A SPATIO-TEMPORAL LOW-RANK TOTAL VARIATION APPROACH FOR DENOISING ARTERIAL SPIN LABELING MRI DATA

Ruogu Fang[*] *Junzhou Huang*[†] *Wen-Ming Luh*[‡]

[*] School of Computing and Information Sciences, Florida International University, Miami, FL, USA
[†] Department of Computer Science and Engineering, University of Texas at Arlington, TX, USA
[‡] Cornell MRI Facility, Cornell University, NY, USA

ABSTRACT

Arterial spin labeling MRI (ASL-MRI) can provide quantitative signals correlated to the cerebral blood flow and neural activity. However, the low signal-to-noise ratio in ASL requires repeated acquisitions to improve the signal reliability, leading to prolonged scanning time. At fewer repetitions, noise and corruptions arise due to motion and physiological artifacts, introducing errors into the cerebral blood flow estimation. We propose to recover the ASL-MRI data from the noisy and corrupted observations at shorter scanning time with a spatio-temporal low-rank total variation method. The low-rank approximation uses the similarity of the repetitive scans, and the total variation regularization considers the local spatial consistency. We compare with the state-of-art robust M-estimator for ASL cerebral blood flow map estimation. Validation on simulated and real data demonstrate the robustness of the proposed method at fewer scanning repetitions and with random corruption.

Index Terms— Low-rank, total variation, arterial spin labeling magnetic resonance imaging, cerebral blood flow

1. INTRODUCTION

Arterial spin labeling (ASL) perfusion MRI is a non-invasive technique that quantifies absolute cerebral blood flow (CBF) by magnetically labeling the arterial blood water. The changes in CBF are believed to be directly linked to neural activity [1]. Compared to the more commonly used blood oxygenation level dependent (BOLD) contrast-based technique, which is a complex function of a number of physiological variables [2], ASL has the potential to more accurately reflect the spatial location and magnitude of neural activation [3]. While ASL is gaining importance in brain-behavior relationship studies, there are a number of limitations that prevent its large-scale application. These limitations include the low signal-to-noise ratio (SNR) (typically less than half of BOLD), poorer temporal resolution and fewer number of slices (3-15 slices for ASL compared to 30-40 slices in whole-brain BOLD). Among these limitations, the low signal-to-noise ratio is the most critical issue and requires repeating the measurements numerous times (usually ≥ 30 pairs) to accumulate enough data for a robust estimation. It is due to the fact that the labeled blood volume is small compared to the brain tissues, and the difference between label/control signals is only around 0.5-2.0% of the control image magnitude [4].

Efforts have been devoted to improve the spatial SNR or temporal stability of ASL perfusion signal by using spatial smoothing [5], temporal filtering [6], wavelet denoising [7], noise regression [8], or robust statistics [9]. These post-processing methods reduce the spatial or temporal noise by using the neighborhood information, or rejecting outliers using statistical analysis. Despite of the efficacy of existing post-processing methods, these approaches exploit only local context in the temporal or spatial domains. Global structure of the ASL sequences, especially due to its repetitive label/control (L/C) pairing, can be further exploited to improve the accuracy at reduced scanning time. Moreover, a single statistical model of image noise, e.g. addictive white noise, is usually assumed, which is often violated in practice. Impulse noise and arbitrary corruption due to motion and physiological artifacts can induce detrimental effects in the final perfusion maps.

Based on the two observations above, we propose a novel spatio-temproal low-rank total variation (STLRTV) algorithm to model the structure of ASL sequences by exploiting both the local and global information in the spatio-temporal domain. Low-rank approximation has been increasingly used in medical image denoising, completion and super-resolution [10, 11], yet it is mostly used for spatial approximation. ASL imaging has multiple repetitions of the same structure over time, and therefore leveraging both the temporal low-rank and spatial total variation properties would help. The mixture of noise would be removed using low-rank approximation and local consistency would be preserved via total variation. We validated this proposed method on simulated and in-vivo ASL-MRI datasets. The experimental results demonstrate the effectiveness and robustness of our approach.

This work was partially supported by CTSC Pilot Award, Weill Cornell Seed Grant, NSF IIS-1423056, CMMI-1434401, CNS-1405985.

2. METHODOLOGY

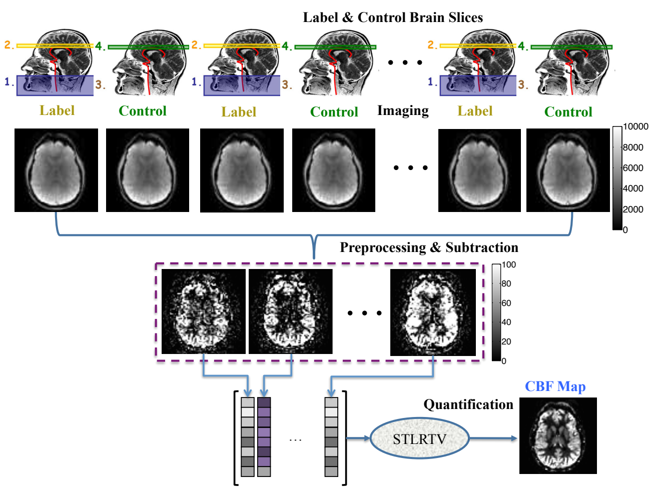

Fig. 1. Overview of our proposed Spatio-Temporal Low-rank Total Variation system.

2.1. Overview of Spatio-Temporal Low-Rank Total Variation Restoration Framework

Fig. 1 shows the proposed framework of Spatio-Temporal Low-rank Total Variation system. We first start with labeling the brain slices using magnetic tagging, and acquire the label and control images. Then we perform image preprocessing including motion removal, co-registration to anatomical image, smoothing, followed by subtraction of the L/C pairs. The difference or perfusion images in the purple dashed box are usually only 1% of the signal strength of the total signal. Although the perfusion images can be directly used for CBF quantification, the accuracy is an issue due to the low signal strength of the delivered blood and physiological artifacts. Therefore, we propose a spatio-temporal low-rank total variation (STLRTV) method to extract the ground-truth perfusion images from the noisy data. The low-rank term models the inherent structure in the temporal repetitions, while the total variation term takes the local information into account. The perfusion images of the whole brain are represented as a 4-D tensor, and the processed tensor using STLRTV is more robust to random corruption and artifacts occurred at fewer repetitions. The final CBF maps are estimated from the processed tensor to improve estimation accuracy and diagnostic decision making.

2.2. Spatio-Temporal Low-Rank Total Variation Formulation

In this section, we introduce the spatio-temporal low-rank total variation method for ASL-MRI data.

The proposed method can be formulated as below:

$$X = \arg\min_X \frac{1}{2}\|Y - X\|_F^2 + \lambda_{\text{tv}}\|X_{(s)}\|_{TV} + \lambda_{\text{rank}}\text{rank}(X_{(t)}) \tag{1}$$

where X is the spatio-temporal ASL tensor to be estimated, Y is the acquired data, $\|X_{(s)}\|_{TV}$ represents the total variation regularization for the spatial dimensions, $\text{Rank}(X_{(t)})$ represents the low-rank regularization for the temporal dimension. λ_{rank} and λ_{TV} are the respective weights for the two terms.

Tensor Total Variation Regularization: The tensor total variation of the spatial 3D data is defined discretely as $\|X\|_{TV} = \sum_{i,j,k}\sqrt{\sum_{d=1}^{3}(\nabla_d X)^2}$, where ∇_d denotes the forward finite difference operator on the d^{th} coordinate. TV regularization is remarkably effective at simultaneously preserving edges while smoothing away noise at homogeneous regions, even at low SNR [12]. Since TV term is also non-smooth, the problem is difficult to solve. Here we use the FISTA [13] algorithm to solve the TV optimization problem.

Temporal Low-Rank Regularization: The rank of the tensor X can be approximated by the trace norm $\|X_{(t)}\|_*$, which is the sum of the singular values of X unfolded in the temporal dimension. For instance, if the 4D tensor X is with size of $N_1 \times N_2 \times N_3 \times T$, the unfolded $X_{(t)}$ is a 2D matrix with size $(N_1 \times N_2 \times N_3) \times T$. The low-rank regularization can be solved by singular value thresholding according to [14] using λ_{rank} as the shrinkage parameter.

2.3. STLRTV Optimization

The cost function in Eq. 1 can be reformulated as with the analysis above:

$$X = \arg\min_X \frac{1}{2}\|Y - X\|_F^2 + \lambda_{\text{tv}}\|X_{(s)}\|_{TV} + \lambda_{\text{rank}}\|X_{(t)}\|_* \tag{2}$$

We use composite splitting technique to solve the problem in Eq. 2 by dividing it into two subproblems with non-smooth terms below. We call the algorithm Composite Splitting Recovery (CSR).

Initialization: We introduce two new variables Z_1 and Z_2 and set their initial values $Z_1^0 = Z_2^0 = Y$.

Subproblem 1: Solve the tensor total variation problem using proximal map [13].

$$\begin{aligned} X_1 &= \arg\min_X \frac{1}{2}\|X - Z_1^{k-1}\|_F^2 + \lambda_{TV}\|X_{(s)}\|_{TV} \\ &= \arg\min_X \text{prox}_{\lambda_{\text{TV}}}(\|X_{(s)}\|_{TV})(Z_1^{k-1}) \end{aligned} \tag{3}$$

where $\text{prox}_\rho(g)(x) := \arg\min_u\{g(u) + \frac{1}{2\rho}\|u - x\|^2\}$.

Subproblem 2: Solve the low-rank problem using singular value thresholding

$$X_2 = \arg\min_X \frac{1}{2}\|X - Z_2^{k-1}\|_F^2 + \lambda_{rank}\|X_{(t)}\|_* \tag{4}$$

as a closed form solution according to [14]:

$$X = \text{fold}_t[SVT_{\lambda_{rank}}(Z_2^{k-1})] \tag{5}$$

Subproblem 3: Update X^k:

$$X^k = (X_1 + X_2)/2 \tag{6}$$

$$Z_1^k = Z_1^{k-1} + X^k - X_1 \tag{7}$$

$$Z_2^k = Z_2^{k-1} + X^k - X_2 \tag{8}$$

Parameters are optimized based on a small set of database. In this work, we set $\lambda_{rank} = 1.0, \lambda_{tv} = 0.5$, and the maximum iteration number is 100. The convergence criterion is when $\|X^k\| - \|X^{k-1}\|/\|Y\| < 10^{-6}$.

3. DATA ACQUISITION AND PREPROCESSING

Pseudo-continuous ASL (pCASL) images were acquired acquired with parameters of FOV=21.6 cm, $72 \times 72 \times 25$ matrix, image acceleration factor=2.5, flip angle=90°, TR=4.5 s, TE=13.2 ms, slice thickness= 5 mm, labeling duration=2.5 s, post label delay time=1.25 s. One hundred and twenty eight label/control image pairs were acquired for each subject. Two subjects are evaluated in our experiments.

All ASL data preprocessing was performed using the SPM8[1] (Statistical Parametric Mapping 8, Wellcome Department of Imaging Neuroscience, University College London, UK) Matlab toolbox and batch scripts from ASL toolbox[2] [15]. The image origins were first set to be the AC-PC line, followed by motion correction customized for ASL MRI. Coregistration on gray matter between ASL and the structure images were performed based on normalized mutual information. A brain mask was generated to remove extracranial voxels by thresholding the mean image with a threshold of 20 percentage of the maximum. The label/control ASL volumes were pair-wise subtracted to obtain a perfusion-weighted series per subject. A general kinetic model [16] was then applied to obtain quantitative ASL CBF maps.

4. EXPERIMENTS

Temporal Low-Rank Approximation: To evaluate the low rank property of the temporal sequences in ASL-MRI images, we select one subject from the ASL data, with size of $72 \times 72 \times 25$ in spatial dimensions and 128 repetitions as the temporal dimension. We then perform singular value decomposition (SVD) on the temporal dimension by stacking the 3D spatial volume at each repetition time as a long vector of a new matrix. As shown in Fig. 2, the eigenvalues decrease exponentially, with many eigenvalues close to zero at the end. We reconstruct the ASL volumes from the largest eigenvalues and compare to the mean volume computed from

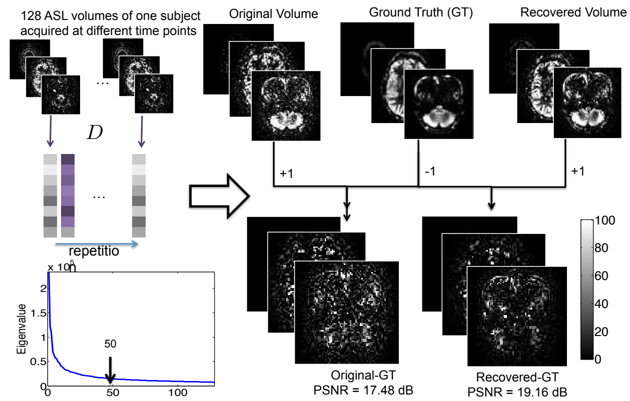

Fig. 2. Low-rank approximation of the ASL data. The left part shows the formation of the large matrix with each repetition sequence as a column and the singular value decomposition. The right part shows the original and recovered volume of one repetition randomly selected from all repetitions. They are compared with the ground truth in the middle.

the average of all repetitions, which is regarded as the ground truth. The most commonly used metric Peak Signal-to-Noise (PSNR) ratio is computed as $\text{PSNR} = 20\log_{10}(\frac{\max\|V_0\|}{\|V_r - V_0\|})$, where V_0 stands for the ground truth volume, and V_r is the recovered volume.

From Fig. 2 we observe that when using the top 50 eigenvalues, the recovered volume has improved PSNR (19.16 dB) compared to the original volume (17.48 dB) relative to the ground truth. When comparing with the original volume, the recovered volume has a high PSNR of 41.56 dB. The result shows that while using only 50 eigenvalues can well approximate the original volume, the low-rank approximation can also remove the random error and noise in the single repetition due to scanning artifacts and white noise, therefore improving the signal fidelity with reference to the ground truth. Our analysis suggests that the ASL data can be, and will benefit from being represented using the temporal low-rank approximations.

Reduced Repetitions: To evaluate the performance of noise removal at fewer repetitions, we use the leave-one-out experimental setting in [17]. Randomly selected N volumes out of the 128 repetitions are used as a noisy input and an average of the remaining $128 - N$ volumes are used as the ground truth, for $N = 1, 2, \ldots, 30$. These datasets are processed with the spatio-temporal low-rank total variation (STLRTV) and Huber's M-estimator [9]. Huber's M-estimator uses robust statistics to deal with the outliers and compute the robust mean. The filtered datasets are averaged over all repetitions and compared with the ground truth, with PSNR computed from the differences as the evaluation metric. Fig. 3 shows the PSNR of M-estimator and STLRTV at difference N.

The results show that while Huber's E-estimator saturates with the increasing N after $N \geq 10$, our proposed method

[1]http://www.fil.ion.ucl.ack.uk/spm
[2]https://www.cfn.upenn.edu/ zewang/ASLtbx.php

continues to improve the performance when the number of repetition increases. STLRTV performs remarkably better than M-estimator, with a margin of at least 5 dB at repetitions fewer than 10, and even larger performance margin of at least 10 dB when the repetitions are more than 10. The overall PSNR of the proposed method is always favorable compared to M-estimator.

Fig. 4 shows the visual result of the average CBF maps using M-estimator and the proposed method. While the average image without processing (noisy) and the outcome of the M-estimator has increased level of random noise and overestimated CBF values due to fewer repetitions, STLRTV method yields CBF map more comparable to the ground truth.

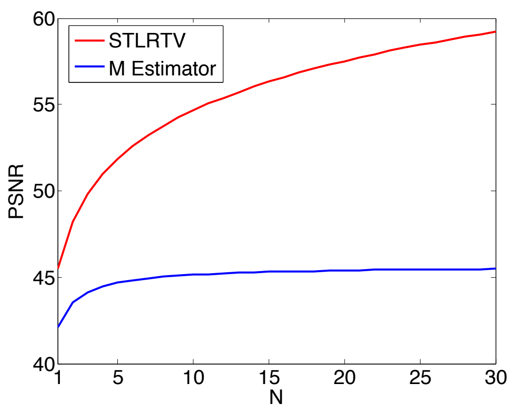

Fig. 3. Peak Signal-to-Noise (PSNR) ratio of the average perfusion image at 1-30 repetitions using the proposed spatio-temporal low-rank total variation (STLRTV) and Huber's M-estimator.

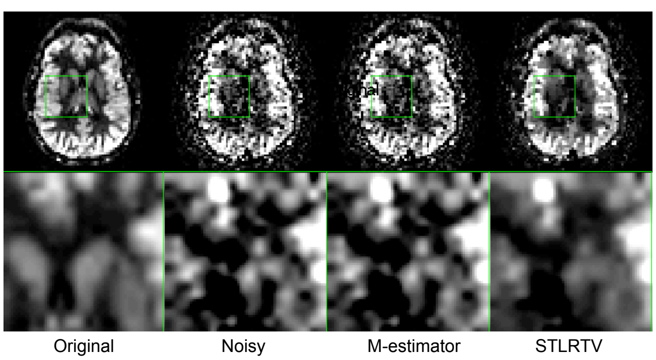

Fig. 4. Visual results for an adult scan in recovery from fewer repetitions with the closed-up view of the selected regions.

Corruption Recovery: In ASL dataset, there are various types of corruptions due to the scanning artifacts, motion, physiological artifacts such as respiratory and pulsation, etc. In order to access the effectiveness of the proposed method in corruption recovery, we remove 50% of the data from the ASL volumes, and recover using the proposed method and the M-estimator, which is designed for robust estimation in case of outliers and corruption. Fig. 5 shows that while the proposed STLRTV method can recover most of the corrupted information, M-estimator fails at the recovery, when the corruption is not in the form of uniform noise, as in [9]. Note that because the 50% missing data is randomly selected from each repetition, the average corrupted data of all repetitions in Fig. 5(b) is totally corrupted. Nevertheless, the proposed STLRTV method could well restore the ASL sequences from the severely corrupted data, while M-estimator which uses robust statistics does not well handle the corruption.

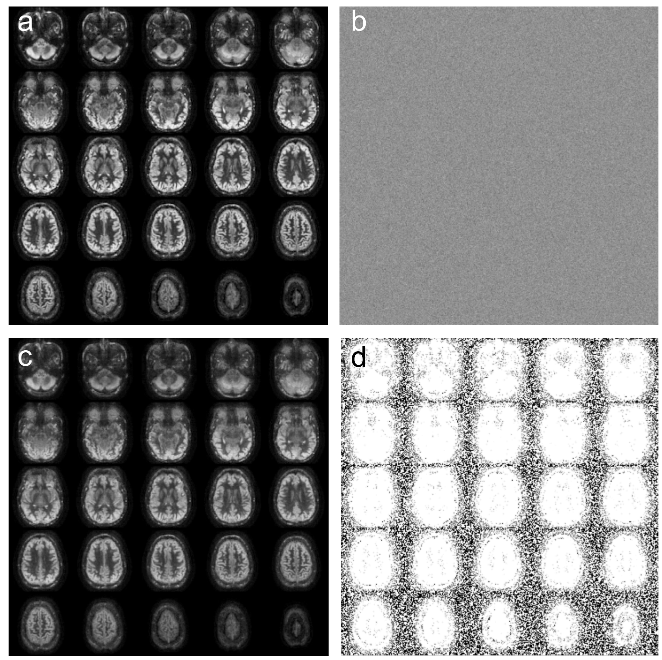

Fig. 5. A healthy subject dataset with corruption of 50% of the volume data. (a) The original average volume. (b) Average volume of the corrupted dataset. Note that *despite only 50% of the data is corrupted in each repetition, the average of all repetitions resulted in corruption distributed over all voxels*. (c) Average volume of the recovered dataset using STLRTV. (d) Average volume recovered using M-estimator.

5. CONCLUSION

In this paper we proposed a spatio-temporal low-rank total variation method to improve the signal quality in arterial spin labeling MRI. For the first time, we show that the joint spatial total variation and temporal low-rank regularization is a viable solution to improve SNR in ASL-MRI. The combination joins the spatial coherence with the temporal similarity for effective signal recovery. Our method outperforms the state-of-art robust estimation method in recovery from reduced repetitions and corrupted data. Further comparison with spatial and temporal filtering methods would be an interesting future work to validate the effectiveness of this joint method.

6. REFERENCES

[1] Charles Smart Roy and CS Sherrington, "On the regulation of the blood-supply of the brain," *The Journal of physiology*, vol. 11, no. 1-2, pp. 85, 1890.

[2] Karla L Miller, Wen-Ming Luh, Thomas T Liu, Antigona Martinez, Takayuki Obata, Eric C Wong, Lawrence R Frank, and Richard B Buxton, "Nonlinear temporal dynamics of the cerebral blood flow response," *Human brain mapping*, vol. 13, no. 1, pp. 1–12, 2001.

[3] Wen-Ming Luh, Eric C Wong, Peter A Bandettini, B Douglas Ward, and James S Hyde, "Comparison of simultaneously measured perfusion and bold signal increases during brain activation with t1-based tissue identification," *Magnetic resonance in medicine*, vol. 44, no. 1, pp. 137–143, 2000.

[4] EC Wong, "Potential and pitfalls of arterial spin labeling based perfusion imaging techniques for MRI," *Functional MRI. Heidelberg: Springer*, pp. 63–69, 1999.

[5] Jiongjiong Wang, Ze Wang, Geoffrey K Aguirre, and John A Detre, "To smooth or not to smooth? ROC analysis of perfusion fMRI data," *Magnetic resonance imaging*, vol. 23, no. 1, pp. 75–81, 2005.

[6] Yashar Behzadi, Khaled Restom, Joy Liau, and Thomas T Liu, "A component based noise correction method (compcor) for bold and perfusion based fMRI," *Neuroimage*, vol. 37, no. 1, pp. 90–101, 2007.

[7] Adnan Bibic, Linda Knutsson, Freddy Ståhlberg, and Ronnie Wirestam, "Denoising of arterial spin labeling data: wavelet-domain filtering compared with gaussian smoothing," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 23, no. 3, pp. 125–137, 2010.

[8] Jeanette A Mumford, Luis Hernandez-Garcia, Gregory R Lee, and Thomas E Nichols, "Estimation efficiency and statistical power in arterial spin labeling fMRI," *Neuroimage*, vol. 33, no. 1, pp. 103–114, 2006.

[9] Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, and Christian Barillot, "Robust estimation of the cerebral blood flow in arterial spin labelling," *Magnetic resonance imaging*, vol. 32, no. 5, pp. 497–504, 2014.

[10] Hien M Nguyen, Xi Peng, Minh N Do, and Zhi-Pei Liang, "Denoising MR spectroscopic imaging data with low-rank approximations," *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 1, pp. 78–89, 2013.

[11] Feng Shi, Jian Cheng, Li Wang, Pew-Thian Yap, and Dinggang Shen, "Low-rank total variation for image super-resolution," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pp. 155–162. Springer, 2013.

[12] David Strong and Tony Chan, "Edge-preserving and scale-dependent properties of total variation regularization," *Inverse problems*, vol. 19, no. 6, pp. S165, 2003.

[13] Amir Beck and Marc Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[14] Jian-Feng Cai, Emmanuel J Candès, and Zuowei Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[15] Ze Wang, Geoffrey K Aguirre, Hengyi Rao, Jiongjiong Wang, María A Fernández-Seara, Anna R Childress, and John A Detre, "Empirical optimization of asl data analysis using an asl data processing toolbox: Asltbx," *Magnetic resonance imaging*, vol. 26, no. 2, pp. 261–269, 2008.

[16] Richard B Buxton, Lawrence R Frank, Eric C Wong, Bettina Siewert, Steven Warach, and Robert R Edelman, "A general kinetic model for quantitative perfusion imaging with arterial spin labeling," *Magnetic Resonance in Medicine*, vol. 40, no. 3, pp. 383–396, 1998.

[17] Jan Petr, Jean-Christophe Ferré, Jean-Yves Gauvrit, and Christian Barillot, "Improving arterial spin labeling data by temporal filtering," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2010, pp. 76233B–76233B.