

基于 SLIC 的图像分割方法研究

刘 洋,周宁宁

(南京邮电大学 计算机学院,江苏 南京 210046)

摘 要:图像分割是将图像分割成若干个特定的、独特的区域,并提取出目标的技术和过程。它是计算机视觉和图像处理中重要的研究内容之一。近年来,以 GrabCut 为代表的基于图论的图像分割方法成为了研究热点。基于图论的图像分割是一种将图像边缘信息和纹理信息相结合的能量最小化方法。GrabCut 算法是一种优秀的基于图论的交互式图像分割算法。该算法具有简单的交互性和良好的分割效果,但也存在分割速度较慢、算法时间复杂度过高以及对背景和前景相似的图片分割不理想等缺陷。超像素块提取是指将图像分割成一组集合,集合的每个超像素块里的像素点都有相似的特征,各个超像素块之间有着明显的区别。超像素有许多便捷的属性,并被广泛地用作基础预处理阶段,以减少各种计算机视觉任务的计算。针对 GrabCut 图像分割方法时间复杂度过高以及对背景和前景相似的图片处理质量不好的问题,提出了基于超像素的 GrabCut 图像分割方法。

关键词:GrabCut;图像分割;超像素;SLIC;K-means

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2019)01-0075-05

doi:10.3969/j.issn.1673-629X.2019.01.016

Research on Image Segmentation Method Based on SLIC

LIU Yang, ZHOU Ning-ning

(School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210046, China)

Abstract: Image segmentation is the technology and process of dividing the image into several specific and unique regions and extracting the target, which is one of the important research contents in computer vision and image processing. In recent years, image segmentation based on graph theory, represented by GrabCut, has become a research hotspot. Image segmentation based on graph theory is an energy minimization method that combines image edge information and texture information. The GrabCut is an excellent interactive image segmentation algorithm based on graph theory. The algorithm has simple interactivity and great segmentation effect, but there are some defects such as slow segmentation, high algorithm time complexity and poor segmentation of similar background and foreground images. Superpixel block extraction refers to the segmentation of the image into a set of collections, where each pixel in each superpixel block has similar characteristics, and there are obvious differences between the superpixel blocks. Superpixels have many convenient properties and are widely used as a basic pre-processing stage to reduce the computation of various computer vision tasks. In view of the high time complexity of GrabCut image segmentation method and the poor quality of image processing with similar background and foreground, we propose a method of GrabCut image segmentation based on hyperpixel.

Key words: GrabCut; image segmentation; superpixel; SLIC; K-means

1 概 述

图像分割是计算机视觉领域的关键技术之一,是一个提取图像中感兴趣或者有意义的像素集合和特征的过程。图像的特征有很多种,可以是像素的灰度值、颜色和纹理等图像原有的特征,也可以是直方图等经过处理后的非图像原有的特征。图像分割的目标就是把一幅图像分成几个子区域,这些子区域的划分是基

于图像的某个特征的,而且相同区域内某个特征很相似,不同区域内的某个特征有着显著区别。图像分割是图像分析以及图像处理必不可少的步骤,在某种意义上,很多后续的图像处理的质量取决于图像分割的质量。

近年来,新的数学理论和研究不断出现,每个数学理论都被学者们进行二次研究,并应用到图像分割领

收稿日期:2018-01-20

修回日期:2018-05-23

网络出版时间:2018-09-21

基金项目:国家自然科学基金(61170322,61373065,61302157);智能电网,国家重点实验室(NO.9)

作者简介:刘 洋(1993-),男,硕士,研究方向为图像处理;周宁宁,博士,副教授,通讯作者,研究方向为图像处理等。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20180920.1535.012.html>

域,所以有许多结合了一些具体的理论、方法和工具的图像分割技术被提出^[1]。由于图像分割技术不存在一般性的理论,因此每当数学界有新的理论和方法问世时,学者们就会把它用于图像分割领域,因此一直都有新的独特图像分割算法被提出。

基于区域的分割方法是根据相似性准则,将图像划分为不同的区域^[2]。一组像素集合可以代表相应的增长区域,遍历种子点,并把它归于符合的种子像素集合,直到所有的像素被归并完成^[2]。分水岭分割方法是一种基于数学形态学和拓扑理论的分割方法,其基本思想是利用类比图像处理的地形测绘,将图像的每个像素灰度值代表点的高程,每一个局部最小值和影响的地区称为集水盆地,盆地的边界形成了一个分水岭。之所以叫分水岭算法,是因为它与水势淹没的过程比较相似,首先淹没的是图像的最低处,然后整幅图像被慢慢淹没,当完全淹没图像时,它会溢出,然后在溢水区域建成。重复这一过程,直到整个图像的所有点被淹没,然后建立一系列水坝,就变成了每一个流域的分水岭。分水岭算法是一种经典而有效的分割方法,但受噪声的影响较大^[3]。

模糊集理论和系统是近年来一个非常流行的研究领域,该方法可以提高模式识别的精确率,可以解决一系列不确定问题,比如不同程度的信息不完整、不精确和模糊等,已成为图像分割的一个重要的数学工具^[4]。

20 世纪 80 年代,人工智能概念被提出,近年来发展迅猛,在图像分割领域也有其用武之地,因此出现了基于神经网络的图像分割技术。利用神经网络进行图像分割,主要是利用神经网络的训练数据确定图像分割所需要的数据,比如节点之间边的权值,然后输入训练完成的数据对图像进行重新分割^[5]。

基于图论的图像分割也是热点研究课题^[6],并且发展迅速,已经有了很多成熟的算法。在该算法中,图像用带权的无向图表示,图像的像素点用无向图的节点表示,将无向图边的权值视为图像像素间的相似度,通过剪切无向图的边,得到一个图割,当一个图割的边集合权重之和最小时,边称为一个最小割,这就是图割的最终结果。换句话说,基于图论的图像分割是一个最优化问题,不断优化经过分割后的边集合的权值来达到一个最小割。基于图论的分割算法首先将图像映射为带权无向图 $G = (V, E)$, 图像中的每个像素点和相邻像素点之间的连接分别用无向图的节点 $N \in V$ 和无向图的每条边 $e \in E$ 表示,边的权值表示相邻像素在灰度、纹理等特征上的相似程度。而图的一个剪切可以用对图的一个分割 s 来表示,被分割的每个区域 $C \in S$ 对应着图中的一个子图。而分割的最佳原则是保持在同一子图之内的相似性最大,而不同子图之

间的相似性保持最小。除去不需要的边,划分成几个子图的无向图,实现分割,这是基于图论图像分割算法的核心和本质。GrabCut 是一种基于图论的分割方法^[7-9],针对 GrabCut 图像分割方法时间复杂度过高以及对背景和前景相似的图片处理质量不好的问题,文中提出一种基于超像素的 GrabCut 图像分割方法。该算法首先对图像进行超像素块的提取,然后对提取后的图像进行 GrabCut 图像分割。

2 GrabCut 算法简介

GrabCut 算法是迭代的 GraphCut 算法。GrabCut 算法是由微软研究院研究的一种经典图割算法,具有良好的分割效果和简单的操作性,因此得到了广泛的应用。GrabCut 是非完全标记的,通过这种标记,初始先用矩形框框出前景区域,那么矩形框外面的区域便是背景区域,然后对其颜色空间建立 GMM,用 GMM 进行迭代来代替之前的估计以完成能量的最小化。

3 基于超像素的 GrabCut 图像分割主要思想

3.1 超像素

超像素为计算图像特征提供了一种便捷的方式^[10-12]。通过获得图像中的冗余来大幅地减轻后续图像处理的复杂度。超像素块集合提取的结果就是根据图像的某个特征,灰度图像可以是灰度值,彩色图像可以是 RGB 颜色通道的欧氏距离,把原始图像分成一组不同的小区域。这些区域内部是基于某个图像特征相似的,相邻区域中的像素点是基于某个特征差异比较大的。计算基于超像素的图像特征比像素更有效。因此,在基于超像素的图像处理任务中,图像原语和冗余大大减少。一般来说,图像超像素分割方法通常用于计算机视觉中的预处理步骤,以提高它们的效率和性能。

在超像素生成方法的大多数论文中,现有的方法基于其基本思想分为两类:基于图论的方法和基于 k-均值的方法。基于图论的超像素生成模型以图像像素为节点,相邻像素之间的相似性作为边权值。超像素是通过最小化图上定义的能量函数来生成的。

基于 k-均值的方法通过对集群中心的粗略初始化和对集群的优化来生成超像素,直到满足一些收敛条件为止。尽管这些现有方法的解决方案是不同的,但它们的目标是相同的。一种理想的超像素生成方法应该是高效和简单的,并且生成的超像素应该具有许多特定的属性。第一,边缘坚持属性,生成的超像素的边缘应尽可能地与图像的对象边界保持一致;第二,同

质性,一个超像素的像素应该是均匀的;第三,紧凑性,产生的超像素集合应该是紧凑适当的。

3.2 算法流程

文中算法是先载入处理的图像,然后用 SLIC 算法提取超像素块,对提取后的超像素进行预处理,最后用 GrabCut 进行分割。具体流程如图 1 所示。

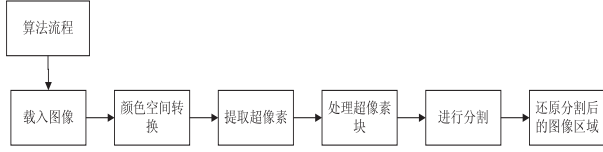


图 1 算法流程

算法步骤如下:

(1) 载入要处理的图像。

(2) 将图像转换为 Lab 颜色空间。

因为 SLIC 算法图像信息是取自于 CIELAB 空间而不是 RGB 空间,这是在 5 维中进行的, $[lab_{xy}]$ 。 $[lab]$ 是指 CIELAB 颜色空间中的像素颜色向量。CIELAB 颜色空间就是指 HSV 空间,亮度空间, xy 指像素的位置。两种颜色的最大的空间是有限的,空间距离依赖于图片大小。

(3) 用 SLIC 算法提取超像素块。

SLIC 生成超像素是通过基于颜色相似度和临近度的像素聚类^[13-16]。SLIC 提供的方法,虽然简单,但它解决了这些问题,并且产生了高质量、紧凑、几乎一致的超级像素。简单的线性迭代聚类(SLIC)在由 L , a , b 值和 x , y 像素坐标定义的 5-d 空间中,执行一个局部的像素集群。SLIC 采用一种距离度量,在超像素形状中实现了紧凑性和规律性,并无缝地适应灰度和彩色图像。SLIC 很容易实现,并且与竞争的方法相比,SLIC 的效率要高得多。对于许多视觉任务,紧凑且高度一致的超像素是尊重图像边界的,当从基于像素的图形切换到超像素时,可以看到显著的速度增加,但是不规则的超像素会使分割效果变差,在超像素位置上从图像中提取的灰度值等局部特征变得不那么有意义。与现有的其他方法相比,SLIC 以较低的计算成本得到类似或更大的性能。其算法思想是,首先将图像从 RGB 颜色空间转换到 Lab 颜色空间,对应每个像素的 (L, a, b) 颜色值和 (x, y) 坐标组成一个 5 维向量 $\mathbf{V} [L, a, b, x, y]$,两个像素的相似性即可由它们的向量距离来度量,相似性与它们之间的向量距离成反比。

SLIC 是基于图像的颜色相似性和在图像平面上的距离来生成超像素。这是在五维的 Lab_{xy} 空间中完成的。Lab 颜色空间中的像素颜色空间,被认为是对小颜色距离的感知一致,而 xy 是像素位置。在这个 5D 空间中,不可能简单地使用欧氏距离,而不使空间距离标准化。数据在这个 5D 空间中聚集像素,引入

了一种考虑超像素大小的新距离度量。使用它,在这 5D 空间中加强了颜色的相似性和像素的接近性,这样预期的集群大小和它们的空间范围几乎是相等的。在 Lab 颜色空间中的欧氏距离对于短距离的感知是有意义的。如果空间像素的距离超过了这个感知颜色距离限制,那么它们就开始超过像素颜色的相似性(产生的超像素不尊重区域边界,只在图像平面上接近)。因此,不用在 5D 空间中使用简单的欧氏规范,而是使用一种距离尺度来定义:

$$\begin{cases} d_{lab} = [(L_k - L_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2]^{0.5} \\ d_{xy} = [(X_i - X_k)^2 + (Y_k - Y_i)^2]^{0.5} \\ D_s = d_{lab} + (m/S) * d_{xy} \end{cases} \quad (1)$$

其中, m 用来调整 d_{xy} 的权值,一般为 1 ~ 20,在算法中设置为 10。

(4) 对提取后的超像素进行预处理。

为了方便后续的分割操作,需要对提取后的超像素块进行一些处理。首先载入提取成超像素块的图像,然后将每个超像素块按照左上角的颜色将整个超像素块统一分配颜色。

(5) 用 GrabCut 进行分割。

设原始图像表示为 $\mathbf{Z} = \{Z_1, Z_2, \dots, Z_N\}$, 使用超像素方法预处理得到的图像表示为 $\mathbf{S} = \{S_1, S_2, \dots, S_N\}$, 其中 $S_i = \frac{1}{m} \sum_{(Z_j \in \mathbf{Z}, Z_i \in M)} Z_j$, m 是当前超像素块所含的像素点个数, M 是当前超像素中像素点的集合。对图像建立 GMM 模型,图割的过程也是一个像素标记的过程,这个标记在算法中用值表示,如果为 1,则表示对应的像素是前景,反之则为背景。 \mathbf{K} 向量代表 GMM 的 k 个高斯分量,反映每一个像属于哪个高斯分量, $k = \{k_1, \dots, k_n, \dots, k_N\}$, $k_n \in \{1, 2, \dots, K\}$ 。图像中的像素都有一个 k 与之对应,从而根据 k 的值,像素要不属于前景像素,就属于背景像素。通常 GrabCut 算法的 Gibbs 能量函数形式^[7]为:

$$E(\alpha, k, \theta, S) = U(\alpha, k, \theta, S) + V(\alpha, S) \quad (2)$$

数据项 U 定义为:

$$U(\alpha, k, \theta, S) = \sum_n D(\alpha, k, \theta, S) \quad (3)$$

其中

$$D(\alpha_n, k_n, \theta_n, S_n) = -\log p(S_n | \alpha_n, k_n, \theta) - \log \pi(\alpha_n, K_n) \quad (4)$$

其中, $p()$ 是高斯概率分布; $\pi()$ 是混合权重系数。

于是有:

$$D(\alpha_n, k_n, \theta_n, S_n) = 0.5 \log \det \sum (a_n, k_n) + 0.5 [S_n - \mu(a_n, K_n) T] \sum (a_n, k_n)$$

$$[S_n - \mu(a_n, k_n)] - \log \pi(a_n, k_n) \tag{5}$$

那么,参数 θ 可表示为:

$$\theta = \{ \pi(\alpha, k), \mu(\alpha, k), \sum(\alpha, k), \alpha = 0, 1, k = 1, 2, \dots, K \} \tag{6}$$

平滑项 V 可用 RGB 空间的欧几里德距离求出:

$$V(\alpha, S) = \gamma \sum_{(m, n) \in C} [\alpha_n \neq \alpha_m] \exp - \beta \| S_m - S_n \|^2 \tag{7}$$

其中, C 表示相邻像素对。

具体的实现步骤如下:

首先,用户在原始图像上标注一个包含目标物的矩形框,算法以“非完全标号”方式来标定图像的背景区域(T_B)和未知区域(T_U),矩形框外的区域为 T_B :

$$\begin{cases} T_B = \varphi \\ T_U = \bar{T}_B \end{cases} \tag{8}$$

然后,根据标定的背景区域和前景区域初始化 GMM,对 T_U 迭代进行直至收敛:

①求 T_U 中每个像素 n 的 GMM 参数 k_n : $k_n = \underset{k_n}{\operatorname{argmin}} D_n(\alpha_n, k_n, \theta, S_n)$ 。

②从图像数据 S 中学习 GMM 参数: $\theta = \underset{\theta}{\operatorname{argmin}} U(\alpha, k, \theta, S)$ 。

③为了参数估计而进行的分割:用最小割解 $\min_{\{\alpha_n; n \in T_U\}} \min_k E(\alpha, k, \theta, S)$ 。

最后,根据所得到的参数对图像构造网络,并用最大流算法进行分割。分割完成后,将预处理后的图像中的超像素块按照坐标还原成原图。

(6)算法执行完毕。

4 实验结果与分析

实验统一在 Win7 64 位操作系统上进行,处理器 Intel I5,内存 4 GB。实验采用 C++语言,在 VS2015 平台上进行。

4.1 实验 1

实验 1 的目的是验证文中算法能够提高 GrabCut 算法的实现速度。

具体的实验流程如图 2 所示。

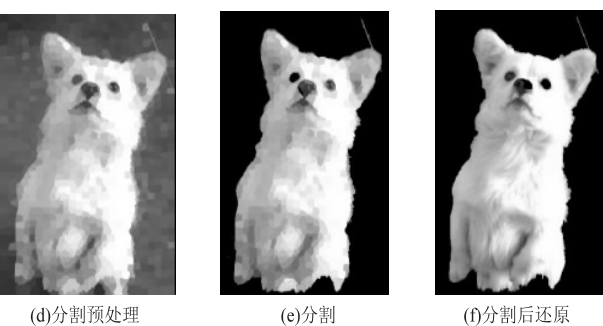


图 2 实验流程

原始的 GrabCut 算法和文中算法的分割效果对比如图 3 和图 4 所示。

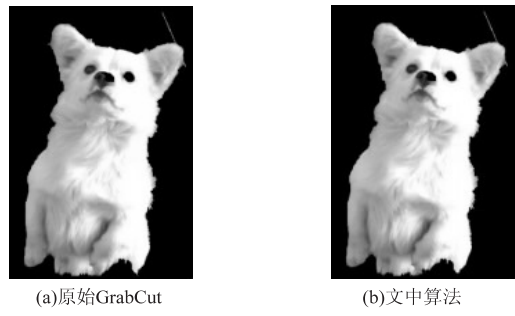


图 3 分割效果对比(1)

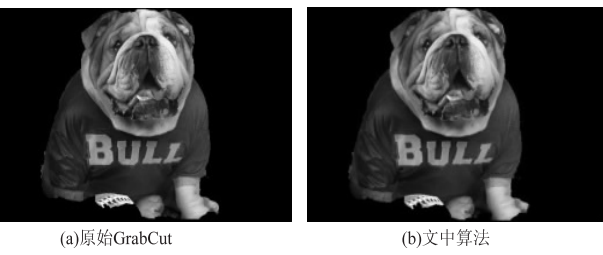


图 4 分割效果对比(2)

其中图 3 是 41 k 的图像,图 4 是 100 k 的图像。

表 1 是原始 GrabCut 算法与使用超像素处理后的 GrabCut 算法运行时间的对比。

表 1 运行时间对比				s
图像	原始 GrabCut	1 000 个 超像素块	500 个 超像素块	200 个 超像素块
41 k 图像	4.1	3.3	2.7	2.1
100 k 图像	7.6	5.7	5.0	3.7

通过对比可知,当处理前景和背景相似的图像时,采用超像素预处理后的 GrabCut 图像分割算法的分割效果和原始的 GrabCut 算法差不多,几乎一致,但是运行速度明显快。并且当设定的超像素块越少时,运行速度越快,这也和常理吻合。但是,过少的超像素块会使分割后的图像失真,效果变差。

4.2 实验 2

实验 2 是为了检验当处理前景和背景相似的图片时,文中算法对原始的 GrabCut 图像分割算法是否有所改进。

原始的 GrabCut 算法和文中算法的对比如图 5 所示。

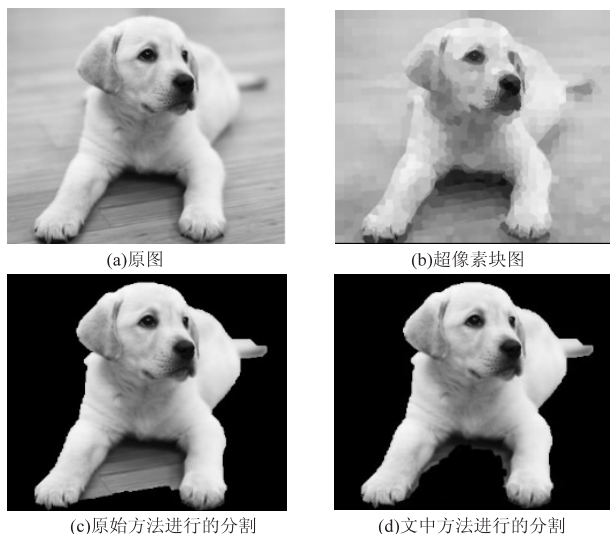


图 5 GrabCut 算法和文中算法的对比

由图 5 可以看出,在处理前景和背景相似的图片时,文中算法比原始的 GrabCut 算法有着较为明显的改进。因为把图片进行超像素块的提取,把每个超像素设为同一个色块,这样处理后,前景和背景之间的特征差距增大,相似度减弱,所以才能提高分割效果。

5 结束语

针对 GrabCut 图像分割算法运行速度慢,以及对于前景背景相似的图片分割效果较差的不足,提出了一种改进方法。实验结果表明,通过对图像进行一些预处理,即对图像进行超像素块的采集,然后对采集后的图像再进行分割,会有效提高算法的分割速度,并在一定程度上改善 GrabCut 图像分割方法对于前景背景相似图像分割效果较差的不足,是一种可行的方法。但是,该方法依然没有完美地解决这一不足,在以后的研究中会继续致力于这个问题。

参考文献:

- [1] SILLA JR C N, FREITAS A A. A survey of hierarchical classification across different application domains[J]. Data Mining and Knowledge Discovery, 2011, 22(1-2): 31-72.
- [2] VICTOR O R, JUAN I G L, NICOLAS S L, et al. An improved watershed algorithm based on efficient computation of shortest paths[J]. Pattern Recognition, 2007, 40(3): 1078-1090.
- [3] PARVATI K, RAO B S P, DAS M M. Image segmentation using gray-scale morphology and marker-controlled watershed transformation[J]. Discrete Dynamics in Nature and Society, 2008, 2008: 384346.
- [4] MASOOLEH M G, MOOSAVI S A S. An improved fuzzy algorithm for image segmentation[J]. Proceedings of World Academy of Science, Engineering and Technology, 2008, 28(4): 400-404.
- [5] ZHANG Junying, FAN Xiuju, DONG Jiyang, et al. Image segmentation based on modified pulse-coupled neural networks[J]. Chinese Journal of Electronics, 2007, 16(1): 119-122.
- [6] 侯 叶. 基于图论的图像分割技术研究[D]. 西安: 西安电子科技大学, 2011.
- [7] TANG Meng, GORELICK L, VEKSLER O, et al. GrabCut in one cut[C]//IEEE international conference on computer vision. Sydney, NSW, Australia: IEEE, 2013: 1769-1776.
- [8] TANG Meng, AYED I B, MARIN D, et al. Secrets of GrabCut and kernel k-means[C]//IEEE international conference on computer vision. Santiago, Chile: IEEE, 2015: 1555-1563.
- [9] JI Zhihua, YU N, WU Shibin, et al. Improved GrabCut for human brain computerized tomography image segmentation[C]//International conference on health information science. [s. l.]: [s. n.], 2016: 22-30.
- [10] VEKSLER O, BOYKOV Y, MEHRANI P. Superpixels and supervoxels in an energy optimization framework[C]//Proceedings of the 11th European conference on computer vision. Heraklion, Crete, Greece: Springer, 2013: 211-224.
- [11] LIU Mingyu, TUZEL O, RAMALINGAM S, et al. Entropy rate superpixel segmentation[C]//IEEE conference on computer vision and pattern recognition. Colorado Springs, CO, USA: IEEE, 2011: 124-156.
- [12] ZHANG Yuhang, HARTLEY R, MASHFORD J, et al. Superpixels via pseudo-Boolean optimization[C]//IEEE international conference on computer vision. Barcelona, Spain: IEEE, 2011: 1387-1394.
- [13] 南柄飞, 穆志纯. 基于 SLIC 融合纹理信息的超像素分割方法[J]. 仪器仪表学报, 2014, 35(3): 527-534.
- [14] 张亚亚, 刘小伟, 刘福太, 等. 基于改进 SLIC 方法的彩色图像分割[J]. 计算机工程, 2015, 41(4): 205-209.
- [15] 周良芬, 何建农. 基于 GrabCut 改进的图像分割算法[J]. 计算机应用, 2013, 33(1): 49-52.
- [16] 伊聪聪, 吴 斌, 张红英. 一种改进的 Grabcut 图像分割方法[J]. 小型微型计算机系统, 2014, 35(5): 1164-1168.