

一种改进的模糊C均值图像分割算法

刘洪普^{1,2}, 杨 乐¹, 侯向丹^{1,2}, 顾军华^{1,2}

(1. 河北工业大学 计算机科学与软件学院 天津 300401;

2. 河北工业大学 河北省大数据计算重点实验室 天津 300401)

摘要: 针对传统的模糊C均值(FCM)图像分割算法效率较低和分割结果对噪声敏感等问题,提出了一种改进的模糊C均值图像分割算法,可以有效地根据图像灰度分布信息选取初始聚类中心,同时充分考虑了邻域像素对于聚类的影响.结果表明,该算法能够根据图像特征自动初始化合适数量的近似聚类中心,对噪声图像具有较好的分割效果.

关键词: 模糊C均值; 初始聚类中心; 图像分割; 灰度直方图; 邻域

中图分类号: TP391.4

文献标志码: A

文章编号: 1671-6841(2017)02-0066-06

DOI: 10.13705/j.issn.1671-6841.2016328

0 引言

图像分割是图像处理中很重要的一部分,也是图像识别过程中一个很重要的子过程,图像分割效果的好坏将直接影响到图像识别的准确率.为取得更好的分割效果,越来越多的研究者将其他领域的理论模型和已有分割技术相融合产生出新的分割方法.其中基于模糊聚类分析的图像分割技术应用非常广泛^[1-3],模糊C均值(FCM)算法^[4]为其中的典型代表,该方法通过图像中各像素点与聚类中心的灰度值的欧氏距离,将不同的点归属于不同的聚类.由于在此过程中需要较少的人工参与,更有利于图像分割自动化的实现.然而FCM算法有其自身的缺点:①该算法无法自动确定聚类数目,需要手动确定聚类中心数目(聚类数);②原始算法只考虑了单个像素的灰度特征,忽略了邻域空间特征,因此分割结果很容易受到噪声的影响^[5].

针对上述缺点,随后出现许多改进的FCM图像分割方法.文献[6-7]中引入了灰度以外的其他特征作为聚类分割的考虑因素,文献[8-9]中引入了图论方法辅助图像分割,文献[10]提出了子图分割再合并的FCM方法,文献[11]中提出了一种能自适应模糊加权指数 m 的FCM方法,这些方法在图像分割方面比传统FCM方法取得了更好的效果,但大都忽略了邻域对于图像分割的影响.文献[12]提出了邻域加权FCM(neighboring weighted FCM, NW-FCM)方法,根据原图邻域像素信息生成新图像,然后根据新图像和原图像进行新的FCM的图像分割,此方法对于含噪声图像具有较好的分割效果,但仍需要人为确定聚类数目并随机初始化聚类中心.因此,本文在考虑图像中像素灰度分布和像素邻域空间信息的基础上,提出一种基于图像灰度分布直方图和空间邻域信息的FCM分割算法,该算法能够自动选取合适的初始聚类中心,并有效减少噪声影响.

1 传统FCM算法

假设图像由含 n 个像素的集合 $P = \{p_1, p_2, \dots, p_n\}$ 组成,图像分割就是将图像中的所有像素点分别划分到 c 个聚类中, $V = \{v_i\} (i = 1, 2, \dots, c)$ 为聚类中心矩阵,每一个聚类对应图像中一个部分的分割结果,通过

收稿日期: 2016-11-01

基金项目: 天津市自然科学基金项目(16JCYBJC15600).

作者简介: 刘洪普(1977—),男,河北廊坊人,讲师,主要从事图像处理及智能算法研究, E-mail: liuii@scse.hebut.edu.cn; 通讯作者: 侯向丹(1976—),女,河北石家庄人,副教授,主要从事智能算法及图像处理技术研究, E-mail: hxd@scse.hebut.edu.cn.

不断迭代使得目标函数不断收敛,最终完成聚类 and 图像分割. 其中 p_j 表示第 j 个像素的灰度值,其目标函数^[7]为

$$J_m = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|p_j - v_i\|^2, \quad (1)$$

式中: $\|\cdot\|$ 表示欧氏距离; m 为模糊加权指数,且 $m \in [1, \infty)$; 定义一个 $c \times n$ 的二维隶属矩阵 U , U 中的元素 u_{ij} 表示第 j 个数据点 p_j 属于聚类 i 的隶属度. 式(1)需满足以下约束条件:

$$\sum_{i=1}^c u_{ij} = 1, u_{ij} \in [0, 1], 0 \leq \sum_{j=1}^n u_{ij} \leq n. \quad (2)$$

根据拉格朗日乘子法,可以求得当目标函数取得最小值时,隶属度和聚类中心的求解式(3)和(4),可通过不断迭代这两个公式求得目标函数的最小值从而完成最终聚类.

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|p_j - v_i\|}{\|p_j - v_k\|} \right)^{2/(m-1)}}, \quad (3)$$

$$v_i = \frac{\sum_{j=1}^n u_{ij}^m p_j}{\sum_{j=1}^n u_{ij}^m}. \quad (4)$$

2 改进的 FCM 算法

2.1 通过直方图寻找初始聚类中心

提出了一种基于图像灰度分布直方图信息寻找初始聚类中心的方法. 该方法通过对原灰度直方图形成新的灰度分布直方图,使图像灰度分布趋势更加明显,再在新的直方图中确定初始聚类中心. 具体方法如下:

1) 求出输入图像的灰度分布直方图数据,并进行以 n 为单位长度的区间合并,用 t 表示合并后的新灰度级, x 表示新灰度级中包含的图像中各级灰度像素点平均数.

图1(a)为文献[13]中图像分割库的一幅图片,用它来提取直方图信息,图1(b)为图1(a)的灰度级直方图,可以清楚地看出图中主要的波峰和波谷,但由于细小的波峰、波谷的存在,程序很难对主要变化趋势进行准确的判断. 为此,以灰度值范围256的约数32、16、8、4作为单位长度 n 的可选值,对原直方图进行灰度区间合并,在新的分布直方图中灰度表示为 t_i , i 的范围为 $1 \sim 256/n$, t_i 的集合用 t 表示,以新的灰度 t_i 中包含的原来各灰度值的像素的数量平均值作为新数量值 x_i , x_i 的集合用 x 表示,这样就形成了一个趋势变化很明显的分布图,更有利于对于主体变化趋势的判断,如图1(c)所示. 可以看出,原直方图灰度区间合并后,灰度分布较为集中的区域($t_1 - t_3$, $t_3 - t_5$, $t_5 - t_8$)更为凸显,有助于划分像素分布集中的灰度区间.

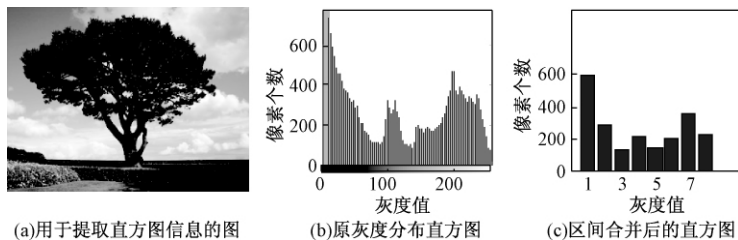


图1 两种灰度分布图的对比

Fig. 1 Comparison of the two kinds of gray level distribution

2) 在区间合并后的直方图中寻找极大值点和极小值点.

$$\text{极大值点: } \text{sgn}(x_{i+1} - x_i) - \text{sgn}(x_i - x_{i-1}) = -2, \quad (5)$$

$$\text{极小值点: } \text{sgn}(x_{i+1} - x_i) - \text{sgn}(x_i - x_{i-1}) = 2. \quad (6)$$

3) 求出合格的分割区间位置 $lseg$. 如果求出的 x 的极大值和极小值一共有 m 个,则可分为 $m-1$ 个变化区间,各区间 t 对应的像素点数 x 变化幅度可表示为 r_j , $j \in (1, m-1)$, 用 $\sigma(t)$ 表示 t 中 x 的标准差,选取满

足 $|r_j| > |\sigma(t)|$ 的下降区间,一直到产生 $c-1$ 个下降区间.用 $lseg_i$ 表示第 i 个分割点位置, $i \in (1, c-1)$.

4) 根据求得的分割区间位置将整个区间 t 划分为 c 个新的区间,步骤如下: (i) 将 $lseg_1$ 作为第一个分割的位置,将 t 分为左右两个部分; (ii) 以下一个分割点 $lseg_2$ 将 (i) 分割后的右侧区域分为左右两部分; (iii) 按照 (ii) 的方法依次类推,直到最后一个分割区域位置 $lseg_{end}$,新灰度区间 t 被分为 c 个子区间.每个子区间表示为 sub_q ,每一个子区间在 t 中的起始位置为 $lsub_q$,当 $q \neq 1$ 时, $lsub_q = lseg_{q-1}$; 当 $q = 1$ 时, $lsub_q = 1, q \in (1, c)$.

5) 求出各分子区间的平均灰度值作为初始聚类中心.将划分后的 t 灰度区间转化为原来灰度区间,求取各区间灰度平均值,由于分成了 c 个子区间,则产生 c 个初始聚类中心, n 为区间合并的单位长度,设分割子区间 sub_q 中 t 的个数为 $numt_q$,可求得 $numt_q = lsub_{q+1} - lsub_q$, $hist(g)$ 表示灰度为 g 的像素点的数量,第 q 个分割区间上的聚类计算公式为

$$c_q = \frac{\sum_{j=0}^{numt_q-1} \sum_{i=1}^n \{ [(lsub_q - 1 + j) \times n + i] \times hist((lsub_q - 1 + j) \times n + i) \}}{\sum_{j=0}^{numt_q-1} \sum_{i=1}^n hist((lsub_q - 1 + j) \times n + i)}. \quad (7)$$

2.2 结合邻域信息的 FCM

传统的 FCM 图像分割算法仅考虑了图像的灰度信息,忽略了空间信息的影响,使算法对噪声非常敏感,很难取得令人满意的分割效果.为此,在文献 [12] 中 NW-FCM 方法的基础上提出了一种改进的 FCM 算法.

在一幅 $h \times n$ 图像的每一个 $r \times r$ 的邻域中,邻域中心坐标为 (x, y) , $p(x+u, y+v)$ ($u, v \in \{-r, 0, r\} \wedge (u, v) \neq (0, 0)$) 是坐标为 $(x+u, y+v)$ 处像素的灰度值, $gap(u, v)$ 定义为邻域像素与中心像素的距离,则有

$$gap(x+u, y+v) = |p(x+u, y+v) - p(x, y)|, (u, v) \neq (0, 0), \quad (8)$$

$\overline{p(x, y)}$ 为邻域像素的平均值, $\overline{gap(x+u, y+v)}$ 为邻域距离的均值,设 $std_p(x, y)$ 为邻域像素的标准差,可构成如下条件判断式:

$$|p(x+u, y+v) - \overline{p(x, y)}| + gap(x+u, y+v) > std_p(x, y) + \overline{gap(x+u, y+v)}. \quad (9)$$

满足 (9) 式的 $p(x+u, y+v)$ 为邻域干扰元素,将此元素定义为 p_{jam} ,这些像素属于异常像素,将会对图像分割造成干扰,故需要删去.令 $p^*(x, y)$ 为邻域权值图像,可以表示为

$$p^*(x, y) = \overline{p_{val}(x, y)} = avg(\sum(p(x+u, y+v)) - \sum(p_{jam}(x+u, y+v))), \quad (10)$$

式中: $avg(\cdot)$ 为求均值运算符; $\sum(\cdot)$ 为求和运算符; $p^*(x, y)$ 为有效邻域元素的均值 $\overline{p_{val}(x, y)}$.

在 FCM 基础上加入邻域像素影响的新目标函数为

$$J_m = \sum_{i=1}^c \sum_{j=1}^N u_{ij}^m \|p_j - v_i\|^2 + a \sum_{i=1}^c \sum_{j=1}^N u_{ij}^m \|p_j^* - v_i\|^2, \quad (11)$$

式中: a 为惩罚因子,控制惩罚项的惩罚效果; $N = h \times n$ 为总像素点数.改进后的迭代公式为

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|p_j - v_i\|^2 + a \|p_j^* - v_i\|^2}{\|p_j - v_k\|^2 + a \|p_j^* - v_k\|^2} \right)^{m-1}}, \quad (12)$$

$$v_i = \frac{\sum_{j=1}^N u_{ij}^m (p_j + ap_j^*)}{(1+a) \sum_{j=1}^N u_{ij}^m}. \quad (13)$$

2.3 改进算法的实现

设 $P = \{p_1, p_2, \dots, p_n\}$ 为原图像中各点,算法具体步骤如下:

- 1) 通过 2.1 中算法求出图像的初始聚类中心 $\{C_i^{(0)}\}$ ($i = 1, 2, \dots, c$);
- 2) 根据式 (8) ~ (10) 计算邻域加权均值图像;
- 3) 用当前聚类中心 C_i 根据式 (12) 计算隶属度函数 u_{ij} ;
- 4) 判断是否满足 $\max |U^{(n+1)} - U^{(n)}| \leq \varepsilon$ 或 $n \geq n_{\max}$,如满足则算法终止;
- 5) 用当前隶属度函数 u_{ij} 根据式 (13) 修正聚类中心,返回 3)。

3 实验结果及分析

为了验证所提出算法的性能,对几幅不同图片进行初始聚类中心测试实验,并通过几幅人工合成图片和真实图片对本文算法的分割性能进行测试.

3.1 自适应聚类中心方法的实验

实验所用部分图片如图2所示,其中图2(b)、2(c)来自文献[14],用图2的3个图片对直方图初始聚类中心方法的准确度进行了测试.由于FCM方法获得的最终聚类中心与实际聚类中心十分接近,因此,将直方图方法获得的初始聚类中心和FCM方法最终得到的聚类中心进行对比,结果如表1所示.

可以看出,当图像比较简单时(如图2(a)),本文直方图方法查找到的聚类中心与传统FCM方法最终收敛到的聚类中心非常接近.对于较复杂图像(如图2(b)、2(c)),本文方法所得聚类中心与实际聚类中心仍比较接近.同时在此过程中,实验中的图片都得到了较为准确的聚类数.因此,相比于随机初始化聚类中心的方法,本文方法具有较好的参考价值,并可省去人为确定聚类数的操作.



图2 实验所用部分图片

Fig.2 The pictures used in the experiment

表1 聚类中心对比结果

Tab.1 The comparison results of cluster centers

图片	FCM 方法	本文直方图方法	聚类数
图片 1	113.473, 209.798, 22.981	111.680, 207.179, 30.790	3
图片 2	169.238, 49.315	190.768, 67.269	2
图片 3	65.604, 180.816	83.742, 195.477	2

3.2 图像分割实验

为了验证分割方法的有效性,通过FCM、NW-FCM和本文算法对于人工合成噪声图片、实物(树木)图片和遥感图像进行了分割测试实验.

首先对于人工合成噪声图片进行了分割测试.测试图片为黑白灰的棋盘图像,其中含有高斯噪声(均值为0,方差为0.01),参数设置为 $c=3$, $\alpha=50$, $r=2(5 \times 5 \text{ 邻域})$.实验结果如图3所示.

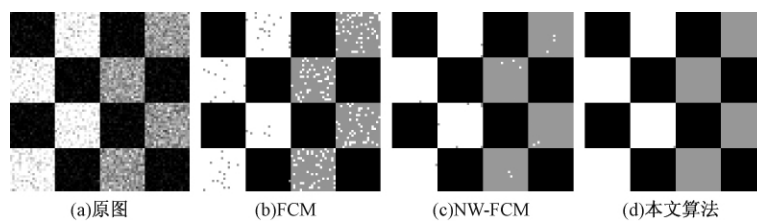


图3 人工合成噪声图片分割结果

Fig.3 The segmentation effect of synthetic noise image

FCM算法没有考虑邻域信息的影响,分割效果很差,分割后的图像中仍然含有较多噪声. NW-FCM算法由于较好地考虑了邻域像素的影响,因此分割后的结果去除了绝大部分噪声,仅在不同灰度像素的交界处有零星噪声.而本文方法由于充分考虑了邻域像素的影响,效果最佳.

对图1(a)树木图片进行了分割测试,结果如图4所示.可以看出,通过传统FCM算法分割的结果在地面见光处的噪声比较多,分割的区域并不连贯;树冠中的细小噪声也没有很好地消除.对于NW-FCM算法,由于考虑了邻域信息的影响,因此在树冠和地面上的分割效果有明显改善.本文方法则更好地去除了噪声的干扰,在树冠和地面处取得了更为连贯的分割区域,分割结果更佳.

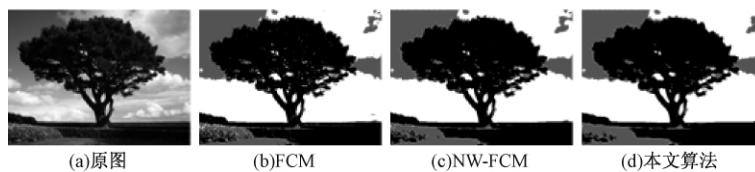


图4 树木图片分割结果

Fig.4 The segmentation results of tree image

对形态更为抽象杂乱的遥感图像进行了分割测试,分割结果如图5所示.可以看出,对于遥感图像分割的结果,NW-FCM和本文算法明显优于FCM方法,这两种方法在分割效果上非常接近,本文算法在局部细节上略优于NW-FCM算法.

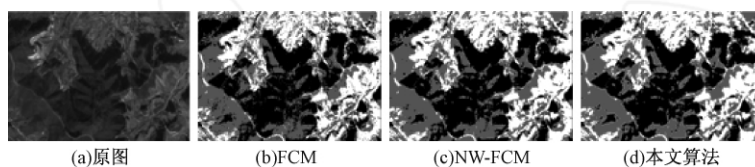


图5 遥感图像分割结果

Fig.5 The segmentation results of remote sensing image

可见本文结合邻域信息的FCM算法对于人工合成噪声图像、实物图像和遥感图像的分割结果都要优于FCM和NW-FCM算法.

4 结束语

针对FCM图像分割算法无法以合理的方法获得适当数目近似初始聚类中心,以及分割结果容易受到噪声影响的缺点,提出了一种通过图像灰度分布直方图信息初始化聚类中心,同时考虑邻域信息的改进FCM算法.实验结果表明,该算法能够较好地确定初始聚类中心的近似位置及数目,有效减少聚类过程噪声的影响,相对于FCM和NW-FCM聚类图像分割算法,本文算法分割效果更好.

参考文献:

- [1] DENG W Q, LI X M, GAO X F, et al. A modified fuzzy C-means algorithm for brain MR image segmentation and bias field correction [J]. Journal of computer science and technology, 2016, 31(3): 501–511.
- [2] JUI S L, LIN C, XU W, et al. Dynamic incorporation of wavelet filter in fuzzy C-means for efficient and noise-insensitive MR image segmentation [J]. International journal of computational intelligence systems, 2015, 8(5): 796–807.
- [3] TU X, GAO J, ZHU C, et al. MR image segmentation and bias field estimation based on coherent local intensity clustering with total variation regularization [J]. Medical & biological engineering & computing, 2016, 54(12): 1807–1818.
- [4] BEZDEK J C, HATHAWAY R J, SABIN M J, et al. Convergence theory for fuzzy C-means: counterexamples and repairs [J]. IEEE transactions on systems, man and cybernetics, 1987, 17(5): 873–877.
- [5] 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 transactions on medical imaging, 2002, 21(3): 193–199.
- [6] NICHAT A M, LADHAKE S A. Brain tumor segmentation and classification using modified FCM and SVM classifier [J]. International journal of advanced research in computer and communication engineering, 2016, 5(4): 73–76.

- [7] 依玉峰,高立群,郭丽.改进FCM在交互式图像分割中的应用[J].中国图象图形学报,2012,17(3):342-348.
- [8] 吴秋红,吴谨,朱磊,等.基于图论和FCM的图像分割算法[J].液晶与显示,2016,31(1):112-116.
- [9] HUANG Q H, LEE S Y, LIU L Z, et al. A robust graph-based segmentation method for breast tumors in ultrasound images [J]. Ultrasonics, 2012, 52(2):266-275.
- [10] 周晓明,李钊,刘雄英.一种基于改进FCM的自动图像分割算法[J].华南理工大学学报(自然科学版),2014,42(3):1-7.
- [11] 李晓冰.基于自适应模糊加权指数的FCM聚类测量图像分割方法[J].红外技术,2013,35(3):146-149.
- [12] 康家银,龚成龙,张文娟.基于邻域加权模糊C均值的遥感影像分割[J].系统仿真学报,2012,24(9):1969-1972.
- [13] ALPERT S, GALUN M, BASRI R, et al. Image segmentation by probabilistic bottom-up aggregation and cue integration [J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 34(2):315-327.
- [14] MARTIN D, FOWLKES C, TAL D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics [C]// Proceedings of 8th IEEE International Conference on Computer Vision. Vancouver, 2001:416-423.

An Improved Fuzzy C-means Algorithm for Image Segmentation

LIU Hongpu^{1,2}, YANG Le¹, HOU Xiangdan^{1,2}, GU Junhua^{1,2}

(1. School of Computer Science and Engineering, Hebei University of Technology, Tianjin 300401, China;

2. Key Laboratory of Big Data Computing of Hebei Province, Hebei University of Technology, Tianjin 300401, China)

Abstract: The traditional fuzzy C-means (FCM) image segmentation algorithm suffered from low efficiency and noise susceptibility. An improved fuzzy C-means algorithm for image segmentation was proposed. In the improved algorithm, the greyscale distribution of the original image was used to define the initial cluster centers; then the effects of the neighboring pixels on clustering were considered. Experimental results validated the effectiveness of the proposed method, demonstrating that the automatically chosen cluster centres enhanced the clustering efficiency. Additionally, the method was robust, with the clustering being less affected by noise in the original image.

Key words: fuzzy C-means; initial cluster center; image segmentation; greyscale histogram; neighborhood

(责任编辑:孔 薇)

