Research Article

An Efficient Method for Detection of Outliers in Tracer Curves Derived from Dynamic Contrast-Enhanced Imaging

Linning Ye ¹ and Zujun Hou²

¹Department of Computer Science, Southwestern University of Finance and Economics, Chengdu 610074, China

²FITPU Healthcare Pte. Ltd., Block 79 Ayer Rajah Crescent, JTC Launch Pad #05-03, Singapore

Correspondence should be addressed to Linning Ye; yeln.t@swufe.edu.cn

Received 17 December 2017; Accepted 6 February 2018; Published 12 April 2018

Academic Editor: Xosé M. Pardo

Copyright © 2018 Linning Ye and Zujun Hou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Presence of outliers in tracer concentration-time curves derived from dynamic contrast-enhanced imaging can adversely affect the analysis of the tracer curves by model-fitting. A computationally efficient method for detecting outliers in tracer concentration-time curves is presented in this study. The proposed method is based on a piecewise linear model and implemented using a robust clustering algorithm. The method is noniterative and all the parameters are automatically estimated. To compare the proposed method with existing Gaussian model based and robust regression-based methods, simulation studies were performed by simulating tracer concentration-time curves using the generalized Tofts model and kinetic parameters derived from different tissue types. Results show that the proposed method and the robust regression-based method achieve better detection performance than the Gaussian model based method. Compared with the robust regression-based method, the proposed method can achieve similar detection performance with much faster computation speed.

1. Introduction

Dynamic contrast-enhanced imaging using computed tomography (DCE-CT) or magnetic resonance imaging (DCE-MRI) is a functional imaging method that can be used for in vivo assessment of tumor perfusion and allows for quantitative estimation of physiological parameters pertaining to tumor microcirculation [1, 2]. Analysis of DCE imaging data can be performed by fitting a tracer kinetic model to the tissue tracer concentration-time curves estimated from the DCE images.

It is well-recognized that noise in the DCE imaging dataset poses a major problem for model-fitting. This is often further compounded by movement artefacts as the dynamic images are acquired over a period of time. Sudden body movements during imaging can result in distinct spikes in the sampled tracer concentration-time curve, on top of the smaller background fluctuations due to image noise. Such spikes in the data points are called outliers (Figure 1) and they

can affect the outcome of model-fitting dramatically. Outliers could also appear in the tracer concentration curves as artefacts due to prior data processing steps, for example, image registration error. In DCE-MRI, tracer concentration is estimated by additional calculations performed on the MR signal and outliers could appear due to certain computational operations, such as division by small numbers contaminated by noise. Outliers are commonly encountered in tracer concentration-time curves sampled from DCE imaging and screening of concentration curves for outliers before model-fitting should be recommended.

Visually, an outlier appears to be “out-of-place” and does not conform to the mass of the data. If one has prior knowledge of the data, then a model can be used to describe the majority of data and outlier detection can be carried out by the approach of hypothesis testing of the suspicious data point against the model. For example, a frequently used model for hypothesis testing of outliers is the Gaussian model [3, 4]. Let $\{y(t_i) \mid i = 1, \dots, n\}$ denote a set of data at

FIGURE 1: Example of an outlier in a tissue tracer concentration-time curve sampled from DCE-CT images (HU: Hounsfield unit).

timepoints t_i . The Gaussian model assumes that the data is sampled from a Gaussian distribution, $G(\mu, \sigma^2)$, with mean μ and standard deviation σ :

$$y(t_i) \sim \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}. \quad (1)$$

For a random variable following the Gaussian distribution, the probability of sampling a value outside the interval $\mu \pm \sigma z_p$ would be $2(1 - p)$, where $z_p = \sqrt{2} \text{erf}^{-1}(2p - 1)$, and $p \in (0, 1)$ is the p -quantile function of the standard normal distribution. Hence, hypothesis testing for outlier detection can be constructed simply as a thresholding step:

$$\left| \frac{y(t_i) - \mu}{\sigma} \right| > z_p, \quad (2)$$

and $y(t_i)$ would be considered an outlier if (2) holds true.

In practice, the Gaussian model would be oversimplified as the observed data could be derived from a physical process and the underlying dynamics could be more complex than being constant over time (approximated by mean μ). A natural extension is the truncated power series:

$$y(t) = \beta_0 + \beta_1 t + \dots + \beta_r t^r + \epsilon, \quad (3)$$

where $\{\beta_0, \dots, \beta_r\}$ are the model parameters and ϵ is the error term which follows the Gaussian distribution $G(0, \sigma^2)$. Although the model in (3) describes a polynomial of power r , it is a linear model with respect to the model parameters $\{\beta_0, \dots, \beta_r\}$. If the model parameters could be estimated from the observed data, then this model will predict a value $\hat{y}(t_i)$ at time t_i , and the residue $e_i = y(t_i) - \hat{y}(t_i)$ follows $G(0, \sigma^2)$. Thus, hypothesis testing for outlier detection can similarly be constructed as

$$\left| \frac{y(t_i) - \hat{y}(t_i)}{\sigma} \right| > z_p. \quad (4)$$

To estimate the model parameters $\beta = (\beta_0, \dots, \beta_r)$, one can proceed to minimize

$$\min_{\beta} \sum_{i=1}^n (y(t_i) - \beta_0 - \beta_1 t_i - \dots - \beta_r t_i^r)^2. \quad (5)$$

Setting the derivatives to zero and solving the resulting linear equations would yield the well-known ordinary least squares (OLS) solution. However, the OLS estimator would be biased if the data is heavily tailed or contains outliers. An unbiased solution can be derived through the inclusion of weights:

$$\min_{\beta} \sum_{i=1}^n w(e_i) (y(t_i) - \beta_0 - \beta_1 t_i - \dots - \beta_r t_i^r)^2, \quad (6)$$

where $w(e_i)$ is a weight function of the residue which is designed to be small for large residues. By this design, outliers would have less influence in determining the solution. This technique has been referred to as the weighted least squares (WLS) method, and, in robust statistics, it is more commonly called robust regression [5–8]. Previous studies have been performed on the design of the weight function and an example is the Welsch function:

$$w_{\text{Welsch}} = \exp \left\{ - \left(\frac{e}{\delta} \right)^2 \right\}, \quad (7)$$

where δ is a scaling parameter stipulated by the user.

Computation is a major issue in practical implementation of robust regression because of the need to iteratively alternate between the steps of least squares estimation and weight calculation, until convergence of solution. In a DCE imaging dataset, multiple concentration-time curves have to be processed in order to generate parametric maps of physiological parameters. A more efficient outlier detection method is warranted.

In this study, a new method for outlier detection is presented, which assumes that the underlying data dynamics can be approximated using a piecewise linear model. A robust clustering algorithm is designed to efficiently construct the piecewise linear model, which is subjected to hypothesis testing to determine the presence of outliers. As the piecewise linear model is derived in a nonrepetitive manner, the proposed method becomes much more efficacious than conventional robust regression-based methods.

2. Methods

2.1. Piecewise Linear Model. Graphically, a piecewise linear function in one dimension consists of line segments and a simple piecewise linear function (Figure 2) is given as follows:

$$\varphi(t) = \begin{cases} \ell_1(t) & \text{for } t \leq t_0 \\ \ell_2(t) & \text{for } t > t_0, \end{cases} \quad (8)$$

where $\ell_k(t) = \beta_{0,k} + \beta_{1,k}t$, $k = 1, 2$ is a line segment with slope $\beta_{1,k}$ and intercept $\beta_{0,k}$. The function could be either

FIGURE 2: Examples of discontinuous (a) and continuous (b) piecewise linear functions.

continuous or discontinuous at the point t_0 (Figure 2). In this study, as the detection process runs in the manner similar to a moving window filter which analyzes the local neighborhood of each data point, a piecewise linear model would be a reasonable (sufficient) approximation of the underlying dynamics in a local neighborhood of the studied point, and hence the piecewise linear model forms the basic assumption of the observed data:

$$y(t) = \varphi(t) + \epsilon + o(t), \quad (9)$$

where $\epsilon \sim G(0, \sigma^2)$ and $o(t)$ models the outlier. It should be pointed out that (9) is just one way to model the non-Gaussian noise arising from practice for which the true generation mechanism as well as the true distribution of the noise is unknown. There exist other approaches to non-Gaussian noise modeling, for example, the mixture Gaussian model [9] and the Laplacian distribution model [10].

2.2. Numerical Implementation for Outlier Detection: A Clustering Approach. If one was to follow the conventional approach for outlier detection based on regression analysis, one would proceed to fit the data with the piecewise linear model and perform hypothesis testing by comparing the discrepancy between observed data and predicted value by the fitted model. In the following a clustering approach is presented so that regression and outlier detection is carried out in one sweep.

The basic idea of data analysis using piecewise linear models is that the data can be classified into two or more groups or clusters, each consisting of elements which share similar features. For example, if the deviation of a data point from a particular cluster centroid is smaller than a threshold, the data point can be assigned to that cluster. For outlier detection, the most interesting part would be to find the cluster to which the suspicious data likely belongs, after which a simple linear fitting would suffice for the desired hypothesis testing. Bearing this in mind, the detection process would consist of two stages: clustering and testing.

2.2.1. Clustering. As the local neighborhood of each data point is assumed to follow the piecewise linear model which is composed of two line segments, the local neighborhood could be classified into two groups or two clusters. To account for the presence of possible outliers, an additional cluster is

allocated, leading to three possible clusters in total. The elements in each cluster are arranged in ascending order of time. To estimate signal deviation within cluster, the average value of the absolute first-order difference of the cluster elements is computed and denoted as δ^i , $i = 1, 2, 3$. If in cluster $\mathbb{C}^i$, the k th element is $(t_{(k)}^i, y_{(k)}^i)$, then the absolute first-order difference between the k th and the $(k+1)$ th element is defined as

$$\left| \frac{y_{(k+1)}^i - y_{(k)}^i}{t_{(k+1)}^i - t_{(k)}^i} \right|. \quad (10)$$

Each cluster, $\mathbb{C}^i$, consists of the following three records:

$$\mathbb{C}^i \rightarrow \begin{cases} y_{(1)}^i, \dots, y_{(n_i)}^i \\ t_{(1)}^i, \dots, t_{(n_i)}^i \\ \delta^i. \end{cases} \quad (11)$$

A key consideration in the clustering process is to determine whether a new data $y(t_k)$ complies with the existing elements of a cluster. If $y(t_k)$ is assigned to a particular cluster $\mathbb{C}^i$, the incremental absolute deviation of the cluster would be

$$\delta_a = \left| \frac{y(t_k) - y_{(n_i)}^i}{t_k - t_{(n_i)}^i} \right|. \quad (12)$$

It is utilized for calculating a distance metric from the original cluster deviation δ^i :

$$\text{dist}_k^i = |\delta^i - \delta_a|. \quad (13)$$

If the distance is less than a predetermined threshold T_{nbs} ,

$$\text{dist}_k^i < T_{\text{nbs}}, \quad (14)$$

then $y(t_k)$ is assigned to $\mathbb{C}^i$ and the records of the cluster are updated accordingly. In the situation that (14) holds true for multiple clusters, the decision would be to choose the cluster which contains the element with the nearest temporal distance (i.e., minimum $t_k - t_{(n_i)}^i$) from $y(t_k)$.

An example is given in Figure 3 to illustrate the clustering process. The data of interest is marked with a cross and eight neighboring points (i.e., half-window size is four) are selected to assist in the inference of whether the data of interest is an

FIGURE 3: An example to illustrate the clustering process, where “x” stands for the data point of interest with the local neighborhood consisting of 4 points to the left and right of “x” and circles and squares represent the two clusters: (a) initialization of the first cluster, (b) next three datasets classified into the first cluster, (c) the fifth data initialized to the second cluster, and (d) the rest classified into the first cluster.

outlier. The clustering starts with assigning the first data from the left to the first cluster as shown in Figure 3(a). After that, the clustering proceeds from left to right and the next three points are found to be close to the first cluster; thus they are classified into the same cluster, as shown in Figure 3(b). When the fifth point (i.e., the point of interest) is encountered, it is found that its distance to the first cluster is too large; therefore, a new cluster is generated to accommodate this point, as shown in Figure 3(c). Similarly, the 4 points to the right of the point of interest are found to be close to the first cluster. Consequently, there are two clusters in this example and the final clustering result is shown in Figure 3(d).

The pseudocode of the clustering algorithm is described as shown in Pseudocode 1.

2.2.2. Hypothesis Testing for Outlier Detection. After data clustering, the next step is to select the right cluster for data fitting. Usually the one with the largest number of elements

would be selected. In the situation of two clusters with equal number of elements, then the cluster containing the element with the nearest temporal distance would be chosen. After data fitting, a value $\hat{y}(t_k)$ can be predicted and compared with the observed data. If the distance is larger than a threshold T_{err} ,

$$|y(t_k) - \hat{y}(t_k)| > T_{err}, \quad (15)$$

then an outlier would be detected.

2.3. Threshold Parameters. There are two threshold parameters, T_{nbs} and T_{err} , involved in the algorithm. These threshold parameters can be determined from the statistics of the dataset, as follows. T_{nbs} is used for assessing cluster compliance by comparing the incremental deviation due to each additional data point with the overall cluster deviation. Note that the underlying signal is usually not constant over time, adding to signal intrinsic deviation, δ_{signal} , on top of deviation

```

Initialization: three empty clusters with deviation 0;
Loop over all data  $k$  from 1 to  $n$ 
    if  $C^1$  is empty, then assign to  $C^1$ 
    elseif  $C^2$  is empty
        if data  $k$  complies with  $C^1$ , then assign to  $C^1$ 
        else assign to  $C^2$ 
    elseif  $C^3$  is empty
        if data  $k$  complies with  $C^1$  or  $C^2$ , then assign accordingly to  $C^1$  or  $C^2$ 
        else assign to  $C^3$ 
    else assign data  $k$  accordingly among  $C^1$ ,  $C^2$  or  $C^3$ 
End of loop

```

PSEUDOCODE 1

due to noise, δ_{noise} . Let $y'(t_k)$ denote the first-order difference of the input signal. Then δ_{signal} is estimated by

$$\delta_{\text{signal}} = \left| \text{median} \{y'(t_k)\} \right|. \quad (16)$$

As for δ_{noise} , it is related to the standard deviation of noise σ . In a similar fashion to [11], the estimation of σ is based on the median of absolute deviation (MAD) of $\{y'(t_k)\}$:

$$\hat{\sigma} = 1.4826 * \frac{\text{MAD} \{y'(t_k)\}}{\sqrt{2}}, \quad (17)$$

where the value of 1.4826 ensures an unbiased estimate when the noise follows a Gaussian distribution, and division by $\sqrt{2}$ transforms the estimation of standard deviation of noise in $\{y'(t_k)\}$ to the standard deviation of noise in $\{y(t_k)\}$. For convenience, let $\hat{\sigma}'$ denote the standard deviation of noise in $\{y'(t_k)\}$, which relates to $\hat{\sigma}$ by

$$\hat{\sigma}' = \sqrt{2}\hat{\sigma}. \quad (18)$$

Then, an estimate of T_{nbs} goes as follows:

$$T_{\text{nbs}} = \delta_{\text{signal}} + \lambda_{\text{nbs}}\hat{\sigma}', \quad (19)$$

where λ_{nbs} plays the same role as z_p , as mentioned earlier in the hypothesis testing approach. λ_{nbs} is set to 1.5 in this study, which implies the assumption that about 93.3% of the noisy data in $\{y'(t_k)\}$ could be inliers.

As T_{err} is used for characterizing $y(t_k) - \hat{y}(t_k)$, which pertains to the noise in $\{y(t_k)\}$, the estimation of T_{err} is straightforward:

$$T_{\text{err}} = \lambda_{\text{err}}\hat{\sigma}, \quad (20)$$

where λ_{err} is set to $2.5\sqrt{2}$.

As pointed out in [12, 13], statistical methods tend to be oversensitive to noise in signal hypothesis testing, particularly when the sample size is limited. In such cases, a simple

solution is to add a small positive number γ to the threshold parameter. Then (19) and (20) are modified as follows:

$$T_{\text{nbs}} = \begin{cases} \delta_{\text{signal}} + \lambda_{\text{nbs}}\hat{\sigma}', & \text{if } \hat{\sigma}' > \hat{\sigma}'_c \\ \delta_{\text{signal}} + \lambda_{\text{nbs}}\hat{\sigma}' + \gamma, & \text{otherwise,} \end{cases} \quad (21)$$

$$T_{\text{err}} = \begin{cases} \lambda_{\text{err}}\hat{\sigma} & \text{if } \hat{\sigma} > \hat{\sigma}_c \\ \lambda_{\text{err}}\hat{\sigma} + \gamma & \text{otherwise,} \end{cases}$$

where the cut-off noise scale $\hat{\sigma}'_c$ is set to 0.5 and γ is a small percentage of the average signal intensity, that is, 5%.

3. Monte Carlo Simulation

Computer simulation experiments were performed to evaluate the performance of the proposed outlier detection method under various noise conditions. Similar to [14], the arterial input function $C_a(t)$ was modeled as a sum of gamma-variate functions:

$$C_a(t) = \Gamma(t, t_0, \alpha_0, \eta_0) + 0.2\tilde{\Gamma}(t, t_1, \alpha_1, \eta_1) + 0.05\tilde{\Gamma}(t, t_2, \alpha_2, \eta_2), \quad (22)$$

where

$$\Gamma(t, t_0, \alpha_0, \eta_0) = (t - t_0)^{\alpha_0} e^{-(t-t_0)/\eta_0}, \quad t \geq t_0 \quad (23)$$

refers to the initial peak of the arterial input function. $\tilde{\Gamma}$ is a gamma-variate function normalized (i.e., scaled) to ensure equal maximal value with the initial peak Γ , such that the second term in (22) describes the first recirculation peak reaching a height about 20% of the initial peak. Following [14], values of the various parameters in the gamma-variate functions are listed in Table 1 and the sampling interval of the arterial input function was set as 1.6 s.

Tissue concentration-time curves $C_{\text{tiss}}(t)$ were generated using the generalized Tofts model [16]:

$$C_{\text{tiss}}(t) = C_a(t) \otimes [K_{\text{trans}} e^{-K_{\text{trans}} t / v_e} + v_p \delta(t)], \quad (24)$$

TABLE 1: Parameters of gamma-variate functions for simulating the arterial input function.

$\Gamma(t, t_0, \alpha_0, \eta_0)$	$t_0 = 32$ s	$\alpha_0 = 3$	$\eta_0 = 5$ s
$\tilde{\Gamma}(t, t_1, \alpha_1, \eta_1)$	$t_1 = 60$ s	$\alpha_1 = 3$	$\eta_1 = 10$ s
$\tilde{\Gamma}(t, t_2, \alpha_2, \eta_2)$	$t_2 = 90$ s	$\alpha_2 = 3$	$\eta_2 = 20$ s

TABLE 2: Kinetic parameters utilized in the simulation study.

Tissue type	K_{trans} (min^{-1})	v_e (ml/ml)	v_p (ml/ml)
(#1) Breast tumor [15]	0.25	0.45	0.06
(#2) Acute multiple sclerosis lesion [16]	0.056	0.18	0.1
(#3) Chronic multiple sclerosis lesion [16]	0.012	0.41	0.06

TABLE 3: Processing time (millisecond) by different outlier detection methods.

	Hampel	Proposed method	Welsch
Processing time (ms)	5	10	130

where $\otimes$ denotes convolution and $\delta(t)$ is the Dirac delta function. Values of kinetic parameters (K_{trans} , v_e , v_p) used for simulating these tracer curves were shown in Table 2.

Gaussian noise $G(0, \sigma_G^2)$ was added to the $C_{\text{tiss}}(t)$ curves with σ_G varying from 0.3 to 2.1 (in steps of 0.3) to simulate various noise conditions. For each Gaussian noise level σ_G , 1000 simulation runs were performed. In each simulation run, the presence of outliers was simulated by adding salt-and-pepper noise [17] at randomly selected time points in $C_{\text{tiss}}(t)$ where the concentration values were increased or reduced (randomly) by 80% of the maximum concentration value in $C_{\text{tiss}}(t)$. The percentage of the number of outliers in each simulated $C_{\text{tiss}}(t)$ curve was kept constant at 10%.

Since the locations of the outliers are known in each simulation, one can study performance of a outlier detector by computing quantitative measures using the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) as follows [18].

(1) Sensitivity is also called true positive rate:

$$\text{Sensitivity} = \frac{\#(\text{TP})}{\#(\text{TP}) + \#(\text{FN})}. \quad (25)$$

(2) Specificity is also called true negative rate:

$$\text{Specificity} = \frac{\#(\text{TN})}{\#(\text{TN}) + \#(\text{FP})}. \quad (26)$$

To compare the performance of the proposed outlier detection method with existing methods, two well-established outlier detectors were also implemented in the Monte Carlo simulations. These are, namely, the Hampel and Welsch methods [4, 11], which are, respectively, based on the Gaussian model and the robust regression approach mentioned before. The Hampel method replaces the parameters of mean and standard deviation in (2) with median and

MAD, respectively, to make the inference more resistant against outliers. The Welsch method detects outliers using the thresholding condition (4) after a robust regression model is derived.

All computations were performed by using Matlab™ (The MathWorks, Natick, MA, USA) and running on a desktop personal computer with 3 GHz CPU and 8 GB memory.

4. Results and Discussion

4.1. Simulation Results. Figure 4(a) shows a simulated noisy tissue concentration-time $C_{\text{tiss}}(t)$ curve where outliers are marked with red squares, which represent the ground truth. The dashed line in Figure 4(a) shows the simulated $C_{\text{tiss}}(t)$ curve prior to the addition of noise. Figures 4(b)–4(d) present the outliers detected by the Hampel method, Welsch method, and the proposed piecewise linear method, respectively. Compared with the ground truth in Figure 4(a), all methods could successfully detect outliers located in the relatively flat regions of the $C_{\text{tiss}}(t)$ curve. However, among the six outliers in upslope region of $C_{\text{tiss}}(t)$, the Hampel Gaussian-based method could not detect any of them, while the Welsch robust regression model succeeded in detecting five outliers with three false positives, and the proposed method also detected five of the outliers but with one false positive.

Figures 5–7 summarize the performance (i.e., sensitivity and specificity) of the Hampel, proposed, and Welsch methods for different tissue types. As illustrated in Figures 5–7, the proposed method generally achieved higher sensitivities than the Hampel method (by about 6–11%) with similar levels of specificity. The performance of the proposed method (in terms of sensitivity and specificity) is close to the performance of the Welsch method.

The average processing time of different methods is given in Table 3. We can find that the Hampel and Welsch methods

FIGURE 4: Example of a simulation run with outlier detection by different methods.

require about 5 and 130 ms, respectively, to process one $C_{tiss}(t)$ curve, while the proposed method takes about 10 ms. The results show that the computation speed of the Hampel method and the proposed method is much faster than the Welsch method.

4.2. Parameter Sensitivity. For the proposed method, threshold parameters were fixed during simulation experiments. Here, we evaluate the sensitivity of different parameters. The tracer concentration-time curves utilized in the evaluation

are simulated using kinetic parameters typical of acute multiple sclerosis. Figure 8 plots the performance index (sensitivity on the left y -axis and specificity on the right y -axis) with respect to threshold T_{err} under noise conditions $\sigma_G = 0.3$ and $\sigma_G = 2.1$. We can find that both sensitivity and specificity vary in a tight range, between 0.996 and 1 for sensitivity and between 0.992 and close to 1 for specificity, respectively, when $\sigma_G = 0.3$ (as shown in Figure 8(a)), which provides much freedom in determining the threshold. When $\sigma_G = 2.1$, as shown in Figure 8(b), the sensitivity drops below 0.9 after the threshold is larger than 20, and the specificity is nearly

FIGURE 5: Comparison of outlier detection with different Gaussian noise conditions, where the tracer concentration-time curves are simulated using kinetic parameters typical of breast tumors.

monotonically increasing from 0.99 to 1, which suggests that a threshold in the wide range between 1.5 and 10 times of σ_G would be reasonable.

Figure 9 shows the parameter sensitivity with respect to λ_{nbs} , which relates to the clustering process in the proposed method. When $\sigma_G = 0.3$, either the sensitivity or the specificity is greater than 0.99 for $\lambda_{nbs} \in [0.5, 3]$. In particular, both performance indexes are greater than 0.995 when $\lambda_{nbs} > 1.5$. When $\sigma_G = 2.1$, the specificity is greater than 0.994 for $\lambda_{nbs} \in [0.5, 3]$, and the sensitivity drops below 0.95 for $\lambda_{nbs} > 2.9$.

The effect of half-window size on the performance of the proposed method is shown in Figure 10 for the same simulation condition in Figure 9, which suggests that the

FIGURE 6: Comparison of outlier detection with different Gaussian noise conditions, where the tracer concentration-time curves are simulated using kinetic parameters typical of acute multiple sclerosis.

moving window size (2 times half-window size plus 1) 9 or 11 would be suitable.

5. Concluding Remarks

This study presents a method for detection of outliers in tracer concentration-time curves derived from DCE imaging. Salient features of the proposed method include (1) a robust clustering method for classification of data, (2) a piecewise linear model being utilized to describe underlying data dynamics from which distance measures of individual data points can be derived, and (3) outlier detection being formulated as a hypothesis testing of distance measures based on

FIGURE 7: Comparison of outlier detection with different Gaussian noise conditions, where the tracer concentration-time curves are simulated using kinetic parameters typical of chronic multiple sclerosis.

objective statistical threshold metrics, allowing automation of the method under different noise conditions without user intervention. Monte Carlo simulation experiments of various tissue types show that the proposed method compares favorably with two well-established techniques in terms of computation speed and performance (sensitivity and specificity). Specifically, the proposed method achieves detection accuracy similar to the robust regression-based method at the computation speed close to the Gaussian model based method.

Although the method has been designed to detect outliers in tracer curves from DCE imaging, it is expected to be also applicable in other scenarios where two main assumptions

FIGURE 8: Parameter sensitivity with respect to threshold T_{err} for the proposed method under noise condition $\{\sigma_G = 0.3 \text{ or } 2.1\}$, where the left vertical axis represents the sensitivity and the right vertical axis represents the specificity.

FIGURE 9: Parameter sensitivity relating to the clustering threshold for the proposed method under noise condition $\{\sigma_G = 0.3 \text{ or } 2.1\}$, where the left vertical axis represents the sensitivity and the right vertical axis represents the specificity.

hold: (1) underlying data can be approximated by a piecewise linear model and (2) outliers are present with noise that follows the Gaussian distribution and is independent and identically distributed.

FIGURE 10: Parameter sensitivity with respect to half-window size for the proposed method under noise condition $\{\sigma_G = 0.3 \text{ or } 2.1\}$, where the left vertical axis represents the sensitivity and the right vertical axis represents the specificity.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] T. S. Koh, S. Bisdas, D. M. Koh, and C. H. Thng, "Fundamentals of tracer kinetics for dynamic contrast-enhanced MRI," *Journal of Magnetic Resonance Imaging*, vol. 34, no. 6, pp. 1262–1276, 2011.
- [2] S. P. Sourbron and D. L. Buckley, "Tracer kinetic modelling in MRI: estimating perfusion and capillary permeability," *Physics in Medicine and Biology*, vol. 57, no. 2, pp. R1–R33, 2012.
- [3] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers, Massachusetts, Mass, USA, 3rd edition, 2012.
- [4] R. K. Pearson, *Mining Imperfect Data: Dealing with Contamination and Incomplete Records*, Society for Industrial and Applied Mathematics, Pennsylvania, Pa, USA, 2005.
- [5] P. J. Huber, "Robust regression: asymptotics, conjectures and Monte Carlo," *The Annals of Statistics*, vol. 1, pp. 799–821, 1973.
- [6] P. W. Holland and R. E. Welsch, "Robust regression using iterative reweighted least squares," *Communications in Statistics: Theory and Methods*, vol. 6, no. 9, pp. 813–827, 1977.
- [7] C.-M. Cheng and S.-H. Lai, "A consensus sampling technique for fast and robust model fitting," *Pattern Recognition*, vol. 42, no. 7, pp. 1318–1329, 2009.
- [8] S. B. Heinrich, "Efficient and robust model fitting with unknown noise scale," *Image and Vision Computing*, vol. 31, no. 10, pp. 735–747, 2013.
- [9] Z. Hou and T. S. Koh, "Image denoising using robust regression," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 243–246, 2004.
- [10] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532–1546, 2000.
- [11] R. R. Wilcox, *Introduction to Robust Estimation and Hypothesis Testing*, Wiley, New York, NY, USA, 1990.
- [12] A. C. Bovik, T. S. Huang, and D. C. Munson Jr., "Nonparametric tests for edge detection in noise," *Pattern Recognition*, vol. 19, no. 3, pp. 209–219, 1986.
- [13] Z. Hou and T. S. Koh, "Robust edge detection," *Pattern Recognition*, vol. 36, no. 9, pp. 2083–2091, 2003.
- [14] K. Murase, "Efficient method for calculating kinetic parameters using T1-weighted dynamic contrast-enhanced magnetic resonance imaging," *Magnetic Resonance in Medicine*, vol. 51, no. 4, pp. 858–862, 2004.
- [15] D. L. Buckley, "Uncertainty in the analysis of tracer kinetics using dynamic contrast-enhanced T1-weighted MRI," *Magnetic Resonance in Medicine*, vol. 47, no. 3, pp. 601–606, 2002.
- [16] P. S. Tofts, "Modeling tracer kinetics in dynamic Gd-DTPA MR imaging," *Journal of Magnetic Resonance Imaging*, vol. 7, no. 1, pp. 91–101, 1997.
- [17] O. Marques, *Practical Image and Video Processing Using MATLAB*, Wiley, Fusionopolis, Singapore, 2011.
- [18] Z. Hou, H.-K. Lam, W.-Y. Yau, and Y. Wang, "A variational formulation for fingerprint orientation modeling," *Pattern Recognition*, vol. 45, no. 5, pp. 1915–1926, 2012.