

一种改进的空间加权模糊 C 均值算法在 遥感影像分割中的应用

张鹏林¹ 黄 丽¹ 吕志勇¹ 周 冰¹

(1 武汉大学遥感信息工程学院,武汉市珞喻路 129 号,430079)

摘 要:针对模糊 C 均值算法对初始聚类中心比较敏感的问题,提出了一种改进的空间加权模糊 C 均值算法。该算法首先通过 BP 神经网络训练已知样本,得出初始隶属矩阵,从而提高了初始聚类中心的可靠性。鉴于空间数据具有自相关的特点,在算法中考虑邻域像元对中心像元的影响,并赋予不同的权重,使算法具有更强的抗噪能力。为验证方法的有效性,将改进的方法用于 SPOT 2.5 m 遥感影像数据的分割,并与 FCM、SFCM 等算法的分割结果进行了对比分析,结果表明,本文方法取得了较好的效果。

关键词:模糊 C 均值;空间信息;图像分割

中图分类号:P237.3

图像分割是图像处理、分析中至关重要的一个环节,因此成为重要的研究热点之一^[1]。图像分割的本质是将同质像元分为同类,不同质像元分为异类。在过去的几十年,模糊 C 均值算法(fuzzy C-means, FCM)作为一种无监督聚类算法,已成功地应用在数据分类和图像分割等领域。FCM 算法通过引入隶属度来软化样本的归属,具有良好的柔韧性,用于挖掘更多的信息^[2]。

遥感影像作为空间数据的一种,像元间表现了较强的相关性,因此,在影像分割过程中,考虑空间自相关性是非常有必要的。然而,FCM 算法没有考虑空间相邻像元间的自相关特点,其对图像中的噪声较为敏感。Keh-Shih 等^[3]提出了加入空间函数的 SFCM 算法在图像分割中的应用,降低了噪声对聚类的影响,改善了图像的分割效果,但其降噪能力有限;Chen 等^[4]提出的算法中考虑了全图每一个像素点邻域内各点对该点的影响;Cai^[5]提出了将空间函数和灰度信息相结合的快速 FCM 算法,提高了精度和效果,但是文中引入的控制参数无法通过实验自动获取;Yu^[6]通过考虑自适应窗口大小对相邻像元的影响,提出了一种改进的自适应空间信息的 FCM 算法对彩色图像进行分割;刘小利等^[7]提出了并行 Tabu 搜索和空间信息约束的模糊聚类,通过构造广义隶

属矩阵实现分割。此外,在构建空间函数,考虑空间自相关关系的同时,Fan^[8]提出了单点逼近型加权 FCM 算法,通过对原始数据进行概率统计和加入样本属性权值来调整数据为均匀分布,并取得不错的效果;唐亮等^[9]提出考虑数据的空间分布特性的模糊 C-均值算法,考虑了每类数据的空间大小赋予不同的权重;Richard^[10]等人研究了 FCM 算法中各数据属性权重的影响,并通过调整数据的分布,以达到更高的聚类精度,但是对于具体的权值设定和调整过程仍没有较好的方法。

针对上述问题,本文提出了一种改进的空间加权模糊 C 均值算法(spatial-weighted fuzzy C-means algorithm, SWFCM)。

1 基于空间加权改进算法

1.1 FCM 算法

1973 年,Bezdek^[11]提出了模糊 C 均值算法,通过引入样本到隶属中心的隶属度,“软化”了样本的归属,使确定每个样本必须“硬”属于某个类变成计算每个样本属于各个类的概率。

在 FCM 算法中,将数据集 $\{x_k | k=1, 2, \dots, n\}$ 划分成 c 个模糊聚类是通过最小化式(1)所示

收稿日期:2013-04-17。

项目来源:国家十二科技支撑计划资助项目(2012BAJ15B04)。

的目标函数完成的。

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

其中, v_i 表示聚类中心; $x_i \in R^d$; $U = [u_{ik}]$ 是数据集的模糊划分矩阵,且满足约束条件 $\sum_{i=1}^c u_{ik} = 1, 1 \leq k \leq n, u_{ik} \in [0, 1]; m \in [1, \infty)$ 为一个加权指数,决定着所有分类的模糊程度。文献[12]认为, $m=2$ 时实验结果最准确。在目标函数和约束条件下,计算得模糊矩阵为:

$$u_{ik} = 1 / \sum_{j=1}^c \left(\frac{\|x_k - x_i\|}{\|x_k - x_j\|} \right)^{2/(m-1)}, \quad 1 \leq i \leq c, 1 \leq k \leq n \quad (2)$$

聚类中心为:

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

FCM 算法中根据式(2)和式(3)不断地迭代更新聚类中心和隶属矩阵,直至 $\|v_i^{(t)} - v_i^{(t-1)}\| \leq \epsilon$ (ϵ 为阈值),算法结束。

1.2 改进的空间加权 FCM(BP-SWFCM)

FCM 算法通过迭代更新聚类中心,以获取最优聚类结果,因此,聚类中心的初值会直接影响迭代次数和聚类结果。而标准 FCM 中,聚类中心是通过随机方式产生的,因此,其对迭代次数和结果的影响比较明显。为了得到可靠的聚类结果,本文提出采用 BP 神经网络训练样本得到隶属函数,然后将实验样本导入得到隶属矩阵,即每个样本对各类所做的贡献。另外,FCM 算法在聚类时,只考虑了在特征空间中像元间的影响,忽略了一些必要的空间信息,导致 FCM 算法陷入局部最优。为解决这一问题,算法中加入空间函数,并根据数据本身的特点赋予相应的权值,改善图像分割的效果。

加入权重的空间函数为:

$$h_{ij} = \sum_{sk \in NB(s,j)} \omega_{sk} u_{sk} \quad (4)$$

式中, ω_{sk} 表示邻域像元到中心像元的权重, $\omega_{sk} = d_{sk} / \sum_{k=1}^n d_{sk}$, d_{sk} 为样本 k 到类 i 的 Euclidean 距离。本文采用 3×3 的窗口进行实验,则模糊划分矩阵为:

$$u'_{ij} = u_{ij}^p h_{ij}^q / \sum_{i=1}^c u_{ij}^p h_{ij}^q \quad (5)$$

式中, p 和 q 为控制 u_{ij} 和 h_{ij} 相对重要性的参数。当 $p=1, q=0$ 时,为模糊 C 均值算法的情形。

1.3 算法在遥感影像分割中的应用过程

根据改进的 FCM 算法,利用 BP 神经网络产生初始隶属矩阵 $U^{(0)} = [u_{ik}]$,将其代入到空间加权 FCM 算法的整体循环迭代过程中,算法具体

步骤如下:

1) 初始化参数。设定初始参数类别数目 C , 模糊指数 $m=2$,最大循环次数为 500,截止误差 $\epsilon = 0.000\ 001$;

2) 确定初始聚类中心和权值矩阵。从遥感图像中读入每个像素点的 RGB 值,存入样本集合 $X = \{x_1, x_2, \dots, x_n\}$ 中,利用已知样本用 BP 进行训练,得到模型;然后将所有样本导入,得到初始隶属矩阵 $U^{(0)} = [u_{ik}]$;

3) 根据式(3)计算聚类中心;

4) 按照式(5)更新模糊划分矩阵;

5) 重复步骤 3)、4),直到 $\|v_i^{(t)} - v_i^{(t-1)}\| \leq \epsilon$ 结束,否则继续。

2 算法在影像分割中的实验

为验证本文基于改进的空间加权 FCM 算法(BP-SWFCM)的聚类效果,实验采用了某区的 SPOT 遥感影像,截取了其中具有代表性的 256×256 大小的图像进行实验,并与 FCM、SFCM 算法的聚类效果进行对比分析。

图 1(a)为原始影像,图 1(b)、1(c)、1(d)分别是用 BP-SWFCM 算法中针对 p 和 q 给定不同值得出的结果;表 1 为设置不同的 p, q 值进行实验所需的时间和迭代次数。实验结果显示,在对空间函数调控时, p, q 的设置对图像分割的影响不是很大,包括分割结束所需的时间和迭代的次数。从图 1(b)、1(c)中可以看出, p 值相同时, q 值越大,图像越平滑;从图 1(b)、1(d)可知,参数 p 越大,可能导致模糊掉图像的一些细节信息。综上所述, p 和 q 值的设定很难控制,正如文献[3]所述,相比之下, $p=1, q=1$ 时,实验结果较好。

表 1 不同 p, q 值的实验结果

Tab. 1 Results of Different p, q Values

原始影像	时间 t/s	迭代次数/次
$p=1, q=0$	34	42
$p=1, q=1$	34	42
$p=0, q=1$	35	42

图 2(a)至图 2(e)分别为 ERDAS 图像分割、FCM 算法、SFCM 算法、BP-SWFCM 算法、NSKFCM 算法[12]效果图。相比之下,用 ERDAS 分割的效果图较差,图像中的噪声较多,分割效果不理想。从图 2(b)、2(c)中可知,加入空间信息的 FCM 算法确实能够减弱噪声对其的影响。在考虑加入空间信息时构造空间函数的问题,学术界至今没有提出一种通用的函数[8-10]。

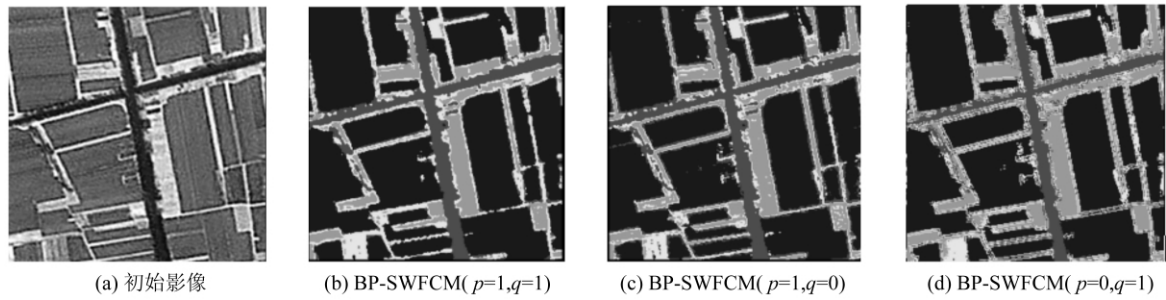
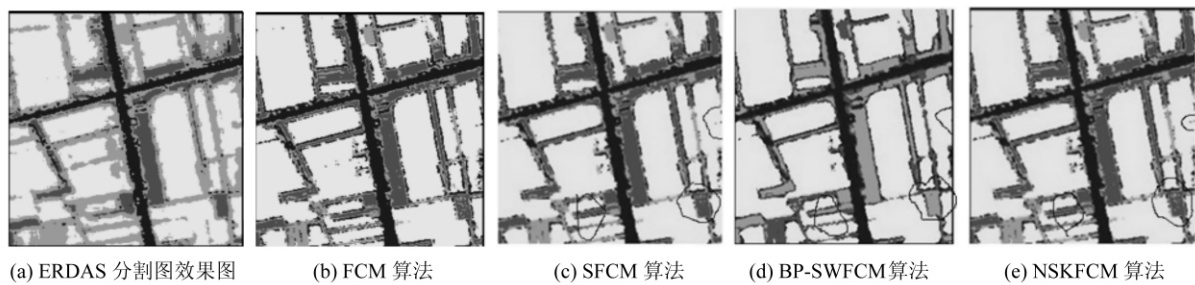
图1 不同 p, q 值的实验结果Fig. 1 Results of Image Segmentation with Different p, q Values

图2 不同算法用于图像分割的比较

Fig. 2 Comparison of Segmentation Results on SPOT 2.5 m Image

本文引用文献[3, 12]中两种不同的定权算法,图2(c)中SFCM^[3]算法是将所有的模板都赋予等值权后的聚类结果;图2(e)中,NSKFCM^[12]算法是基于核函数的空间FCM算法,在核函数的基础上加入空间信息,算法中根据邻域对中心像元的影响引入空间定权因子 α ;本文构造的空间函数方法如图2(d)所示。比较图2(c)~2(e)可以明显看出,本文方法得到的效果更平滑,很多孤立的点都被去除,分割结果较好。同时也说明就空间函数定权的问题,采用不同的定权方式得到的效果不同。

为充分证明本文方法的抗噪能力,文中采用人工合成图,利用Matlab加入椒盐噪声,如图3(a)所示。图3(c)与图3(b)相比较可知,BP-SWFCM方法有了不少的改善。表2是两种方法的抗噪能力评价。

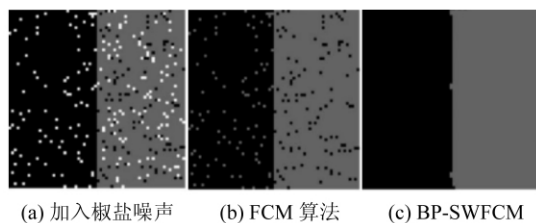


图3 加入椒盐噪声后的图像分割结果

Fig. 3 Segmentation Results After Add Salt and Pepper Noise

表2 抗噪能力评价

Tab. 2 Anti-noise Capability Evaluation

	FCM	BP-SWFCM
误分个数/个	2 340	203
正确个数/个	63 196	65 332
正确比/%	95.43	99.69

3 结 语

本文采用BP神经网络算法训练函数模型,得到了可靠的初始聚类中心,克服了FCM算法中对初始聚类中心敏感的问题。同时,选用BP算法训练还结合考虑了遥感数据量大的特点;在引入空间函数定权算法时,通过实验对比分析可知,本文方法的抗噪能力更强。另外,本文只是对分割过程中引入的空间特征进行了初步探讨,今后将在聚类中心的可靠性和空间数据的分布特点两方面进行研究。

参 考 文 献

- [1] Bezdek J C, Hall L O, Clarke L P. Review of MR Image Segmentation Techniques Using Pattern Recognition[J]. Med Phys, 1993,20(4): 1 033-1 048
- [2] Bezdek J C. Pattern Recognition with Fuzzy Objective Function Algorithms[J]. New York: Plenum Press, 1981:43-93
- [3] Chuang K S, Tzeng H L, Chen S R, et al. Fuzzy C-

- Means Clustering with Spatial Information for Image Segmentation[J]. *Computerized Medical Imaging and Graphics*, 2006, 30:9-15
- [4] Chen S C, Zhang D Q. Robust Image Segmentation Using FCM with Spatial Constraints Based on New Kernel-induced Distance Measure[J]. *IEEE Transactions on Systems Man Cybernet*, 2004, 34(4): 1 907-1 916
- [5] Cai W L, Chen S C, Zhang D Q. Fast and Robust Fuzzy C-Means Clustering Algorithms Incorporating Local Information for Image Segmentation[J]. *Pattern Recognition*, 2007, 40(3): 825-838
- [6] Yu Zhiding, Zou Ruobing, Yu Simin. A Modified Fuzzy C-Means Algorithm with Adaptive Spatial Information for Color Image Segmentation[C]. *IEEE Symposium on Computational Intelligence for Image Processing*, Nashville, TN, 2009
- [7] 刘小利,朱国宾,李清泉,等. 基于并行 Tabu 搜索和空间信息约束的遥感影像模糊聚类[J]. *武汉大学学报·信息科学版*, 2009, 34(5): 527-530
- [8] Fan Jianchao, Han Min, Wang Jun. Single Point Iterative Weighted Fuzzy C-Means Clustering Algorithm for Remote Sensing Image Segmentation[J]. *Pattern Recognition*, 2009, 42:2 527-2 540
- [9] 唐亮,黄培之,谢维信. 顾及数据空间分布特性的模糊 C-均值聚类算法研究[J]. *武汉大学学报·信息科学版*, 2003, 28(4): 476-479
- [10] Nock R, Nielsen F. On Weighting Clustering[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(8): 1 223-1 235
- [11] Bezdek J C. *Fuzzy Mathematics in Pattern Classification*[J]. Ithaca: Cornell University, 1973
- [12] Yang Yue, Guo Shuxu, Tian Runlan, et al. A Novel Kernel-based Fuzzy C-Means Algorithm with Spatial Information for Image Segmentation[J]. *Energy Procedia*, 2011, 13: 485-491

第一作者简介:张鹏林,副教授,博士,博士生导师。主要从事遥感图像处理、模式识别、特征提取等方面的研究。

E-mail: zpl@whu.edu.cn

An Improved Weighted Fuzzy C-Means Algorithm with Spatial Information for Remote Sensing Image Segmentation

ZHANG Penglin¹ HUANG Li¹ LV Zhiyong¹ ZHOU Bing¹

(¹ School of Remote Sensing and Information Engineering, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

Abstract: In this paper, we propose an improved spatial-weighted fuzzy C-means algorithm since the traditional fuzzy C-means algorithm is more sensitive to the initial cluster centers than other commonly deployed algorithms, such as the FCM, SFCM algorithms. BP neural network algorithms are used to train samples and obtain an initial membership matrix, thus increasing the reliability of the initial cluster centers. Since spatial data have patterns of spatial auto-correlation, the neighboring pixels will contribute to the center pixel with different weights to robustly handle noises. Segmentation experiments were conducted using SPOT 2.5 meters remote sensing images to verify the effectiveness of our algorithm. In comparison with FCM, SFCM algorithms, the experimental results show that the proposed method obtains better results.

Key words: fuzzy C-means; spatial information; image segmentation

About the first author: ZHANG Penglin, associate professor, Ph.D., Ph.D supervisor. His major researches are remote sensing and medical image processing, pattern recognition, image feature extraction.

E-mail: zpl@whu.edu.cn