

改进的自适应模糊 C 均值聚类图像分割算法

朱红霞, 王黎明

(中北大学 电子测试技术国家重点实验室, 山西 太原 030051)

摘 要: 分析自适应 FCM 图像分割算法的优缺点, 提出了改进的自适应模糊 C 均值聚类算法. 利用像素的空间邻域信息构造二维直方图, 作为自适应 FCM 的聚类样本, 降低了样本空间的维数, 解决了自适应 FCM 收敛速度慢, 对噪声敏感等问题. 通过与自适应 FCM 算法对含噪图像的分割结果以及分割速度, 性能的对比分析, 证明该算法收敛速度快, 分割精度高, 对噪声有较强的鲁棒性.

关键词: 二维直方图; FCM 算法; 自适应; 图像分割

中图分类号: TP391.41

文献标识码: A

文章编号: 1000-7180(2010)12-0087-03

An Improved Method for Image Segmentation with Auto-adaption Fuzzy C-Means Clustering

ZHU Hong-xia, WANG Li-ming

(National Key Laboratory for Electronic Measurement Technology, North University of China, Taiyuan 030051, China)

Abstract: Based on the analysis of advantages and disadvantages of the self-adaptive FCM image segmentation algorithm, an improved self-adaptive fuzzy C-means clustering algorithm is proposed. By construct a two-dimension histogram with pixel neighborhood space information, as a cluster sample for self-adaptive FCM, the dimension of sample space is reduced, and then, the problem of self-adaptive FCM on low rate of convergence and sensitive to noise is solved. At last, the improved algorithm is improved to have higher rate of convergence, higher accuracy of separation, and more robustness to noise, by contrast to self-adaptive FCM algorithm on image segmentation to a noise contained image.

Key words: two-dimensional histogram; fuzzy c-means clustering algorithm; auto-adaptation; image segmentation

1 引言

图像分割是指把图像分成各具特性的区域并提出感兴趣的目标. 至今已提出了多种图像分割方法, 但并没有一种适合所有图像的通用分割算法. 近年来, 比传统的硬方法保留更多的原始图像信息的模糊分割算法引起人们的广泛关注, 尤其是模糊 C 均值算法作为一种无需训练样本、无监督聚类算法已成功应用在图像分析、医疗诊断、目标识别和图像分割等领域.

传统的 FCM 算法^[1]分割结果受初始聚类中心和聚类数目的影响比较大, 自适应 FCM^[2]有效地解决了这一问题, 但其收敛速度慢, 且未能考虑各个像

素点的邻域信息, 抗噪性能比较差. 然而图像在成像过程中不可避免地受到不同噪声的干扰, 采用自适应 FCM 算法并不能得到满意的分割结果.

文中提出一种改进的自适应模糊 C 均值图像分割算法. 该方法利用二维直方图描述像素的邻域关系, 将二维直方图与自适应模糊 C 均值结合来实现图像的分割, 有效提高了聚类速度和分割精度, 实验结果表明该算法的有效性.

2 传统的 FCM 算法

FCM 聚类分割实现方法是根据图像像素和聚类中心的加权相似性测度, 对目标函数进行迭代优化以确定最佳聚类. 根据文献[3-4]中 FCM 算法, 定

义下面的目标函数式(1),用拉格朗日乘数法求解出使其极小的两个迭代公式(2)和(3)进行数据聚类。

$$J_m(U, V) = \sum_{i=1}^c \sum_{k=1}^n (u_{ik})^m d_{ik}^2 \quad (1)$$

$$\mu_{ik} = \frac{1}{\sum_{j=1}^c (d_{ik}/d_{jk})^{2/(m-1)}} \quad (2)$$

$$v_i = \frac{\sum_{k=1}^n (u_{ik})^m x_k}{\sum_{k=1}^n (u_{ik})^m} \quad (3)$$

像素集 $X = \{x_1, x_2, \dots, x_n\} \in R^n$ 分为 c 类, $m \in [1, \infty)$ 是一个加权指数,常取值为 2, d_{ik}^2 是第 k 个样本到第 i 类的欧氏距离, u_{ik} 为 X 中任意样本 x_k 对 i 类的隶属度。

3 改进的自适应模糊 C 均值聚类(FCM)算法

3.1 自适应 FCM 算法

由于传统的 FCM 算法初始聚类中心是随机取值,且要求聚类数目 C 的先验知识^[4],而关于数据空间分布及结构的先验知识是很少的,这一要求限制了 FCM 算法的实际应用。自适应 FCM 算法针对 FCM 算法上存在的不足,在算法结构上进行了具体的改进,增加了聚类有效性函数的分析,解决了 FCM 算法中初值选择问题,动态确定聚类中心的个数,可以克服主观性^[4]。

自适应 FCM 算法步骤如下:

(1) 聚类中心的初值选取:每一个数据点作为一个潜在的聚类中心点;并令迭代次数 $L = 0$,模糊指数 $m = 2$ 。

(2) 对每个样本点 x_i 定义测度^[5]:

$$P_i = \sum_{j=1}^n e^{-\alpha \|x_i - x_j\|^2}$$

P_i 为 x_i 到其他任一样本点的距离函数。计算各 P_i 值,选择对应 P_i 值最高的点 x_i 作为第一个聚类中心 x_1^* ;从形成下一聚类中心的候选样本点中剔除 x_1^* 点。

(3) 计算决定下一个聚类中心点, $P_i = P_i - P_1^* e^{-\beta \|x_i - x_1^*\|^2}$, 其中, $\beta = \frac{4}{r_b^2}$, $r_b = 1.5r_a$, 距 x_1^* 越近的样本点,其相应的 P_i 值就越小,成为下一个聚类中心的可能性就越小。同样选择对应 P_i 值最高的点 x_i 作为聚类中心,并从形成一聚类中心的候选数据中剔除该点。

(4) 按照(3)的方法计算其它的聚类中心点,将经过 k 次计算后, $P_i = P_i - P_k^* e^{-\beta \|x_i - x_k^*\|^2}$, 直到 $\|P_k^* - P_1^*\| < \epsilon$ 为止。根据上述步骤,可计算出初始聚类中心和聚类数目 C 。

(5) 按下式计算出 U^L :

当 $I_j = \Phi$ 时,

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}^2}{d_{kj}^2} \right)^{1/(m-1)}} \quad (4)$$

当 $I_j \neq \Phi$ 时,

$$\forall i \in I_j, u_{ij} = 0, \text{ 且 } \sum_{i \in I_j} u_{ij} = 1 \quad (5)$$

(6) 修正隶属度矩阵 U^L :

对元素 x_j , 设它对第 P 类的隶属度最大, 值为 u_{pj} ; 对第 s 类的隶属度次最大, 值为 u_{sj} 。修正后, 样本 x_j 对第 P 类的隶属度为: $u_{pj} = u_{pj} + (1 - \alpha)u_{sj}$, 对第 s 类的隶属度为: $u_{sj} = \alpha u_{sj}$ ($\alpha \in [0, 1]$ 为置吸引力抑制因子), u_{pj} 对其余各类的隶属度不变。

(7) 用修正后的隶属度矩阵计算下一次迭代中

$$\text{心 } V^{L+1}: v_i^{(L+1)} = \frac{\sum_{j=1}^N (u_{ij}^{(L)})^m x_j}{\sum_{j=1}^N (u_{ij}^{(L)})^m}$$

(8) 判断是否满足循环终止条件: 达到限定的

最大循环次数或给定的阈值 ϵ , $\sum_{i=1}^c \sum_{j=1}^n \|u_{ij}^{L+1} - u_{ij}^L\| < \epsilon$ 。如果不满足终止条件, 根据(5)中的公式继续迭代。参数更新公式为: $L = L + 1$ 。

3.2 改进的自适应 FCM 算法

由于图像样本数据量很大(一幅 256×256 的图像, 样本数据量为 65536), 故自适应 FCM 算法收敛速度很慢, 且自适应 FCM 算法的主要目的在于将向量的样本空间点按照某种距离度量划分成 c 个子空间, 但在图像分割中未能考虑各个点的灰度特征及其邻域像素的关联程度, 即使在同一个聚类中, 像素点和噪声点的灰度相差很大, 但两点的隶属度可能相同, 因此该算法的抗噪性能很差^[6]。

文中通过构造合理的二维直方图作为样本空间, 不仅考虑了像素点的信息, 而且还用到了邻域点的相关信息, 可以有效地抑制噪声。且对于任何一幅图像, 其二维直方图的样本数据量均为 256×256 , 对于数据量大的图像, 缩小了样本空间, 从而减少了聚类时间。

3.3 二维直方图

把图像看成一个二维灰度函数 $F(x, y)$, 包含 M

$\times N$ 个像素点,灰度取值范围为 $0 \sim L-1$ (灰度图像的级数,常取值为256). $G(x,y)$ 是 $F(x,y)$ 的平滑图像,显然平滑图像的灰度取值范围仍为 $0 \sim L-1$.对于一幅图像,用 P_{ij} 表示相应的“灰度-邻域均值对”出现的频数,这里: $0 \leq P_{ij} \leq M \times N, \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P_{ij} = M \times N$,那么 $\{P_{ij}, i, j = 1, 2, \dots, L\}$ 就是该图像的关于点灰度-区域灰度均值的二维直方图^[5].图1(c)和图1(d)分别是图1(b)的一维和二维直方图.可以看出,在强噪声干扰下,一维直方图是单峰的,物体和背景的分布区很不明显,但二维直方图由于利用了图像邻域的相关信息,物体和背景的双峰分布仍明显地得到保留.

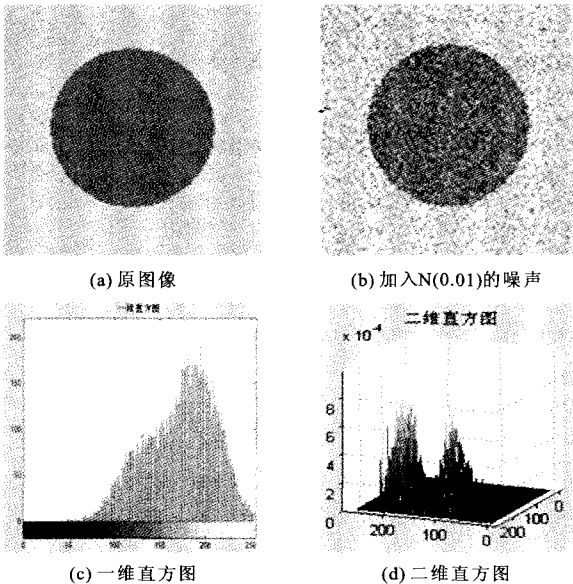


图1 图像及其直方图

4 实验结果及分析

为验证算法的有效性,用仿真图像和自然图像做实验,分别比较了自适应FCM算法^[6]与文中算法的图像分割效率.测试环境为:CPU 2.80GHz,512MB内存,Matlab7.0仿真工具实现.

4.1 分割时间的比较:

表1所示是两种算法对不同尺寸图像进行分割时的运行时间比较.可见新算法利用二维直方图作为样本空间,图像分割速度明显提高,尤其是图像尺寸越大,这种优势越明显.

4.2 分割性能的比较:

图2(a)所示为标准cameraman测试图像,图像尺寸为 256×256 ,灰度级为256,图2(b)为加上的噪声后的图像,这里分别用自适应FCM算法和文

中算法进行分割实验,分割结果如图2(c)和图2(d)所示.显然,由于文中算法考虑了空间邻域信息,有效抑制了噪声的干扰,分割结果较前者更为理想.

表1 不同尺寸图像分割中两种算法的分割时间及迭代次数对比

图像大小	自适应FCM算法 时间/秒	新算法 时间/秒	自适应FCM算法 迭代次数	新算法 迭代次数
128×128	3.3661	11.2250	37	41
256×256	5.8757	9.4354	41	35
512×512	18.3873	9.7868	45	36
1024×1024	71.2970	13.4563	48	46

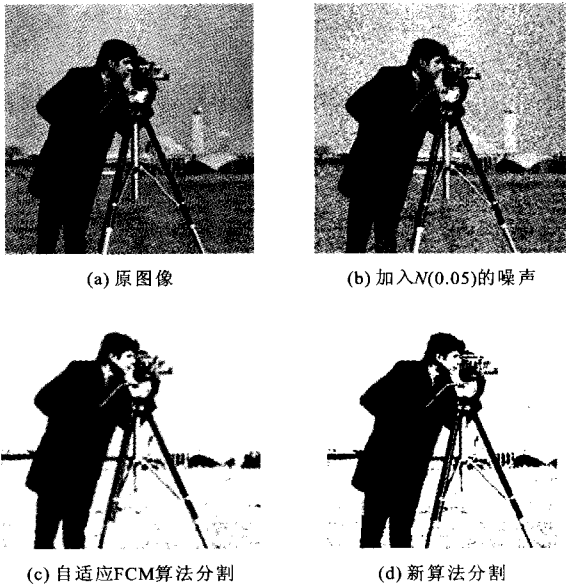


图2 两种算法对自然图像的分割

图1(a)是一个 128×128 大小的仿真图像,它由物体(圆)和背景两部分组成,其中背景灰度为192,物体灰度为128.图1(b)是图1(a)叠加了高斯噪声后的图像.图3(a)是采用自适应FCM算法分割的结果,图3(b)是采用改进的自适应FCM算法分割的结果.

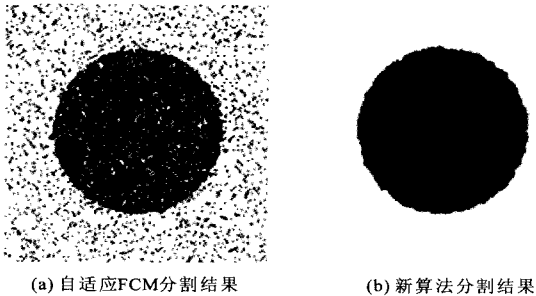


图3 两种算法对含噪图像的分割结果

CSMC 0.35 μm 标准 CMOS 混合工艺模型,仿真结果表明,实现的带隙基准功耗低,温度系数小,具有良好的 PSRR 特性,满足实际模拟集成电路产品的设计指标要求。

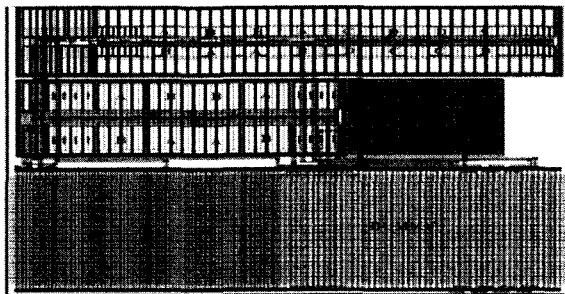


图6 设计的带隙基准的版图

参考文献:

- [1] Widlar R. New developments in IC voltage regulators [J]. IEEE Journal of Solid—State Circuits, 1971, 6(1): 2—7.
- [2] 李凡阳, 刘文平. 一种带温度保护和曲率补偿的带隙电压基准结构[J]. 微电子学与计算机, 2008, 25(3): 74—76.
- [3] Inyeol L, Gyudong K, Wonchan K. Exponential curvature compensated BiCMOS bandgap reference[J]. IEEE Journal

of Solid—State Circuits, 1994, 29(11): 1396—1403.

- [4] 盛庆华, 张亚君, 王红义. 一种线性补偿的带隙基准电路[J]. 微电子学与计算机, 2007, 24(1): 169—172.
- [5] Leung K N, Mok P K T, Leung C Y. A 2V 23uA 5.3ppm/ $^{\circ}\text{C}$ curvature compensated CMOS bandgap reference[J]. IEEE Journal of Solid—State Circuits, 2003, 38(31): 561—564.
- [6] Razavi B. Design of analog CMOS integrated circuits [M]. New York: McGraw Hill, 2001: 377—404.
- [7] Palumbo G. Voltage references: from diodes to precision high—order bandgap circuits [M]. New York: IEEE Press, 2002: 29—33.
- [8] 刘珂, 杨海钢, 尹韬. 采用电阻补偿的高 PSRR 基准电压源设计[J]. 微电子学, 2008, 38(2): 192—196.

作者简介:

姚舜 男, (1984—), 硕士研究生. 研究方向为模拟集成电路设计.

顾晓峰 男, (1971—), 博士, 教授, 博士生导师. 研究方向为半导体器件物理、集成电路设计.

梁海莲 女, (1979—), 讲师. 研究方向为集成电路设计.

冯向光 男, (1969—), 博士, 教授, 高级工程师. 研究方向为集成电路设计.

(上接第 89 页)

表 2 给出了两种算法分割性能及聚类中心的对比, 显然新算法的分类正确率高于自适应 FCM 算法.

表 2 含噪图像的两种算法分割性能的比较

分割算法	聚类数目	聚类中心	错分点数	错误率/%
自适应 FCM 算法	3	(0.4616, 0.6720, 0.8335)	2193	13.39
新算法	3	(0.5115, 0.6463, 0.7830)	223	1.36

5 结束语

文中充分利用了图像灰度的邻域空间信息, 并将 FCM 算法的聚类数目及初始化聚类中心进行自适应处理, 使含噪图像的分割精度^[7]明显提高, 对于像素样本空间较大的图像降低了图像的分类时间, 实现了快速精确分割.

参考文献:

- [1] Zhang D Q, Chen S C, Pan Z S, et al. Kernel—based fuzzy clustering incorporating spatial constraints for image segmentation[C]// Proc. International Conference

on Machine Learning and Cybernetics. Washington, 2003(4): 2189—2192.

- [2] Feng Yanqiu, Chen Wufan, Liang Bin, et al. A new algorithm for image segmentation based on gibbs random field and fuzzy c—means clustering[J]. Acta Electronica Sinica, 2004, 32(4): 645—647.
- [3] 焦远锋, 李万龙, 郑山红. 一种新的两阶段 FCM 聚类算法[J]. 微电子学与计算机, 2009, 26(4): 40—42.
- [4] 王洪春, 彭宏. 基于模糊 C—均值的增量式聚类算法[J]. 微电子学与计算机, 2007, 24(6): 156—161.
- [5] Chen Songcan, Zhang Daoqiang. Robust image segmentation using FCM with spatial constraints based on new kernel—induced distance measure[J]. IEEE Transactions on systems, man, and cybernetics, 2004, 34(4): 1907—1916.
- [6] 宫改云. FCM 参数研究及其应用[D]. 西安: 西安电子科技大学, 2004: 2—11.
- [7] 李云松, 李明. 基于灰度空间特征的模糊 C 均值聚类图像分割[J]. 计算机工程与设计, 2007, 28(6): 1358—1363.

作者简介:

朱红霞 女, 硕士研究生. 研究方向为图像处理与识别.