Joint Reconstruction of Multi-contrast MR Images for Multiple Sclerosis Lesion Segmentation

Pedro A Gómez[1,2,3], Jonathan I Sperl[3], Tim Sprenger[2,3], Claudia Metzler-Baddeley[4], Derek K Jones[4], Philipp Saemann[5], Michael Czisch[5], Marion I Menzel[3], Bjoern H Menze[1]

[1]Computer Science, Technische Universität München, Munich, Germany
[2]Medical Engineering, Technische Universität München, Munich, Germany
[3]GE Global Research, Munich, Germany
[4]CUBRIC, School of Psychology, Cardiff University, Cardiff, Wales, UK
[5]Max Plank Institute of Psychiatry, Munich, Germany
pedro.gomez@tum.de

Abstract. A joint reconstruction framework for multi-contrast MR images is presented and evaluated. The evaluation takes place in function of quality criteria based on reconstruction results and performance in the automatic segmentation of Multiple Sclerosis (MS) lesions. We show that joint reconstruction can effectively recover artificially corrupted images and is robust to noise.

1 Introduction

Multi-contrast MR imaging enables the quantification of metrics that provide information on tissue microstructure. In the domain of neuroimaging, these metrics deepen our understanding of the brain in both health and disease, and could potentially assess the early onset of neurological disorders, such as Multiple Sclerosis (MS). Quantitative metrics are obtained from different MRI techniques, generating multiple contrasts and a wide-range of information regarding tissue microstructure. Obtaining this information, however, comes at the expense of long acquisition times and low signal to noise ratios (SNR).

One possibility for overcoming the limitation of long scan times is through accelerated data acquisitions by compressed sensing (CS). In Diffusion Spectrum Imaging (DSI), acceleration by CS has been successfully demonstrated [1] and is currently being validated in clinical settings. A different approach is to use spatial context to increase data quality without further incrementing acquisition times. One of these methods, presented by Haldar et al. [2], takes advantage of structural correlations between datasets to perform a statistical Joint Reconstruction. This is achieved by incorporating gradient information from all contrasts into the regularization term of a maximum likelihood estimation.

In this study we evaluate the performance of Joint Reconstruction under different noise levels. Furthermore, we investigate the performance of this approach using a metric that evaluates the segmentation accuracy of MS lesions –

i.e., the tasks the images were acquired for – rather than focusing on the common reconstruction error calculated from image intensities.

2 Materials and Methods

2.1 Data acquisition

Five volunteers were scanned with a 3T GE HDx MRI system (GE Medical Systems, Milwaukee, WI) using an eight channel receive only head RF coil. MRI datasets were acquired for a HARDI protocol, a mcDESPOT [3] protocol, and a high resolution T1 weighted anatomical scan (FSPGR). The HARDI protocol consisted of 60 gradient orientations around a concentric sphere with $b = 1200$ s/mm^2 and 6 baseline images at b=0. HARDI datasets were corrected for motion using FSL's FLIRT and FNIRT [4] and both HARDI and mcDESPOT were rigidly registered to the T1 anatomical scan with FLIRT [3].

Seven MS patients were scanned with a CS-DSI acquisition protocol using a GE MR750 scanner (GE Medical Systems, Milwaukee, WI). The CS-DSI protocol comprised of 514 volumes acquired on a Cartesian grid with maximal b-value = 3000 s/mm^2. Additionally, high resolution T1, T2, and FLAIR contrasts were acquired. DSI volumes were co-registered to the first b=0 image, corrected for motion using FLIRT and FNIRT, and a brain mask was obtained using BET [4]. T1, T2 and FLAIR images were down-sampled to the same resolution as the DWIs and all of the volumes were once again co-registered with each other. Finally, for every patient, 11 slices were selected and lesions were manually labelled using a basic region growing algorithm on thresholding FLAIR intensity values.

2.2 Multi-constrast joint reconstruction

In a first experiment we want to evaluate whether Joint Reconstruction can effectively remove noise and maintain data quality in datasets of our multi-contrast sequence. To this end, we studied the reconstruction error under different noise level and optimized the necessary regularization parameters.

After data acquisition and pre-processing, volunteer datasets were artificially corrupted with homogeneous Rician noise and reconstructed using Joint Reconstruction. Then, or a given set of M images, the reconstructed data $\hat{\mathbf{x}}$ was obtained from the corrupted data $\mathbf{y}$ using:

$$\left\{\hat{\mathbf{x}}^1, \hat{\mathbf{x}}^2, \ldots, \hat{\mathbf{x}}^M\right\} = \underset{\{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^M\}}{\arg\min} \sum_{m=1}^{M} \mu_m^2 \left\|\mathbf{F}_m \mathbf{x}^m - \mathbf{y}^m\right\|_2^2 + \Phi\left(\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^M\right), \tag{1}$$

where $\mathbf{F}$ is the Fourier encoding operator, μ is a parameter that adjusts data consistency, and $\Phi(\cdot)$ is a regularization term. As in [2], we define the regularization term as the finite differences over all images. We have to optimize μ and Φ as a function of data quality.

2.3 Lesion segmentation using random forests and multi-contrast image features

In a second experiment we evaluate the performance of a Joint Reconstruction for our sequence using *not* the reconstruction performance of the images, but the DICE scores of an automatic lesion segmentation algorithm. Here, we compare the DICE scores of the ground truth patient datasets with corrupted and jointly reconstructed versions of the datasets.

Random forests have already been implemented to segment MS lesions in multi-contrast MR images, achieving performance comparable to other state of the art segmentation methods [5]. We also propose the use of discriminative classifiers within a random framework to classify voxels, but, given the nature of our patient data, replace context rich features with scalar diffusion features calculated from the CS-DSI protocol.

The feature vector consists of a total of 27 features: three structural MRI intensity channels (T1, T2, and FLAIR), eight diffusion features and 16 kurtosis features. Diffusion features were estimated from the Eigenvalue decomposition of the diffusion tensor $\mathbf{D} \in \mathbb{R}^{3x3}$, while kurtosis features were estimated from projections of the fourth order kurtosis tensor $\mathbf{W} \in \mathbb{R}^{3x3x3x3}$ into spherical and elliptical coordinates. Both tensors were calculated by fitting the data to the diffusional kurtosis model defined in [6] and to a version of the model with a coordinate system rotated into the main directions of diffusion.

The classification task with random forests was accomplished using Matlab's (The Mathworks, Inc) Statistics Toolbox. For this work, a total of 30 trees were grown from four randomly selected datasets and the trained forest was fit to the other three patients. Every tree received a randomly subsampled dataset of voxels and lesion voxels where weighted to proportional to non-lesion voxels.

3 Results

3.1 Experiment 1: Optimization of Joint Reconstruction parameters

In the first experiment we determine regularization parameters of the Joint Reconstruction algorithm that are optimal for our imaging sequence. We use the high resolution volunteer data set.

Volunteer datasets were artificially corrupted with homogeneous Rician noise and reconstructed with different parameter settings. The three regularization parameters, which control for data consistency, regularization, and sensitivity of edge detection, were optimized in function of the remaining noise fraction (RNF) of the reconstructed images, and the root mean square error (RMSE) and structural similarity index (SSIM) [7] of these images to the original raw data.

Table 1 shows exemplary results for a given parameter set with optimized regularization parameters, and Fig. 1 provides a visual comparison of each of the reconstructed contrasts. In this example, Joint Reconstruction was able to remove more than 75% of the artificially added Rician noise, leading to RNF computations between 17.7 and 24.7%.

Table 1. Quality metrics estimated for different jointly reconstructed datasets. Every dataset was individually corrupted with $\sigma = 4\%$ homogeneous Rician noise and jointly reconstructed using Eq. 1.

	Protocol		
Quality criteria	T1	MCDESPOT	DWI
$\sigma_{\hat{\mathbf{x}}}$ [%]	0.992	0.981	0.981
RNF	0.177	0.238	0.247
RMSE	0.090	0.050	0.042
SSIM	0.711	0.683	0.772

3.2 Experiment 2: Evaluation of MS lesion segmentation accuracy

The second experiment evaluated whether Joint Reconstruction can effectively remove noise without losing critical information, such as the borders between lesions and non-lesions. We evaluate the scores on the patient data set.

For five different noise levels, the following was done: homogeneous Rician noise was added to all of the images to corrupt them, images were subsequently reconstructed using Joint Reconstruction, two different kurtosis and diffusion models were fit to the corrupted and reconstructed datasets, and lesion segmentation was performed. The experiment was repeated over 10 iterations and a mean DICE score for every noise level was obtained. Fig. 2 shows the segmentation results of an exemplary dataset and Fig. 3 displays the general performance and robustness to noise.

As seen in Figs. 2 and 3, Joint Reconstruction has a significant impact on segmentation results. At low noise levels, jointly reconstructed datasets yield lower DICE scores than raw data and even noisy datasets. This is most likely do to the fact that Joint Reconstruction has a smoothing effect and that, for certain parameter settings, small edge structures are ignored and blurred out. These small edge structures include the boundary between lesions and non-lesions, especially since this boundary is not completely clear or the same in the multiple

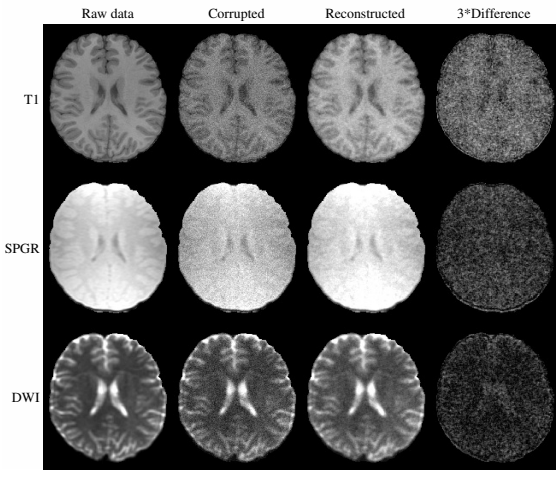

Fig. 1. Reconstructed datasets from a noisy input. Rows show, from top to bottom, three different acquisition protocols: T1, SGPR, and DWI. Columns, from left to right, display: raw data, data corrupted with $\sigma = 4\%$ homogeneous Rician noise, reconstructed data, and absolute difference between the raw data in the first column and the reconstructed data in the third column multiplied times three for better visualization.

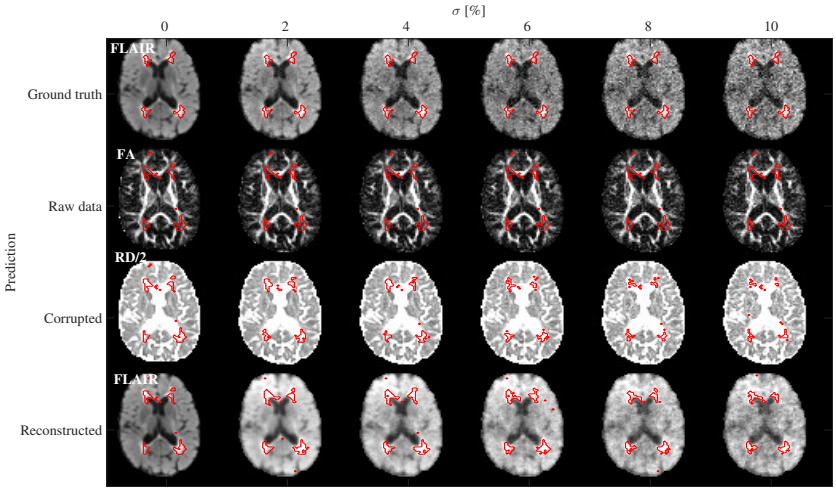

Fig. 2. Segmentation performance with respect to noise. Each row shows a different contrast, indicated in white letters, and the labeled lesions for ground truth (top row) plus predictions on raw data, corrupted data and reconstructed data (bottom three rows). Note that fractional anisotropy (FA) and radial diffusivity (RD) maps weren't directly corrupted, but estimated from corrupted data. RD is shown divided by two for better visualization.

contrasts. As noise levels increase, DICE scores of corrupted datasets decrease while reconstructed datasets maintain similar values.

4 Discussion

In this work, Joint Reconstruction was evaluated for multi-contrast MR images according to multiple criteria and the role of the method on lesion segmentation was further studied. From this analysis, it was established that Joint Reconstruction has a significant impact on lesion segmentation, especially at low noise

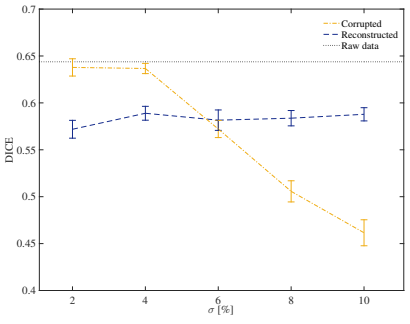

Fig. 3. DICE scores for corrupted and reconstructed datasets as a function of noise levels. Plots show mean ± standard deviation of 10 iterations and the black line indicates the average DICE score obtained from raw data.

levels, where over-smoothing can lead to decreased performance of the segmentation algorithm. On the other hand, Joint Reconstruction proved to be robust to noise, and at higher noise levels, was able to remove noise while still capturing the differences between lesions and non-lesions.

Parameter settings play a crucial role on the Joint Reconstruction framework. Optimizing parameters with respect to the reconstruction errors may not lead to the parameter set that is optimal for lesion segmentation. Furthermore, data quality of each particular dataset also affects the optimal parameter set. Consequently, future work will focus on developing novel, disease-specific and data-adaptive metrics that effectively discriminate between normal state and disease and that can be used to optimize the entire imaging pipeline from data acquisition to analysis.

4.1 Acknowledgements.

This work was funded by the European Commission under Grant Agreement Number 605162.

References

1. Menzel MI, Tan ET, Khare K, Sperl JI, King KF, Tao X, et al. Accelerated diffusion spectrum imaging in the human brain using compressed sensing. Magn Reson Med. 2011;66:1226–33.
2. Haldar JP, Wedeen VJ, Nezamzadeh M, Dai G, Weiner MW, Schuff N, et al. Improved diffusion imaging through SNR-enhancing joint reconstruction. Magn Reson Med. 2013;69(1):277–289.
3. Deoni SCL, Rutt BK, Arun T, Pierpaoli C, Jones DK. Gleaning multicomponent T1 and T2 information from steady-state imaging data. Magn Reson Med. 2008;60:1372–1387.
4. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. Neuroimage. 2012;62:782–790.
5. Geremia E, Clatz O, Menze BH, Konukoglu E, Criminisi A, Ayache N. Spatial decision forests for MS lesion segmentation in multi-channel magnetic resonance images. Neuroimage. 2011;57:378–390.
6. Jensen JH, Helpern JA, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. Magn Reson Med. 2005;53:1432–1440.
7. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: From error visibility to structural similarity. IEEE Trans Image Process. 2004;13:600–612.