

融合万有引力和局部熵的FCM算法*

崔兆华^{1,2}, 高立群¹, 马红宾³, 李洪军⁴

(1. 东北大学 信息科学与工程学院, 沈阳 110819; 2. 解放军第65041部队, 沈阳 110113; 3. 白求恩国际和平医院 院务部营房科, 石家庄 050081; 4. 白城医学高等专科学校 公共教学部, 吉林 白城 137000)

摘要: 为了克服传统模糊C-均值(fuzzy C-means, FCM)聚类算法特征描述单一、易受复杂灰度影响而出现误分割的缺点, 将万有引力和图像局部熵融入到FCM算法。算法首先引入图像局部信息熵来描述节点(像素点)间的特征, 同时计算新节点的同质值; 其次, 将该同质值看做新节点的质量, 节点之间通过万有引力算子形成关联, 使节点灰度特征和节点空间位置特征有效结合, 以此解决传统FCM算法节点特征描述孤立的缺陷。最后, 对三类典型的灰度分布不均的医学图像进行仿真实验, 结果表明改进算法获得了更加精确的分割结果。

关键词: 图像分割; 模糊C-均值聚类算法; 图像局部熵; 万有引力算子

中图分类号: TP391 **文献标志码:** A **文章编号:** 1001-3695(2013)12-3828-03

doi:10.3969/j.issn.1001-3695.2013.12.081

New FCM algorithm fused universal gravitation and local entropy

CUI Zhao-hua^{1,2}, GAO Li-qun¹, MA Hong-bin³, LI Hong-jun⁴

(1. School of Information Science & Engineering, Northeastern University, Shenyang 110819, China; 2. 65041 Troops of PLA, Shenyang 110113, China; 3. Dept. of Barracks, Bethune International Peace Hospital, Shijiazhuang 050081, China; 4. Baicheng Medical College, Baicheng Jilin 137000, China)

Abstract: To overcome the shortcomings of the traditional fuzzy C-means (FCM) clustering algorithm which were simple image feature description and easy distributed by complex grey influence with wrong segmentation, this paper proposed an improved FCM algorithm for image segmentation, combined with universal gravitation principle and local entropy theorem. Firstly, it introduced the image local entropy to accurately measure image node property between two adjacent nodes, and meanwhile computed the node homogeneous value. Then it was taken as the node quality, formed closely relationship using gravity algorithm which made the node grey feature and spatial position combine effectively. The above method solved the problem of the description of node feature isolation of traditional FCM algorithm. Finally, the simulation results show that the presented algorithm can obtain more precise segmentation results from three types of grey non-uniform distribution medical images.

Key words: image segmentation; FCM clustering algorithm; image local entropy; gravity algorithm

0 引言

图像分割是指将图像中具有特殊意义的不同区域分开, 这些区域是相互不相交的, 每个区域满足灰度、纹理、颜色等特征的某种相似性准则^[1]。目前主要的图像分割方法有直方图阈值法、特征空间聚类、基于区域的方法、边缘检测方法、模糊方法、神经网络方法六类^[1]。其中, 模糊方法由于符合人眼视觉的主观性而应用广泛, 目前已经提出了很多种基于模糊聚类的图像分割算法^[2]。

模糊聚类算法中最为著名的是模糊C-均值聚类(FCM)算法, 它通过对耗散函数的迭代优化实现对像素的划分, 根据每个像素属于不同区域的程度将像素分到某个区域中^[3]。作为一种无监督的聚类算法, FCM已成功地应用于医学图像分割^[4,5]。但是, 传统的FCM算法^[6]主要有两点不足: a) 由于采用经典的欧氏距离测度, 从而导致算法仅对球形数据聚类形式

鲁棒性好, 大大降低了算法对其他形式数据的应用范围; b) 进行图像分割时, 仅仅在特征空间中进行聚类, 而忽略了图像中像素点间的空间邻域信息, 使得传统FCM算法对结构复杂图像和含噪声图像比较敏感。针对以上缺点, 研究者提出了许多FCM的改进算法, 其中具有代表性的有:

a) 采用其他距离测度公式来代替欧氏距离度量以增强FCM的抗噪性能。如文献[4]中利用一种新颖的非相似度量指标代替原有的距离度量将局部空间上下文信息与非局部信息结合, 并将其用于脑核磁共振图像的分割, 取得了较好的效果。文献[7]引入特征散度来进行距离测度, 并结合聚类有效性指数确定初始聚类数目。

b) 结合区域信息, 克服FCM仅利用图像灰度信息的缺点。典型的有文献[5], 该方法利用分水岭分割方法对原图像进行预分割, 再对小区建立包含邻域信息的聚类目标函数, 充分发挥了分水岭算法的优势, 提高了分割精度。文献[8]不但考虑了像素的空间属性还考虑了像素的纹理属性, 利用Gabor函

收稿日期: 2013-01-15; **修回日期:** 2013-03-18 **基金项目:** 国家自然科学基金资助项目(51005042); 国家“863”计划资助项目(2012AA062002); 中央高校基本科研业务费资助项目(N100403005)

作者简介: 崔兆华(1981-), 女, 吉林白城人, 助理工程师, 博士研究生, 主要研究方向为图像分割(neuczh2006@sina.com); 高立群(1949-), 男, 教授, 博士, 主要研究方向为图像处理、智能优化; 马红宾(1977-), 男, 助理工程师, 学士, 主要研究方向为数字化营房管理; 李洪军(1963-), 男, 讲师, 学士, 主要研究方向为数据建模、数据结构。

数模拟人眼视觉提取图像的纹理特征,并以加权的形式与图像灰度特征共同作用,较精确地实现了图像的分割。

综上所述,改进距离测度方法,拓宽图像特征提取范围是增强图像分割准确性的有效途径。本文基于这种思想,在充分考虑到图像像素点的灰度、空间位置信息的基础上,提出一种融合万有引力算子和图像局部熵的 FCM 图像分割算法。

1 融合万有引力和局部熵的 FCM 算法

传统 FCM 算法衡量节点间相似性采用单一的灰度特征,为了提高 FCM 算法对复杂灰度分布图像的分割精度,本文结合万有引力算子及图像局部熵有效地表达了节点灰度特征和空间分布的关系。

1.1 万有引力算子

将图像中的每个像素看成宇宙中的一个物体,这样,一幅 $M \times N$ 的图像看成是由 $n = M \times N$ 个质量均匀的物体(节点)组成的集合,每个节点对周围的节点都存在引力,反之,这个节点也接受周围节点给它的力,这个力即通过万有引力来计算。

牛顿万有引力定律是一个向量的形式,它由引力方向和引力大小来确定。以一幅 8×8 大小的图像为例来说明图像节点对应的万有引力。其中,小黑点代表图像中节点的分布情况,大黑点为图像上任意两个质量分别为 m_1 和 m_2 的节点,这两个节点之间形成的万有引力如图 1 所示。

图 1 中节点 1 和节点 2 之间的引力可以用公式表述为

$$f_{1,2} = G \frac{m_1 m_2 \hat{r}_{2,1}}{\|r_{2,1}\|^2} = G \frac{m_1 m_2 r_{2,1}}{\|r_{2,1}\|^3} \quad (1)$$

其中: G 为万有引力常数; r_1 、 r_2 分别是两个节点的位置向量; $r_{2,1} = \|r_2 - r_1\|$ 表示两节点的距离; $\hat{r}_{2,1} = (r_2 - r_1) / \|r_2 - r_1\|$ 为节点 1 和 2 的单位向量。同样,物体 1 和 2 的向量引力可以表示为

$$f_{2,1} = -f_{1,2} \quad (2)$$

目前万有引力算子的效果已在图像边缘检测和随机游走图像分割算法中得到了验证(详见文献[9,10])。因此,本文在基于模糊聚类的图像分割中采用万有引力算子提取灰度和空间特征的公式采用式(1)的形式,在下文将予以详细叙述。

1.2 图像局部熵计算同质值

在基于灰度图像的分割中,图像的聚类结果一般为几块均匀性区域,所以均匀性在图像分割中起到重要作用。本文即采用各节点邻域的平均同质值(也即图像局部熵)来描述节点的质量,计算每个区域内像素的平均同质值,同质值大的像素表示该点和周围邻域像素灰度差小,分在同一区域的概率大;反之,则概率小。根据这种原则,本文即以图像局部熵来计算节点同质值。Shannon^[11]于 1948 年提出熵的概念。熵可以度量概率系统的不确定程度,设 X 是取有限个值的随机变量, $p_i = P\{X = x_i\}$ ($i = 1, 2, \dots, n$), 则 X 的 Shannon 熵定义为

$$H(X) \cong \sum_{i=1}^n p_i \log_2 \frac{1}{p_i} = - \sum_{i=1}^n p_i \log_2 p_i \quad (3)$$

其中: $p_i \geq 0$, $\sum_{i=1}^n p_i = 1$ 。

当系统确定时,某一事件 x_i 的概率为 1,其余概率为 0,熵值为 0 最小;反之,系统越不确定,熵值越大。

熵不仅可以度量概率系统的确定程度,而且可以度量灰度的均匀性程度^[12]。灰度越均匀,熵值越小;反之,灰度越复杂,

熵值越大。本文根据熵可以度量灰度均匀性程度的思想来计算图像灰度的同质值也即节点的质量。对于一幅大小为 $M \times N$ 的图像,选择其局部窗口大小为 $L \times S$,局部熵定义为

$$H_f = - \sum_{i=1}^L \sum_{j=1}^S P(i,j) \log[P(i,j)] \quad (4)$$

$$P(i,j) = f(i,j) / \sum_{i=1}^L \sum_{j=1}^S f(i,j) \quad (5)$$

其中: $f(i,j)$ 为图像灰度值, $P(i,j)$ 表示灰度分布。

如果把每个节点看做物体,节点邻域的平均同质值看做该节点的质量,某一节点与聚类中心节点之间的相互作用力可以通过公式描述如下:

$$G_k(i,j) = G \frac{H_{f(i,j)} H_{v_k}}{\|D\|^2} \quad (6)$$

其中: $G_k(i,j)$ 是节点 $f(i,j)$ 和聚类中心节点 v_k 的万有引力; $H_{f(i,j)}$ 、 H_{v_k} 分别为节点 $f(i,j)$ 和聚类中心点 v_k 的平均同质值; G 为像素万有引力的计算系数。 D 为从 $f(i,j)$ 到 v_k 的归一化距离,表示为

$$\|D\| = \|f(i,j) - v_k\| / \|f(i,j)\| + \|v_k\| \quad (7)$$

以上方法不仅考虑了每个节点的灰度特征,而且也考虑了节点间的距离关系,并在距离和节点灰度之间获得了有效的关系函数。

1.3 融合万有引力和熵的 FCM 算法

融合万有引力算子和熵的 FCM 算法实现步骤如下:

a) 对于输入图像,计算像素点在 x 和 y 方向上的梯度幅值 and 变化方向,以计算该像素点的万有引力特征分量。

b) 计算熵函数,用于估算图像像素分布的均匀程度(也即同质性)。将同质值作为该点像素的质量来计算图像邻域像素的万有引力。

c) 根据式(6)获得新的描述像素特征的函数,以此代替图像单一像素灰度特征来表征图像。

d) 应用模糊 C-均值聚类算法对用式(6)代替的图像进行分割。具体详细过程如下:

用以上获得的新的特征向量来衡量像素点间的相似性。

定义目标函数为

$$J(U,V) = \sum_{i,j,k=1}^c (u_k(i,j))^m (G_k(i,j))^2 \quad (8)$$

其中: $U = [u_k(i,j)]$ 为模糊隶属度矩阵; $V = \{v_k(k = 1, 2, \dots, c)\}$ 为聚类中心集合; $u_k(i,j)$ 是 $I(i,j)$ 对于第 k 类的隶属度函数,且满足 $0 \leq u_k(i,j) \leq 1$ 和 $\sum_{k=1}^c u_k(i,j) = 1$; $m \in [1, \infty)$ 是一个控制聚类结果的权指数; $G_k(i,j)$ 为新的距离公式,由式(6)给出。

隶属度函数和聚类中心更新公式分别如式(9)(10)所示:

$$u_k(i,j) = 1 / \sum_{p=1, p \neq k}^c (G_k(i,j) / G_p(i,j))^{2/(m-1)} \quad (9)$$

$$v_k = \sum_{i,j} [(u_k(i,j))^m G_k(i,j)] / \sum_{i,j} (u_k(i,j))^m \quad (10)$$

基于以上改进的模糊聚类的图像分割就是将图像像素点当成具有质量的节点,节点的信息和距离构建的关系函数看做节点新的特征,通过迭代来确定对应的隶属度函数值和聚类中心,使目标函数 J 取最小,最终获得图像分割结果。

2 结果与分析

为验证本文算法的有效性和可行性,对三类典型的医学图

像进行仿真实验,并将本文算法的仿真结果与传统 FCM 算法^[6]、KFCM 算法^[13]、文献[8]提出的 VPBFCM 算法进行仿真比较。实验参数设置为:第一类和第二类图像聚类中心数目 c 取 3,第三类图像聚类中心数目取 2;权指数 m 取 2; ε 取 1×10^{-5} 。实验均利用 MATLAB 7.0.4 语言编程实现,实验环境为 Intel® Core™2 Quad CPU Q9400 @ 2.66 GHz,3.50 GB 内存的 PC 机。

2.1 细胞图像仿真分析

图 2 为 340×224 大小的宫颈细胞图像,由细胞核、细胞质及背景区域构成,其中细胞核是要提取的目标。由于有三类物体,所以实验中聚类中心数目 c 取 3。由图 2(a)可知,图像中目标区域和另外两类区域的对比度比较明显,但是由于细胞质被细胞核中的液体包围给提取细胞核目标带来一定困难。图 2(b)~(e)都是采用分三类进行聚类后再对图像二值化的分割结果。可以看出,应用本文算法能够精确地提取细胞核,这是由于采用本文改进的 FCM 算法不仅考虑了图像的灰度特征,同时结合了图像节点的空间位置信息,综合了以上特征对结果的影响。文献[8]算法能相对准确地、不受细胞质干扰提取出目标区域。传统 FCM 和 KFCM 算法最终都受到细胞质干扰而出现误分割。

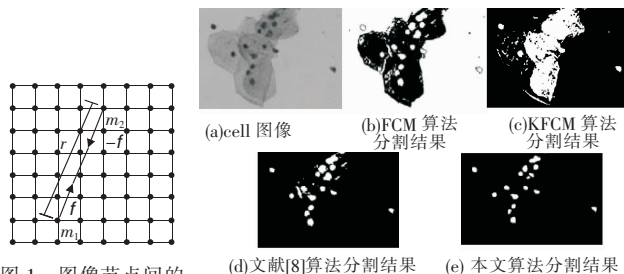


图 1 图像节点间的万有引力

图 2 各算法对 cell 图像分割结果

2.2 含噪头颅 CT 图像仿真分析

图 3(a)为 256×256 大小的含噪头颅 CT 图像,白色为头颅的白质区域,也即脑神经纤维部分,是需要分割的目标区域。该区域目标的组织结构较为复杂,为精确分割带来一定难度。应用传统的 FCM 和 KFCM 算法即使类别数设定为 3 也分割不出目标区域并且不能消除噪声的干扰。文献[8]将灰度特征融合了提取纹理信息的 Gabor 滤波器,一定程度上克服了噪声的干扰,能够提取出目标区域,但是轮廓线条较粗。而应用本文改进的 FCM 算法更加精确地提取出了灰质区域,有利于医生对影像进一步的精确判读和诊断。

2.3 血管图像仿真分析

为了进一步验证本文算法对更为复杂的医学图像同样适用,采用血管图像进行仿真实验。图 4(a)所示为 111×110 大小的血管图像,整幅图像灰度分布不规则,目标边界模糊不清,细小的血管目标区域与背景区域灰度较接近,即使通过人眼视觉感知也很难精确定位细小血管轮廓区域。从分割结果来看,采用 KFCM 算法的分割结果存在明显的过分割现象,将背景错分为目标;文献[8]提出的改进算法虽然能保留一些目标的细节,但是由于受到模糊边界的影响,同样将小部分背景错分为目标,对于主血管区域的提取区域扩大了;传统的 FCM 算法分割的结果和本文融合万有引力及局部熵的 FCM 算法分割的结果都是在主血管方向上得到了较好的目标区域,但是,传统 FCM 将细小血管当成背景而漏掉。采用本文算法分割效果较

理想,基本保留了图像中大多数的目标区域,最大限度地提取了图像的细节信息。

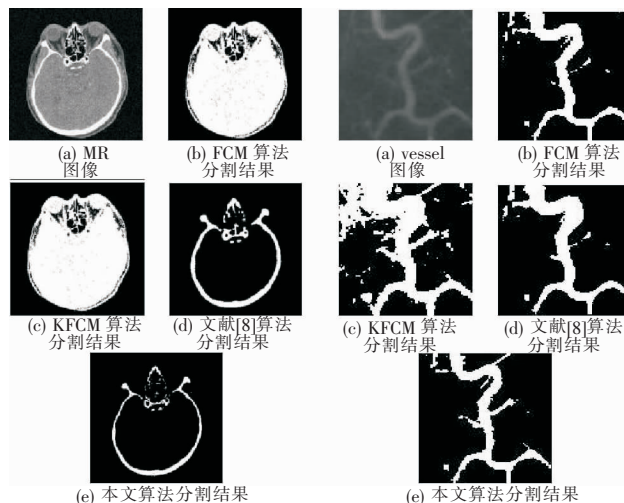


图 3 各算法对 MR 图像分割结果

图 4 各算法对 vessel 图像分割结果

3 结束语

本文将万有引力定律和图像局部熵引入到 FCM 算法中,综合考虑了图像灰度特征和像素空间分布对聚类结果的影响,克服了传统 FCM 算法在图像分割中仅仅应用节点灰度信息描述特征的缺陷,大大降低了噪声对分割结果的干扰。仿真结果表明,本文算法对灰度分布复杂的图像分割精度更高。

参考文献:

- [1] 林开颜,吴军辉,徐立鸿.彩色图像分割方法综述[J].中国图象图形学报,2005,10(1):906-910.
- [2] 王骏,王士同,邓超红.聚类分析研究中的若干问题[J].控制与决策,2012,27(3):321-328.
- [3] 李旭超,刘海宽,王飞,等.图像分割中的模糊聚类方法[J].中国图象图形学报,2012,17(4):447-458.
- [4] WANG Jian-zhong, KONG Jun, LU Ying-hua, et al. A modified FCM algorithm for MRI brain image segmentation using both local and non-local spatial constraints[J]. Computerized Medical Imaging and Graphics, 2008, 32(8):685-698.
- [5] 龚劬,姚玉敏.基于分水岭和改进的模糊聚类图像分割[J].计算机应用研究,2011,28(12):4773-4775.
- [6] BEZDEK J C, HATHAWAY R J. Convergence and theory for fuzzy C-means clustering: counterexamples and repairs[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1987, 17(5):873-877.
- [7] 王向阳,王春花.基于特征散度的自适应 FCM 图像分割算法[J].中国图象图形学报,2008,13(5):906-910.
- [8] 潘改,高立群,依玉峰.基于视觉感知的模糊 C-均值聚类算法[J].东北大学学报:自然科学版,2012,33(12):1690-1694.
- [9] SUN Gen-yun, LIU Qin-huo, LIU Qiang, et al. A novel approach for edge detection based on the theory of universal gravity[J]. Pattern Recognition, 2007, 40(10):2766-2775.
- [10] 郭丽,高立群,片兆宇.一种基于滑降和随机游走图像分割算法[J].计算机辅助设计与图形学学报,2009,21(8):1149-1154.
- [11] SHANNON C E. A mathematical theory of communication[J]. Bell System Technical Journal, 1948, 27:379-423.
- [12] YUE X D, MIAO D Q, ZHANG N, et al. Multiscale roughness measure for color image segmentation[J]. Information Sciences, 2012, 216(20):93-112.
- [13] GIZOLAMI M. Mercer kernel-based clustering in feature space[J]. IEEE Trans on Neural Networks, 2002, 13(3):780-784.