Application-Driven MRI: Joint Reconstruction and Segmentation from Undersampled MRI Data

Jose Caballero[1], Wenjia Bai[1], Anthony N. Price[2], Daniel Rueckert[1], and Joseph V. Hajnal[2]

[1] Department of Computing, Imperial College London, UK
[2] Division of Imaging Sciences and Biomedical Engineering Department, King's College London, St. Thomas' Hospital, London, UK
{jose.caballero06,w.bai,d.rueckert}@imperial.ac.uk,
{anthony.price,jo.hajnal}@kcl.ac.uk

Abstract. Medical image segmentation has traditionally been regarded as a separate process from image acquisition and reconstruction, even though its performance directly depends on the quality and characteristics of these first stages of the imaging pipeline. Adopting an integrated acquisition-reconstruction-segmentation process can provide a more efficient and accurate solution. In this paper we propose a joint segmentation and reconstruction algorithm for undersampled magnetic resonance data. Merging a reconstructive patch-based sparse modelling and a discriminative Gaussian mixture modelling can produce images with enhanced edge information ultimately improving their segmentation.

1 Introduction

Magnetic resonance imaging (MRI) produces highly detailed images with excellent soft tissue contrast. In some cases images are not an end in themselves, but rather a means of access to clinically relevant parameters which are obtained as post-processing steps, such as segmentation or tissue characterisation. Acquiring the data necessary to produce images is a time consuming process that can impose significant demands on the patient, but at the same time a lot of information is discarded during post-processing. In cases where the clinically relevant parameters sought are known a priori, the design of image acquisition and reconstruction would ideally be application-driven, such that they are tailored to the information necessary to determine them within some reliability standards.

Compressed sensing (CS) [5, 9] has emerged as an effective way of reducing acquisition time. Incoherently acquiring a fraction of the data normally needed, perfect recovery is possible imposing a sparsity condition on the image, hence allowing scan acceleration without disrupting quantitative analysis. However, it is possible that better measurements could be achieved from fast acquisition data by directly focusing on the final analysis to be performed and treating any reconstruction as an enabling step rather than a distinct endpoint to be achieved first.

P. Golland et al. (Eds.): MICCAI 2014, Part I, LNCS 8673, pp. 106–113, 2014.
© Springer International Publishing Switzerland 2014

This concept of application-driven MRI is broad, given that each imaging design is dependent on the information required from the image. This paper focusses on the case of image segmentation, which is central to many applications of medical image analysis. Segmentation enables the extraction of parameters such as hippocampal volume from brain scans to diagnose or monitor Alzheimer's disease, or ventricular morphology from cardiac cine data revealing cardiac function or ventricular mass.

The problem of segmentation from undersampled measurements has been studied for images that are easily represented in a low dimensional labelled space. Examples are hyperspectral images, which can be described by a few distinct spectral signatures [15], and tomography, where images are well represented as piece-wise constant elements [8, 12]. In the case of MRI it is more challenging to provide a simple description of different regions. An arsenal of segmentation methods is available, based amongst others on region-growing [7], atlas registration [4] or deformable models [10], but the vast majority assume a fully acquired image, despite the growing interest on fast MRI acquisition supported by CS.

We propose a joint optimisation which balances reconstruction fidelity and segmentation performance using a patch-based dictionary sparsity model and a Gaussian mixture model (GMM) on intensity values. Patch-based dictionary sparsity has been proposed as an effective way of exploiting MRI redundancy for reconstruction, notably enabling the use of adaptive dictionaries [2, 11]. The GMM term promotes simpler reconstructions by penalising deviations from the mean values of a few Gaussian distributions. Using 2D brain phantom images and cardiac cine scans, we show that reconstructed images benefit from enhanced edges ultimately improving their segmentation. The paper focusses on a reconstruction-segmentation process but eventually anticipates expanding this to consider optimal sampling strategies for the acquisition.

The paper is organised as follows: Section 2 introduces the necessary background on CS for MRI using patch-based dictionaries and GMM segmentation, and describes the joint reconstruction-segmentation method proposed. Section 3 presents experiments performed on MRI data, visually and quantitatively analysing the benefits of jointly segmenting while reconstructing.

2 Methods

2.1 Patch-Based Compressed Sensing MRI

CS targets perfect data recovery from incomplete samples. Provided incoherent sampling, perfect reconstruction can in theory be guaranteed with the assumption that the data is redundant when represented in a sparsity transform domain [5]. For MRI, transforms such as total variation, wavelets or temporal Fourier transforms apply depending on the imaging modality [9]. Recently, the use of redundant patch-based dictionaries in the image domain has been proposed to sparsely represent images [11] and dynamic data [2,3], and have gained particular attention for their additional ability of becoming adaptive through dictionary learning (DL) algorithms [6, 13].

Let $\mathbf{y}$ be an undersampled acquisition from a scan of size N. Given a dictionary $\mathbf{D}$ with N_a atoms of size N_p, a CS reconstruction $\mathbf{x}$ is given by

$$\min_{\mathbf{x},\mathbf{\Gamma}} \ \|\mathbf{F}_u\mathbf{x}-\mathbf{y}\|^2+\frac{\lambda}{N_p}\sum_{n=1}^{N}\|\mathbf{R}_n\mathbf{x}-\mathbf{D}\boldsymbol{\gamma}_n\|^2 \ \text{ s.t. } \ \|\boldsymbol{\gamma}_n\|_0\leq T \ \ \forall n, \tag{1}$$

where $\mathbf{F}_u$ is an undersampled Fourier transform, $\mathbf{R}_n$ is a patch extraction operator, λ is a weighting parameter, T is a sparsity threshold, and $\boldsymbol{\gamma}_n$ is the sparse representation of patch $\mathbf{R}_n\mathbf{x}$ organised as column n of matrix $\mathbf{\Gamma}$. The dictionary can be a non-adaptive sparsifying dictionary (e.g. discrete cosine transform (DCT) or wavelets), or can be included as a free variable in the problem letting it become adaptive [2,3,11]. Solving this problem directly is difficult as it is non-convex in both variables, but they can be individually updated as a least-squares and orthogonal matching pursuit (OMP) problem iteratively [3,11].

Although CS reconstructions have demonstrated great potential, in practice images can only be assumed to be approximately sparse. Problems are typically encountered at high acceleration rates in regions of the image that do not adhere to the sparse model, generally resulting in lost fine details and smoothed edges [14]. These reconstruction errors are prone to propagate into post-processing stages, compromising the accuracy of quantitative measurements.

2.2 Gaussian Mixture Model Segmentation

Image segmentation based on a GMM assumes image features that are randomly generated from an underlying mixture of Gaussian distributions [1]. In some cases, similar pixel intensities correspond to the same tissue class, and can therefore be used to label data. This typically applies for brain scans, where grey matter, white matter and cerebrospinal fluid show different intensities, or cardiac cine data, where blood pool regions have a characteristic high intensity.

Assuming $\mathbf{x}$ to be MRI data, each pixel x_n can be expressed by a mixture of K Gaussians as $\mathrm{P}(x_n)=\sum_{k=1}^{K}\pi_k\mathcal{N}(x_n|\mu_k,\sigma_k)$, denoting $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$ and $\boldsymbol{\pi}$ as the means, standard deviations and mixture weightings of all Gaussians. The GMM is formally found by maximising the log-likelihood of the posterior distribution

$$\max_{\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\pi}} \ \ln \mathrm{P}(\mathbf{x}|\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\pi}) = \max_{\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\pi}} \ \sum_{n=1}^{N}\ln\left(\sum_{k=1}^{K}\pi_k\mathcal{N}(x_n|\mu_k,\sigma_k)\right). \tag{2}$$

The solution of this maximisation is complicated by the sum inside the logarithm, but the Expectation-Maximisation (EM) [1] algorithm can iteratively find it by alternating the update of the model $(\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\pi})$ and a latent variable $r_{nk}\equiv\mathrm{P}(z_k=1|x_n)$. This term provides a labelling of pixels with an associated uncertainty as the probability that Gaussian k generated data point x_n.

2.3 Segmentation-Driven MRI (SegMRI)

The following task formulates reconstruction imposing a model on the data that weights the deviation from a patch-based sparse and a GMM representation:

$$\min_{\substack{\mathbf{x},\boldsymbol{\Gamma}\\ \boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\pi}}} \|\mathbf{F}_u\mathbf{x}-\mathbf{y}\|^2+\frac{\lambda}{N_p}\sum_{n=1}^{N}\|\mathbf{R}_n\mathbf{x}-\mathbf{D}\boldsymbol{\gamma}_n\|^2-\beta\ln \mathrm{P}(\mathbf{x}|\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\pi}) \text{ s.t. } \|\boldsymbol{\gamma}_n\|_0\leq T \ \ \forall n. \tag{3}$$

If the new tuning parameter is set to $\beta = 0$, the problem is the same as in (1) and is purely reconstructive as the GMM model is not used. For $\beta > 0$, the reconstructed intensities need to be consistent with the GMM emerging as a by-product of the reconstruction. The larger the penalisation of this term is, the closer intensities are to the means of Gaussians, so effectively it homogenises intensities within a region and reduces the number of intensities sitting around the boundaries of two regions. At the extreme $\beta \to \infty$, each pixel intensity is set to the mean of the Gaussian that is the most likely to have generated it.

Problem (3) is also non-convex, but can be solved alternating a conjugate gradient (CG) update of image $\mathbf{x}$, OMP coding of the sparse modelling $\boldsymbol{\Gamma}$, and the EM update of the GMM parameters. The method is summarised in algorithm 1. At each iteration the uncertainty in segmentation is reduced as r_{nk} tends to an indicator function for each n pulling intensities closer to their attributed Gaussian mean. The EM step can eventually become unstable ($\sigma_k \to 0$) as a result, so a minimum variance that the Gaussians in the GMM can achieve is set to prevent this. Global convergence is not proven, but small K and constrained variances reduce the chance of EM local minima, and patch-based MRI reconstruction is empirically robust [2, 3, 11]. The computational bottleneck is OMP ($\sim 2NT(N_p+N_a)$ operations/iteration), although patch coding is parallelisable.

Algorithm 1. SegMRI reconstruction

Input: Acquisition $\mathbf{y}$, number of Gaussians K, sparse model parameters (T, N_p, N_a).
Output: Reconstructed k-space $\mathbf{x}$, GMM parameters ($\boldsymbol{\mu}$, $\boldsymbol{\sigma}$, $\boldsymbol{\pi}$).
Initialise: $\mathbf{x}_0 = \mathbf{F}_u^{\mathrm{H}}\mathbf{y}, \boldsymbol{\Gamma}_0 = \mathbf{0}, t = 0, \sigma_{0,k} = 0.5\ \forall k$, $\boldsymbol{\mu}_0$ randomly chosen from $\mathbf{x}$.
repeat
- **1.** $t \leftarrow t+1$
- **2.** Update sparse coding with OMP using $\mathbf{x}_{(t-1)}$:
  $\boldsymbol{\Gamma}_t \leftarrow \arg\min_{\boldsymbol{\Gamma}} \sum_{n=1}^{N}\|\mathbf{R}_n\mathbf{x}_{(t-1)}-\mathbf{D}\boldsymbol{\gamma}_n\|^2$ s.t. $\|\boldsymbol{\gamma}_n\|_0\leq T$
- **3.** Update GMM parameters with EM using $\mathbf{x}_{(t-1)}$:
  $\{\boldsymbol{\mu}_t, \boldsymbol{\sigma}_t, \boldsymbol{\pi}_t\} \leftarrow \arg\max_{\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\pi}} \ln \mathrm{P}\left(\mathbf{x}_{(t-1)}|\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\pi}\right)$
- **4.** Update reconstruction given the GMM and sparse coding:
  $\mathbf{x}_t \leftarrow \arg\min_{\mathbf{x}} \|\mathbf{F}_u\mathbf{x}-\mathbf{y}\|^2+\frac{\lambda}{N_p}\sum_{n=1}^{N}\|\mathbf{R}_n\mathbf{x}-\mathbf{D}\boldsymbol{\gamma}_n\|^2-\beta\ln \mathrm{P}\left(\mathbf{x}|\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\pi}\right)$

until $\frac{\|\mathbf{x}_{t-1}-\mathbf{x}_t\|^2}{\|\mathbf{x}_t\|^2} \leq 10^{-4}$;

3 Experiments and Results

In order to assess the benefits of jointly segmenting while reconstructing images, accelerated data is reconstructed with the SegMRI method in (3) and the CS method in (1). The latter reconstruction is segmented with a GMM as a separate step and both results are compared with a GMM segmentation of the fully

sampled data. All tests use real-valued images, but the extension to complex-valued data is possible with complex-valued dictionaries [2, 11] and GMMs [1]. Parameters are fixed throughout using DCT dictionaries with $N_p = 8^2$ and $N_p = 4^3$ for 2D and cine tests respectively, $N_a = 196$, $T = 5$, $\lambda = 10^{-3}$ and $\beta = 10^{-9}$. The minimum standard deviation to prevent instability in EM is set to $\sigma_k > 10 \times \sigma_{min}$, with σ_{min} the lowest standard deviation from the fully sampled image GMM, which can be known from prior scans. Experiments were performed on Matlab, Intel Core i72600 CPU, 3.4 GHz, 8 GB RAM.

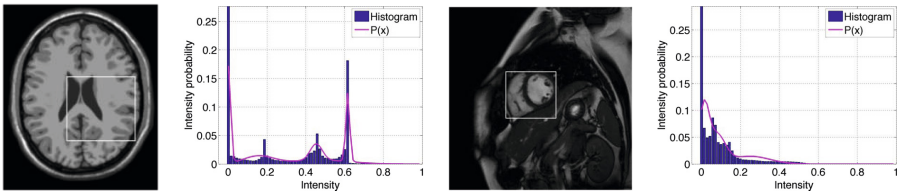


Fig. 1. Test data examples. From left to right: 2D phantom image, its histogram and GMM, one frame from a cardiac cine scan, and its histogram and GMM.

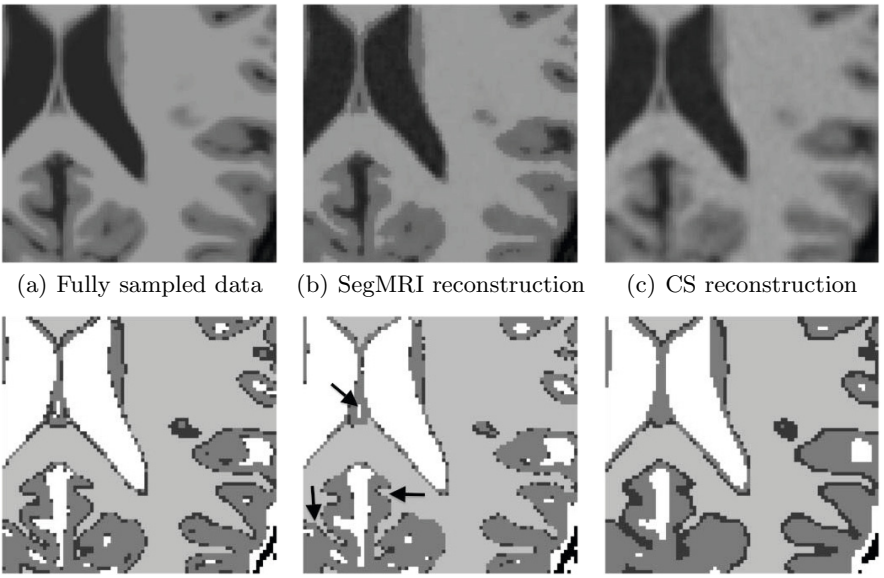
3.1 2D Brain Phantom Experiments

In this experiment the method is evaluated for 2D image reconstruction using sagittal, axial and coronal slices from a brain phantom[1]. A test image is shown in figure 1 alongside its histogram. Notice that the histogram has 4 distinct peaks and that it is well fitted by a 5 component GMM (including one for outliers).

Consider a 5-fold accelerated acquisition of this image. Performing a patch-based CS reconstruction recovers coarse structure but misses fine details and edges as shown in figure 2. When this reconstruction undergoes GMM segmentation, these errors propagate resulting in significantly more partial volume effects compared to the the fully sampled image segmentation. The result from SegMRI is more accurate, for instance at locations highlighted by arrows in figure 2(e). A stable result was found after 30 iterations of 2 seconds each.

Table 1 presents the mean and standard deviation of total misclassified pixels using 18 test images (6 from each plane) at several acceleration factors. As expected, the segmentation accuracy tends to decrease in both cases with increasing acceleration. However, notice that the SegMRI result outperforms the segmentation from the CS reconstruction, which reaffirms the visual inspection of figure 2. We use these percentages in a paired t-test and display the mean difference between CS+GMM and SegMRI errors (μ), t-values (t) and p-values (p). The very low p-value of the full set reveals high statistical significance.

[1] Phantom from BrainWeb (`http://brainweb.bic.mni.mcgill.ca/brainweb/`)



(a) Fully sampled data (b) SegMRI reconstruction (c) CS reconstruction

(d) Fully sampled GMM (e) SegMRI segmentation (f) CS+GMM

Fig. 2. Brain phantom reconstruction and segmentation.

Table 1. Pixel misclassification (mean±std%) and t-test for brain phantom

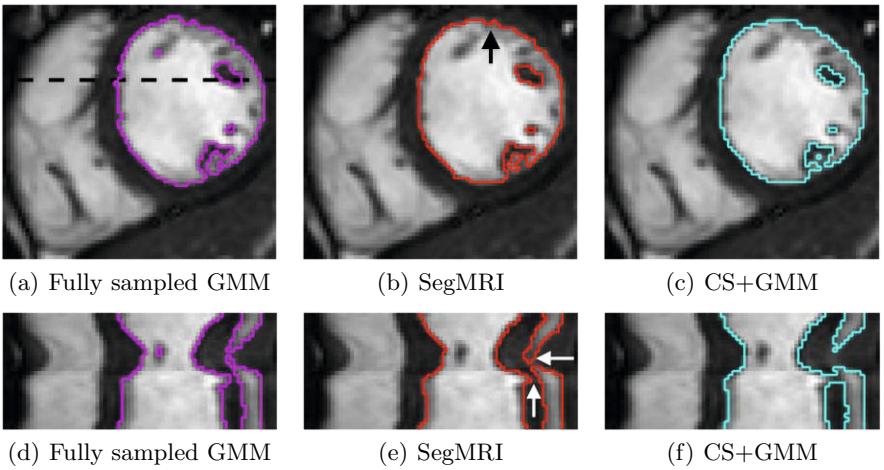
Acceleration	2	4	6	8	10	12	14	**Full set**
CS+GMM	5.1±0.9	10.2±1.4	12.8±1.7	14.6±1.9	15.5±2.1	16.3±2.3	17.0±2.4	$\mathbf{13.1 \pm 4.3}$
SegMRI	3.8±1.0	8.3±1.7	10.8±1.9	12.4±2.2	13.4±2.3	14.0±2.3	14.5±2.4	$\mathbf{11.0 \pm 4.1}$
μ	1.25	1.93	2.07	2.20	2.10	2.28	2.52	**2.05**
t	11.12	10.70	12.19	11.34	11.13	13.53	10.47	**27.59**
p ($\times 10^{-9}$)	3.2	5.7	0.8	2.4	0.8	0.2	7.8	$\mathbf{5.3 \times 10^{-46}}$

3.2 Cardiac Cine Experiments

The same experiment is performed on 7 cardiac cine MRI data, a frame of which is shown in figure 1, using 3D (2D+t) dictionaries. Considering left ventricular structure analysis, the choice $K = 3$ for all tests ensures a clear left ventricle segmentation due to the high intensity difference between blood pool and myocardium. We perform CS reconstruction followed by GMM segmentation on 10-fold accelerated scans and compare them with SegMRI segmentation.

Table 2 shows the 7 patient dataset results of pixel classification for the left ventricle. Again, SegMRI consistently outperforms the disjoint segmentation procedure. Although quantitative differences are small, visual inspection shows that SegMRI is able to capture fine details also revealed in the segmentation run on the fully sampled data. An example is given in figure 3 with the boundaries of

the left ventricle segmentations overlaid onto the fully sampled data on the slice plane and across the temporal profile of the dashed line in figure 3(a). The t-test performed also implies high statistical significance in these results. For cine data, convergence was attained in 20-30 iterations of 60 seconds each.



(a) Fully sampled GMM (b) SegMRI (c) CS+GMM

(d) Fully sampled GMM (e) SegMRI (f) CS+GMM

Fig. 3. Left ventricle segmentation overlaid onto fully sampled data

Table 2. Pixel misclassification (mean±std%) and t-test for left ventricle structure

Acceleration	2	4	6	8	10	12	14	**Full set**
CS+GMM	2.5±1.5	4.8±3.0	6.0±3.9	6.8±4.2	7.3±4.5	7.8±4.7	7.9±4.9	$\mathbf{6.2 \pm 4.4}$
SegMRI	2.0±1.1	4.0±2.6	5.1±3.4	5.9±3.8	6.3±4.2	6.8±4.4	7.0±4.7	$\mathbf{5.3 \pm 4.0}$
μ	0.49	0.78	0.88	0.88	0.90	1.04	0.86	**0.83**
t	3.87	4.55	4.64	4.70	5.68	6.45	5.72	**13.3**
p ($\times 10^{-3}$)	8.3	3.9	3.6	3.3	1.3	0.6	1.2	$\mathbf{9.5 \times 10^{-15}}$

4 Conclusion

We have proposed a joint reconstruction-segmentation framework for undersampled fast acquisitions of MRI data. Imposing a GMM term in the reconstruction provides a segmentation that degrades less with increasing undersampling compared to a separated reconstruction and segmentation processing. Extensions of this work could look at natural additions to GMM segmentations such as Markov random fields and spatial priors for improved robustness. An interesting direction could also further integrate segmentation and reconstruction by performing region specific dictionary learning informed by intermediate segmentations.

The traditional analysis pipeline is burdened with high amounts of data that is time-consuming and expensive to acquire and process, and in scenarios where

the information needed from the images is not the images themselves, but information derived from them, an application-driven scan could provide a more efficient and accurate solution. Image analysis tools, other than the segmentation example discussed in this paper, might also benefit from symbiotically combining post-processing with acquisition and reconstruction.

References

1. Bishop, C.M.: Pattern recognition and machine learning. Springer, New York (2006)
2. Caballero, J., Price, A.N., Rueckert, D., Hajnal, J.V.: Dictionary learning and time sparsity in dynamic MR data reconstruction. IEEE Trans. Med. Imag. 33(4), 979–994 (2014)
3. Caballero, J., Rueckert, D., Hajnal, J.V.: Dictionary learning and time sparsity in dynamic MRI. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part I. LNCS, vol. 7510, pp. 256–263. Springer, Heidelberg (2012)
4. Cabezas, M., Oliver, A., Lladó, X., Freixenet, J., Cuadra, M.B.: A review of atlas-based segmentation for magnetic resonance brain images. Comput. Meth. Prog. Bio. 104(3), 158–177 (2011)
5. Candès, E.J., Romberg, J., Tao, T.: Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory 52(2), 489–509 (2006)
6. Elad, M., Aharon, M.: Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans. Image Process. 15(12), 3736–3745 (2006)
7. Haralick, R.M., Shapiro, L.G.: Image segmentation techniques. Comput. Vision Graph. 29(1), 100–132 (1985)
8. Hsiao, I.T., Rangarajan, A., Gindi, G.: Joint-MAP reconstruction/segmentation for transmission tomography using mixture-models as priors. IEEE Nucl. Sci. Symp. Conf. Rec. 3, 1689–1693 (1998)
9. Lustig, M., Donoho, D., Pauly, J.M.: Sparse MRI: The application of compressed sensing for rapid MR imaging. Magn. Reson. Med. 58(6), 1182–1195 (2007)
10. McInerney, T., Terzopoulos, D.: Deformable models in medical image analysis: A survey. Med. Image Anal. 1(2), 91–108 (1996)
11. Ravishankar, S., Bresler, Y.: MR image reconstruction from highly undersampled k-space data by dictionary learning. IEEE Trans. Med. Imag. 30(5), 1028–1041 (2011)
12. de Sompel, D.V., Brady, M.: Simultaneous reconstruction and segmentation algorithm for positron emission tomography and transmission tomography. In: ISBI 2008, pp. 1035–1038. IEEE (2008)
13. Tosic, I., Frossard, P.: Dictionary learning. IEEE Signal Process. Mag. 28(2), 27–38 (2011)
14. Virtue, P., Uecker, M., Elad, M., Lustig, M.: Predicting image quality of undersampled data reconstruction in the presence of noise. In: Proc. 21st Annual Meeting ISMRM, Salt Lake City, Utah, USA, p. 2668 (2013)
15. Zhang, Q., Plemmons, R., Kittle, D., Brady, D., Prasad, S.: Joint segmentation and reconstruction of hyperspectral data with compressed measurements. Appl. Opt. 50(22), 4417–4435 (2011)