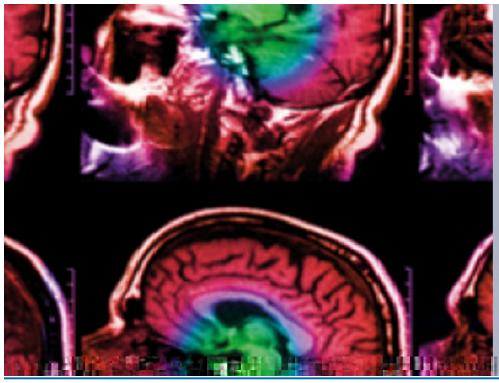
PAPER

Reducing Poisson noise and baseline drift in x-ray spectral images with bootstrap Poisson regression and robust nonparametric regression

To cite this article: Feng Zhu *et al* 2013 *Phys. Med. Biol.* **58** 1739



IOP PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. **58** (2013) 1739–1758 doi:10.1088/0031-9155/58/6/1739

Reducing Poisson noise and baseline drift in x-ray spectral images with bootstrap Poisson regression and robust nonparametric regression

Feng Zhu[1], Binjie Qin[1], Weiyue Feng[2], Huajian Wang[2], Shaosen Huang[1], Yisong Lv[3] and Yong Chen[4]

[1] School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China
[2] CAS Key Laboratory of Nuclear Analytical Techniques and CAS Key Lab for Biomedical Effects of Nanomaterials and Nanosafety, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, 100049, People's Republic of China
[3] Department of Mathematics, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China
[4] School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China

E-mail: bjqin@sjtu.edu.cn

Received 30 July 2012, in final form 7 January 2013
Published 22 February 2013
Online at stacks.iop.org/PMB/58/1739

Abstract

X-ray spectral imaging provides quantitative imaging of trace elements in a biological sample with high sensitivity. We propose a novel algorithm to promote the signal-to-noise ratio (SNR) of x-ray spectral images that have low photon counts. Firstly, we estimate the image data area that belongs to the homogeneous parts through confidence interval testing. Then, we apply the Poisson regression through its maximum likelihood estimation on this area to estimate the true photon counts from the Poisson noise corrupted data. Unlike other denoising methods based on regression analysis, we use the bootstrap resampling method to ensure the accuracy of regression estimation. Finally, we use a robust local nonparametric regression method to estimate the baseline and subsequently subtract it from the x-ray spectral data to further improve the SNR of the data. Experiments on several real samples show that the proposed method performs better than some state-of-the-art approaches to ensure accuracy and precision for quantitative analysis of the different trace elements in a standard reference biological sample.

(Some figures may appear in colour only in the online journal)

1. Introduction

1.1. Various applications of x-ray spectral imaging

X-ray spectral imaging has been used for more than half a century to identify and quantify the elemental composition of a wide variety of geological, biological and medical target

samples (Jenkins *et al* 1995). Recently, due to the advent of the third generation synchrotron radiation facility, x-ray spectral imaging has provided quantitative imaging of trace elements in a biological sample with high sensitivity (sub-mg kg^{-1}) and high spatial resolution (sub-μm to nm). More and more researchers in the field of biomedicine and life science are showing great interest in this technology (Gherase and Fleming 2011). In the analysis of diseases such as Parkinson's disease and Alzheimer's disease, x-ray spectral images are useful when the quantitative imaging of the element spatial distribution is needed to study the disease development (Popescu *et al* 2009, Wang *et al* 2010). Qin *et al* (2011) used synchrotron radiation x-ray spectral images to explore the spatial association of copper in rat aortic media. Furthermore, as an interdisciplinary science complementary to genomics and proteomics, a new research subject called metallomics has been developed recently and is receiving great attention as a new frontier in the investigation of trace elements in biology (Mounicou *et al* 2009). However, the accurate quantitative analysis is badly affected by the Poisson noise and baseline errors that are inherent in the x-ray spectral imaging. In particular, denoising x-ray spectral images that have low photon counts pose a big challenge in the quantitative analysis of trace elements in biomedicine, which is also a focus of this study.

1.2. Typical procedure of x-ray spectral imaging

Different elements in a sample emit different scattered characteristic x-ray beams of many different energies when the sample is scanned and irradiated by the incident beams (such as x-ray, electrons) at every scanning location; each of these beams goes to the photon counting detector, and the intensities (i.e. photon counts at the detector) of these characteristic beams are proportional to the contents of the elements. This phenomenon is the very foundation of the analysis based on x-ray spectral imaging. Based on this physical law, a typical x-ray energy versus the intensity spectrum divided by thousands of energy channels can be collected. Due to the physical nature of characteristic beams, only one or a few elements will be present at a particular energy channel of the spectrum when these elements are scanned at particular scanning location.

Scanning electron microscopy with an energy dispersive x-ray spectrometer (SEM-EDS) and energy dispersive micro synchrotron-based x-ray fluorescence (μSXRF) imaging are two commonly used methods to study the interactions of trace elements and single cells in a natural system for the reason that they have relatively high sensitivity and high spatial resolution (Twining *et al* 2003). Under energy-dispersive configuration, both methods use the same analytic procedure to quantitatively analyze the spectral images. The apparatus of SEM-EDS is cheaper and smaller than that of μSXRF imaging, but the monochromaticity, detection sensitivity and spatial resolution of μSXRF imaging are much higher than those of SEM-EDS (Van Grieken and Markowicz 2002). In this paper, the two data sets used in our experiments are produced by these two methods.

Before using these spectral imaging methods for the quantitative analysis in biomedicine, a specific analyte (or standard sample) in the form of either solid or solution is placed on the scanning platform to collect the multidimensional x-ray spectral data. The acquisition of multidimensional x-ray spectral data is a typical Poisson process (Boulanger *et al* 2010), which makes the raw x-ray spectral data corrupted by Poisson noise. Besides, instrument-based systematic errors will also lead to a continuous, slowly varying baseline in the acquired spectral data (Twining *et al* 2003). Figure 1(a) displays such typical raw x-ray spectral data polluted with Poisson noise and a systematic baseline. Many hardware-based efforts have been made to improve the signal-to-noise ratio (SNR) such as the insurance of 90° geometry between the incident and scattered beams (Geraki *et al* 2004). In this paper, we proposed a

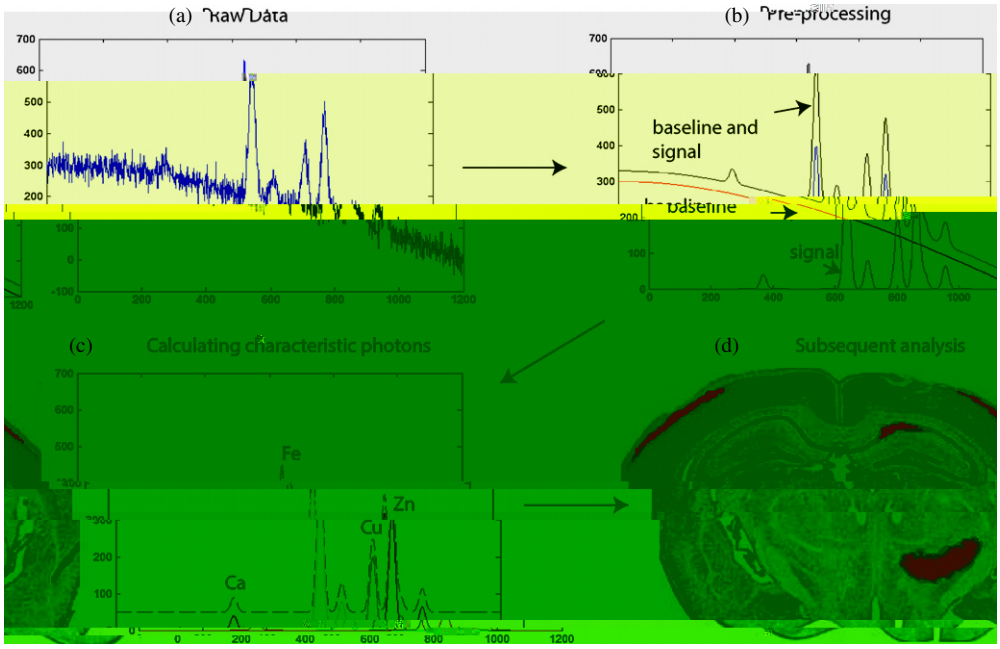


Figure 1. The typical processing procedure for x-ray spectral image data: (a) simulated raw x-ray spectral data which are corrupted with Poisson noise and baseline; (b) using pre-processing methods to reduce Poisson noise and baseline, baseline (red line) is subtracted from the denoised data (black line) and the clean signal (blue line) is obtained (here the baseline is lowered manually in order to give a clear show); (c) separating different characteristic peaks and using these peaks to calculate the characteristic x-ray intensities of different elements; here we simulated the KL peak of Ca, KL and KM peaks of Fe, Cu, and Zn; the pre-processed data (black dash line) are raised manually in order to give a clear display; (d) the calculated x-ray intensities can be used in the subsequent analysis such as mapping certain elements in biological samples.

novel software-based method that can reduce Poisson noise and baseline in the x-ray spectral data by means of signal processing. The desired effect of our method can be seen from the pre-processing procedure in figure 1(b) with baseline (red line) subtracted from the denoised data (black line). The relatively pure spectral data (blue line) in figure 1(b) are the output of our proposed method.

The subsequent procedure is to calculate the intensities (photon counts) of characteristic x-ray beams of different elements from the spectrum. In this step, the signals of different elements need to be separated and then integrated (figure 1(c)). The original data (black dash line) have been separated as the sum of Ca (red peaks), Fe (green peaks), Cu (blue peaks) and Zn (purple peaks). Methods such as iterative least-squares fitting of a mathematical model combined with Monte Carlo simulations (Bekemans *et al* 2003), and baseline-corrected spectra fitting to a summed exponentially modified Gaussian (EMG) peak model with a sigmoidal baseline (Twining *et al* 2003) are typically used in the separation of different elements' characteristic peaks. After the separation procedure, each element's peak areas are integrated to be equal to the number of photon counts for each element, which is then used for the quantitative mapping of element spatial distribution to disease development (figure 1(d)). The whole data processing procedure is displayed in figure 1.

Recently, a multi-platform open source software tool called PyMCA for the analysis of energy-dispersive x-ray fluorescence spectra has been developed (Sole *et al* 2007). This

software tool has combined many well-performed algorithms in it, which can be used to calculate the intensities of characteristic x-ray beams (as displayed in figure 1(c)) in our standard biological samples. In the next step (figure 1(d)), we demonstrate the use of these characteristic photon intensities so that we can obtain the quantitative amounts of different trace elements in standard biological samples (Wang *et al* 2010).

1.3. Review of Poisson denoising and baseline removal

As has been mentioned above, we will describe a new signal processing method that deals with Poisson noise and baseline in the x-ray spectral image data. Our method is generally applicable to x-ray spectral images that have low photon counts and therefore pose a big challenge in the quantitative analysis of trace elements in biomedicine.

There is extensive literature on Poisson denoising methods which can be generally divided into three classes. The first use multiscale analysis technique (Zhang *et al* 2008, Luisier and Blu 2008, Wang 2007, Spring and Clegg 2009) such as wavelet analysis. After the noisy signal being decomposed into noise and the useful signal by wavelet transform, the inversely transformed signals will be free from noise. Since the Poisson statistics are generally more difficult to be tackled than the Gaussian ones, the variance stabilizing transform is integrated into the multiscale analysis framework to transform the noise model from Poisson to Gaussian (Anscombe 1948, Spring and Clegg 2009, Makitalo and Foi 2011, Zhang *et al* 2008, Palakkal and Prabhu 2012). In general, these algorithms usually require prior knowledge of noise to set up appropriate parameters. The second class of methods estimate the true photon counts directly through statistical means, such as Bayesian inference combined with multiscale analysis (Timmerman and Nowak 1999, Kolaczyk 1999, Lefkimmiatis *et al* 2009), hypothesis testing (Kolaczyk 2000), maximum likelihood estimation and regression analysis. In addition, there are variational approaches for Poisson denoising (Le *et al* 2007, Chan and Chen 2007, Bonettini and Ruggiero 2011, Zhou and Li 2013). Regression analysis methods (Boulanger *et al* 2008, Kervrann and Trubuil 2004) have a good performance when they deal with low photon count data. However, the statistical methods often give bad results when the sample size is small. Some resampling algorithms (Haynor and Woods 1989) have been introduced to overcome this small sample problem. Among these algorithms, the bootstrap method (Dahlbom 2002) allows the estimation of the sampling distribution of almost any statistic using only very simple methods. Considering the fact that multidimensional x-ray spectral data are usually low photon count data containing a very high level of Poisson noise, we propose Poisson regression with the bootstrap resampling method to remove the Poisson noise in the multidimensional x-ray spectral data. To the best of our knowledge, we are the first to propose the bootstrap Poisson regression methods for x-ray spectral image denoising.

In order to further improve the performance of x-ray spectral imaging, we should also remove the baseline in the x-ray spectra. The baseline is caused by systematic errors such as the nonlinear response of the detector (Van Grieken and Markowicz 2002). The continuous and low varying baseline in x-ray spectra can be treated as superposition on the original data (Ruckstuhl *et al* 2001). Various methods in the literature have been proposed to estimate the baseline of a spectrum. Liland *et al* (2010) give an overview of the baseline correction for the multivariate calibration of spectra. Here, we choose the robust nonparametric regression methods based on the algorithm proposed by Ruckstuhl *et al* (2001) for baseline removal since this method is both effective and robust. Being different from the original method, our algorithm uses a new regression kernel for the flexibility of second-order robust regression.

The remainder of this paper is organized as follows. Section 2 will explain the method in detail. Section 2.2 introduces the bootstrap Poisson regression method. Section 2.3 outlines

the robust regression methods for baseline removal. Section 3 gives the experimental results of our method in comparison with other state-of-the-art methods using standard samples including biological samples. In section 4, we briefly summarize our method and future research directions.

2. Methods

2.1. Data acquisition and data structure

In this work, we use alloy wire data and two standard biological samples (bovine liver NIST 1577a and pig liver GBW 08551) as real samples in our multidimensional x-ray spectral imaging experiments. Both wire sample and biological samples are placed on an acquisition platform with a micron spot of incident beam focused on each scanning location in the samples. An energy dispersive x-ray spectrometer counts the characteristic photons emitted from each scanning sample point. Then, the incident beam spot will move to the next scanning location for continuous data acquisition. The scanning time of each scanning point is the same. Based on the above description, the underlying structure of the acquired multidimensional x-ray spectral data can be taken as a 2D–1D structure. The 2D part of the multidimensional data refers to the two-dimensional images that are formed from the total scanning points of the same energy channel, while the 1D part refers to the whole spectrum for all energy channels at a single scanning point.

2.2. Poisson denoising

Photon counting x-ray spectral image data are typical Poisson distributed data. That is to say the photon count data Y_i ($i = 1, \ldots, N$, with N being the total number of scanning points in each image of the 2D part of the data) follow a distribution with a density function $f(Y_i; \lambda_i) = \frac{\lambda_i^{Y_i} \mathrm{e}^{-\lambda_i}}{Y_i!}$, where λ_i is the desired noise free photon counts, which can be estimated as the mathematical expectation of Y_i. However, it is impossible to calculate the expectation of $f(Y_i; \lambda_i)$ (λ_i) through Y_i itself. One method to solve the problem is to find more data from the same distribution ($f(Y_i; \lambda_i)$). As for x-ray spectral data, these identically distributed data can be found through the following analysis.

In modern high-resolution x-ray spectral imaging, an incident beam is scanned as a nanoscale spot in a raster pattern across the sample's surface to make the scanning points get fairly close to each other to offer sufficient details of the sample. The nanoscale spot is so small that the neighboring scanning points around an interest point can be assumed to belong to a homogeneous region of the sample. In the biomedical imaging, the homogeneous regions are not equally extended to all directions from a point of interest, so that these irregular homogeneous regions usually introduce local discontinuities at some directions in the sample. Therefore, we need to check which neighboring points belong to the same homogeneous part as the current scanning point of interest. A proper size of neighborhood should be determined so that there are sufficient number of homogeneous points that are chosen to recover the noise free photon counts λ_i. Furthermore, the bootstrap resampling method is used to produce sufficient candidates for good estimation of λ_i. With these spatial-domain local homogeneity assumption and sampling strategies, we have local photon count data that are independent and identically distributed (IID) for the following Poisson regression analysis.

By analyzing the above-mentioned spatial-domain features of the x-ray spectral data, we can use Poisson regression analysis to estimate the λ_i from these IID data. Poisson regression assumes that the response variable Y_i has a Poisson distribution and assumes that the logarithm

of its expected value (λ_i) can be modeled by a linear combination of unknown parameters. We suppose that the total number of homogeneous data in the local Poisson regression area is m, and these homogeneous data that belong to the same distribution as that of Y_i in the neighbourhood of Y_i are marked as $Y_{i,j}$ (Y_i is included in $Y_{i,j}$, $j = 1, \ldots, m$). Therefore, the logarithmic forms of the expectation λ_i for the $Y_{i,j}$ can be fitted with a linear model function:

$$\log(\lambda_i) = a_i x_j + b_i + \varepsilon_i; \qquad \lambda_i \approx \hat{\lambda}_i = \mathrm{e}^{a_i x_j + b_i}, \quad j = 1, \ldots, m, \tag{1}$$

where x_j is an auxiliary explanatory variable for each point of $Y_{i,j}$, $\hat{\lambda}_i$ is the estimator of the true expectation λ_i, a_i and b_i are unknown parameters to be calculated, and ε_i is the white noise with a fixed variance and zero mean. The explanatory variable x_j has no physics meaning, but it acts as a mathematic auxiliary tool for Poisson regression. To determine the explanatory variable x_j, the $Y_{i,j}$ are sorted increasingly or decreasingly so that the explanatory variable x_j can simply be the sorting index. Therefore, the noise free photon counts λ_i are approximated as

$$\lambda_i = E(Y_i) \approx \mathrm{e}^{a_i x_{j'} + b_i}, \tag{2}$$

where $x_{j'}$ is the corresponding explanatory variable to Y_i.

With the above-mentioned general scheme in mind, we should first choose the proper irregular homogeneous regions so that the Y_i's neighbouring photon count data have the same distribution function as that of Y_i. Here, we use the significance test to solve this problem. The acquired photon counts Y_i are assumed to have a Poisson distribution. The data $Y_i^{'}$ in the Y_i's neighbourhood should share the same cumulative distribution function:

$$F(u) = \mathrm{e}^{-Y_i} \sum_{k=0}^{u} \frac{Y_i^k}{k!}. \tag{3}$$

With equation (3) we can calculate the lower and upper bound photon counts of the confidence interval of level 95%. The data $Y_i^{'}$ that have the photon counts outside the confidence interval will be taken as the data samples having the different distributions from that of Y_i.

After choosing the homogeneous data that are a candidate for choosing $Y_{i,j}$ in the Poisson regression analysis of Y_i, the performance of estimator $\hat{\lambda}_i$ in equation (1) is then dependent on the $Y_{i,j}$'s size (or bandwidth) m and can vary at each point of the photon count data sequence according to image contents. A small bandwidth will make local regression analysis sensitive to noises and outliers, while large bandwidth will create a large approximation error in local regression. In order to optimally estimate the bandwidth m, we analyze the performance of the estimator and consider the usual local L_2 risk (Boulanger *et al* 2008) defined as

$$\mathcal{R}(\hat{\lambda}_i, \lambda_i) = E[(\hat{\lambda}_i - \lambda_i)^2], \tag{4}$$

where λ_i is the unknown expectation. The local risk $\mathcal{R}(\hat{\lambda}_i, \lambda_i)$ is defined at each point and then differs from usual global performance measures that integrate errors on the whole images. This local risk of the candidates selected from the significance test reaches its minimum at each point. Choosing a new larger m that does not increase the number of candidates means that the local discontinuity appears; the previous smaller m is then considered as the optimal size. Otherwise, the local risk should be minimized to choose the optimal m as follows.

Boulanger *et al* (2008) have given in detail the solution of the minimization problem described in equation (4), which designs a sequence of increasing bandwidth: $M = \{m^{(n)}(x_j), n \in [0, N]; m^{(n-1)}(x_j) < m^{(n)}(x_j)\}$. And then this sequence is used to detect the optimal bandwidth m^* for the local smoothing:

$$m^*(x_j) = \sup_{m^{(n)}(x_j) \in M} \{n' < n : \left|\hat{\lambda}_i^{(n)} - \hat{\lambda}_i^{(n')}\right| < \vartheta \upsilon_{n'}(x)\}, \tag{5}$$

where $\vartheta = 2\sqrt{2}$ is a positive constant and $\upsilon_n^2(x)$ is the variance of the data Y_i in the regression context. Boulanger *et al* (2008) have proved that m^* in the sequence of M that satisfies equation (5) will be the optimal bandwidth m that minimizes the risk described in equation (4). Typically, choosing eight or more homogeneous data points in the local regression area will ensure satisfactory estimation of λ_i.

In order to calculate the values of a_i and b_i in equation (1), we use the principle of maximum likelihood estimation to compute the set of parameters (a_i, b_i) that make the following log-likelihood function value as large as possible:

$$l(a_i, b_i) = \sum_{j=1}^{m} Y_{i,j}(a_i x_j + b_i) - e^{a_i x_j + b_i} - \log(Y_{i,j}!). \tag{6}$$

Unfortunately, directly computing a_i and b_i is difficult since equation (6) has no closed-form solution. An iterative weighted least-squares method (Davison and Hinkley 1997) can be used to estimate a_i and b_i. At each iteration an adjusted responses vector $\boldsymbol{z}_i = (\ldots, z_{i,j}, \ldots)$ is regressed on the x_j with elements $z_{i,j}$ being expressed as

$$z_{i,j} = w_j + (Y_{i,j} - e^{w_j}) * \frac{1}{w_j}, \tag{7}$$

where the weight w_j is given by the estimators $\hat{a}_i$ and $\hat{b}_i$ (corresponding to a_i and b_i)

$$w_j = \hat{a}_i x_j + \hat{b}_i. \tag{8}$$

The result of each iteration given in the form of matrix is

$$\begin{bmatrix} \hat{a}_i \\ \hat{b}_i \end{bmatrix} = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} \boldsymbol{z}_i, \tag{9}$$

where $\mathbf{X}$ is the matrix of x_j and $\mathbf{W}$ is the diagonal matrix of weights w_j (i.e. $\mathbf{W}[s,t] = w_j$, when $s = t$; $\mathbf{W}[s,t] = 0$, when $s \neq t$).

So far, the estimators $\hat{a}_i$ and $\hat{b}_i$ are ready to be substituted into equation (2) for estimating the noise free photon counts λ_i. However, the accuracy of statistical estimation depends on the sample size. In practice, one needs as many samples as possible to ensure high estimation accuracy in statistical analysis. However, it is often not easy to obtain many samples. Efron (1979) has introduced bootstrap resampling methods to deal with this problem. Here, we also use the bootstrap resampling methods to increase the accuracy of statistic analysis.

To apply the bootstrap method to enhance the accuracy of estimating λ_i, we need to generate new sample data $Y_{i,j}^*$ from the original data $Y_{i,j}$. Specifically, we generate $Y_{i,j}^*$ from the estimated expectation $\hat{\lambda}_i$ by using the following equation:

$$Y_{i,j}^* = \hat{\lambda}_i + \hat{\lambda}_i^{1/2} * \varepsilon_j^*, \qquad j = 1, \ldots, m, \tag{10}$$

where $\varepsilon_1^*, \ldots, \varepsilon_m^*$ are adjustment parameters that determine the performance of the whole resampling method. It has been demonstrated that the residuals of Poisson regression can be used as these adjustment parameters for bootstrap resampling methods (Davison and Hinkley 1997). In order to deduce the residuals of Poisson regression, we first introduce the $\mathbf{H}$ matrix that is derived from equations (7)–(9):

$$\mathbf{H} = \mathbf{X}(\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}. \tag{11}$$

With this definition of matrix $\mathbf{H}$, the standardized Pearson residuals can be written as

$$r_{Pj} = \frac{Y_{i,j} - \hat{\lambda}_i}{\{\hat{\lambda}_i(1 - h_j)\}^{1/2}}, \quad j = 1, \ldots, m, \tag{12}$$

where h_j is the jth diagonal element of the hat matrix $\mathbf{H}$.

The standardized Pearson residuals are further expressed as mean-adjusted Pearson residuals by $r_{Pj} - \bar{r}_P$, where $\bar{r}_P$ is the mean of r_{Pj}. These mean-adjusted Pearson residuals have all the qualities that bootstrap resampling method needs. With the adjustment parameters $\varepsilon_1^*, \ldots, \varepsilon_m^*$ being sampled from these mean-adjusted, standardized Pearson residuals according to the bootstrap resampling rule, the new sample data $Y_{i,j}^*$ can be obtained as additional data to implement Poisson regression. Thus, the desired expected value λ_i (noise-free photon counts in equation (2)) can be estimated with higher accuracy. Theoretically, the number of new samples generated from bootstrap methods needs to be infinity. In practice, 300 or more samples can ensure good results with pleasant accuracy.

2.3. Baseline drift removal

Baseline drift in x-ray spectral images is caused by hardware-based systematic errors such as the nonlinear response of the detector. Based on the feature of a slowly varying local continuous baseline, it is convenient to remove the baseline drift from the spectrum of each scanning point by using a robust local regression method (Ruckstuhl *et al* 2001).

After the Poisson regression analysis, the spectral data of one scanning point can be modeled as follows:

$$V(c_k) = g(c_k) + s(c_k) + \varepsilon_k, \quad k = 1, \ldots, K, \tag{13}$$

where $V(c_k)$ is the processed data by Poisson regression, $g(c_k)$ is the baseline, $s(c_k)$ is the desired signal at the energy channel c_k, ε_k represents the measurement errors with zero mean and variance ξ, and K is the total number of points in the whole spectrum of a single scanning point.

In order to separate the three components in equation (13), a locally weighted scatter plot smoothing method (Cleveland 1979) can be used. With the data $V(c_k)$ at a energy channel c_k, we suppose that the bandwidth of the local regression context around c_k is n. In the n defined local regression context, the Poisson regression processed data t_{k+i} within this context can be defined with the following modified equation (from equation (13)):

$$t_{k+i} = g(c_{k+i}) + E_{k+i}; \quad E_{k+i} = s(c_{k+i}) + \varepsilon_{k+i}; \quad i = -n/2, \ldots, n/2, \tag{14}$$

where $g(c_{k+i})$ is the baseline function that has enough smoothness and E_{k+i} is the sum of signal $s(c_{k+i})$ and error ε_{k+i}. The baseline function $g(c)$ can be approximated by using second-order Taylor's formula at point c_k:

$$\begin{aligned} g(c) &= g(c_k) + g'(c_k)(c - c_k) + g''(c_k)(c - c_k)^2 + \mathrm{O}((c - c_k)^2) \\ g(c) &\approx \hat{g}(c) = \beta_0 + \beta_1(c - c_k) + \beta_2(c - c_k)^2 \end{aligned}, \tag{15}$$

where $\hat{g}(c)$ is the baseline estimated by the regression model. The parameter vector $\beta(c_k) = [\beta_0, \beta_1, \beta_2]^T$ can be calculated by incorporating a weight scheme into the local least-squares problem to decreases the influence of data points in proportion to their distance from c_k. That is,

$$\beta(c_k) = \arg\min_{\boldsymbol{\beta}} \sum_{i=-n/2}^{n/2} K\left(\frac{c_{k+i} - c_k}{h}\right)\{t_{k+i} - [\beta_0 + \beta_1(c_{k+i} - c_k) + \beta_2(c_{k+i} - c_k)^2]\}, \tag{16}$$

where $K[(c_{k+i} - c_k)/h]$ is a unimodal symmetric non-negative weight function that is zero outside the c_k's neighbourhood, which is defined by $c_k \pm h$ with h being half of the bandwidth of the regression context. The estimation performance is not greatly dependent on the choice of the weight function K (Ruckstuhl *et al* 2001). Here, we choose a logarithmic function to reduce the influence of the quadratic part in equation (15), so that the weight of data in the

energy channels far from the current channel will drop quickly to have a little effect on the estimation while the data in the close energy channels will have high weight:

$$K(u) = 1 - \log[(e-1)u + 1]. \tag{17}$$

Equation (16) is used to estimate $\beta(c_k)$ initially; next we use the residuals of this estimation to assign robustness weight $w_r(c_{k+i})$ to each point, such that the points with large residuals will receive small robustness weights. The baseline curve $g(c)$ is then refined by performing a weighted least-squares fit, according to

$$\beta(c_k) = \arg\min_{\boldsymbol{\beta}} \sum_{i=-n/2}^{n/2} w_r(c_{k+i}) K\left(\frac{c_{k+i} - c_k}{h}\right) \{t_{k+i} - [\beta_0 + \beta_1(c_{k+i} - c_k) + \beta_2(c_{k+i} - c_k)^2]\}. \tag{18}$$

This fit is repeated iteratively to converge with the weights $w_r(c_{k+i})$ always being calculated from the previous iteration. On the choice of various $w_r(c_{k+i})$, we use Tukey's bisquare weights

$$w_r(c_{k+i}) = \{\max[1 - (r_{k+i}/b)^2, 0]\}, \tag{19}$$

where $r_{k+i} = [t_{k+i} - \hat{g}$ / . ( ) (. . . ( ) / . ( ) / .

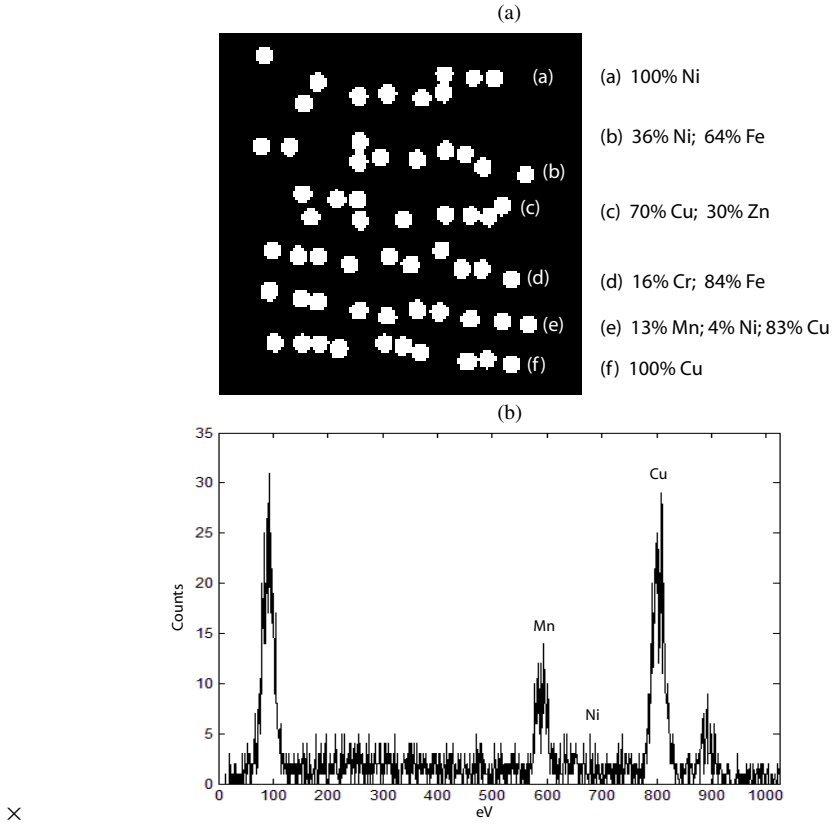


Figure 2. (a) The different positions of the wires with the different wire compositions and component concentrations: (a) 100% Ni, (b) 36% Ni, 64% Fe, (c) 70% Cu, 30% Zn, (d) 16% Cr, 84% Fe, (e) 13% Mn, 4% Ni, 83% Cu, (f) 100% Cu ((a)–(f) are referred to the rows of dots in figure 2(a)). (b) A single-pixel spectrum from the Cu/Mn/Ni Wire.

and component abundances. The image dimensions are 128 × 128 pixels, and a complete 1024-channel spectrum is collected at each pixel. Figure 2(b) shows a typical single-pixel spectrum for the Cu/Mn/Ni wire. The discrete nature of the data is clearly evident, and the SNR is sufficiently low so that the presence of Ni cannot be clearly discriminated from the background.

3.1.2. Results and analysis. In order to test the bootstrap Poisson regression with the robust nonparametric regression baseline removal (BPR-RR) method, we introduce other four state-of-the-art methods for comparison purpose. The first method uses traditional Anscombe transform combined with Wiener filter (ATW) to remove the Poisson noise and local medians

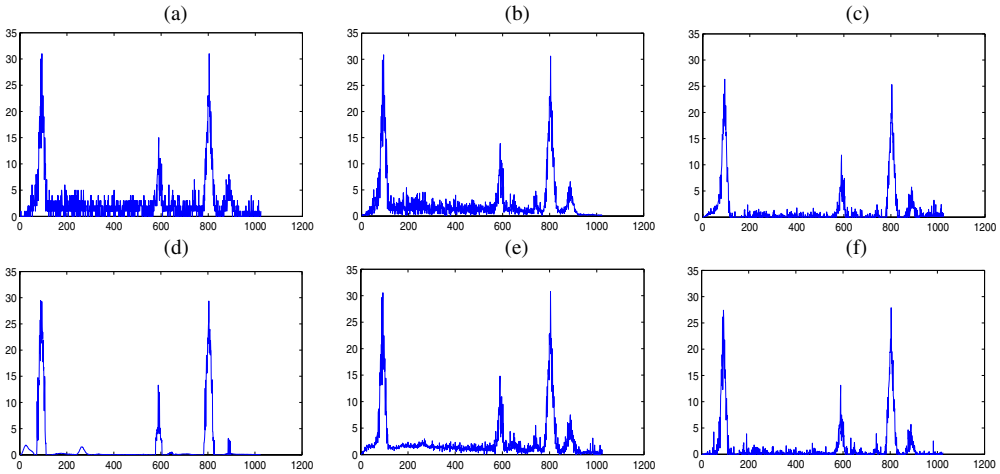


Figure 4. Typical single-pixel spectra of Mn/Ni/Cu alloy wire (line e in figure 2(a), the line displayed in figure 3) for (a) original data, (b) BLS-GSM method, (c) ATW-LM method, (d) MPG-ALS method, (e) SURE-LET method and (f) BPR-RR method (vertical axis: photon counts; horizontal axis: eV).

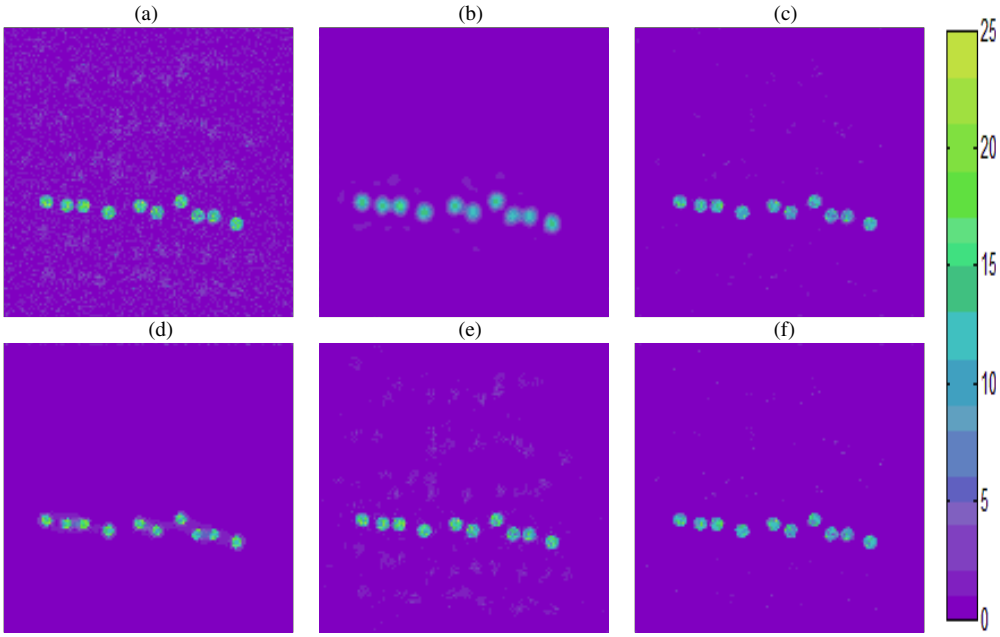
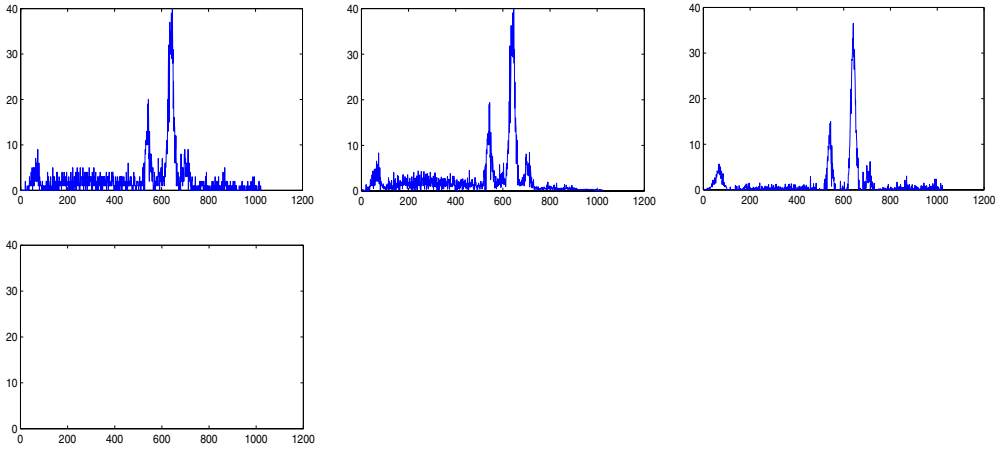


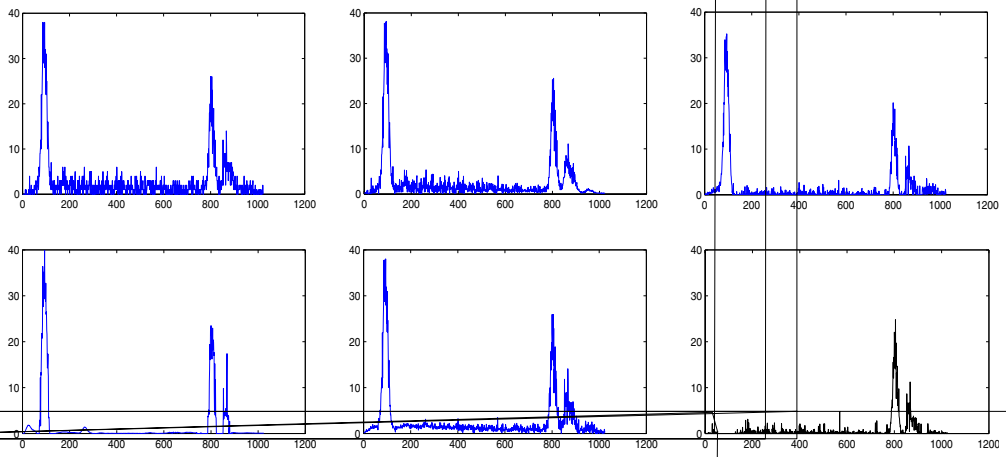
Figure 5. The x-ray spectral images at the energy channel of Cr's KL2 line energy (5.41 keV) for (a) original data, (b) BLS-GSM method, (c) ATW-LM method, (d) MPG-ALS method, (e) SURE-LET method and (f) BPR-RR method.

SURE-LET method also has the worst baseline drift removing performance compared with the other methods (figures 4(e), 6(e) and 8(e)).

Among all the five methods, the BPR-RR and ATW-LM methods achieve the best performance in terms of both denoising and preserving the original photon counts.


40
30
20
10
0
0 200 400 600 800 1000 1200

40
30
20
10
0
0 200 400 600 800 1000 1200


0 10 20 30 40
0 200 400 600 800 1000 1200

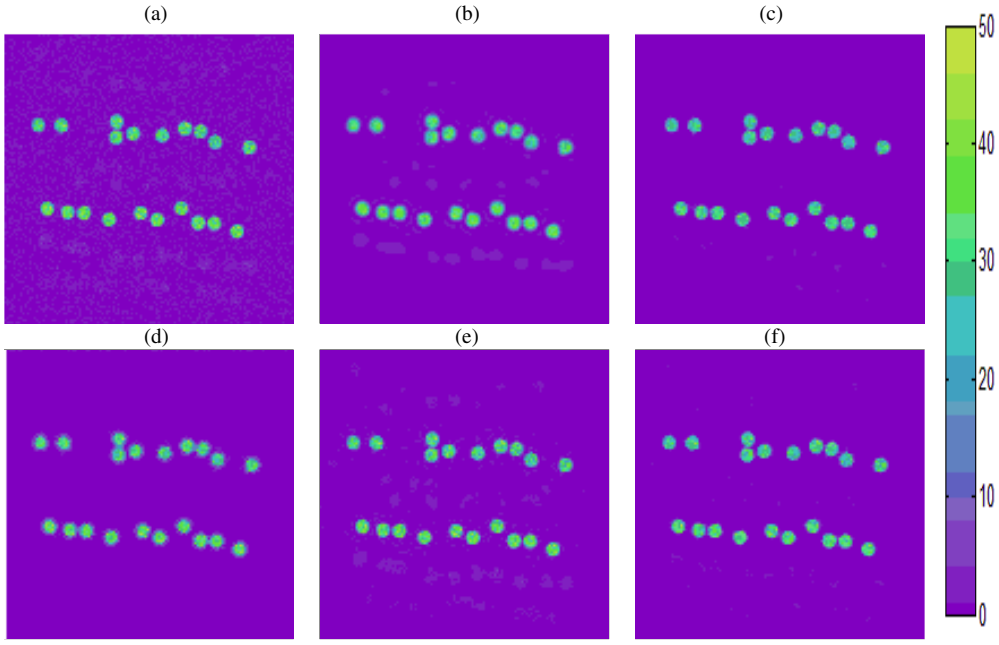


Figure 9. The x-ray spectral images at the energy channel of Fe's KL2 line energy (6.39 keV) for (a) original data, (b) BLS-GSM method, (c) ATW-LM method, (d) MPG-ALS method, (e) SURE-LET method and (f) BPR-RR method.

of standard image processing methods such as threshold, it is easy to pick up the data in the background area. The signal mean is easily calculated by using the data on the areas of wire dots. Finally, the SNR of pre-processed data is computed as a ratio of the signal mean to the noise standard deviation (the square root of noise variance).

Figure 10 shows the SNRs of the images in the KL2 line energy channels of the six component elements. Figure 10 displays that component elements with small SNRs still have the relative small SNRs in the processed data. The MPG-ALS method and BPR-RR method can enhance the SNRs of low-concentration elements (Cr, Mn and Zn) to almost the same level as the SNRs of high-concentration elements (Ni, Fe and Cu). The ATW-LM method, SURE-LET method and BPR-RR method have similar high computed SNRs, while our method has the highest computed SNRs among these methods. (All the elements' line energy data are referred to the database of the National Institute of Standards and Technology of the United States.)

3.2. Standard biological sample experiment

3.2.1. Experimental data. The two commercially available standard biological samples, bovine liver (NIST 1577a) and pig liver (GBW 08551) with known element concentrations, are imaged with μSXRF at the beamline BL15U at Shanghai Synchrotron Radiation Facility (Shanghai, China). These two samples were separately pressed into an $\sim$5 mg cm^{-2} tablet, and then sandwiched between double Mylar films. The image dimensions of bovine liver and pig liver are 11×11 and 10×10, respectively. Both samples are scanned with a complete 2048-channel spectrum at each pixel. More details about the beamline station and the preparation of samples can be found in Wang *et al* (2010).

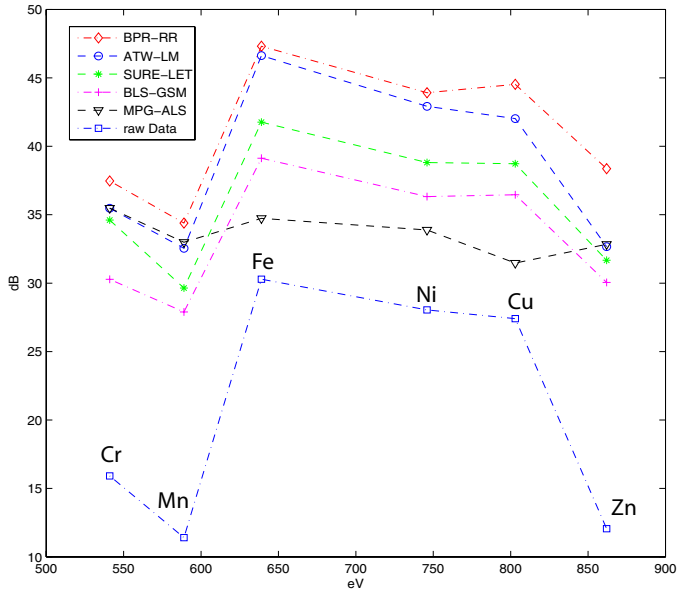


Figure 10. The signal-to-noise ratios estimated from the x-ray spectral images of each element's KL2 line energy. The elements from the left to right are Cr (5.41 keV), Mn (5.89 keV), Fe (6.39 keV), Ni (7.46 keV), Cu (8.03 keV) and Zn (8.62 keV).

3.2.2. Results and analysis. Based on the results of section Section 3.1, we choose the ATW-LM method, SURE-LET method and our BPR-RR method to preprocess the two standard biological samples. However, the SURE-LET method failed to process the spectra data of these samples because the quantity of these two samples data is too small to be processed by the SURE-LET method. All the parameters of the other methods were set as in section Section 3.1.

We further use the method described in Marco *et al* (1999) to implement the quantitative analysis of these two biological samples. In μSXRF imaging, the varied current intensity of x-ray and the differences in thickness and density of thin biological samples can cause considerable errors and undesirable precision.

Table 1. The ratios R_{ic} obtained from the raw bovine liver (NIST 1577a) data and the data preprocessed by the ATW-LM and BPR-RR methods.

	$R_{ic} \times 10^2$		
	Raw data	ATW-LM	BPR-RR
Ca	0.75 ± 0.04	0.69 ± 0.04	0.77 ± 0.05

to separate and integrate different elements' characteristic peaks and Compton peaks in the raw liver data, BPR-RR and ATW-LM method preprocessed liver data.

Table 2 shows that Poisson denoising and baseline drift removal methods do have a positive effect on the accuracy and precision for the quantitative analysis of the different trace elements. Among the three different preprocessing methods, the BPR-RR method has produced best quantitative analysis of element concentrations that are in good agreement with the certified values. The BPR-RR method has obtained high accuracy and high precision for quantitative analysis of the trace element Ca and Cu. Although the quantitative analysis of Zn is not as good as those of Ca and Cu, the BPR-RR method still has good accuracy and good precision for quantitative analysis of trace element Zn. The possible cause of the decrease of accuracy and precision for quantitative analysis of Zn is already mentioned in section Section 3.1.2. Both bovine liver and pig liver contain the elements Cu and Zn. The peaks of the spectra of Cu and Zn are so close that even an excellent tool such as PyMCA is not able to separate these two elements in the spectra thoroughly and precisely. As for the element Fe, the BPR-RR method has achieved good accuracy but poor precision for quantitative analysis due to the calculated standard deviation uncertainty (± 109) being significantly larger than the certified standard deviation uncertainty (± 40). One explanation for this is that the concentration of element Fe in the pig liver (GBW 08551) is very high to introduce large photon counts in the spectrum. Large photon counts assume an approximately normal distribution rather than the Poisson distribution (Haight 1967). Therefore, the method that is designed to deal with the low photon count data (e.g. Cu, Ca and Zn) will not have the excellent denoising effect for the high photon count data of element Fe in the pig liver.

4. Conclusion

In this paper, we have introduced a bootstrap Poisson regression and robust nonparametric regression method to reduce Poisson noise and baseline drift in the multidimensional x-ray spectral data. The proposed method is very effective to improve the SNRs of the raw x-ray spectral data. The comparison with other competing methods shows that the BPR-RR method offers performance better than some state-of-the-art approaches. By using two standard biological samples and applying Compton peak standardization in μSXRF quantitative imaging, the BRP-RR preprocessing method can ensure satisfying accuracy and precision for quantitative analysis of the trace elements in biological samples. The concentrations of Ca, Fe, Cu and Zn in the standard reference material (GBW 08551, pig liver) determined by BRP-RR preprocessing and subsequent quantitative imaging method were in good agreement with the certified values. This work can be extended along several directions in the future. First, the bootstrap samples can be sampled through nonparametric ways to converge more quickly to the statistic features of the original raw data. Second, Poisson regression can be computed within the Bayesian framework, where *a priori* information about the expectation of photon count data can be introduced to improve its performance. Third, our algorithm cannot achieve satisfactory performance in the case of extreme local discontinuities (or local nonhomogeneity) in all directions from the current scanning point of interest in the sample, though, such a case is a rare occurrence in modern high-resolution biomedical x-ray spectral imaging. At last, the Poisson denoising and baseline removal method are automatically adaptive not only to the low-concentration elements but also to the high-concentration elements that produce high photon count data (such as the element Fe in liver data), while our proposed method is apt to deal with the low photon count data.

Acknowledgments

The authors would like to thank Dr Paul Kotula for providing the wires' EDS dataset. This work was partially performed at the Shanghai Synchrotron Radiation Facility (SSRF) in China and supported by the National Basic Research Program of China (973 Program 2010CB834300, 2011CB933403), National Natural Science Foundation of China (61271320, 21175136 and 60872102) and the small animal imaging project (06-545). The authors would also like to thank the anonymous reviewers whose contributions considerably improved the quality of this paper.

References

Anscombe F J 1948 The transformation of Poisson, binomial and negative-binomial data *Biometrika* **35** 246–54

Bekemans B, Vincze L, Somogyi A, Drakopoulos M, Kempenaers L, Simionovici A and Adams F 2003 Quantitative x-ray fluorescence analysis at the ESRF ID18F microprobe *Nucl. Instrum. Methods Phys. Res.* B **199** 396–401

Bonettini S and Ruggiero V 2011 An alternating extragradient method for total variation-based image restoration from Poisson data *Inverse Problems* **27** 095001

Boulanger J, Kervrann C, Bouthemy P, Elbau P, Sibarita J B and Salamero J 2010 Patch-based nonlocal functional for denoising fluorescence microscopy image sequences *IEEE Trans. Med. Imaging* **29** 442–54

Boulanger J, Sibarita J B, Kervrann C and Bouthemy P 2008 Non-parametric regression for patch-based fluorescence microscopy image sequence denoising *5th IEEE Int. Symp. on Biomedical Imaging* pp 748–51

Chan R H and Chen K 2007 Multilevel algorithm for a Poisson noise removal model with total-variation regularization *Int J. Comput. Math.* **84** 1183–198

Cleveland W S 1979 Robust locally weighted regression and smoothing scatterplots *J. Am. Stat. Assoc.* **74** 829–36

Dahlbom M 2002 Estimation of image noise in PET using the bootstrap method *IEEE Trans. Nucl. Sci.* **49** 2062–66

Davison A C and Hinkley D V 1997 *Bootstrap Methods and their Application* (Cambridge: Cambridge University Press) chapter 7 p 327

Efron B 1979 Bootstrap methods: another look at the jackknife *Ann. Stat.* **7** 1–26

Eilers P H 2004 Parametric time warping *Anal. Chem.* **76** 404–11

Friedrichs M S 1995 A model-free algorithm for the removal of baseline artifacts *J. Biomol. NMR* **5** 147–53

Geraki K, Farquharson M J and Bradley D A 2004 X-ray fluorescence and energy dispersive x-ray diffraction for the quantification of elemental concentrations in breast tissue *Phys. Med. Biol.* **49** 99–110

Gherase M R and Fleming D E 2011 A calibration method for proposed XRF measurements of arsenic and selenium in nail clippings *Phys. Med. Biol.* **56** 215–25

Haight F A 1967 *Handbook of the Poisson Distribution* (New York: John Wiley & Sons Inc.)

Haynor D R and Woods S D 1989 Resampling estimates of precision in emission tomography *IEEE Trans. Med. Imaging* **8** 337–43

Jenkins R, Gould R W and Gedcke D 1995 *Quantitative X-Ray Spectrometry* 2nd edn (New York: Marcel Dekker)

Keenan M R 2007 Multivariate analysis of spectral images composed of count data *Techniques and Applications of Hyperspectral Image Analysis* chapter 5 (West Sussex, UK: John Wiley & Sons Inc.)

Kervrann C and Trubuil A 2004 An adaptive window approach for Poisson noise reduction and structure preserving in confocal microscopy *IEEE Int. Symp. on Biomedical Imaging* vol 1 pp 788–91

Kolaczyk E D 1999 Bayesian multiscale models for Poisson processes *J. Am. Stat. Assoc.* **94** 920–33

Kolaczyk E D 2000 Nonparametric estimation of intensity maps using Haar wavelets and Poisson noise characteristics *Astrophys. J.* **534** 490–505

Kotula P G, Keenan M R and Michael J R 2003 Automated analysis of SEM X-ray spectral images: a powerful new microanalysis tool *Microsc. Microanal.* **9** 1–17

Le T, Chartrand R and Asaki T J 2007 A variational approach to reconstructing images corrupted by Poisson noise *J. Math. Imaging Vis.* **27** 257–63

Lefkimmiatis S, Maragos P and Papandreou G 2009 Bayesian inference on multiscale models for Poisson intensity estimation: applications to photon-limited image denoising *IEEE Trans. Image Process.* **18** 1724–41

Liland K H, Almoy T and Mevik B H 2010 Optimal choice of baseline correction for multivariate calibration of spectra *Appl. Spectrosc.* **64** 1007–16

Luisier F and Blu T 2008 SURE-LET multichannel image denoising: interscale orthonormal wavelet thresholding *IEEE Trans. Image Process.* **17** 482–92

Makitalo M and Foi A 2011 Optimal inversion of the anscombe transformation in low-count poisson image denosing *IEEE Trans. Image Process.* **20** 99–109

Marco L M, Greaves E D and Alvarado J 1999 Analysis of human blood serum and human brain samples by total reflection x-ray fluorescence spectrometry applying Compton peak standardization *Spectrochim. Acta* B **54** 1469–80

Mounicou S, Szpunar J and Lobinski R 2009 Metallomics: the concept and methodology *Chem. Soc. Rev.* **38** 1119–38

Palakkal S and Prabhu K M M 2012 Poisson image denoising using fast discrete curvelet transform and wave atom *Signal Process.* **92** 2002–17

Popescu B F *et al* 2009 Mapping metals in Parkinson's and normal brain using rapid-scanning x-ray fluorescence *Phys. Med. Biol.* **54** 651–63

Qin Z, Toursarkissian B and Barry L 2011 Synchrotron radiation x-ray fluorescence microscopy reveals a spatial association of copper on elastic laminae in rat aortic media *Metallomics* **3** 823–28

Ruckstuhl A F, Jacobson M P, Field S W and Dodd J A 2001 Baseline subtraction using robust local regression estimation *J. Quant. Spectrosc. Radiat. Transfer.* **68** 179–93

Sole V A, Papillon E, Cotte M, Walter Ph and Susini J 2007 A multiplatform code for the analysis of energy-dispersive x-ray fluorescence spectra *Spectrochim. Acta* B **62** 63–8

Spring B Q and Clegg R M 2009 Image analysis for denoising full-field frequency-domain fluorescence lifetime images *J. Microsc.* **235** 221–37

Timmerman K E and Nowak R D 1999 Multiscale modeling and estimation of Poisson processes with application to photon-limited imaging *IEEE Trans. Inform. Theory* **45** 846–62

Twining B S, Baines S B, Fisher N S, Maser J, Vogt S, Jacobsen C, Tovar-Sanchez A and Sanudo-Wilhelmy S A 2003 Quantifying trace elements in individual aquatic protist cells with a synchrotron x-ray fluorescence microprobe *Anal. Chem.* **75** 3806–16

Van Grieken R E and Markowicz A A 2002 *Handbook of X-ray Spectrometry* 2nd edn (New York: Marcel Dekker) chapter 3

Wang H J, Wang M, Wang B, Meng X Y, Wang Y, Li M, Feng W Y, Zhao Y L and Chai Z F 2010 Quantitative imaging of element spatial distribution in the brain section of a mouse model of Alzheimer's disease using