RECONSTRUCTION OF DYNAMIC UNDER-SAMPLED MRI USING SELF-SIMILARITY AMONG 1D TEMPORAL SNIPPETS

Esben Plenge[1], *Mitchell A. Cooper*[2], *Martin R. Prince*[2], *Yi Wang*[2], *Pascal Spincemaille*[2], *Michael Elad*[1]

[1] Division of Computer Science, Technion - Israel Institute of Technology, Haifa, Israel
[2]Department of Radiology, Weill Cornell Medical College, New York, NY

ABSTRACT

This paper introduces a new empirical model for dynamic MRI and shows its application to MRI reconstruction. The model proposes that short 1D signals, so-called *snippets*, along the image's temporal dimension are sparse under non-linear transformation using a compact dictionary trained on the data itself. We employ this model to the problem of reconstructing dynamic abdominal MRI and validate its efficacy on a dynamic computational phantom and on an *in vivo* dynamic MRI sequence. We show how the approach extends and outperforms a state-of-the-art reconstruction algorithm.

Index Terms— Dynamic MRI, Compressed Sensing, Dictionary Learning, Sparse Coding, Abdominal MRI

1. INTRODUCTION

This paper introduces a new MR image model and its application to reconstruction of dynamic magnetic resonance imaging (MRI) of the human abdomen. Abdominal MRI is used, among other things, to detect malignant liver lesions after the injection of a contrast agent. For this, high spatial and temporal resolution is needed. A major challenge in such settings is patient motion, which leads to blurring and other image artifacts that impair the underlying physiological information. A common way of addressing this problem is to ask the patient to lay still and to hold his/her breath during the acquisition. For various reasons, not all patients are able to do so. To complicate matters further, other types of motion are present in abdominal MRI, most notably peristalsis of the gastrointestinal tract.

From a technical perspective, motion artifacts in MRI appear due to limited temporal resolution. The time it takes to sample a single time frame of a dynamic sequence is long enough for significant global and local displacements to take place in the subject. The only way to decrease the acquisition time (given a fixed imaging sequence) is to reduce the number of sample points, that is, decrease the resolution. Any dynamic MRI acquisition thus involves a basic trade-off between spatial and temporal resolution. Decreasing the spatial resolution is undesirable for obvious reasons: details of anatomical structures and physiological processes are lost by increasing the pixel size. The temporal and spatial resolution are therefore carefully balanced according to the diagnostic or other requirements of an acquisition.

The research leading to these results has received funding from the European Research Council under European Union's Seventh Framework Program, ERC Grant agreement no. 320649.

A particularly successful way of improving the spatio-temporal trade-off in MRI has been multi-channel acquisition, or *parallel imaging* [1, 2]. By using multiple coils, several images are acquired simultaneously. The individual coil images are under-sampled along the phase encoding direction and thus acquired faster. Due to the variation in the coil sensitivity fields, the resulting aliasing can be unfolded into a full field-of-view image by the use of computational post-processing. By exploiting image redundancies along the temporal dimension, these concepts have been further leveraged for dynamic imaging [3]. At the core of these computational methods is the following image encoding model

$$\mathbf{y} = \mathbf{UFSx} \tag{1}$$

where $\mathbf{x}$ is the vectorized image, $\mathbf{y} = [\mathbf{y}_1^T \mathbf{y}_2^T ... \mathbf{y}_C^T]^T$ is the concatenation of the data acquired from C coils ($\mathbf{y}^T$ denotes the non-conjugate transpose of $\mathbf{y}$), $\mathbf{U}$ is a sampling operator, $\mathbf{F}$ is the Fourier transform, and $\mathbf{S} = [\mathbf{S}_1 \mathbf{S}_2 ... \mathbf{S}_C]^T$ is a coil sensitivity operator.

More recently, the framework of compressed sensing (CS) and its application to MRI has further improved the spatio-temporal trade-off in dynamic MRI [4, 5, 6]. CS theory asserts that MR images can be under-sampled (below the Nyquist rate), thus decreasing acquisition time, without sacrificing resolution in the reconstructed image. The CS MRI problem can be formulated as follows:

$$\min_x \mathbf{\Phi}(\mathbf{\Psi x}) \quad s.t. \quad ||\mathbf{UFSx} - \mathbf{y}||_2 < \epsilon \tag{2}$$

where $\mathbf{x}$ and $\mathbf{y}$ are vectors containing the reconstructed image and the acquired k-space data, respectively, $\mathbf{\Psi}$ is a sparsifying operator, and $\mathbf{\Phi}$ is a sparsity promoting function, often the L_p ($p \leq 1$) norm. The error threshold ϵ is set according to the image noise level.

CS relies on two things: sparsity of the signal in the transform domain ($\mathbf{\Psi}$), and on the under-sampling artifacts of $\mathbf{\Psi x}$ being incoherent (noise-like). Wavelets and finite difference transforms are widely used as sparsifying transforms in CS MRI, and shrinkage methods for reducing the artifacts in the sparse domain. A different class of transforms use the MR data itself for signal representation [7, 8, 9, 10, 11]. These methods use sparse representations of spatial image patches to reduce under-sampling artifacts in the CS reconstruction.

Great advances have been achieved with respect to sampling efficiency and reconstruction quality of dynamic MRI. Certain under-sampling artifacts, however, remain difficult to get rid off. One of them is the fluctuation of the image intensities (*ripples*) over time. In this work we explore the

application of a new type of image-based transform designed to reduce these temporal ripples in dynamic MRI of the human abdomen. Previous patch-based approaches to dynamic MRI reconstruction have used spatial [11] or spatiotemporal [9, 10] image patches. Here we propose using the self-similarity among short 1D temporal signals (*snippets*) which are retrieved by collecting a number of temporally adjacent signal values at each spatial voxel position. We compress the information contained in these snippets into a sparse coding dictionary which we subsequently use for sparse representation of the snippets themselves.

2. METHOD

2.1. Problem formulation

The CS problem in Eq.2 is often given in its Lagrangian form and can be generalized to allow multiple regularization terms:

$$\min_{x} ||\mathbf{UFSx} - \mathbf{y}||_2^2 + \sum_{q=1}^{Q} \lambda_q \mathbf{\Phi}_q(\mathbf{\Psi}_q) \tag{3}$$

in which λ_q determines the weight of the qth regularizing term. In [11], the dynamic (2D+time) reconstruction problem includes both spatial and temporal regularization terms:

$$\begin{aligned}\mathbf{P1}: \mathbf{x}^*_{t=1,\ldots,T} = arg \min_{\mathbf{x},\{\alpha_{i,j,t}\}} \{||\sum_{t=1}^{T} \mathbf{U}_t\mathbf{FSx}_t - \mathbf{y}_t||_2^2 \\ + \lambda_1 \sum_{t=1}^{T}\sum_{i,j} ||\mathbf{R}^p_{i,j}\mathbf{x}_t - \alpha_{i,j,t}\mathbf{P}_{i,j}\mathbf{D}^p_t||_2^2 \\ + \lambda_2 \sum_{t=1}^{T} ||\mathbf{x}_t - (\mathbf{x}_{t-1} + \mathbf{x_{t+1}})/2||_2^2\}\end{aligned} \tag{4}$$

The first term ensures fidelity between the acquired data of temporal frame t and the data synthesized by the acquisition model $\mathbf{U_tFS}$ from the solution estimate of the tth frame $\mathbf{x}_t$. The second term promotes local proximity between adjacent slices up to a scaling of the intensities: $\mathbf{R}^p_{i,j}$ extracts an $n_p \times n_p$ spatial patch from location i, j in $\mathbf{x}_t$, $\mathbf{P}_{i,j}\mathbf{D}^p_t$ is a local dictionary of patches that is specific to image position i, j and frame t, and $\alpha_{i,j,t}$ is a set of weights that yields the least squares solution of the term. Allowing local scaling in this way is particularly useful in the case of perfusion imaging where image contrast changes locally over time. The third term promotes proximity among adjacent slices in an spatially global fashion.

It is well known that self-similarity among image patches can be exploited for signal enhancement in MRI [12]. Based on the observation of high similarity among temporal 1D snippets in dynamic MRI, we propose to extend **P1** by an additional regularization term (for compactness we abuse notation and keep in mind that **P2** should be minimized with respect to $\mathbf{x}$, $\{\alpha_{i,j,t}\}$, and $\{\beta_{i,j,t}\}$):

$$\begin{aligned}\mathbf{P2}: \mathbf{P1} + \lambda_3 \sum_{i,j} ||\beta_{i,j}||_0 \\ s.t. \quad ||\mathbf{R}^s_{i,j,t}\mathbf{x} - \mathbf{D}^s\beta_{i,j,t}||_2^2 < \epsilon \quad \forall i,j,t\end{aligned} \tag{5}$$

Given a trained dictionary $\mathbf{D}^s$, this term promotes sparsity of the coefficient vector $\beta_{i,j,t}$ subject to a proximity constraint between temporal snippets of the full reconstruction estimate $\mathbf{x}$ and their approximation using $\mathbf{D}^s$. The snippets are of size $1 \times 1 \times n_s$ and are extracted from all locations i, j, t in $\mathbf{x}$ using the operator $\mathbf{R}^s_{i,j,t}$.

The acronym of the method of [11] is PROUD. In the following we shall thus refer to the proposed method by PROUD with extended regularization, or, PROUDer.

2.2. Method implementation

In [11], **P1** was solved for each t by an iterative scheme similar to the one proposed in [8]: the method alternates between solving for α and $\mathbf{x}_t$: first the current $\mathbf{x}_t$ estimate is "denoised" using $\mathbf{D}^p_t$. Next, the denoised estimate is added to the adjacent frames $\mathbf{x}_{t-1}$ and $\mathbf{x}_{t+1}$ in a weighted average. Finally, the residual $(\mathbf{U}_t\mathbf{FSx}_t - \mathbf{y}_t)$ is back-projected onto the solution estimate in a gradient descent step.

The method was initialized by an iteration that included only the two first terms of **P1**. The spatial dictionary $\mathbf{D}^p_t$ in all iterations was constructed by concatenating the temporal frame $\mathbf{x}_{t-1}$ and a reference frame. The reference frame was created by averaging $\mathbf{U}^H_t\mathbf{F}^H\mathbf{S}^H\mathbf{y}_t$ over all time points t ($\mathbf{M}^H$ denotes the conjugate transpose of $\mathbf{M}$). $\mathbf{P}_{i,j}$ was constrained to extract $n_p \times n_p$ patches from the dictionary frames only at position i, j. A patch size of $n_p = 7$ was found experimentally to yield good results.

The new regularization term was incorporated into the above scheme by applying an additional "denoising" step at the end of each iteration. All snippets $\mathbf{R}^s_{i,j,t}\mathbf{x}$ were extracted from $\mathbf{x}$ and replaced by their approximations $\mathbf{D}^s\beta_{i,j,t}$. Here, the error threshold ϵ was chosen experimentally as a multiple of the background standard deviation, and a sparse $\beta_{i,j,t}$ was found by orthogonal matching pursuit [13]. The reconstruction was reassembled by simply averaging the temporally overlapping "denoised" snippets back into $\mathbf{x}$:

$$\mathbf{x} = \left[\sum_{i,j,t} (\mathbf{R}^s_{i,j,t})^T \mathbf{R}^s_{i,j,t}\right]^{-1} \sum_{i,j,t} (\mathbf{R}^s_{i,j,t})^T \mathbf{D}^s \beta_{i,j,t} \tag{6}$$

The dictionary $\mathbf{D}^s$ was trained using K-SVD [14]. The $1 \times 1 \times n_s$ training signals were extracted at either all image locations i, j (for the phantom dataset), or for a large randomly selected subset thereof (for the *in vivo* MRI dataset). n_s was experimentally chosen to be 5 for the phantom experiments and 11 for the MRI data experiments. The number of atoms in the dictionary was set to 20 in both experiments. A simple way to accommodate K-SVD training on complex signals is to concatenate the real and imaginary components of the signals into vectors of length $2n_s$. This was done prior to both training and reconstruction. The image acquisition matrix $\mathbf{U}_t\mathbf{F}$ was implemented by a non-uniform fast Fourier transform [15, 16] and coil sensitivity maps $\mathbf{S}$ were constructed by dividing low-resolution coil images by their root sum of squares. The coil images were obtained by merging all sampled time frames into a single image per coil.

2.3. Experimental setup

The proposed method was validated using two data sets. For quantitative experiments we used a computational phantom of the abdominal region (Fig. 1). The phantom was constructed by segmenting a 3D reference frame of a 4D abdominal MRI scan into aorta, portal vein, liver, kidney, spine, colon, and abdominal wall regions. A cartoon frame was created by applying intensities to the segmented areas resembling the appearance of the acquisition. The remaining 139 temporal 3D frames were non-rigidly registered to the reference frame, and the deformation parameters subsequently applied to the cartoon frame. To reduce the computational burden in our experiments, we used a single cropped 2D slice over time ($256 \times 180 \times 140$). Before simulating the under-sampled acquisitions, additive zero-mean Gaussian noise was added to the phantom. Experiments were performed with noise of standard deviations 0, 10, 20, and 30. A single coil acquisition was simulated by $\mathbf{y}_t = \mathbf{U}_t\mathbf{F}\mathbf{S}\mathbf{x}$, where $\mathbf{S}$ implemented a uniform sensitivity field and $\mathbf{U}_t$ was a variable density spiral trajectory with golden angle rotation of the spiral leaves between each time frame t [17]. As measures of reconstruction quality we used peak signal-to-noise ratio (PSNR) and the mean structural similarity index measure (mean SSIM) (averaged over the time frames).

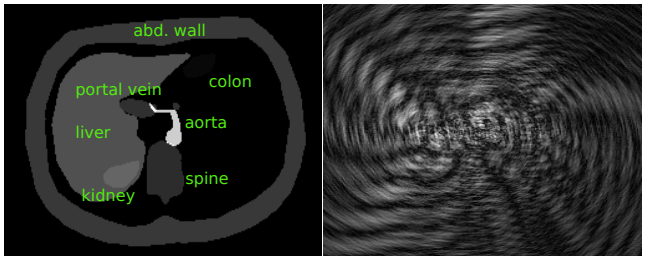

Fig. 1. *Left:* Temporal frame of the phantom. *Right:* The corresponding phantom frame after acquistion simulated by $\mathbf{y}_t = \mathbf{U}_t\mathbf{F}\mathbf{S}\mathbf{x}$, and reconstruction by $\mathbf{x} = (\mathbf{U}_t\mathbf{F}\mathbf{S})^H\mathbf{y}_t$.

For qualitative *in vivo* experiments we used an 8-coil 4D liver perfusion acquisition which use a golden angle spiral LAVA sequence with a stack of variable density spiral trajectories [17]. As for the phantom, a 2D+time image (reconstruction grid size: $256 \times 256 \times 240$) was extracted and used for the experiments.

3. RESULTS

Fig. 2 shows the PSNR and mean SSIM of the phantom reconstruction experiments for varying noise levels. It is seen how PROUDer improves the reconstruction considerably over PROUD, for all noise-levels. PSNR improvements after 11 iterations are 0.7, 1.7, 2.1, and 1.4 dB for noise-levels of 0, 10, 20, and 30, respectively. The corresponding improvements in mean SSIM are 0.26, 0.50, 0.49, and 0.44. It is also seen how PROUDer nearly converges after 1 iteration. The computational overhead of the dictionary-based snippet "denoising" step in PROUDer is thus more than compensated for.

In Fig. 3, a single time frame of the PROUD and PROUDer reconstructions are shown on the left. There

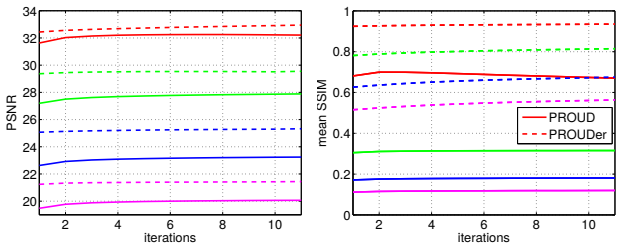

Fig. 2. PSNR and mean SSIM of PROUD and PROUDer reconstructions for simulated acquisitions with noise standard deviations of 0 (*red*), 10 (*green*), 20 (*blue*), 30 (*magenta*).

is hardly any visible difference between these two frames. If the window/level is adjusted, it can be seen how PROUDer reduces the background noise more. In the center and on the right, temporal cross-sections of the reconstructions are shown. Here the effect of the snippet-based regularization in PROUDer can be appreciated: the temporal intensity fluctuations (ripple artifacts) present in the PROUD reconstruction are much less visible in PROUDer after 1 iteration and practically gone after 6.

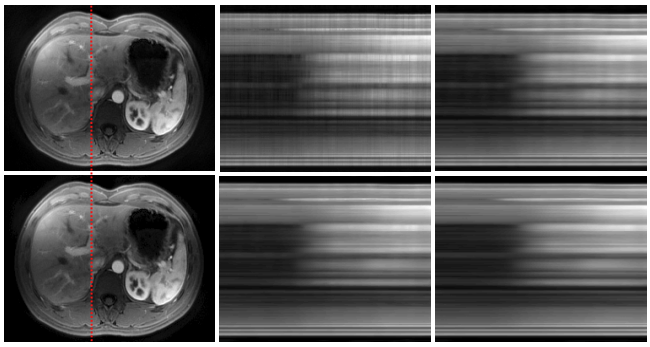

Fig. 3. *Top/bottom:* Reconstruction using PROUD/ PROUDer. *Left to right:* Reconstructed frame, temporal cross-section at dashed line after 1 iteration, and 6 iterations.

In studies of contrast enhanced MRI of the liver, it is common practice to consider the time-intensity curves of the aorta and the portal vein. In Fig. 4, the efficient removal of the ripples in these curves by PROUDer is again seen clearly.

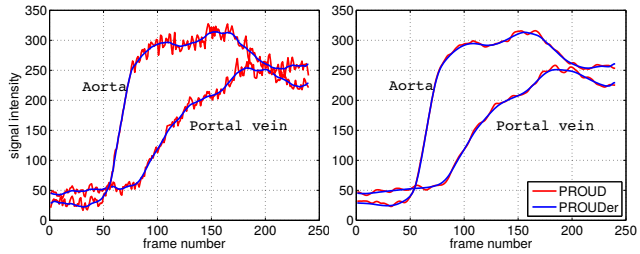

Fig. 4. Time-intensity curves extracted from aorta and portal vein regions in PROUD and PROUDer reconstructions after 1 (*left*) and 6 (*right*) iterations.

4. DISCUSSION

In this paper, we have proposed a new empirical model for dynamic MRI according to which short temporal snippets can be represented sparsely using a trained dictionary. We have shown that the model can be used to regularize an abdominal MRI reconstruction problem. Our reconstruction method compares favourable, both quantitatively and qualitatively, with the state-of-the-art method which it extends, especially in dealing with the temporal intensity fluctuations which are a well-known artifact in spiral under-sampling schemes.

The results presented here are preliminary but point out multiple directions of future research: it will be interesting to apply and validate the snippet-based regularization in other settings: to other dynamic MRI reconstruction applications (e.g. cardiac cine), in conjunction with other sampling schemes (e.g. Cartesian and stack-of-stars sampling), and incorporated into other CS algorithms (e.g. the recent matrix recovery approaches [18, 19]). Also, ways of automatically determining parameters such as snippet length n_s, dictionary size, and the error threshold ϵ should be studied, both in order to avoid tedious tuning work, but also to learn more about the properties of the model.

5. REFERENCES

[1] K.P. Pruessmann, M. Weiger, M.B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magn Reson Med*, vol. 42, no. 5, pp. 952–962, 1999.

[2] M.A. Griswold, P.M. Jakob, R.M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn Reson Med*, vol. 47, no. 6, pp. 1202–1210, 2002.

[3] J. Tsao, P. Boesiger, and K.P. Pruessmann, "k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations," *Magnetic Resonance in Medicine*, vol. 50, no. 5, pp. 1031–1042, 2003.

[4] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn Reson Med*, vol. 58, no. 6, pp. 1182–1195, 2007.

[5] H. Jung, K. Sung, K.S. Nayak, E.Y. Kim, and J.C. Ye, "k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI," *Mag Reson Med*, vol. 61, no. 1, pp. 103–116, 2009.

[6] R. Otazo, D. Kim, L. Axel, and D.K. Sodickson, "Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI," *Magn Reson Med*, vol. 64, no. 3, pp. 767–776, 2010.

[7] M. Doneva, P. Bornert, H. Eggers, C. Stehning, J. Senegas, and A. Mertins, "Compressed sensing reconstruction for magnetic resonance parameter mapping," *Magn Reson Med*, vol. 64, no. 4, pp. 1114–1120, Oct 2010.

[8] S. Ravishankar and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE Trans Med Imag*, vol. 30, no. 5, pp. 1028–1041, 2011.

[9] J. Caballero, A.N. Price, D. Rueckert, and J.V. Hajnal, "Dictionary learning and time sparsity for dynamic MR data reconstruction," *IEEE Trans Med Imag*, vol. 33, pp. 979–994, 2014.

[10] Y. Wang and L. Ying, "Compressed sensing dynamic cardiac cine MRI using learned spatiotemporal dictionary," *IEEE Trans Biomed Eng*, vol. 61, no. 4, pp. 1109–1120, Apr 2014.

[11] M.A. Cooper, T.D. Nguyen, B. Xu, M.R. Prince, M. Elad, Y. Wang, and P. Spincemaille, "Patch based reconstruction of undersampled data (PROUD) for high SNR and high frame rate contrast enhanced liver imaging," *Magn Reson Med, accepted for publication*, 2014.

[12] E. Plenge, D.H.J. Poot, W.J. Niessen, and E. Meijering, "Super-resolution reconstruction using cross-scale self-similarity in multi-slice MRI," in *Proceedings of the 16th MICCAI, Part III*. 2013, pp. 123–130, Springer-Verlag.

[13] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer, 1st edition, 2010.

[14] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans Sig Proc*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[15] J.A. Fessler, "On NUFFT-based gridding for non-Cartesian MRI," *J Magn Reson*, vol. 188, no. 2, pp. 191–195, Oct 2007.

[16] M. Lustig and J.M. Pauly, "SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary k-space," *Magn Reson Med*, vol. 64, no. 2, pp. 457–471, 2010.

[17] M.D. Agrawal, P. Spincemaille, K.W. Mennitt, B. Xu, Y. Wang, S.P. Dutruel, and M.R. Prince, "Improved hepatic arterial phase MRI with 3-second temporal resolution," *J Magn Reson Imag*, vol. 37, no. 5, pp. 1129–1136, 2013.

[18] B. Zhao, J. P. Haldar, A. G. Christodoulou, and Z. P. Liang, "Image reconstruction from highly undersampled (k, t)-space data with joint partial separability and sparsity constraints," *IEEE Trans Med Imaging*, vol. 31, no. 9, pp. 1809–1820, Sep 2012.

[19] R. Otazo, E. Candes, and D.K. Sodickson, "Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components," *Magn Reson Med*, pp. n/a–n/a, 2014.