Computed Tomography Perfusion Imaging Denoising Using Gaussian Process Regression

Fan Zhu[1,2], Trevor Carpenter[2], David Rodriguez Gonzalez[1,2], Malcolm Atkinson[1], Joanna Wardlaw[2]

[1]Data-Intensive Research Group, School of Informatics
[2]Brain Research Imaging Centre, Division of Clinical Neuroscience
University of Edinburgh
Edinburgh, UK

E-mail: F.Zhu@ed.ac.uk

Abstract.
Objective: Brain perfusion weighted images acquired using dynamic contrast studies have an important clinical role in acute stroke diagnosis and treatment decisions. However, Computed Tomography (CT) images suffer from low contrast-to-noise ratios (CNR) as a consequence of the limitation of the exposure to radiation of the patient. As a consequence, the developments of methods for improving the CNR are valuable.
Methods: The majority of existing approaches for denoising CT images are optimized for 3D (spatial) information, including spatial decimation (spatially weighted mean filters) and techniques based on wavelet and curvelet transforms. However, perfusion imaging data is 4D as it also contains temporal information. Our approach using Gaussian process regression (GPR), which takes advantage of the temporal information, to reduce the noise level.
Results: Over the entire image, GPR gains a 99% CNR improvement over the raw images and also improves the quality of haemodynamic maps allowing a better identification of edges and detailed information. At the level of individual voxel, GPR provides a stable baseline, helps us to identify key parameters from tissue time-concentration curves and reduces the oscillations in the curve.
Conclusion: GPR is superior to the comparable techniques used in this study.

Keywords: Perfusion Imaging, CT, Gaussian Process Regression, Noise Reduction.

1. Introduction

Perfusion weighted images acquired using dynamic contrast studies may potentially have an important clinical role in acute stroke diagnosis and treatment decisions. Analysis of perfusion imaging provides quantitative measurements of haemodynamic parameters such as cerebral blood volume (CBV) [1, 2] and mean transmit time (MTT) [3, 4]. Sajjad (2008) [5] stated that the results of perfusion imaging analysis enable cerebral tissue information to be displayed that can been used in the management of acute strokes and tumors. Parameters such as CBV and time to peak (TTP) are

important in the context of thrombolysis. Therefore, improving the precision of these parameters may help to determine the lesion area when applying ordinal (per patient) threshold values. In the research setting, perfusion data has often been acquired using MRI [6, 7]. However, computed tomography (CT) imaging remains the most easily accessible and widely available modality for acute stroke patients [1, 8].

Since X-ray radiation increases the risk of inducing cancer [9, 10], CT scanning is constrained by a trade-off between image quality and the amount of radiation exposure to patients. Therefore CT imaging suffers from a low contrast to noise ratio (CNR) making post processing of the acquired images (to produce perfusion parametric maps) problematic. The noise leads to difficulty when attempting to estimate parameters such as cerebral blood flow (CBF), cerebral blood volume (CBV) and time to peak (TTP) [11]. The quality of these haemodynamic maps decreases dramatically with the increase of noise level [12]. The majority of current noise reduction approaches for CT images are optimised for 3D (spatial only) information, including spatial decimation (weighted mean filters, Gaussian filters [13, 14, 15]), techniques based on wavelet transforms [16, 17, 18] and curvelet transforms [19, 20, 21]. However, perfusion imaging also contains temporal information and attempts to use the temporal information to reduce noise have started recently [22].

In this paper, we propose Gaussian process regression based [23, 24] approaches, which make use of the 4D information to reduce noise. We also compare the performance of some spatial information based approaches and our Gaussian process regression based methods in terms of the quality of signal, CNR and the quality of the produced haemodynamic parametric maps.

2. Methods

2.1. Gaussian Process Regression

A Gaussian process is a generalization of the Gaussian probability distribution. Whereas a probability distribution describes random variables which are scalars or vectors, a stochastic process governs the properties of functions. A Gaussian process is a collection of random variables, any finite number of values that have a joint Gaussian distribution, and is widely used to solve regression problems [24]. Gaussian process regression was designed to be a prediction tool. It is usually used to predict the unknown dependent variable for any given independent variables based on known but noisy observation of the dependent and independent variables. The key idea of our method is to re-predict all given noisy dependent variables in order to reduce the noise, the details of which will be stated below.

A Gaussian process is specified by its mean and variance. In our algorithm, the denoised value is considered as the expected mean value and its variance is used to calculate the confidence interval.

In the case of CT imaging, we do not have access to the function values themselves,

but only to the noisy observations:

$$Y = f(T) + \epsilon \tag{1}$$

where $T = \{t_1, t_2, \cdots, t_n\}$ is an input vector which is the time series in this case, f is the function, $Y = \{y_1, y_2, \cdots, y_n\}$ are noisy observation values and ϵ represents Gaussian noise with variance μ_n^2.

The function that measures how one observation relates to another in Gaussian process is a covariance function $k(t_p, t_q)$, where t_p and t_q are time series. Generally, if $t_p \approx t_q$, $k(t_p, t_q)$ will approach its maximum, indicating $f(t_p)$ is almost perfectly correlated with $f(t_q)$, as we expect when the function is continuous and smooth. On the other hand, if t_p is distant from t_q, $k(t_p, t_q)$ will be very small, indicating no linear relationship between $f(t_p)$ and $f(t_q)$.

To estimate f_*, the expected mean value of $f(t_*)$, for an arbitrary t_* in equation 1, three matrices are required: $K(T, T)$ in Equation 2 is used to define the correlation between independent variables; $K(T_*, T)$ in Equation 3 is used to measure the covariance between observation and unknown points; $K(T_*, T_*)$ in Equation 4 is the covariance between unknown points:

$$K(T,T) = \begin{bmatrix} k(t_1,t_1) & k(t_1,t_2) & \cdots & k(t_1,t_n) \\ k(t_2,t_1) & k(t_2,t_2) & \cdots & k(t_2,t_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(t_n,t_1) & k(t_n,t_2) & \cdots & k(t_n,t_n) \end{bmatrix} \tag{2}$$

$$K(T_*,T) = \begin{bmatrix} k(t_*,t_1) & k(t_*,t_2) & \cdots & k(t_*,t_n) \end{bmatrix} \tag{3}$$

$$K(T_*,T_*) = [k(t_*,t_*)] \tag{4}$$

The two key terms for Gaussian process regression can be calculated as:

$$f_* = K(T_*,T)K(T,T)^{-1}[y_1, y_2, \cdots, y_n]^{transpose} \tag{5}$$

$$var(f_*) = K(T_*,T_*) - K(T_*,T)K(T,T)^{-1}K(T_*,T)^{transpose} \tag{6}$$

where $var(f_*)$ is its variance. Using this formula to calculate f_* for whole time series, we can reconstruct $f(x)$ in Equation 1 with its variance function $var(f)$.

2.2. Characterization of Noise in CT Images

As each voxel in CT images is acquired by adding values from a number of different projectors together, no matter the distribution of noise produced at each detector, the final noise in each voxel follows a normal distribution [25]. Therefore, the assumption of Gaussian process regression holds, and the CT perfusion images can be denoised properly using Gaussian process regression.

2.3. Denoising Using Gaussian Process Regression (GPR)

The key rationale for using Gaussian process regression (*GPR*) to denoise CT images relies on the continuity of tissue time-concentration curves (temporal information). The first step is to obtain the time-concentration curve from a fixed voxel, which is considered as a noisy observation, and thus build observed y_i of the points on vector Y at time t_i in equation 1, where t_i is a sampling time in T.

Typically the covariance functions which are used in the denoising algorithm have parameters with the assumed form [24]:

$$k(t_p, t_q) = \mu_f^2 exp(-\frac{1}{2l^2}(t_p - t_q)^2) + \mu_n^2 \delta_{pq} \tag{7}$$

where l is the length-scale, μ_f^2 is termed the signal variance and μ_n^2 is the noise variance (ϵ in Equation 1). δ_{pq} is a Kronecker delta which equals to one if $p = q$ and equals to zero otherwise. More details about parameters and their selection will be given in section 3.2.

According to Equations 2, 3 and 5, for each voxel in the tissue time-concentration curve, the Gaussian process regression filtering equation is defined as follows:

$$f^{'}(t_x) = [k(t_x, t_1)k(t_x, t_2)\cdots k(t_x, t_n)] \begin{bmatrix} k(t_1,t_1) & k(t_1,t_2) & \cdots & k(t_1,t_n) \\ k(t_2,t_1) & k(t_2,t_2) & \cdots & k(t_2,t_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(t_n,t_1) & k(t_n,t_2) & \cdots & k(t_n,t_n) \end{bmatrix}^{-1} \begin{bmatrix} f(t_1) \\ f(t_2) \\ \vdots \\ f(t_n) \end{bmatrix} \tag{8}$$

where t_i is the sampling time of the i_{th} time point, $f(t_i)$ is the raw data and $f^{'}(t_i)$ is the denoised result for the t_i time point. For each voxel, Equation 8 needs to be applied for each time point, that is t_x needs to iterate through t_1 to t_n, to reconstruct the tissue time-concentration curve.

2.4. Denoising Using Multiple Observations Gaussian Process Regression (MGPR)

As a CT brain image has a low contrast-to-noise ratio but a high spatial resolution (usually 512×512 per slice), a spatial decimating noise reduction method called multiple observation Gaussian process regression (MGPR) is also investigated. First of all, images are divided into many small blocks of voxels and each block is called a kernel. After that, in a manner similar to other spatial decimating methods, in MGPR, a new tissue time-concentration curve of the target voxel is calculated from all of the voxels in the same kernel. These adjusted voxels within the same kernel will be treated as multiple noisy observations in Gaussian process regression. Thus the key idea of MGPR is that each sampling time point of any voxel has been observed multiple times instead of once.

For example, if the decimation kernel size is 2×2, the input vector will be built as follows:

$$T = \{t_1, t_1, t_1, t_1, t_2, t_2, t_2, t_2, \cdots, t_n, t_n, t_n, t_n\} \tag{9}$$

$$Y = \{y_1', y_1'', y_1''', y_1'''', y_2', y_2'', y_2''', y_2'''', \cdots, y_n', y_n'', y_n''', y_n''''\} \tag{10}$$

where t_i is still the sampling time and $y_i', y_i'', y_i''', y_i''''$ denote four voxels in the same 2×2 kernel at time point i. The multiple observation Gaussian process regression filtering equation is the same as Equation 8 but with input vectors four times larger than GPR.

The decimation factor can be 1×1, 2×2, 3×3 or even larger. The GPR method mentioned in section 2.3 corresponds to the case when the kernel size is 1×1. The MGPR used in experiments has a decimation factor of 2×2 or 3×3. Furthermore, MGPR uses the same covariance functions and parameters as GPR.

2.5. Patients and Imaging Acquisition

We used data from 10 patients (6 female and 4 male) with a median age of 80 years (maximum 88 years, minimum 47 years, mean 75 years) recruited in a prospective randomised clinical trial. All patients had a radiologically confirmed diagnosis of ischemic stroke. The average time to imaging from onset was 2 hours and 17 minutes (with a maximum of 5.5 and a minimum of 1.5 hours).

2.6. Other Methods In Comparison

2.6.1. TIPS Bilateral Filter Bilateral filter is a smoothing and edge preserving filter. Mendrik *et al.* [22] have recently developed a bilateral filter based method called Time-Intensity Profile Similarity (TIPS) that has been shown to produce higher quality CBF than Gaussian, 3D bilateral or 4D bilateral filter. In their method, they also make use of temporal information. However, TIPS only uses temporal information in the weighting function to determine the similarity function that determines whether two voxels are similar, but it does not utilise the continuity property.

Following the results of Mendrik's research, the standard deviation (μ_d) in their Gaussian closeness function, which determines which distance is considered close, is set to 4; the standard deviation (μ_ζ) in their TIPS function, which determines up to which sum of squared difference measure the time-intensity profiles are still considered similar, is set to 6. The kernel size we used is 5×5.

2.6.2. Mean Filter The spatial mean filter (MEAN), consists of replacing the value of each voxel with the mean value of its neighbours, including itself. It is a simple and intuitive method of smoothing images. Mean filtering can reduce the variation of intensity between adjacent voxels and thus reduce noise in images. In mean filtering, the first step is to group voxels by their coordinates. After that, the average value in a group will be assigned to every voxel within the group. In our experiment, the size of each group is 3×3. Mean filter is a 3D filter and is performed time point by time point.

2.6.3. Mean & GPR Filter Since the results of GPR and MGPR have different resolutions, it is not straightforward to compare these two methods. Therefore, a method called MEAN & GPR is introduced. Essentially, this method is a mean filter followed by a GPR denoising. The mean filter used in this method also has a 3×3 kernel. Mean

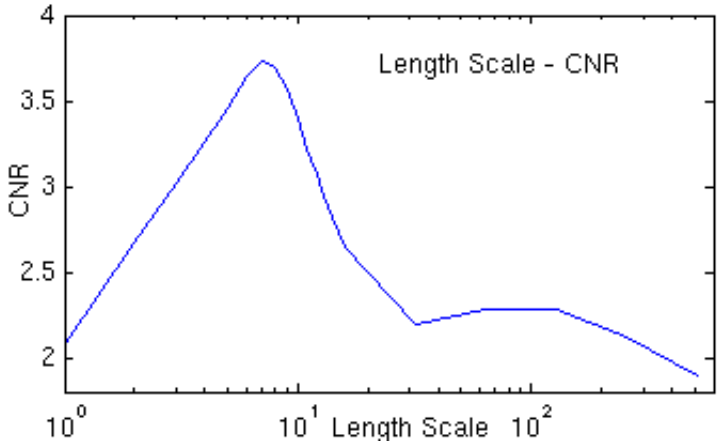


Figure 1: Length Scale Factor

This figure indicates the relationship between length scale l and CNR. Different length scales are applied to a same dataset in order to find out how CNR changes with the length scale.

& GPR can deliver results which have the features of GPR denoising as well as the same resolution as MGPR.

3. Results

3.1. Contrast to Noise Ratio

A quantitative parameter of effective signal to noise is contrast-to-noise ratio (CNR). It is used to measure the noise level relative to the signal change. CNR is defined as one value for the whole image of a scan. It is defined as the reciprocal of the coefficient of variation [26].

$$\text{CNR} = \frac{\mu - e}{\sigma} \tag{11}$$

where μ is the signal mean value, e is the bias and σ is the standard deviation of the noise. Since there is no contrast material injection in the first ten seconds of each scan, the first few seconds of the signal are pure noise. Therefore, the standard deviation of the first ten seconds is considered as the standard deviation of the noise. The bias, e, is set to a baseline value, which is the trimmed mean of the intensity values of the first ten seconds.

Furthermore, only voxels in the non-background area are used to calculate the CNR, thus the background area, with zero intensity values, does not affect the CNR.

3.2. Covariance Function Selection

The covariance function in equation 7 is a squared exponential covariance function. The length scale l, the signal variance μ_f^2 and the noise variance μ_n^2 are required to be set so as to optimize the marginal likelihood. If the length scale is too short, the output will

be very unstable, as it will be overfitted. Conversely, if the length scale is too long, the output will be over smoothed and this will lead to inaccurate results.

The selection of the length scale l is based on Figure 1. Since the peak at $l = 8$ represents the turning point between over smoothing and insufficient smoothing, the parameter l is set to 8 in order to achieve the largest CNR. The signal variance μ_f^2 is optimized to 1.1 and the noise variance μ_n^2 is optimized to 0.5.

3.3. Total Variation

In order to measure the level of oscillation in the data, total variation is introduced. The average total variation, V is determined as follows:

$$V_{overall} = \frac{\sum_{v=1}^{NumberofVoxels} \sum_{t=2}^{time} |f(t) - f(t-1)|}{\text{Number of Voxels}} \tag{12}$$

where $time$ is the number of sampling time points, $f(t)$ represents the intensity value at given time interval.

A variant to Equation 12 is also introduced to measure the oscillation level in the baseline period. As there is no injection of contrast material in the first 10 seconds in each scan, the oscillation level in the baseline period is then calculated from the first ten time intervals and satisfies the following equation:

$$V_{baseline} = \frac{\sum_{v=1}^{NumberofVoxels} \sum_{t=2}^{10} |f(t) - f(t-1)|}{\text{Number of Voxels}} \tag{13}$$

3.4. Standard Deviation

Another method to measure the oscillation level is to use standard deviation. The following equation is used to evaluate the overall smoothness:

$$SD_{overall} = \frac{\sum_{v=1}^{NumberofVoxels} SD(v, time)}{\text{Number of Voxels}} \tag{14}$$

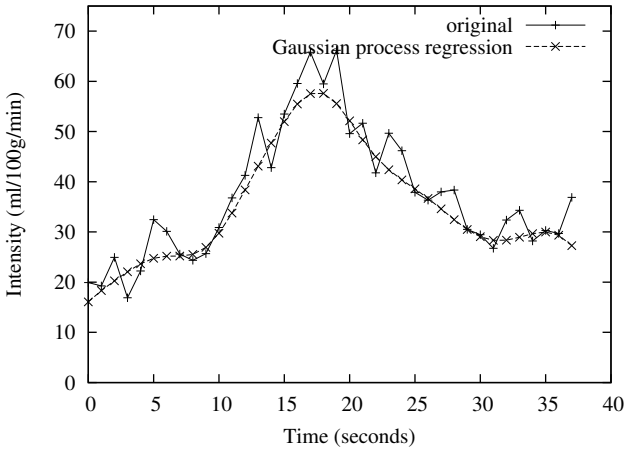
where $SD(v, time)$ is the standard deviation of the intensity values among all of the time points for the v_{th} voxel. Thus $SD_{overall}$ stands for the average standard deviation of all of the voxels (excluding background) in the brain. Another term, $SD_{baseline}$, is defined to quantify the smoothness of the baseline period. $SD_{baseline}$ is similar to $SD_{overall}$ but the SD is calculated from the first ten voxels instead of all of the time points:

$$SD_{baseline} = \frac{\sum_{v=1}^{NumberofVoxels} SD(v, 10)}{\text{Number of Voxels}} \tag{15}$$

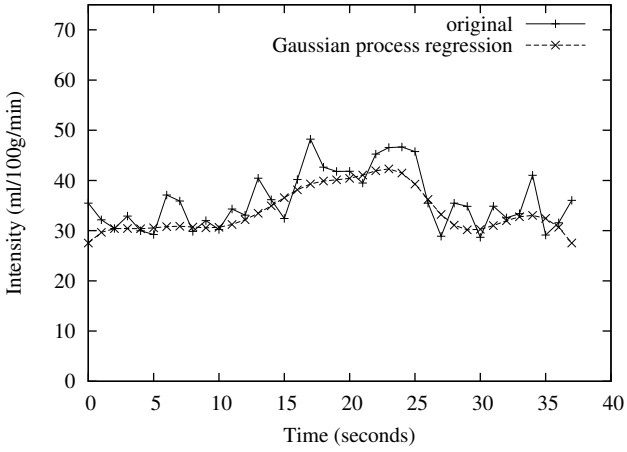
where $SD(v, 10)$ is the standard deviation of the intensity values of the first ten time points for the v_{th} voxel.

3.5. Gaussian Process Regression Denoising Results

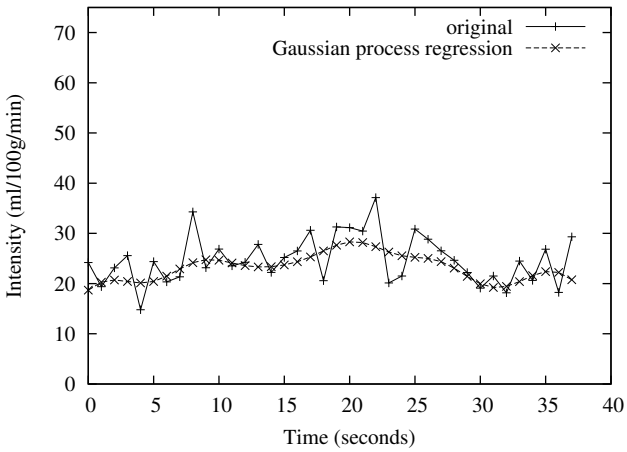
Typical results of GPR noise reduction are shown in Figure 2, and quantitive results are also available in Table 1. From this figure, it can be seen that GPR handles noise well for all of the grey matter, white matter and arteries. In the arterial time series (Figure



(a) Artery



(b) Grey Matter



(c) White Matter

Figure 2: Results of GPR

The figures show the raw and GPR processed time series of three randomly selected voxels corresponds in an artery, grey matter and white matter respectively. The y-axis represents the intensity value on each sampling time point.

Table 1: Raw Vs. GPR Denoised

	Raw	GPR	GPR / Raw
$V_{overall}$	182.2 (33)	40 (8.8)	22%
$V_{baseline}$	44 (11)	9.6 (3.9)	22%
$SD_{overall}$	8.0 (1.1)	6.5 (1.0)	81%
$SD_{baseline}$	5.0 (1.7)	3.2 (1.3)	64%

Comparative measurements of oscillation levels using the five parameters mentioned in Sections 3.3 and 3.4. The background area with straight zero intensity values is exclude from the calculation. The table represents the average (with the standard deviation in brackets) value among ten subjects.

2a) where the CNR is the highest, the application of GPR can perfectly fit the expected shape of the time-concentration curve. In the grey matter (Figure 2b) and the white matter (Figure 2c) where the CNRs are lower than in the arteries, GPR also produces significant improvement.

More specifically, there are three improvements when looking at the level of individual time series.

First, the oscillations after denoising are much smaller than they are in the raw data. The average (standard deviation) of total variation, calculated using Equation 12, among all the ten subjects before GPR is 182 (33). The total variation falls dramatically to 40 (8.8) with a reduction of 78% after applied GPR noise reduction. $SD_{overall}$ also drops by 19%. However, it is problematic to use $V_{overall}$ and $SD_{overall}$, since there is a increase in signal after the contrast arrival and a drop after it reaches the maximum concentration, so that the expected values are not zero. In raw data $V_{overall}$ is even larger than the maximal expected total variation†; GPR delivers results with $V_{overall}$ far less than raw data.

Second, GPR helps in determining parameters, such as time to peak (TTP)‡, bolus arrival time§ and Tmax (or peak height) from the output curves. As showed in Figures 2b and 2c, in the raw data, there are multiple peaks with similar peak values due to the noise and some of them are far away from the real peak (for example, the one at 8 seconds in the white matter). Even in the artery (Figure 2a)), the oscillations can still lead to a bias in the TTP. GPR ameliorates the situation and makes the determination more consistent.

Third, the baseline calculated from the first few voxel values¶ can be calculated more accurately due to the stable onset of the tissue time-concentration curve. Quantitive measurements of the oscillations during the baseline period are the $V_{baseline}$ and $SD_{baseline}$ obtained in Equations 13 and 15. GPR decreased the $V_{baseline}$ from 44 (11) to 9.6 (3.9). Meanwhile, the $SD_{baseline}$ reduces from 5.0 (1.7) to 3.2 (1.3). As the

† In an artery, the baseline is usually 20-30 and the peak is usually 70-100. The amplitude in an artery is also the largest among all of the tissue types.

‡ The time corresponding to the maximum contrast variation.

§ The time it takes for an injected bolus of contrast material to arrive at a given region of the brain.

¶ Usually the trimmed mean value of the first five, or a few more, seconds.

Table 2: Contrast-to-Noise Ratio

	Raw	TIPS	GPR	MEAN 3	GPR & MEAN 3	MGPR 3×3
Subject 1	1.58	1.78	3.70	1.86	3.14	3.79
Subject 2	1.32	1.43	3.09	1.57	2.83	3.28
Subject 3	2.16	2.53	3.29	2.64	2.61	3.75
Subject 4	1.28	1.34	1.93	1.32	1.80	1.91
Subject 5	1.80	2.07	4.32	2.27	3.99	4.18
Subject 6	2.10	2.39	5.44	2.64	5.12	5.12
Subject 7	2.04	2.30	3.60	2.47	2.85	3.87
Subject 8	1.36	1.53	3.51	1.62	3.70	3.31
Subject 9	1.14	1.20	1.35	1.24	1.15	1.40
Subject 10	1.07	1.15	1.76	1.23	1.39	1.85

Comparison of the CNR of raw and denoised images for all of the ten subjects. The CNR displayed are the average CNR of the whole brain images (black background outside the brain is excluded). Column *Raw* is for raw data; column *TIPS* is for TIPS filter; column *GPR* is for Gaussian process regression; column *MEAN 3* is for $[3 \times 3]$ spatial weighted mean filter; column *GPR & MEAN 3* is to apply Gaussian process regression to $[3 \times 3]$ spatial weighted mean filter denoised data and column *MGPR 3 × 3* is for $[3 \times 3]$ MGPR.

injection of contrast material occurs after ten seconds, the signal in the first ten seconds is expected to be constant. Thus the expected $V_{baseline}$ and $SD_{baseline}$ values are both zero. The decrease by 78% and 36% for $V_{baseline}$ and $SD_{baseline}$ respectively indicates the reduction in oscillations achieved by using GPR.

3.6. Multiple Observations Gaussian Process Regression Denoising

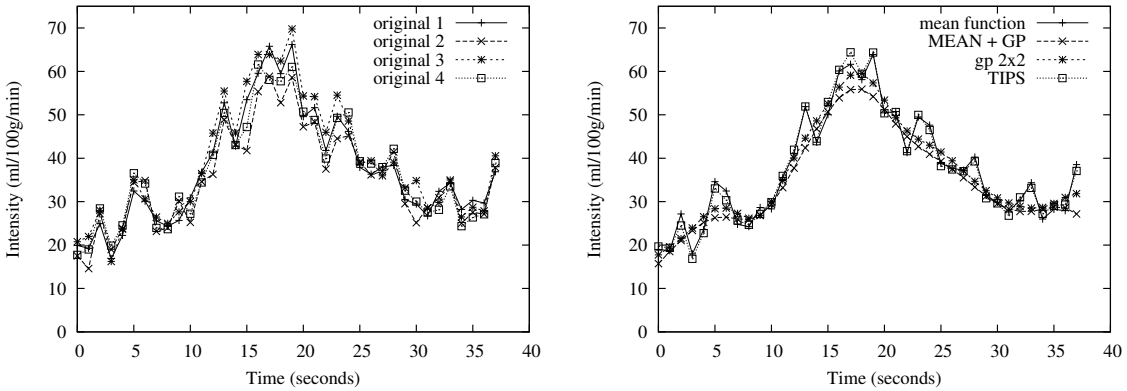
To evaluate the performance of spatial decimation based noise reduction methods, the results of four methods are presented. The first is the weighted mean filter (MEAN). The second is to add an extra GPR step after method one (MEAN & GPR). The third is MGPR as described in section 2.4. The fourth is TIPS filter described in section 2.6.1.

Figure 3 shows how these four methods differ in their denoising efficiency in the context of 2×2 spatial decimation. For all of the three different tissue areas, spatially weighted mean filter does not perform well, as it exhibits oscillations. Both MGPR and Mean & GPR deliver smooth results for the artery data; MGPR is the best for grey matter data, as it delivers a stable baseline, as well as retaining information about the signal peak; in the white matter data, with the lowest CNR, MGPR still delivers a better baseline and peak value compared with the other two methods.

In conclusion based on data from 10 subjects, spatial mean filter does not perform as well as the other methods in this comparison. MGPR is better than MEAN & GPR because MGPR provides a higher CNR and does not lose detailed distribution information in the spatial mean filter step.

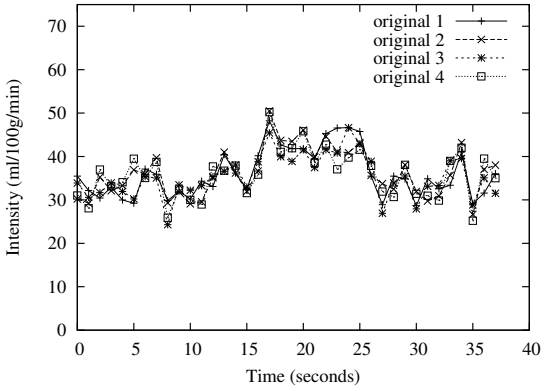
3.7. CNR Improvement

Tables 2 and 3 show that our basic GPR method gives a 99% higher CNR on average than the CNR for raw data. When using spatial decimation, MEAN, to obtain higher CNR, the approach using the spatially weighted mean filter only gives an 18% improvement for

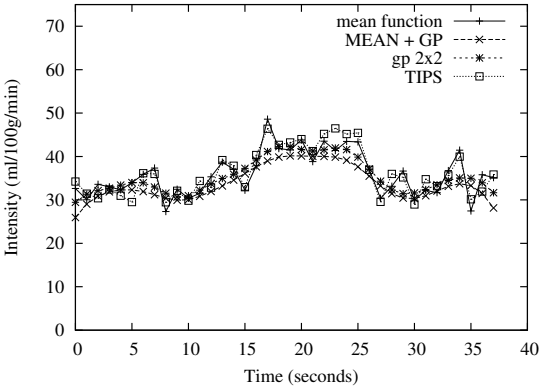


(a) Artery Raw

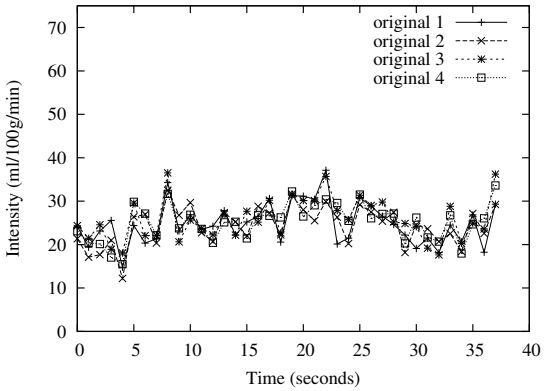
(b) Artery Denoised



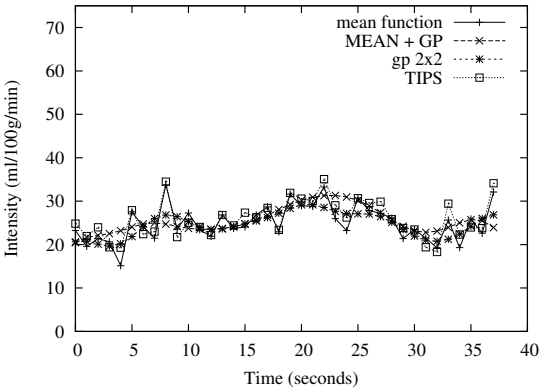
(c) Grey Matter Raw



(d) Grey Matter Denoised



(e) White Matter Raw



(f) White Matter Denoised

Figure 3: Comparison of Four Denoising Methods

The figures on the left show the raw data of the four (2×2) adjacent voxels. The ones on the right show the results processed by different denoising methods: spatially weighted MEAN filter (mean function), weighted mean filter plus regular Gaussian process regression (*MEAN + GP*), *MGPR* (gp 2×2) and TIPS bilateral filter (TIPS). Figures from top to bottom are for tissue types artery, grey matter and white matter. The y-axis represents the intensity value on each sampling time point.

3×3 spatial decimation. Furthermore, the MEAN & GPR method shows its advantages in that it also improves the CNR by about 78%, which is much better than MEAN alone,

Table 3: CNR Improvement

Optimized results	Raw	TIPS	GPR	MEAN 3	GPR & MEAN 3	MGPR 3×3
Mean	1	1.11	1.99	1.18	1.78	2.02
Standard Deviation	-	0.042	0.516	0.072	0.589	0.457

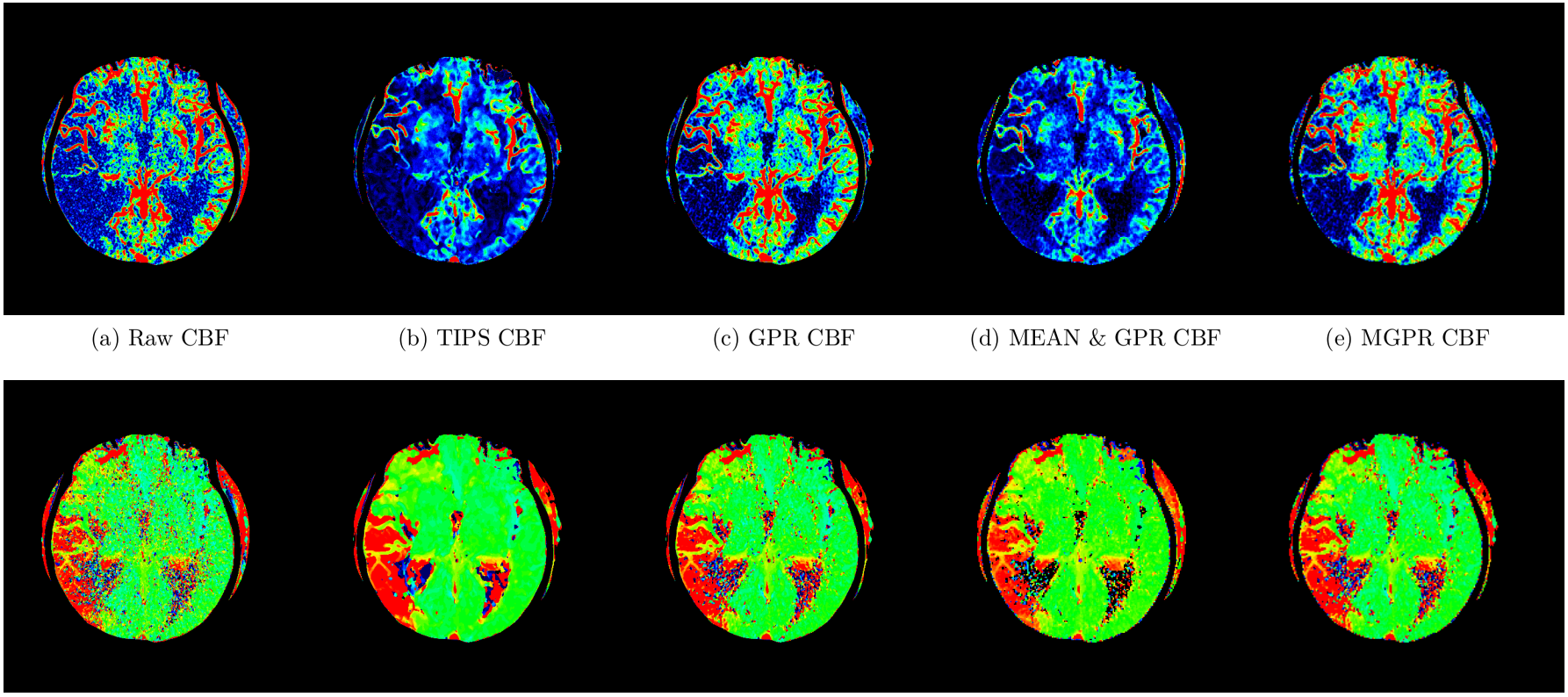
The average improvement of the CNR for the methods in Table 2 relative to the raw image. The CNR of each raw image is set to 1 for normalization.

but smaller than the CNR obtained using the GPR method. The TIPS bilateral filter method only gains 11% CNR improvement, which is much smaller than that obtained by our GPR.

A much better solution is to use the MGPR method which more than doubles the CNR. In our study, MGPR is always the best method, because it delivers the highest CNR among all of the decimation methods that we investigated.

3.8. Qualitative Results

CNR is not the only criterion to measure perfusion image quality. Another very important criterion is the quality of the paramatric maps. Figure 4 shows the result of comparing the quality of CBF and TTP maps when using GPR. The CBF value for each voxel was calculated using a truncated singular value decomposition (SVD) method [27, 28] with a threshold set to 0.15 [29]. In the raw CBF map (Figure 4a), the lesion area at the left bottom may not be clear enough to be distinguished and the image is blurred; the TIPS bilateral filter (Figure 4b) produces less blurred images. GPR denoising gives the best result; the lesion area can be noticed more easily and the image is much clearer (Figure 4c). GPR denoising also demonstrates its advantage in TTP maps; the edges and detailed information can be identified better than in raw data and in TIPS bilateral filter denoised data, and consequently better than Gaussian, 3D bilateral or 4D bilateral filter. For the low flow areas, the white matter and lesion area, GPR provide very good results. Spatial decimation based GPR methods (MEAN & GPR and MGPR) provide relatively good solutions. However, in the full size CBF images in Figure 5, results from spatial decimation based method will become a little blurred as their resolution is reduced by a factor of 9. Compared to TIPS and MEAN & GPR, GPR (Figure 5d) does not reduce the size of vascular features. GPR processed CBF is also less blurred than the the maps produced by MGPR (Figure 5f).

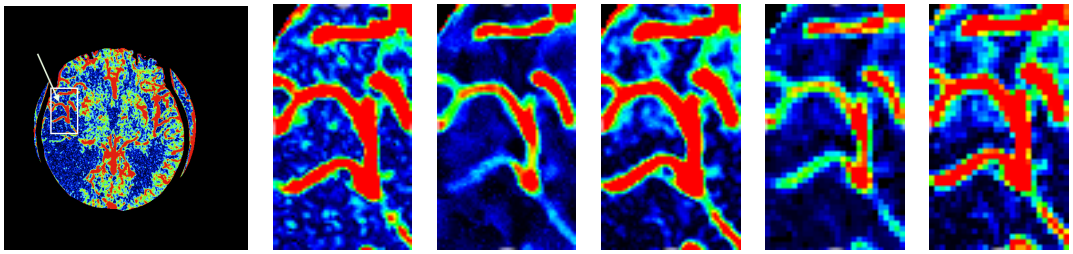


(a) Raw CBF (b) TIPS CBF (c) GPR CBF (d) MEAN & GPR CBF (e) MGPR CBF

(f) Raw TTP (g) TIPS TTP (h) GPR TTP (i) MEAN & GPR TTP (j) MGPR TTP

Figure 4: CBF and TTP Maps

Figures from left to right are raw, *TIPS denoised, GPR denoised, GPR & MEAN 3* and *MGPR* 3×3 denoised maps. The top row is for CBF maps and the bottom row is for TTP maps. Results from the *MEAN 3* method are worse than the raw maps and are not presented here for comparison.



(a) Zoom In Area (b) Raw (c) TIPS (d) GPR (e) MEAN & GPR (f) MGPR

Figure 5: Full Size CBF

This figure contains full size CBF images of the selected rectangle shown in Figure 5a. parts of the full size CBF images. Figures 5b to 5f are a 60×100 area of full sized 512×512 images.

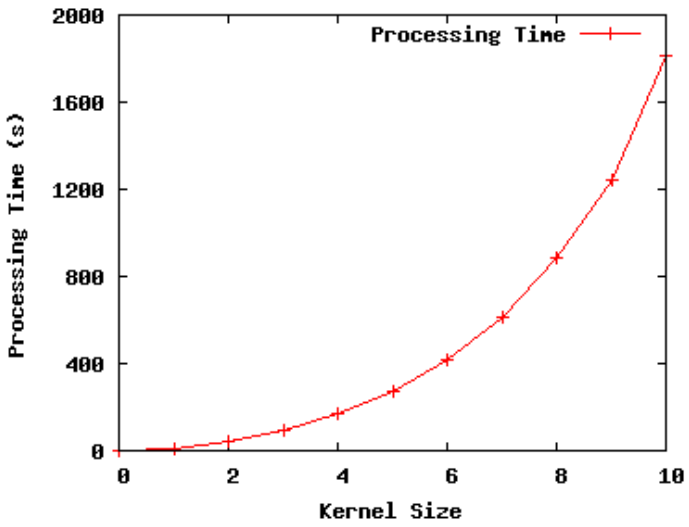


Figure 6: Processing Time

Kernel 1 is an equivalent method to GPR, kernel i ($i \neq 1$) means method MGPR with decimating factor $[i \times i]$.

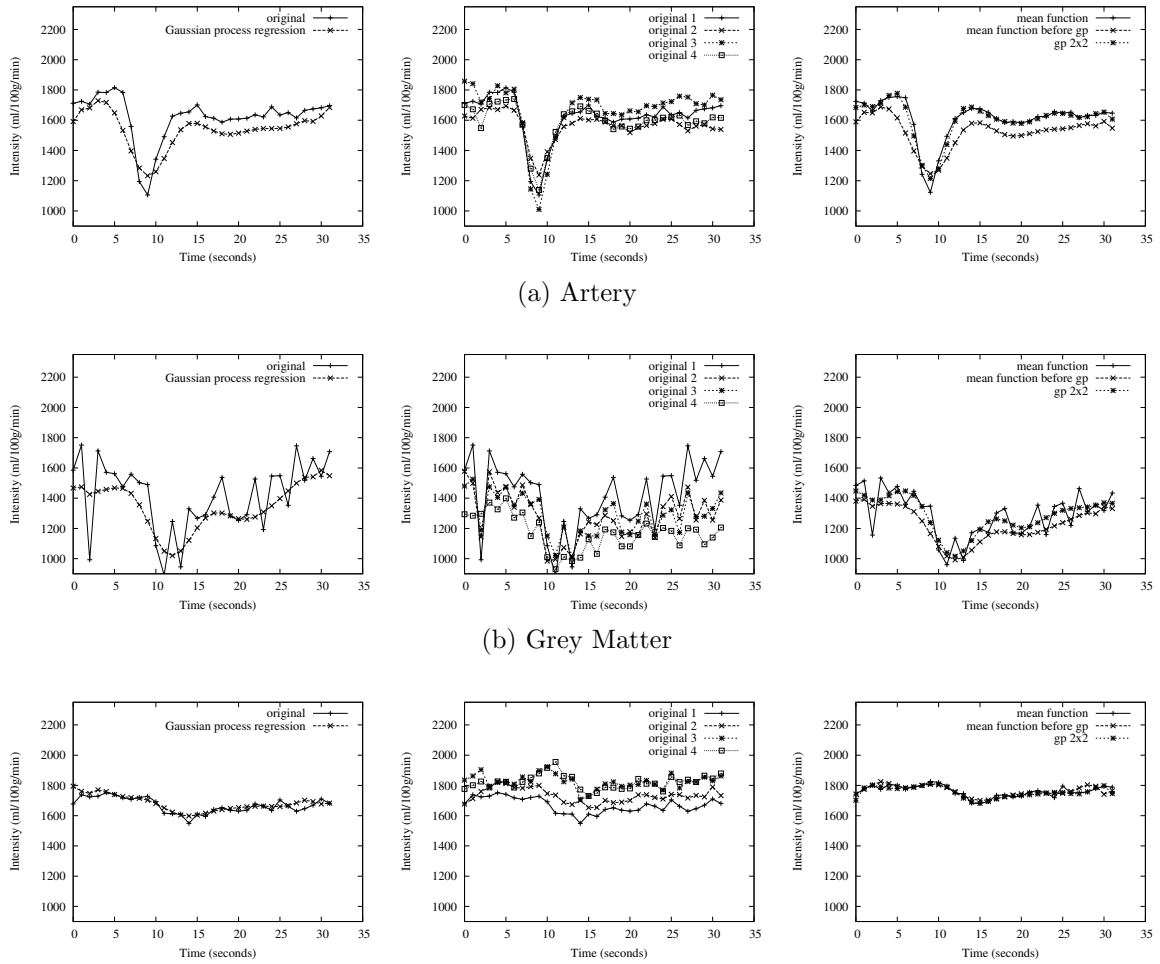
3.9. Processing Time

Figure 6 illustrates the differences in processing time between our Gaussian process regression based methods for the entire volume. Increasing the size of the voxel block will dramatically increase the time needed for processing. GPR only takes 11.8 seconds and it takes more than half an hour if we use $[10 \times 10]$ MGPR. The real processing time fits the expectation well:

$$Time(df) = Time(1) \times df^2 \tag{16}$$

where df is decimation factor described in section 2.4.

In comparison, the processing time for 1×1 MEAN and 1×1 GPR & MEAN is less than 10 seconds and the processing time for TIPS is about 40 seconds.



(a) Artery

(b) Grey Matter

(c) White Matter

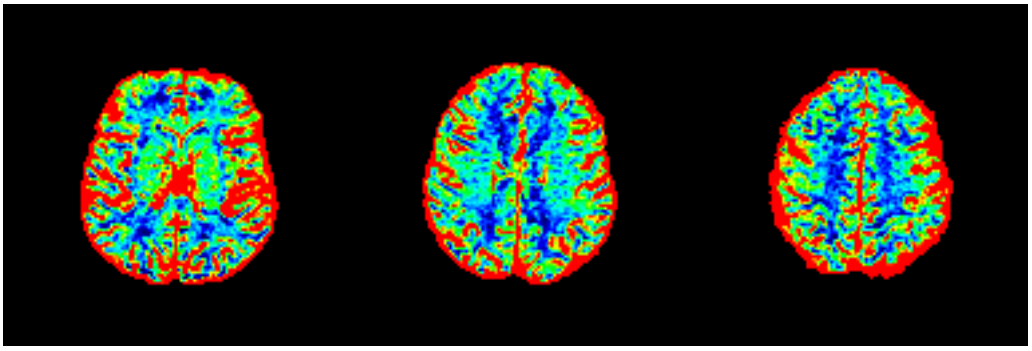
Figure 7: MR Data Denoising Using Gaussian Process

The figures on the left represent the impact of GPR. The figures in the middle are the raw data for spatial decimation denoising of 2×2 kernels in different tissue types. The figures on the right are results using spatially weighted mean filter, GPR & MEAN and MGPR. The y-axis represents the intensity value on each sampling time point.

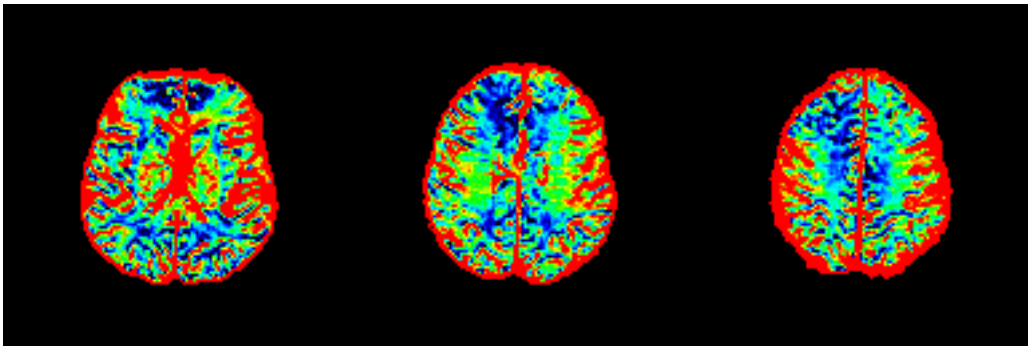
3.10. Magnetic Resonance Imaging Data

Magnetic Resonance imaging (MRI) data has much higher CNR but a lower resolution than CT imaging data. As shown in the left column of Figure 7, using GPR to denoise tissue time-concentration curve gives a much smoother result than the original. Furthermore, all three methods give reasonable baseline, peak value and TTP. In grey matter and white matter, both MEAN & GPR and MGPR methods provide results with less oscillations than the weighted mean filter approach.

Figure 8 illustrates an example of the improvement achieved by GPR in CBF maps. A lesion area can be identified at the anterior cerebral artery (top left hand side of the images) in the CBF maps delivered by GPR denoise which is not easy to spot in the raw CBF. GPR processed CBF also provides more details in white matter and the lesion area.



(a) Raw



(b) GPR Denoised

Figure 8: Comparison of Raw and GPR denoised CBF for MRI
Visualisation of the GPR impact on the CBF for MRI data. Each column is a pair and all three pairs are different slices from a same scan.

4. Conclusions

Results of GPR without spatial decimation show 99% improvement in CNR over raw data. Our spatial decimation based GPR also shows great improvement compared with the spatially weighed mean filter. For 3×3 spatial decimation, our MGPR and MEAN & GPR achieve 102% and 78% higher CNR than raw data respectively, compared to the 18% improvement by simply using a weighted mean filter. Considering the individual voxel, GPR denoised tissue time-concentration curve has much smaller oscillations than raw data, which helps to distinguish parameters, such as baseline value and TTP, more easily and with better accuracy. For haemodynamic parametric maps, our GPR methods provide a much better solution with clearer edges and more detailed information than other approaches. These results show that Gaussian Process Regression based methods handle noise better than other comparable techniques used.

Acknowledgments

The authors would like to thank the Scottish Funding Council (SINAPSE) and the Tony Watson Scholarship for providing the funds and training and the School of Informatics, University of Edinburgh for their laboratory facilities. The patient data came from

Multi-Centre Acute Stroke Imaging Study which was funded by the Translational Medicine Research Collaboration (TMRC).

The work was carried out at the Brain Research Imaging Centre (www.bric.ed.ac.uk), Division of Clinical Neurosciences, University of Edinburgh, a core area of the Wellcome Trust Clinical Research Facility and part of the SINAPSE collaboration funded by the Scottish Funding Council and the Chief Scientist's Office.

This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF) (www.ecdf.ed.ac.uk). The authors wish to express their gratitude to Dr. Jano van Hemert for his guidance, assistance and support. The authors would like to thank both reviewers for their valuable comments and suggestions.

References

[1] M. Wintermark, J.P. Thiran, P. Maeder, P. Schnyder, and R. Meuli. Simultaneous measurement of regional cerebral blood flow by perfusion CT and stable xenon CT: a validation study. *American journal of neuroradiology*, 22(5):905, 2001.

[2] M. Wintermark, M. Sesay, E. Barbier, K. Borbely, W.P. Dillon, J.D. Eastwood, T.C. Glenn, C.B. Grandin, S. Pedraza, J.F. Soustiel, et al. Comparative overview of brain perfusion imaging techniques. *Stroke*, 36(9):e83, 2005.

[3] R.E. Latchaw, H. Yonas, G.J. Hunter, W.T.C. Yuh, T. Ueda, A.G. Sorensen, J.L. Sunshine, J. Biller, L. Wechsler, R. Higashida, et al. Guidelines and recommendations for perfusion imaging in cerebral ischemia. *Stroke*, 34(4):1084, 2003.

[4] M. Wintermark, N.J. Fischbein, W.S. Smith, N.U. Ko, M. Quist, and W.P. Dillon. Accuracy of dynamic perfusion CT with deconvolution in detecting acute hemispheric stroke. *American journal of neuroradiology*, 26(1):104, 2005.

[5] Z. Sajjad. Perfusion imaging in ischaemic stroke. *JPMA. The Journal of the Pakistan Medical Association*, 58(7):391, 2008.

[6] C.S. Rivers, J.M. Wardlaw, P.A. Armitage, M.E. Bastin, T.K. Carpenter, V. Cvoro, P.J. Hand, and M.S. Dennis. Persistent infarct hyperintensity on diffusion-weighted imaging late after stroke indicates heterogeneous, delayed, infarct evolution. *Stroke*, 37(6):1418, 2006.

[7] C.S. Rivers, J.M. Wardlaw, P.A. Armitage, M.E. Bastin, P.J. Hand, and M.S. Dennis. Acute ischemic stroke lesion measurement on diffusion-weighted imaging important considerations in designing acute stroke trials with magnetic resonance imaging. *Journal of Stroke and Cerebrovascular Diseases*, 16(2):64–70, 2007.

[8] T.E. Mayer, G.F. Hamann, J. Baranczyk, B. Rosengarten, E. Klotz, M. Wiesmann, U. Missler, G. Schulte-Altedorneburg, and H.J. Brueckmann. Dynamic CT perfusion imaging of acute stroke. *American journal of neuroradiology*, 21(8):1441, 2000.

[9] A.B. de Gonzalez and S. Darby. Risk of cancer from diagnostic X-rays: estimates for the UK and 14 other countries. *The Lancet*, 363(9406):345–351, 2004.

[10] A.J. Einstein, M.J. Henzlova, and S. Rajagopalan. Estimating risk of cancer associated with radiation exposure from 64-slice computed tomography coronary angiography. *JAMA: the journal of the American Medical Association*, 298(3):317, 2007.

[11] T.A. Bronikowski, C.A. Dawson, and J.H. Linehan. Model-free deconvolution techniques for estimating vascular transport functions. *International journal of bio-medical computing*, 14(5):411–429, 1983.

[12] H.J. Wittsack, AM Wohlschlager, EK Ritzl, R. Kleiser, M. Cohnen, RJ Seitz, and U. Moodder. CT-perfusion imaging of the human brain: advanced deconvolution analysis using circulant singular value decomposition. *Computerized Medical Imaging and Graphics*, 32(1):67–77, 2008.

[13] A.P. Witkin. Scale-space filtering. *Readings in computer vision: issues, problems, principles, and paradigms*, pages 329–332, 1987.

[14] L. Østergaard, D.F. Smith, P. Vestergaard-Poulsen, S.B. Hansen, A.D. Gee, A. Gjedde, and C. Gyldensted. Absolute cerebral blood flow and blood volume measured by magnetic resonance imaging bolus tracking: comparison with positron emission tomography values. *Journal of Cerebral Blood Flow & Metabolism*, 18(4):425–432, 1998.

[15] Acute Stroke Imaging Standardization Group - Japan (ASIST-Japan). Perfusion mismatch analyzer (PMA). software,, Nov 2006.

[16] G Pajares and J Manueldelacruz. A wavelet-based image fusion tutorial. *Pattern Recognition*, 37(9):1855–1872, 2004.

[17] Anja Borsdorf, Rainer Raupach, and Joachim Hornegger. *Wavelet Based Noise Reduction by Identification of Correlations*, volume 4174. Springer Berlin / Heidelberg, 2006.

[18] Anja Borsdorf, Rainer Raupach, and Joachim Hornegger. Separate CT-reconstruction for orientation and position adaptive wavelet denoising. *Informatik aktuell*, pages 232–236, 2007.

[19] Jean-Luc Starck, EJ Candes, and DL Donoho. The curvelet transform for image denoising. *Image Processing, IEEE Transactions on*, 11(6):670–684, 2002.

[20] R. Sivakumar. Denoising of computer tomography images using curvelet transform. *ARPN Journal of Engineering and Applied Sciences*, 2(1):21–26, 2007.

[21] S A Hyder and R Sukanesh. An efficient algorithm for denoising MR and CT images using digital curvelet transform. *Adv Exp Med Biol*, 696:471–480, 2011.

[22] A.M. Mendrik, E. Vonken, B. van Ginneken, H.W. de Jong, A. Riordan, T. van Seeters, E.J. Smit, M.A. Viergever, and M. Prokop. TIPS bilateral noise reduction in 4D CT perfusion scans produces high-quality cerebral blood flow maps. *Physics in Medicine and Biology*, 56:3857, 2011.

[23] C.E. Rasmussen. Gaussian processes in machine learning. *Advanced Lectures on Machine Learning*, pages 63–71, 2004.

[24] C.E. Rasmussen and C.K.I Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[25] B.R. Whiting, P. Massoumzadeh, O.A. Earl, J.A. OSullivan, D.L. Snyder, and J.F. Williamson. Properties of preprocessed sinogram data in x-ray computed tomography. *Medical physics*, 33:3290, 2006.

[26] J.T. Bushberg, J.A. Seibert, E.M. Leidholdt Jr, J.M. Boone, and E.J. Goldschmidt Jr. The essential physics of medical imaging. *Medical Physics*, 30:1936, 2003.

[27] L Ostergaard, RM Weisskoff, DA Chesler, C Gyldensted, and BR Rosen. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. part i: Mathematical approach and statistical analysis. *Magn Reson Med*, 36(5):715–25, 1996.

[28] L Ostergaard, AG Sorensen, KK Kwong, RM Weisskoff, C Gyldensted, and BR Rosen. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. part ii: Experimental comparison and preliminary results. *Magn Reson Med*, 36(5):726–36, 1996.

[29] R. Wirestam, L. Andersson, L. Ostergaard, M. Bolling, JP. Aunola, A. Lindgren, B. Geijer, S. Holtås, and F. Ståhlberg. Assessment of regional cerebral blood flow by dynamic susceptibility contrast MRI using different deconvolution techniques. *Magn Reson Med*, 43(5):691–700, 2000.