

基于局部阈值和聚类中心迭代的肺结节检测算法

陈侃 李彬 田联房

(华南理工大学自动化科学与工程学院 广州 510640)

摘要 肺部疾病通常以肺结节的形式表现出来。为了对肺部疾病进行诊断治疗,需要对肺结节进行准确的检测。提出了基于局部阈值和聚类中心迭代的肺结节检测算法。首先,对肺实质图像采用局部阈值算法,提取感兴趣区域(ROIs),并且计算 ROIs 的形态特征、灰度特征和纹理特征;其次,结合规则、聚类中心迭代和欧式距离,对 ROIs 进行分类。实验结果表明,所提算法能够较好地检测出孤立性结节、低对比度结节和粘连肺壁结节。

关键词 肺结节,局部阈值,聚类中心迭代,欧式距离

中图分类号 TP391.41 文献标识码 A

Pulmonary Nodules Detection Algorithm Based on Local Threshold and Iterative of Clustering Center

CHEN Kan LI Bin TIAN Lian-fang

(School of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract Lung disease is usually showed in the form of pulmonary nodules. In order to diagnose lung disease, pulmonary nodules must be detected accurately. This paper proposed pulmonary nodules detection algorithm based on local threshold and iterative of clustering center. First, local algorithm was proposed for extracting interested region(ROIs) in lung parenchyma image. Morphological features, gray features and texture features were calculated in interested region (ROIs). Second, rule, iterative algorithm of clustering center and Euclidean distance were used to classing ROIs. Experiment results show that the proposed algorithm can better detect isolation nodules, low contrast nodules and nodules which adhere to lung wall.

Keywords Pulmonary nodules, Local threshold, Iterative of clustering center, Euclidean distance

临床研究表明,肺癌是导致人类死亡的最大恶性肿瘤之一^[1]。肺癌在早期阶段常常表现为肺结节的形式。检查并且治疗肺结节,可以提高肺癌患者的生存率^[2]。随着医学影像技术的发展,CT 广泛运用在肺部疾病检测,并且成为肺结节检测的有效设备。CT 扫描会产生大量的切片,极大增加了影像学医生的劳动强度。为了缓解影像学医生的工作量,急需产生计算机辅助系统(CAD),帮助影像学医生初步检测肺结节^[3]。肺部 CAD 是一种有效的系统,可以精确检测出肺结节的位置,并且提高阅片速度。

在 CAD 系统中,肺结节计算机辅助自动检测是一项热点研究主题。现在,许多学者提出了肺结节检测算法:灰度阈值法^[4],因为肺结节像素 CT 值通常比肺实质像素 CT 值高,所以用灰度阈值法可以识别肺结节初始的候选点;圆形滤波加强法^[5,6],因为肺结节在 2 维平面上通常表现为圆形结构,而血管和气管在 2 维平面上,通常表现为线状结构,所以用圆形加强滤波器可以加强肺结节,从而消除线性结构对象;模板匹配法^[7,8],用模板匹配方法检测肺结节,其优点是只需要定义结节的一些参考模型(如球形、半圆形等),即可通过快速模板匹配检测到结节,但是模型匹配方法都是基于各种模型假设提出的结果,并不是所有的肺部 CT 图像都满足假设条件,因

此有一定的局限性。以上列举的算法检测肺结节的精度很大程度上依赖参数的选择,并且不同实验数据选取不同的参数。本文提出的基于局部阈值和聚类中心迭代的肺结节检测算法,能够较好地检测出孤立性结节、低对比度结节和粘连肺壁结节。

1 算法描述

基于局部阈值和聚类中心迭代的肺结节检测算法流程图如图 1 所示。本文提出的算法包括以下 4 个步骤:(1)对原始的 CT 图像进行肺实质分割;(2)采用局部阈值算法,分割感兴趣区域(ROIs),并且计算 ROIs 的形态特征、灰度特征和纹理特征;(3)基于规则肺结节剔除;(4)采用聚类中心迭代算法和欧式距离,对 ROIs 进行分类。

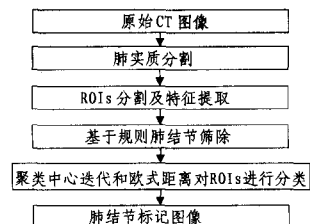


图1 本文算法处理步骤

到稿日期:2011-03-15 返修日期:2011-05-25 本文受广东省教育部产学研结合项目(2009B090300057),教育部博士点基金资助项目(200805610018),广州市番禺区科技攻关项目(2009-Z-108-1),华南理工大学中央高校基本科研业务费(2009ZM0077)资助。

陈侃(1984-),男,博士生,主要研究方向为医学图像处理与模式识别;李彬(1979-),男,博士,讲师,主要研究方向为科学计算可视化、图像处理与模式识别,E-mail:binleermnboy@yahoo.com.cn(通信作者)。

1.1 肺实质分割

为了提高肺结节检测的速度,同时为防止肺外器官对检测造成影响,在肺结节检测之前,需要进行肺实质分割。本文采用综合的方案,结合全局阈值法、边界检测、区域增长和形态学处理,实现肺实质分割。先分割左右肺叶以及去除气管和支气管,而后通过形态学处理的方法获取没有空洞并带有光滑边界的肺实质。

1.2 ROIs 分割和特征值提取

1.2.1 ROIs 分割

通常,ROIs 与背景之间的灰度具有模糊性。为了很好地保持 ROIs 边界的形状和分割对比度低的 ROIs,本文采用局部阈值算法产生局部阈值^[9],并根据局部阈值对肺实质图像进行分割,提取 ROIs。局部阈值算法具体实现如下:

1)采用高斯滤波,对肺实质图像进行平滑处理;

2)对平滑后的肺实质图像,每点像素进行以下处理:

a)设置一个局部窗口 $w * h$,其中 w, h 分别为局部窗口的宽度和高度;

b)求取局部窗口最大灰度值 g_{max} 和最小灰度值 g_{min} ;

c)计算局部阈值 T :

$$T = \frac{(g_{max} + g_{min})}{2} \quad (1)$$

d)进行二值化处理:如果该点的灰度值大于局部阈值 T ,则将该点的像素设置为白色;否则,将该点的像素设置为黑色。

1.2.2 特征值提取

在进行 ROIs 分类时,特征提取是一个关键的步骤。好的特征组合必须满足以下两个方面要求:(1)特征组合可以很好地描述 ROIs 的特征;(2)这些特征具有几何变换不变性。根据以上两个方面,本文提取 ROIs 的形态特征、灰度特征和纹理特征^[10]:(1)ROIs 的形态特征包括周长、面积、直径、圆形度、矩形度、细长度和扁度;(2)ROI 的灰度特征包括灰度均值和灰度方差;(3)纹理是对图像的像素灰度级在空间上的分布模式的描述,纹理特征是图像局部性质的统计。设图像灰度级为 L ,由于 ROIs 通常比较小,本文分别考虑 $0^\circ, 45^\circ, 90^\circ, 135^\circ$ 4 个方向,相隔 $\delta=2$ 的像素距离的一对像素,选用灰度共生矩阵的 4 个参数:能量、对比度、熵和逆差矩。

1.3 基于规则、聚类中心迭代和欧式距离的分离器

1.3.1 基于规则的分类器

根据肺结节在医学中的征象,肺结节通常表现为肺内类圆形病变。因此,可以利用这个医学征象,根据 ROIs 的直径、大小和圆形度,去除大量的非肺结节。基于如下规则筛除假阳性结节:

规则 1 如果候选结节的直径 $a > T_a$,则可以判断该区域为非肺结节区域;

规则 2 如果候选结节的面积 $A > T_A$,则可以判断该区域为非肺结节区域;

规则 3 如果候选结节的圆形度 $S < T_s$,则可以判断该区域为非肺结节区域。

其中, T_a, T_A 和 T_s 为经验阈值。

1.3.2 聚类中心迭代算法

本文采用聚类中心迭代算法实现同类训练样本的聚类。

定义目标函数为

$$d = (x_i - c)^T * (x_i - c) \quad (2)$$

式中, d 表示第 i 类训练样本到本类聚类中心的距离, x_i 表示第 i 个训练样本, c 表示本类聚类中心。

求目标函数的最小值,目标函数关于聚类中心 c 求导:

$$\frac{\partial d}{\partial c} = -(x_i - c) \quad (3)$$

将式(3)代入聚类中心迭代方程,表示如下:

$$c(k+1) = c(k) - \frac{1}{k+1} * \frac{\partial d}{\partial c} \Big|_{c=c(k)} \quad (4)$$

式中, $c(k+1)$ 表示第 $k+1$ 步的聚类中心, $c(k)$ 表示第 k 步聚类中心。

聚类中心迭代法的算法步骤:

1)初始化聚类中心 $c(0)$;

2)输入第 k 次迭代样品,根据式(2)计算 d ;

3)根据式(4)计算第 $k+1$ 步的聚类中心;

4)循环执行 2)、3),直到第 $k+1$ 步的聚类中心变化非常小;

5)获得最终的聚类中心 $c=c(k+1)$ 。

1.3.3 欧式距离分类器

采用欧式距离对 ROIs 进行分类:

$$Mdis = \sqrt{(x_i - c)^T * (x_i - c)} \quad (5)$$

式中, $Mdis$ 表示样本 x_i 到聚类中心的距离。

欧式距离分类算法,步骤如下:

1)根据规则进行候选肺结节的初删除;

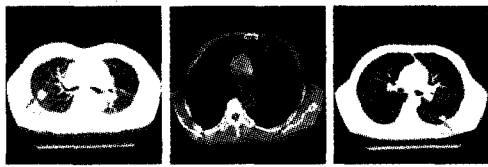
2)根据训练样本,计算肺结节样本的聚类中心 c_T 和非肺结节的聚类中心 c_N ;

根据式(5)计算未知分类的候选结节 x ,分别计算与聚类中心 c_T 和聚类中心 c_N 的距离。如果候选结节 x 到聚类中心 c_N 的距离小于其到聚类中心 c_T 的距离,则筛除候选结节。

2 实验结果及分析

本文采用的实验图像取自广州某三甲医院。CT 肺图像的分辨率为 $512 * 512$,存储格式为 16 位的 DICOM 图像。

本文选用的病例包含了 44 幅图像,肺结节的大小从 0.2~3cm,肺结节位置不定,有孤立的,也有与肺壁粘连和低对比度的。其中 50 个肺结节被医生标志。聚类中心迭代算法采用的训练样本包括 145 个样本,其中有 30 个结节样本和 115 个非结节样本。基于规则的分离器中,经验阈值 T_a, T_A 和 T_s 分别为 6, 30 和 0.3。实验使用 3 个病例,如图 2 所示。(a)病例 1 中包含了一个孤立性的肺结节,(b)病例 2 中包含了一个低对比度的肺结节,(c)病例 3 中包含了一个与肺壁粘连的肺结节。



(a) 病例 1 (b) 病例 2 (c) 病例 3

图 2 实验原图

采用肺结节检测的敏感性(Sensitivity),分析本文提出的基于局部阈值和聚类中心迭代的肺结节检测算法。分别使用 2 个不同的局部窗口分割 ROIs,敏感性如表 1 所列。从表 1 可以看出,选择合适的局部窗口可以提高算法检测的敏感性。

表 1 肺结节检测的敏感性

局部窗口大小 (w * h)	阳性结节的数目 (TPs)	假阴性结节的数目(FNs)	本文算法敏感性 (sensitivity)
31 * 31	32	18	0.64
51 * 51	39	11	0.78

肺结节检测的敏感性定义为

$$Sensitivity = \frac{TP}{TP + FN} \quad (6)$$

式中,TP 为检测的真阳性肺结节, FN 为漏检真阳性肺结节。

本文造成检测敏感性低的主要原因有:(1)肺结节太小,医学征像和提取的特征很难与血管区分,从而造成小结节漏检。(2)算法是基于 2D 空间,在 2D 空间中,肺结节、血管和气管具有很多相似的特征,从而造成肺结节漏检。(3)结节因与肺壁相粘连,分割出的区域形状不规则,也会产生漏检。

对 3 种类型的肺结节检测结果如图 3 所示。(a)为孤立性结节检测结果,(b)为低对比度结节检测结果,(c)为粘连肺壁结节检测结果。从实验检测结果可以观察到,本文提出的算法可以较好地检测肺孤立性肺结节、低对比度的肺结节和粘连肺壁性的肺结节。



(a) 病例 1 检测结果 (b) 病例 2 检测结果 (c) 病例 3 检测结果

图 3 检测的结果

结束语 本文提出了基于局部阈值和聚类中心迭代的肺结节检测算法。首先,采用局部阈值算法分割 ROIs,并且计算 ROIs 的形态特征、灰度特征和纹理特征。其次,根据肺结节在 CT 图像的表现形式,采用基于规则的方法消除非结节

的区域并对训练样本采用聚类中心迭代算法,分别求其肺结节聚类中心和肺非肺结节聚类中心。最后,根据欧式距离,对未知分类的候选结节进行分类。实验结果表明,本文提出的算法能够较好地检测出孤立性结节、低对比度结节和粘连肺壁结节。

参考文献

- [1] American Cancer Society. Cancer facts and figures 2005[EB/OL]. <http://www.cancer.org>
- [2] Liu Y, Yang J Z, et al. A Method of Pulmonary Nodule Detection Utilizing Multiple Support Vector Machines[C]//IEEE International Conference on Computer Application and System Modeling. 2010, v10-118-v1-1216
- [3] Noriaki M, Hyoungseop K, Yoshiinori I, et al. Automatic Detection of Lung Nodules in Temporal Subtraction Images by Use of Shape and Density Features [C]// IEEE Fourth International Conference on Innovative Computing. 2009;1288-1292
- [4] Zhao B, Gamsu G, Ginsburg M S. Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm [J]. J Appl Clin Med Phys, 2003, 4(3): 248-260
- [5] Li Q. Recent progress in computer-aided diagnosis of lung nodules on thin-section CT [J]. Computerized Medical Imaging and Graphics, 2007, 31(4/5): 248-257
- [6] Baek T, Kim J S, Na Y H, et al. Pulmonary Nodules: Automated Detection on CT Images with Morphologic Matching Algorithm- Preliminary Results [C]//Radiology. 2005, 236: 259-299
- [7] Dehmshiki J, Ye X, Lin X, et al. Automated detection of lung nodules in CT images using shape-based genetic algorithm [J]. Computerized Medical Imaging and Graphics, 2007, 31(6): 408-417
- [8] Dehmshiki J, Ye X, et al. A Hybrid Approach for Automated Detection of Lung Nodules in CT Images [C]//IEEE International Symposium on Biomedical Imaging; Macro to Nano. 2006: 506-509
- [9] Tong L J, Chen K, et al. Document Image Binarization Based on NFCM[C]//IEEE Confence Image and Signal Processing. 2009: 1-5
- [10] Zhang J, Li B, Tian L F. Lung nodule classification combining rule-based and SVM [C]//5th IEEE International conference on Bio-Inspired Computing: Theories and Applications (BIC-TA). 2010, 2: 1033-1036

(上接第 287 页)

- [13] Deb K. An efficient constraint handling method for genetic algorithms[J]. Computer Methods in Applied Mechanics and Engineering, 2000, 186(2-4): 311-338
- [14] Jiménez F, Verdegay J L. Evolutionary techniques for constrained optimization problems[C]// Zimmermann HJ, ed. Proc. of the 7th European Congress Intelligence Techniques and Soft Computing(EUFIT'99). Berlin; Springer-Verlag, 1999

- [15] Ghasabi-Oskoei H, Mahdavi-Amiri N. An efficient simplified neural network for solving linear and quadratic programming problems[J]. Applied Mathematics and Computation, 2006, 175(1): 452-464
- [16] Runarsson T P, Yao X. Search biases in constrained evolutionary optimization[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2005, 35(2): 233-243