Parallel Acquisition & Compressed Sensing

Sebastian Kozerke, PhD
Institute for Biomedical Engineering, University and ETH Zurich, Zurich, Switzerland
Imaging Sciences and Biomedical Engineering, King’s College London, UK

Introduction

Resolution and scan time are key parameters in every MR experiment and of particular concern in cardiovascular applications necessitating fast imaging to capture time-varying objects within a limited acquisition window and/or total scan time. To this end it is not surprising that developments of parallel imaging and compressed sensing have mainly been driven by requirements from cardiovascular imaging.

Assuming that particular imaging properties such as contrast and repetition time should remain unchanged, the only option to reduce the acquisition window and/or scan time is to decrease the number of interleaves or phase-encodes $N$. While resolution is not to be compromised, a reduction in $N$ translates into reduced density of sampling of spatial frequencies in k-space and hence is referred to as undersampling. As k-space sampling density and image field-of-view are inversely proportional, the consequences of undersampling are aliasing artifacts in the image domain. This aliasing may be ignored or filtered in particular applications and at low undersampling rates (1,2). In general, however, dedicated treatment in image reconstruction is required to estimate the missing k-space data.

To facilitate further discussion, the object defined in two or three spatial dimensions is assumed to consist of a discrete set of complex-valued voxels stacked in vector $\hat{\rho}$. Accordingly, the process of MR encoding reads:

$$\vec{d} = \Gamma \cdot FT \cdot S \cdot \hat{\rho} + \vec{n}$$  \hspace{1cm} [1]

with $\vec{d}$ stacking the measured k-space data for all receiver coils, $\Gamma$ denoting the undersampling operator, $FT$ the Fourier transform operator, coil sensitivities $S$ and noise $\vec{n}$. By introducing matrix $E = \Gamma \cdot FT \cdot S$ the forward model of MR encoding simplifies to:

$$\vec{d} = E \cdot \hat{\rho} + \vec{n}$$  \hspace{1cm} [2]

Image reconstruction now refers to solving the inverse problem to [2] i.e. finding image $\vec{\imath}$ that is consistent with the acquired data $\vec{d}$ within the uncertainty given by the noise variance $\varepsilon = \frac{1}{N-1} \sum_{n} n_{i}^{2}$:

$$\|\vec{d} - E \cdot \vec{\imath}\|_{2}^{2} \leq \varepsilon$$  \hspace{1cm} [3]

Traditionally, image reconstruction in standard and parallel imaging has been considered a linear problem. For an overdetermined problem, image $\vec{\imath}$ is found using the normal equation $E^{H} \vec{d} = E^{H}E \vec{\imath}$ according to:

$$\vec{\imath} = (E^{H}E)^{-1}E^{H} \vec{d}$$  \hspace{1cm} [4]

Besides data $\vec{d}$, equation [4] requires knowledge of the undersampling pattern $\Gamma$ and coil sensitivities $S$. While the interleave/phase-encode table of the pulse program defines matrix $\Gamma$, the coil sensitivities $S$ have to be estimated from a pre-scan, from the data itself or from fully sampled reference lines acquired with the actual data.

Without undersampling ($R=1$), equation [4] simplifies to the inverse Fourier transform resulting in sensitivity-weighted images for all receive coils:

$$S\vec{\imath} = E^{H} \vec{d} = FT^{-1} \vec{d}$$  \hspace{1cm} [5]

Here the size of vectors $\vec{d}$ and $\vec{\imath}$ are equal. Without taking into account any prior knowledge about the object imaged it is implied that one needs to acquire as many k-space samples $\vec{d}$ as there are voxels $\vec{\imath}$ to reconstruct.
Parallel imaging

To reduce the number of interleaves/phase-encodes $N$, undersampling is employed in conjunction with coil array detection. If the reduction factor $R$ is smaller than or equal to the number of independent coil elements $N_c$ the linear reconstruction problem reads (3):

$$\tilde{t} = (E^H\psi^{-1}E)^{-1}E^H\psi^{-1}\tilde{d}$$  \hspace{1cm} [6]

Here the noise covariance matrix $\psi$ is included to account for differences in coupling of the coil elements to the object and potential mutual coupling among coil elements. Prior knowledge about the object may be incorporated to further regularize equation [6] to avoid small values in the inverse and hence excessive amplification of noise. Such prior knowledge may, for example, identify background voxels that are known to contain noise only (3). By introducing the regularization matrix $\Theta$ the image reconstruction problem is now written as:

$$\tilde{t} = (E^H\psi^{-1}E + \lambda\Theta^{-1})^{-1}E^H\psi^{-1}\tilde{d}$$  \hspace{1cm} [7]

Parameter $\lambda$ permits adjusting the influence of the prior information. If the reduction factor exceeds the number of independent coil elements i.e. $R>N_c$ the reconstruction problem becomes underdetermined and necessitates regularization:

$$\tilde{t} = \Theta E^H(E\Theta E^H + \lambda\psi)^{-1}\tilde{d}$$  \hspace{1cm} [8]

While the equations above are directly applicable to image-domain parallel imaging approaches as introduced with SENSE (3), one may apply the Fourier transform to equation [6] to arrive at an equivalent k-space representation as used in GRAPPA and variants thereof (4,5) providing unfolded images for each coil element:

$$S\tilde{t} = FT^{-1}(FT((E^H\psi^{-1}E)^{-1}E^H\psi^{-1})\tilde{d}) = FT^{-1}(F\tilde{d})$$  \hspace{1cm} [9]

The practical fact that coil sensitivities are spatially smooth functions and hence have a compact representation in k-space requires only few entries in matrix $F$ to reconstruct an image to good approximation (6).

Limits

Regardless of the particular parallel imaging method used, there are limits to the maximum undersampling factor $R$ achievable in practice. The “ease” of solving the inverse problem is expressed by the spatially dependent geometry factor $g$ (3):

$$g = \sqrt{(E^H\psi^{-1}E)_{i,j}(E^H\psi^{-1}E)_{j,i}} \geq 1$$  \hspace{1cm} [10]

Accordingly, the signal-to-noise ratio (SNR) in the reconstructed image will be inversely proportional to both the $g$-factor and the square root of the reduction factor $R$:

$$SNR \propto \frac{1}{g(x)\sqrt{R}}$$  \hspace{1cm} [11]

Intuitively, the ability of surface coils to encode spatial information is given by the orthogonality of their spatial sensitivity patterns, which is essentially determined by fundamental electrodynamics (7,8). It has been demonstrated that beyond a reduction factor of about 4, the $g$-factor increases exponentially when reducing the number of phase-encodes along one dimension in Cartesian imaging. If reduction of phase-encodes along two orthogonal dimensions is possible, the critical reduction factor can roughly be squared yielding a critical reduction factor of 16. These theoretical limits are now being approached with the wider available of large coil arrays featuring 32 and more independent receive elements (9-11).

By including prior information about the image to expect, small entries in the inverse in equations [7] and [8] can be avoided and, accordingly, excessive noise amplification reduced. The prior knowledge matrix $\Theta$ may be populated based on estimates from low-resolution data, which can be obtained separately or interleaved with the undersampled acquisition. In k-t SENSE and related techniques (12-15) the signal estimates in $\Theta$ have been referred to as training data.
Implementation

Besides the reconstruction domain (cp. equations [6] and [9]), the main difference between the various implementations relates to the process of measuring coil sensitivities $S$. Sensitivities can be obtained explicitly using separate pre-scans (3), or implicitly using auto-calibration data acquired interleaved with the undersampled acquisition (16) or using the temporal average of data undersampled in k-t space (17). Radial and spiral trajectories inherently provide a densely sampled k-space center and hence provide coil sensitivity data implicitly (18, 19).

A number of implementation issues of parallel imaging have, however, been identified in relation to explicit and implicit calibration (20-22). Inaccuracies of coil sensitivities may arise from positional inconsistencies in particular when explicit calibration is used. Likewise, partial volume effects, aliasing artifacts and noise in the sensitivity maps can compromise reconstruction quality.

A solution to address inconsistencies between sensitivities and data includes sensitivities as an unknown into the reconstruction process. In JSENSE sensitivities are modeled by a low-order polynomial function and its coefficients are estimated as part of image reconstruction (23) hence treating both the image and the sensitivities as unknowns. However, reconstruction errors occur if the model fails to describe the coil sensitivities.

An alternative to modeling sensitivities explicitly is to recast the reconstruction task as a set of linear equations in k-space (24):

$$\text{minimize} \| (G - I) \vec{d}_c \|_2^2 \text{ s.t. } \| \vec{d} - D \cdot \vec{d}_c \|_2^2 \leq \varepsilon \tag{12}$$

Here $G$ is a calibration matrix used to synthesize a point from its neighborhood. It allows extracting coil sensitivities $S$. Vector $\vec{d}_c$ contains the reconstructed Cartesian k-space samples and operator $D$ relates the Cartesian k-space samples to the data acquired on an (arbitrary) trajectory.

Non-linear reconstruction

While the image reconstruction task up to this point as been considered a linear problem, recent work has indicated that regularized non-linear inversion can address some of the implementation issues associated with parallel imaging (25). Moreover, it has been questioned if the image reconstructed using the pseudo-inverse in equation [4] is really the best choice among all possible images satisfying the data consistency constraint in equation [3].

Instead of expressing image reconstruction as a linear problem, it may also be modeled as a non-linear operator equation with operator $F$ simultaneously mapping the complex-valued voxels $\vec{\rho}$ and the coil sensitivities $\vec{s}$ for each of the $N_c$ coils to k-space data $\vec{d}$:

$$\vec{d} = F(\vec{\rho}, \vec{s}_1 \ldots \vec{s}_{N_c}) \tag{13}$$

The non-linear system can be inverted by linearizing the functional around an initial guess $(\vec{\rho}_0, \vec{s}_1 \ldots \vec{s}_{N_c})'$ and updating the solution iteratively using the Gauss-Newton method. Since equation [13] is highly underdetermined even without undersampling, regularization is required. To this end, sensitivities $\vec{s}$ can be assumed to be smooth functions in space (25) and hence the number of coefficient to solve for is reduced. Addition variational penalties may be introduced to enforce piece-wise constant images (26). One implementation of a variational penalty is given by the total variation (TV) defined by the spatial gradient operator $V$:

$$TV(\vec{I}) = \sum_{i} |\nabla I_i| \tag{14}$$

Using the iteratively regularized Gauss-Newton (IRGN) method in conjunction with the TV constraint, equation [13] can be solved. Up to 10-fold undersampling of 2D radial imaging of the beating heart has been demonstrated outperforming previous parallel imaging implementations significantly (26).

Compressed sensing

Like in parallel imaging, compressed sensing (27,28) aims to reconstruct images from fewer samples than traditionally required by Shannon’s sampling theorem. While parallel imaging in general does not require
assumptions about the object to be imaged, compressed sensing makes explicit use of the fact that objects are sparse or can be sparsified using a suitable transform operation.

In general, three conditions have to be met for compressed sensing to work. First, the image or a transform thereof is required to be sparse i.e. only a subset of all entries in \( \mathbf{r} \) are expected to assume values different from zero (or noise). While this condition is fulfilled for angiographic data directly, dedicated transform operations are required to sparsify other objects. The Wavelet transform has been shown particularly efficient in compressing image data of various anatomies (27). Besides Wavelets, the temporal Fourier transform or other temporal transforms may be employed (29). The second condition of compressed sensing requires incoherence between the sparse representation of the object and the data points sampled in the sensing domain. This requirement can be well approached by using random sampling patterns in the sensing domain. In practice it is implemented by randomly sampling N/R profiles in k-space. The third requirement relates to the deployment of \textit{non-linear reconstruction} to select among all possible images \( \mathbf{r} \) the one that is most probable and closest to the desired object. Mathematically this minimization problem can be written as:

\[
\begin{align*}
\text{minimize} & \quad \| \phi \mathbf{r} \|_p + \lambda \cdot \text{TV}(\mathbf{r}) \\
\text{s.t.} & \quad \| \mathbf{d} - E \cdot \mathbf{r} \|_2^2 \leq \varepsilon
\end{align*}
\]

Here \( \phi \) represents the transform operator to compress the object and \( \| \|_p \) the \( L_p \)-norm. Most implementations choose \( p = 1 \) as it converts the reconstruction task into a convex optimization problem. Instead of considering the constrained optimization problem as posed in equation [15], the unconstrained problem in Lagrangian form may be used instead:

\[
\begin{align*}
\arg\min & \quad \| \mathbf{d} - E \cdot \mathbf{r} \|_2^2 + \lambda_1 \| \phi \mathbf{r} \|_1 + \lambda_2 \cdot \text{TV}(\mathbf{r})
\end{align*}
\]

Here parameters \( \lambda_1 \) and \( \lambda_2 \) determine the relative weights of sparsity and the total variation penalty relative to required data consistency.

\textit{Limits}

For most objects in cardiovascular imaging the sparsity condition can only be approximated. It has been demonstrated that reduction factors between 2 and 4 are well applicable in practice. However, at larger reduction factors image quality starts degrading significantly. There has been theoretical evidence that good image reconstruction quality is obtained if the number of k-space points to sample exceeds the number of sparse coefficients representing the object by a factor of 3-5 (27). Accordingly, compressed sensing performance is inherently dependent on the actual object to be imaged.

Another issue relates to the limited “randomness” that can be created with the small number of interleaves/phase-encodes available in practical cardiovascular imaging protocols. To judge the suitability of a chosen random sampling pattern, the side-lobe-to-peak ratio of the point-spread function has been proposed (27). Finally, the optimal choice of the parameter(s) in equations [15,16] depends on the application and often requires application-specific tuning.

\textit{Implementation}

A number of variants of the original Sparse MRI method (27) have been proposed over recent years. In order to address image degradation at higher reduction factors, the combination of compressed sensing with parallel imaging has been a logical extension. While in its most simple implementation compressed sensing and parallel imaging are applied sequentially (30), distributed compressed sensing integrates the signals from all coils into vector \( \mathbf{d} \) (31) and hence permits exploiting additional joint sparsity introduced by the differences in receive sensitivities among coil elements. A number of recent cardiovascular applications of distributed compressed sensing have demonstrated reduction factors of 8-10 for cine and contrast-enhanced perfusion imaging (31-32).

\textit{References}


Pruessmann KP. Encoding and reconstruction in parallel MRI. NMR in Biomedicine 2006;19(3):288-299.


