

一种新的基于区域和边界的图象分割方法

杜啸晓 杨新 施鹏飞

(上海交通大学 图象处理与模式识别研究所 上海 200030)

摘要 用分水岭方法进行图象分割时,容易造成图象的过度分割,从而使得物体的轮廓线掩埋在杂乱的分水岭线中,为了克服这种过度分割问题,提出了一种保持边界的非线性扩散方法,该方法就是先对图象进行平滑,在去噪声的同时,即减少了梯度图象中区域最小值的数目,而分水岭分割后的图象区域数目与它相同;然后,根据初始分水岭分割的结果,使用区域灰度相似性和边界强度相结合的准则,进行由底向上的层次融合,从而较好地解决了过度分割的问题.实验结果表明,该方法可提供精确且封闭的区域轮廓线.

关键词 分水岭 图象分割 图象平滑 偏微分方程

中图法分类号: TP391.41 **文献标识码:** A **文章编号:** 1006-8961(2001)08-0755-05

A New Image Segmentation Method Based on Region and Edge

DU Xiao-xiao, YANG Xin, SHI Peng-fei

(Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong Univ., Shanghai 200030)

Image segmentation based on watershed method always results in over-segmentation and makes the contours of the objects buried in the irrelevant watershed lines. In this paper, we first smooth the image while preserving the edge by nonlinear diffusion method, which can reduce the noise in the image and at the same time the numbers of the region minimums of the gradient image that is equal to the region numbers of watershed segmentation. Then from the result of the initial watershed segmentation which is organized by the Region Adjacency Graph (RAG), we execute the bottom-up hierarchical region merging according to the region average gray value similarity and the edge strength criterions that can settle the over-segmentation problem well. The region average gray value is a rough characteristic about the region while the edge strength criterion is local. Similarity criterion of the average gray value of the region is first used in Merging operation. Then upcast the RAG and continue to merge according to edge strength. The results of the experiment using 2D real images show that this method can provide accurately localized and closed region contours.

Keywords Watershed, Image segmentation, Image smoothing, Partial differential equations

0 引言

图象处理中,图象分割是一个非常关键的过程,因为通过分割,将一幅图象划分成 N 个不同的区域,从而可为后继的分析,如识别或基于区域的压缩等提供基础.一般来说,图象的分割方法可分为两种:一是基于边缘检测的方法,如Canny算法^[1],虽然它是一种最优的线性边缘检测算子,但是由于用基于边缘检

测的方法所得到的轮廓线不能保证是封闭的,因此需要进行边缘跟踪处理,这将使得边缘定位不精确;另一种是基于区域的方法,如金字塔方法^[2],该方法中的区域一般事先给定,如正方形区域,但由于物体的轮廓线往往是任意形状的,因此,这种基于固定区域形状的方法,其分割速度较慢.使用分水岭方法进行图象分割之所以引起人们重视^[3-4]的原因,一是其计算速度较快,二是物体轮廓线的封闭性和定位的精确性.但是它存在过度分割的问题.为了解决这个问题,

基金项目:国家自然科学基金项目(No. 69931010)

收稿日期:2000-08-18;改回日期:2000-11-20

本文介绍了一种基于边和区域的混合方法,提出了区域之间的边界强度准则.

1 保持边界的非线性扩散

在图象分割中,分水岭变换使用的是梯度图象,由于梯度图象易受噪声的干扰,使得梯度图象中的区域最小值过多,而分水岭变换后得到的图象区域数目与梯度图象的区域最小值数目相同(这是形成过度分割的根本原因),因此需要对原始图象进行平滑.本文采用非线性扩散方法^[6]来进行平滑.其数学模型为:

$$\frac{\partial u}{\partial t} - g(|DG_s * u|)((1 - h(|Du|))\Delta u + h(|Du|)|Du| \operatorname{div} \frac{Du}{|Du|}) = 0 \quad (1)$$

式中, $u(x, y, 0) = u_0(x, y)$; $t \in [0, T]$; $u_0(x, y)$ 为初始的灰度图象; $u(x, y, t)$ 为在时间 t 时的平滑图象; $*$ 表示卷积运算; G_s 为高斯平滑核; $|DG_s * u|$ 为梯度 $|Du|$ 的局部估计; $g(\cdot)$ 为非增实函数,且 $s \rightarrow \infty$ 时, $g(\cdot)$ 趋向于 0; Δu 为拉普拉斯算子. 式(1)中的因子 $|Du| \operatorname{div} \frac{Du}{|Du|}$ 为退化扩散项,它表示只在与梯度方向正交的方向上进行扩散,而在梯度方向上不进行扩散,这就保证了图象边缘不被模糊. $DG_s * u$ 表示梯度图象的高斯平滑,因子 $g(|DG_s * u|)$ 用于边缘的增强,它表示扩散速度,如果 $DG_s * u$ 的绝对值较小,那么由于 $g(\cdot)$ 为非增函数,因此此处的扩散速度 $g(|DG_s * u|)$ 相对较强,反之,在图象的边缘点上, $|DG_s * u|$ 的值相对较大,即此处扩散速度相对较小. 由此可见,退化扩散模型的解使得图象边缘得以保持,而灰度变化不大的地方更加平滑. $h(\cdot)$ 为光滑非降函数,当 $s > E_0$ 时, $h(s) = 1$, 此时模型为退化扩散方程,它在与梯度正交的方向上进行平滑; 当时 $s < E_0$ 时, $h(s) = 0$, 此时模型为热扩散方程,它在各个方向上进行相同的扩散,这对于去除孤立噪声是很有作用的,这里, E_0 为预设定的门限. 上述模型在角点处易造成模糊,从而使得角点处的定位不准确. 本文通过对扩散速度加以限制,较好地解决了上述问题. 具体来说,就是与 $DG_s * u$ 正交的方向上计算二阶导数,当它大于某门限 T 时,令扩散速度趋于 0. 这样就可使得平滑后的图象中角点定位准确.

2 分水岭变换及分割方法

将一幅图象视为地形曲面,其灰度值对应地形高度. 图象 u 中,连通区域 M 具有的区域最小值 h 定义如下^[4]: 设 M 的邻域为 $N(M)$, 对于 $(x, y) \in M$, $u(x, y) = h$; 对于 $(x, y) \in N(M)$, $u(x, y) > h$. 大家知道,分水岭变换是由浸水过程来模拟的. 设由图象 u 描述的地形曲面为 S , 在 S 的每个区域最小值处刺一小孔,并将其垂直浸入水中,水则从区域最小值处的孔中进入曲面 S . 这样从不同的区域最小值处进入的水将逐步填满各自对应的区域. 为避免不同区域的水相互融合,在它们之间垒起水坝. 随着水的相继浸入,最后,只有水坝露出水面,这些水坝即构成了图象的分水岭线,水坝所围的区域称为贮水盆地. 每个贮水盆地对应一个区域最小值.

在区域 A 中, p_1, p_2 两点之间的测地距离 $d_A(p_1, p_2)$ 定义为:

$$d_A(p_1, p_2) = \inf\{l(P), P \text{ 为点 } p_1, p_2 \text{ 之间的路径, 且此路径上的所有点属于 } A, l(P) \text{ 为路径长度.}\}$$

令 $B = \bigcup B_i \subset A$, B_i 为 B 内的连通部分, 则 B_i 在 A 内的测地影响区 $iZ_A(B_i)$ 定义为

$$iZ_A(B_i) = \{p \in A, \forall j \in [1, k] / \{i\}, d_A(p, B_i) < d_A(p, B_j)\}$$

即由 A 中符合这些条件的点组成的集合,这些点与 B_i 之间的测地距离小于与 B 中其他连通部分之间的测地距离. 为方便叙述,记 $C(M)$ 为与区域最小值 M 对应的贮水盆地. 设

$$\begin{aligned} T_h(u) &= \{p, u(p) \leq h\} \\ C_h(M) &= \{p \in C(M) | u(p) \leq h\} \\ &= C(M) \cap T_h(u) \end{aligned}$$

式中, $u(p)$ 为点 p 的灰度值, $T_h(u)$ 为由图象 u 中灰度值小于、等于 h 的点组成的集合; $C_h(M)$ 表示与 M 对应的贮水盆地中那些灰度值小于、等于 h 的点组成的集合. 这样区域 A 中,测地影响区的骨架为区域 A 中不属于测地影响区的点集,即

$$SKIZ_A = A / iZ_A(B), iZ_A(B) = \bigcup_{i \in [1, k]} iZ_A(B_i)$$

上述浸水过程的模拟是一个循环过程: 令循环的初始集合为 $X_{\lambda_{\min}} = T_{\lambda_{\min}}(u)$, 即图象 u 中高度为最小高度 h_{\min} 的那些点的集合. 显然 $X_{\lambda_{\min}} \subseteq T_{\lambda_{\min+1}}(u)$, 设 Y 为 $T_{\lambda_{\min+1}}(u)$ 的一连通部分, 则 Y 和 $Y \cap X_{\lambda_{\min}}$ 具有如下 3 种包含关系^[5]:

(1) $Y \cap X_{h_{\min}} = \emptyset$, 此时, Y 为 u 的新的最小值。

(2) $Y \cap X_{h_{\min}} \neq \emptyset$, 且连通. 此时, Y 属于与区域极小值 $Y \cap X_{h_{\min}}$ 对应的贮水盆地, $Y \cap C_{h_{\min+1}} (Y = X_{h_{\min}})$.

(3) $Y \cap X_{h_{\min}} \neq \emptyset$, 且不连通, 则区域 Y 包含图象 u 中不同的区域最小值, 记为 $Z_1, Z_2 \dots Z_k$, 设其中一个为 Z_i , 则 $C_{h_{\min+1}}(Z_i) = iZ_Y(Z_i)$, 即由 Z_i 形成的、高度小于 $h_{\min+1}$ 的贮水盆地元素集合等于区域 Y 中 Z_i 的影响区元素的集合。

综合上述关系, 则下次循环的初始集合为: $X_{h_{\min+1}} = \min_{h_{\min+1}} \cup iZ_{T_{h_{\min+1}}}(X_{h_{\min}})$. 重复上述循环过程, 直到图象的最高高度 h_{\max} , 于是贮水盆地为 $X_{h_{\max}}$, 而分水岭则为它的补集。

本文采用的分水岭算法包含排序和浸水两个过程, 其中排序过程按照象素的灰度进行排序, 以便直接访问同一高度的象素; 在浸水过程中, 假设高度小于等于 h 的象素的贮水盆地已经标记出来了, 则在处理高度为 $h+1$ 的象素时, 将这一层中与已标记的贮水盆地相邻的象素送入先入先出(FIFO)队列, 再由这些象素开始, 根据测地距离, 将已经标记的贮水盆地扩展至 $h+1$ 层; 然后, 通过扫描, 将 $h+1$ 层中可能存在的区域最小值赋予新的标号; 最后, 在分水岭变换的结果中, 同一标号的象素属于同一贮水盆地, 而将距不同贮水盆地距离相等的象素标为分水岭点。

图象分割的目的是将图象分割成不同的物体, 即提取出物体的轮廓线. 由于物体的轮廓线是灰度变化比较强烈的图象部分, 而梯度图象正好描述了每个象素点的灰度变化情况, 因此, 可使用梯度图象的分水岭变换来进行图象的分割。

3 层次区域融合方法

图象经过保边界平滑后, 即可求梯度图象, 再经梯度图象的分水岭变换得到初始分割结果. 这样虽可在一定程度上减轻过度分割, 但对复杂图象来说, 由于上述过程依然存在过度分割的情况, 因此还需要在上述分水岭变换的基础上, 进行区域的融合, 以便得到有意义的分割结果。

区域融合是在整幅图象划分成 n 个区域的基础上进行的. 在分水岭变换中, 除了 n 个区域的标号外, 还存在分水岭点, 因此, 为了对图象进行 n 个区域划分, 必须将这些分水岭点分配至其相邻的区域中. 分

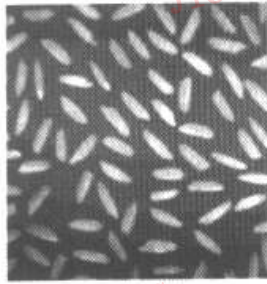
水岭点的分配准则是灰度相似性, 即在分水岭点的领域中, 寻找与其灰度最接近的象素, 且此象素不是分水岭点, 然后将分水岭点的标号设置为此象素的标号. 这样处理后的图象又称为镶嵌板图象; 最后通过对镶嵌板图象进行扫描, 来获取区域邻接图数据结构 RAG. 这种划分为 n 个区域的图象, 其 RAG 为一无向图 $G=(V, E)$, 其中, V 表示区域节点, $V=\{1, 2, \dots, n\}$; E 表示区域节点之间的边, $E \subset V \times V$.

关于区域融合的准则, 在文献[7]中, 采用了区域灰度相似性和最小描述长度准则. 正如该文献中所述的, 最小描述长度准则虽然是与参数无关的, 但过度分割的情况依然存在, 而区域灰度相似性准则是基于这样的假设, 即初始分水岭变换所形成的每个区域, 其灰度特性具有一定的齐次性. 设两相邻区域 $R_i, R_j, i \neq j$, 若 $|\mu(R_i) - \mu(R_j)| < T_1$, 则 R_i 与 R_j 进行融合, 其中, $\mu(R_i)$ 为区域 R_i 的灰度平均值, T_1 为一门限. 另外, 本文还引入边的强度准则, 若设相邻区域 R_i 与 R_j 之间的边为 l , 则边界的强度定义如

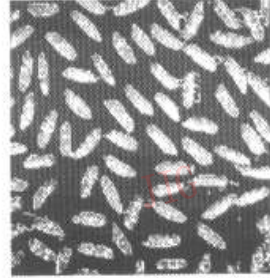
下: $Str(l) = \frac{1}{N} \sum_{p \in l} |Grad(p)|$, 其中, p 为边上 l 的点, $Grad(p)$ 为象素 p 的梯度, 这样, 当 $Str(l) < T_2$ 时, 则将两区域进行融合, 其中 T_2 为门限. 对于某区域来说, 区域灰度相似性准则是全局性的, 而边界的强度准则是局部性的. 另外, 由于初始分水岭变换得到的分割区域个数较多, 且区域面积相对较小, 因而每个区域的灰度齐次性较强, 此时, 可应用灰度齐次性准则来进行初始融合; 此后, 由于区域面积的增加, 灰度齐次性较弱, 即区域内象素的灰度方差较大, 此时则不宜采用灰度相似性准则, 而应采用边界强度准则进行融合。

4 实验结果及分析

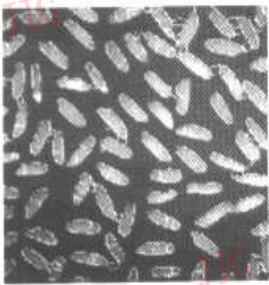
图 1 为大米图象及分割结果图象, 其中, 图 1(a) 为原始图象; 图 1(b) 为经分水岭分割后的大米图象, 由该图可看到过度分割的情况, 此时的区域个数为 732 个; 图 1(c) 为原始大米图象经保边界平滑后, 再进行分水岭变换的结果图象, 此时的区域个数为 224, 由此可见, 平滑过程使得区域个数减少很多; 图 1(d) 为在灰度相似性准则下进行融合的结果图象, 门限为 20, 此时的区域个数为 142; 图 1(e) 为在图 1(d) 的基础上, 根据边界强度准则进行融合的结果图象, 门限为 11, 此时的区域个数为 95, 由此即



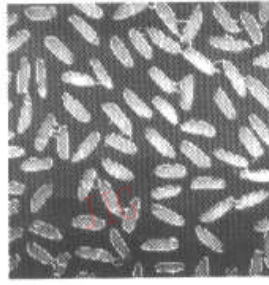
(a) 原始大米图象



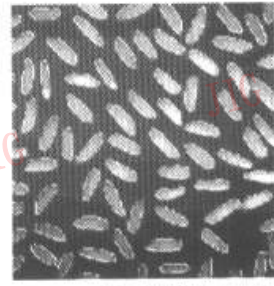
(b) 经分水岭分割后的大米图象



(c) 经过保边界平滑后的分水岭分割大米图象



(d) 根据图1(c)的区域邻接图,用灰度相似性准则融合后的大米图象

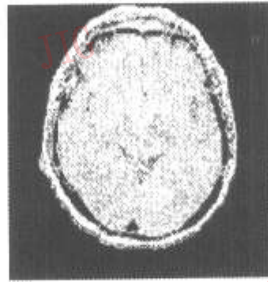


(e) 根据图1(d)的区域邻接图,用边界强度准则融合后的大米图象

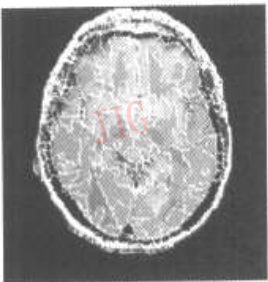
图1 大米图象融合结果



(a) 原始大脑图象



(b) 经分水岭分割后的大脑图象



(c) 经过保边界平滑后的分水岭分割大脑图象



(d) 根据图2(c)的区域邻接图,用灰度相似性准则融合后的大脑图象



(e) 在图2(d)的基础上,用边界强度准则融合后的大脑图象

图2 大脑图象融合结果

得到了大米的轮廓线.图 2 为原始及分割结果大脑图象,其中,图 2(a)为原始大脑图象,图 2(b)为原始大脑图象经分水岭分割后的结果图象.此时的区域个数为 2 217 个;图 2(c)为大脑图象经保边界平滑后的分水岭分割结果图象,其区域个数为 743;图 2(d)为在灰度由边界强度准则进行融合的结果图象,门限为 40,此时的区域个数为 289;图 2(e)为在图 2(d)的基础上,根据边界强度准则进行融合的结果图象,门限为 9,其区域个数为 218.实验结果表明,本文的方法获得了较好的分割结果.

5 结 论

本文首先使用非线性扩散方法对图象进行保边界平滑,然后对平滑后的图象进行分水岭分割,得到初始的分割区域,最后采用灰度相似性和边界强度准则进行区域的融合,从而较好地解决了分水岭变换方法中的过度分割问题,而且在交互式环境下,本文的方法对复杂图象的分割是很有有效的.

参 考 文 献

- 1 Canny J. A computational approach to edge detection. IEEE Trans. On Pattern Anal. Machine Intell, 1986, PAMI-8, (6): 679~698.
- 2 Ballard D H, Brown C M. Computer Vision. Prentice-Hall, Inc, New Jersey, 1982.
- 3 Laurent Najman, Michel Schmitt. Geodesic saliency of watershed contours and hierarchical segmentation. IEEE Trans. on Pattern Anal. Machine Intell. 1996, 18(12):1163~1173.
- 4 Paul T Jackway. Gradient watersheds in morphological scale-space. IEEE Trans. On Image Processing, 1996, 5 (6): 913~921.

- 5 Vincent L, Soille P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Trans. on Pattern Anal. and Mach. Intell. 1991,13(6):583~598.
- 6 Alvarez L, Lions P-L, Morel J-M. Image selective smoothing and edge detection by nonlinear diffusion I. SIAM Number. Anal. 1992,29(3):845~866.
- 7 Kostas Haris, Efstratiadis S N, Maglaveras N. Watershed-based image segmentation with fast region merging. In: Proceedings of IEEE International Conference on Image Processing, Chicago, IL, USA, 1998,3:338~342.



杜啸晓 1971 年生,分别于 1993 年和 1996 年获哈尔滨工程大学计算机系学士和硕士学位,现为上海交通大学图象处理和模式识别研究所博士研究生.主要研究方向为视频运动估计、运动分割等.



杨 新 1982 年获西北工业大学硕士学位,1995 年获布鲁塞尔自由大学工学博士学位,1995 年至 1996 年任西北工业大学计算机系副教授,主要从事时间序列图象中运动物体识别研究,已在国内外杂志和学术会议上发表文章 10 余篇.



施鹏飞 1940 年生,1965 年研究生毕业于上海交通大学,现任上海交通大学图象处理与模式识别研究所所长,教授,博士生导师.研究领域为图象分析、模式识别、智能技术与系统.已发表论文 80 余篇.