HAL
open science

Robust Cerebral Blood Flow Map Estimation in Arterial Spin Labeling

Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot

To cite this version:

Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot. Robust Cerebral Blood Flow Map Estimation in Arterial Spin Labeling. International Workshop on Multimodal Brain Image Analysis (MBIA), held in conjunction with MICCAI 2012, Oct 2012, Nice, France. pp.215-224. inserm-00724974

HAL Id: inserm-00724974

https://inserm.hal.science/inserm-00724974

Submitted on 16 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Robust cerebral blood flow map estimation in Arterial Spin Labeling

Camille Maumet[1,2,3,4], Pierre Maurel[1,2,3,4], Jean-Christophe Ferré[1,2,3,4,5], Christian Barillot[1,2,3,4]

1. University of Rennes 1, Faculty of medecine, F-35043 Rennes, France
2. INSERM, U746, F-35042 Rennes, France
3. CNRS, IRISA, UMR 6074, F-35042 Rennes, France
4. Inria, VISAGES project-team, F-35042 Rennes, France
5. CHU Rennes, Department of Neuroradiology, F-35033 Rennes, France

Abstract. Non-invasive measurement of Cerebral Blood Flow (CBF) is now feasible thanks to the introduction of Arterial Spin Labeling (ASL) Magnetic Resonance Imaging (MRI) techniques. To date, the low signal-to-noise ratio of ASL gives us no option but to repeat the acquisition in order to accumulate enough data to get a reliable signal. Perfusion signal is usually extracted by averaging across the repetitions. However, due to its zero breakdown point, the sample mean is very sensitive to outliers. A single outlier can thus have strong detrimental effects on the sample mean estimate.
In this paper, we propose to estimate robust ASL CBF maps by means of M-estimators to overcome the deleterious effects of outliers. The behavior of this method is compared to z-score thresholding as recommended in [8]. Validation on simulated and real data is provided. Quantitative validation is undertaken by measuring the correlation with the most widespread technique to measure perfusion with MRI: Dynamic Susceptibility weighted Contrast (DSC).

1 Introduction

Arterial Spin Labeling (ASL), a Magnetic Resonance Imaging (MRI) technique introduced in the early 1990s, allows non-invasive quantification of Cerebral Blood Flow (CBF) [2]. Contrary to Dynamic Susceptibility weighted Contrast (DSC), the most widespread technique to measure perfusion with MRI, ASL does not rely on the injection of an exogenous contrast agent.

During the ASL acquisition, blood water, used as an endogenous tracer, is labeled with a radio-frequency pulse in the neck of the patient. After a delay of a few hundred of milliseconds, called inversion time, a labeled image of the brain is acquired. A control image is acquired without prior labeling and the difference between control and label image leads to a perfusion weighted image. A model is then applied to this image to obtain a quantification of CBF. Besides the absence of allergic reaction risk compared to DSC, ASL is particularly well suited for longitudinal studies or studies on patients with difficult venous access

such as children. However this comes at the cost of low signal to noise ratio (SNR) and lower spatial resolution than DSC.

Due to the low SNR of the ASL sequence, a single pair of control and label image is not sufficient to measure perfusion. The acquisition is usually repeated several times, leading to R pairs of images (usually $R \geq 30$). Perfusion information is then usually extracted by pair-wise subtracting the control and label images and averaging across the repetitions.

Though sample average, as an unbiased estimate of mean, ensures convergence as R grows, it has a zero breakdown point and is thus very sensitive to outliers. In particular, sudden subject motion may not be correctly corrected by registration and cause strong corolla-shaped artefacts. To avoid the detrimental effects that a few abnormal repetitions could have in the final perfusion map, it is often suggested to ignore the volumes corresponding to the motion peaks using an appropriate threshold [7]. However the choice of these thresholds is empirical and there is no common rule across studies nor automatic methods to tune these ad-hoc parameters. In [8], the authors proposed an automatic algorithm for outlier rejection in ASL perfusion series based on z-score thresholding at the volume (or slice) level. Their method produced satisfactory results on a qualitative validation based on ratings made by medical experts.

How to appropriately deal with outliers has been largely studied in the statistical literature and a large range of methods has emerged. Z-score is known to be sensitive upon sample size and is suffering from masking effects when more than one outliers is present in the serie [6]. Indeed, in a dataset containing more than one outlier, the standard deviation estimate will be artificially inflated which may prevent z-score based outlier detection. On the other hand, M-estimators are robust techniques to estimate location and scale in the presence of outliers [5]. We focus on Huber's M-estimator [3], as it is the most widely used.

In this paper, we propose to estimate robust ASL CBF maps by means of Huber's M-estimator. This method is compared to z-thresholding as proposed in [8]. Validation is undertaken by measuring the voxel-to-voxel correlation between ASL CBF maps and DSC CBF maps as an affine relationship is expected between these estimates of CBF [9].

Section 2 presents the statistical methods and the validation procedure. Section 3 presents the results on simulated data and on real datasets from patients diagnosed with brain tumors.

2 Material and Methods

2.1 Robust CBF map creation

Starting from a perfusion-weighted serie, namely a 4D volume made of the R repetitions obtained after pair-wise subtracting the control and label scans, the objective is to compute a single perfusion weighted volume. This section presents z-score thresholding and M-estimators as statistical method to compute robust CBF maps.

Z-score thresholding In [8], an outlier rejection algorithm based on z-scores is proposed in order to remove outliers from the perfusion-weighted serie. The outlier rejection is performed both on a volume-by-volume and a slice-by-slice basis. For each volume (respectively slices) v, the mean μ_v and standard deviation σ_v of in-brain voxel intensities is computed. Assuming a gaussian distribution of μ_v and σ_v, a volume is then rejected if:

$$\begin{aligned} &|\mu_v| > \text{avg}_{i=1}^{R}(\mu_i) + 2.5\ \text{avg}_{i=1}^{R}(\mu_i), \\ &\text{or}\ \ \sigma_v > \text{avg}_{i=1}^{R}(\sigma_i) + 1.5\ \text{std}_{i=1}^{R}(\sigma_i) \end{aligned} \tag{1}$$

where avg and std stand for the sample mean and sample standard deviation. The constants 1.5 and 2.5 were determined empirically. To avoid over-filtering, series verifying $\ln(\max_{i=1}^{R}(\sigma_i) - \min_{i=1}^{R}(\sigma_i)) < 1$ are not searched for outliers. Once the outliers are identified, the perfusion map is then computed by averaging the remaining repetitions.

M-estimators Another solution to deal with outliers is to employ robust statistics such as M-estimators, which will not be overly influenced by outliers. In [3], M-estimators are defined, given a function ρ, as solutions $\hat{\theta}$ of:

$$\hat{\theta} = \underset{\theta}{\text{argmin}}\Big(\sum_{i=1}^{n} \rho(x_i, \theta)\Big). \tag{2}$$

If ρ is differentiable, and ψ is its derivative then eq. (2) can be solved by finding the root of:

$$\sum_{i=1}^{n} \psi(x_i, \theta) = 0. \tag{3}$$

The sample average can be seen as an M-estimator with $\rho(x_i, \theta) = (x_i - \theta)^2$ and $\psi(x_i, \theta) = 2(x_i - \theta)$ leading to $\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{N}$.

The M-estimator of location proposed by Huber in [3] is defined by:

$$\psi(x_i, \theta) = \gamma\Big(\frac{x_i - \theta}{\sigma}\Big) \qquad \text{where}\ \ \gamma(x) = \begin{cases} -k, & x < -k \\ x, & -k < x < k \\ k, & x > k \end{cases} \tag{4}$$

k will be set to 1.345 throughout this paper corresponding to 95% efficiency in gaussian data [4]. Likewise, σ is estimated by a robust estimator: the median absolute deviation divided by 0.6745. Huber's M-estimator is applied voxel by voxel on the perfusion weighted serie to obtain the robust perfusion weighted map.

2.2 Validation

Simulated data In order to assess the efficiency of each technique, we generated simulated data with a known quantity of outliers based on two real datasets. Outliers were drawn from a uniform distribution with extrema (-100;100). These values were determined empirically. Indeed, in an uncorrupted perfusion-weighted

map, the values usually range between -10 and +10 and voxel standard deviation can in fact be up to 50. Also, by looking at the values of identified outliers in a real dataset, we found values as big as 300 in absolute value.

As data corruption usually affects multiple voxels per volume [8], outlier simulation was undertaken by corrupting from 0% to 50% percent of the volumes. We will refer lately to these volumes as outlier volumes. Then, 2%, 20% or 50% of the voxels in each outlier volume were replaced by random outliers leading to low, medium and high level of volume corruption respectively. Each simulation was repeated 30 times in order to get estimates of the standard deviation. Simulated data were based on two real datasets as described below.

The first dataset was a perfusion-weighted serie with a large number of repetitions, R=250, from a healthy subject. The perfusion-weighted map obtained by averaging the 250 repetitions was considered as the ground truth. The 60 first volumes of the serie were extracted and used as dataset for robust CBF map estimation. The quality of the maps produced by each method was measured in term of sum of square difference (SSD) with the estimated ground truth.

The second dataset was built on the perfusion-weighted map of one patient diagnosed with a brain tumor. The original ASL CBF map of this subject presented few artefacts identified by visual inspection and a very low level of motion (<0.5mm and <0.2° in all directions). As DSC is currently the reference method to estimate perfusion with MRI, the quality of the maps produced by each method was measured by computing the Pearson linear correlation coefficient with the DSC CBF map. This assumes an affine relationship between CBF maps produced by ASL and DSC [9].

Experiments on real clinical data sets The efficiency of both algorithms was estimated on a dataset of 14 perfusion-weighted maps of patients diagnosed with brain tumors. The quality of the ASL CBF map was assessed by voxel-to-voxel correlation with the DSC CBF map.

2.3 Data

Data : 14 patients diagnosed with brain tumors were involved in this study. Data acquisition was performed on a 3T Siemens Verio MR scanner with a 32-channel head-coil. Patients were scanned in the context of clinical practice. The imaging protocol included a 3D T1-weighted anatomical sequence (TR: 1900ms, TE: 2.27ms, FOV: 256 x 256 x 176mm^3, flip angle: 9°, resolution: 1 x 1 x 1mm^3), a PICORE Q2TIPS sequence with crusher gradients (TR: 3000ms, TE: 18ms, FOV: 192 x 192mm^2, flip angle: 90°, in plane resolution: 3x3mm^2, slice thickness: 7mm, inter-slice gap: 0.7mm, TI: 1700ms, TI_{wd}: 700ms, R = 60), a DSC sequence (GRE EPI, TR: 1500ms, TE: 30ms, FOV: 230 x 230mm^2, flip angle: 90°, in plane resolution: 1.8 x 1.8mm^2, slice thickness: 4mm, inter-slice gap: 1.2mm) and 3D T1-weighted post gadolinium sequence (TR: 1900ms, TE: 2.27ms, flip angle: 9°, FOV: 250 x 250 x 176mm^3, resolution: 1 x 1 x 1mm^3).

1 healthy subject was involved in this study, the imaging protocol included a 3D T1-weighted anatomical sequence (same parameters as above) and a PI-CORE Q2TIPS sequence with crusher gradients (TR: 2500ms, TE: 19ms, flip angle: 90°, in plane resolution: 3x3mm^2, slice thickness: 7mm, inter-slice gap: 0.7mm, TI: 1800ms, TI_{wd}: 700ms, R = 250).

Pre-processing : Image pre-processing was performed using SPM8 (Wellcome Department of Imaging Neuroscience, University College, London) Matlab toolbox. A six-parameter rigid-body registration of the ASL volumes was carried out in order to reduce undesired effects due to subject motion. Coregistration on grey matter map was then performed based on normalised mutual information. The average of unlabeled volumes was used to estimate the geometrical transformation to apply to each volume.

The 60 unlabeled and labeled ASL volumes were pair-wise subtracted in order to obtain a perfusion weighted serie per subject. Robust ASL perfusion-weighted map was then carried out as described in section 2. A standard kinetic model [1] was then applied in order to obtain quantitative ASL CBF maps.

The DSC images were processed using MR manufacturer software by manually choosing an arterial input function to calculate CBF and mean transit time maps. Similarly to ASL, DSC CBF maps were coregistered on grey matter maps.

3 Results

3.1 Validation on simulated data

Dataset with 250 repetitions: Figure 1 presents the simulation study based on a healthy subject data. The performances of sample average, z-score thresholding [8], and Huber's M-estimator are assessed by measuring the SSD of the ASL CBF map with the ground truth estimated by averaging a large number of repetitions.

As described in fig. 1, with a medium or a high level of corruption, z-score thresholding and Huber's M-estimator perform equally and better than averaging until 20% of volumes are corrupted. If more than 20% of the volumes are affected by outliers, then M-estimators provide better estimates than both z-score thresholding and averaging. The robust M-estimator CBF map is closer to the ground truth and less sensitive in an increase in the number of outliers. The same behavior is observed with a low number of corrupted voxels per volume except that the separation point is at 5% of corrupted volumes instead of 20%. The lower performances of z-thresholding when the number of corrupted volumes exceed 20% (or 5% with low corruption) is a consequence of the masking effect which penalize this estimator when several outliers are present in the serie. Moreover, the performance of Huber's M-estimator always depicts a smaller variance than z-thresholding.

Both Huber's M-estimator and z-score thresholding provide better estimate than the sample average. As the level of corruption per volume decreases, the

separation point between Huber's M-estimator and z-score thresholding tends to become lower. This can probably be explained by the fact that the method proposed in [8] is based on a global mean and standard deviation estimate per volume (or slice) and is therefore less suited to detect lowly corrupted volumes.

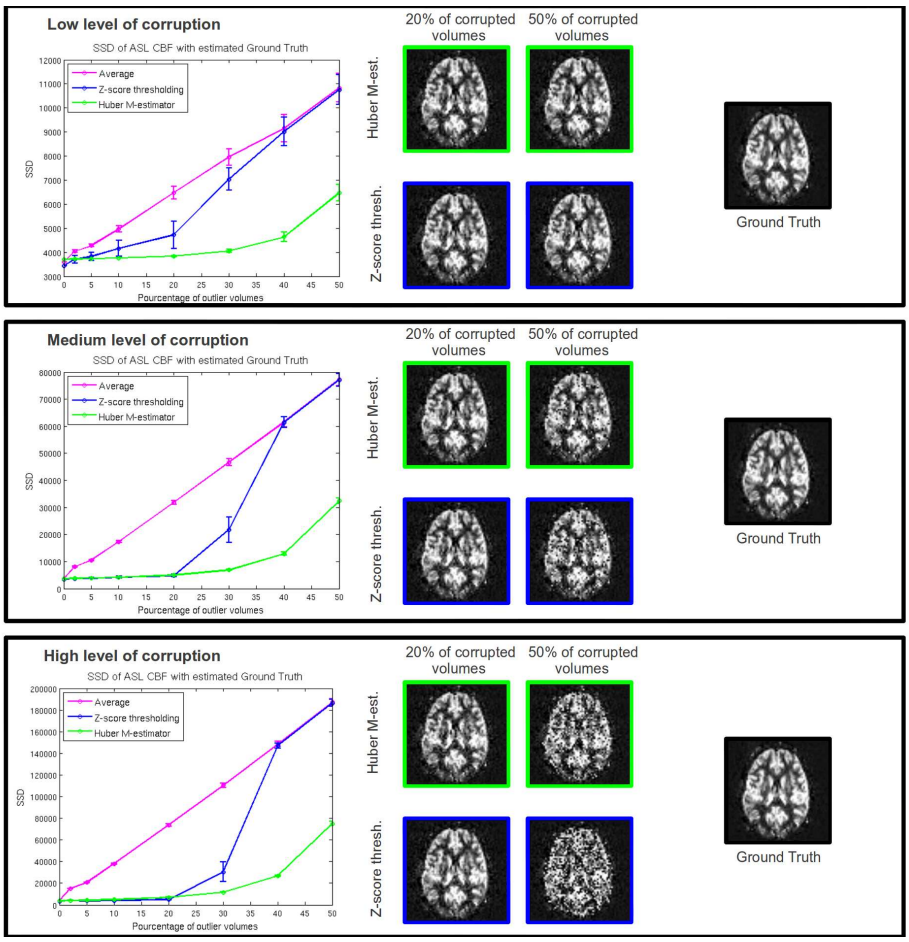

Fig. 1. Healthy subject dataset with simulated outliers: SSD of ASL CBF map, computed by M-estimator, z-score thresholding [8] and sample average, with the estimated ground truth. Low, medium and high level of volume corruption, from 0% to 50% of corrupted volumes. In all configuration Huber's M-estimators is either better or as good as z-thresholding to estimate robust CBF maps. In the presence of outliers, Huber's M-estimator is always more accurate than the sample average.

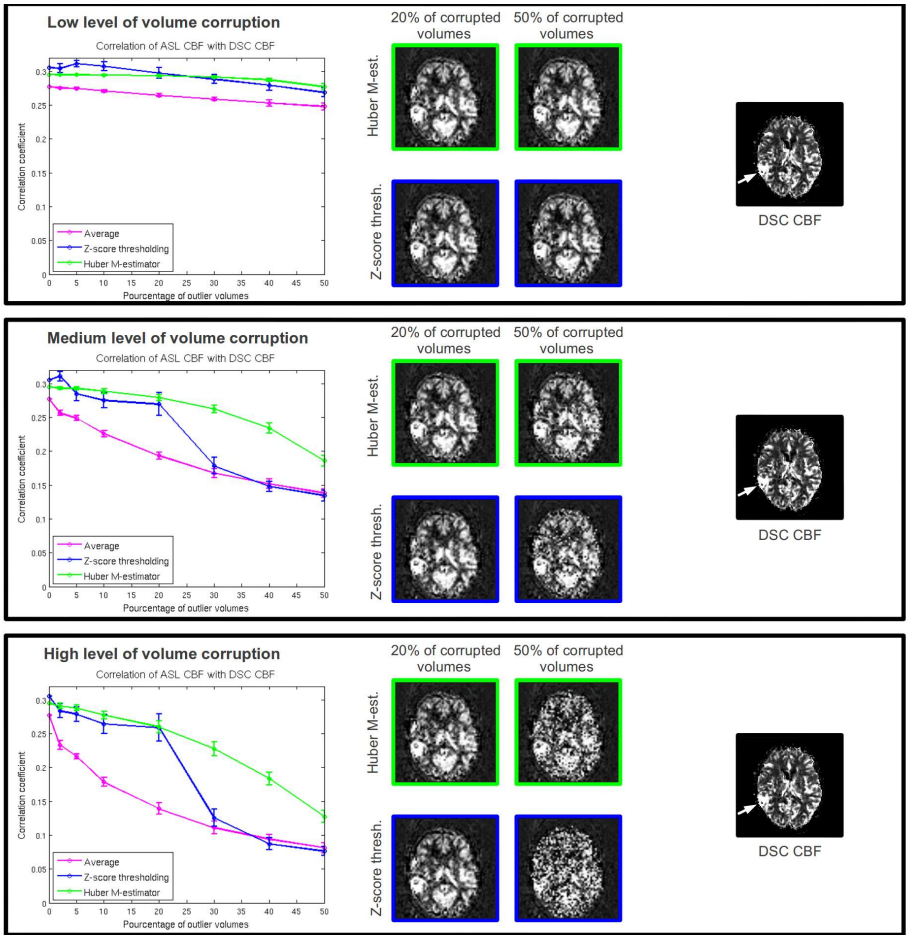

Fig. 2. Patient dataset with simulated outliers: correlation of ASL CBF map, computed by M-estimator, z-score thresholding [8] and sample average, with the DSC CBF map. Low, medium and high level of volume corruption, from 0% to 50% of corrupted volumes. The white arrow points the tumor site. A similar behavior as for healthy subject simulation (fig. 1) is observed outlining that correlation with DSC CBF is a valuable indicator to measure the quality of the ASL CBF estimates.

Simulation based on pathological data: Figure 2 presents the simulation study based on pathological data of a subject suffering from a brain tumor. The performances of sample average, z-score thresholding [8], and Huber's M-estimator are assessed by measuring the correlation coefficient of ASL CBF with DSC CBF.

The simulation involving a high level of volume corruption leads to very similar results than the one obtained in the previous section on healthy subject data. Both Huber's M-estimator and z-thresholding perform better than

averaging until 20% of corrupted volumes. After this threshold, z-thresholding performances drop until reaching the same correlation as the sample average for 30% of outlier volumes. This result suggests that correlation with DSC is a good measure of ASL CBF map quality. For medium level of volume corruption, the same tendency is observable.

With a low level of volume corruption, the trend is less clear. Overall the correlation coefficient seems much less affected by the increasing number of outliers. Z-score thresholding and Huber's M-estimator are both better estimator of the mean than the sample average. Z-score thresholding however displays a higher variance in its performance estimates. In comparison with the previous simulation study, there is probably a higher level of noise in the so-called "uncorrupted" pathological data than in the "uncorrupted" healthy subject data. The inherent higher level of noise in pathological data might prevent the correct detection of low level of volume corruption.

3.2 Validation on real data

Table 1 presents the correlation coefficient obtained for 14 patients diagnosed with brain tumors. Overall, there is a significant improvement of both Huber's M-estimator (p=0.007) and z-score thresholding (p=0.010) over the sample average (paired two sample t-test). In this dataset, there was no significant difference between the two filtering methods (paired t-test p=0.84).

Table 1. Real clinical dataset: correlation coefficient with DSC CBF map of ASL CBF map computed by M-estimator, z-score thresholding [8] and sample average in 14 patients diagnosed with brain tumors. Last column: mean and standard deviation across subjects.

Patients	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Mean ± std.
Huber M-est.	.45	.32	.29	.51	.52	.34	.28	.12	.14	.27	.35	.16	.17	.17	.29 ± .13
z-score thresh.	.45	.24	.27	.53	.51	.35	.28	.15	.14	.30	.35	.18	.16	.20	.29 ± .13
Average	.46	.25	.20	.42	.52	.31	.25	.12	.14	.25	.32	.13	.17	.12	.26 ± .13

Fig. 3 presents an example of robust ASL CBF maps in which motion artefacts are significantly reduced by both Huber's M-estimator and z-thresholding.

4 Conclusion

We studied the ability of Huber's M-estimator to compute robust CBF maps in ASL. The behavior of this estimator was studied in both simulated and real clinical datasets and compared to an outlier removal technique based on z-thresholding previously introduced in the ASL literature [8].

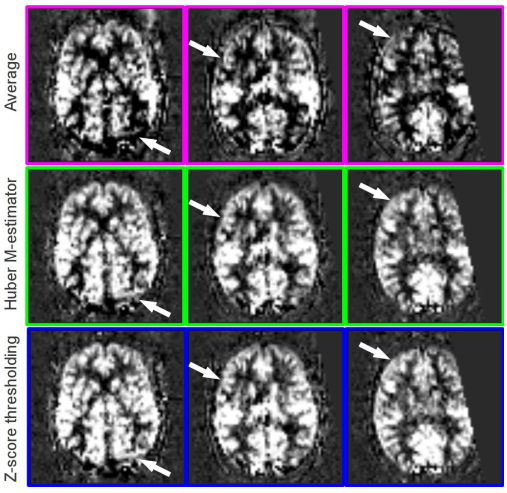

Fig. 3. Example of robust CBF map in one patient: three contiguous axial slices are depicted. White arrows outlines large artefacts presents in the averaged perfusion-weighted map and correctly corrected by both z-score thresholding and M-estimator.

Out of this study, it is confirmed that outlier filtering, either via outlier removal or M-estimation, provides more robust CBF maps than the sample average. Though, on real clinical datasets, both robust methods performed equally, the simulation study clearly stated the superior robustness of M-estimators over z-score thresholding. Overall Huber's M-estimates are either as good as or better than z-thresholding and are always less variable.

As M-estimators are able to deal with a broader range of outliers, we recommend the use of M-estimators as robust method to compute ASL CBF maps. This study focused on patients diagnosed with brain tumors, as DSC sequence is part of their routine clinical protocol. Other pathologies might be related with different outlier patterns and a larger validation study on real datasets is therefore needed in order to outline the cases in which M-estimator will have a significantly better behavior than z-thresholding. Future work will also investigate the effect of other types of M-estimators like Tukey's Biweight.

References

1. Buxton, R.B., Frank, L.R., Wong, E.C., Siewert, B., Warach, S., Edelman, R.R.: A general kinetic model for quantitative perfusion imaging with arterial spin labeling. Magnetic Resonance in Medicine 40(3), 383–96 (1998)
2. Detre, J.A., Leigh, J.S., Williams, D.S., Koretsky, A.P.: Perfusion imaging. Magnetic Resonance in Medicine 23, 37–45 (1992)
3. Huber, P.: Robust estimation of a location parameter. The Annals of Mathematical Statistics (1964)
4. Krasker, W.S., Welsch, R.E.: Efficient bounded-influence regression estimation. Journal of the American Statistical Association 77(379), 595–604 (1982)

5. Rousseeuw, P., LeRoy, A.: Robust Regression and Outlier Detection. Wiley series in probability and mathematical statistics. Probability and mathematical statistics, Wiley-Interscience (2003)
6. Shiffler, R.: Maximum Z scores and outliers. The American Statistician (1988)
7. Sidaros, K., Olofsson, K., Miranda, M.J., Paulson, O.B.: Arterial spin labeling in the presence of severe motion. Journal of Cerebral Blood Flow and Metabolism (2005)
8. Tan, H., Maldjian, J.A., Pollock, J.M., Burdette, J.H., Yang, L.Y., Deibler, A.R., Kraft, R.A.: A fast, effective filtering method for improving clinical pulsed arterial spin labeling MRI. Journal of Magnetic Resonance Imaging 29(5), 1134–9 (2009)
9. Warmuth, C., Günther, M., Zimmer, C.: Quantification of Blood Flow in Brain Tumors: Comparison of Arterial Spin Labeling and Dynamic Susceptibility-weighted Contrast-enhanced MR Imaging. Radiology 228(4), 523–532 (2003)