum_{i=1}^c \frac{1}{(y_k - v_i)^2 + \beta} - 1}$

$$v_i = \frac{\sum_{k=1}^N u_{ik}^p y_k}{\sum_{k=1}^N u_{ik}^p} \quad (6)$$

这样就得到了一种新的聚类算法。即 MOP-FCM。其具体执行步骤如下:

(1) 初始化参数:根据聚类对象设置聚类数 c , 给控制二级目标项作用权重的参数 β 赋值, 设定迭代数 N_1 。

(2) 用传统的模糊 C 均值聚类算法得到图像的

初始分割(可以先通过 C 均值算法寻找初始聚类中心和隶属度的初始值。

(3) 利用式(5)计算并更新各个隶属度 u_{ik} 。

(4) 利用式(6)计算并更新各聚类中心 $v_i, i=1, 2, \dots, c$ 。

(5) 重复第 3 步和第 4 步直到满足收敛条件。一般采用的收敛条件是迭代步数 $n = N_1$, 或各聚类中心的变化的绝对值小于一个很小的正数。

(6) 按照最大隶属度原则完成最后的分割。

参数 β 的取值与图像的信噪比有较大的关系, 在图像的分割过程中需要选择合适的取值使得分割效果最好。

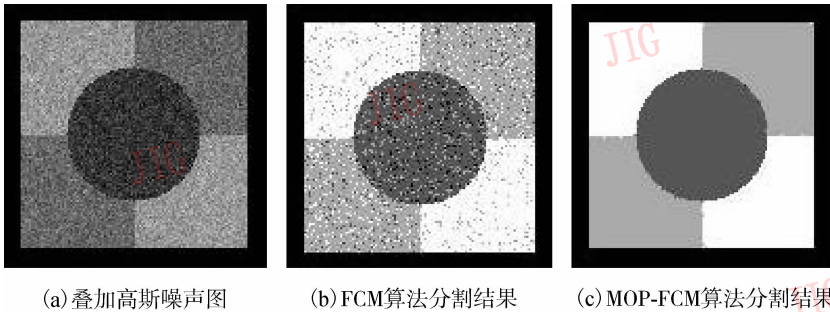


图 1 模拟方块图分割

Fig. 1 Segmentation of a synthetic image

从图 1 可以看出, FCM 算法与 MOP-FCM 算法相比分割结果不够理想, 表现为各个灰度值区域内都有噪声污点。

4.2 模拟脑部 MR 图像

模拟脑部 MR 图像的数据取自 BrainWeb 网站。其模拟脑部 MR 图像数据库常被作为图像分割的金标准。每组数据集由 $181 \times 217 \times 181$ 个像素组成, 层厚为 1mm, T1 加权。设定分割类别为灰质、白质、脑脊液和背景等 4 类。图 2 中, (a), (b) 分别是未添加噪声的模拟脑部 MR 图像(横断面成像的层数 $Z = 72$)及其硬分割标准结果, (c), (e), (g), (i) 分别表示在(a)上迭加了 3%, 5%, 7%, 9% 噪声和 20% 不均匀性的模拟脑部 MR 图像。(d), (f), (h), (j) 分别表示 MOP-FCM 算法对噪声图像分割结果。对于图 2 (c), (e), (g), (i), β 的取值分别为 120, 500, 1 000, 2 000。实验结果表明即使分割较高噪声的图像, MOP-FCM 算法能够得到比较理想的结果。

分别用 MOP-FCM 聚类算法与传统 FCM 聚类算法对模拟脑部 MR 图像分割比较。为了量化地评估各种算法的分割效果, 定义错分率为

4 实验结果

利用 MOP-FCM 聚类算法和 FCM 聚类算法对模拟方块图和脑部 MR 图像进行分割比较。实验中, 参数取值如下: $p = 2, R_k$ 取 2D 图像的 8 邻域, $N_r = 9$, 迭代步数 $N_1 = 5$ 。

4.1 模拟方块图

数据集采样 128×128 像素组成的模拟方块图。图 1 列出了 FCM 算法和 MOP-FCM 算法对叠加了高斯噪声图像的分割结果。实验中取 β 值为 3 000。

$$MCR = \frac{\text{错误分割的像素数量}}{\text{总像素数量}} \times 100\%$$

$MCR \in [0, 1]$, MCR 越小表示分割结果越好。

表 1 给出了实验中, 两种图像分割算法对迭加了不同噪声的模拟脑部 MR 图像的进行分割的错分率。

表 1 仿真脑部 MR 图像迭加不同噪声图像的错分率

Tab. 1 Misclassification ratio (MCR) for simulated MR results

单位: %				
算法	3% 噪声	5% 噪声	7% 噪声	9% 噪声
FCM	3.99	5.21	7.25	9.92
MOP-FCM	3.87	4.60	5.36	6.09

从实验数据可以看出, MOP-FCM 聚类算法, 对不同噪声的图像都优于传统的 FCM 分割算法, 特别在分割噪声较大图像时优越性更明显, 分割效果较好。

4.3 临床脑部 MR 图像

数据集采用 GE 公司的 Magnetom 1.5T 磁共振系统对右岛叶前部海绵状血管瘤进行 T1 MRI 成像, 其参数重复时间 $TR = 500\text{ms}$, 回波时间 $TE = 20\text{ms}$, $251 \times 184 \times 22$ 像素, 层厚为 8mm。在运行 MOP-FCM 算法前, 先去除了脑壳和背景, 提取了脑组织部分。设定分割类别为病灶、灰质、白质、脑脊

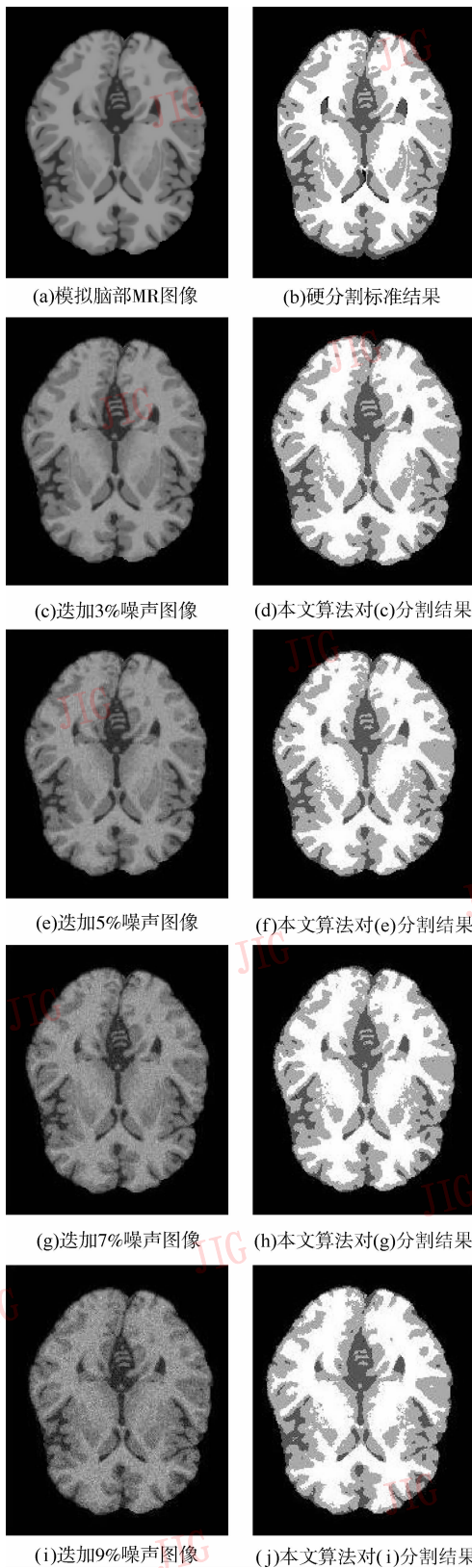
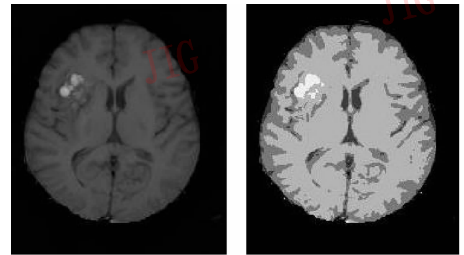


图 2 MOP-FCM 算法对迭加了不同噪声的模拟脑部 MR 图像的分割结果

Fig. 2 The segmentation results about simulated brain MR image corrupted by different noises utilizing the MOP-FCM algorithm

液和背景等 5 类。图 3 为用 MOP-FCM 算法分割临床脑部 MR 图像的结果。

从图像上可以看出, MOP-FCM 算法能很好地分割出图像中病灶区域。



(a)真实MR图像 (b)MOP-FCM算法对(a)分割结果

图 3 MOP-FCM 算法对真实脑部 MR 图像的分割结果

Fig. 3 The segmentation results about brain MR image utilizing the MOP-FCM algorithm

5 结 论

MOP-FCM 聚类算法对传统的 FCM 算法进行改进,引入了隶属度函数的二级目标,在两级隶属度函数同时作用下,使目标函数达到极小值来控制邻域作用,从而合理地利用了空间信息。该算法不仅继承了传统 FCM 聚类算法无监督、算法简单、运算速度快、稳健性好等优点,而且在有效去除噪声,提高图像分割精度方面,有显著提高,是一种比较稳健有效的图像分割方法。

参考文献 (References)

- 1 Zhang D Q, Chen S C. A novel kernelized fuzzy C-means algorithm and segmentation of MRI data [J]. Artificial Intelligence in Medicine, 2004, 32(1):37 ~ 50.
- 2 Pham D L, Prince J L. Adaptive fuzzy segmentation of magnetic resonance images [J]. IEEE Transactions on Medical Image, 1999, 18(9):737 ~ 752.
- 3 Pham D L. Spatial models for fuzzy clustering [J]. Computer Vision and Image Understand, 2001, 84(2):285 ~ 297.
- 4 Ahmed M N, Yamany S M, Mobamed N, et al. A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data [J]. IEEE Transactions on Medical Image, 2002, 21(3): 193 ~ 199.
- 5 Liew A W C, Yan H. An adaptive spatial fuzzy clustering algorithm for 3-D MR image segmentation [J]. IEEE Transactions on Medical Image, 2003, 22(9): 1063 ~ 1075.
- 6 Bezdek J C, Pal S K. Fuzzy Models for Pattern Recognition [M]. Piscataway, NJ, USA, IEEE Press, 1999.
- 7 Song Xue-feng, Wei Xiao-ping. Operations Research [M]. Nangjing: Southeast University Press, 2003: 108 ~ 123. [宋学锋, 魏晓平. 运筹学[M]. 南京:东南大学出版社,2003:108 ~ 123.]