Published in final edited form as:
Conf Proc IEEE Eng Med Biol Soc. 2010 ; 2010: 3390–3393. doi:10.1109/IEMBS.2010.5627934.

PSF Model-Based Reconstruction with Sparsity Constraint: Algorithm and Application to Real-Time Cardiac MRI

Bo Zhao, **Justin P. Haldar**, and **Zhi-Pei Liang**
Department of Electrical and Computer Engineering and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, 1406 West Green Street, Urbana, IL 61801

Bo Zhao: bozhao1@illinois.edu

Abstract

The partially separable function (PSF) model has been successfully used to reconstruct cardiac MR images with high spatiotemporal resolution from sparsely sampled (k,t)-space data. However, the underlying model fitting problem is often ill-conditioned due to temporal undersampling, and image artifacts can result if reconstruction is based solely on the data consistency constraints. This paper proposes a new method to regularize the inverse problem using sparsity constraints. The method enables both partial separability (or low-rankness) and sparsity constraints to be used simultaneously for high-quality image reconstruction from undersampled (k,t)-space data. The proposed method is described and reconstruction results with cardiac imaging data are presented to illustrate its performance.

I. INTRODUCTION

Several model-based methods have been proposed to enable reconstruction from highly undersampled (**k**, *t*)-space data [1]–[5]. The partially separable function (PSF) model assumes that the (**k**, *t*)-space MR signal *s*(**k**, *t*) is partially separable to the *L*th order [4] such that

$$s(\mathbf{k},t)=\sum_{\ell=1}^{L}u_\ell(\mathbf{k})v_\ell(t), \tag{1}$$

where $\{u_\ell(\mathbf{k})\}_{\ell=1}^{L}$ and $\{v_\ell(t)\}_{\ell=1}^{L}$ can be treated as spatial and temporal basis functions, respectively. Due to strong spatiotemporal correlations in cardiac MR signals, *L* can usually be chosen to be very small [6], which significantly reduces the number of degrees of freedom of the model.

Consider the following Casorati matrix formed from the samples of *s*(**k**, *t*),

$$\mathbf{S}=\begin{bmatrix} s(\mathbf{k}_1,t_1) & \dots & s(\mathbf{k}_1,t_M) \\ \vdots & \ddots & \vdots \\ s(\mathbf{k}_N,t_1) & \dots & s(\mathbf{k}_N,t_M) \end{bmatrix}, \tag{2}$$

where N represents the number of k-space points at each time instant for a total of M instants in an 'ideal' data set $\{s(\mathbf{k}_n, t_m)\}_{n,m=1}^{N,M}$. If $s(\mathbf{k}, t)$ is Lth-order partially separable, it is easy to show that the rank of $\mathbf{S}$ is upper bounded by L [4] and $\mathbf{S}$ can be decomposed as

$$\mathbf{S}=\mathbf{U}_k\mathbf{V}_t, \tag{3}$$

where $\mathbf{V}_t \in \mathbb{C}^{L\times M}$ is a basis for the temporal subspace of $\mathbf{S}$, and $\mathbf{U}_k \in \mathbb{C}^{N\times L}$ is a matrix of k-space coefficients. With sparse sampling, many entries of $\mathbf{S}$ are missing. The actual measured data can be expressed as

$$\mathbf{d}=\Omega(\mathbf{S})+\xi, \tag{4}$$

where $\mathbf{d} \in \mathbb{C}^p$ is the measured data, Ω: $\mathbb{C}^{N\times M} \rightarrow \mathbb{C}^p$ is the sparse sampling operator, and $\xi \in \mathbb{C}^p$ represents measurement noise. In this paper, we assume that $\mathbf{V}_t$ is known (i.e., predetermined from navigator signals as is done in [4], [6]) and our goal is to determine $\mathbf{U}_k$ from $\mathbf{d}$. The basic PSF method solves the following unconstrained problem:

$$\widehat{\mathbf{U}}_k = \arg\min_{\mathbf{U}_k\in\mathbb{C}^{N\times L}} \left\|\mathbf{d}-\Omega(\mathbf{U}_k\mathbf{V}_t)\right\|_2. \tag{5}$$

A necessary condition for (5) to be well-posed is that each k-space location is sampled at least L times. In practice, we need much more measurements to avoid a poorly conditioned inverse problem. In the previous work, spatial-spectral support information has been exploited to improve the conditioning of the problem and reduce reconstruction artifacts [7], [8].

In this paper, we use an ($\mathbf{x}$, f)-domain sparsity constraint to regularize the inverse problem. An ($\mathbf{x}$, f)-domain sparsity constraint has been used in previous work [9]–[11]. Our method is the first attempt to use both explicit partial separability and ($\mathbf{x}$, f)-domain sparsity for image reconstruction from highly undersampled ($\mathbf{k}$, t)-space data. The rest of paper is organized as follows: Section II describes the proposed formulation and image reconstruction algorithm; Section III presents an application of the proposed method to real-time cardiac MRI and discussion of related work, followed by a brief conclusion in Section IV.

II. PSF solution with sparsity constraint

Incorporating an ($\mathbf{x}$, f)-domain sparsity constraint, the PSF reconstruction problem can be formulated as

$$\hat{\mathbf{U}}_s = \arg\min_{\mathbf{U}_s\in\mathbb{C}^{N\times L}} \|\mathbf{d}-\Omega(\mathscr{F}_s\mathbf{U}_s\mathbf{V}_t)\|_2^2 + \lambda\,\|\mathbf{U}_s\mathbf{V}_t\mathbf{F}_t\|_1\,, \tag{6}$$

where $\mathscr{F}_s$ represents a spatial Fourier transform operator, $\mathbf{F}_t$ is a temporal Fourier transform matrix,

$$\mathbf{U}_s = \mathscr{F}_s^{-1}\mathbf{U}_k$$

, and λ is a regularization parameter. In the above expression, both $\mathscr{F}_s$ and $\mathbf{F}_t$ are assumed to be orthonormal, and $\|\cdot\|_1$ of a matrix is defined as the sum of the magnitudes of all entries in the matrix. The $\ell 1$ penalty (or $\ell 1$ regularization) in (6) has been widely used for finding sparse solutions to ill-posed inverse problems [12].

We propose to solve (6) using an additive half-quadratic regularization algorithm [11], [13], [14] and a continuation procedure [11], [15], [16]. First, the non-differentiable $\ell 1$ norm is replaced by a differentiable approximation

$$\left\|\mathbf{U}_s\mathbf{V}_f\right\|_1 \approx \sum_{n=1}^{N}\sum_{m=1}^{M}\phi(|(\mathbf{U}_s\mathbf{V}_f)_{n,m}|), \tag{7}$$

where $\phi(t)$: $\mathbb{R} \rightarrow \mathbb{R}$ is the Huber function defined as

$$\phi(t) = \begin{cases} t^2/2\alpha, & \text{if} |t| \le \alpha, \\ |t| - \alpha/2, & \text{if} |t| > \alpha. \end{cases} \tag{8}$$

Note that $\phi(t)$ is a continuously differentiable function, and the accuracy of the approximation in (7) is controlled by the parameter α. As $\alpha \rightarrow 0$, the Huber function approaches $|t|$. Furthermore, it can be shown that $\phi(t) = \min_{g\in\mathbb{R}}\{\frac{1}{2\alpha}(t-g)^2 + |g|\}$. Therefore, (7) can be rewritten as

$$\sum_{n=1}^{N}\sum_{m=1}^{M}\phi(|\mathbf{U}_s\mathbf{V}_f|_{n,m}) = \min_{\mathbf{G}\in\mathbb{C}^{N\times M}}\frac{1}{2\alpha}\left\|\mathbf{U}_s\mathbf{V}_f - \mathbf{G}\right\|_F^2 + \left\|\mathbf{G}\right\|_1, \tag{9}$$

where $\mathbf{G}$ is an auxiliary matrix and $\|\cdot\|_F$ denotes the Frobenius norm. Using (7) and (9), the problem in (6) can be approximated as

$$\{\hat{\mathbf{U}}_s, \hat{\mathbf{G}}\} = \arg\min_{\{\mathbf{U}_s,\ \mathbf{G}\}} \|\Omega(\mathscr{F}_s\mathbf{U}_s\mathbf{V}_t) - \mathbf{d}\|_2^2 + \frac{\lambda}{2\alpha}\left\|\mathbf{U}_s\mathbf{V}_f - \mathbf{G}\right\|_F^2 + \lambda\left\|\mathbf{G}\right\|_1, \tag{10}$$

We solve (10) using an alternating optimization procedure, which first optimizes $\mathbf{U}_S$ with a fixed value of $\mathbf{G}$, and then optimizes $\mathbf{G}$ with a fixed value of $\mathbf{U}_S$. This procedure is repeated until $\mathbf{U}_S$ and $\mathbf{G}$ converge. At the ℓth iteration and for a fixed $\mathbf{U}_s^{(\ell-1)}$, minimizing (10) over $\mathbf{G}$ is accomplished using the following soft-shresholding operation

$$\mathbf{G}_{n,m}^{(\ell)}=\begin{cases} 0, & \text{if} |\mathbf{Q}_{n,m}|<\alpha, \\ \frac{\mathbf{Q}_{n,m}}{|\mathbf{Q}_{n,m}|}(|\mathbf{Q}_{nm}|-\alpha), & \text{if} |\mathbf{Q}_{n,m}| \geq \alpha, \end{cases} \tag{11}$$

where $\mathbf{Q}=\mathbf{U}_s^{(\ell-1)}\mathbf{V}_f$. With a fixed $\mathbf{G}^{(\ell)}$, minimizing (10) with respect to $\mathbf{U}_s$ is equivalent to

$$\hat{\mathbf{U}}_s = \arg\min_{\mathbf{U}_s\in\mathbb{C}^{N\times L}} \|\Omega(\mathscr{F}_s\mathbf{U}_s\mathbf{V}_t)-\mathbf{d}\|_2^2+\frac{\lambda}{2\alpha}\left\|\mathbf{U}_s\mathbf{V}_f-\mathbf{G}^{(\ell)}\right\|_F^2 \tag{12}$$

By changing variables $\mathbf{U}_k = \mathscr{F}_s\mathbf{U}_s$ and $\mathbf{B}^{(\ell)} = \mathbf{F}_s\mathbf{G}^{(\ell)}\mathbf{F}_t$, we can convert (12) into

$$\widehat{\mathbf{U}}_k= \arg\min_{\mathbf{U}_k\in\mathbb{C}^{N\times L}} \left\|\Omega(\mathbf{U}_k\mathbf{V}_t)-\mathbf{d}\right\|_2^2+\frac{\lambda}{2\alpha}\|\mathbf{U}_k\mathbf{V}_t-\mathbf{B}^{(\ell)}\|_F^2. \tag{13}$$

Though (13) is a large-scale quadratic optimization problem, it has a simple decoupled structure, which enables very efficient evaluation. Specifically,

$$\Omega(\mathbf{U}_k\mathbf{V}_t)=\begin{bmatrix} \Omega_1(\mathbf{u}_k^1\mathbf{V}_t) \\ \vdots \\ \Omega_N(\mathbf{u}_k^N\mathbf{V}_t) \end{bmatrix}, \tag{14}$$

where $\mathbf{u}_k^n \in \mathbb{C}^{1\times L}$ represents the nth row of $\mathbf{U}_k$, and Ω_n is the sampling operator that takes samples from the nth row of $\mathbf{U}_k\mathbf{V}_t$. Thus, (13) is equivalent to solving

$$\widehat{\mathbf{u}}_k^n= \arg\min_{\mathbf{u}_k^n\in\mathbb{C}^{1\times L}} \left\|\Omega_n(\mathbf{u}_k^n\mathbf{V}_t)-\mathbf{d}_n\right\|_2^2+\frac{\lambda}{2\alpha}\|\mathbf{u}_k^n\mathbf{V}_t-\mathbf{b}_n^{(\ell)}\|_F^2 \tag{15}$$

for $1\leq n\leq N$, where $\mathbf{d}_n$ is the measured data from the nth row of $\mathbf{S}$, and $\mathbf{b}_n^{(\ell)}$ is the nth row of $\mathbf{B}^{(\ell)}$. In (15), each quadratic optimization problem is equivalent to

$$\left(\mathscr{A}_n^*\mathscr{A}_n+\frac{\lambda}{2\alpha}\mathscr{B}_n^*\mathscr{B}_n\right)\mathbf{u}_k^n=\mathscr{A}_n^*\mathbf{d}_n+\frac{\lambda}{2\alpha}\mathscr{B}_n^*\mathbf{b}_n^{(\ell)} \tag{16}$$

for $1 \leq n \leq N$, where $\mathscr{A}_n(\cdot) = \Omega_n(\cdot\mathbf{V}_t)$, $\mathscr{B}_n(\cdot) = \cdot\mathbf{V}_t$, $\mathscr{A}_n^*$ and $\mathscr{B}_n^*$ are the adjoint operators of $\mathscr{A}_n$ and $\mathscr{B}_n$, respectively. Since each normal equations in (16) only has L unknowns, it can be solved very efficiently. For example, ideally the conjugate gradient method can solve (15) or (16) in only L iterations if the exact arithmetic is assumed [17]. Moreover, the completely decoupled structure is quite suitable for parallel computation for further computational accelerations.

The above alternating optimization procedure gives an optimal solution to (10). To obtain a good approximate to (6), a continuation scheme is used. The continuation procedure starts with a large value of α, which causes the alternating algorithm to converge rapidly [14]. Then we decrease the value of α in (10) and use the previous solution as a 'warm start' for the new alternating optimization. The above procedure is repeated until the cost function of (10) well approximates the cost function in (6).

III. Results and Discussion

A. Simulation Results

The proposed method has been validated systematically using both human and rat cardiac phantoms. These phantoms were used to simulate free-breathing cardiac imaging experiments with variable heartbeat periods. Snapshot images from two phantoms in the diastolic phase of the cardiac cycle are shown in Fig. 1. In the experiments, 200 phase encodings and 256 samples per readout were used with the human phantom, while 256 phase encodings and 256 samples per readout were used with the rat phantom.

The data acquisition scheme shown in Fig. 2 was used to collect measurements. This sampling pattern consists of acquiring training data in the center of *k*-space and imaging data in the other region of *k*-space (See Fig. 2). The training data is used for estimating $\mathbf{V}_t$, while the imaging data is used to solve $\mathbf{U}_s$. For the simplicity of discussion, we define the number of data frames as M/N_p, where N_p represents the number of phase encodings at each time instant. In the simulation, 11 data frames of data were of data were generated for both phantoms.

In addition to the proposed method (order-16 PSF with sparsity constraint), two other reconstruction methods were applied to process this data: order-6 PSF (the best-performing basic PSF method for the data set tested) and $\ell 1$-regularized reconstruction without PSF modeling [10]. For both methods including $\ell 1$-regularization, the regularization parameters were chosen manually to optimize the performance.

Some representative snapshot images reconstructed from the human and rat data sets (zoomed-in to the heart) are shown in Fig. 3 and 4, respectively. It is clear that the proposed method has a much better spatial resolution than the other two methods, particularly in the myocardium and the blood pool, where the other two methods produced significant artifacts. Relative reconstruction error defined as, i.e.,

$$\text{error} = \frac{\left\| \mathbf{S} - \widehat{\mathbf{U}}_k \mathbf{V}_t \right\|_F}{\left\| \mathbf{S} \right\|_F}, \tag{17}$$

are shown in Table I, which are consistent with the improved performance of the proposed method.

Tests were also done to compare the amount of data required for (a) $\ell 1$ regularized reconstruction without PSF modeling, (b) basic PSF with $L = 16$, and (c) the proposed method with $L = 16$. Reconstruction errors as a function of the amount of acquired data are shown in Fig. 5, and reconstructed images are shown in Fig. 6. As expected, reconstruction performance improves with the increasing data for both basic PSF and the proposed method. However, the proposed method can reconstruct higher quality images especially when the number of data frames is small compared with the basic PSF with the same model order. For

$\ell 1$ regularized reconstruction without PSF modeling, the reconstructed images suffer from significant blurring artifacts although the reconstructions are stable with different amount of measurements.

B. Discussion

Simultaneous use of rank and sparsity constraints was also considered in [11]. Instead of imposing an explicit rank constraint, the algorithm in [11] minimizes the nuclear norm. We have observed that the explicit rank constraint can be more effective than the nuclear norm rank constraint [18]. In terms of computation, the proposed algorithm has the following advantages over the $\ell 1$ and nuclear norm minimization algorithm. The algorithm in [11] requires one SVD (or partial SVD) at each iteration, which is much more time-consuming than solving the decoupled least squares problems in (16). In addtion, for the proposed method both spatial and temporal Fourier transforms can be computed on the spatial and temporal basis, instead of the whole data matrix, which is more efficient.

IV. Conclusion

A new method has been described for reconstruction of highly undersampled ($\mathbf{k}$, t)-space data. This method simultaneously enforces partial separability (low-rankness) and encourages sparsity in the ($\mathbf{x}$, f)-domain. The resulting reconstruction problem is efficiently solved using an additive half-quadratic optimization procedure. The proposed method has been validated with simulated data, and shows significantly better performance than the existing methods that use rank or sparsity constraints alone.

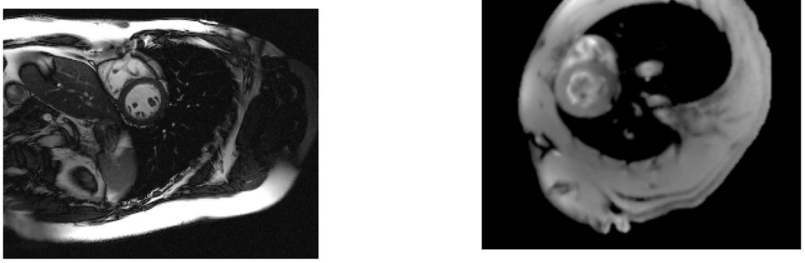
Acknowledgments

This work was supported in part by the following research grants: NIHP41-EB-00197, NIH-P41-RR-02395, and NSF-CBET-07-30623.

References

1. Liang ZP, Jiang H, Hess CP, Lauterbur PC. Dynamic imaging by model estimation. Int J Imaging Syst Tech. 1997; 8:551–557.
2. Tsao J, Boesiger P, Pruessmann KP. k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploring spatiotemporal sparsity. Magn Reson Med. 2003; 50:1031–1042. [PubMed: 14587014]
3. Lustig M, Santos JM, Donoho DL, Pauly JM. k-t SPARSE: high frame rate dynamic mri exploiting spatio-temporal sparsity. Proc Int Symp Magn Reson Med. 2006:2420.
4. Liang ZP. Spatiotemporal imaging with partially separable functions. Proc IEEE Int Symp Biomed Imaging. 2007:988–991.
5. Aggarwal N, Bresler Y. Patient-adapted reconstruction and acquisition dynamic imaging method (PARADIGM) for MRI. Inverse Probl. 2008; 24:045015.
6. Brinegar C, Wu YJL, Foley LM, Hitchens TK, Ye Q, Ho C, Liang ZP. Real-time cardiac MRI without triggering, gating, or breath holding. Proc IEEE Eng Med Bio Conf. 2008:3381–3384.
7. Brinegar C, Zhang H, Wu YJL, Foley LM, Hitchens TK, Ye Q, Pocci D, Lam F, Ho C, Liang ZP. Real-time cardiac MRI using prior spatial-spectral information. Proc IEEE Eng Med Bio Conf. 2009:4383–4386.
8. Christodoulou AG, Brinegar C, Haldar JP, Zhang H, Wu YJL, Foley LM, Hitchens TK, Ye CHQ, Liang ZP. High-resolution cardiac MRI using partially separable functions and weighted spatial smoothness regularization. Proc IEEE Eng Med Bio Conf. 2010 to be published.
9. Gamper U, Boesiger P, Kozerke S. Compressed sensing in dynamic MRI. Magn Reson Med. 2008:365–373. [PubMed: 18228595]

10. Jung H, Sung K, Nayak KS, Kim EY, Ye JC. k-t FOCUSS: a general framework for high resolution dynamic MRI. Magn Reson Med. 2009:103–116. [PubMed: 19097216]
11. Goud S, Hu Y, Jacob M. Real-time cardaic MRI using low rank and sparsity penalties. Proc IEEE Int Symp Biomed Imaging. 2010:988–991.
12. Donoho DL. Compressed sensing. IEEE Trans Inf Theory. 2006; 52(4):1289–1306.
13. Geman D, Yang C. Nonlinear image recovery with half-quadratic regularization. IEEE Trans Image Process. 1995; 4(7):932–946. [PubMed: 18290044]
14. Nikolova M, Ng M. Analysis of half-quadratic minimization methods for signal and image recovery. SIAM J Sci Comput. 2005; 27(3):937–965.
15. Madsen K, Bruun H. A finite smoothing algorithm for linear ℓ1 estimation. SIAM J Optimization. 1993; 3(2):223–225.
16. Becker, S.; Bobin, J.; Candès, EJ. NESTA: A fast and accurate first-order method for sparse recovery. Preprint. 2009. http://arxiv.org/abs/0904.3367
17. Björck, Å. Numerical methods for least squares problems. Philadephia, PA: SIAM; 1996.
18. Haldar JP, Hernando D. Rank-constrained solutions to linear matrix equations using PowerFactorization. IEEE Signal Process Lett. 2009; 16:584–587.



(a) human phantom (b) rat phantom

Fig. 1.
Snapshot images in the human and rat phantoms.

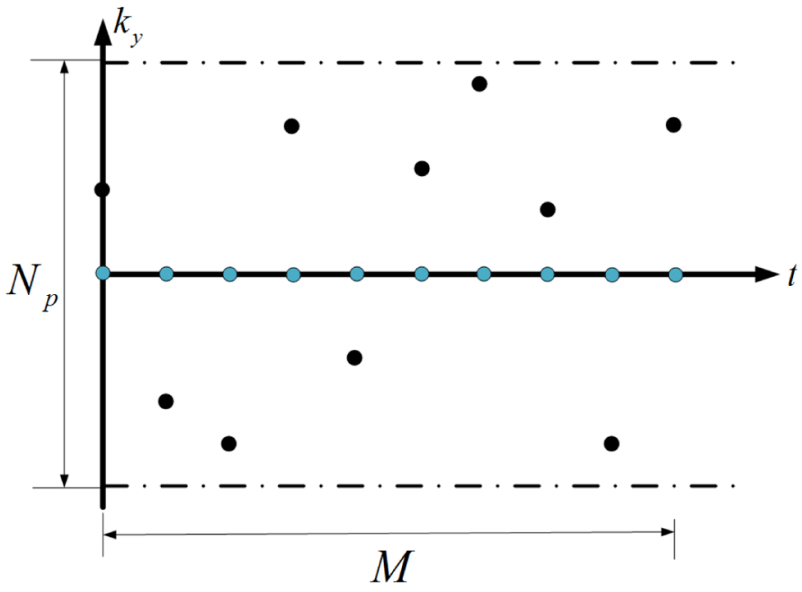


Fig. 2.
Data acquisition scheme in the simulation. Blue dots represent the training data, while black dots represent the imaging data.

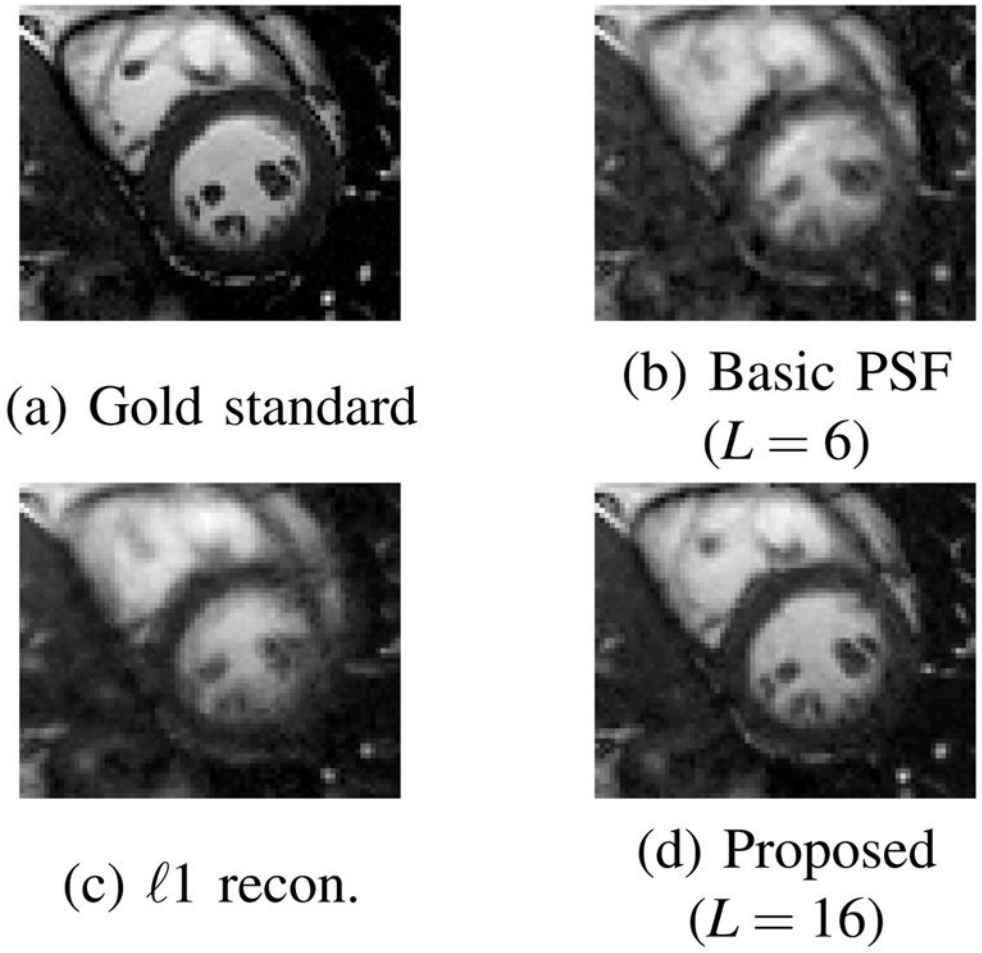


Fig. 3.
Reconstructions of the human phantom (zoomed-in to the heart) using 11 data frames.

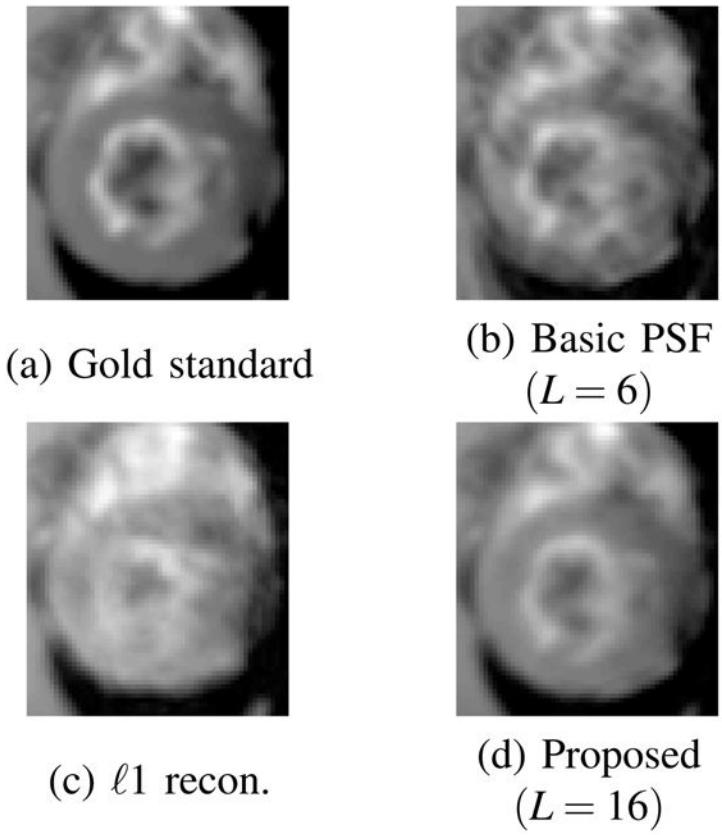


Fig. 4.
Reconstructions of the rat phantom (zoomed-in to the heart) using 11 data frames.

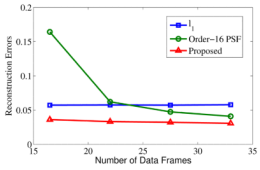


Fig. 5.

Reconstruction errors versus the number of data frames for the human phantom. The errors for ℓ1 regularized reconstruction almost do not change with the increasing amount of data. Basic PSF reconstruction needs a large amount of measured data for accurate reconstruction, while the proposed method can yield high-quality reconstructions with much less data.

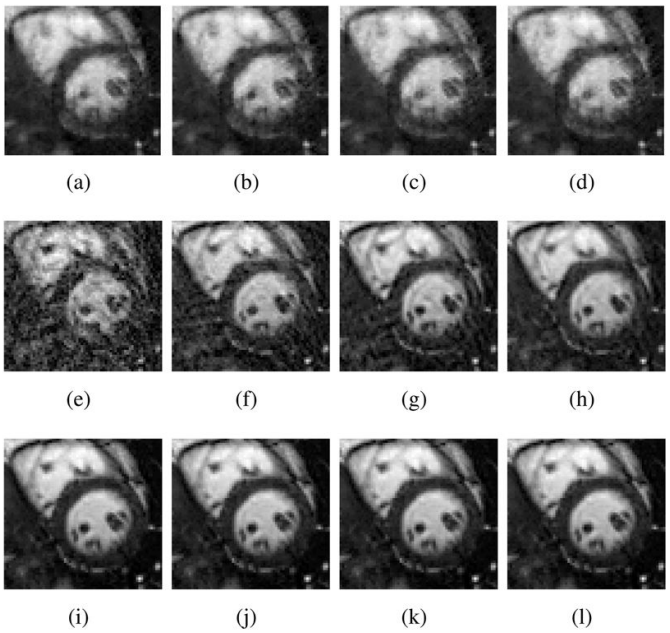


Fig. 6.

Reconstructions of the human phantom (zoomed-in to the heart). (a)–(d) show $\ell 1$ regularized reconstructions without PSF modeling using 16.5, 22, 27.5, 33 data frames, respectively; (e)–(h) show basic order-16 basic PSF reconstructions using 16.5, 22, 27.5, and 33 data frames, respectively; (i)–(l) show reconstructions with the proposed method (order-16 PSF with sparsity constraint) using 16.5, 22, 27.5, and 33 data frames, respectively.

TABLE I

Relative reconstruction errors.

Phantom	Order-6 PSF	ℓ1 recon.	Proposed
Human	6.45%	5.98%	**4.01**%
Rat	4.94%	6.90%	**2.10**%