

合理利用空间信息的 MRI 脑部图像分割

王媛媛

(西安电子科技大学 电子工程学院, 陕西 西安 710071)

摘要: 在合理利用空间信息的基础上, 提出了一种更准确, 紧致性和分离性更好的分割算法。该算法首先定义一个空间函数, 并在其中引入一个控制参数, 该参数可以对噪声点、边缘点以及区域内的点进行区别对待, 然后用空间信息更新隶属度。实验结果表明, 该算法效果要明显优于 sFCMpq 算法及其改进算法(EsFCMpq)。

关键词: 图像分割; 模糊 c -均值聚类; 邻域信息; MRI 脑部图像

中图分类号: TP399

文献标识码: A

文章编号: 1674-7720(2011)19-0031-04

The wise use of spatial information for MRI brain segmentation

Wang Yuanyuan

(Department of Electronic Engineering, Xidian University, Xi'an 710071, China)

Abstract: Based on the wise use of spatial information, this paper proposes a more accurate segmentation algorithm with samples compact within one cluster and separated among different clusters. The algorithm firstly defines a spatial function by introducing a control parameter, which can make the noise points, edge points and points within the region be treated differently. Then it uses the spatial information to update the membership function. Experimental results show that the proposed algorithm outperforms the sFCMpq algorithms and its improved algorithm EsFCMpq in various cluster validity functions.

Key words: image segmentation; FCM clustering algorithm; neighborhood information; MRI brain image

图像分割是图像分析和模式识别的首要问题, 它是图像分析和模式识别系统的重要组成部分, 并决定图像的最终分析质量和模式识别的判别结果^[1]。医学图像分割长期以来一直是图像处理的研究热点, 由于人体解剖结构的复杂性、组织器官形状的不规则性、不同个体的差异性等原因, 使得到目前为止, 还无法得到一种能对所有图像进行有效分割的分割算法。目前, 图像分割算法主要包括基于边界、基于阈值、基于模糊集理论、基于区域的方法。由于 MR 图像成像设备获取图像的不确定性或模糊性, 造成不同个体组织之间难以找到清晰的边界, 而模糊聚类法是一种有效的方法。在脑部 MRI 图像的分割中, 最具代表性的算法是模糊 c -均值聚类算法(FCM)。传统的 FCM 算法由 DUNN J C^[2]提出, 后来由 BEZDEK J C^[3]进行改进。FCM 算法采用迭代优化目标函数, 最终获得对数据集的模糊划分。该算法的缺点是仅利用了灰度信息的聚类算法, 没有考虑相关像素之间的相关性, 未能利用图像的空间信息, 这就导致了图像

分割的不准确性^[4-5]。近几年来, 很多文献都着力于利用图像空间信息的改进的 FCM 算法, 提高了对低信噪比图像的分割精度^[6-7]。目前, 结合空间信息的 FCM 算法主要有两种, 一种是改进目标函数, 在目标函数中加入空间信息; 另一种是改进隶属度函数, 在隶属度函数中加入空间信息。本文提出的算法是后一种情况。本算法首先定义一个空间函数, 在空间函数中引入一个控制参数, 该参数可以对噪声点、边缘点以及区域内的点都进行区别对待, 然后用空间信息更新隶属度。实验结果表明, 该算法的效果要明显优于 sFCMpq 算法及其改进算法(EsFCMpq)。

1 算法介绍

1.1 经典 FCM 算法

FCM 算法是通过目标函数进行迭代优化, 进而对数据样本进行模糊聚类的一种方法, 分类结果用一个模糊隶属度矩阵 $U = \{u_{ik}\} \in R^{cn}$ 来表示。对于图像分割, 数据样本集就是 N 个像素, 通过 FCM 算法把这 N 个像素

分成 C 个类, 得到 C 个类中心和模糊隶属度矩阵, 其中对于 u_{ik} , 它表示第 k 个像素划分为第 i 个类的程度, 即隶属度。FCM 的目标函数^[6]定义为:

$$J_m(\mathbf{U}, \mathbf{V}) = \sum_{k=1}^N \sum_{i=1}^C (u_{ik})^m d^2(x_k, v_i) \quad (1)$$

其中, $\sum_{i=1}^C u_{ik} = 1, u_{ik} \in [0, 1]; V = \{v_1, v_2, \dots, v_C\}$ 表示样本的 C 个聚类中心; $d^2(x_k, v_i) = \|x_k - v_i\|^2$, 表示第 k 个样本到第 i 类中心的欧式距离; m 为模糊加权指数 ($1 \leq m \leq \infty$), m 值越大, 分类的模糊程度越高, 通常情况下, 取 $m=2$ 。

FCM 的实质就是一个将目标函数 $J_m(\mathbf{U}, \mathbf{V})$ 最小化的迭代收敛过程。为了使 $J_m(\mathbf{U}, \mathbf{V})$ 取得最小值, 隶属度和聚类中心用式(2)、式(3)进行更新:

$$u_{ik} = \left[\sum_{j=1}^C \left(\frac{d^2(x_k, v_j)}{d^2(x_k, v_i)} \right)^{\frac{1}{m-1}} \right]^{-1} \quad (2)$$

$$v_i = \frac{\sum_{k=1}^N \sum_{i=1}^m u_{ik} x_k}{\sum_{k=1}^N \sum_{i=1}^m u_{ik}}, 1 \leq i \leq C \quad (3)$$

1.2 sFCMpq 算法及其改进算法

传统的 FCM 聚类算法仅利用了灰度信息, 没有考虑相邻像素之间的相关性, 导致了图像分割的不准确性。图像中相邻的像素具有相似的特征值, 并且它们属于同一类的概率也非常大^[9]。根据这一特点, 参考文献[10]提出了一种 sFCMpq 算法, 在利用灰度信息的同时, 也考虑了空间邻域信息。该方法首先定义了一种空间函数:

$$h_{ij} = \sum_{k \in \Omega(x_j)} u_{ik} \quad (4)$$

其中, $i=1, 2, \dots, C; j=1, 2, \dots, N; \Omega(x_j)$ 表示在空域以像素 x_j 为中心的矩形窗口; u_{ik} 表示像素 x_k 属于第 i 类的程度, 即隶属度; 空间函数 h_{ij} 表示像素 x_j 属于第 i 类的可能性。然后用空间函数对隶属度进行更新, 公式如下:

$$u'_{ij} = \frac{\mu^p h_{ij}^q}{\sum_{k=1}^C \mu^p h_{ik}^q} \quad (5)$$

其中, 参数 p 和 q 用于控制两个函数的相对重要性, 仿真结果表明 $p=0, q=2$ 时可以得到最优的聚类结果^[10]。

参考文献[7]对 sFCMpq 算法定义的空间函数进行改进。在本文中, 该算法记为 EsFCMpq。在引入空间信息的时候, 对不同的点进行了区分对待^[11], 用一个控制参数 α_k 控制空间信息, 改进的空间函数为:

$$h_{ij} = \sum_{k \in \Omega(x_j)} \alpha_{jk} u_{ik} \quad (6)$$

$$\alpha_{jk} = \begin{cases} 1, & |x_k - \bar{m}| < \sigma_j \\ 0, & |x_k - \bar{m}| > \sigma_j \end{cases} \quad (7)$$

其中, $k \in \Omega(x_j)$, \bar{m} 表示原图像中以像素 x_j 为中心的邻域

$\Omega(x_j)$ 内像素的均值, σ_j 是邻域 $\Omega(x_j)$ 内像素的方差值, $|x_k - \bar{m}|$ 表示的是 x_k 与 \bar{m} 之间的绝对距离。

1.3 本文算法

sFCMpq 算法定义的空间函数是在规定的矩形窗口邻域中各个像素的隶属度之和, 再利用对其进行归一化均值以更新隶属度, 这个过程相当于一个平滑的过程, 可使因为偏差场造成的图像灰度分布不均匀引起的误发现象得到改善。但是, 在区域和区域的边界处, sFCMpq 算法认为每个像素受到邻域中各个像素影响程度是相同的, 这样, 在边缘处会出现过平滑的现象, 造成边缘区域分类的错误^[9]。参考文献[9]提出的改进算法 (EsFCMpq 算法) 在定义的空间函数中引入了一个控制参数 α_{ik} , 其作用是如果经过判断得出某一点为噪声点或者边缘点, 则在对其隶属度更新的过程中不予考虑; 反之, 则说明像素点是一个区域的内部点, 用其对隶属度进行更新, 这样可以使区域内部更加均匀。该方法实质上是引用了阈值平均的思想, 通过设置阈值来减少由于平均带来的边缘模糊问题。

EsFCMpq 算法通过比较像素点和邻域均值的差的绝对值与邻域方差的关系来确定控制参数, 如果差的绝对值大于邻域方差, 便认为该点是噪声或者是边缘点, 对该点不予考虑; 反之, 则认为是区域内的点。该算法认为区域内的点对中心像素的影响程度是相同的, 但实际上, 在要处理像素的邻域内, 各个像素均对要处理的中心像素有一定的影响, 只是影响的程度不同。从这个角度来说, EsFCMpq 算法的分割结果不够精确, 所以本文算法在 EsFCMpq 算法的基础上进行了改进:

$$h_{ij} = \sum_{k \in \Omega(x_j)} p(x_j, x_k) u_{ik} \quad (8)$$

由式(8)可以看出, 本文的控制参数为 $p(x_j, x_k)$, 用它来代替 α_{ik} 。

h_{ij} 表示加入邻域像素空间信息后, 中心像素 x_j 属于第 i 类的可能性。 h_{ij} 和 u_{ik} 的值都在 $[0, 1]$ 之间。如果 $\Omega(x_j)$ 中所有的像素都属于第 i 类, 即 $u_{ik}=1$, 其中 x_k 为邻域像素, 则 $h_{ij}=1$ 。也就是说, 中心像素 x_j 主要被它的邻域像素影响, 根据式(8), 可以得出 $\sum_{k \in \Omega(x_j)} p(x_j, x_k)=1$ 。与此同时, $p(x_j, x_k)$ 还要满足以下条件: x_k (邻域像素) 与 x_j (中心像素) 的距离越大, $p(x_j, x_k)$ 的值应该越小。

因此, 控制参数 $p(x_j, x_k)$ 可以写成如下形式:

$$p(x_j, x_k) = \left(\sum_{l \in \Omega(x_j)} \frac{d^2(x_j, x_l)}{d^2(x_j, x_j)} \right)^{-1} \quad (9)$$

在确定控制参数的过程中, 引入模糊的思想, 即假设待处理的像素为 x_j , 其邻域为 $\Omega(x_j)$, 在 $\Omega(x_j)$ 中, x_j 为中心像素, 当邻域内的像素 x_k 与其距离越大, 说明 x_k 影响 x_j 的程度越小; 反之, 若 x_k 与 x_j 距离越小, 说明其影响 x_j 的程度越大。边缘点或者噪声点与中心像素点的距

离最大,则对中心像素的影响最小;而其他非噪声或者边缘点对中心像素的影响程度由它们自身与中心像素之间的距离而决定。 $p(x_j, x_k)$ 的取值0~1之间的数表示,数值越接近1,表示 x_k 与中心像素 x_j 的相关性越大,在考虑邻域信息时,该点对中心像素的影响力大;相反,数值越接近0,表明相关性越小。 $p(x_j, x_k)$ 的值不采用单纯的0、1来表示,体现了模糊划分的特点,提高了准确性。为了提高算法的收敛速度,引入FCM的快速算法确定初始聚类中心^[13]。具体的执行步骤如下:

- (1) 确定聚类数目 C ,模糊加权指数 m ,迭代终止阈值 ε ,最大迭代次数 $item_{max}$;
- (2) 运用快速FCM算法,计算初始聚类中心 V_0 ;
- (3) 用式(8)、式(9)和式(5)对隶属度更新后,再对聚类中心进行更新;
- (4) 重复步骤(3)直到各聚类中心收敛或者迭代次数达到 $item_{max}$;
- (5) 根据隶属度最大原则对图像进行分割。

2 聚类有效性函数

为了定量评价分割的性能,给出 V_{pc} 和 V_{pe} 两个参数, V_{pc} 为分割系数, V_{pe} 为分割熵,当 V_{pc} 达到最大或者 V_{pe} 达到最小时,得到最佳的分割结果^[14-15]。但是 V_{pc} 和 V_{pe} 缺少与特征属性的直接联系,而 V_{fs} 和 V_{sb} 是基于特征结构的有效性函数。利用数据集的特征结构,一种鲁棒性好的聚类分割结果要求在一个类中样本是紧凑的,在不同类之间,样本是分离的。而 V_{fs} 和 V_{sb} 就是从特征结构方面来评价聚类性能的参数。 V_{fs} 为Fukuyama-Sugeno函数, V_{sb} 为Xie-Beni函数,当 V_{fs} 或者 V_{sb} 达到最小时,就能得到较好的聚类结果^[16-17]。具体计算公式如下:

$$\left\{ \begin{array}{l} V_{pc}(U) = \frac{\sum_{j=1}^N \sum_{i=1}^C u_{ij}^2}{N} \\ V_{pe}(U) = -\frac{1}{N} \left\{ \sum_{j=1}^N \sum_{i=1}^C [u_{ij} \log u_{ij}] \right\} \\ V_{fs}(U, V, X) = \sum_{i=1}^C \sum_{j=1}^N u_{ij}^m (\|x_j - v_i\|^2 - \|\bar{v}_i - \bar{v}\|^2) \\ V_{sb}(U) = \frac{\sum_{i=1}^C \sum_{j=1}^N u_{ij}^2 \|x_j - v_i\|^2}{N \times (\min_{i \neq k} \{\|v_i - v_k\|^2\})} \end{array} \right. \quad (10)$$

其中, $\bar{v} = \frac{1}{C} \sum_{i=1}^C v_i$

3 实验结果

实验数据分别是lena灰度图像和真实的人体脑部MRI图像,取 $m=2$ 、 $\varepsilon=1e-6$ 、 $N_k=9$,lena灰度图像的分类数 $C=2$,MRI脑部图像的分类数 $C=4$ 。对图像分别用sFCMpq算法、改进的sFCMpq算法以及本文算法进行分割,分割结果用4个聚类有效性函数进行描述。

对图1分别用sFCMpq、EsFCMpq和本文算法进行分割后的结果如图2~图4所示。



图1 lena 原图像



图2 sFCMpq 分割结果



图3 改进的 sFCMpq (EsFCMpq) 的分割结果



图4 本文算法的分割结果

使用聚类有效性参数对算法的性能进行比较,结果如表1所示。

表1 三种算法的比较

	V_{pc}	V_{pe}	V_{fs}	V_{sb}
sFCMpq	0.89497	0.17174	-6.82E+07	0.097333
EsFCMpq	0.91036	0.14825	-6.99E+07	0.097644
本文算法	0.91519	0.14006	-7.39E+07	0.093002

对图5分别用sFCMpq、EsFCMpq和本文算法进行分割的结果如图6~图8所示,分割的结果依次为背景、脑白质、脊髓液、脑灰质。从分割结果可以看到,sFCMpq、EsFCMpq算法的分割效果都没有本文算法进行分割的效果好,sFCMpq算法使区域内部数据比较均匀,但是在一些边界处出现了过平滑的现象,一些细节及边缘未能检测到,比如像脊髓液等的检测没有后两种方法的效果好;EsFCMpq引入控制因子进行改进,结果有了一定的改善,区域内部数据更加均匀,但它没有考虑到邻域内除了边缘点和噪声点外其他像素对中心像素影响的程度是不一样的,分割结果不准确,有的脊髓液及边缘没有被分割出来。

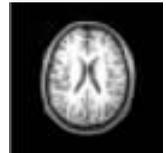


图5 原始脑部MRI图像



图6 sFCMpq 分割结果



图7 改进的 sFCMpq (EsFCMpq) 的分割结果



图8 本文算法的分割结果

使用聚类有效性参数对算法的性能进行比较,结果如表 2 所示。

表 2 三种算法的比较

	V_{pe}	V_{pc}	V_{fs}	V_{sb}
sFCMpq	0.915 53	0.143 18	-6.47E+08	0.056 167
EsFCMpq	0.936 31	0.108 5	-6.79E+08	0.041 819
本文算法	0.950 24	0.086 897	-7.10E+08	0.041 101

以上结果表明,无论是真实图像还是合成图像,从 v_{pe} 和 v_{pc} 两个参数来看,本文算法在分割精确性上优于 sFCMpq 和 EsFCMpq 算法;从 v_{fs} 和 v_{sb} 两个参数看,本文算法在紧致性和分离性上要优于 sFCMpq 和 EsFCMpq 算法。

传统的 FCM 算法分割并不理想,原因在于它只考虑了图像的灰度信息。本文算法既考虑了灰度信息又合理地利用了图像的空间信息。在空间信息统计中引入一个改进的控制参数来区分噪声、边缘点和区域内部的点,并对区域内部的点进行区别对待,既能控制邻域信息的使用,避免边缘过平滑的现象,又能更加合理地利用空间信息。实验结果表明,与 sFCMpq、EsFCMpq 算法相比,该算法分割结果的精确性更高,分割结果有更好的紧致性和分离性,是一种鲁棒性更好的聚类算法。

和 EsFCMpq 存在的问题一样,由于加入了空间信息,并且引入了控制参数,在计算量上要比 FCM、sFCMpq、EsFCMpq 都有所增加,这是该算法存在的问题。参考文献

- [1] CHENG H D, JIANG X H, SUN Y, et al. Color image segmentation: advances and prospects [J]. Pattern Recognition, 2001, (34):2259-2281.
- [2] DUNN J C. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters [J]. Journal of Cybernetics, 1973, 3:32-57.
- [3] BEZDEK J C. Pattern recognition with fuzzy objective function algorithms [M]. New York: Plenum Press:1981.
- [4] 余学飞.基于模糊理论的医学图像分割算法研究[D].广州:南方医科大学,2009.
- [5] 辛学刚,卢振泰,陈武凡.融入空间信息的医学图像优质分割[J].计算机工程与应用,2009,45(34):225-226.
- [6] TOLIAS Y A, PANAS S M. Image segmentation by a fuzzy clustering algorithm using adaptive spatially constrained membership functions[J]. IEEE Transactions on Systems, Man and Cybernetics, 1998, 28(3):359-369.
- [7] MOHAMED N A. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data [J]. IEEE Transactions on Med Image, 2002, 21(3):193-199.
- [8] CHEN W J, GIGER M L, BICK U. A fuzzy c-means

(FCM)-based approach for computerized segmentation of breast lesions in dynamic contrast-enhanced MR images [J]. Academic Radiology, 2006, 13(1):63-72.

- [9] 张兰,王珂,杨文宏.一种结合空间信息的 FCM 算法对脑 MR 图像的分割 [J]. 计算机工程与应用, 2007, 43(26): 203-205.
- [10] CHUANG K S, TZENG H L, CHENS, et al. Fuzzy c-means clustering with spatial information for image segmentation [J]. Computerized Medical Imaging and Graphics, 2006(30):9-15.
- [11] 李斌,陈武凡.基于模糊聚类空间模型的非均匀 MR 图像分割[J].医疗卫生设备,2006,27(2):3-4.
- [12] LUNG H V, KIM J M. A generalized spatial Fuzzy C-Means algorithm for medical image segmentation [J]. IEEE International Conference on Fuzzy Systems, FUZZ-IEEE, 2009:409-414.
- [13] 李志梅,肖德贵.快速模糊 C 均值聚类的图像分割方法 [J]. 计算机工程与应用, 2009, 45(12):187-189.
- [14] BEZDEK J C. Cluster validity with fuzzy sets [J]. Journal of Cybernetics, 1974, 8(3):58-73.
- [15] BEZDEK J C. Mathematical models for systematic and taxonomy [C]. Proceedings of 8th International Conference on Numerical Taxonomy, 1975:143-166.
- [16] XIE X L, BENI G. A validity measure for fuzzy clustering [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13(8):841-847.
- [17] FUKUYAMA Y, SUGENO M. A new method of choosing the number of clusters for the Fuzzy C-Means Method [C]. Proceedings of 5th Fuzzy Systems Symposium, 1989:247-250.
- [18] POPESCU M, BEZDEK J C, KELLER J M, et al. A new cluster validity measure for bioinformatics relational datasets [C]. IEEE International Conference on Fuzzy Systems, 2008: 726-731, .
- [19] BALAFAR M A, RAMLI A R, MASJOHOR S, et al. Compare different spatial based Fuzzy -C-Mean (FCM) extensions for MRI image segmentation [J]. The 2nd International Conference on Computer and Automatic Engineering (ICCAE), 2010, 5(1):609-611.
- [20] AYECH M W, KALTI K E, AYE B E [C]. International Conference on Pattern Recognition, 2010:2306-2309.

(收稿日期:2011-07-20)

作者简介:

王媛媛,女,1984年生,硕士研究生,主要研究方向:医学图像处理。