

光子计数探测器投影数据噪声的 p -范分布
统计分析郭杰¹, 蔡爱龙², 王少宇², 郑治中¹, 李磊², 闫簇^{2*}¹战略支援部队信息工程大学基础部, 河南 郑州, 450001;²战略支援部队信息工程大学信息工程学院, 成像与智能处理河南重点实验室, 河南 郑州 450001

摘要 光子计数型能谱计算机断层扫描(CT)已成为CT成像领域的最新技术,建立精确的噪声模型可以为开发高效的能谱重建算法和降低辐射剂量提供理论支撑。通过对光子计数型能谱CT投影数据噪声进行深入分析和理论推导,提出了能谱CT投影数据噪声的 p -范分布模型。首先,综合考虑光子计数探测器能量采集中的光子统计波动和电子热噪声,利用贝叶斯公式推导噪声分布模型。然后经过投影数据的密度函数拟合验证、拟合优度检验等实验环节对所提理论模型进行检验。结果表明,相比于传统分布模型,所提分布模型能够更加精准地刻画光子计数型能谱CT成像机理和物理过程。最后,对投影数据观测序列进行时间序列分析,并将预测得到的结果用于修复异常值。从仿真实验结果和实际数据实验结果可以看出,该预测值具有良好的修复效果。

关键词 光子计数型能谱计算机断层扫描; 投影数据噪声; 拟合优度检验; 一元 p -范分布; 时间序列分析

中图分类号 TP391.41

文献标志码 A

DOI: 10.3788/AOS231610

1 引言

能谱计算机断层扫描(CT)利用物质在不同能量下衰减系数的差异获取丰富的能谱信息,具有显著的物质识别和分析能力,尤其是光子计数型能谱CT,可以实现更低的电子噪声和更高的分辨率,是CT成像的最新发展技术方向。光子计数型能谱CT通过先进的半导体硬件以设置阈值的方式实现“一次扫描,多能量成像”。从双能CT到能谱CT,其对多种材料的分类能力有所提高。从积分型能谱到光子计数型能谱CT,其对X射线能量信息的区分能力有所拓展,可以获得高精度的能量信息。但是光子计数探测器中的噪声严重削弱了图像重建质量,影响光子计数型能谱CT技术的应用。如何精确刻画光子计数探测器中投影数据噪声的统计特性,有的放矢地设计能谱重建算法,有效抑制噪声已成为能谱CT成像领域的研究热点之一。

能谱的多色性和对数变换使得基于比尔-朗伯定律很难建立符合能谱CT成像机制的分布模型。为了解决这个问题,学者们提出了有针对性的噪声分布模型并设计了科学高效的求解策略。La Rivière等^[1-2]对没有经过系统校准和对数变换的原始投影数据噪声统计特性展开研究,采用泊松分布进行建模、统计迭代重

建算法进行求解,明显提升了图像重建质量,但是未能考虑X射线的多色性。Lu等^[3-4]对低剂量CT投影数据的噪声进行统计分析后指出,经过系统校准和对数变换后的投影数据噪声服从非平稳高斯分布,且均值和方差存在非线性关系。该模型为后续统计迭代重建算法提供了有力支撑。对于超低剂量X射线CT成像来说,光子数目严重不足,投影数据中电子噪声占据主导地位,导致数据中的非正值给对数变换带来了挑战。为了克服这个困难,Foi等^[5]提出了一种混合泊松高斯分布模型,可以分别模拟光子传感和电子热噪声,有效处理测量数据中的非正值。近年来,Ye等^[6]针对此问题,又提出了一种基于移位泊松分布的模型,使得似然函数能够更好地表征观测统计信息的原始测量数据,此外,他们通过设计代理函数有效地解决了非凸目标函数的优化问题。为了考察基于模型驱动的图像重建算法,Fu等^[7]对比了几种对数变换前后的分布模型,并评估了它们的精度。

实际X射线产生的是具有连续能量分布的多色能谱,X射线衰减信息可以表示为不同能量光子衰减信息的加权平均,是非线性模型。使用基于单能、窄束X射线建立的模型会导致重建图像产生硬化伪影、金属伪影等虚假信息,严重削弱图像重建质量。为此,Whiting等^[8-9]通过严格的理论推导,利用特征函数得

收稿日期: 2023-09-28; 修回日期: 2023-11-03; 录用日期: 2023-11-17; 网络首发日期: 2023-12-12

基金项目: 国家自然科学基金(62271504,62101596,62201616)、国家重点研发计划(2020YFC1522002)

通信作者: *ybspace@hotmail.com

到概率密度函数(PDF)并提出,对数变换前的多色 X 射线衰减系数服从复合泊松分布(CP)。该分布模型能够比较完整地刻画能谱 CT 成像物理过程,但是模型中涉及卷积运算,计算很难处理。2020 年, Gonzalo 等^[10]更加深入地研究了探测器中多色 X 射线束的统计特性,继而推导出对数变换后的投影数据服从耿贝尔分布的表达式,并且提出利用非中心伽马分布作为耿贝尔分布的近似分布模型。

相比传统 CT^[11-14],光子计数型能谱 CT 投影数据受到的噪声更加严重,也更加复杂,尤其在光子通量极少的情形下,投影数据噪声明显。本文探索了光子计数型能谱 CT 投影数据噪声的统计特性,创新性地提出噪声数据的一元 p -范分布模型。具体来说:首先,综合考虑光子通量和电子热噪声,利用贝叶斯公式从理论上推导出投影数据噪声服从一元 p -范分布;然后,通过观察几种常见分布对投影数据的密度函数的拟合情况,发现一元 p -范分布拟合程度最好,尤其是在光子通量较少的情形下;最后,利用统计检验对提出的一元 p -范分布模型进行统计推断,实验结果表明,投影数据在各个数据集上均符合一元 p -范分布。所提光子计数探测器投影数据噪声的分布模型扩充了光子计数型能谱 CT 成像模型理论框架,为设计高效的光子计数型能谱 CT 图像重建算法提供了理论支撑。

2 模型与方法

2.1 光子计数探测器投影数据噪声统计特性理论分析

根据比尔-朗伯定律,对于某能量的 X 射线,其与物质发生相互作用后的强度衰减规律可以表示为

$$I(E) = I_0(E) \int_0^{E_{\max}} S(E) \exp\left[-\int_L \mu(l, E) dl\right] dE, \quad (1)$$

式中: $I_0(E)$ 是能量为 E 的 X 射线发射的光子数量; $I(E)$ 是光子在与物质相互作用衰减后透射到探测器中的光子数量; E_{\max} 是 X 射线最大能谱; $S(E)$ 是 X 射线能谱密度,且满足 $\int_0^{E_{\max}} S(E) dE = 1$; $\mu(l, E)$ 是线衰减系数,与 X 射线光子的能量 E 和位置 l 有关; L 表示 X 射线穿过物体的积分路径。

多能谱 CT 成像主要受到两种噪声的干扰^[15],即 X 射线光子统计波动噪声和 CT 成像系统的电子热噪声。光子统计波动噪声是探测器收集的 X 射线光子数量不足引起的,可以用复合泊松分布描述。电子热噪声是探测器光电二极管和其他电子元件的电子涨落导致的,通常认为其服从高斯分布。因此,对于多能谱 X 射线,如果忽略光束硬化效应,光子计数探测器获得的 CT 透射数据可以用复合泊松分布 $[C_p(\lambda, \theta)]$ 和统计独立的高斯分布 $[N(\mu_e, \sigma_e^2)]$ 来描述。传输数据模型可以表示为

$$I(E) = C_p(\lambda, \theta) + N(\mu_e, \sigma_e^2) \stackrel{\Delta}{=} Q + \varepsilon, \quad (2)$$

式中: Q 、 ε 分别是探测器接收到的 X 射线的光子通量和电子热噪声;复合泊松分布中的参数 λ 表示穿过患者的 X 射线光子数的期望值,用于度量探测器收集到的平均辐射强度,如果扫描中的电压、电流确定,该参数可以反映辐射剂量水平;参数集 θ 表示 X 射线能谱; ε 刻画电子热噪声,它服从均值和方差分别为 μ_e 、 σ_e^2 的高斯分布。在 CT 系统中,电子热噪声的平均值 μ_e 可以在每次扫描之前利用空白扫描测量数据来确定,并从测量强度中减去,因此不妨假设 $\mu_e = 0$ 。根据特征函数与 PDF 之间的傅里叶变换关系,不难从光子通量的特征函数推导其 PDF,即得到式(2)中复合泊松分布的解析表达式:

$$f(E) = \exp(-\lambda) \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} S^{k \otimes}(E), \quad (3)$$

式(3)和文献[8]中导出的结果保持一致,这里的符号“ $k \otimes$ ”表示将能谱进行 k 次卷积运算。复合泊松分布的概率密度可以视为 X 射线能谱卷积的加权和,其权重可以表示为泊松分布的分布律表达式。从式(3)可以看出,复合泊松分布的 PDF 涉及能谱的卷积运算,计算很难处理,实际应用中通常采用较易处理的分布来近似复合泊松分布。文献[16]考虑用伽马分布近似复合泊松分布,经过文献[10]的分析发现,这两种分布具有相同的前两阶矩,因此可以认为它们的均值和方差是相同的,而且伽马分布的可加性可以进一步简化式(3)中的 k 重卷积运算。伽马分布的 PDF^[17]表示为

$$g(x; \alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x), & x \geq 0 \\ 0, & x < 0 \end{cases}, \quad (4)$$

式中: $\alpha > 0$ 、 $\beta > 0$ 分别为形状参数和尺度参数; $\Gamma(\cdot)$ 为伽马函数。

用 Y 表示给定的理想投影数据,沿用式(2)中的记号,贝叶斯公式表示为

$$P(I, Q|Y) = P(Q|Y)P(I|Q, Y), \quad (5)$$

式中:等式右端的第一项表示在给理想投影数据时探测器接收到的光子通量的概率,由上述分析可知,其遵循 X 射线光子统计波动规律,即服从复合泊松分布,且可以用伽马分布近似表示;第二项表示在已知投影数据和光子通量的条件下电子热噪声的条件概率,近似服从高斯分布。整理式(5)得到

$$P(I, Q|Y) = \frac{\beta^\alpha}{\Gamma(\alpha)} Q^{\alpha-1} \exp(-\beta Q) \frac{1}{\sqrt{2\pi} \sigma_e} \times \exp\left[-\frac{(I-Q)^2}{2\sigma_e^2}\right], \quad (6)$$

经过化简,式(6)可以转化为

$$P(I, Q|Y) = \frac{\tilde{\lambda}\beta\omega}{\sqrt{\pi} \sigma_e} \exp\left[-\left(\tilde{\lambda} \frac{|I-Q|}{\sigma_e}\right)^2\right], \quad (7)$$

式中: $\tilde{\lambda}$ 为常数; $\omega = \frac{t^{\alpha-1} \exp(-t)}{\Gamma(\alpha)}$, $t = \beta Q$ 。文献 [18] 中的一元 p -范分布 PDF 表示为

$$f(x) = \frac{p\eta}{2\sigma\Gamma(1/p)} \exp\left[-\left(\eta \frac{|x-\mu|}{\sigma}\right)^p\right], \quad (8)$$

式中: 参数 $p > 0$; $\eta = [\Gamma(3/p)/\Gamma(1/p)]^{1/2}$; μ, σ 分别为一元 p -范分布的均值和标准差。对比式(7)和式(8)可以发现, 当 $p=2(\Gamma(1/2) = \sqrt{\pi})$ 时, 两者具有相似的表达形式, 因此, 考虑用一元 p -范分布刻画光子计数探测器投影数据的噪声分布。首先, 观察前期大量的扫描实验数据拟合结果可以发现, 一元 p -范分布对投影数据密度函数的拟合程度是最好的, 即最接近真实投影数据分布, 尤其是在能量通道光子数目不足的情况下。其次, 大量的研究和观测误差数据处理经验表明, p -范分布可以视为带有随机误差的观测数据的一般分布, 在误差为单峰、对称的假设条件下, 可以假设观测噪声数据服从 p -范分布。通过选择合适的 p 值, 可以使噪声的 p -范分布模型较高斯分布更加接近噪声数据的真实分布, 因此, 采用 p -范分布描述噪声数据的统计特性有很大的优越性。根据 p 值选取的不同, p -范分布可以表示不同的分布, 比如常用的正态分布、拉普拉斯分布都是这一分布的特殊情形^[18]。最后, 根据线性模型的稳健估计理论^[19], 当观测误差为 p -范分布, 且 $0 < p < 2$ 时, 得到的似然估计是稳健估计。综上所述, 将光子计数探测器投影数据噪声建模为一元 p -范分布。尽管用 p -范分布描述观测误差具有显著的优势, 但是从式(8)可以看到, 其 PDF 的计算相对比较复杂, 在应用于实际数据处理时存在一定的困难。因此, 可以参考文献[20]中的做法, 选择容易处理的常见分布或者常见分布的组合近似表示 p -范分布, 比如高斯分布和拉普拉斯分布的线性组合, 或者高斯分布和均匀分布的线性组合。本研究采用高斯分布与拉普拉斯分布的线性组合来简化 p -范分布的计算。

2.2 拟合优度检验的基本原理

为了利用拟合优度检验对所提投影数据噪声模型进行统计推断, 简要介绍拟合优度检验原理^[21]。拟合优度检验要解决的问题为: 给定一组数据, 验证它是否来自某一既定分布。设 X_1, X_2, \dots, X_n 为独立同分布样本, 拟合优度检验中, 原假设 H_0 表示 X_1 的分布为 $F(x)$, 备择假设 H_1 表示 X_1 的分布不是 $F(x)$, 其中, $F(x)$ 为已知分布。构造拟合优度检验统计量 $D = D(X_1, X_2, \dots, X_n; F)$ 。当 H_0 为真时, 检验统计量的分布是已知的, 比如 χ^2 分布。若显著性水平 $\alpha = 0.05$, 则根据 χ^2 分布的分位点确定阈值 D_α 。将检验统计量的观测值 $D_0 = D(x_1^0, x_2^0, \dots, x_n^0; F)$ 与阈值 D_α 进行比较, 若 $D_0 \geq D_\alpha$ 或者 $P(D_0) = P\{D(x_1^0, x_2^0, \dots, x_n^0; F) \geq D_\alpha\} \leq \alpha$, 则拒绝 H_0 ; 反之, 则无法拒绝 H_0 。直观上讲, $P(D_0)$ 越小, 用指定分布

来拟合观测数据的分布越不合适, 则越有可能拒绝原假设; $P(D_0)$ 越大, 拒绝原假设的可能性越小。换句话说, 检验统计量的观测值越大, 用指定分布来拟合观测数据的分布越不合适, 越有可能拒绝原假设; 检验统计量的观测值越小, 越无法拒绝原假设。

2.3 观测数据时间序列分析原理

为了利用时间序列分析深入研究观测数据噪声的特性, 需要介绍时间序列分析的一般概念^[22]。将某个固定角度下得到的投影数据视为时间序列进行分析, 目的是为了考察光子计数探测器的响应是否均匀, 光子计数成像系统是否存在系统性误差, 最终利用观测数据序列分析的结果进行预测。时间序列分析过程主要包括平稳性检测、建立时间序列模型以及模型预测等环节。其中, 建立时间序列模型的主要步骤为模型识别、模型拟合及模型诊断。模型识别主要从时间序列模型中选择适合观测数据的模型; 模型拟合即找到给定模型的未知参数的估计; 最后通过残差分析对模型进行诊断, 判断模型对数据的拟合是否合理等。

一般地, 如果时间序列模型满足关系式:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}, \quad (9)$$

则称 $\{Y_t\}$ 为自回归滑动平均混合过程, 记为 ARMA(p, q), 阶数分别为 p 和 q 。 $e_t, e_{t-1}, \dots, e_{t-q}$ 是均值为 0、方差为 δ_e^2 的独立同分布的随机变量序列。回归和滑动系数分别满足 $|\phi_j| < 1 (j = 1, 2, \dots, p)$ 、 $|\theta_i| \leq 1 (i = 1, 2, \dots, q)$ 。

3 实验结果及分析

3.1 光子计数探测器投影数据噪声统计特性分析的实验设计

为了验证上述观测投影数据噪声近似服从 p -范分布的论断, 开展实验采集实际数据进行检验。通过在同一个扫描角度下对扫描物体进行 999 次的重复扫描 (帧速率为 10 frame/s, 约历时 100 s), 获得扫描投影数据。然后对重复扫描所得的光子计数探测器投影数据进行统计建模和分析 (主要考虑经过系统校准但没有经过对数变换的投影数据)。实验中所采用的模体分别为搭配 10 mg/cm³ 碘棒的牛骨以及搭配 5 mg/cm³ 碘棒的兔头。光源型号为滨松 L12721, 最大能量为 300 keV, 对牛骨扫描的系统电压设置为 140 kVp, X 射线管的电流强度为 0.6 mA。对兔头扫描的系统电压设置为 120 kVp, X 射线管的电流强度为 0.6 mA。旋转中心与 X 射线源之间的距离为 650 mm, X 射线源到探测器的距离为 1160 mm。探元尺寸为 0.1 mm × 0.1 mm。所用的牛骨标本数据集如图 1 所示, 其中, 数据集 A、B、C 依次表示空白扫描区域、牛骨扫描区域和碘棒扫描区域。对于兔头标本数据集, 也做相同的标注。为了节省篇幅, 图 2 展示了几种具有代表性的结果。



图 1 牛骨标本数据集

Fig. 1 Dataset of bovine bone

3.2 光子计数探测器投影数据噪声的 PDF 图形拟合分析

为了深入分析光子计数探测器投影数据噪声的统计特性,对同一个角度多次重复扫描的投影数据进行 PDF 图像绘制,并用高斯分布、泊松分布、伽马分布和一元 p -范分布进行拟合,观察其拟合程度,结果如图 2 所示。从图 2 可以看出,对于不同的扫描模体和数据集,一元 p -范分布在大多数情形下都能对采集的投影数据进行比较准确的建模。通过仔细分析实际数据发现,在光子通量极少的情况下,一元 p -范分布相比其他

分布拟合程度达到最优,如图 2(a)、(b)所示。图 2(a)、(b)所对应的数据集是光子穿过牛骨或者兔骨部分得到的投影数据,骨头密度比较大,因此,从发射源发射出的 X 射线光子很难穿透该部分扫描物体到达探测器,从而使投影数据产生了很高的噪声。在能量通道光子数目不足的情况下,电子热噪声会占据显著优势,此时投影数据中包含负值,无法进行对数变换。文献[6]采用移位泊松分布进行建模,但需要设计一系列的二次代理函数,计算开销比较大。根据以上现状分析,利用所提一元 p -范分布对投影数据进行建模是一

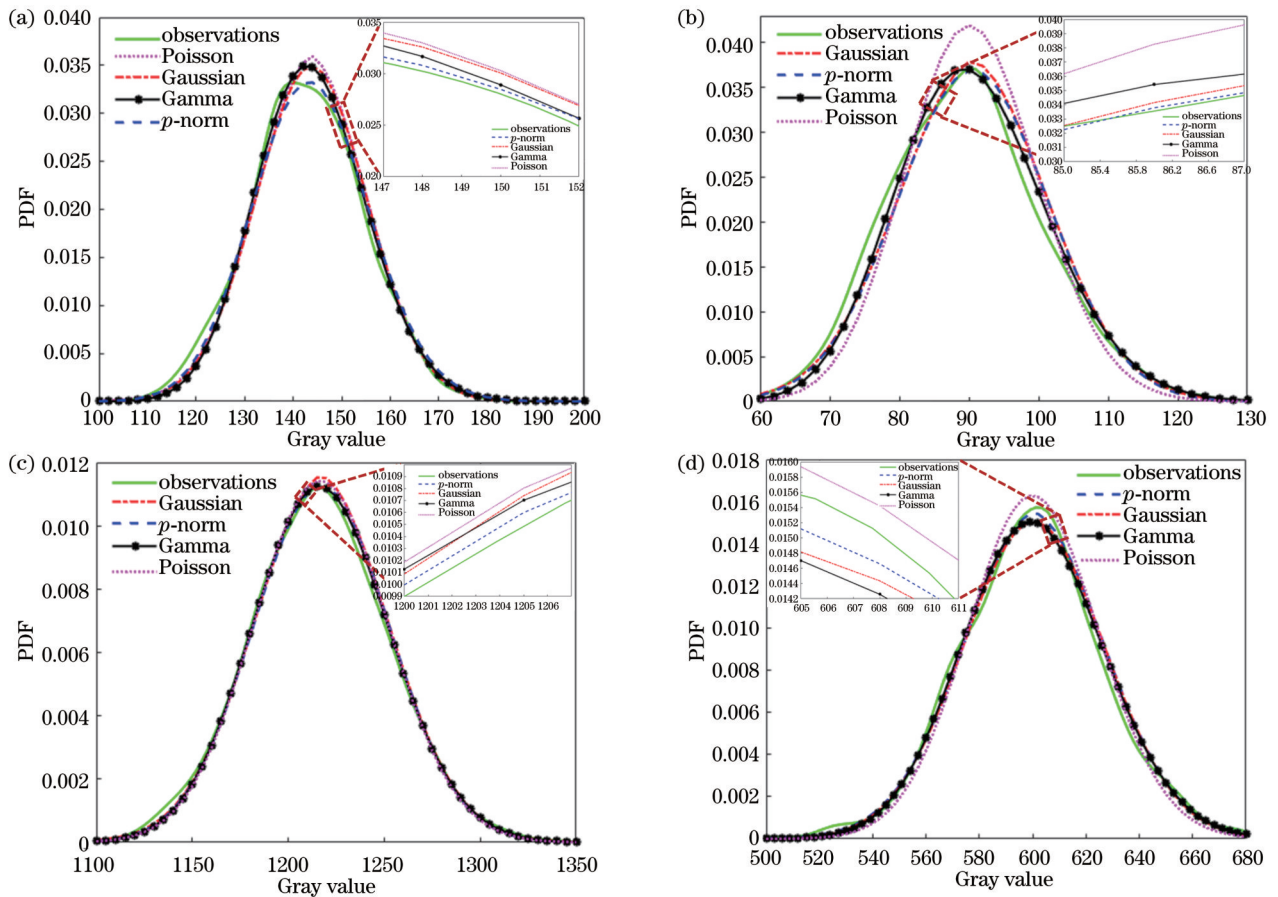


图 2 不同数据集投影数据噪声的 PDF 拟合结果。(a) 牛骨的 B 数据集;(b) 兔头的 B 数据集;(c) 牛骨的 C 数据集;(d) 牛骨的 A 数据集

Fig. 2 PDF fitting results of projected data noise in different datasets. (a) Bovine bone B dataset; (b) rabbit head B dataset; (c) bovine bone C dataset; (d) bovine bone A dataset

个较好的选择。在光子通量较少或极少的情况下,泊松分布的拟合效果是最差的,如图 2(b)所示。因此,在出现光子饥饿的情形下用泊松分布刻画观测数据的统计特性会严重削弱能谱 CT 图像重建质量。

在光子通量较少的情况下,一元 p -范分布相比其他分布拟合程度较优,如图 2(c)所示。在高光子通量的情况下,一元 p -范分布和高斯分布对观测数据的拟合基本一致,如图 2(d)所示,这一点通过中心极限定理可以得到解释。从另一个侧面也说明在误差允许的范围内用高斯分布近似计算一元 p -范分布的合理性,或者像文献[20]所述,用高斯分布和拉普拉斯分布的线性组合可以近似表示一元 p -范分布。从数据处理的经验来看,通常高斯分布的权重远远高于拉普拉斯分布的权重。从实验数据的 PDF 还可以发现,伽马分布对投影数据噪声的拟合程度较泊松分布稍好一些。这

也从实验角度验证了能谱 CT 观测数据分布模型用伽马分布近似复合泊松分布的合理性。

3.3 光子计数探测器投影数据噪声 p -范分布的拟合优度检验

为了进一步验证所提投影数据噪声的 p -范分布统计特性,采用拟合优度检验^[21],结果如表 1 所示。从拟合优度检验结果可以看到:当 $p < 0.05$ 时,拒绝原假设,即不符合指定分布;当 $p > 0.05$ 时,无法拒绝原假设,认为符合指定分布。但是 p 值越小,拟合程度越弱;相反, p 值越大,拟合程度越好。由上述分析可知:在光子通量最少的数据集 B 中,一元 p -范分布拟合程度最优;在光子通量较少的数据集 C 中,一元 p -范分布拟合程度较优;在光子通量最多的数据集 A 中,这几种分布的拟合程度几乎差不多,因此,从便于计算的角度考虑,此时选用正态分布建模观测数据较为合适。

表 1 拟合优度检验结果
Table 1 Results of goodness-of-fit test

Dataset	Assessment	Gaussian	Poisson	Gamma	p -norm
Bovine bone-B	p -value	0.1310	0.1418	0.1330	0.3326
	Test statistic	9.8522	10.3419	11.2365	7.3026
	Degrees of freedom	6	8	7	8
	Conclusion	Yes	Yes	No	Yes
Rabbit head-B	p -value	0.5093	0.2059	0.5018	0.5337
	Test statistic	5.0356	8.5148	5.3334	3.0215
	Degrees of freedom	6	7	6	6
	Conclusion	Yes	Yes	Yes	Yes
Bovine bone-C	p -value	0.0327	0.0001	0.0221	0.3125
	Test statistic	13.7325	18.9702	14.7714	10.1526
	Degrees of freedom	6	7	6	7
	Conclusion	No	No	No	Yes
Rabbit head-A	p -value	0.9247	0.9075	0.9524	0.8652
	Test statistic	1.9446	2.7450	1.6020	3.2531
	Degrees of freedom	6	7	6	6
	Conclusion	Yes	Yes	Yes	Yes
Rabbit head-C	p -value	0.0510	0.0612	0.0503	0.0501
	Test statistic	29.0262	23.1069	16.8697	25.3261
	Degrees of freedom	6	7	7	5
	Conclusion	Yes	Yes	Yes	Yes

3.4 光子计数探测器投影数据噪声时间序列分析

首先,为了对观测时间序列进行预测,必须进行平稳性检测,通常利用图检法和单位根检验法进行平稳性检测。根据检测结果可以判断,不论是牛骨标本还是兔头标本,原始观测序列都呈现非平稳性,一阶差分之后的序列则呈现平稳性,结果如图 3(a)、(b)所示。由克拉美分解定理^[22]可知,时间序列主要受到确定性影响和随机性影响两种因素的作用,而确定性影响可以由多项式决定,多次差分之后即为常数。从上述分析可知,观测序列呈一阶差分平稳,说明光子计数探测

器的响应是均匀的,或者光子计数成像系统不存在系统误差。对一阶差分之后的平稳序列进行模型识别,考察一阶差分序列的自相关函数图和偏自相关函数图[图 3(c)、(d)],从图 5 中可以看出,当滞后的阶数大于等于 2 时,自相关系数或者偏自相关系数都落在二倍标准差范围内,两者均为截尾。因此可以确定一阶差分序列模型为自回归滑动平均混合过程。结合赤池信息准则^[22]确定模型的阶数为 $p=1, q=2$ 。为了检测模型拟合的优劣,通过残差分析进行模型诊断。在分位数-分位数(Q-Q)图中,如果 X 的分布是正态分布,则

样本数据呈现线性关系。使用蓝色加号('+')标记绘制 X 中的每个数据点,两条红线代表理论分布的参考线,一条实线连接数据的第一、第三分位数,一条虚线将实线延伸至样本数据的两端,如图 3(e)所示。最后,利用所得模型进行预测,结果如图 3(f)所示。由文献[22]可知,对于差分之后的 ARMA 模型,当回归系数满足一定条件时,预测值与观测序列均值相差无几,

这一点可以从图 3(f)所示的预测值中看出,其中,黑色线表示预测结果,它基本靠近观测序列的均值附近,这为观测序列中异常值的修复提供了理论支撑。比如,光子计数探测器制作工艺的局限导致的拼缝、探测器的计数率达不到要求带来的坏点甚至坏片,便可以利用坏点或拼缝等相邻点的预测值对这类异常值进行修复。

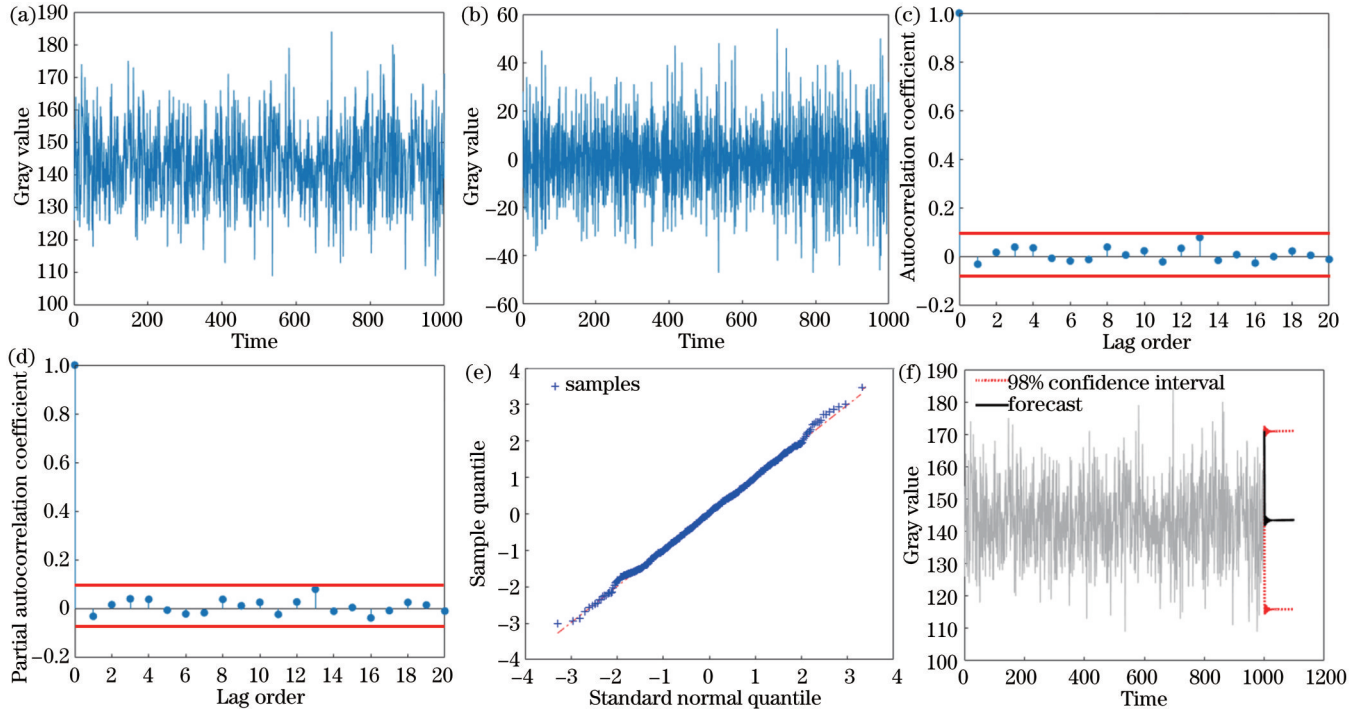


图 3 投影数据观测时间序列分析流程。(a) 原始序列时序;(b) 一阶差分序列时序;(c) 差分序列自相关图;(d) 差分序列偏自相关图;(e) 差分序列模型诊断;(f) 模型预测

Fig. 3 Flowchart for analysis of observation sequences in projection data. (a) Time series of the observations; (b) time series of the first-order differenced sequence; (c) autocorrelation plot of the differenced sequence; (d) partial autocorrelation plot of the differenced sequence; (e) diagnostic of the differenced sequence model; (f) model prediction

为了验证观测序列预测值在修复异常值中的可行性,进行仿真数据实验和实际数据实验。对于仿真数据,采用文献[23]中的仿真小鼠模体,所用模体成分主要有两部分:骨头和软组织,如图 4(a)所示。仿真产生的 120 kVp 的 X 射线能谱被分成 4 个不同的能量通道。射线源到光子计数探测器和旋转中心的距离分别为 967 mm 和 300 mm。光子计数探测器由 512 个像素组成,且每个像元大小为 0.37 mm,最终得到 360° 的投影数据。首先,固定角度对小鼠模体进行锥束扫描,得到 300 张真值投影图像。为了仿真实际投影数据,加入强度为 1×10^6 的泊松噪声;为了仿真探测器的坏点效果,加入异常值。然后,利用上述预测结果对异常值进行修复,结果如图 4(b)、(c)、(d)所示,依次为投影真实数据、加入噪声和异常值扰动的投影数据和修复后的投影数据。从图 4(d)所示感兴趣区域的修复结果可以看到,异常值区域的修复结果接近真值数据,与真值之间的误差约为 7.2×10^{-5} 。

对于实际扫描实验,采集实际数据的扫描参数和仿真数据实验相同。选用牛骨数据集和三个能量通道的投影数据重建图像,展示的能量通道为 20~30 keV。投影过程模拟圆形轨迹扇形束扫描方式,并在 0~360° 范围内采集 720 frame 不同能量下的投影数据。CT 重建图像尺寸为 512 pixel \times 512 pixel。实际数据实验思想为:利用异常值相邻点的预测值修复投影数据中的异常值,然后利用滤波反投影法(FBP)重建图像。实验中,对比算法直接利用 FBP 算法对实际投影数据进行图像重建。对比两种思路所重建的图像质量,图 4(e)、(f)分别代表能量通道为 20~30 keV 的异常值修复前后的重建图像以及感兴趣区域图像。在该组实验中,由于前期已经对投影数据进行了较为细致的环状伪影校正和几何伪影校正等预处理工作,因此消除了绝大部分由拼缝导致的坏点的影响,即在图像重建环节只能探测到投影数据中的极少量异常值,异常值修复前后的 FBP 重建图像差异不是特别明显。但是从

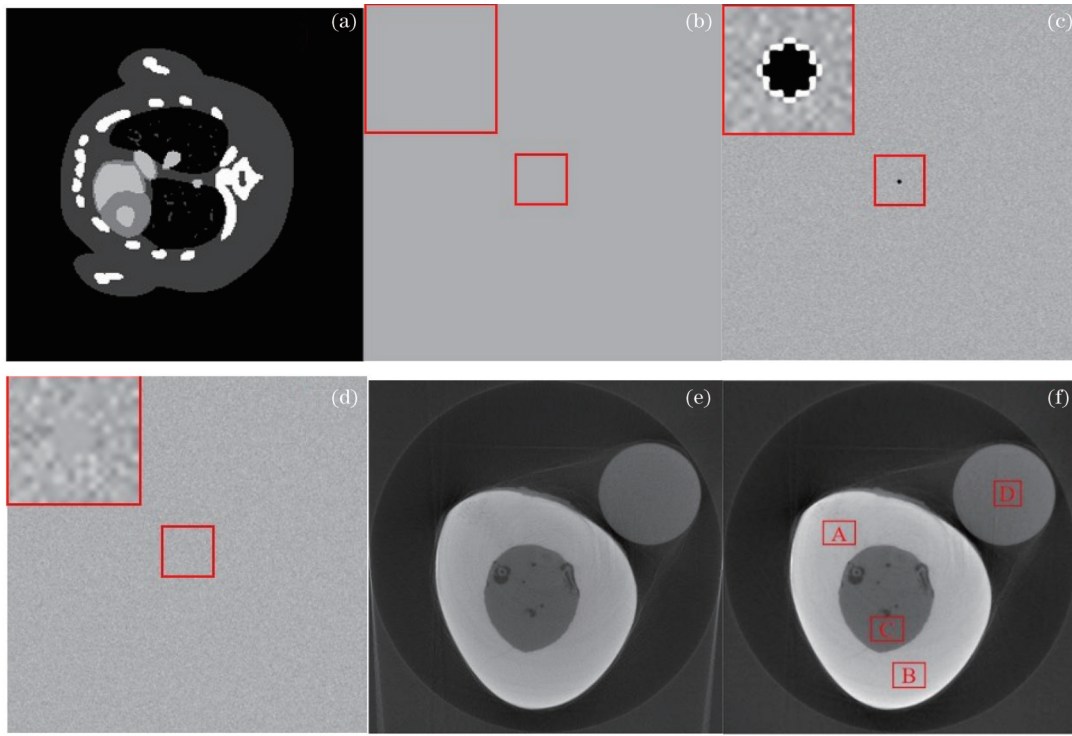


图 4 修复结果。(a) 仿真小鼠模体(窗宽为 $[0.142\ 0.145]\text{ cm}^{-1}$);(b) 真值投影图像;(c) 加入噪声和扰动的投影图像;(d) 修复后的投影图像(展示窗宽为 $[0.345\ 0.355]\text{ cm}^{-1}$);(e) 异常值修复前重建图像结果(窗宽为 $[0\ 0.35]\text{ cm}^{-1}$);(f) 异常值修复后重建图像结果(窗宽为 $[0\ 0.35]\text{ cm}^{-1}$)

Fig. 4 Restoration results. (a) Mouse thorax phantom (display window is $[0.142\ 0.145]\text{ cm}^{-1}$); (b) ground truth projection image; (c) projection image with added noise and outliers; (d) projection image after outlier restoration (display window is $[0.345\ 0.355]\text{ cm}^{-1}$); (e) reconstruction results before outlier restoration (display window is $[0\ 0.35]\text{ cm}^{-1}$); (f) reconstruction results after outlier restoration (display window is $[0\ 0.35]\text{ cm}^{-1}$)

表 2 所示感兴趣区域的定量评估指标可以看到(均值±标准差),异常值修复后的图像在均值和标准差方面均有所改善。均值即图像的平均等效衰减系数,该指标刻画了重建图像的精度;噪声标准差则度量了算法对

噪声的抑制程度。表 2 中的 A、B、C、D 分别表示图 4(f) 中的不同区域,从这 4 个区域可以看出,不论是平滑区域还是纹理组织部位,异常值修复后图像的标准差均有不同程度的降低,表明算法有效抑制了噪声。

表 2 图 4(f) 中感兴趣区域的均值和标准差比较

Table 2 Comparison of the mean and standard deviation for the region of interest in Fig. 4(f)

Region of interest	A	B	C	D
Before restoration	0.0042 ± 0.0024	0.0036 ± 0.0022	0.0149 ± 0.0208	0.0460 ± 0.0033
After restoration	0.0039 ± 0.0018	0.0033 ± 0.0014	0.0145 ± 0.0204	0.0440 ± 0.0028

4 结 论

探索了光子计数型能谱 CT 投影数据噪声的统计特性,提出了采用一元 p -范分布来建模投影数据噪声。充分考虑能谱 CT 成像中的光子统计波动噪声和电子热噪声,利用贝叶斯公式导出投影数据噪声的一元 p -范分布模型,通过拟合实际数据的密度函数与统计检验验证了该分布模型。一元 p -范分布能充分刻画观测误差的统计规律,尤其是在光子数目不足的情形下,此外,一元 p -范分布在拟合数据的真实分布时能够达到最优。所建立的光子计数探测系统投影数据的概率统计模型不仅可以深入分析探测系统的性能,还可以在

仿真实验时准确地模拟噪声,也可以为统计重建算法中似然函数的优化提供更加准确的目标函数。所提一元 p -范分布虽然能充分刻画光子计数型能谱 CT 的统计特性,但是一元 p -范分布 PDF 的计算比较困难,需要根据 p 值的选取将其转化为高斯分布和拉普拉斯分布的线性组合,从而进行近似计算。综上,探索了光子计数探测器投影数据的噪声统计特性,丰富了关于 X 射线能谱 CT 成像系统的理论成果,为多能谱图像重建算法的设计优化提供了理论支撑。

参 考 文 献

- [1] La Rivière P J, Billmire D M. Reduction of noise-induced streak artifacts in X-ray computed tomography through spline-based

- penalized-likelihood sinogram smoothing[J]. IEEE Transactions on Medical Imaging, 2005, 24(1): 105-111.
- [2] La Rivière P J, Bian J G, Vargas P A. Penalized-likelihood sinogram restoration for computed tomography[J]. IEEE Transactions on Medical Imaging, 2006, 25(8): 1022-1036.
- [3] Lu H B, Hsiao I T, Li X, et al. Noise properties of low-dose CT projections and noise treatment by scale transformations[J]. IEEE Nuclear and Plasma Sciences Society, 2001, 4(10): 1662-1666.
- [4] Lu H B, Li X, Li L H, et al. Adaptive noise reduction toward low-dose computed tomography[J]. Proceedings of SPIE, 2003, 5030: 759-766.
- [5] Foi A, Trimeche M, Katkovnik V, et al. Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data[J]. IEEE Transactions on Image Processing, 2008, 17(10): 1737-1754.
- [6] Ye S Q, Ravishankar S, Long Y, et al. SPULTRA: low-dose CT image reconstruction with joint statistical and learned image models[J]. IEEE Transactions on Medical Imaging, 2020, 39(3): 729-741.
- [7] Fu L, Lee T C, Kim S M, et al. Comparison between pre-log and post-log statistical models in ultra-low-dose CT reconstruction[J]. IEEE Transactions on Medical Imaging, 2017, 36(3): 707-720.
- [8] Whiting B R. Signal statistics of X-ray computed tomography[J]. Proceedings of SPIE, 2002, 4682: 53-60.
- [9] Whiting B R, Massoumzadeh P, Earl O A, et al. Properties of preprocessed sinogram data in X-ray computed tomography[J]. Medical Physics, 2006, 33(9): 3290-3303.
- [10] Gonzalo V S F, Raúl S J E. Statistical characterization of the linear attenuation coefficient in polychromatic CT scans[J]. Medical Physics, 2020, 47(11): 5568-5581.
- [11] 冀东江, 渠刚荣, 胡春红, 等. 基于同步辐射 CT 图像重建的对比度增强方法[J]. 激光与光电子学进展, 2020, 57(22): 221024.
- Ji D J, Qu G R, Hu C H, et al. Contrast enhancement method based on synchrotron radiation CT image reconstruction[J]. Laser & Optoelectronics Progress, 2020, 57(22): 221024.
- [12] 黎玲, 金恒, 刘杰, 等. 基于工业 CT 图像的自适应三维网格模型重建[J]. 光学学报, 2023, 43(3): 0334001.
- Li L, Jin H, Liu J, et al. Adaptive 3D mesh model reconstruction based on industrial CT images[J]. Acta Optica Sinica, 2023, 43(3): 0334001.
- [13] 刘昊鑫, 赵源萌, 张存林, 等. 基于改进 U-net 的牙齿锥形束 CT 图像重建研究[J]. 中国激光, 2022, 49(24): 2407207.
- Liu H X, Zhao Y M, Zhang C L, et al. Study on tooth cone beam CT image reconstruction based on improved U-net network[J]. Chinese Journal of Lasers, 2022, 49(24): 2407207.
- [14] 魏屹立, 杨子元, 夏文军, 等. 基于子空间投影和边缘增强的低剂量 CT 去噪[J]. CT 理论与应用研究, 2022, 31(6): 721-729.
- Wei Y L, Yang Z Y, Xia W J, et al. Low dose CT denoising based on subspace projection and edge enhancement[J]. Computerized Tomography Theory and Applications, 2022, 31(6): 721-729.
- [15] 郎云霞, 孔慧华, 牛晓伟. 基于主成分分析的多能谱 CT 图像分析方法研究[J]. CT 理论与应用研究, 2022, 31(6): 749-760.
- Di Y X, Kong H H, Niu X W. Research on multi-energy spectrum CT image analysis method based on principal component analysis[J]. Computerized Tomography Theory and Applications, 2022, 31(6): 749-760.
- [16] Seri R, Choirat C. Comparison of approximations for compound Poisson processes[J]. ASTIN Bulletin, 2015, 45(3): 601-637.
- [17] 茆诗松. 高等概率统计[M]. 北京: 高等教育出版社, 2002.
- Mao S S. Advanced probability and statistics[M]. Beijing: Higher Education Press, 2002.
- [18] 孙海燕, 胡宏昌. P-范分布及其抽样分布[J]. 应用概率统计, 2003, 19(4): 424-428.
- Sun H Y, Hu H C. P-norm distribution and its sampling distributions[J]. Chinese Journal of Applied Probability and Statistics, 2003, 19(4): 424-428.
- [19] Huber P J. Robust regression: asymptotics, conjectures and Monte Carlo[J]. The Annals of Statistics, 1973, 1(5): 799-821.
- [20] 孙海燕. p-范分布的近似表示[J]. 武汉大学学报(信息科学版), 2001, 26(3): 222-225.
- Sun H Y. Approximate representation of the p-norm distribution [J]. Geomatics and Information Science of Wuhan University, 2001, 26(3): 222-225.
- [21] 韦博成. 参数统计教程[M]. 北京: 高等教育出版社, 2006.
- Wei B C. A course in parametric statistics[M]. Beijing: Higher Education Press, 2006.
- [22] Cryer J D, Chan K S. 时间序列分析及应用[M]. 潘红宇, 译. 北京: 机械工业出版社, 2011.
- Cryer J D, Chan K S. Time series analysis with applications[M]. Pan H Y, Transl. Beijing: Chinese Machinery Press, 2011.
- [23] Ducros N, Abascal J, Sixou B, et al. Regularization of nonlinear decomposition of spectral X-ray projection images[J]. Medical Physics, 2017, 44: e174-e187.

Statistical p -Norm Distribution Analysis of Noise in Photon-Counting Detector Projection Data

Guo Jie¹, Cai Ailong², Wang Shaoyu², Zheng Zhizhong¹, Li Lei², Yan Bin^{2*}

¹Basic of Department, PLA Strategic Support Force Information Engineering University, Zhengzhou 450001, Henan, China;

²Henan Key Laboratory of Imaging and Intelligent Processing, Department of Information Systems Engineering, PLA Strategic Support Force Information Engineering University, Zhengzhou 450001, Henan, China

Abstract

Objective Spectral computed tomography (CT) is a technology that utilizes the differences in attenuation coefficients of substances across different channels, which can demonstrate significant capabilities in material identification and analysis. Particularly, photon-counting spectral CT, which significantly curtails electronic noise and enhances resolution, signifies the latest technological advancements in CT imaging. However, effects such as photon starvation, charge sharing, and

pulse pile-up engender severe noise in photon-counting spectral CT, directly undermining the image reconstruction quality and hampering the applications of photon-counting spectral CT technology. Our paramount research focus lies in accurately characterizing the statistical properties of projection data noise in photon-counting detectors, designing precise spectral CT reconstruction algorithms, and suppressing noise.

Methods Initially, a theoretical analysis is conducted on the statistical noise characteristics in the projection data of photon-counting detectors. Specifically, by comprehensively considering the statistical distribution of photon flux and electronic noise in the projection data, where photon flux can be characterized by a compound Poisson distribution and approximated by a Gamma distribution, and electronic noise follows a Gaussian distribution. A theoretical noise distribution model of projection data is derived by the Bayesian formula. Subsequently, a statistical inference is carried out on the proposed theoretical noise distribution model of projection data. On the one hand, the probability distribution of the noise is fitted via actual data experimentation. On the other hand, a goodness-of-fit test is conducted on the theoretical noise distribution model. Ultimately, by adopting time series analysis for prediction, the predicted values are employed to restore outliers in the projection data.

Results and Discussions We derive a rigorous theoretical noise distribution model in photon-counting spectral CT projection data (Eq. 9), bearing a similar expression to the univariate p -norm distribution. The rationality of characterizing the noise distribution of projection data using univariate p -norm distribution is then analyzed from three perspectives. By fitting the probability distribution of the actual data, the proposed univariate p -norm noise distribution model aligns more closely with the actual data than Gaussian, Poisson, and Gamma distributions, especially under extremely low photon flux, and the fitting degree of the proposed noise distribution model is optimal (Fig. 2). A goodness-of-fit test is conducted on the proposed noise distribution. The results are shown in Table 1. The proposed noise distribution is consistent with various collected datasets and consistency is the best in datasets with low photon flux. Lastly, the restoration of outliers using predicted values shows clear improvement from both visual images (Fig. 4) and quantitative assessments (Table 2). The proposed univariate p -norm distribution aptly characterizes the statistical properties of photon-counting spectral CT. However, the probability density function of the univariate p -norm distribution is challenging to calculate, and it should be transformed into a linear combination of Gaussian distribution and Laplace distribution for approximation, according to the p -value selection.

Conclusions We investigate the statistical noise characteristics in the projection data of photon-counting spectral CT, and propose to employ univariate p -norm distribution to model the projection data noise. The distribution is verified by fitting actual data probability density functions and statistical inference tests. The univariate p -norm distribution can fully characterize the statistical law of observational errors. Especially under the insufficient number of photons, the univariate p -norm distribution can reach optimal when fitting the actual data distribution. The statistical probability model of projection data from the devised photon-counting detection system allows for an in-depth analysis of the system performance and accurate noise simulation during simulation experiments, and provides an accurate objective function for optimizing the likelihood functions in statistical iteration reconstruction. We explore the statistical noise characteristics of projection data in photon-counting detectors, enrich the theoretical results of X-ray spectral CT imaging systems, and provide theoretical support for the design and optimization of multi-spectral image reconstruction.

Key words photon-counting spectral computed tomography; projection data noise; goodness-of-fit test; univariate p -norm distribution; time series analysis