

# 基于模糊 C 均值聚类的遥感图像分割方法

刘一超, 全吉成, 王宏伟, 孙林

(空军航空大学 特种专业系, 吉林 长春 130022)

**摘要:** 在遥感图像分割领域, 模糊 C 均值聚类算法得到了广泛的应用。但存在计算量大、易受噪声干扰等缺点。针对以上缺点对快速模糊 C 均值聚类算法进行了改进。首先利用一维灰度直方图进行快速模糊 C 均值聚类降低计算量; 然后在此基础上根据像素的邻域特性构造新的隶属度函数; 最后根据新的隶属度函数对每个像素进行分类。实验结果表明, 该算法能快速有效地分割图像, 并具有较强的抗噪能力。

**关键词:** 遥感; 模糊 C 均值聚类; 隶属度

中图分类号: TP751.1

文献标识码: A

文章编号: 1674-7720(2011)01-0034-04

## Remote sensing image segmentation based on fuzzy C-means clustering algorithm

Liu Yichao, Quan Jicheng, Wang Hongwei, Sun Lin

(Department of Specialty, Aviation University of Air Force, Changchun 130022, China)

**Abstract:** FCM clustering algorithm is widely applied to automated image segmentation. But standard FCM algorithm has many problems, such as great amount of calculation and easy noise interference. This paper proposes a modified fast FCM algorithm for image segmentation. With the modified algorithm, images can be mapped to gray-scale histogram space from pixel space, on the basis of which membership function can be improved by the full use of pixel's neighborhood feature. The experiment shows that new algorithm is effective in image segmentation and has good performance of resisting noise.

**Key words:** remote sensing; FCM; membership

在遥感图像处理领域, 图像分割是一个最具有挑战性同时也是最有活力的研究方向, 是进行遥感图像目标识别的前提。图像分割是将图像中所有的像素按照某个准则分成不重叠的、相似的区域。模糊 C 均值聚类算法 FCM(Fuzzy C Mean)在图像分割中是一种重要的方法。FCM 算法的一个重要属性是它引入了模糊性来判别图像中每一个像素的归属。与其他图像分割算法相比, FCM 能够更多地保留原始图像信息。而传统的 FCM 算法存在以下两个缺点: 一是没有考虑图像的空间信息, 使得它对图像中的噪声非常敏感; 二是当图像很大时, 算法的运行时间会很长。为了克服传统 FCM 的这两个缺点, 国内外学者提出了很多改进的算法。为了提高 FCM 的抗噪能力, 许多学者引入局部空间信息<sup>[1-5]</sup>, 使得每个像素的所属类别不仅与自身的灰度相关还受到相邻像素灰度的影响, 使改进算法虽然能够降低传统 FCM

对噪声的敏感, 但其运行时间远远高于传统的 FCM 方法。

针对以上缺点, 本文提出一种改进的模糊 C 均值聚类算法: 首先利用一维灰度直方图进行快速模糊 C 均值聚类降低计算量; 然后在此基础上根据像素的邻域特性构造新的隶属度函数; 最后根据新的隶属度函数对每个像素进行分类。

### 1 模糊 C 均值聚类

#### 1.1 传统 FCM 算法

设有  $n$  个样本, 记为  $X = \{x_i, i = 1, 2, \dots, n\}$ , 将其分成  $c$  类。设  $V_i = \{i = 1, 2, \dots, c\}$  为样本的聚类中心,  $u_{ik} = \{i = 1, 2, \dots, c; k = 1, 2, \dots, n\}$  是第  $k$  个样本对第  $i$  类的隶属度, 且满足  $0 \leq u_{ik} \leq 1$  及  $0 \leq \sum_{k=1}^n u_{ik} \leq n$ , 则 FCM 的目标函数为:

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

式中,  $U=\{u_{ik}\}$ ,  $V=(v_1, v_2, \dots, v_c)$ ;  $m>1$  为常数, 其约束条件为:

$$\sum_{i=1}^c u_{ik}=1 \quad \forall k=1, 2, \dots, n$$

FCM 的目标函数可以通过迭代更新隶属度矩阵  $U$  和  $V$  聚类中心得到最优解:

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left( \frac{U_{j-}V_i}{U_{j-}V_k} \right)^{\frac{2}{m-1}}} \quad i \leq c \quad j \leq n \quad (2)$$

$$V_i = \frac{\sum_{k=1}^n (u_{ik})^m x_k}{\sum_{k=1}^n (u_{ik})^m} \quad i \leq c \quad (3)$$

不断更新式(2)、式(3)直到聚类中心  $V_i$  的变化值小于  $\varepsilon(\varepsilon>0)$  时, 迭代停止。

### 1.2 基于空间约束的 FCM 算法(FCM\_S)

为了增强算法对噪声的抑制能力, 得到满意的分割效果, 很多研究人员在传统的模糊 C 均值聚类算法的基础上加入局部空间信息来提高算法的性能。参考文献[6] 通过改进传统模糊 C 均值聚类算法的目标函数, 提出了一种基于空间约束的模糊 C 均值聚类算法 FCM\_S (Fuzzy C-Means Clustering Algorithm with Spatial Constraints)。该算法不仅考虑了像素灰度信息, 还考虑了邻域像素的影响, 增强了算法对噪声的抑制能力。FCM\_S 的目标函数定义为:

$$J_m = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \|x_k - v_i\|^2 + \frac{\alpha}{N_R} \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \sum_{r \in N_k} \|x_r - v_i\|^2 \quad (4)$$

式中,  $x_k$  为第  $k$  个像素的灰度值,  $v_i$  为第  $i$  类的中心值,  $u_{ik}$  为第  $k$  个像素属于第  $i$  类的隶属度,  $N_k$  为像素  $x_k$  的邻域,  $x_r$  为像素  $x_k$  的邻域像素点,  $N_R$  为邻域所包含的点数, 参数  $\alpha$  控制着邻域像素对中心像素的影响权重。每一个像素  $x_k$  都满足约束条件  $\sum_{i=1}^c u_{ik}=1$ 。为了使目标函数达到局部最优值, 必须满足两个必要不充分条件:

$$u_{ik} = \frac{\left( \|x_k - v_i\|^2 + \frac{\alpha}{N_R} \sum_{r \in N_k} \|x_r - v_i\|^2 \right)^{-1/(m-1)}}{\sum_{j=1}^c \left( \|x_k - v_j\|^2 + \frac{\alpha}{N_R} \sum_{r \in N_k} \|x_r - v_j\|^2 \right)^{-1/(m-1)}} \quad (5)$$

$$v_i = \frac{\sum_{k=1}^N u_{ik}^m \left( x_k + \frac{\alpha}{N_R} \sum_{r \in N_k} x_r \right)}{(1+\alpha) \sum_{k=1}^N u_{ik}^m} \quad (6)$$

式(5)和(6)的缺点是在每一次的迭代过程中, 计算邻域像素都会耗费大量时间。为了减少计算量, 参考文献[2]提出一种改进的 FCM\_S 算法(称为 FCM\_S1), 简化了邻域像素的计算:

$$J_m = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \|x_k - v_i\|^2 + \alpha \sum_{i=1}^c \sum_{j=1}^N u_{ik}^m \|\bar{x}_k - v_i\|^2 \quad (7)$$

式中,  $\bar{x}_k$  是以  $x_k$  为中心的邻域中像素的均值, 它可以通过生成原始图像的均值滤波图像被提前计算出来。

式(4)中的  $\sum_{r \in N_k} \|x_r - v_i\|^2 / N_R$  被替换为  $\|\bar{x}_k - v_i\|^2$ , 明显减少了算法的执行时间。则该算法的两个必要不充分条件改写为:

$$u_{ik} = \frac{\left( \|x_k - v_i\|^2 + \alpha \|\bar{x}_k - v_i\|^2 \right)^{-1/(m-1)}}{\sum_{j=1}^c \left( \|x_k - v_j\|^2 + \alpha \|\bar{x}_k - v_j\|^2 \right)^{-1/(m-1)}} \quad (8)$$

$$v_i = \frac{\sum_{k=1}^N u_{ik}^m (x_k + \alpha \bar{x}_k)}{(1+\alpha) \sum_{k=1}^N u_{ik}^m} \quad (9)$$

FCM\_S1 不仅减少了算法的执行时间, 还提高了对高斯噪声的鲁棒性。然而该算法并不适用于存在椒盐噪声的图像, 因此参考文献[2]将算法中的均值滤波替换为 中值滤波即算法 FCM\_S2, 作为改进算法的补充。

### 1.3 改进的 FCM 算法(MFCM)

为了克服噪声对 FCM 算法的影响, 参考文献[7]引入空间函数  $h_{ij} = \sum_{x_j \in NB(x_i)} u_{ij}$ , 其中  $NB(x_j)$  是以像素  $x_j$  为中心、大小为  $3 \times 3$  的邻域,  $h_{ij}$  表示像素  $x_j$  属于第  $i$  类的概率。改进的隶属度函数为:

$$u_{ij} = u_{ij} h_{ij} / \sum_{k=1}^c u_{ij} h_{ij} \quad (10)$$

通过引入  $h_{ij}$ , 可以明显降低噪声对图像的影响, 但是该算法的运行时间比传统 FCM 算法加长了。

## 2 本文改进的快速 FCM 算法

传统 FCM 算法运行时间受到图像大小的制约, 为了提高 FCM 的运算效率, Yang 等人将图像一维直方图运用到算法中, 节约了 FCM 的运行时间。算法如下<sup>[8]</sup>:

定义  $G=\{L_{\min}, L_{\min+1}, \dots, L_{\max}\}$  为图像的灰度级,  $L_{\max}$  为最大的灰度级, 灰度尺度大小为  $L_{\max}-L_{\min}$ 。对于一幅大小为  $S \times T$  的图像,  $f(s, t)$  表示图像上  $(s, t)$  这一点的灰度值,  $His(g)$  表示图像中灰度级是  $g$  的像素个数,  $g \in G$ 。则图像的一维灰度统计直方图函数为:

$$His(g) = \sum_{s=0}^{S-1} \sum_{t=0}^{T-1} \delta(f(s, t)-g) \quad (11)$$

式中, 函数  $\delta(0)=1, \delta(g \neq 0)=0$ 。新的目标函数定义为:

$$J_m = \sum_{g=0}^{L_{\max}-1} \sum_{i=1}^c (u_{ig})^m His(g) d^2(g, v_i) \quad (12)$$

最小化目标函数, 从而得到快速 FCM(FFCM)算法的隶属度矩阵  $U$  和聚类中心  $V$ 。

$$u_{ig} = 1 / \sum_{j=1}^c (d_{ig}^2 / d_{ij}^2)^{\frac{2}{m-1}} \quad (13)$$

$$v_i = \left( \sum_{g=0}^{L_{max}-1} (u_{ig})^m \text{His}(g) \right) / \left( \sum_{g=0}^{L_{max}-1} (u_{ig})^m \text{His}(g) \right) \quad (14)$$

然而快速模糊 C 均值聚类算法同样会受到噪声的影响。为了克服这一问题,本文提出一种改进的快速模糊 C 均值聚类算法。其基本思想是:首先利用快速模糊 C 均值聚类方法得到图像的分类矩阵  $U$ ;然后充分利用像素的邻域特性,对隶属度函数进行改进,从而抑制图像中噪声对聚类的影响,重新生成新的分类矩阵;最后按照新的分类矩阵对图像进行分割。本文改进的隶属度函数如下:

$$n_{i^*j} = \begin{cases} \text{Max}(u_{ij}) & i^* = \text{max} \\ 0 & i^* \neq \text{max} \end{cases} \quad (15)$$

$$h_{ij} = \frac{\sum_{j \in \text{NB}(x_j)} n_{i^*j}}{\sum_{i=1}^c \sum_{j \in \text{NB}(x_j)} n_{i^*j}} \quad (16)$$

$$u_{ij}^* = \frac{u_{ij} \times h_{ij}}{\sum_{i=1}^c u_{ij} \times h_{ij}} \quad (17)$$

式中,  $\text{NB}(x_j)$  为像素  $x_j$  的  $3 \times 3$  邻域,  $i \in (1, c), i^* \in (1, c), j \in \text{NB}(x_j), \text{max}$  为具有最大隶属度的类的索引,  $\text{max} \in (1, c)$ 。

式(15)表示保留邻域中每个像素的最大隶属度。式(16)表示由邻域像素中的最大隶属度构成新的空间函数,邻域像素中属于同一类的数目越多,则中心像素  $x_j$  属于这一类的空间函数值越大。式(17)表示由新的空间函数和原始隶属度函数组合构成新的隶属度函数。在同质区域内,新的空间函数仅仅起到增强原始隶属函数的作用,聚类结果保持不变。而对于图像中的噪声,在聚类过程中可以起到抑制的作用。

### 3 试验结果与分析

为了验证本文算法的有效性,分别利用传统 FCM、FCM\_S1、FCM\_S2、MFCM 与本文算法进行比较和分析。实验中阈值  $\varepsilon=0.001$ , 加权指数  $m=2$ , 邻域采用  $3 \times 3$  的窗口,算法 FCM\_S1 和 FCM\_S2 中的参数  $\alpha$  指定为 2。

本文采用划分系数和划分熵两个聚类评价参数对算法进行评价。

划分系数定义如下<sup>[9]</sup>:

$$V_{pc} = \frac{1}{n} \sum_{i=1}^c \sum_{j=1}^n u_{ij}^2 \quad (18)$$

对于硬划分  $V_{pc}=1$ ,而对于模糊划分  $V_{pc}<1$ 。一般来说,希望模糊划分越明显越好,因此  $V_{pc}$  越大模糊聚类效果越好。

划分熵定义如下<sup>[10]</sup>:

$$V_{pe} = -\frac{1}{n} \sum_{i=1}^c \sum_{j=1}^n (u_{ij} \log u_{ij}) \quad (19)$$

式中,令  $\log(0)=0$ ,显然对于硬划分  $V_{pe}=0$ ,而对于模糊划分  $V_{pe}>0$ 。 $V_{pe}$  越小,模糊聚类效果越好。

图 1 是一幅  $255 \times 255$  的原图像,由油罐、阴影和背

景组成,因此聚类时类别数  $C=3$ 。

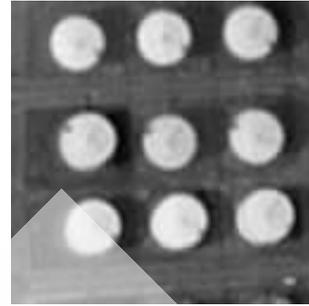


图 1 原图像

实验结果对各种聚类分割算法进行了分析比较,其结果分别如图 2、图 3、表 1 所示。

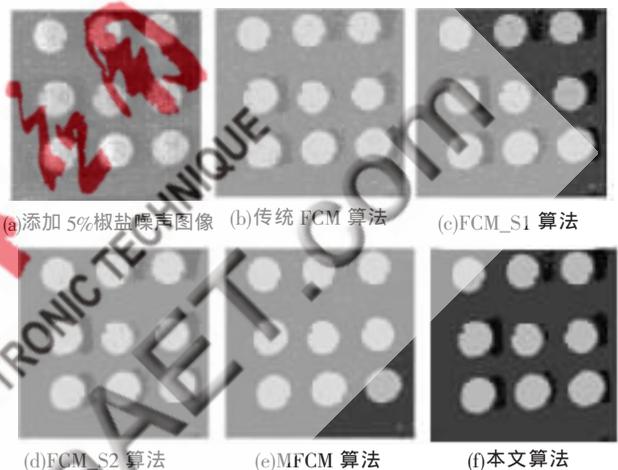


图 2 各种聚类分割算法比较

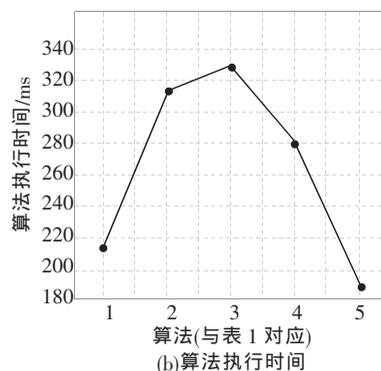
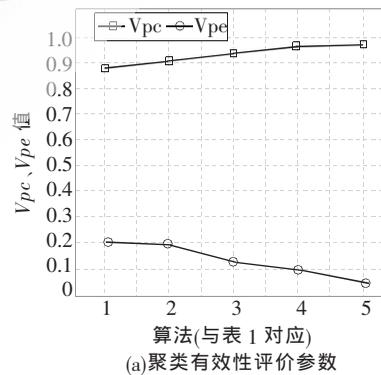


图 3 算法评价参数比

表 1 算法的有效性评价

聚类分割算法	$V_{pc}$	$V_{pc}$	执行时间/ms
FCM	0.8766	0.195 0	210.78
FCM_S1	0.9022	0.184 9	307.09
FCM_S2	0.9340	0.117 5	323.81
MFCM	0.9602	0.088 5	276.2
本文算法	0.975 1	0.042 0	188.42

本文提出了一种改进的快速 FCM 算法。首先利用一维灰度直方图进行快速模糊 C 均值聚类从而降低计算量,并在此基础上,充分考虑像素的邻域特性,对隶属度函数做一定改进。实验结果表明,本文算法分割速度明显高于传统的 FCM 算法和改进的 FCM 算法,并对噪声具有较强的鲁棒性。

## 参考文献

- [1] PHAM D L. Fuzzy clustering with spatial constraints [C]. in: IEEE Proceedings of the International Conference Image Processing, New York, 2002: 65–68.
- [2] CHEN S C, ZHANG D Q. Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure [J]. IEEE Transaction. Systems Man Cybernet, B 2004, 34(4): 1907–1916.
- [3] CAI W, CHEN S, ZHANG D Q. Fast and robust fuzzy C2means clustering algorithms incorporating local information for image segmentation [J]. Pattern Recognition, 2007, 40(3): 825–838.
- [4] SZILÁGYI L, SZILÁGYI S M, BENYÓ Z. A modified FCM algorithm for fast segmentation of brain MR images. Analysis and Design of Intelligent Systems Using Soft Computing

Techniques[M]. Heidelberg: Springer, 2007: 119–127.

- [5] 王志兵,鲁瑞华.改进的基于模糊 C 均值聚类的图像分割算法[J].西南大学学报(自然科学版)2009, 31(3): 169–172.
- [6] AHMED M N, YAMANY S M, MOHAMED N, et al. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data [J]. IEEE Transaction. Med. Imaging, 2002, 21: 193–199.
- [7] CHUANG K S, TZENG H L, CHEN S. Fuzzy c-means clustering with spatial information for image segmentation[J]. Computerized Medical Imaging and Graphics, 2006, 30: 9–16.
- [8] YANG Y, ZHENG C X, LIN P. Fuzzy clustering with spatial constraints for image thresholding [J]. Optica Applicata, 2005, 35: 309–315.
- [9] BEZDEK J C. Cluster validity with fuzzy sets [J]. Cybern, 1974, 3: 58–73.
- [10] BEZDEK J C. Mathematical models for systematic and taxonomy [C]. In: Proceedings of eight international conference on numerical taxonomy, San Francisco, 1975: 143–166.

(收稿日期: 2010–09–02)

## 作者简介:

刘一超,男,1985年生,硕士研究生,主要研究方向:数字图像处理。

全吉成,男,1960年生,教授,硕士研究生导师,主要研究方向:数字图像处理和资源管理。