

An Image Segmentation Algorithm Based On Fuzzy C-Means Clustering

ZHANG Xinbo, JIANG Li

College of Information and Electronic Engineering, ZheJiang Gongshang University, Hangzhou, china

e-mail: zxb@mail.zjgsu.edu.cn, jl321123@163.com

Abstract: Image segmentation algorithm based on fuzzy c-means (fcm) clustering is an important algorithm in the image segmentation fields. However, it doesn't consider the special information but the gray information. That made the segmentation unideal. Therefore, we proposed an image segmentation algorithm based on weighed fcm. Moreover we optimized the parted pattern and initialized the number of the classification by the validity functions. The experimental result illustrates that the proposed method can realize the segmentation on the noised image or complicated image.

Keywords: fuzzy C means; image segmentation; weighted; validity functions

一种基于加权模糊 C 均值聚类的图像分割算法

张新波, 姜丽

浙江工商大学信息与电子工程学院, 杭州, 中国, 310018

e-mail: zxb@mail.zjgsu.edu.cn, jl321123@163.com

【摘要】基于模糊 C-均值(FCM)的图像分割方法是一种极其重要和应用相当广泛的分割算法。但这种算法仅利用了图像的灰度信息, 而没有考虑空间信息, 使得其对于含噪图像的分割效果不理想。因此, 本文考虑邻域像素的影响, 设计权值, 提出一种基于加权 FCM 的图像分割算法。此外, 对于较复杂的图像, 我们还通过算法的有效性函数来初始聚类类别数, 优化分割模型。实验结果表明, 本文提出的算法能有效的实现对含噪图像和复杂图像的分割。

【关键词】模糊 C-均值; 图像分割; 加权; 有效性函数

1 引言

图像分割就是指把图像分解成各具特性的区域并提取出感兴趣的目标的技术和过程, 它是由图像处理到图像分析的一个关键步骤。它在图像工程中占有非常重要的位置, 已在计算机视觉、模式识别和医学图像处理等实际应用中得到了广泛的应用^[6]。

在众多的分割算法中, 基于模糊 C 均值聚类(FCM)的图像分割方法是一类极其重要和应用相当广泛的算法。然而, 基于 FCM 的图像分割方法, 除利用图像的灰度信息外并没有考虑图像的空间信息, 这使得其对于含噪图像的分割效果不太理想。因此, 在进行分割时合理地考虑邻域像素对中心点的不同作用, 能够有效的改进对含噪图像的分割结果^[2,7,10]。

本文, 正是在上述思想上, 先用图像邻域内像素的平均值来代替其中心点, 再通过“距离偏差”来度量邻域模板内像素点对其中心点的影响, 设计权值, 提出了一种基于加权 FCM 的图像分割算法。该方法能充分考虑图像中邻域像素的影响, 对于含噪图像, 能很好的将目标从背景中提取出来, 有很好的分割效果。同时, 对于较复杂的图像, 本文还通过传统 FCM 算法的有效性函数, 提出先初始化复杂图像聚类类别数, 再对图像分割的思想, 这样不仅优化了聚类参数原型, 也能在一定程度上改善分割结果。

2 模糊 C 均值图像分割算法

模糊 C 均值聚类算法是一种迭代优化的无监督分类方法。对于样本集为 $X = \{x_1, x_2, \dots, x_n\} \subset R^n$ 的待分类图像 (n 是样本集中元素的个数, C 为样本的分类数), Bezdek^[1,3]将模糊 C 划分问题归纳为一个基于下述约

基金资助: 浙江省自然科学基金资助项目 (Y1080264); 浙江省高校青年教师资助计划项目

束条件下的求目标函数极值的问题:

$$\min\{J_m(U,V)\} = \min\left\{\sum_{j=1}^n \sum_{i=1}^C u_{ij}^m (d_{ij})^2\right\} \quad (1)$$

其中, $m \in [1, +\infty]$ 是一个加权系数, 控制分类矩阵 U 的模糊程度; d_{ij} 表示样本 x_j 与第 i 类的聚类中心 v_i 之间的距离. 模糊 C 划分用由隶属度 u_{ij} 构成的模糊矩阵 $U=[u_{ij}]_{C \times n}$ 来表示, u_{ij} 表示第 $j(j=1,2,\dots,n)$ 个样本点对第 $i(i=1,2,\dots,C)$ 个子集类的隶属度, 且 u_{ij} 满足如下条件:

$$\begin{aligned} u_{ij} &\in [0,1], \forall i,j; \sum_{i=1}^C u_{ij} = 1, \forall j; \\ 0 &< \sum_{j=1}^n u_{ij} < n, \forall i; \end{aligned} \quad (2)$$

根据拉格朗日乘法可求得 u_{ij} 和 v_i :

$$\begin{aligned} u_{ij} &= \left[\sum_{k=1}^C \left(\frac{d_{ij}}{d_{kj}} \right)^{\frac{2}{m-1}} \right]^{-1}, 1 \leq i \leq C, 1 \leq j \leq n \\ v_i &= \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m}, i=1,2,\dots,C \end{aligned} \quad (3)$$

FCM 图像分割算法, 就是通过不断的迭代来获得最佳模糊分类矩阵 U 和聚类中心矩阵 V , 再根据最大隶属度原则来确定最优的聚类划分, 从而得到最后的分割图像.

3 基于加权模糊 C 均值的图像分割算法

3.1 算法的设计

对于待分类的含噪图像, 充分考虑邻域像素的影响, 从两方面对传统的 FCM 图像分割算法进行改进.

- 对大小为 $M \times N$ 的含噪图像使用 3×3 的模板, 并用邻域内像素的平均值来代替其中心点(图像的边缘点不考虑), 对含噪图像进行预处理;

$$\begin{aligned} x(i,j) &= \frac{1}{9} (x(i-1,j-1) + x(i-1,j) \\ &+ x(i-1,j+1) + x(i,j-1) + x(i,j) \\ &+ x(i,j+1) + x(i+1,j-1) + x(i+1,j) \\ &+ x(i+1,j+1)) \end{aligned} \quad (4)$$

其中, i,j 表示第 i 行第 j 列.

- 把邻域内像素对中心点的影响, 定义为图像中各像素点的权值, 改进 FCM 算法的目标函数.

定义1: 对大小为 $M \times N$ 的含噪图像, 使用 3×3 的模板, 计算图像中第 $i+r_1$ 行第 $j+r_2$ 列个像素点对第 i 行第 j 列个像素点(即模板中心点)的影响, 称

$$t(i+r_1, j+r_2) = 1 - \frac{|x(i+r_1, j+r_2) - x(i, j)|}{\sum_{r_1=-1}^1 \sum_{r_2=-1}^1 |x(i+r_1, j+r_2) - x(i, j)|} \quad (i=1,2,$$

$\dots, M; j=1,2,\dots,N)$ 为距离偏差, 式中 $r_1=-1,0,1; r_2=-1,0,1$. 可以看出, $t(i+r_1, j+r_2) \in [0,1]$ 且其值越大影响越大;

定义 2: $i=1,2,\dots,M; j=1,2,\dots,N$ 称:

$$wd(i, j) = \left(\sum_{r_1=-1}^1 \sum_{r_2=-1}^1 \frac{1}{t(i+r_1, j+r_2)} \right)^{-1}$$

为图像中各个像素点所对应的初始权重. 对 $wd(i,j)$ 作归一化处理, 得权重 w_{ij} 且

$$w_{ij} = \frac{wd(i, j)}{\sum_{i=1}^M \sum_{j=1}^N wd(i, j)}$$

可以看出, w_{ij} 越大, 模板内的领域像素对中心点的影响就越大. 同样, 图像边缘像素点的权值不通过上式定义考虑(这里用其相邻行或列的像素点的权值来代替).

此时, 对于待分类的图像样本 $X=[x_{ij}]$ ($i=1,2,\dots,M; j=1,2,\dots,N$) 考虑加权信息, 将 FCM 算法的目标函数改写为:

$$J'_m(U,V) = \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^C w_{ij} u_{kij}^m d_{kij}^2 \quad (5)$$

根据拉格朗日乘法, 隶属度 u_{kij} 和聚类中心 v_k 为:

$$\begin{aligned} u_{kij} &= \left[\sum_{s=1}^C \left(\frac{d_{kij}}{d_{sij}} \right)^{\frac{2}{m-1}} \right]^{-1} \\ v_k &= \frac{\sum_{i=1}^M \sum_{j=1}^N w_{ij} (u_{kij})^m x_{ij}}{\sum_{i=1}^M \sum_{j=1}^N w_{ij} (u_{kij})^m} \end{aligned} \quad (6)$$

3.2 仿真实验验证

为了验证上述分割算法的有效性, 在同一硬件条件和 matlab 平台上, 我们分别采用传统 FCM 算法、文献[7]中提出的由二维直方图构造权值的算法和本文提出的分割算法, 对两幅图像进行试验.

图 1(a) 是一幅 128×128 大小的叠加了 $N(0,25)$ 的高斯白噪声实际图像“鹅”. 图 1(b)、图 1(c) 和图 1(d) 为对应的三种算法的分割结果.



Figure 1(a). The image with noise
图 1(a). 加了噪声的原图

观察这些实验得到的图形直观结果, 我们可以看到, 本文中提出的新算法能获得较好的分割结果.



Figure 1(b). The segmentation result by FCM algorithm
图 1(b). FCM 算法的分割结果



Figure 1(c). The segmentation result by algorithm in literature[7]
图 1(c).文献[7]中算法的分割结果



Figure 1(d). The segmentation result by algorithm in this paper
图 1(d). 本文中算法的分割结果



Figure 2(a). The image with noise
图 2(a). 加了噪声的原图



Figure 2(b). The segmentation result by FCM algorithm
图 2(b). FCM 算法的分割结果



Figure 2(c). The segmentation result by algorithm in literature[7]
图 2(c).文献[7]中算法的分割结果

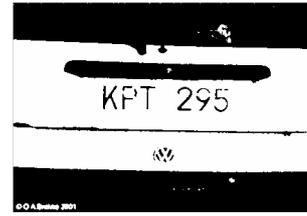


Figure 2(d). The segmentation result by algorithm in this paper
图 2(d). 本文中算法的分割结果

图 2(a)是一幅 302×436 大小，同样叠加了 N(0,25) 的高斯噪声的汽车图像。三种算法的分割结果见图 2(b)、图 2(c)和图 2(d)。

4 一种针对复杂图像的改进分割算法

基于 FCM 的图像分割算法的效果，有时也依赖于算法中聚类原型参数的选择。因此，对于一幅相对复杂的图像，如果能获得对图像的先验认识（比如：初始聚类类别数），这样不仅可以优化聚类原型参数，也可以改善图像的分割效果。

本文中，我们先用传统 FCM 算法的有效性函数(划分系数 $F(U;C)$ 和划分熵 $H(U;C)$)对算法进行有效性分析，并选择使划分系数最大和划分熵最小的 C 值作为聚类类别数。再用上文提出的改进算法对图像进行分割。

模糊划分系数 $F(U;C)$ 和划分熵 $H(U;C)$ 的定义如下^[1]：

$$F(U;C) = \frac{1}{n} \sum_{i=1}^c \sum_{j=1}^n u_{ij}^2$$

$$H(U;C) = -\frac{1}{n} \sum_{i=1}^c \sum_{j=1}^n u_{ij} \ln(u_{ij}) \quad (7)$$

例如，对于一相对复杂的实际图像 lena，由表 1 可见，lena 图像的最佳聚类类别数 $C=6$ 。在确定了聚类类别数后，再分别用上文的三种算法对图像进行分割。图 3(a)是一幅 256×256 大小的实际图像“lena”，也叠加了 N(0,25)的高斯白噪声。图 3(b)、图 3(c)和图 3(d)是给出的实验结果，可以看出，此方法可以改善图像的分割效果。

Table 1. The $F(U;C)$ and $H(U;C)$ of lena
表 1. lena 图像的有效性函数 $F(U;C)$ 和 $H(U;C)$

函数	C=4	C=5	C=6	C=7	C=8	C=9
F	0.8034	0.79222	0.81037*	0.78176	0.78096	0.77227
H	0.1121	0.11481	0.10551*	0.11666	0.11570	0.11916



Figure 3(a). The image with noise
图 3(a). 加了噪声的原图



Figure 3(b). The segmentation result by FCM algorithm
图 3(b). FCM 算法的分割结果



Figure 3(c). The segmentation result by algorithm in literature[7]
图 3(c).文献[7]中算法的分割结果



Figure 3(d). The segmentation result by algorithm in this paper
图 3(d). 本文中算法的分割结果

此外，对于本文中用到的每幅图像，我们运行三种算法各5次，并取5次的运行时间平均值作为各算法对于每个图像的分割时间，见表2。从运行时间上看，本算法也是可行的。

5 结论

图像分割是图像处理研究的一个重要方面,图像

Table 2. Segmentation time of three kinds algorithm
表 2. 三种算法对于各个图像的分割时间

分割时间 (s)	FCM 算法	文献[7]中算法	本文中算法
实际图像“鹅”	1.657	1.186	1.265
汽车图像	98.203	100.063	82.781
图像“lena”	45.344	64.8878	32.2748

分割的质量直接影响到更高层的图像分析和理解。利用 FCM 聚类算法进行图像的分割，能有效地减少人为的干预，对于存在模糊性的图像比较适合。但是，由于传统的 FCM 算法中未能考虑邻域信息的影响，因此算法的抗噪声能力较差。本文提出了一种加权的 FCM 分割算法，该方法在 FCM 算法中引入邻域像素的影响，设定权重,完善了分割模型。另外，对于较复杂的图像，通过传统 FCM 算法的有效性函数，不仅可以获得聚类原型参数，也可以改善图像的分割效果。实验结果表明本文提出的算法是相当有效的，其分割效果优于传统的 FCM 分割算法，也优于文献[7]中提出的算法，对噪声具有较强的鲁棒性,有一定的实际意义。

References (参考文献)

- [1] Bezdek J C. Pattern Recognition with Fuzzy Objective Function Algorithms [M].New York: Plenum Press, 1981.
- [2] JIANG Ai-ping, YANG Yue-hua, Yang Xing-quan. Multifractal Image Segmentation Based on Multi-weight Method [J], Journal of Image and Graphics., 2007, 12(10), P1889-1892 (Ch). 蒋爱平,杨悦华,杨兴全. 基于多重加权法的多重分形图像分割研究. 中国图像图形学报. 2007, 12(10):1889-1892.
- [3] Gao Xinbo. Fuzzy Cluster Analysis and its Application[M]. Xi'an: Publishing House of XIDIAN University,2004.1(Ch). 高新波.模糊聚类分析及其应用[M].西安: 西安电子科技大学出版社. 2004.1.
- [4] SONG Yu-chen, ZHANG Yu-ying, MENG Hai-dong. Research based on euclid distance with weights of clustering method[J]. Computer Engineering and Applications.2007,43(4): 179- 180 (CH). 宋宇辰,张玉英, 孟海东.一种基于加权欧氏距离聚类方法的研究[J].计算机工程与应用.2007,43(4): 179-180.
- [5] R J. Hathway and J. C. Bezdek. Optimization of clustering criteria by reformulation. IEEE Transactions on Fuzzy System[J]. May 1995;vol. 3, no. 2, pp.241-245.
- [6] Refael C.Gonzalez,Richard E.Woods.Digital Image processing(second edition)[M]. Pearson Education Asia Limited And Publishing House of Electronics Industyt:2007.8
- [7] YANG Run-ling, GAO Xin-bo. A Fast Automatic Image Segmentation Algorithm Based on Weighting Fuzzy c-Means Clustering[J]. Journal of Image and Graphics, 2007, 12(12): 2105-2112(CH). 杨润玲,高新波. 基于加权模糊 C 均值聚类的快速图像自动分割算法[J]. 中国图像图形学报.2007,12(12):2105-2112.
- [8] LI Zheng-zhou; PENG Su-jing; WANG Yun; LIU Guo-jin. Gray Image Segmentation Algorithm Based on Mean Shift[J]. Acta Photonica Sinica, 2007, 36(SUP): 286-289(CH). 李正周,彭素静,王允,刘国金. 基于均值偏移的灰度图像分割方法[J]. 光子学报. 2007, 36(SUP): 286-289.
- [9] LI Yan, SUN Jing-guang, ZHANG Xin-jun. Research on a Fast Segmentation Algorithm with FCM Clustering[J]. Computer & Information Technology. 2008.3(CH). 李岩,孙劲光,张新君. 一种基于 FCM 聚类的图像快速分割算法研究[J]. 计算机与信息技术. 2008.3
- [10] LIU Jian-zhuang. Image fuzzy cluster segmentation Based on Two Dimension Histogram [J]. Acta Electronica Sinica, 1992,20(9): 40~46(CH). 刘健庄.基于二维直方图的图像模糊聚类分割方法[J].电子学报, 1992,20(9): 40~46.