

# 几种优化 FCM 算法聚类中心的方法 对比及仿真

朱 然,李积英

(兰州交通大学 电子与信息工程学院,甘肃 兰州 730070)

**摘 要:**模糊 C-均值聚类(FCM)算法由于能够很好地解决像素分类的不确定性而得到广泛应用,但是聚类中心的初始化对其分割效果有很大的影响。文中以初始聚类中心为重点,分别用  $K$  均值算法、遗传算法、蚁群算法、粒子群优化算法优化 FCM 算法初始聚类中心,将优化后的结果作为 FCM 的初始聚类中心,并利用 MATLAB 软件进行了实验仿真。通过实验结果对比分析,不仅优化后的运算时间有所减少,而且所得的聚类中心更加稳定,使得分割出的目标也更加完整、清晰,验证了改进算法的有效性。

**关键词:**模糊 C-均值聚类;聚类中心;智能算法;参数优化

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2015)05-0017-04

doi:10.3969/j.issn.1673-629X.2015.05.005

## Contrast and Simulation of Several Clustering Centers of Optimized FCM Algorithms

ZHU Ran, LI Ji-ying

(School of Electronic & Information Engineering, Lanzhou Jiaotong University,  
Lanzhou 730070, China)

**Abstract:** Fuzzy C-Means clustering (FCM) algorithm was widely used because of effectively solving the uncertainty of pixel classification, but the initialization of clustering center had a great influence on its segmentation effect. So, focused on initial clustering center, respectively use  $K$ -means algorithm, genetic algorithm, ant colony algorithm, particle swarm optimization algorithm to optimize the FCM algorithm initial clustering center and take optimized result as its initial clustering center. MATLAB software was applied to carry out the experimental simulation. Through comparing and analyzing the results of the experiment, not only reduce the optimized operation time, but the clustering center is more stable, making segmentation effect is more complete and clear, which verifies the validity of the improved algorithm.

**Key words:** FCM; cluster centers; intelligent algorithms; parameter optimization

## 0 引言

模糊聚类是图像分割中有效的分割手段,得到了众多学者的关注和研究,其中 FCM 算法是应用和研究最多的算法<sup>[1-6]</sup>。文献[7]通过引入簇的数目来调整聚类中心的个数,即首先根据初始的聚类中心进行聚类,将聚类结果作为簇的个数,当两个簇的中心距离小于设定的阈值时,则合并两个簇,然后再随机生成聚类中心,重新进行聚类,这样就可以对新进入的数据进行分类,从而得到更精确的聚类中心。文献[8]利用粒子群优化算法的全局性对 FCM 算法的初始聚类中心

进行了优化,并成功地应用到卫星云图的聚类上。文献[9]通过分析数据样本,计算出样本的权重,将具有代表性的样本作为 FCM 算法的初始聚类中心,从而增强了算法的鲁棒性。

## 1 FCM 算法

FCM 算法是基于目标函数最小化的算法,其目标函数如下:

$$J_q = \sum_{i=1}^n \sum_{j=1}^m u_{ij}^q d(x_i, \theta_j) \quad (1)$$

收稿日期:2014-07-03

修回日期:2014-10-07

网络出版时间:2015-04-22

基金项目:甘肃省自然科学基金项目(1310RJZA057)

作者简介:朱 然(1989-),男,硕士研究生,研究方向为图像分割;李积英,副教授,研究方向为图像分割。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20150422.1005.014.html>

其中,  $u_{ij}$  是样本  $x_i$  对第  $j$  个聚类中心的隶属度;  $q$  是模糊因子;  $d$  是数据集  $x_i$  到第  $j$  个聚类中心的距离;

$$\sum_{j=1}^n u_{ij} = 1, 0 \leq \sum_{j=1}^n u_{ij} \leq n.$$

利用目标函数构造拉格朗日函数,并分别对聚类中心和隶属度求偏导,可得对应的表达式为:

$$\theta_j = \frac{\sum_{i=1}^N U_{ij}^q x_i}{\sum_{i=1}^N U_{ij}^q} \quad (2)$$

$$U_{ij} = \frac{1}{\sum_{k=1}^m \left( \frac{d(x_i, \theta_j)}{d(x_i, \theta_k)} \right)^{2/(l-1)}} \quad (3)$$

FCM 算法的步骤为:

Step1: 初始化参数, 包括加权指数  $q$ , 聚类个数  $C$ , 最大迭代次数  $T_{\max}$  以及允许的最小误差  $\varepsilon$ ;

Step2: 初始化隶属度矩阵  $U_0$  以及迭代次数  $l = 0$ ;

Step3: 按照公式(2)计算、更新聚类中心  $\theta_j$ ;

Step4: 按照公式(3)计算、更新隶属度矩阵  $U_{l+1}$ ;

Step5: 比较最近两次隶属度矩阵之差是否小于设定的阈值或者迭代次数达到最大值, 即:  $\max\{U(l) - U(l+1)\} < \varepsilon$  或者  $l > T_{\max}$ , 则结束, 否则,  $l = l + 1$ , 返回 Step3 继续迭代, 直到满足结束的条件为止。

## 2 几种改进初始聚类中心算法的介绍

### 2.1 K 均值算法

K 均值(K - Means)算法是 FCM 算法的基础, 与 FCM 算法不同的是, K 均值不涉及隶属度矩阵的问题, 对于待分类样本, 处理的结果只有“0”和“1”两种, 即属于这个类和不属于这个类<sup>[10-11]</sup>。K 均值算法的目标函数为:

$$J = \sum_{i=1}^K \sum_{x \in C_i} \|x - C_i\|^2 \quad (4)$$

其中,  $K$  是聚类个数;  $x$  是聚类样本;  $C_i$  是第  $i$  类的聚类中心。

K 均值算法流程如下:

Step1: 参数初始化, 聚类个数  $K$ , 聚类中心  $C$ 。

Step2: 设定  $l = 1$ , 如果  $\|x_j - C_p^{(l)}\| < \|x_j - C_i^{(l)}\|$ ,  $i, p = 1, 2, \dots, K, i \neq p$ , 那么  $x \in C_p^{(l)}$ , 即  $x$  是聚类中心为  $C_p^{(l)}$  的样本集。

Step3:  $C_i^{(l+1)} = \frac{1}{N} \sum_{x \in C_i^{(l)}} x, i = 1, 2, \dots, K$ 。

Step4: 如果  $C_i^{(l+1)} \neq C_i^{(l)}, i = 1, 2, \dots, K$ , 令  $l = l + 1$ , 返回 Step2 重新计算; 如果  $C_i^{(l+1)} = C_i^{(l)}, i = 1, 2, \dots, K$ , 那么算法收敛, 迭代结束。

### 2.2 遗传算法优化 FCM 初始聚类中心

遗传算法模仿自然界生物的遗传和进化现象, 将选择、交叉、变异机理作为算法的操作算子, 按照遗传和进化机制来完成对问题最优解的自适应搜索<sup>[12-13]</sup>。遗传算法优化 FCM 算法的流程为:

Step1: 参数初始化, 包括种群规模  $N$ , 交叉概率  $P_c$ , 变异概率  $P_m$ , 终止进化的迭代次数  $T$  以及初始聚类数目  $C$ 。

Step2: 设置迭代次数  $t = 0$ , 并随机选取初始种群。

Step3: 计算个体适应值, 其适应函数  $f = \frac{1}{J_q}$ 。

Step4: 进行选择、交叉、变异, 获得新一代的种群, 判断新种群是否满足停止迭代准则,  $t = t + 1$ 。

Step5: 如果  $t$  大于最大迭代次数, 则停止迭代, 输出聚类个数和聚类中心, 否则返回 Step2 继续迭代。

### 2.3 蚁群算法优化 FCM 初始聚类中心

蚁群算法模拟仿生学中蚂蚁觅食的过程, 将蚁群算法应用到图像分割中的原理就是将图像中的每一个像素视为蚂蚁, 将聚类中心视为“食物源”, 蚂蚁按照信息素更新机制选择路径, 寻找食物源, 找到食物源的同时也就完成了像素分类的问题, 从而完成图像的分割<sup>[14]</sup>。蚁群算法有着较好的全局性和鲁棒性。

蚁群算法优化 FCM 的流程为:

Step1: 随机选取  $M$  个点作为 FCM 的初始聚类中心, 初始化其他参数, 包括最大迭代次数、隶属度矩阵、允许的最大迭代误差。

Step2: 计算每个像素点到聚类中心的欧氏距离  $d_{ij}$ ,  $d_{ij} = \|P(X_i - X_j)\|^2$ 。

Step3: 计算信息素浓度  $\tau_{ij}(t)$ 。

$$\tau_{ij}(t) = \begin{cases} 1 & d_{ij} \leq r \\ 0 & d_{ij} \geq r \end{cases} \quad (5)$$

Step4: 计算转移概率  $P_{ij}(t)$ , 若  $P_{ij}(t)$  大于  $P_0$ , 说明像素  $X_i$  属于聚类中心为  $X_j$  的像素群, 需要将  $X_i$  合并到  $X_j$  所属的类中。

$$P_{ij}(t) = \frac{\tau_{ij}^\alpha(t) \eta_{ij}^\beta(t)}{\sum_{s=1}^j \tau_{sj}^\alpha(t) \eta_{sj}^\beta(t)} \quad (6)$$

Step5: 按公式(7)重新计算聚类中心。

$$O_j = \frac{1}{J_q} \sum_{k=1}^j X_k, X_k \in C_j \quad (7)$$

Step6: 计算总体误差  $\varepsilon = \sum_{j=1}^k D_j$ , 如果满足  $\varepsilon \leq \varepsilon_0$ , 则输出聚类数  $c$  和聚类中心  $v_j$ , 其中  $D_j$  为:

$$D_j = \sum_{k=1}^j \left( \left( \sum_{i=1}^2 x_{ki} - O_{ji} \right)^2 \right)^{\frac{1}{2}} \quad (8)$$

### 2.4 粒子群优化 FCM 初始聚类中心

粒子群算法与蚁群算法相似, 都是通过群体智能

来搜索最佳的结果。不同的是,蚁群算法是按照信息素来更新进化的方向,而粒子群算法则是根据“个体最优”和“全局最优”来进化<sup>[15-17]</sup>。在搜索过程中,按照公式(9)、(10)来更新自己的位置和速度。

$$v_{id}^{k+1} = v_{id}^k + c_1 r_1 (p_{id} - x_{id}^k) + c_2 r_2 (p_{gd} - x_{id}^k) \tag{9}$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \tag{10}$$

其中,  $k$  是迭代次数;  $c_1$ 、 $c_2$  是学习因子,  $r_1$ 、 $r_2$  是  $[0,1]$  内的随机数。

粒子群优化算法的流程为:

Step1: 设定合适的阈值、最大的迭代次数以及种群的初始化;

Step2: 计算样本数据对每个粒子的隶属度矩阵  $U$ 。

Step3: 按照式(11)计算每个粒子的适应值,此时会产生个体最优值和全局最优值。

$$f = \frac{1}{1 + J_q(U, V)} \tag{11}$$

表 1 几种算法所得聚类中心及收敛时间比较

算法	所得聚类中心	运行时间/s
FCM	(57.871 4 102.870 5 154.528)	78.3
K 均值优化算法	(61.432 1 119.038 3 159.981 9)	28.8
遗传优化算法	(76.656 3 119.500 0 141.857 1)	36.2
蚁群优化算法	(56.506 1 119.501 2 165.456 1)	69.9
粒子群优化算法	(57.874 4 112.953 2 154.528 5)	64.5

通过实验数据可以看出,几种改进的算法所得到的聚类中心比较接近,与标准 FCM 算法相比,不仅运算时间有所减少,而且所得的聚类中心更加稳定。

3.2 分割效果对比

图 1 中(a)为原始图像,(b)为 FCM 算法分割效果图。可以看出,FCM 算法能够分割出图像中的物体,但是细节信息并没有得到保留;



(a)

Step4: 比较每个粒子所经历过最好的个体最优值和全局最优值,若较好,则替代当前的个体最优值和全局最优值。

Step5: 按照速度公式和位置公式,产生新的粒子群。

Step6: 如果得到足够好的适应值或达到最大的迭代次数,则结束迭代,否则返回 Step3。

3 实验结果对比分析

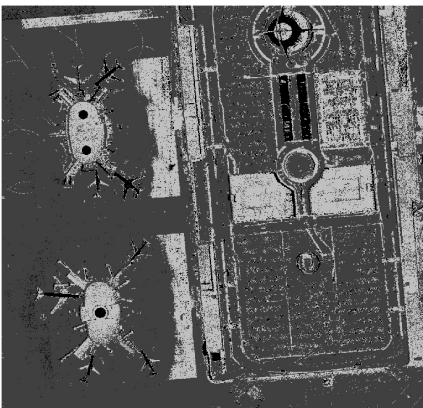
实验运行平台: 联想 Intel(R) Core(TM) i3-2120 CPU, 2 G 内存。Matlab7.0 环境下,选取一幅 1 204×1 204 的遥感图像为实验对象,聚类数目均为 3,各算法的运行参数分别参见文献[7,10-11,14]。

3.1 实验数据分析

算法所得聚类中心及收敛时间比较如表 1 所示。

图 2 中(a)为 K 均值优化算法分割效果,(b)为遗传优化算法分割效果,(c)为蚁群聚类优化算法分割效果,(d)为粒子群优化算法分割效果。

与 FCM 算法分割效果图相比较,可以看出几种改进算法所分割出的目标相较于标准的 FCM 算法所分割的目标更加完整、清晰。



(b)

图 1 原图与 FCM 分割效果图

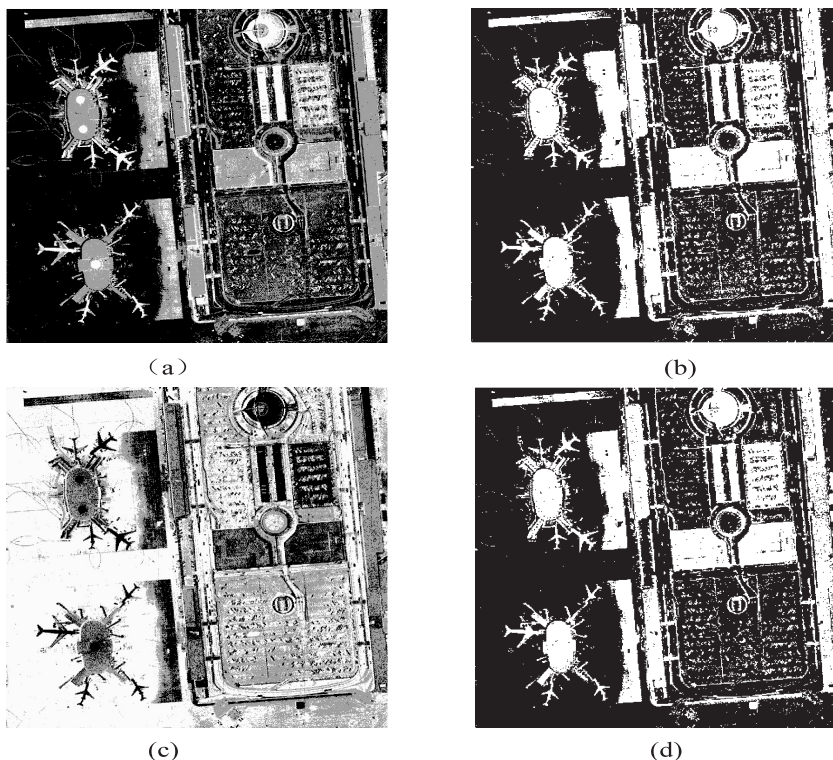


图 2 几种改进方法的分割效果图

## 4 结束语

初始参数的选取对 FCM 算法的分割效果有很大影响,利用其他智能算法得到初始聚类中心,在一定程度上起到了优化 FCM 算法的目的,但是如何降低优化算法的复杂度将是后续工作的重点。

### 参考文献:

- [1] Balafar M A, Rahmaamli A B D, Iqbalsaripan M, et al. Medical image segmentation using fuzzy C-mean (FCM) and user specified data[J]. Journal of Circuits, Systems, and Computers, 2010, 19(1): 6-10.
- [2] Zanaty E A. Determining the number of clusters for kernelized fuzzy C-means algorithms for automatic medical image segmentation[J]. Egyptian Informatics Journal, 2012, 13: 39-58.
- [3] Wang Qian, Zhang Qipeng, Zhou Wei. Study on remote sensing image segmentation based on ACA FCM[C]//Proc of 2010 2nd international conference on multimedia and computational intelligence. [s. l.]: [s. n.], 2012: 1286-1291.
- [4] 张 辉. 稳健模糊 C-均值聚类算法在图像分割中的应用研究[J]. 计算机工程与科学, 2010, 32(6): 45-47.
- [5] 张 静, 许高峰. 基于优化分块颜色直方图及模糊 C 聚类的彩色图像检索方法[J]. 计算机工程与科学, 2011, 33(8): 106-111.
- [6] 黄晓莉, 曾黄麟, 王秀碧. 基于模糊 C 均值聚类的粗集理论连续属性的离散化新算法[J]. 重庆邮电大学学报: 自然科学版, 2006, 18(5): 650-652.
- [7] 刘坤明, 罗 可. 改进的模糊 C 均值聚类算法[J]. 计算机工程与应用, 2009, 45(21): 97-98.
- [8] 李秀馨, 王敬东, 徐焯晔, 等. 基于改进 FCM 算法的卫星云图聚类方法研究[J]. 红外技术, 2013, 35(3): 150-154.
- [9] 江克勤, 施培蓓. 优化初始中心的模糊 C 均值(FCM)算法[J]. 合肥工业大学学报: 自然科学版, 2009, 32(5): 762-764.
- [10] 李 伟. 改进的基于模糊 C-均值聚类的图像分割算法[J]. 哈尔滨工业大学学报: 自然科学版, 2013, 29(4): 457-461.
- [11] 张明恒, 王华莹, 郭 烈. 基于改进 K-Means 算法的车辆识别方法[J]. 计算机技术与发展, 2012, 22(5): 53-56.
- [12] 王晓军, 熊 潇. 基于改进遗传算法的工作流调度研究[J]. 计算机技术与发展, 2013, 23(7): 108-111.
- [13] 徐月芳. 基于遗传模糊 C-均值聚类算法的图像分割[J]. 西北工业大学学报, 2002, 20(4): 549-553.
- [14] 杨立才, 赵莉娜, 吴晓晴. 基于蚁群算法的模糊 C 均值聚类医学图像分割[J]. 山东大学学报: 工学版, 2007, 37(3): 51-54.
- [15] Xue Zhaohui, Du Peijun, Su Hongjun. Harmonic analysis for hyperspectral image classification integrated with PSO optimized SVM[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 6(7): 2131-2146.
- [16] 孙越泓. 基于粒子群优化算法的图像分割研究[D]. 南京: 南京理工大学, 2010.
- [17] 陈晓霞, 廖家平, 赵熙临, 等. 基于禁忌粒子群优化的 FCM 聚类方法[J]. 湖北工业大学学报, 2013, 28(2): 45-48.



# 几种优化FCM算法聚类中心的方法对比及仿真

作者：[朱然](#)，[李积英](#)，[ZHU Ran](#)，[LI Ji-ying](#)  
作者单位：[兰州交通大学 电子与信息工程学院, 甘肃 兰州, 730070](#)  
刊名：[计算机技术与发展](#)  
英文刊名：[Computer Technology and Development](#)  
年，卷(期)：2015(5)

引用本文格式：[朱然](#). [李积英](#). [ZHU Ran](#). [LI Ji-ying](#) 几种优化FCM算法聚类中心的方法对比及仿真[期刊论文]-[计算机技术与发展](#) 2015(5)