

ROC and LROC analysis of 3D CT performance for lesion detection

SHI Cheng, XING Yu-xiang*

(Key Laboratory of Particle and Radiation Imaging of the Ministry of Education, Department of Engineering Physics, Tsinghua University, Beijing 100084, China)

[Abstract] ROC curve is introduced to evaluate X-ray CT systems and to assess the detection performance by quantitative analysis. There are some limitations with traditional ROC method, since it only focuses on signal known exactly (SKE) situation. LROC method, which considers the location of defect, is closer to the actual situation. Both kinds of results were got by using ROC and LROC methods, and were compared with the physic meaning of ROC and LROC model. In general, LROC method was better than ROC, and the assumption of LROC model did not hold in some cases. There were also some limitations for LROC in practice. In present, combining ROC and LROC method is the way to get a nice result.

[Key words] ROC curve; Localization; Tomography, X-ray computed

用 ROC 和 LROC 方法观察三维 CT 成像识别缺陷的性能

施 成, 邢宇翔*

[清华大学工程物理系 粒子技术与辐射成像教育部重点实验室(清华大学), 北京 100084]

[摘 要] 将 ROC 引入 CT 检测系统来对缺陷识别的效果进行定量分析。传统的 ROC 方法由于只针对缺陷位置固定的情形, 存在自身的局限性, 而考虑了缺陷定位的 LROC 方法更符合实际。本文基于两种方法获取的实验数据进行研究, 并结合 LROC 模型和 ROC 模型的物理意义进行比较分析。结果表明, 经综合比较, LROC 方法优于 ROC 方法, 但 LROC 模型的假设前提在一些情形下并不成立, 因此在实际应用中也存在自身的局限性。目前来看, 将 ROC 实验和 LROC 实验结合使用, 是提高诊断效果的最优方案。

[关键词] ROC 曲线; 定位; 体层摄影术, X 线计算机

[中图分类号] R814.42 **[文献标识码]** A **[文章编号]** 1003-3289(2012)02-0361-06

CT 断层重建已在医疗和工业领域得到广泛应用。在 CT 成像过程中, 采样不可避免地受到采样率、滤波函数、射束硬化、噪声等影响, 使物体的断层结构无法得到精确还原, 易致漏诊、误诊。

ROC 方法^[1]以每一个检测结果作为诊断界值

(cut-off point), 计算得到真阳性率(true positive fraction, TPF)和假阳性率(false positive fraction, FPF), 以 FPF(即特异度)为横坐标、TPF(即敏感度)为纵坐标绘制曲线, 称 ROC 曲线。ROC 方法并不考虑缺陷位置, 但在实际工业检测和医疗诊断中, 通常需要定位缺陷。如果采用经典 ROC 模型, 在检测中会出现实验结果和实际背离的情形, 所以需要针对缺陷位置可变建立模型。本研究采用定位 ROC(localization receiver operating characteristic, LROC)方法, 对 ROC 曲线进行改良, LROC 曲线以定位正确的 TPF 为纵坐标, 以 FPF 为横坐标, 得到 LROC 曲线。

[基金项目] 国家自然科学基金(60772051)。

[作者简介] 施成(1986—), 男, 福建福清人, 硕士。研究方向: CT 系统性能评估。E-mail: shic04@mails.tsinghua.edu.cn

[通讯作者] 邢宇翔, 清华大学工程物理系 粒子技术与辐射成像教育部重点实验室(清华大学), 100084。E-mail: xingyx@mail.tsinghua.edu.cn

[收稿日期] 2011-05-08 **[修回日期]** 2011-10-18

本研究分别使用 ROC 和 LROC 方法观察三维 CT 成像在缺陷识别方面的性能,对两种方法得到的结果进行对比分析,并对其适用性和有效性进行讨论。

1 双正态 ROC 模型

双正态模型^[2-3]目前在 ROC 分析中最为常用,该法假设正常组和异常组的实验结果均符合正态分布,然后根据实验结果拟合 ROC 曲线。

假设异常组和正常组的测试值 x 的分布分别满足以下表达式,其中参数 a, b 表示两组分布的关系:

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \exp\left[-\frac{x^2}{2}\right] && \text{正常组} \\ g(x) &= \frac{b}{2\pi} \exp\left[-\frac{(bx-a)^2}{2}\right] && \text{异常组} \end{aligned} \quad (1)$$

以 5 分类法为例。对实验材料进行分类,将相邻分类材料间的截断点依次记为 $\eta_1, \eta_2, \eta_3, \eta_4, \eta_5$ 的值越大,诊断阳性的可信度越高。由此异常组被分类至第 i 类的概率为:

$$P_{i/p} = G(\eta_i) - G(\eta_{i-1}) \quad (2)$$

其中 $G(\cdot)$ 为式(1)中异常组对应的概率分布函数。正常组被分类至第 i 类的概率为:

$$P_{i/n} = F(\eta_i) - F(\eta_{i-1}) \quad (3)$$

其中 $F(\cdot)$ 为式(1)中正常组对应的概率分布函数。式(2)、(3)中下标“ p ”“ n ”分别表示“异常”和“正常”,以下同。正常组中总例数记为 N_n ,在第 i 个分类里,该分类的频数为 $k_i, \sum_{i=1}^5 k_i = N_n$;异常组中总例数记为 N_p ,在第 i 类的频数为 $l_i, \sum_{i=1}^5 l_i = N_p$ 。

正常组和异常组资料得到的综合似然函数为:

$$L(k_{i=1, \dots, 5}, l_{i=1, \dots, 5} | a, b, \eta_{i=1, \dots, 4}) = (P_{1/n})^{k_1} (P_{2/n})^{k_2} \dots (P_{5/n})^{k_5} (P_{1/p})^{l_1} (P_{2/p})^{l_2} \dots (P_{5/p})^{l_5} \quad (4)$$

转化为对数形式为:

$$\ln L = \sum_{i=1}^5 k_i \ln(P_{i/n}) + \sum_{i=1}^5 l_i \ln(P_{i/p}) \quad (5)$$

最大化此对数似然函数,可得到 6 个参数 ($a, b, \eta_{i=1, \dots, 4}$) 的估计值。绘制对应的 ROC 曲线,得到其曲线下面积(area under curve, AUC):

$$AUC = \Phi\left(\frac{a}{\sqrt{1+b^2}}\right) \quad (6)$$

其中 $\Phi(\cdot)$ 表示标准正态分布的概率分布函数。AUC 是 ROC 实验的重要诊断指标,较高的 AUC 值表示诊断效果较好。

2 局部正态 LROC 模型

LROC 是针对缺陷定位而提出的方法。有多种不同方法能够获取 LROC 曲线^[4-6],其中应用较广泛的是 Swensson 模型^[4],称为“局部正态 LROC 模型”。局部正态 LROC 模型假设对图像上存在缺陷和不存在缺陷的位置的观察结果分别对应一个正态分布。

正常组或异常组的无缺陷位置满足标准正态分布:

$$f(x) = \frac{1}{2\pi} \exp\left[-\frac{x^2}{2}\right] \quad (7)$$

而对异常组中有缺陷位置满足如下分布:

$$g(x) = \frac{b}{2\pi} \exp\left[-\frac{(bx-a)^2}{2}\right] \quad (8)$$

同样以 5 分类法实验为例,将相邻分类材料间的截断点依次记为 $\eta_1, \eta_2, \eta_3, \eta_4, \eta_5$ 的值越大,诊断阳性的可信度越高。

正常组中总例数记为 N_n ,在第 i 类的频数记为 $N_{i/n}, \sum_{i=1}^5 N_{i/n} = N_n$;异常组中被正确定位的总例数记为 N_c ,在第 i 个分类里,该分类的频数为 $N_{i/c}$;异常组中被错误定位的总例数记为 \tilde{N}_c ,在第 i 个分类里,该分类的频数为 $\tilde{N}_{i/c}$,异常组的总例数为 $N_p = N_c + \tilde{N}_c$ 。正常组第 i 类的概率可表示为:

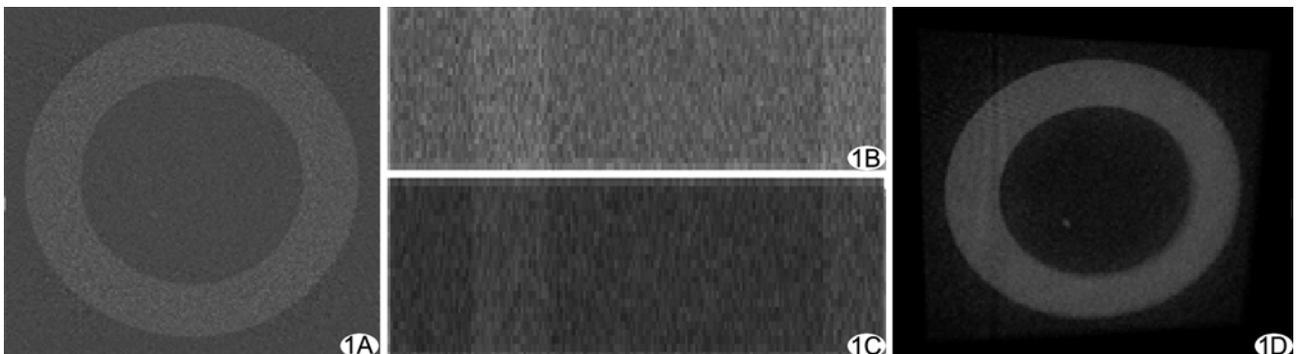


图 1 ROC 实验使用的三维 CT 图像在 3 个互相垂直方向上的切平面图(A~C)和 3D 视图(D)

$$P_{i|n} = \text{prob}(x_{\max} < \eta_i) - \text{prob}(x_{\max} < \eta_{i-1}) = F(\eta_i) - F(\eta_{i-1}) \quad (9)$$

综合所有类并进行对数变换后得到对数似然函数:

$$\ln(L_n) = N_{i|n} \times \ln[F(\eta_i)] + \sum_{i=2}^4 N_{i|n} \times \ln[F(\eta_i) - F(\eta_{i-1})] + N_{5|n} \times \ln[1 - F(\eta_1)] \quad (10)$$

异常组正确定位的图像第类的概率为:

$$P_{i|c} = \text{Prob}\{x_i > x_{\max} \& x_i > \eta_i\} = \int_{\eta_i}^{\infty} g(x) F(x) dx \quad (11)$$

异常组错误定位的图像第类的概率为:

$$\tilde{P}_{i|c}(x_i) = \text{Prob}\{x_{\max} > x_i \& x_{\max} > \eta_i\} = \int_{\eta_i}^{\infty} G(x) f(x) dx \quad (12)$$

综合式(11)、(12),统计所有类的情况,得到异常组的对数似然函数:

$$\begin{aligned} \ln(L_p) = & N_{1|c} \times \ln\left[\int_{-\infty}^{\eta_1} g(x) F(x) dx\right] + \sum_{i=2}^4 N_{i|c} \\ & \times \ln\left[\int_{\eta_{i-1}}^{\eta_i} g(x) F(x) dx\right] + N_{5|c} \\ & \times \ln\left[\int_{\eta_1}^{\infty} g(x) F(x) dx\right] \\ & + \tilde{N}_{i|c} \times \ln\left[\int_{-\infty}^{\eta_1} G(x) f(x) dx\right] + \sum_{i=2}^4 \tilde{N}_{i|c} \\ & \times \ln\left[\int_{\eta_{i-1}}^{\eta_i} G(x) f(x) dx\right] + \tilde{N}_{5|c} \\ & \times \ln\left[\int_{\eta_1}^{\infty} G(x) f(x) dx\right] \end{aligned} \quad (13)$$

综合正常和异常组的情况,得到全局似然函数:

$$\ln(L) = \ln(L_n) + \ln(L_p) \quad (14)$$

最大化式(14)中的似然函数,得到该模型下的参数估计,由这些参数可以分别绘制 ROC 曲线和 LROC 曲

线,其中的 LROC 的 AUC 为:

$$\text{AUC} = \Phi\left[\frac{a}{\sqrt{1+b^2}}\right] \quad (15)$$

局部正态 LROC 模型和双正态 ROC 模型的假设前提存在差异,会对最后的结果产生影响,将通过具体实验结合分析。

3 三维 CT 成像的 ROC 和 LROC 实验

在对 ROC 模型和 LROC 模型的研究中,主要关注以下两个问题:①两种模型在物理意义上的差异带来的影响;②双正态 ROC 模型和局部正态 LROC 模型的适用范围。

为方便描述,双正态 ROC 模型下,称缺陷位置已知的 ROC 曲线为 ROC-SKE(signal known exactly), AUC 记为 AUC_{R-SKE} ,缺陷位置可变情况下的 ROC 曲线表示为 ROC-SPV(signal position variable), AUC 记为 AUC_{R-SPV} ;在局部正态 LROC 模型下,得到的 LROC 曲线记为 L-LROC, AUC 为 AUC_{L-L} ,同时得到的 ROC 曲线记为 L-ROC, AUC 记为 AUC_{L-R} 。

3.1 圆轨道锥束 CT 成像的缺陷检测实验 针对不同噪声下圆轨道锥束 CT 成像缺陷检测^[7]进行研究。先对物体进行数值投影 $\bar{g}(\beta, a, b)$,在投影域加入不同等级的高斯噪声^[7]:

$$\bar{g}(\beta, a, b) \sim \text{Gaussian}(\bar{g}(\beta, a, b), \alpha \sqrt{g(\beta, a, b)}) \quad (16)$$

其中 α 为控制噪声等级的参数,然后使用 FDK 方法重建,一个典型的样本重建结果如图 1 所示。实验中取 $\alpha=0.3 \sim 0.8$,即对不同噪声的重建图像进行 ROC 试验。 α 越大,实验图像的 SNR 越低,缺陷的识别难度越大。

LROC 实验中,缺陷被随机置于第 1 到第 4 象限

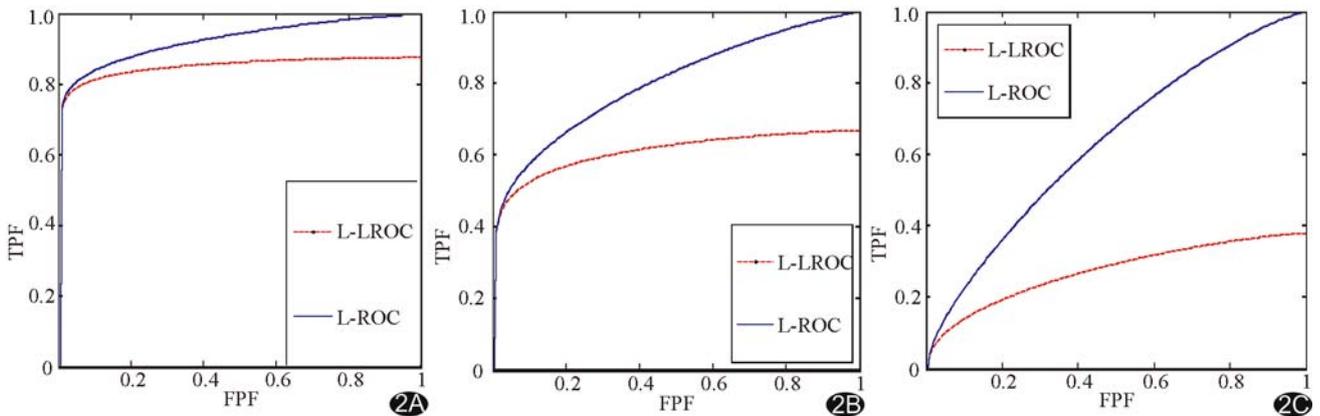


图 2 不同噪声水平下 LROC 实验结果 A~C 分别为噪声水平 $\alpha=0.3, 0.5, 0.7$ 时的实验结果

的对称位置上。当观察者判定的缺陷位置与实际缺陷位置位于同一象限时,系统记录缺陷被正确识别,否则记录缺陷被错误定位。对于无缺陷的图像,实验结果只记录 5 分类等级,与缺陷位置的标记无关。完成实验得到 3 组 5 分类数据,分别为有缺陷正确定位、有缺陷错误定位和无缺陷。同样对噪声水平 $\alpha=0.3\sim 0.8$ 的图像进行 LROC 实验,对所获结果与 ROC 实验进行比较。

由于噪声增大可降低有缺陷图像的 SNR,随着噪声等级增加,LROC 曲线的 AUC 也应随之减小。图 2 给出了 $\alpha=0.3, 0.5, 0.7$ 时的 LROC 曲线,与预期相符。

物理意义上, AUC_{L-R} 和 AUC_{R-SPV} 具有一致的含义,在实际实验中,二者的结果也基本吻合。双正态 ROC 模型下,缺陷位置不固定的重建图像实验结果如图 3 所示, AUC_{L-R} 变化曲线和 AUC_{R-SPV} 变化曲线趋势保持一致,均较好地反映了其物理意义。

对于 AUC_{R-SKE} 曲线,由于缺陷位置固定,同样噪声水平下,诊断的正确率应高于位置不定的情形,即理论上 AUC_{R-SKE} 应当大于 AUC_{L-R} 和 AUC_{R-SPV} 。实验结果印证了这一判断。绝大多数情况下, AUC_{L-R} 和 AUC_{R-SPV} 曲线均位于 AUC_{R-SKE} 曲线下方。

缺陷位置固定和缺陷位置可变所引起的诊断差异也是本实验关注的内容。图 3 显示,噪声较小时,大部分缺陷均能被正确识别,因此 AUC_{R-SPV} 、 AUC_{L-R} 与 AUC_{R-SKE} 结果类似;噪声较大时,缺陷识别均较差, AUC_{R-SPV} 、 AUC_{L-R} 与 AUC_{R-SKE} 的差距也不明显。因此,LROC 的意义更多体现在噪声水平适中的情形,且缺陷位置是否明确会对实验结果产生比较明显的影响。

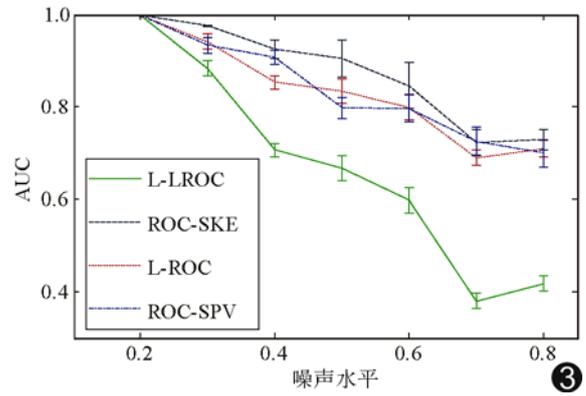


图 3 LROC 模型和 ROC 模型下实验结果对照

综上所述,LROC 方法比 ROC 方法更加适用于检测缺陷位置不确定的图像。

3.2 使用 LROC 进行单层螺旋 CT 的探测器长度优化

通过分析单层螺旋 CT 的探测器长度对于重建图像质量的影响^[8],探讨 ROC 方法和 LROC 方法的差异。

用模拟程序生成含缺陷和不含缺陷的无噪声投影数据,然后通过光子数假设在投影数据中加入高斯噪声^[9]。在 R-SKE 方法的实验结果中^[8],当探测器长度为 8~10 mm (1~1.25 倍螺距),缺陷的识别效果最佳。

同样以缺陷随机放置的 LROC 实验结果与之对照。图 4 为探测器长度分别为 4、8、10 mm 时得到的 LROC 曲线。随着探测器长度的增加,LROC 曲线下面积先上升后下降,与 R-SKE 模型较为一致。在探测器长度为 8 mm 时,LROC 模型下 AUC 最大,诊断效果最好。

进一步对双正态 ROC 模型和局部正态 LROC 模

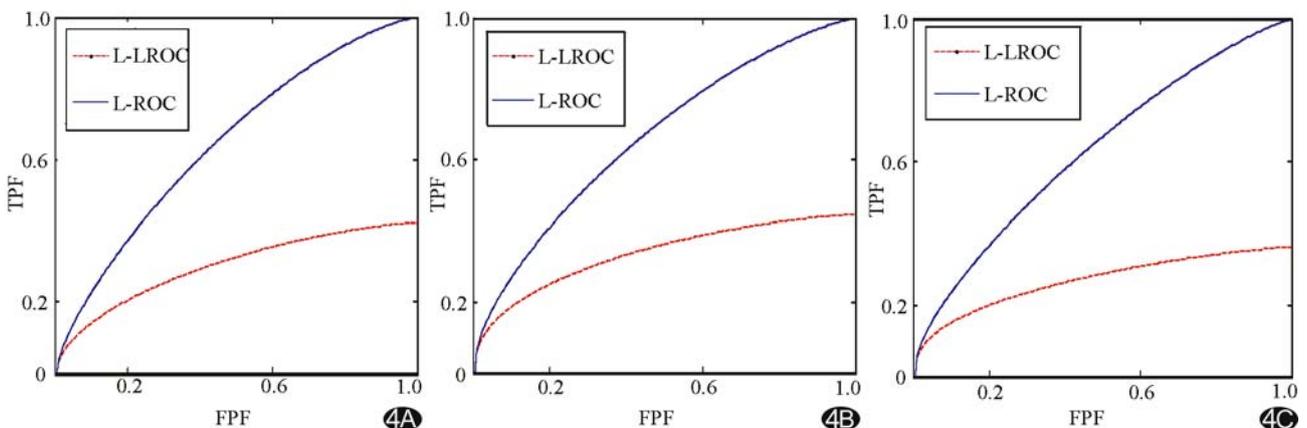


图 4 LROC 模型下 SSHCT 实验结果 A~C. 分别为探测器厚度为 4、8、10 mm 时的实验结果

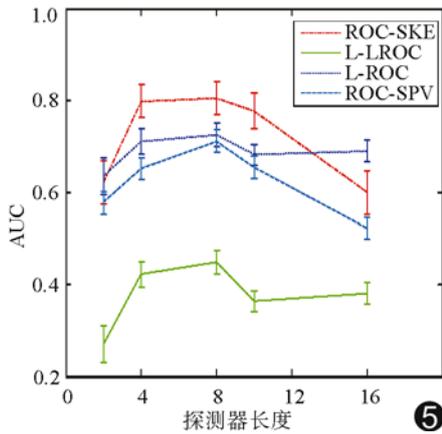


图 5 双正态 ROC 模型和局部正态 LROC 模型下实验结果对比 SNR 诊断效果比较图

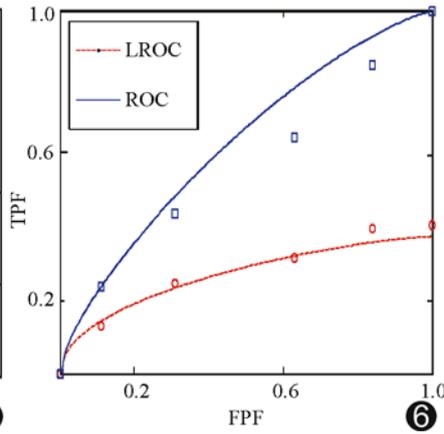


图 6 探测器长度为 16 mm 时 LROC 实验结果

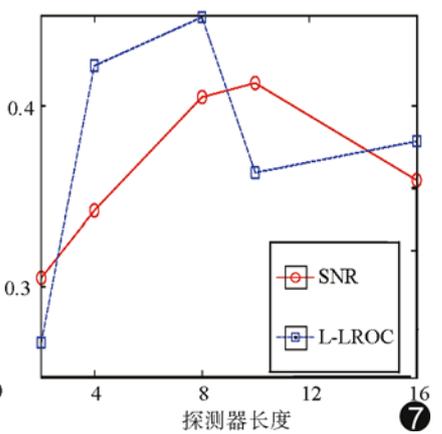


图 7 L-LROC 和 SNR 诊断效果比较图

型下的 AUC 实验结果进行对比,如图 5 所示。一方面,不同探测器长度下, AUC_{R-SPV} 与 AUC_{R-SKE} 对比均相对偏小,这是缺陷位置不确定所必然导致;另一方面,在大多数情况下, AUC_{L-R} 与 AUC_{R-SPV} 吻合,但当探测器长度为 16 mm 时出现明显偏差。经过分析发现,当探测器长度增至 16 mm (2 倍螺距) 时, AUC_{L-R} 明显大于 AUC_{R-SPV} 。这是由于对有缺陷图像的缺陷位置观察和对无缺陷位置的观察同时进行,观察结果互相影响,即得到的两组测试量不再相互独立,不符合局部正态 LROC 模型的假设;尤其当有缺陷图像缺陷位置的分布和无缺陷位置的分布接近时,彼此影响更为明显, $1 - AUC_{L-R} = AUC_{L-L}$ 的关系也不复成立。当探测器长度为 16 mm 时,这一现象非常明显。

探测器长度为 16 mm 时的实验数据见表 1,其曲线拟合结果和各个截断点比较见图 6。 AUC_{L-L} 存在一定偏误,但仍大致反映出实际诊断情况;而 AUC_{L-R} 的 AUC 则严重偏误,不再具备实际意义。所以,当有缺陷图像缺陷位置和无缺陷位置的诊断相互影响较大时, AUC_{L-R} 和 AUC_{L-L} 的关系无法确定,在此条件下,结果为 $1 - AUC_{L-R} > AUC_{L-R} - AUC_{L-L}$,说明在 LROC 模型假设下, AUC_{L-R} 被高估,而 AUC_{L-L} 被低估。

表 1 探测器长度为 16 mm 时 5 分类数据

等级	FP	ALL	CL	IL	FPF	TPF	TP&CL
1	12	11	1	10	1	1	0.41
2	16	15	6	9	0.84	0.85	0.40
3	24	16	5	11	0.63	0.65	0.32
4	15	15	9	6	0.31	0.44	0.25
5	8	18	10	8	0.11	0.24	0.13
合计	75	75	31	44	—	—	—

最后对 LROC 方法与 SNR、HTC 等图像评估指标^[10-11]进行对比。在 ROC 实验中,这些评估指标获得的结果与 AUC_{R-SKE} 有较高的一致性^[8],但与 LROC 实验结果严重偏离,以 SNR (图 7) 为例,无法取代 AUC_{L-L} 作为诊断指标。究其原因,主要在于 SNR 是在假定缺陷位置已知的情况下得到的结果。目前领域内的 HTC 等分析方法无法替代 LROC 分析。

4 结束语

ROC 方法只对缺陷位置确定或不考虑缺陷位置的情形进行诊断,因此存在局限性,对于特别关注缺陷位置的情况,需要引入新的方法,即 LROC 方法进行研究。

本研究结果显示,传统 ROC 模型下,由于缺陷位置已经确定,因此缺陷识别的效果最优,但与实际检测有所脱节;R-SPV 和 L-ROC 在物理意义上完全一致,实际实验结果也基本一致,正常情况下 AUC_{R-SPV} 与 AUC_{L-R} 均位于 AUC_{R-SKE} 之下;在 LROC 模型下,由于有缺陷位置图像缺陷位置的识别和无缺陷位置的识别并不独立,在相互干扰较为严重的情形下, AUC_{L-R} 严重失真,不再具有实际意义。综上可知,不关注缺陷位置时, AUC_{R-SPV} 更具有实际意义;而需要关注缺陷位置时,使用 AUC_{L-L} 更加合适。

总的来说,局部正态 LROC 模型在双正态 ROC 模型的基础上进行了改进,但由于假设的前提条件并不总能满足,因此同样存在局限性。改进 LROC 模型应当考虑有缺陷图像的缺陷位置和无缺陷位置识别的相关性,才能得到更准确反映其物理意义的模型。目前情况下,综合采用 ROC 和 LROC 实验是评估诊断效果或系统性能的有效方法。

[参考文献]

[1] Metz CE. Fundamental ROC analysis//Beutel J, Knudel K, Van Metter R, et al. Handbook of medical imaging, Volume 1:Physics and sychophysics. Bellingham: SPIE Press, 2000:754-769.

[2] Pesce LL, Metz CE. Reliable and computationally efficient maximum-likelihood estimation of "proper" binormal ROC curves. Academic Radiology, 2007, 14(7):814-829.

[3] Barrett HH, Myers KJ. Foundations of image science. New York: John Wiley & Sons, 2004.

[4] Swensson RG. Unified measurement of observer performance in detecting and localizing target objects on images. Med Phys, 1996, 23(10):1709-1725.

[5] Chakraborty DP. Statistical power in observer-performance studies: Comparison of the receiver operating characteristic and free-response methods in tasks involving localization. Academic Radiology, 2002, 9(2):147-156.

[6] Metz CE. Receiver operating characteristic analysis: A tool for the quantitative evaluation of observer performance and imaging systems. J Am Coll Radiol, 2006, 3(6):413-422.

[7] Shi C, Xing YX. ROC analysis of 3D X-ray CT performance for lesion detection. IEEE NSS-MIC, 2009:3731-3735.

[8] Shi C, Xing YX. Optimization of detector thickness for single slice helical CT with ROC study. Proc SPIE, 2010:7627.

[9] Wang J, Li TF, Lu HB, et al. Noise reduction for low-dose single-slice helical CT sinogram. IEEE trans Nucl Sci, 2006, 53(3): 1230-1237.

[10] Frey EC, Gilland KL, Tsui BMW. Application of task-based measures of image quality to optimization and evaluation of three-dimensional reconstruction-based compensation methods in myocardial perfusion SPECT. IEEE Trans Med Imag, 2002, 21(9):1040-1050.

[11] Abbey CK, Barrett HH. Human- and model-observer performance in ramp-spectrum noise: Effects of regularization and object variability. J Opt Soc Am A, 2001, 18(3):473-488.

“前沿 CT 成像技术临床应用与研究征文大赛”通知

近年来,CT 设备取得长足发展,能谱 CT、双源 CT 和宽探测器的诞生,引领能谱成像、低剂量高清成像、心脏成像及大范围灌注成像等前沿技术应用于临床各个领域。放射科医师根据临床需要开展了更加广泛的研究,相信伴随研究成果的不断总结和发表,临床应用前景会越来越受到同行的关注,也将会丰富影像诊断的手段。为更好地理解 and 认知前沿技术理论,展示在影像诊断中的新成果,《中国医学影像技术》期刊社与通用电气医疗集团联合举办“前沿 CT 技术临床应用与研究征文大赛”。

1. 征文内容:上述前沿 CT 技术在影像诊断和辅助治疗中的新方法、新发现、新理论以及优化检查新方案、影像诊断新技术的临床和实验研究。

2. 征文要求:论文撰写要遵循临床研究设计的基本要求,写作规范参见《中国医学影像技术》论著格式。要求 5000 字以内,未在国内外杂志上公开发表。

3. 投稿方式:通过《中国医学影像技术》投稿系统网上投稿,网址为 <http://www.cjmit.com>。(注意:请在填写“中文题目”一栏时,在中文题目后面用括号标注“征文大赛”字样,以免与其他稿件混淆。)

4. 征文截止时间:2012 年 6 月 30 日。

5. 评选办法:由影像学领域的相关著名专家组成专家评审委员会独立评审。评审标准将从论文的创新性、科学性、实用性及可读性等全面衡量。

6. 奖项设置:本次征文大赛将产生一等奖 1 名:资助参加国际学术会议或短期国际交流一次;二等奖 3 名:资助参加国内学术会议 1 次;三等奖 6 名,资助参加地区性学术会议 1 次。所有获奖者将获赠 2013 年全年《中国医学影像技术》和通用电气 CT 影像杂志(JCT)。

7. 获奖论文作者名单将在 2012 年 10 月前举办的“前沿 CT 成像技术征文大赛论坛”上揭晓,并刊登于《中国医学影像技术》杂志上;获奖论文将择优发表于《中国医学影像技术》或《中国介入影像与治疗学》杂志上。“征文大赛”入选论文将全部收录在由主办单位编辑印刷的论文集中,作者均可获得“前沿 CT 技术临床应用与研究征文大赛”论文收录证书。

主办单位保留对本次征文活动的最终解释权。

《中国医学影像技术》期刊社

2011 年 10 月 27 日