

基于 EM 和 Mean-shift 的肺结节分割

孙申申¹⁾ 李宏¹⁾ 侯欣然²⁾ 康雁^{1),2)} 赵宏^{1),2)}

¹⁾(东北大学 信息科学与工程学院,沈阳 110004) ²⁾(国家数字化医学影像设备工程技术研究中心,沈阳 110079)

摘要 针对结节与血管相连且两者 CT 值相近造成分割困难的问题,提出了一种基于期望最大(estimation maximum, EM)的自适应带宽参数选择的方法,并采用均值漂移(Mean-shift)算法解决结节分割。与基于统计分析规则的带宽选择方法和基于最优化的带宽选择方法相比,该方法能直接求得正确带宽参数,且时间复杂度低。应用血管梯度的法向量方向服从正态分布,而结节梯度的法向量方向服从均匀分布,建立血管粘连型结节模型,并用期望最大估计模型参数,根据均匀分布的权重和带宽选择定理确定带宽参数。该方法对仿真数据和 CT 数据(19 个粘连血管性肺结节)进行评估实验,都取得了正确的分割结果。结果表明,该方法对分割粘连血管型结节是有效的。

关键词 肺结节 期望最大 均值漂移 梯度法向方向分布特征

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2009)10-2016-07

Pulmonary Nodule Segmentation Based on EM and Mean-shift

SUN Shen-shen¹⁾, LI Hong¹⁾, HOU Xin-ran²⁾, KANG Yan^{1),2)}, ZHAO Hong^{1),2)}

¹⁾(College of Information Science and Engineering, Northeastern University, Shenyang 110004)

²⁾(National Engineering Research Center for Digital Medical Imaging Device, Shenyang 110079)

Abstract Aiming at solving the segmentation problem caused by the connection of lung nodule and vessel, a new adaptive bandwidth selection method based on EM is proposed and we apply it into nodule segmentation. Compared it to the method of bandwidth chosen based on statistical analysis rule or optimized rule, it has some advantages such as low time complexity and correct bandwidth in accord with a real problem. The vertical orientation vectors of vessel's gradient was constructed as to normal distribution and the vertical orientation vectors of nodule's gradient as uniform distribution, we modeled the nodule connected vessel, estimated model parameters by EM, and extracted bandwidth values in Mean-shift based on the weight of uniform distribution and bandwidth selection theorem. The proposed method was tested on synthetic data and the clinical chest CT volumes, and all the results were correct. The results revealed that the proposed method is successful in segmentation lung nodules connected vessel.

Keywords lung nodule, EM, Mean-shift, vertical orientation vectors of gradient's distribution feature

1 引言

肺癌是死亡率最高的癌症。肺结节是肺癌的影像表现形式。在临床中,诊断肺结节的良、恶性意义

重大。肺结节的生长速度是鉴别良、恶性的指标。计算机辅助诊断(computer aided detection, CAD)系统为定量地诊断肺结节良、恶性提供了新的手段。它应能自动地帮助医生分割影像上已检测出的结节,目的是测量它的体积和计算一段时间内它的倍

基金项目:国家自然科学基金项目(60671050,60771067)

收稿日期:2008-04-22;改回日期:2008-09-01

第一作者简介:孙申申(1980 ~),女。东北大学信息学院计算机应用技术博士研究生。主要研究方向为计算机辅助诊断。

E-mail:sunshenshen@126.com

增率。对比同一个病人不同时间的两幅影像,如果其中结节体积有明显增大,则该结节为恶性结节。精确地分割各类结节的边缘则是必须解决的问题。结节的类型有孤立型、血管粘连型、胸壁粘连型、毛玻璃型和空洞型。其中,血管粘连型结节恶性概率最大,而这类结节的 CT 像素灰度值通常与血管上的 CT 像素灰度值接近,仅考虑灰度信息将无法准确分割。因此,精确地分割血管粘连型肺结节是有意义的,并且具有很大难度。

目前,国内外在这一方向上做了广泛的研究。文献[1]用形态学中的腐蚀和膨胀操作分割血管粘连型肺结节,在把血管和结节分开的同时,结节的毛刺也被腐蚀掉了,毛刺是判断结节良、恶性的又一重要指标,可见这个方法的应用使肺结节丢失了重要的信息。用形状限制的动态规划方法逐层得到结节大致轮廓线,再用 EM 方法求得结节的精确边缘^[2],但没有实现真正意义上的 3 维分割,同时由于优化方法的应用导致算法速度慢。高斯模板匹配方法过于严格,当结节影像被噪声污染时不能得到太好的分割结果^[3],需要结合非参数密度估计方法(如 Mean-shift)来提高估计的准确性和自适应性。然而,Mean-shift 算法需要自适应确定带宽参数,不适当的带宽参数会导致过分割和欠分割现象。文献[3]用迭代步长法求得最佳带宽参数,但带宽选择范围过长,迭代次数过多。文献[4]在其基础上通过先选择初始带宽参数再用迭代步长法求得最佳带宽参数,缩短了带宽参数的选择范围。上述两个方法都是在一定的参数选择范围内,逼近带宽参数的最优值,其分割结果将会受到步长的影响。如果步长太大,不能得到很好的分割效果,步长太小,计算速度就会变慢。

本文提出了一种直接求得 Mean-shift 带宽参数的方法,从而解决了粘连血管的肺结节精确分割的问题。血管里充满着向同一方向流动的血液,这就使得血管上像素梯度的垂直方向基本一致,服从正态分布,而结节上像素梯度的垂直方向服从均匀分布。首先利用关系矩阵求得梯度的法向量方向,然后,假设粘连血管型结节像素梯度的垂直方向服从正态分布和均匀分布的混合概率分布,基于期望最大拟合出参数和权重。均匀分布的权重可以被看作是结节上像素个数占粘连血管型结节总像素个数的比重。最后,根据均匀分布的权重求得核函数的带宽参数,代入均值漂移聚类分割出精确的肺结节边

缘。人造数据和肺结节真实数据的实验结果表明,该方法快速而有效地解决了粘连血管型结节的分割问题。

2 基于 EM 的结节权重估计

2.1 预处理

首先,基于 Hessian 矩阵^[5]的圆点滤波器检测出球状候选点。设点 x Hessian 矩阵的特征值 $\lambda_1(x), \lambda_2(x), \lambda_3(x)$ 且 $|\lambda_1(x)| \leq |\lambda_2(x)| \leq |\lambda_3(x)|$, 候选点集合: $s = \{x \mid |\lambda_3(x)|^2 / |\lambda_1(x)| > \tau\}$ 。自适应的阈值 τ 优于文献[5]中固定的阈值, τ 为滤波图像的方差,如图 1 所示。这样,不但可使结节候选点不漏,也不会造成候选点过多,给“去假阳”步骤减少了不必要的负担。

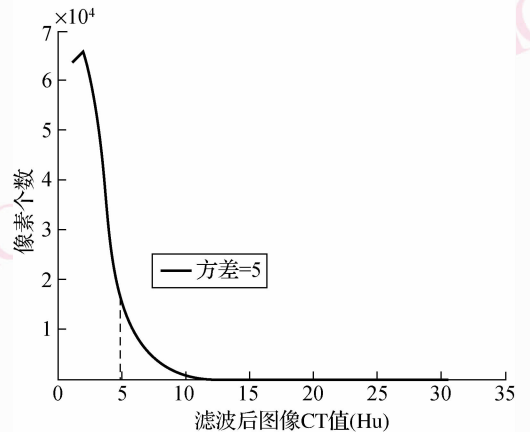


图 1 圆点滤波图像阈值选取

Fig. 1 Choose threshold of dot filter image

其次,区域生长出该结节候选点的中心点 c 。自适应确定以 c 为中心的感兴趣区域^[6] (region of interest, ROI)。肺 CT 影像上高亮区域(包括血管和结节)的 CT 值通常大于 -750 Hu, 而肺实质区像素的 CT 值小于这个阈值,如图 2 所示。ROI 的初始尺寸为 $3 \times 3 \times 3$, 逐渐扩大 ROI 的尺寸直到高亮区域的像素数目占整个 ROI 的像素数目的比率小于给定阈值。设 W_c^k 为第 k 次迭代以 c 为中心高亮区域像素个数。则 ROI xy 方向上的尺寸为

$$r_{xy}^k = W_c^k / k^3 \quad (1)$$

CT 影像像素间距为 x_{mm} mm, 影像间的层厚为 z_{mm} mm, 则 ROI 在 z 方向上的尺寸为

$$r_z^k = x_{mm} \times r_{xy}^k / z_{mm} \quad (2)$$

再次,以 c 为种子点和迭代阈值法得到的阈值

在自适应 ROI 内生长出粘连血管型结节的前景区。

最后,把 ROI 沿 z 方向最大密度投影(maximum intensity projection, MIP)到 2 维平面上,投影后前景区的像素个数 N_{uv} ,如图 3 所示。

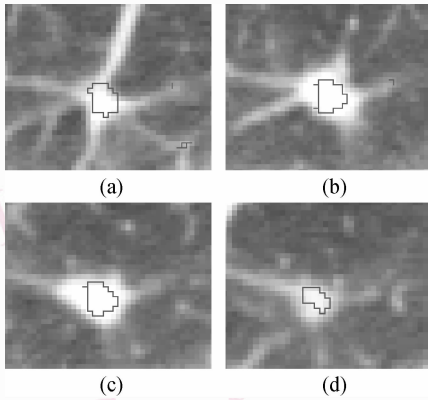


图 2 一个 3 维血管粘连型结节四层数据
Fig. 2 A 3D connected vessel nodule



图 3 图 2 结节的 MIP

Fig. 3 Maximum intensity project of figure 2's nodule

2.2 基于关系矩阵(2D)的血管流向特征提取

文献[7]中给出了 3D 的关系矩阵,这里导出 2D 的关系矩阵。在 MIP 图像上以像素 j 为中心的 3×3 窗中的 2 维高斯梯度向量 $\zeta_j = \{g_i\}_{i=1}^n$,其中 $n = 3^2$, g_i 为像素 i 的梯度向量。由于 ROI 尺寸小,高斯尺度也应该小。垂直于 ζ_j 的向量 v 被定义为投影到 ζ_j 中

所有向量的平方和最小 $E(v) = \frac{1}{n} \sum_{i=1}^n ((g_i)^T v)^2 = v^T G G^T v$,其中, $G = \frac{1}{\sqrt{n}} [g_1, \dots, g_n]$, $G G^T$ 是 2×2 关系矩阵。 λ_1, λ_2 为 $G G^T$ 的特征值且 $\lambda_1 < \lambda_2$,与 λ_1 对应的特征向量是 e_1 ,则像素 i 的梯度的法向量为

$$v_i = \sqrt{\lambda_2} \times e_1 \quad (3)$$

v_i 表示血管流动的方向。

虽然特征向量有唯一的方向,但是特征向量不是唯一的。为了让向量充满 $[0, 2\pi]$ 区间,需要重新定义特征向量的象限:

- 若 $g_i(x_0, y_0) \in \text{I}$ 则 $v_i(x_0, y_0) \in \text{II}$
- 若 $g_i(x_0, y_0) \in \text{II}$ 则 $v_i(x_0, y_0) \in \text{III}$
- 若 $g_i(x_0, y_0) \in \text{III}$ 则 $v_i(x_0, y_0) \in \text{IV}$
- 若 $g_i(x_0, y_0) \in \text{IV}$ 则 $v_i(x_0, y_0) \in \text{I}$

为了保证 ROI 区域内血管流向一致,需要把一些向量的方向转换到其相反的方向。在指向同一方向最多的向量中任取一个向量 $Arbv$,如果它和其他向量 v_i 之间的夹角在 $(\frac{\pi}{2}, \pi]$ 范围内,把向量 v_i 转到相反的方向 $-v_i$ 。也就是说,

$$v_i = \begin{cases} -v_i & -1 \leq \frac{v_i \cdot Arbv}{|v_i| \times |Arbv|} < 0 \\ v_i & \text{其他} \end{cases} \quad (5)$$

如果直方图不连续,把部分角度加 2π 后使得直方图连续并使所有角度在 $[0, \pi]$ 范围内。

2.3 粘连血管型结节模型

假设结节和血管都服从高斯分布,则结节梯度的法向量方向是指向四面八方的,服从均匀分布,血管梯度的法向量方向是沿血管流向指向的,服从正态分布,如图 4(b)所示。

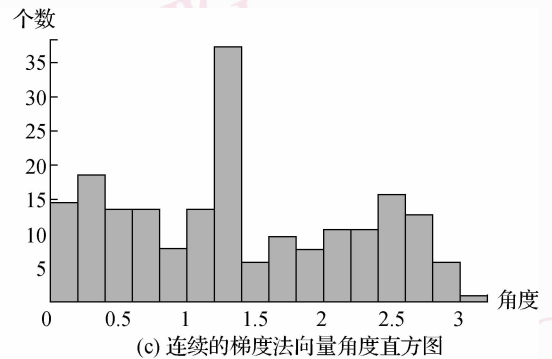
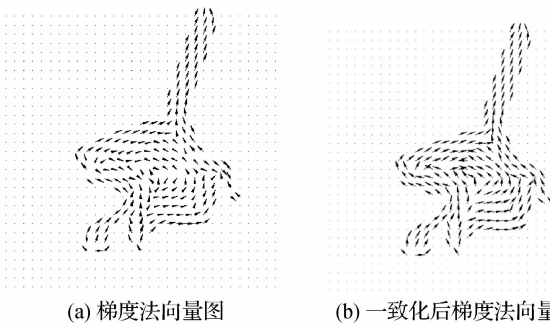


图 4 流向特征方向图

Fig. 4 Flow feature direction figure

因此,建立粘连血管型结节的模型如下:

$$f(x) = \omega_1 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-u)^2}{2\sigma^2}\right) + \frac{\omega_2}{\pi} \quad (6)$$

式中, ω_1 和 ω_2 为权重, 并有 $\omega_1 + \omega_2 = 1$, u 和 σ 分别为正态分布的均值和方差。

EM 是一种无监督参数估计的方法^[8], 用 EM 估计模型参数时需要首先给出模型参数的初始值, 而其估计出的结果对初始值敏感, 所以, 本文先用 K-means 聚类算法 (初始聚类数 $C = 2$), 求出类均值、方差和权重作为 EM 估计的初始值。

EM 是针对噪声数据集或数据不完整的最大似然估计方法。对于数据集 X 的对数似然函数:

$$\begin{aligned} L(\Theta | X) &= \log \prod_{i=1}^N p(x_i | \Theta) \\ &= \sum_{i=1}^N \log \left(\omega_1 \times \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_i - u_1)^2}{2\sigma_1^2}\right) + \frac{\omega_2}{\pi} \right) \end{aligned} \quad (7)$$

找到使 $(L(\Theta) | X)$ 最大的 Θ 。 $\Theta^{(s)}$ 是参数 Θ 在第 s 次迭代的估计值, 后验期望概率为

$$\begin{aligned} p(l = 1 | x_i, \Theta^{(s)}) &= \frac{\omega_1^{(s)} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_i - u_1^{(s)})^2}{2\sigma_1^{(s)2}}\right)}{\omega_1^{(s)} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_i - u_1^{(s)})^2}{2\sigma_1^{(s)2}}\right) + \omega_2^{(s)} \frac{1}{\pi}} \end{aligned} \quad (8)$$

$$\begin{aligned} p(l = 2 | x_i, \Theta^{(s)}) &= \frac{\omega_2^{(s)} \frac{1}{\pi}}{\omega_1^{(s)} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_i - u_1^{(s)})^2}{2\sigma_1^{(s)2}}\right) + \omega_2^{(s)} \frac{1}{\pi}} \end{aligned} \quad (9)$$

参数估计公式为

$$\omega_1^{(s+1)} = \frac{1}{N} \sum_{i=1}^N p(l = 1 | x_i, \Theta^{(s)}) \quad (10)$$

$$\omega_2^{(s+1)} = \frac{1}{N} \sum_{i=1}^N p(l = 2 | x_i, \Theta^{(s)}) \quad (11)$$

$$u_1^{(s+1)} = \frac{\sum_{i=1}^N x_i \times p(l = 1 | x_i, \Theta^{(s)})}{N\omega_1^{(s+1)}} \quad (12)$$

$$\begin{aligned} \sigma_1^{(s+1)} &= \sqrt{\frac{\sum_{i=1}^N p(l = 1 | x_i, \Theta^{(s)}) (x_i - u_1^{(s+1)}) (x_i - u_1^{(s+1)})}{N\omega_1^{(s+1)}}} \end{aligned} \quad (13)$$

根据均匀分布的权重 ω_2 , 可以确定结节像素个数占整个粘连血管型结节像素个数的比例, 也就是说, 可以预测待分割物体中所含结节的像素个数。

3 基于 Mean-shift 的结节精确分割

Mean-shift 聚类算法是一种基于 Parzen 窗技术的非参数核密度估计方法。它是寻找窗内密度极大值像素 (也称模式点) 的过程。对于每个样本点, 迭代 Mean-shift 过程, 收敛到相同模式点的像素可认为属于同一聚类簇。

在 Mean-shift 算法中, Parzen 窗的尺寸称作带宽。不同特征有不同的带宽参数 H 。它是一个重要的参数, 它对分割结果起着重要的作用。如果带宽 H 选取太小, 则偏移量小, 算法收敛速度很快, 从而导致过分割现象; 如果带宽 H 选取太大, 则偏移量大, 算法收敛速度很慢, 从而导致欠分割现象。目前, 常用的带宽计算方法有: 快速算法、交叉验证法和插入规则法。然而, 这些方法大多依赖于未知密度的拉普拉斯变换或在计算偏差时引入高阶导数, 都不适合解决实际问题。

3.1 新的带宽参数选择方法

带宽选择定理^[9]: 假设样本服从正态分布 $N(u, \Sigma)$, 采用正态核函数进行均值漂移, 如果带宽 $H = \Sigma$, 则均值漂移向量 $m(x)$ 的模取最大值。

以 c 为中心, 尺寸为 $r_{xy} \times r_{xy} \times r_z$ 的 3D ROI CT 数据的特征空间 $X_i = (x_i, y_i, z_i, r_i) \in \mathbf{R}^4$, 其中 $r_i = I(x_i, y_i, z_i)$, x_i, y_i, z_i 为空间位置特征, r_i 为灰度特征。根据均匀分布的权重 ω_2 , 可知 MIP 的前景区中结节的像素个数是 $N_{uu} \times \omega_2$ 。结节在 x, y 方向上的跨度是 $S_{XY} = \text{int}(\sqrt{N_{uu} \times \omega_2})$, 在 z 方向上的跨度是 $S_Z = \text{round}(S_{XY} \times x_{mm}/z_{mm})$ 。在 r 方向上的跨度是 $S_R = CT_{\max} - CT_{\min}$, CT_{\max}, CT_{\min} 分别为 N_{uu} 中像素的最大和最小 CT 值。假设结节服从各向异性的正态分布, x 方向的带宽参数 h_x, y 方向上的带宽参数 h_y, z 方向上的带宽参数 h_z 和灰度的带宽参数 h_r , 根据带宽选择定理当结节区域在 x 方向上 (或 y, z 方向上或灰度) 的直径为 d 时, x 方向上 (或 y, z 方向上或灰度) 的均值漂移带宽应为 $d/4$, 这是因为服从正态分布 $N(u, \sigma)$ 物体的直径等于 4σ ,

$$\begin{cases} h_x = S_{XY}/4 \\ h_y = S_{XY}/4 \\ h_z = S_z/4 \\ h_r = S_R/4 \end{cases} \quad (14)$$

3.2 Mean-shift 分割结节过程

对于 $r_{xy}^k \times r_{xy}^k \times r_z^k$ 的 3D ROI 数据集 $X_i \in \mathbf{R}^4$, $i = 1, \dots, N_{wu}$ 则样本点 x 处的多变量高斯核概率密度估计函数为^[10]

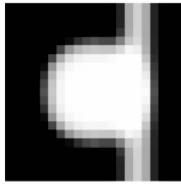
$$\hat{f}(X) = \frac{1}{N_{wu} (2\pi)^2 |H|^{1/2}} \sum_{i=1}^{N_{wu}} \exp\left(-\frac{1}{2} D^2(X, X_i, H)\right)$$

$$D^2(X, X_i, H) = (X - X_i)^T H^{-1} (X - X_i)$$

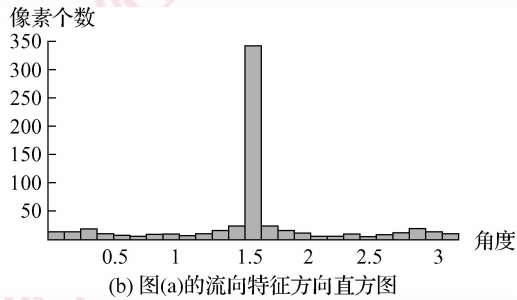
$$H = \text{diag}(h_x, h_y, h_z, h_r) \quad (15)$$

特征空间中,密度的极大值点为聚类中心点。也就是,密度函数梯度等于零的点,

$$\nabla \hat{f}(X) = \frac{1}{N_{wu} (2\pi)^2} \sum_{i=1}^{N_{wu}} \frac{H^{-1} (X_i - X)}{|H|^{-1/2}} \exp\left(-\frac{1}{2} D^2(X, X_i, H)\right) \quad (16)$$



(a) 血管粘连型结节人造数据MIP图像



(b) 图(a)的流向特征方向直方图

图 5 人造血管粘连型结节 MIP 图像及流向特征方向直方图

Fig. 5 Maximum intensity projection of synthetic nodule which connected vessel and flow feature orientation histogram

4.2 算法在 CT 结节数据上的评估

评估实验是在 $0.6\text{mm} \times 0.6\text{mm} \times 2\text{mm}$ 的临床多层肺 CT 影像上进行的,包含 19 个病例中的 19 个粘连血管型结节,其结果全部符合医生手工勾画出

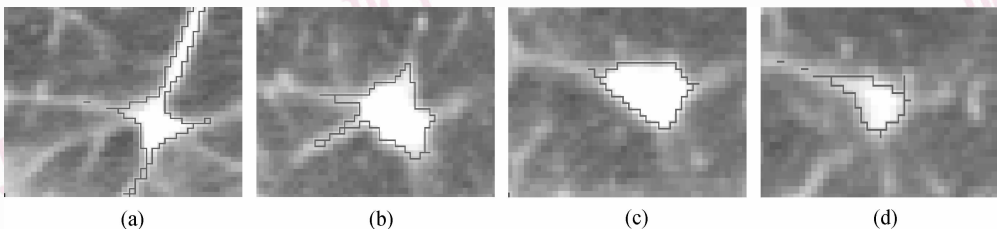


图 6 区域生长算法分割图 2 结节各层结果图

Fig. 6 Segmentation figure 2 by region grow

$$m(X) = \frac{\sum_{i=1}^{N_{wu}} X_i \exp\left(-\frac{1}{2} D^2(X, X_i, H)\right)}{\sum_{i=1}^{N_{wu}} \exp\left(-\frac{1}{2} D^2(X, X_i, H)\right)}$$

是 Mean-shift 的偏移向量。

Mean-shift 向量总是指向密度增长最快的方向。Mean-shift 过程通过迭代下面两步,保证了收敛到聚类中心:

- (1) 计算 Mean-shift 向量 $m(X)$ 。
- (2) 平移窗口,窗口中心位置移到 $X^{t+1}, X^{t+1} = X^t + m(X^t)$ 。

对于每个数据点,运行上述过程,收敛到相同聚类中心的点的集合被认为属于同一个聚类簇。

4 实验仿真

4.1 算法在仿真数据上的评估

一个半径为 6,尺度为 1 的 3D 高斯球和一个半径为 2,高度为 21,尺度为 1 的 3D 高斯圆柱相叠加模拟与血管粘连型结节。MIP 后如图 5(a)所示,向量方向直方图如图 5(b)所示。

的边缘。图 6 是区域生长算法对图 2 结节的各层分割结果。图 7 是本文算法分割图 2 结节的各层结果。两者比较可以看出,本文算法能有效地把结节和血管分割开。

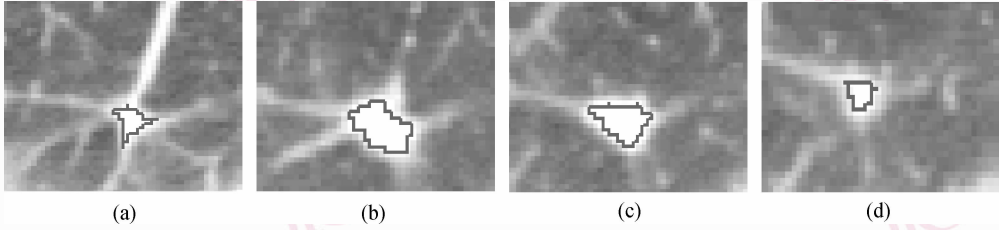


图 7 本文算法分割图 2 结节各层结果

Fig. 7 Segmentation figure 2 by the proposed method

本文算法分割图 8 中的 4 个结节带宽参数选择范围与文献[3]~[4]中算法分割这些结节时带宽选择范围的对比如表 1 所示,图中绿色表示结节,红色表示血管。可以看出,本文算法提出了一个直接

选择带宽参数的新方法而不同于文献[3]~[4]用迭代最优化的方法在带宽范围内选出最佳的带宽参数。

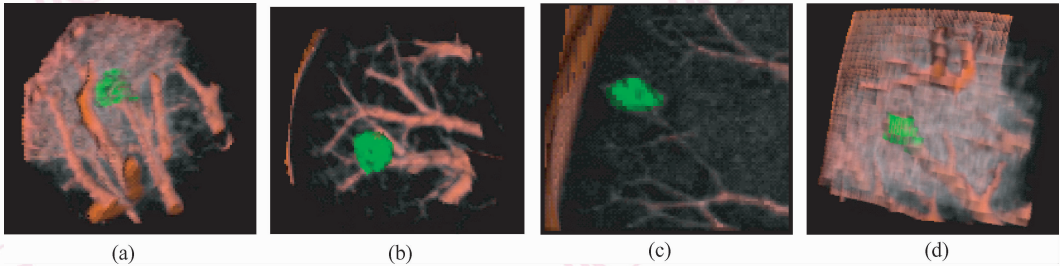


图 8 4 个粘连血管型结节 3D 可视化的结果

Fig. 8 3D visualization of the four nodules connected vessel

表 1 带宽选择范围的对比

Table1 Comparison band-width chosen range

图像	文献[3]算法	文献[4]算法	本文算法
图 8(a)	0.5 ~ 13	2.75 ~ 9.25	3
图 8(b)	0.5 ~ 13	3.0 ~ 9.75	3.25
图 8(c)	0.5 ~ 13	3.0 ~ 10.25	3.25
图 8(d)	0.5 ~ 13	2.75 ~ 9.25	3

分割,在时间和算法的复杂度上比目前现有方法都有了大幅度的改进。肺结节形态各异,除了粘连血管型肺结节外还包括孤立型肺结节、粘连肺壁型肺结节和低密度型肺结节,其中,前两种类型结节用区域生长方法可以得到精确的分割效果,而低密度型肺结节的分割是个难点,需要以后做进一步地研究。

参考文献 (References)

- 1 Kostis W J,Reeves A P, Yankelewitz D F, *et al.* Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images [J]. IEEE Transactions on Medical Imaging, 2003, **22**(10): 1259-1274.
- 2 Ning X, Narendra A, Ravi B. Automated lung nodule segmentation using dynamic programming and EM based classification [J]. SPIE Medical Imaging, 2002, **4684**(5): 666-676.
- 3 Okada K, Comaniciu D, Krishnan A. Robust anisotropic Gaussian fitting for volumetric characterization of pulmonary nodules in multislice CT [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, **25**(2): 281-288.
- 4 Sun Shen-shen, Li Hong, Hou Xin-ran, *et al.* Automatic segmentation

5 结 论

提出了一种新的带宽自动选取均值漂移算法。与基于统计分析规则的带宽选择方法和基于最优化迭代的带宽选择方法相比,该方法能直接求得正确带宽参数,且时间复杂度低。针对灰度特征不能区分结节和血管的问题,仿真数据实验和临床 CT 数据实验表明,提出的梯度法向方向分布特征是有效的,在几秒钟内就能实现血管粘连型肺结节的精确

- of pulmonary nodules in CT images [A]. In: Proceedings of the IEEE International Conference on 1st Bioinformatic and Biomedical Engineering [C], Wuhan, 2007: 790-793.
- 5 Li Q, Sone S, Doi K. Selective enhancement filters for nodules, vessels, and airway walls in two-and three-dimensional CT scans [J]. *Medical Physics*, 2003, **30**(8): 2040-2051.
- 6 Gady A, Samuel G, Armato I I I, *et al.* Vessel tree reconstruction in thoracic CT scans with application to nodule detection [J]. *IEEE Transactions on Medical Imaging*, 2005, **24**(4): 486-499.
- 7 Gady A, Changhua W. Probabilistic modeling based vessel enhancement in thoracic CT scans [A]. In: Proceedings of International Conference on Computer Vision and Pattern Recognition [C], San Diego, CA, USA, 2005, **2**: 684-689.
- 8 Jun Z, James W M, Fellow, *et al.* Maximum-likelihood parameter estimation for unsupervised stochastic model-based image segmentation [J]. *IEEE Transactions on Image Processing*, 1994, **3**(4): 404-420.
- 9 Comaniciu D. An algorithm for data-driven bandwidth selection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(2): 281-288.
- 10 Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(5): 603-619.