

# 修正与融合的肺部分割算法

窦圣昶, 赵海, 朱宏博, 王彬  
(东北大学 计算机科学与工程学院, 辽宁 沈阳 110169)

**摘 要:** 为了能在适当的计算复杂度下获得较好的分割效果,在总结现有方法优缺点的基础上提出了将简单方法进行修正与融合的肺部分割方法. 首先使用聚类的方法将 CT 图像中的像素点分为亮与暗两类,获得灰度值不同的小区域,然后区分出初始肺部的大致区域,最后根据 CT 图像的具体特性,使用识别出的初始肺部区域边界附近像素点的梯度以及灰度信息对边界进行修正. 使用上述综合方法对 CT 图像的处理结果表明,在计算复杂度不是太大的情况下能获得较好的分割效果.

**关 键 词:** CT 图像;肺结节;融合;聚类;梯度

中图分类号: TP 391.4      文献标志码: A      文章编号: 1005-3026(2018)11-1562-05

## Modify and Integrate Lung Segmentation Algorithm

DOU Sheng-chang, ZHAO Hai, ZHU Hong-bo, WANG Bin  
(School of Computer Science & Engineering, Northeastern University, Shenyang 110169, China. Corresponding author: DOU Sheng-chang, E-mail: doushengc@qq.com)

**Abstract:** In order to obtain better segmentation results under acceptable computational complexity, the advantages and disadvantages of the existing methods are summarized, and then the lung segmentation method which modifies and integrates the simple methods is proposed. Firstly, the clustering method is used to classify the pixels in CT images into two categories, and the small regions with different gray values are obtained. And then the initial region of the lung is identified. Finally, according to the specific characteristics of the CT image, the boundary is corrected by using the gradient and the gray information of the pixels near the initial lung boundary. The processing results of CT images using the above synthesis method show that the segmentation results can be better when the computational complexity is not too large.

**Key words:** CT images; pulmonary nodules; integrate; clustering; gradient

肺癌是现在致死率最高的癌症,由于肺癌的治疗效果与发现的时期密切相关,因此,早发现早治疗成为降低肺癌致死率的重要方法<sup>[1]</sup>. 高风险人群的肺癌筛查是实现早发现的重要手段,基于 CT 影像的筛查方式以其便捷、有效的特性成为肺癌筛查的主要方式.

随着肺癌筛查人群的增多,生成的 CT 图像正处于急速膨胀之中,传统的筛查方法是靠医生人工读取 CT 图像识别肺癌及其前期病变,识别的效果受医生的经验以及疲劳程度等状态影响较大,大量 CT 图像的辨识工作加重了医生的工作量,使得医生无法在有效的时间内完成所有 CT 图像的

诊断,并保证没有漏诊的病例. 帮助医生进行 CT 图像判读的计算机辅助系统成为很多医学以及计算机研究者们研究的重要领域<sup>[2-3]</sup>.

计算机辅助诊断系统是对肺部 CT 图像进行辨识,识别出候选的肺结节,供医生参考,系统可以分为 3 个功能部分,即肺部区域分割、肺部结节识别、肺结节的性质判断. 肺部区域分割功能是后续处理步骤的基础,其分割的效果直接影响到后续处理部分的可靠性<sup>[4-5]</sup>. 现有的 CT 图像上的肺部分割方法多是直接基于灰度阈值的分割方法<sup>[6-7]</sup>,其执行效果并不理想,在左右肺叶接近的 CT 图像上难以将左右肺区域分割开. 近期有学者

选择使用神经网络的方法对肺部区域进行分割<sup>[8-9]</sup>,然而,其计算复杂度急速增长,需要过于强大的计算硬件以及很长的计算时间,同时用于训练的数据量非常大,难以在实际中使用。

本文提出了基于多方法融合的肺部区域识别算法,以提高肺部区域识别的有效性与准确性。

# 1 肺分割建模

含有较多空气的肺部区域对 X 射线的吸收较少,而包围肺部区域的胸墙对 X 射线有更强的吸收效果,人体其他组织与肺部区域在 CT 图像上表现为不同的灰度值,肺部以及体外的空气区域表现为亮度较暗的区域,如图 1 所示,同时,身体之外的物体以及身体内的不同部分对灰度上的区分有较大的干扰,因此基于灰度值的肺部分割方法可以作为初始的预处理方法使用<sup>[10]</sup>。本文首先使用了灰度聚类的方法对整个图像进行处理,以获得初始的处理结果,然后根据人体组织对 X 射线的吸收特性从处理结果中分割出肺部区域。

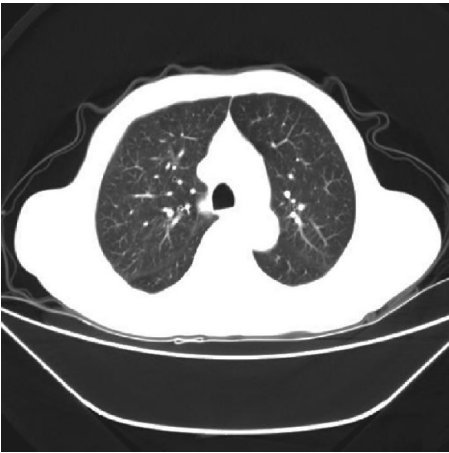


图 1 原始 CT 图像  
Fig. 1 Original CT image

## 1.1 CT 图像的聚类处理

CT 图像上的聚类处理是将图像中的像素点按照灰度值进行聚类以获得初始像素点分类信息的处理方法及其过程。Senthilkumar 等<sup>[7]</sup>使用最小最大值的方法来确定初始的聚类中心,该方法是一个递归的过程,在此问题中初始聚类中心的选择对结果没有影响,对求解过程的影响同样很小,因此本文在使用该方法的过程中在保证结果相同的情况下舍弃了不必要的步骤。本文的算法过程:

1) 设置初始聚类中心。设置两个类别的聚类中心  $c1$  与  $c2$ ,使得  $c1 < c2$ 。

2) 计算图像  $I$  的每个点到两个聚类中心的距离矩阵  $d1$  和  $d2$ :

$$d1 = \sqrt{(I - C1)^2}, \quad (1)$$

$$d2 = \sqrt{(I - C2)^2}. \quad (2)$$

其中: $I$  是图像矩阵; $C1$  与  $C2$  分别是由聚类中心  $c1$  和  $c2$  扩展成的与  $I$  等大小的矩阵。

3) 由对应像素点处距离  $d1$  与  $d2$  的大小关系获得两个聚类结果矩阵  $K1$  与  $K2$ :

$$K1(i,j) = \begin{cases} I(i,j), & d1 < d2; \\ 0, & \text{其他}. \end{cases} \quad (3)$$

$$K2(i,j) = \begin{cases} I(i,j), & d1 > d2; \\ 0, & \text{其他}. \end{cases} \quad (4)$$

4) 由矩阵  $K1$  与  $K2$  的值更新聚类中心  $c1$  与  $c2$ 。

$$c1_{\text{update}} = \frac{1}{L1} \sum_{i=1}^m \sum_{j=1}^n (K1(i,j) \neq 0), \quad (5)$$

$$c2_{\text{update}} = \frac{1}{L2} \sum_{i=1}^m \sum_{j=1}^n (K2(i,j) \neq 0). \quad (6)$$

其中, $L1$  与  $L2$  分别是  $K1$  与  $K2$  中不等于零的像素点的个数。

5) 当  $c1$  与  $c1_{\text{update}}$  以及  $c2$  与  $c2_{\text{update}}$  间的差别足够小时结束,否则重复执行 2)~4) 步。

聚类算法流程图如图 2 所示。

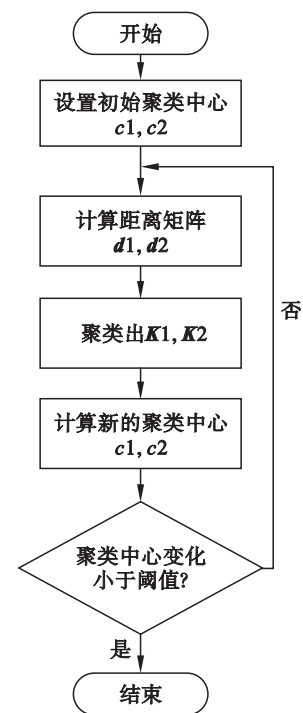


图 2 聚类流程图  
Fig. 2 Clustering flow chart

使用上述算法对图 1 进行聚类,结果如图 3 所示.在图 3 中,将图 1 中的区域聚类成两个区域,图 3a 是聚类结果中的  $K1$  矩阵的可视化显示图像,原来灰度值较小的区域保持原值,灰度值较大的区域像素点的灰度值被置为零,图 3b 是聚类结果中的  $K2$  矩阵的可视化显示图像,保留原值的是灰度值较大的区域,其他区域被置为零,两图之和即是原来的 CT 图像.

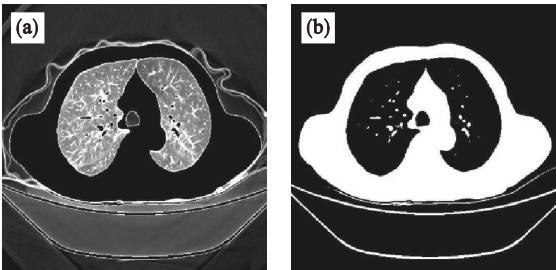


图 3 聚类结果  
Fig. 3 Clustering result  
(a)— $K1$ ; (b)— $K2$ .

1.2 肺部区域的初始分割

对 CT 图像进行聚类处理后获得的是离散点的形式,由于图像灰度的变化特性,除个别明显区域外,还不能形成统一的连通识别结果区域,同时,由于肺部高亮度肺结节以及血管的影响,难以直接由聚类结果识别出肺部区域.本文使用基于胸墙的连通高亮度大区域特性实现胸墙的识别,首先识别出胸墙区域,进而实现肺部区域的初始识别.

在 CT 图像中,包围肺部区域的胸墙区域由于对肺部的支撑作用,是由质地较密的组织组成,在图像中表现为较大区域的连续高亮度区域,因此可以使用大面积的联通区域特性对其进行识别,识别出的胸墙区域如图 4 所示.

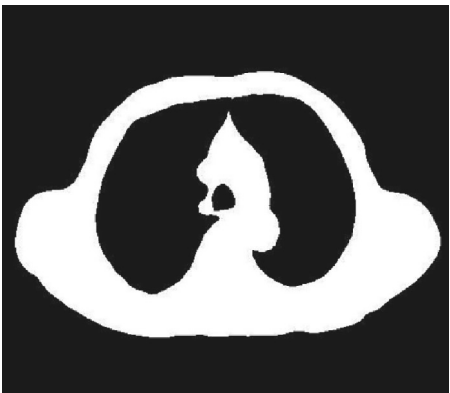


图 4 胸墙区域图  
Fig. 4 Chest wall area figure

在图 4 中已经没有了影响肺部区域分割的小

的区域,并且将肺部 CT 图像分割成了明显的 3 部分区域,对图 4 进行处理能很容易获得初始的肺部区域.

2 肺区域边界修正模型

由图 4 可以明显地看出,使用聚类后的结果分割出的肺部区域在左右肺比较近的部分无法区分左右肺叶,会将其识别为一个区域,这会对后续的处理造成影响,因此需要对其进行处理,本文提出了基于梯度向量场与边界灰度相结合的肺边界修正算法.

在识别出包含肺部的大致区域后,在识别结果中存在两个问题,即左右肺叶的分割问题以及肺部区域模糊边缘处会形成错误边界的问题.然而单一的基于梯度的方法无法解决该问题,本文提出将梯度与灰度相结合的肺边界修正方法,将像素点的梯度和灰度值结合在一起作为新的特征值用于边界识别.像素点的特征值为梯度与灰度的加权和:

$$T = aV + bG . \tag{7}$$

式中: $V$  为像素的灰度值; $G$  为像素点处的梯度值;实验中对  $a$  和  $b$  取值为常数 1. 在初始边界上寻找初始的种子节点,作为边界扩展的初始点,在该像素点的梯度向量一侧寻找  $T$  值最大的像素点,以获得下一个边界点.递归执行上述过程,直到识别出的边界点形成环,即获得了包含目标区域的边界点序列.算法的伪代码如下,结果如图 5 所示.

算法描述
Begin(算法开始)
输入初始分割肺部区域图像 Image
计算图像的梯度 Gradient
选取初始边界 Edges 上的种子节点 $P$
While 新边界 NewEdges 未形成封闭区域
计算种子节点一侧 $180^\circ$ 内的像素点的特征值 $T$
选取 $T$ 值最大的节点作为下一个种子节点
End While
End(算法结束)

分析实际的 CT 图像可知,在分割出的肺部区域的左右两边,边界坐标值点具有清晰的边界以及区域特性,同时初始寻找简单,适于作为种子点,实验中使用了该点作为种子点.

该方法的优势是,识别出的边界一定是连续的封闭边界,不需要后续的边界连续化操作.同时

由于该方法的基础是前期获得初始预处理边界,因此能获得较好的边界轮廓,识别的肺部区域更精确.通过识别出的边界可以分割出准确的肺部区域范围.

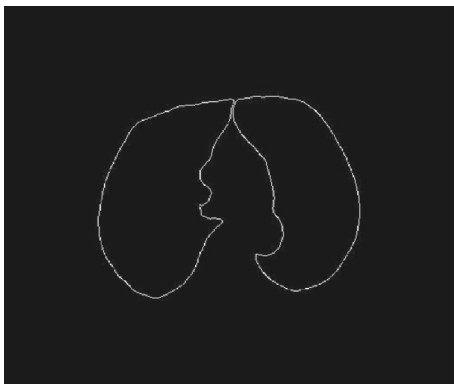


图 5 肺部边界图  
Fig. 5 Lung boundary figure

### 3 实验过程与结果

#### 3.1 实验过程

文中使用从盛京医院获得的肺部 CT 图像作为处理对象. CT 图像经过如前所述的聚类、肺部识别以及边界修正后获得识别出的肺部区域,整体的框架如图 6 所示.



图 6 算法整体框图  
Fig. 6 Block diagram of the algorithm

#### 3.2 实验结果

实验中对多种 CT 图像使用本文的融合方法进行肺部区域的分割识别,对几种代表性的肺部 CT 图像的识别结果如图 7 所示.

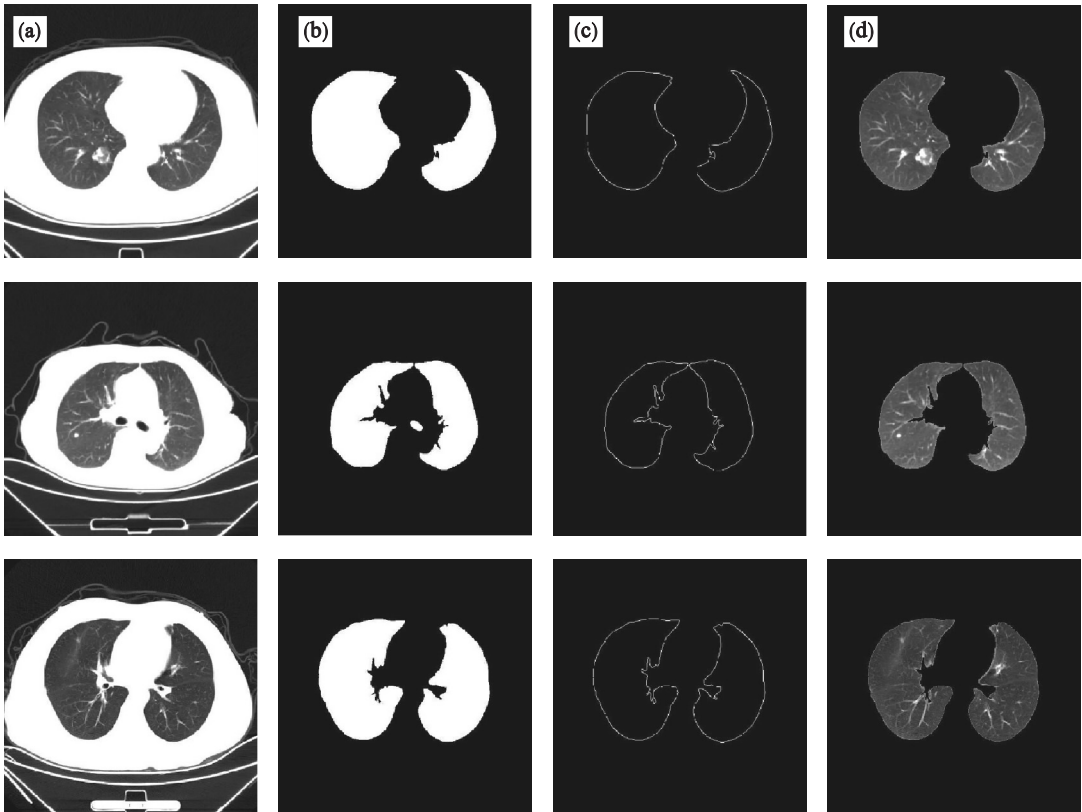


图 7 多实例肺部分割结果  
Fig. 7 Multi-instance lung segmentation results

(a)—原始 CT 图像; (b)—聚类法获得的初始肺部区域; (c)—修正后的边界; (d)—最后识别结果.

由图中可以看出,本文将多种方法进行融合的肺部分割方法能很好地分割出 CT 图像中的肺部区域,为后续的精确定义提供基础.

#### 3.3 性能分析

在聚类以及肺部区域初始识别的部分,是对

图像的像素点进行遍历操作,其计算复杂度为  $O(n)$ ,最后的边界修正部分同样是在像素上的操作,遍历初始肺部区域的边界点,因此其做大遍历的像素点数量不会超过图像中的像素点总数,计算  
(下转第 1571 页)