

文章编号: 1001- 2486(2010) 02- 0051- 05

# 结合峰值检测的高噪声图像快速聚类分割方法\*

赵 晖, 林成龙, 唐朝京

(国防科技大学 电子科学与工程学院, 湖南 长沙 410073)

**摘要:** 针对高噪声污染图像, 提出一种结合峰值检测算法的快速聚类分割方法(FG-ImSeg)。根据平行线投影分割算法将二维直方图映射到一维空间, 利用峰值检测算法检测图像像素点的聚类结果, 调整映射模型的平行线宽度, 使直方图符合双峰分布特性, 最后利用加权模糊  $c$  均值聚类算法实现图像的分割。实验结果证明了该方法是快速有效的。

**关键词:** 快速聚类分割; 平行线投影分割; 峰值检测; 模糊  $c$  均值聚类

中图分类号: TP391. 41 文献标识码: A

## Fast Clustering Segmentation Method Combining Peak Testing for Image with High Noise

ZHAO Hui, LIN Cheng-long, TANG Chao-jing

(College of Electronic Science and Engineering, National Univ. of Defense Technology, Changsha 410073, China)

**Abstract:** Aiming at image with high noise, a fast clustering segmentation approach combining peak testing is proposed in this paper. Two-dimension histogram was converted into one-dimension with parallel projection segmentation algorithm. Peak testing algorithm was used to test the clustering result of the pixels in noisy image. Parallel width in projection model was adjusted according to the testing result. As the histogram is approximated by double-peak distribution, noisy image can be segmented with weighted Fuzzy  $c$ -Means clustering algorithm. Experimental results proved that the proposed approach is fast and effective.

**Key words:** fast clustering segmentation; parallel projection segmentation; peak testing; fuzzy  $c$  means clustering

模糊  $c$  均值聚类方法(FCM) 在图像的聚类分割中得到了广泛应用<sup>[1-4]</sup>。早期的研究主要根据图像的一维灰度直方图实现分割, 然而一维直方图仅仅反映了图像的灰度分布特性, 没有体现空间相关特性, 虽然分割速度较快, 但当信噪比降低时, 分割的效果十分不理想。利用二维直方图实现图像的聚类分割是一种有效的方法, 由于充分利用了图像的灰度值信息和邻域的空间信息, 二维直方图分割方法比一维直方图分割方法具有更强的抗噪性能<sup>[5-6]</sup>。但二维直方图的引入大大增加了运算复杂度, 增加了运算时间, 降低了处理效率。一些研究方法<sup>[7-8]</sup> 针对传统二维直方图中由判决区域造成误分割的情况, 将方形判决区域改为斜方向判决区域, 减少了误判率, 提高了运算速度, 但仍没有从根本上提高运算效率。

为了进一步提高噪声图像的分割精度, 减少运算时间, 提出一种在高噪声条件下结合峰值检测算法的图像快速聚类分割方法(FG-ImSeg), 该方法在得到双峰检测结果之后, 运用基于直方图的加权模糊  $c$  均值聚类算法(HistFCM) 实现了图像的分割。

### 1 平行线投影分割算法

图像的目标点和背景点在二维直方图中都分布在对角线的附近。图 1 给出了含有高斯噪声  $N(0, 100)$  和  $N(0, 1300)$  的图像对应的二维直方图, 两幅图中点的分布都是以对角线为中心的, 在含有低噪声的(a)图中, 点的分布较紧凑, 而在含有高噪声的(b)图中, 点的分布比较分散。

\* 收稿日期: 2009- 09- 09

基金项目: 国家部委资助项目(51329060101)

作者简介: 赵晖(1980—), 男, 博士生。

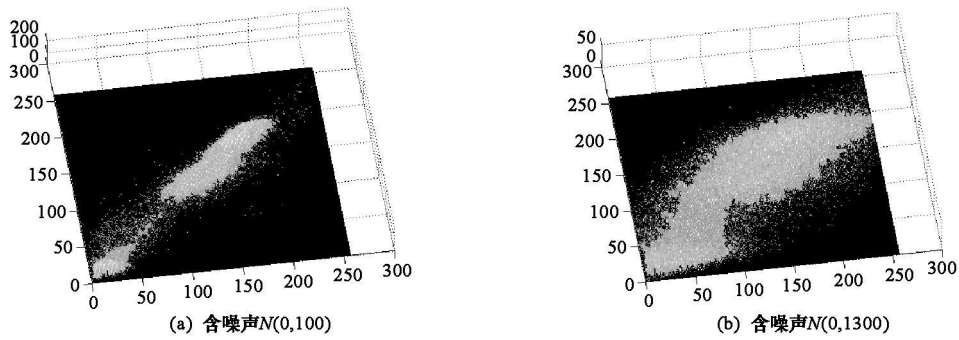


图1 含噪声图像的二维直方图

Fig. 1 Two-dimension histogram of noisy image

图1(b)中,有相当一部分点距离对角线较远,离对角线越远的点,图像中相应像素的灰度值与邻域平均灰度值的差距越大。由于图中点分布范围较大,若按照传统的二维聚类分割方法,会有较多点被误分割。利用平行线投影分割算法(PPSeg)<sup>[9]</sup>,将平行线内的点映射到对角线上,将二维直方图映射为一维,如图2所示。在该方法中,关键问题是对角线与其两侧平行线距离(即 $N$ 值)的选取,具体选取方法将在第3节中论述。

## 2 加权模糊聚类图像分割方法

利用FCM方法对图像中的不同属性点聚类,实现图像的分割。传统的FCM算法中,将样本集 $X = \{x_1, x_2, \dots, x_n\}$ 分为 $c$ 类,样本 $x_k$ 对 $i$ 类的隶属度为 $u_{ik}$ ,通过最小化目标函数 $J_m(u, v)$ 实现聚类。

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

隶属度矩阵 $U = \{u_{ik}\} \in R^{n \times c}$ 满足:

$$\begin{cases} u_{ik} \in [0, 1], & 1 \leq i \leq c, 1 \leq k \leq n \\ \sum_{i=1}^c u_{ik} = 1, & 1 \leq k \leq n \\ 0 < \sum_{k=1}^n u_{ik} < n, & 1 \leq i \leq c \end{cases} \quad (2)$$

其中, $2 \leq c \leq n$ , $m$ 为加权指数,Pal<sup>[10]</sup>的研究表明, $m$ 最佳选择范围为 $[1.5, 2.5]$ 。为了提高图像聚类的运算速度,提出基于直方图的加权 $c$ 均值聚类算法(HistFCM)。对于一幅图像 $f(x, y)$ ,样本集 $X = \{l \in f(x, y)\}$ ,具有 $L$ 灰度层级。对于具有相同灰度级的像素点,在聚类时隶属度一定是相等的。设灰度级为 $l \in L$ 的一维直方图为 $H_l$ ,则式(1)可以改写为

$$J_m(u, v, l) = \sum_{l=0}^{L-1} \sum_{i=1}^c (u_{il})^m H_l \|l - v_i\|^2 \quad (3)$$

隶属度矩阵和聚类中心分别为

$$u_{il} = \frac{1}{\sum_{j=1}^c \left( \frac{\|l - v_i\|}{\|l - v_j\|} \right)^{\frac{2}{m-1}}} \quad (4)$$

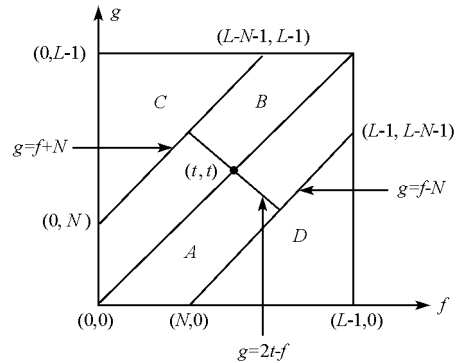


图2 平行线投影分割方案

Fig. 2 Parallel projection segmentation method

$$v_i = \frac{\sum_{l=0}^{L-1} (u_{il})^m H_l \cdot l}{\sum_{l=0}^{L-1} (u_{il})^m H_l} \quad (5)$$

不断迭代更新式(4)和式(5),直到式(3)收敛。HistFCM 算法大大降低了聚类算法的复杂度,节约了计算时间。

### 3 结合峰值检测的聚类分割方法

对于 PPSeg 算法,在进行聚类分割时,要选取恰当的  $N$  值(图2),如果  $N$  值过小,将会损失较多有用的背景和目标点,在分割时造成误判;如果  $N$  值过大,平行线内会包含过多的噪声点,此时,直方图的双峰分布特征不明显,同样容易造成误分割。因此提出一种结合峰值检测和 PPSeg 算法、HistFCM 算法的 FG-ImSeg 方法,如图3所示。

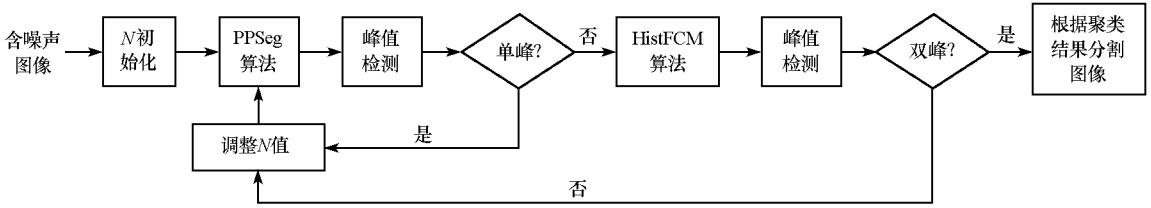


图3 结合峰值检测的快速聚类分割方法(FG-ImSeg)流程图

Fig.3 Flow chart of fast clustering segmentation approach combining peak testing

图中的 HistFCM 算法中,聚类数  $c = 2$ , 加权指数  $m = 2$ 。另外,  $N$  的初值选取十分重要,经验证,当  $N = L \cdot 50\%$  时,图2中  $C$  区和  $D$  区中的像素点主要为噪声点和边缘点,因此设初始值  $N = L \cdot 50\%$ , 调整  $N$  值时,平行线每次向对角线方向平移  $L \cdot 5\%$ , 迭代过程为

$$N_{(i+1)} = N_{(i)} - L \cdot 5\% \quad (6)$$

借鉴文献[11]的做法,引入聚类趋势检验的概念来实现峰值检测。聚类趋势检验中用到了半数框架和  $\alpha$ -显著性检验两个算法。

单峰检测算法1 建立半数框架

(1) 设模式集  $X = \{x_1, x_2, \dots, x_n\} \subset R^p$  是  $p$  维实数空间  $R^p$  中的一个子集,计算均值向量  $m$ :

$$m = \left( \sum_{i=1}^n x_i \right) / n \quad (7)$$

(2) 计算模式样本  $x_i$  与均值向量  $m$  间的距离  $d_i = \|x_i - m\|$ , 并从小到大排列  $d_1, d_2, \dots, d_n$ , 取  $r = d_{\lfloor n/2 \rfloor}$ ,  $\lfloor \cdot \rfloor$  为取整符号;

(3) 球  $(m, r)$  包含的区域即为抽样始点设置域,  $k$ -近邻模式  $P_1, P_2$  允许在球外寻找。

单峰检测算法2  $\alpha$ -显著性检验

(1) 在给定模式集  $X$  中建立半数框架  $(m, r)$ , 令  $s = 0$ ;

(2) 在球  $(m, r)$  中随机设置  $M$  个抽样始点, 计算标准化的统计量  $T_k$ :

$$T_k = \left[ \frac{1}{M} \sum_{i=1}^M \frac{U_i^p(k)}{U_i^p(k) + V_i^p(k)/2} - 0.5 \right] \times \sqrt{12M} \quad (8)$$

(3) 如果  $T_k \geq T(\alpha)$ , 则  $s = s + 1$ , 其中  $\alpha$  为选定的显著水平,  $T(\alpha)$  为相应的临界值, 满足标准的正态分布;

(4) 重复(2)、(3)步  $T$  次(一般  $T = 100$ ), 计算检验大小  $s = s/T$ , 如果  $s \gg \alpha$ , 则认为  $X$  为多峰分布模式, 具有可分性; 反之, 如果  $s$  与  $\alpha$  在同一数量级, 则认为  $X$  为单峰分布模式。

## 4 实验结果与分析

采用结合峰值检测的聚类分割方法,给出了高噪声条件下图像的分割结果。本实验的环境为 Celeron CPU 2.00GHz, 内存 2.00GB, 用 Matlab 7.6 编程实现。根据图 3, 由 PPSeg 算法得到的一维直方图需要满足两个条件: (1) 不具备单峰分布特性; (2) 聚类的两个子集分别具备单峰分布特性, 即直方图具备双峰分布特性。图 4 为含高噪声图像的峰值检测结果。

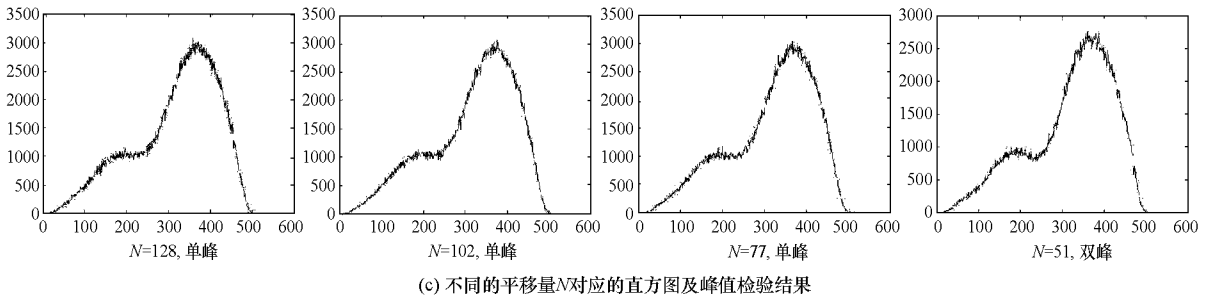
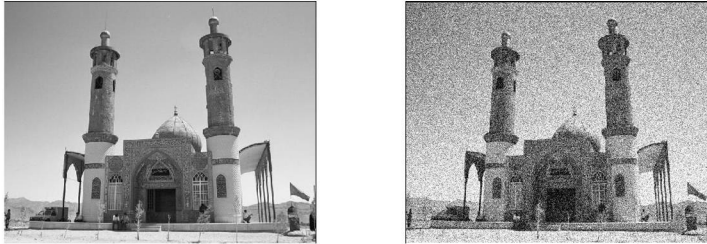


图 4 含高斯噪声  $N(0, 1300)$  清真寺图像的峰值检测结果  
Fig. 4 Peak testing result of "Temple" with Gaussian noise  $N(0, 1300)$

设定灰度级  $L=256$ , 根据式(6),  $N$  在计算过程中递减, 图中仅列出当  $N=128, 102, 77, 51$  时的峰值分布情况。当  $N > 51$  时, 对直方图的峰值检测结果皆为单峰, 继续检测; 当  $N=51$  时, 峰值检测结果为双峰, 检测停止, 此时检测迭代次数为 7。最后根据聚类结果实现图像分割。

针对图 4(b) 中的高噪声图像, 将各种不同的聚类分割方法进行比较, 分割结果如图 5 所示, 其中 (c) 为文献[4]中的基于空间邻域信息的二维模糊聚类图像分割方法(2DFCM\_Aver), (d) 为结合峰值检测的快速聚类分割方法 FC-ImSeg。

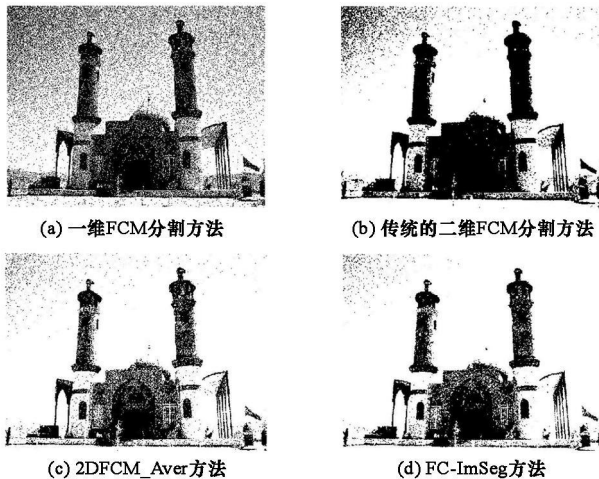


图 5 图像聚类分割结果

Fig. 5 Image segmentation result by clustering

可见, 一维 FCM 分割方法效果最差, 噪声点没有有效去除; 传统的二维 FCM 分割方法, 由于判决区域是正方形, 造成部分误判, 分割错误率较高; 2DFCM \_Aver 方法较(b)中的分割效果好, 但在对目标的边缘分割上有一定误判; FG-ImSeg 方法克服了上述缺点, 噪声点去除彻底, 目标与背景特征清晰, 边缘平滑性较好。

表 1 给出了四种聚类分割方法的性能。一维 FCM 方法的错分割率最高, 相当于 FG-ImSeg 方法的 4 倍, FG-ImSeg 方法用于单次检测的运算时间应与一维 FCM 方法接近, 由于进行了 7 次迭代检测, 其运算时间接近于一维 FCM 方法的 7 倍。和传统二维 FCM 方法和 2DFCM \_Aver 方法相比, FG-ImSeg 方法的性能无论在错分割率还是运算时间上都较优。错分割率可从图 5 的分割结果中得到印证。

表 1 四种分割方法的性能比较

Tab. 1 Comparison of performance between 4 segmentation methods

|             | 错分割点数  | 错分割率(%) | 运算时间(s) |
|-------------|--------|---------|---------|
| 一维 FCM      | 27 014 | 6.87    | 1.20    |
| 传统二维 FCM    | 14 195 | 3.61    | 296.54  |
| 2DFCM _Aver | 11 167 | 2.84    | 23.15   |
| FG-ImSeg    | 6724   | 1.71    | 7.02    |

## 5 结论

针对高噪声图像, 提出了一种结合峰值检测的快速聚类分割方法(FG-ImSeg)。该方法利用 PPSeg 算法将二维问题映射为一维问题, 并用 HistFCM 算法实现对图像的聚类分割, 用峰值检测方法确定映射模型中平行线的位置, 通过检测迭代过程, 得到具有双峰性质的一维映射直方图。FG-ImSeg 方法不仅继承了二维聚类分割方法的抗噪性和精确性, 并且继承了一维聚类分割方法的时效性。经实验验证, 与其他聚类分割方法相比较, FG-ImSeg 方法误分割率最低, 具有快速、准确的分割特性。在将来的工作中, 根据不同的分割需求, 可以通过增加 HistFCM 算法中的聚类数, 实现图像的多目标聚类分割。

## 参考文献:

- [1] 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]. 中国图象图形学报, 2007, 12(12): 2105-2112.
- [2] 高新波, 李洁, 姬红兵. 基于加权模糊  $c$  均值聚类与统计检验指导的多阈值图像自动分割算法[J]. 电子学报, 2004, 32(4): 661-664.
- [3] 王丹, 吴孟达. 粗糙模糊  $c$ -均值算法及其在图像聚类中的应用[J]. 国防科技大学学报, 2007, 29(2): 76-80.
- [4] 余锦华, 汪源源, 施心陵. 基于空间邻域信息的二维模糊聚类图像分割[J]. 光电工程, 2007, 34(4): 114-119.
- [5] Wu S, Amin A. Automatic Thresholding of Gray Level Using Multi-stage Approach[C]//Proceedings of 2003 IEEE International Conference on Document Analysis and Recognition, Edinburgh, Scotland, 2003: 1238-1242.
- [6] Brink A D. Thresholding of Digital Image Using Two-dimensional Entropies[J]. Pattern Recognition, 1992, 25(8): 803-808.
- [7] 梁光明, 刘东华, 李波, 等. 用于显微细胞图像的二维自适应阈值分割算法的优化[J]. 中国图象图形学报, 2003, 8(7): 764-768.
- [8] 吴一全, 潘磊, 吴文怡. 二维直方图区域斜分阈值分割及快速递推算法[J]. 通信学报, 2008, 29(4): 77-83.
- [9] Zhao H, Tang C J, Yu T. Fast Thresholding Segmentation for Image with High Noise[C]//Proceedings of 2008 IEEE International Conference on Information and Automation, Zhangjiajie, China, 2008: 290-295.
- [10] Pal N P, Bezdek J C. On Cluster Validity for the Fuzzy G-means Model[J]. IEEE Trans. Fuzzy Systems, 1995, 3(3): 370-379.
- [11] 高新波, 裴继红, 谢维信. 基于统计检验指导的聚类分析方法[J]. 电子科学学刊, 2000, 26(1): 6-12.