

中图法分类号: TP391.41 文献标志码: A 文章编号: 1006-8961(2011)08-1402-06

论文索引信息: 刘阳, 赵大哲, 刘积仁. 基于3维图像增强的肺结节识别 [J]. 中国图象图形学报, 2011, 16(8): 1402-1407

基于3维图像增强的肺结节识别

刘阳^{1),2)}, 赵大哲¹⁾, 刘积仁²⁾

¹⁾(东北大学医学影像计算教育部重点实验室, 沈阳 110004) ²⁾(东北大学信息科学与工程学院, 沈阳 110004)

摘要: 孤立性肺结节的检测是肺癌早期诊断的关键。针对传统点增强滤波器虽然对结节增强具有很好的敏感性, 但是却产生很多假阳性区域的问题, 提出一种通过计算3维增强密度指数和判别规则来识别肺结节的方法。首先采用自适应双边滤波器对CT图像序列进行降噪和平滑处理; 然后计算对应的Hessian矩阵及其特征值得到预增强系数, 并获得感兴趣区域, 通过对预增强系数的分析来构造3维增强密度指数; 最后应用判别规则对感兴趣体进行识别。针对两个肺部CT图像数据集对该方法进行了测试, 结果表明, 在识别孤立性肺结节方面该方法是有有效的。

关键词: 点增强; 自适应双边滤波; Hessian矩阵; 增强密度指数

Recognition of pulmonary nodules based on 3D image enhancement

Liu Yang^{1),2)}, Zhao Dazhe¹⁾, Liu Jiren²⁾

¹⁾(Northeastern University Key Laboratory of Medical Image Computing,
Ministry of Education, Shenyang 110004 China)

²⁾(School of Information Science & Engineering, Northeastern University, Shenyang 110004 China)

Abstract: The detection of solitary pulmonary nodules (SPN) is proven to be of critical importance in early-stage lung cancer diagnosis. Aiming at reducing the false positive regions caused by the dot enhancement filter which is sensitive to the lung nodules, a new recognition method based on calculation of three-dimensional (3D) enhancement density index and decision rule is proposed. An adaptive bilateral filter is applied to reduce the noisy and smooth CT image sequences. Then, the pre-enhancement coefficients and volume of interest (VOI) are obtained by computing the Hessian matrix and corresponding eigenvalues. After analyzing the distribution of pre-enhancement coefficients, 3D enhancement density index is constructed. Finally, a decision rule is adopted to identify nodule candidates. The proposed method is tested on two lung CT image sets. The experimental results illustrate the efficiency of the proposed algorithm.

Keywords: dot enhancement; adaptive bilateral filter; Hessian matrix; enhancement density index

0 引言

卫生部最新资料显示, 过去30年间, 肺癌死亡率在我国上升了465%, 已取代肝癌成为我国首位恶性肿瘤死亡原因。有研究指出, 肺癌的早期检测

和治疗可以提高肺癌5年存活率^[1]。孤立性肺结节(SPN)被定义为孤立的肺实质(组织)病变, 直径不大于3cm, 轮廓近似为球体, 是早期发现肺癌的重要线索。早期1肺癌(1期)的癌结节手术切除有良好的预后, 其5年生存率可达75%, 且肿瘤越小、发现越早, 生存率越高。因此, 减少肺癌死亡率, 肺结

收稿日期: 2010-05-07; 修回日期: 2010-09-28

基金项目: 国家自然科学基金项目(60671050, 61001047)。

第一作者简介: 刘阳(1981—), 女。东北大学信息科学与工程学院计算机应用技术博士研究生, 主要研究方向为医学图像处理。E-mail: ly.liuy@neusoft.com。

节的早期检测很关键。

近年来,CT成像被认为是对肺癌早期检测手段中最有前途的技术之一。可对于放射医生而言,随着CT成像分辨率的不断增加,从大量CT切片图像数据中检测微小的肺癌区域是一个困难并且耗时的任务。因此,计算机辅助检测(CAD)肺结节方法将会为放射医生提供帮助,也成为医学图像处理领域的一个研究热点。

目前,国内外学者在这一应用领域展开了广泛研究,针对肺部CT图像中肺结节的提取,提出了多种识别方法。其中,利用滤波器对结节区域进行增强是比较典型的方法之一。Shigemoto等人^[2]结合3维结节模型提出了一种模板匹配方法来进行结节增强,但是该方法需要事先提供多种尺寸、多个方向的匹配模板。文献[3-4]提出了基于3维高斯模型的结节增强检测算法,不过仍需先构造不同尺寸的匹配模板。由于在3维CT图像中,肺结节在形态上具有多尺寸、多尺度以及似球形等特征,因此,Li等人^[5]提出了基于3维Hessian矩阵的多尺度点增强滤波器来对结节区域进行增强操作。该方法回避了需要准备多个匹配模板的过程,且有效地去除了血管等形态组织的干扰,但是对于血管交叉点等体素分布类似球体的区域抑制效果不是很好,依然会产生很多的假阳性区域。

本文提出一种基于增强密度指数来进行肺结节检测的方法。首先采用改进的3维点增强滤波方法,得到结节预增强系数,并通过3维区域增长方法对结节候选区进行标识;然后计算增强密度指数,得到预增强区域的增强密度指数;最后通过判定规则去除假阳性区域。

1 提取预增强结节区域

为表述方便,首先做如下约定:令 $g(x,y,z)$ 代表3维CT图像序列,其中 (x,y,z) 为体素坐标点位置, $g(x,y,z)$ 为点 (x,y,z) 的灰度值;定义序列点类别集合 $L = \{nd, bg\}$, nd 代表结节点, bg 代表背景点。为实现图像区域各向具有相同尺度,本文方法均针对已完成插值,具有各向同性的图像序列进行。

首先对CT图像中的噪声点进行降噪处理,同时对肺内软组织进行平滑操作,以提升结节区域特异性。双边滤波器(bilateral filtering)是一种非线性、非迭代、简单的滤波算法,在平滑图像的同时具有很好

的边缘保持特性,在处理相邻各像素值时,不仅考虑到距离上的邻近关系,同时考虑灰度上的相似性,通过对二者的非线性组合,自适应滤波。因此,选用3维双边滤波器对CT图像序列进行3维滤波操作。

对图像序列 g 进行时空双边滤波的加权系数可以表示为

$$W(\xi, \eta, \varphi; x, y, z) = W_D((\xi, \eta, \varphi), (x, y, z)) \cdot W_R(g_{\xi, \eta, \varphi}, g_{x, y, z})$$

式中, $W_D((\xi, \eta, \varphi), (x, y, z))$ 和 $W_R(g_{\xi, \eta, \varphi}, g_{x, y, z})$ 为标准3维高斯滤波函数,因此滤波函数为

$$W(\xi, \eta, \varphi; x, y, z) = K \cdot \exp\left\{-\frac{1}{2}\left[\left(\frac{\xi-x}{\sigma_x}\right)^2 + \left(\frac{\eta-y}{\sigma_y}\right)^2 + \left(\frac{\varphi-z}{\sigma_z}\right)^2\right]\right\} \cdot \exp\left[-\frac{(g_{\xi, \eta, \varphi} - g_{x, y, z})^2}{2\sigma_R^2}\right] \quad (1)$$

式中, σ_R 表示在窗口中图像通过高斯函数进行滤波时的灰度标准差, K 为归一化系数; $\sigma_x, \sigma_y, \sigma_z$ 为高斯函数标准差,其值越大高斯函数的能量就越集中,为保证滤波不改变结节区域的像素值分布趋势,选用各向同性3维高斯函数,即 $\sigma_x = \sigma_y = \sigma_z = \sigma$ 。令经过双边滤波器得到的图像序列为 GA 。

在进行肺结节提取时,对于算法性能产生干扰最大的肺内组织器官是血管和支气管。因此,需对肺部CT图像序列中的管状结构进行弱化处理,使得结节区域相对应的得以增强。本文采用文献[5]提出的方法,对 GA 序列求取其Hessian矩阵及对应特征值并进行分析,以得到序列点的局部特性。

$$H_p = \begin{pmatrix} g_{xx} & g_{xy} & g_{xz} \\ g_{yx} & g_{yy} & g_{yz} \\ g_{zx} & g_{zy} & g_{zz} \end{pmatrix} \quad (2)$$

式中, $g_{ab} = \frac{\partial^2 g}{\partial a \partial b}$ 。在 H_p 的3个特征值 $\lambda_1, \lambda_2, \lambda_3$ 中,幅值最大的特征值对应的特征向量代表 p 点曲率最大的方向,幅值的大小与在特征向量方向上的伸缩长度相对应。Hessian矩阵中,特征值小于零表示凸形,大于零表示凹形。肺结节在CT图像中相对于背景是明亮区域,为几何图像中的凸形,结合结节区域具有似球性的特征,对于Hessian矩阵的3个正交向量,需要关注在3个方向上均为凸形的点。令 $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$, 定义结节预增强系数

$$E_p = \frac{\eta(\lambda_1, \lambda_2, \lambda_3) |\lambda_3|^2}{|\lambda_1|} \quad (3)$$

在式(3)中,当 $\lambda_1 < 0 \wedge \lambda_2 < 0 \wedge \lambda_3 < 0$ 时,函数 $\eta(\lambda_1, \lambda_2, \lambda_3)$ 值为 1, 否则为 0。 E_p 反映了对结节候选区的增强程度。

然后,遍历序列图像中各体素点,通过 3 维区域增长方法,来对预增强感兴趣体进行标识。

2 计算增强密度指数

单纯通过预增强系数来进行结节候选区的判定,会产生比较多的假阳性干扰。针对真、假阳性区域的预增强系数值,本文随机选取 5 个带有结节标识金标准的肺部 CT 图像序列,在经过预增强后,对 3 927 个结节区域体素和 82 394 个假阳性区域体素进行预增强系数分析。这里,对预增强系数进行归一化处理,然后等步长划分为 20 个分析区域,对于每个区域中归一化的预增强系数值的出现频率进行分析,结果如图 1 所示。由于结节区域往往符合球体形状且像素值呈放射性衰减形式分布,所以通过增强方法得到的结节区域预增强系数值在高值区域分布较假阳性区域要更加广泛。在图 1 中为表现更多预增强系数在高值区的分布细节,略去了预增强系数较低的部分。

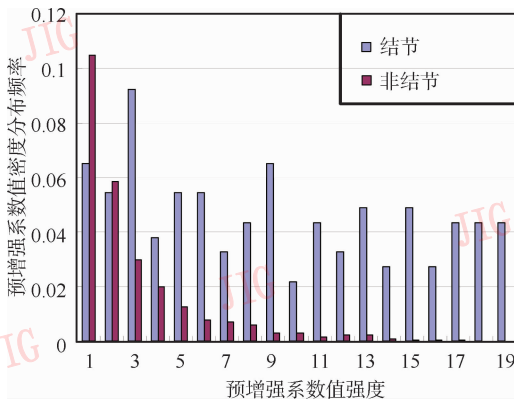


图 1 预增强系数频率分布柱状图

Fig. 1 Histogram of frequency distribution for pre-enhancement coefficient

图 1 反映出结节区域预增强系数值的分布,在高值区其分布密度要远高于假阳性区域值。但是,由于在预增强阶段,假阳性区域的数量要远大于结节区域,所以单纯通过阈值方法来去除预增强阶段产生的假阳性区域,效果并不明显。

如图 2 所示为结节区域和假阳性区域预增强系数网线图。通过对网线图的分析可以看出,在预增强系数值及分布形态方面,二者不具有明显的区分标志。但是如果将关注点放在一个小的 3 维局部区域中,二者的系数值分布会显著不同。

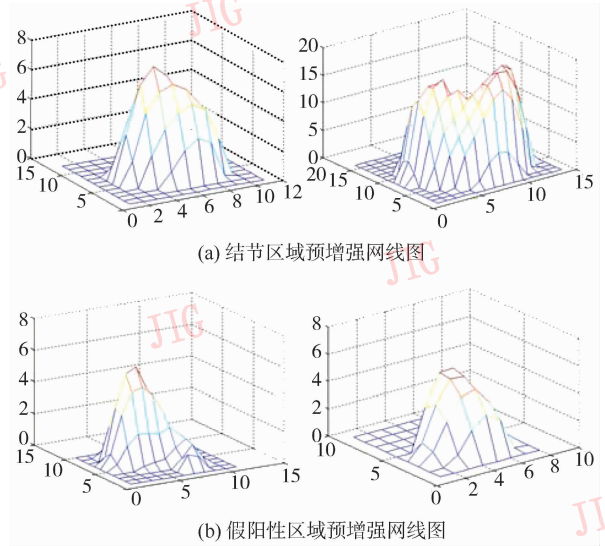


图 2 结节与假阳性区域预增强系数网线图

Fig. 2 Mesh images of pre-enhancement coefficients for nodule and false positive regions

因此,采用局部增强密度指数来对预增强系数进行分析。首先构造一个模板 Ω , Ω 为 $N \times N \times N$ 的 3 维区域,其中 $N = 2\sigma + 1$ 。令 $p(x, y, z)$ 为 Ω 的中心,在 Ω 内定义 p 点增强密度指数为

$$D_p = \frac{\sum_{q \in \Omega} B(q)}{S} \quad (4)$$

式中, S 为区域 Ω 中的体素数量,

$$B(q) = \begin{cases} 1 & E_q > t \\ 0 & \text{其他} \end{cases} \quad (5)$$

t 为预增强系数阈值。通过对预增强系数的分析,结节区域的预增强系数值在图像上要明显高于假阳性区域,那么在阈值选取时首先将背景点去掉,即仅考虑 $E_q (> 0)$ 的体素,在这些体素中,由于结节区域为明亮区域,可以被看做前景点,将假阳性区域视为背景点,这样预增强系数阈值 t 采用 OTSU 单阈值算法来自适应获取。

最后,根据实验结果,引入结节区域判定规则

$$g_p = \begin{cases} nd & D_p \geq 0.25 \\ bg & \text{其他} \end{cases} \quad (6)$$

可以看出,通过对局部增强密度指数的计算和

判定规则,可以排除预增强过程中产生的孤立凸形点,并且对于狭长凸形像素点起到了修剪作用。

整个算法基本步骤如下:

- 1) 使用式(1)对图像序列进行自适应双边滤波;
- 2) 逐像素计算 Hessian 矩阵(式(2))以及其对应特征值 $\lambda_1, \lambda_2, \lambda_3$;
- 3) 计算预增强系数(式(3));
- 4) 逐像素计算增强密度指数(式(4));
- 5) 利用判定规则(式(6))对结节区域像素进行识别。

由于本文方法采用多尺度自动计算,因此,最终需将多尺度识别结果进行叠加。

3 实验及结果

本文算法在真实含结节肺部 CT 图像序列中的应用效果图如图 3 所示:图(a)为原始含结节 CT 图像,箭头所指为金标准标注结节;图(b)为预增强系数值对应的灰度级图像,可以看出,结节区域被增强的同时,一些假阳性区域也随之出现;图(c)为经过增强密度指数分析和判定规则后,得到的肺结节识别结果。虽然本文算法得到的结节形态与原始结节形态略有差异,但是在尽可能减少假阳性区域的同时,结节主体区域被很好地保留下来。

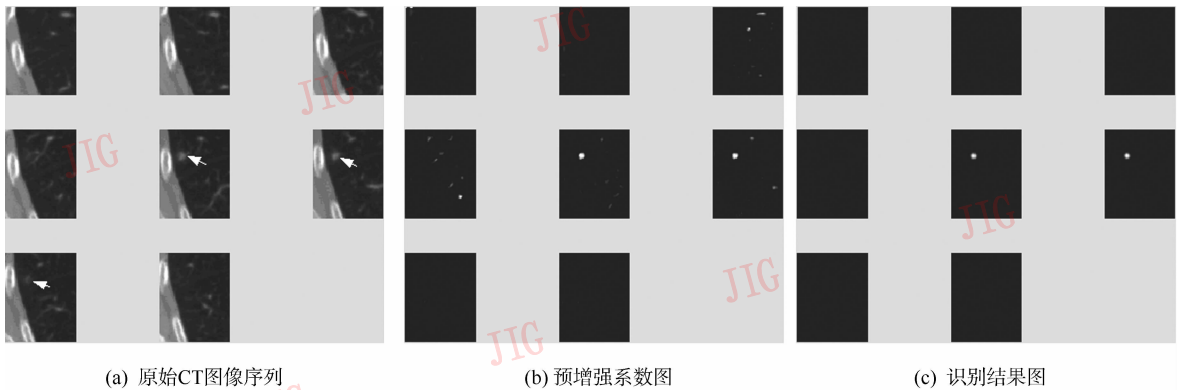


图3 本文方法识别结果图

Fig. 3 Recognition results by the proposed method

本文算法在 LIDC^[6] 数据中第 1 数据集和 NSRCT-LUNG 数据集(表 1)上进行了肺结节识别测试。使用的计算机配置为 CPU:AMD Phenom 9750 2.41 GHz × 4, 3.25 GB 内存,显卡 GF9800 GT;操作系统 Windows XP SP2。

表 1 实验数据集信息表

Tab. 1 Information table of experimental data set

数据集名称	数据集信息
LIDC	病例数:23
	SPN 数量:21
	肺结节尺寸:[4 mm, 14 mm]
NSRCT-LUNG	病例数:9
	SPN 数量:10
	肺结节尺寸:[2.5 mm, 19 mm]

为提高检测算法的运行效率,对实验数据首先进行了肺实质提取,然后使用插值技术获得在

各方向具有相同尺度的影像数据,本文的插值尺度为 0.5 mm/像素。在使用双边滤波器进行去噪平滑操作时,参数 σ 取值很重要,因为如果 σ 取值较小,那么高斯核频带就比较窄,信号边缘的定位精度较高,但平滑作用小,对噪声的抑制能力差;反之,当 σ 取值较大时,其对信号的平滑作用大,去噪声效果较好,但会使信号边缘变得模糊,出现严重的边缘移位现象。在 3 维 CT 图像中,如果肺结节半径为 σ ,那么结节占据了高斯核约 33% 的能量分布;如果结节半径为 2σ ,那么结节占据了约 87% 的能量分布;如果结节半径为 3σ ,那么可以占据约 99% 的能量分布。因此,对于半径为 r 的肺结节, $\sigma = r/3$ 为最佳。由于不同患者的肺结节在 CT 图像上的尺寸有所不同,为了在预增强阶段获得较高的检测敏感性,需要对所有尺寸的结节进行预增强,因此需要针对不同尺寸的结节采用不同的增强尺度,即引入多尺度增强方法。理

论上,尺度选择越多增强效果越好,但是会增加计算时间,提高检测代价,在权衡预增强敏感性和算法运行时间后,本文采用 5 个尺度对肺实质图像进行预增强,这 5 个尺度分别为 1 mm、2 mm、3 mm、5.5 mm 和 9 mm。然后,针对每一个尺度计算产生的预增强系数来计算增强密度指数,其中模板 Ω 的尺寸要与当前采用的增强尺度相匹配,这样每一个尺度所构造的增强密度指数对对应尺度的对象有最好的增强作用,多尺度方法在每一个像素的输出取 5 个增强指数输出的最大值。最后采用结节区域判定规则对肺结节区域进行标识。

ROC (free-response receiver operating characteristic) 曲线是用来评价一种模型有效性的方法,被广泛应用于肺结节识别算法评估。图 4 为本文识别方法的 ROC 评估曲线。

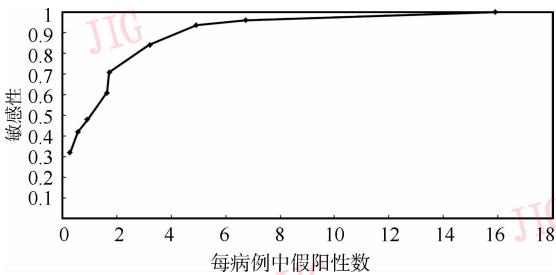


图 4 本文识别方法的 ROC 曲线

Fig. 4 Free-response receiver operating characteristic curve for our recognition method

由于本文算法基于 3 维肺部 CT 图像来进行,所以对比了其他 3 维结节识别算法的性能,比较结果如表 2 所示。可以看出 Paik 等人^[7]的实验数据集仅局限于大于 6 mm 的结节,没有提供对小尺寸结节的检测结果,而 Zhao 等人^[8]的算法实验是在模拟结节基础上进行的, Hara^[9]和 Osman 等人^[10]的算法取得了非常好的效果,但是 Osman 等人的实验数据集数据比较少,且结节形态尺寸分布比较相似,而 Hara 等人没有对其实验数据集进行详细描述,且结节尺寸分布限于小于 7 mm 的数据,所以无法进行同等级结果的比对。由图 4 中可以看出,在兼顾检测敏感性和假阳性率指标的基础上,本文算法在敏感性为 0.935 的情况下,假阳性率为 4.91 假阳性数/例 (FPs/case);如果将敏感性降低到 0.71,假阳性率为 1.72 FPs/case。

表 2 3 维结节识别算法性能对照表

Tab.2 Comparison table of various 3D nodule recognition methods

算法	实验数据集信息	性能
Paik 方法 ^[7]	8 例病例,未知结节数,结节尺寸大于 6 mm	敏感性:80%, 1.3 FPs/case 敏感性:90%, 5.6 FPs/case
Zhao 方法 ^[8]	266 个仿真结节	敏感性:84.2%, 5 FPs/case
Hara 方法 ^[9]	未知结节数,结节尺寸小于 7 mm	敏感性:94%, 2.05 FPs/case
osman 方法 ^[10]	6 例病例,6 个结节,结节尺寸 [3.5 mm, 7.3 mm]	敏感性:100%, 0.83 FPs/case
本文方法	32 例病例,31 个 SPN,结节尺寸 [2.5 mm, 19 mm]	敏感性:71.0%, 1.72 FPs/case, 敏感性:93.5%, 4.91 FPs/case

4 结 论

提出一种基于 3 维点增强滤波器的肺结节识别方法。与传统的点增强滤波器相比,该方法可以更加有效地抑制假阳性区域,使得识别结果更加准确。通过对肺部区域进行自适应双边滤波,来降低肺部 CT 图像中的噪声干扰,对血管等管状结构区域进行平滑弱化;进而计算其对应的 Hessian 矩阵并求取特征值加以分析,得到预增强系数;通过对预增强系数的分析,计算密度增强指数,然后利用判定规则,完成肺结节的识别。最后,针对真实肺部 CT 图像序列进行了本文方法的性能实验,结果表明了方法的有效性。需要指出的是,虽然本文方法不能精确分割肺结节区域,但是准确的识别结果可以为后续的分割提供有力的参考。

参考文献 (References)

- [1] Raymond E L, Robert T O, Ted G. Clinical Oncology [M]. 2st ed. Atlanta; USA; Emily Pualwan, 2001;269-296.
- [2] Shigemoto K, Takizawa H, Yamamoto S, et al. Efficient recognition method for lung nodule shadows in X-ray CT images using 3-D object models and template matching techniques [J]. Medical Imaging Technology, 2003, 21(2):147-156.
- [3] Lee Y, Hara T, Fujita H, et al. Automated detection of pulmonary nodules in helical CT images based on an improved templatematching technique [J]. IEEE Trans. Med. Imaging,

- 2001, 20(7):595-604.
- [4] Yamamoto M, Ishida T, Kawashita I, et al. Development of computer-aided diagnostic system for detection of lung nodules in three-dimensional computed tomography images[J]. Radiological Technology, 2006, 62(4):555-564.
- [5] Li Q, Sone S, Doi K. Selective enhancement filters for nodules, vessels, and airway walls in two- and three-dimensional CT scans [J]. Med. Phys. , 2003, 30(8):2040-2051.
- [6] Samuel G, Armato III, Geoffrey M, et al. Lung image database consortium-developing a resource for the medical imaging research community[J]. Radiology, 2004, 232(9):739-748.
- [7] Paik S, Beaulieu F, Rubin D, et al. Surface normal overlap: a computer-aided detection algorithm with application to colonic polyps and lung nodules in helical CT[J]. IEEE Trans. Med. Imaging, 2004, 23(6):661-675.
- [8] Zhao B, Gordon G, Michelle G, et al. Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm [J]. Journal of Applied Clinical Medical Physics, 2003, 4(3):248-260.
- [9] Hara T, Hirose M, Zhou X, et al. Nodule detection in 3D chest CT images using 2nd order autocorrelation features [C]// Proceedings of the 27th Annual International Conference of the Engineering in Medicine and Biology Society. Piscataway, USA: IEEE, 2005:6247-6249.
- [10] Osmana O, Ozekesa S, Osman N. Lung nodule diagnosis using 3D template matching[J]. Computers in Biology and Medicine, 2007, 37(8):1167-1172.