

文章编号:1671-6833(2017)06-0064-06

基于 Mean Shift 聚类的多级阈值化方法

程全¹, 刘晓青¹, 刘玉春¹, 王志良²

(1. 周口师范学院 机械与电气工程学院, 河南 周口 466001; 2. 北京科技大学 计算机与通信工程学院, 北京 100083)

摘要:为了解决多级阈值化技术中所选阈值的数量通常不能预先确定的问题,提出一种基于 Mean Shift 聚类技术的新型多级阈值化方法.首先,通过使用 Mean Shift 技术探寻出潜在的模式中心,应用迭代的阈值选择方法来自动确定相邻模式中心的各个阈值;然后,采用多级阈值化对图像进行分割;最后,通过实验验证了基于 Mean Shift 聚类技术分割的图像相对于原始图像的对比度有了很大提高.该方法通过简单修改程序参数就能够灵活控制分割精度,可以广泛应用于单阈值分割、多级阈值分割和有损压缩等技术中.

关键词:多级阈值化; 图像分割; 迭代阈值化; 分割质量评估

中图分类号: TP391 **文献标志码:** A **doi:**10.13705/j.issn.1671-6833.2017.06.009

0 引言

图像分割是进行图像理解和目标识别的一个必不可少的步骤,图像分割技术分为软分割和硬分割^[1].聚类算法是软分割方法中最主要的图像分割算法,主要有两种:一种是 K-means 聚类^[2],其优点是易于实现但需要预先指定待分割图像中需要形成的群集的数量及中心.另一种是模糊聚类^[3],当图像中不同物体之间没有明显的边界时,可以基于相似性准则采用模糊聚类的方法进行处理,但是它对噪声特别敏感,并且模糊关系的确定也非常繁琐.软分割计算起来相当费时;而硬分割是一种基于亮度信息的分割方法,它假设图像的直方图含有一个或多个峰值^[4],不需要图像的先验信息、计算简单,实时性高,但它忽略了图像的空间信息,阈值选择不当,经常会导致过分割或欠分割.实际上,该方法可以被认为是一个多级阈值化问题,其中所选取的多个阈值对应于相邻显著峰值的峰谷点.

多级阈值化技术将图像的直方图分为几个部分,当某一图像像素的特征值位于由某一对相邻阈值确定的范围时,该像素就被分到某一相应的部分.为了解决多级阈值分割难题,近年来群智能技术被用来寻找最优阈值. Raja 等人^[5]采用一个改进的粒子群(PSO)算法用于恶性肿瘤热图像的多

阈值分割. Sarkar 等^[6]基于差分进化算法(DE)和最小交叉熵提出了一种新颖的多阈值彩色图像分割方法.

尽管上述方法可以确定出最优阈值,但是不能提供对应于图像直方图中潜在模式数量最优阈值数.因此,笔者提出了一种新型的最优多级阈值化方法,特别适用于具有多模态直方图的图像.首先使用 Mean Shift 程序确定出直方图的各个模式中心;然后使用一种迭代的阈值选取方法来确定相邻模式中心的各个阈值;最后利用已经确定的多个阈值对图像进行多阈值化分割.

1 算法部分

1.1 Mean Shift

Mean Shift 是一种非参数化的聚类方法^[7],其最大优点是它无须预先给出聚类的数量,且对聚类的形状没有任何限制.作为一个迭代的模式探寻方法,Mean Shift 程序利用核来计算观察窗口的特征加权平均,并可以准确地定位出聚类的中心并完成特征空间的划分.

对于 d 维空间 R^d 上的一个 n 点的数据集合 $\{x_i\}_{i=1, \dots, n}$, 点 x 处的多变量核密度估计为:

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{x-x_i}{h}\right\|^2\right), \quad (1)$$

式中: $k(x)$ 是侧面轮廓函数; h 是窗口半径. 令式

收稿日期:2017-05-05; 修订日期:2017-08-10

基金项目:国家自然科学基金资助项目(61401526);河南省自然科学基金资助项目(152300410134)

作者简介:程全(1978—),男,河南沈丘人,周口师范学院副教授,主要从事智能控制研究,E-mail:quan8888@126.com.

(1) 的梯度为 0, 得到密度函数的驻点为

$$x = \frac{\sum_{i=1}^n x_i \cdot g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}, \quad (2)$$

式中: $g(x) = -k'(x)$ 是核函数 $G(x)$ 的侧面轮廓函数.

连续迭代计算 Mean Shift 矢量, 并进行移位直到程序收敛, 就能定位出局部模态. 迭代的方程为

$$y_{j+1} = \frac{\sum_{i=1}^n x_i \cdot g\left(\left\|\frac{y_j-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{y_j-x_i}{h}\right\|^2\right)}, j = 1, 2, \dots, \quad (3)$$

式中: y_1 为核窗口的初始位置中心; y_{j+1} 为 y_j 处利用核 G 和窗口半径 h 计算出的加权平均.

对于 Epanechnikov 核为

$$K_E(x) = \begin{cases} (2c_d)^{-1}(d+2)(1-x^T x), & \text{if } x^T x < 1; \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

式中: c_d 为单位 d 维球体的体积.

Mean Shift 的公式为

$$y_{j+1} = \frac{1}{n_{y_j, x_i \in S_h(y_j)}} \sum x_i - y_j, \quad (5)$$

式中: $S_h(y_j)$ 为中心位于 y_j ; 半径为 h ; 包含 n_{y_j} 个数据点的超球体.

选择图像像素的灰度值作为特征空间, Mean Shift 程序可以探寻出其潜在的概率密度模态. 算法的步骤描述如下:

(1) 将特征空间划分为 n 个非重合的同等大小部分. 每一部分的大小为 $(G_{\max} - G_{\min})/n$, 其中 G_{\max} 和 G_{\min} 分别为图像像素灰度值的最大值和最小值. 为了避开低密度区域, 要保证每一部分中像素的数量不少于一个阈值 T_1 .

(2) 运行 Mean Shift 程序 n 次以得到 n 个收敛点, Mean Shift 程序和核半径为 $h = (G_{\max} - G_{\min})/n$.

(3) 将距离小于一个预设阈值 T_2 的相邻收敛点合并为一点, 求出 m ($m < n$) 个潜在的概率密度模态中心.

为了将图像灰度分为 m 个类, 比如 $0 \sim 255$, 考虑到图像灰度范围的上下限已确定, 需要再确定 K 个阈值, 满足 $K = m - 1$. 在下文中采用一种迭代阈值选取方法来确定 K 个阈值.

1.2 多级阈值化

为了使分割方法更具鲁棒性, 阈值应该能

够自动地被确定出来. 因此, 自动阈值选择是图像分割的一个关键步骤. Ridler and Calvard^[8]提出了一种迭代的阈值选择方法, 该方法描述如下.

给定一幅图像, 假定 $I(i, j)$ 是 (i, j) 处像素的灰度值, 其中 $1 \leq i \leq M$ and $1 \leq j \leq N, 0 \leq I(i, j) \leq L - 1, L$ 为图像的灰度级别. 首先计算出图像的灰度直方图, 假定希望能找到一个阈值 T 并利用该阈值将图像分割为一个二进制图像. 在二进制图像中, 所有灰度值低于 T 的像素, 其灰度值都被 A 所取代; 所有灰度值高于 T 的像素, 其灰度值都被 B 所取代. 将误差函数定义为二进制图像和原始图像相应像素灰度值差的平方和. 使用积分以方便表达, 则误差函数可以表示为:

$$e^2 = \int_0^T (k - A)^2 h(k) dk + \int_T^{L-1} (k - B)^2 h(k) dk, \quad (6)$$

式中, k 为像素的灰度值; $h(k)$ 为灰度值在直方图中出现的次数. 分别令 e^2 对 T, A 和 B 的偏导数为 0 可以得到:

$$T = \frac{A + B}{2}; \quad (7)$$

$$A = \frac{\int_0^T k \cdot h(k) dk}{\int_0^T h(k) dk} = \mu_1; \quad (8)$$

$$B = \frac{\int_T^{L-1} k \cdot h(k) dk}{\int_T^{L-1} h(k) dk} = \mu_2. \quad (9)$$

实际上, A 和 B 分别为直方图被阈值 T 分成的两部分的均值.

注意到阈值 T 仅由两部分的均值 A 和 B 确定, 而这两部分的均值 A 和 B 又只有当 T 确定了以后才能被计算出来. 所以需要使用迭代算法: 首先选定初始阈值做为起始点; 然后利用此阈值将直方图分为两个部分, 分别计算出两部分的均值, 并将阈值更新为两部分均值的和的一半, 该过程反复执行直到阈值收敛; 最后利用多级阈值化方法图像的灰度范围将被 K 个阈值分割为 $K + 1$ 个部分.

给定原始图像 $I(i, j)$, 利用 Mean Shift 程序确定阈值的数量 K , 令 $T_{\min} = 0$ 及 $T_{\max} = L - 1$, 利用迭代阈值选取方法计算出 K 个阈值 $\{T(i) | i = 1, 2, \dots, K\}$, 其中 $T(i) < T(i + 1)$. K 个阈值将图像直方图分割为 m 个互不重叠的区域:

$$J(i,j) = \begin{cases} 0, & T_{\min} \leq I(i,j) \leq T(1); \\ \frac{L-1}{K+1}, & T(1) < I(i,j) \leq T(2); \\ \frac{2 \cdot (L-1)}{K+1}, & T(2) < I(i,j) \leq T(3); \\ \dots & \dots \\ L-1, & T(K) < I(i,j) \leq T_{\max}, \end{cases} \quad (10)$$

其中, J 为分割图像.

2 分割质量评估

图像分割质量的评估没有一个标准的方法, 最常用的是主观评估, 这种方法主观性太强, 而非监督性评估方法由于可以克服上述缺点, 受到越来越多的关注. Haralick 等^[9] 给出了高质量图像分割结果的四条准则: ①分割出的各个区域内部像素的特性要尽量一致; ②相邻区域之间像素特性的差异要大; ③区域内部没有孔洞; ④区域边界在空间位置上要精确, 并且不能支离破碎.

前两条被用于审查图像中目标的特性, 被称为特性准则; 后两条被称为语义准则, 被用来衡量人如何将一个区域识别为一个目标^[10].

Zhang^[11] 引入期望区域熵作为区域内一致性的度量. 假定原始图像被分割成 N 个区域, 采用 R_j 来表示区域 $j(1 \leq j \leq N)$ 中的像素集合; S_j 表示区域 j 的面积, S_{jm} 表示对应于区域 j 的原始图像中具有灰度值 $m(0 \leq m \leq L-1)$ 的像素个数, 那么区域 j 的熵定义为

$$H(R_j) = - \sum_{m=0}^{L-1} \frac{S_{jm}}{S_j} \log \frac{S_{jm}}{S_j}. \quad (11)$$

而分割的期望区域熵为

$$Hr(I) = \sum_{j=1}^N \frac{S_j}{S_I} H(R_j), \quad (12)$$

式中: S_I 是图像 I 的面积, 表示图像内部具有更高的相似度.

Biswas 等^[12] 提出了用 IQI (image quality index) 来度量区域间的差异. 对于一个行数和列数分别为 M 和 N 的图像 I , 令 $I(i,j)$ 表示 (i,j) 处像素的灰度值. 则 IQI 被定义为

$$IQI = \frac{\sum_{i=1}^M \sum_{j=1}^N c(i,j)}{M \times N - \sum_{i=1}^M \sum_{j=1}^N h(i,j)}; \quad (13)$$

$$c(i,j) = \frac{|B - I(i,j)|}{B}; \quad (14)$$

$$h(i,j) = \frac{1}{8} \sum_{\substack{p=i-1, \\ p \neq i}}^{i+1} \sum_{\substack{q=j-1, \\ q \neq j}}^{j+1} \exp(-|I(i,j) - I(p,q)|). \quad (15)$$

IQI 反映的是沿着区域间边界图像像素性质的平均对比度, 值越大表明图像的平均对比度越大.

3 实验结果

选择两组复杂度不同的图像来验证该多阈值图像分割方法的有效性. 第一组图像是一幅合成图像“Gray8”; 第二组图像包含从 Berkeley Segmentation Dataset^[13] 选出来的自然图像 35 010、126 007、241 004、113 016、385 039、368 016. 笔者提出的多阈值分割方法包括 3 个参数: 核半径 h 控制着分割的敏感度; 阈值 T_1 暗示了特征空间中最小可以接受的概率密度; 阈值 T_2 对应于相邻两个收敛点的最小距离. 实验选取 $h = T_2 = (G_{\max} - G_{\min})/10$, 其中, G_{\max} 和 G_{\min} 分别为图像像素灰度值的最大值和最小值; $T_1 = 0.001 \times (M \times N)$, 其中 M 和 N 分别为图像的行列数.

图 1 和图 2 给出了两组图像的分割结果. 中间图像显示了直方图分割的过程示意图, 矩形表示整个灰度直方图中被划分成的几个互不重叠的部分, 每个矩形的中心是其 Mean Shift 程序核窗口的起始点, 每一个五角星代表了相应 Mean Shift 程序的收敛点, 如果相邻收敛点的距离小于核窗口的半径则将此相邻收敛点合并以得到模态中心, 各模态中心用黑色三角形表示, 相邻模态中间的阈值由迭代阈值选取方法选定并由短划线表示其位置.

图 1 所示的 8 个离散的灰度值被成功地检测出来(图 1(b)), 并给出了比较精确的分割结果(图 1(c)). 对于图 2 中的 2 幅自然图像, 也给出了令人满意的分割结果, 并且相对于原始图像, 分割图像的对比度提高了许多.

采用本文方法得到的图像的分割结果如图 3 所示, 从左到右分别为“Lena”“Cameraman”和“Baboon”; 其中第一行显示了原始图像, 第二行为笔者所提多阈值分割方法得到的图像分割结果, 其中合理选择分割半径后得到 $K=6$ 个分割阈值.

将本文分割方法与典型的多阈值分割方法 PSO 及 DE 算法进行了比较. 分割图像的性能参数选择前文所述的期望区域熵 Hr 和图像质量指标 IQI , 比较结果见表 1. 可以看出, 对于对比度比较明显的自然图像, 选择同样数量的阈值数是时,

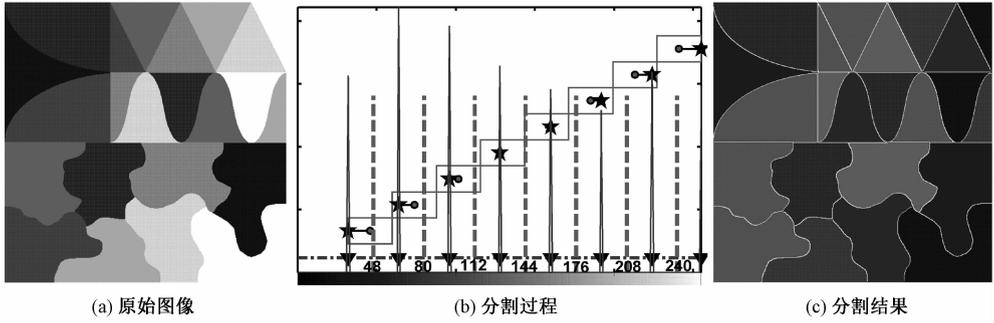


图 1 Gray8 图像分割情况

Fig. 1 Gray8 image segmentation

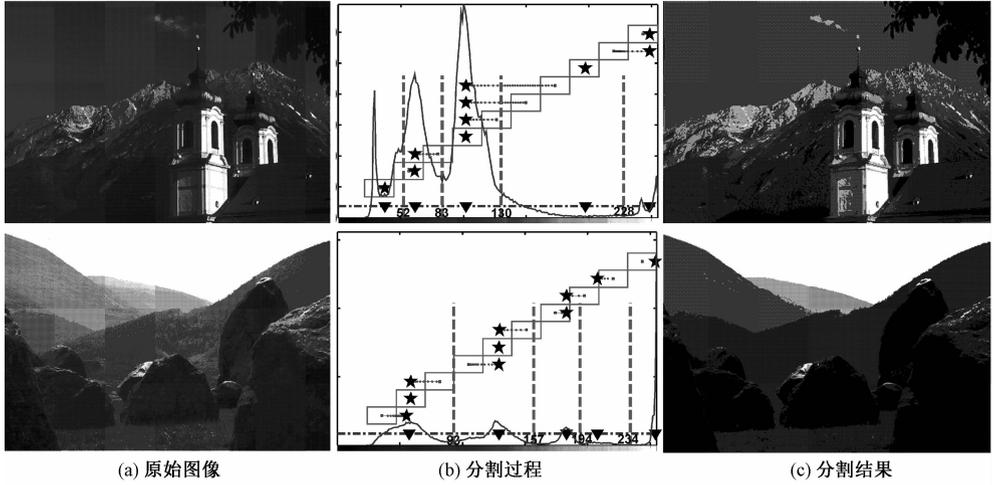


图 2 自然图像分割过程及结果

Fig. 2 Segmentative process and result of image segmentation

表 1 3 种方法的期望区域熵和图像质量指标比较

Tab. 1 Expected region entropy and image quality index for three method

<i>Hr</i>			<i>IQI</i>		
本文方法	PSO	DE	本文方法	PSO	DE
6. 258	6. 176	5. 984	1. 159	1. 073	1. 177
9. 730	9. 663	9. 274	1. 843	1. 470	1. 732
4. 911	4. 997	4. 773	0. 438	0. 457	0. 461

表 2 3 种方法的图像分割期望区域熵性能稳定性比较

Tab. 2 Expected region entropy performance stability of image segmentation for three method

测试图像	本文方法		PSO		DE	
	均值	标准差	均值	标准差	均值	标准差
Lena	6. 303	0. 618	6. 190	1. 874	6. 017	0. 976
Cameraman	9. 679	0. 590	9. 472	0. 489	9. 305	0. 732
Baboon	5. 017	1. 803	5. 135	1. 309	4. 931	2. 390
35010	11. 37	0. 210	11. 408	0. 430	11. 140	0. 790
126007	5. 943	0. 255	5. 810	2. 320	5. 798	2. 018
241004	6. 374	0. 387	6. 193	1. 358	6. 158	2. 041
113016	5. 179	0. 068	5. 104	0. 289	5. 092	1. 034
385039	7. 385	0. 189	7. 272	1. 487	7. 490	1. 376
368016	6. 941	0. 576	6. 577	1. 965	6. 782	1. 069

本文方法的性能优于 PSO 和 DE 算法. 但是对于纹理较丰富的图像如 Baboon, 本文方法稍逊于 PSO 和 DE 算法.

本文分割算法在 Berkeley Segmentation Dataset 测试数据上获得了良好的分割效果. 为了展示本文分割方法的性能, 如图 3 所示, 选取其中的几幅典型图像 35 010、126 007、241 004、113 016、385 039、368 016, 实验选取分割阈值个数为 $K = 6$.

与各群智能算法一样, 本文方法也具有随机性. 表 1 和表 2 为几种方法的最优结果. 分割算法对每幅图像重复运行 50 次后, 分别计算目标函数的均值和方差, 结果如表 2, 3 所示, 可以看出本文的图像分割算法具有良好的稳定性.

对于 $K = 6$, 将本文方法的计算时间与 PSO 和 DE 方法进行了比较, 结果如表 4 所示. 结果显示, 大部分情况下本文算法运行时间最少.

笔者采用的多阈值化方法用到了 3 个参数: 核半径 h 、阈值 T_1 和阈值 T_2 . T_1 和 T_2 对分割结果的影响要远小于 h . 一般来说, 若选取较大的 T_1 值, 则 Mean Shift 程序只能探测出最重要的模



图3 原始测试图像分割的比较

Fig. 3 The compare of original test images and segmented images

表3 3种方法的图像分割 IQI 性能稳定性比较

Tab. 3 Image quality index performance stability of image segmentation for three method

测试图像	本文方法		PSO		DE	
	均值	标准差	均值	标准差	均值	标准差
Lena	1.148	0.765	1.079	1.494	1.157	1.376
Cameraman	1.870	0.613	1.526	0.421	1.772	1.232
Baboon	0.419	1.490	0.441	1.689	0.450	1.660
35010	0.519	0.471	0.528	0.474	0.535	0.609
126007	1.280	0.598	1.197	2.501	1.206	2.103
241004	1.102	0.608	1.097	1.169	1.137	1.942
113016	0.739	0.236	0.728	0.376	0.699	1.284
385039	1.363	0.160	1.309	1.790	1.475	1.018
368016	2.031	0.376	1.947	1.483	1.962	1.215

表4 3种方法的计算时间比较

Tab. 4 Computation time for three method s

测试图像	本文方法	PSO	DE
Lena	1.485	1.871	2.138
Cameraman	1.995	2.395	2.017
Baboon	1.938	1.850	1.749
35010	2.116	2.287	2.431
126007	1.649	1.927	1.818
241004	2.017	2.251	2.514
113016	1.729	1.954	2.137
385039	1.652	1.798	1.819

态,这通常会造成分割;而若选取较小的 T_1 值,则 Mean Shift 程序还能探测出次要的模式,这通常会造成分割。实际上, T_1 选取为整个图像大小的 1% 就可以得到较好的自然图像分割结果。

而 T_2 对分割结果影响较小,通常令 T_2 等于核窗口的半径,因此我们主要讨论 h 的影响。

作为一种非参数化的密度估计方法, Mean Shift 程序的估计精度会受到核半径 h 的影响;通常较小的核半径值会使程序得到太多模式;而较大的核半径值会生成平滑的密度和较少数量的模式。最常用的方法是选择不同的核半径,将算法运行多次直到分割性能达到最优。Comaniciu^[14] 用比较复杂的方式解决了核半径的选择问题,实验显示,当核半径选为 $(G_{\max} - G_{\min})/10$ 时,就可以得到较好的分割结果,分割图像的对比度提高了许多。可以看到,随着核半径的增加, H_r 越来越小,表明随着图像被划分的区域数目的增多,各个区域内像素特性的差异越来越小; IQI 也越来越小,表明区域内像素特性的差异越来越小,尽管如此,任何一个分割结果的 IQI 都要比原始图像的 IQI 大。因此,从图像压缩的角度来看,本文方法也具有一定的参考价值。

4 结论

为了解决多级阈值化图像分割方法中阈值的数量不能预先确定这一难题,提出了一种新型的基于聚类技术的多级阈值化方法。首先使用 Mean Shift 技术搜索到潜在的模式中心点;然后应用迭代的阈值选择方法来确定相邻模式中心的各个阈值;最后采用多级阈值化方法对图像进行分割,并通过大量实验得到了验证。此外,还应用期望区域熵 H_r 和图像质量指标 IQI 对分割结果进行了评估,显示了本文方法的有效性。本文方法可以通过简单地修改程序参数来控制分割的精度,并且无

须预先确定直方图中潜在的模态数量,因此,本文方法可应用于单阈值分割、多级阈值分割和有损压缩等技术中。

参考文献:

[1] TOKAS N, KARKRA S, PANDEY M K. Comparison of digital image segmentation techniques a research review [J]. International journal of computer science and mobile computing, 2016(5):215-220.

[2] DHANACHANDRA N, MANGLAM K, CHANU Y J. Image segmentation using K-means clustering algorithm and subtractive clustering algorithm[J]. Procedia computer science, 2015,54: 764-77.

[3] Gómez D, Yáñez J, GAUDA C. Fuzzy image segmentation based upon hierarchical clustering[J]. Knowledge based system,2015, 87(c):26-37.

[4] OTSU N. A threshold selection method from grey-level histograms[J]. IEEE transaction on systems man cybernetics,2014,9(1):62-66.

[5] RAJA N S M, SUKANYA S A, NIKITA Y. Improved PSO based multi-level thresholding for cancer infected breast thermal images using otsu[J]. Procedia computer science, 2015,48:524-529.

[6] SARKAR S, DAS S, CHAUDHURI S S. Amultilevel color image thresholding scheme based on minimum

cross entropy and differential evolution [J]. Pattern recognition letters,2015,54:27-35.

[7] CHENG Y. Mean shift, mode seeking, and clustering [J]. IEEE transactions on pattern analysis and machine intelligence,2012,17(7):790-799.

[8] RIDLER T W, CALAARD S. Picture thresholding using an iterative selection method[J]. IEEE transactions on systems man & cybernetics,2010,8(8):630-632.

[9] HARALICK R, SHAPIRO L. Survey: image segmentation techniques[J]. Computer vision, graphics and image processing,2013(29):100-132.

[10] 余金煌,陶同赞. 小子域滤波在高密度电法图像处理中的应用[J]. 水电技术,2015(1):5-10.

[11] ZHANG H, FRITTS J E, GOLDMAN S A. An entropy-based objective evaluation method for image segmentation [C]//Proc SPIE Electronic Imaging,2014: 38-49.

[12] BISWAS S, LOVELL B C. Bezier and splines in image processing and machine vision[M]. London: Springer, 2009.

[13] Berkeley segmentation dataset [EB/OL]. [2017-03-01]. <http://www.cs.berkeley.edu/projects/vision/bsds>. 2012.

[14] COMANICIU D. An algorithm for data-driven bandwidth selection[J]. IEEE transactions on pattern analysis and machine intelligence,2013,25(2):281-288.

Based on the Mean Shift Clustering Multilevel Threshold Method

CHENG Quan¹, LIU Xiaoqing¹, LIU Yuchun¹, WANG Zhiliang²

(1. Mechanical and Electrical Engineering, Zhoukou Normal University, Zhoukou 466001, China; 2. School of Computer & Communication Engineering, Beijing University of Science & Technology, Beijing 100083, China)

Abstract: In order to solve the problem that the number of selected thresholds in multilevel thresholds cannot be usually predetermined, a novel multi-level thresholding method based on Mean Shift Clustering technique was proposed. Using Mean Shift technology to explore the potential mode center, the various thresholds of adjacent to the mode center was automatically determined by using of iterative threshold selection method, and then the method of multi-level threshold was used for image segmentation. The experimental results showed that, relative to the original image, contrast of the image split with Mean Shift clustering technique was greatly improved. This method could control the segmentation precision flexibly by simply modifying parameters of the program, and could be widely used in the technology of single threshold segmentation, multi-level threshold segmentation and detrimental compression and other technologies.

Key words: multilevel thresholding; image segmentation; iterative threshold selection; segmentation evaluation