Non-Uniform Sampling Methods for MRI

Steven Troxler

Non-Uniform Sampling Methods for MRI

Steven Troxler

(1) Arizona State University, Tempe AZ 85287-1804 USA.
Steven.Troxler@asu.edu

1. Introduction

Simple Cartesian scans, which collect Fourier transform data on a uniformly-spaced grid in the frequency domain, are by far the most common in MRI. But non-Cartesian trajectories such as spirals and radial scans have become popular for their speed and for other benefits, like making motion-correction easier [12]. A major problem in such scans, however, is reconstructing from nonuniform data, which cannot be performed by a standard fast Fourier transform (FFT) as in the Cartesian case.

Here, we briefly describe the most common reconstruction methods and the non-uniform fast Fourier transform (NFFT) needed to complete the computations quickly. We then give an overview of several current methods for choosing a density compensation function (DCF) and suggest some possible improvements.

2. Reconstruction Methods

The most common method for nonuniform reconstruction in MRI is the Riemann approach, which approximates the integral difining the inverse (continuous) Fourier transform using a Riemann sum

$$f_{\boldsymbol{w}}(\boldsymbol{x}) = \sum_{j=1}^{J} w_j \hat{f}(\boldsymbol{\xi}_j) e^{2\pi i \boldsymbol{\xi}_j \cdot \boldsymbol{x}}, \tag{1}$$

where $\boldsymbol{x} \in \mathbb{Z}_N^d$ are the pixel locations and $\boldsymbol{\xi}_j$, $j = 1, ..., J$, are the frequency locations at which we measure the Fourier transform (we assume $J \geq N^d$). As the subscript $\boldsymbol{w}$ suggests, this approach requires finding appropriate weights w_j for each sample point in the reconstruction, a major theoretical problem. An alternative method, called implicit discretization (ID), assumes that the image itself is a sum of evenly spaced delta impulses at the pixel points of the final image, so that its Fourier transform is a finite-dimensional, harmonic trigonometric polynomial. We can then find a least-squares solution to the resulting system of equations

$$\hat{f}(\boldsymbol{\xi}_j) = \sum_{\boldsymbol{x} \in \mathbb{Z}_N^d} f(\boldsymbol{x}) e^{-2\pi i \boldsymbol{x} \cdot \boldsymbol{\xi}_j} \tag{2}$$

This model, which is known to have negligible error (the model error is the Gibb's error that would appear in a Cartesian reconstruction), has the important advantage of not depending on our arbitrary choice of weights.

These two approaches can be described in terms of matrix algebra as follows: Let G be a $J \times N^d$ matrix given by

$$\boldsymbol{G}_{j,\boldsymbol{x}} = e^{-2\pi i \boldsymbol{\xi}_j \cdot \boldsymbol{x}}.$$

Then we see immediately that

$$\boldsymbol{f}_{\boldsymbol{w}} = \boldsymbol{G}^* \boldsymbol{W} \tilde{\boldsymbol{f}}, \tag{3}$$

where $\boldsymbol{f}_{\boldsymbol{w}}$ is the $N^d \times 1$ vector, indexed by $\mathbb{Z}_N^d$, whose $\boldsymbol{x}$th entry is $f_{\boldsymbol{w}}(\boldsymbol{x})$, $\tilde{\boldsymbol{f}}$ is the $J \times 1$ vector of measurements whose jth entry is $\hat{f}(\boldsymbol{\xi}_j)$, and $\boldsymbol{W}$ is the $N^d \times N^d$ diagonal matrix with diagonal equal to $\boldsymbol{w}$. Once we have $\boldsymbol{w}$, whose determination is the main problem of interest, the remaining issue is one of computational complexity, since $\boldsymbol{G}^*$ is a very large unstructured matrix.

Fortunately, there is a fast method for computing products called the nonuniform fast Fourier transform (NFFT), based on the approximate factorization

$$\boldsymbol{G} \approx \boldsymbol{C}_\phi \boldsymbol{F} \boldsymbol{D}_\phi, \tag{4}$$

where $\boldsymbol{C}_\phi$ is a sparse, banded $N^d \times J$ matrix of convolution interpolation coefficients which depends on our choice of convolution kernel ϕ, $\boldsymbol{F}$ is the uniform $M^d \times M^d$ DFT matrix for some $M > N$, products of which are rapidly computed via the FFT, and $\boldsymbol{D}_\phi$ is an $M^d \times N^d$ modified diagonal deconvolution matrix, also depending on ϕ, whose extra rows are zero. Since it is easy to compute products with all three factors, this algorithm can be used to quickly approximate matrix products involving either $\boldsymbol{G}$ or $\boldsymbol{G}^*$. The theory of the NFFT, as applied to MRI, was first laid out in [11] and [8]. Later, [4] found bounds on the errors for Gaussian interpolation, and [23] and [5] gave general estimates and gave sharper bounds for Gaussian kernels. The most complete discussion of NFFT theory is given in [16], while [15] presents many of the proofs. Practical considerations like computational load and numerical stability were addressed in [3] and [17], while [1] and [6] presented two methods of efficient interpolation using Kaiser-Bessel and Gaussian kernels.

In matrix form, the ID problem attempts to find a least-squares solution to the problem

$$\tilde{\boldsymbol{f}} = \boldsymbol{G} \boldsymbol{f}.$$

The ordinary least squares solution $\boldsymbol{f}_{OLS}$ satisfies the normal equation

$$\boldsymbol{G}^*\boldsymbol{G}\boldsymbol{f}_{OLS} = \boldsymbol{G}^*\hat{\boldsymbol{f}}. \tag{5}$$

Although the matrix $\boldsymbol{G}^*\boldsymbol{G}$ is far too large to invert, it is symmetric, so we may use iterative methods like conjugate gradients to find the solution. The resulting solution typically has excellent quality, but convergence is often slow, making ordinary least squares expensive.

Conjugate gradients converges fastest when $\boldsymbol{G}^*\boldsymbol{G}$ is close to the identity, which is unfortunately rarely the case unless the sampling density is reasonably close to unity. In order to improve the convergence of conjugate gradients, we introduce the weighted least squares problem, which finds the least squares solution to

$$\boldsymbol{W}^{1/2}\tilde{\boldsymbol{f}} = \boldsymbol{W}^{1/2}\boldsymbol{G}\boldsymbol{f}$$

by solving the normal equations

$$\boldsymbol{G}^*\boldsymbol{W}bvecG\boldsymbol{f}_{WLS} = \boldsymbol{G}^*\boldsymbol{W}\hat{\boldsymbol{f}},$$

where $\boldsymbol{W}$ is the modified diagonal density compensation matrix used for the Riemann method. We expect an improvement in convergence because we know that the Riemann method gives much better results with $\boldsymbol{W}$ than without, which means $\boldsymbol{G}^*\boldsymbol{W}\boldsymbol{G}$ approximate the identity much better than $\boldsymbol{G}^*\boldsymbol{G}$. From a signal processing perspective, this has the additional benefit that we weight errors heavier at highly isolated observations of the Fourier transform, which heuristically contain more information about the objective function than less isolated observations.

For either method, then, determining an appropriate value of $\boldsymbol{w}$ is important. It is more essential in the Riemann approach, where a poor choice of $\boldsymbol{w}$ will lead to useless results. The ID method is known to converge quite well after only a few iterations, even when a very rough approximation to $\boldsymbol{w}$ is used, but the better $\boldsymbol{w}$, the fewer iterations are required. It is worth noting that the first iteration, which always moves in the direction of the residual, is actually just a rescaling of the Riemann solution.

3. Determination of an optimal DCF

3.1 Algebraic and Analytic Approaches

Since the equation

$$\tilde{\boldsymbol{f}} = \boldsymbol{G}\boldsymbol{f},$$

used directly in the CG reconstruction, provides an accurate mathematical model for the measurements which does not depend on the choice of a sampling density $\boldsymbol{w}$, the clearest method of evaluating a DCF $\boldsymbol{w}$ is to require that

$$\tilde{\boldsymbol{f}} \approx \boldsymbol{G}\boldsymbol{f}_{\boldsymbol{w}},$$

where

$$\boldsymbol{f}_{\boldsymbol{w}} = \boldsymbol{G}^*\boldsymbol{W}\tilde{\boldsymbol{f}}.$$

This is the same as requiring that

$$\boldsymbol{G}^*\boldsymbol{W}$$

approximate the pseudoinverse $(\boldsymbol{G}^*\boldsymbol{G})^{-1}\boldsymbol{G}^*$ of $\boldsymbol{G}$.

The weighted conjugate gradient method described at the end of the previous section, whose first iteration performs best when the matrix is as close to the identity as possible, leads to a similar but slightly simpler condition, that

$$\boldsymbol{G}^*\boldsymbol{W}\boldsymbol{G} \approx \boldsymbol{I},$$

in the sense that the eigenvalues of $\boldsymbol{G}^*\boldsymbol{W}\boldsymbol{G}$ be as closely clustered as possible. Several techniques have been proposed to use these conditions to find an algebraically ideal DCF via use of a singular value decomposition or some similar approach [22], [20]. These methods, however, tend to have high computational complexity. This is a problem if the same trajectory is not always used, as is the case in many MRI applications in which iterative reconstruction is used to compensate for field inhomogeneities and other measurement imperfections. Moreover, although such algebraic methods generally give workable results, other methods which take analytic considerations into account often perform better empirically. Possible reasons why the theoretically optimal algebraic solutions fail to give the best results include numerical instability and ill-conditioning. In some cases, the algebraic approaches even result in DCF's with negative weights at some points. This contradicts our intuition, and empirical studies indicate that such DCF's tend to perform relatively poorly.

The simplest analytic approaches to determining $\boldsymbol{w}$ are based on the fact that the goal of the Riemann method is to approximate a Riemann sum. For radial and analytic spiral trajectories, which may be smoothly parameterized, methods have been proposed which use the Jacobian of a change-of-coordinates [10], [7]. These techniques give very good results for certain spirals, although for radial trajectories they tend to underweight points near the center. An alternative analytic method, which works for arbitrary nonuniform sampling schemes, is to construct a Voronoi diagram, which partitions the sampled part of frequency space into polygons about each sample point, and weight the samples according to the area or volume of those polygons [19]. This typically results in a good image for radial trajectories. With other trajectories, the results are generally inferior to alternative point-spread-function methods, although it was demonstrated in [9] that performing a few iterations of the weighted conjugate gradient method using Voronoi weights produces an excellent image.

3.2 The Point Spread Function

Most of the best-performing methods for determining the DCF when the trajectory is anything other than an analytic spiral are based on analysis of the *point-spread-function* (PSF). The PSF is defined as the inverse Fourier transform $\check{w}$ of the DCF, where we view the DCF as a distribution on $\mathbb{R}^d$ defined by $w := \sum_j w_j \delta_{\boldsymbol{\xi}_j}$. The PSF $\check{w}$ is then given by

$$\check{w}(\boldsymbol{x}) = \sum_{j=1}^{J} w_j e^{2\pi i \boldsymbol{x}\cdot\boldsymbol{\xi}_j}. \tag{6}$$

This is what the algorithm would produce if the true object were a delta impulse located at zero. The observed data would be a vector of all ones, so the reconstruction would

be the result of applying G^* to $\boldsymbol{w}$ itself, i.e., the function defined by (6).

If f is a more general object, it follows from the convolution theorem (for distributions) that the reconstructed function $f_{\boldsymbol{w}}$ will be equal to the convolution $f * \check{w}$ of the actual object f with the PSF. The more closely the PSF resembles a delta impulse, the better the reconstruction.

It is important to note that, since the PSF is a (nonharmonic) trigonometric polynomial, it will not decay at infinity. Clearly, then, the best that we can hope for is that $\check{w}$ will resemble a delta impulse in some compact neighborhood of the origin. Recall that, by accepting the ID model as having negligible error, we are assuming that f is a finite-dimensional vector defined on $\mathbb{Z}_N^d$ which we associate with a distribution supported on $\mathbb{Z}_N^d$ for notational convenience when dealing with convolutions. Since the terms $\check{w}(\boldsymbol{z})f(\boldsymbol{x}-\boldsymbol{z})$ defining

$$f_{\boldsymbol{w}}(\boldsymbol{x}) = \check{w} * f(\boldsymbol{x}) = \sum_{\boldsymbol{z}\in\mathbb{Z}^d} \check{w}(\boldsymbol{z}) f(\boldsymbol{x}-\boldsymbol{z}) \qquad (7)$$

are nonzero only if $(\boldsymbol{x}-\boldsymbol{z}) \in \mathbb{Z}_N^d$, and we only want to find the reconstruction $f_{\boldsymbol{w}}(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{Z}_N^d$, we conclude that the only values of $\boldsymbol{z}$ for which $\check{w}(z)$ matter are $\boldsymbol{z} \in \mathbb{Z}_{2N}^d$. It is also worth noting that not all points $\boldsymbol{z} \in \mathbb{Z}_{2N}^d$ appear equally often in the convolution defining $\boldsymbol{f}_W$. The origin will appear in one term of every sum, whereas values of $\boldsymbol{z}$ near the edge of $\mathbb{Z}_{2N}^d$ will appear only occasionally.

For notational convenience, let A be the field of view $[-N/2, N/2]^d$ and let B be the region of optimization $[-N, N]^d$. PSF optimization techniques find some computational way of minimizing the error

$$E = \check{w} - \delta$$

over this region of optimization B.

By carefully looking at (7), we see that the frequency with which a PSF error at $\boldsymbol{x}$ actually occurs in the final image is proportional to $p = \chi A * \chi A$. Since errors are unavoidable and we would like to minimize the important errors, we introduce the weighted error, given by

$$E = p\check{w} - \delta, \qquad (8)$$

where p is called the *error profile*. This error can be expressed in the Fourier domain as

$$\hat{E} = \hat{p} * w - 1 = \hat{\chi}_A^2 * w - 1. \qquad (9)$$

Our goal is to minimize these errors, thereby minimizing the error in the final reconstruction $f_{\boldsymbol{w}} = \check{w} * f$.

Although this optimal kernel p was suggest only recently in [13], convolution techniques for minimizing the Fourier domain PSF error $\hat{E}$ have been used for some time. In one of the early gridding papers, Jackson et. al. proposed taking w to be equal to

$$w_1 = \frac{w_0}{\phi * w_0}, \qquad (10)$$

where w_0 is a DCF of unity (in distributional form) and ϕ is the gridding kernel [8]. This method predates PSF techniques, and was instead motivated by the intuitive idea that $\phi * w_0$ would give a reasonable, estimate of the sampling density. Later researchers noted, however, that if we ϕ with $\hat{p}$, we would expect this ratio correction to make $w_1 * \hat{p}$ closer to unity than $w_0 * \hat{p}$ regardless of the initial density w_0 [14]. An iterative technique, based on this observation, starts with a constant DCF w_0 and takes

$$w_{i+1} = \frac{w_i}{\hat{p} * w_i}. \qquad (11)$$

Since $\hat{p}$ can be effectivly truncated, each iteration can be computed quickly, particularly if an efficient sorting algorithm is used to avoid time-consuming searches for the nonzero terms $\hat{p}(\boldsymbol{\xi_k} - \boldsymbol{\xi}_j)w(\boldsymbol{\xi}_j)$ in the convolution [13].

Another iterative algorithm, aimed at the same goal of achieving $\hat{E} = 1$, uses an additive correction instead of a ratio-based correction, taking

$$w_i = w_{i-1} + \sigma(1 - \hat{p} * w_{i-1}),$$

where $\sigma \in (0, 1)$ is a parameter controlling convergence [18]. Taking σ close to 1 may result in the fastest convergence, but could also lead to instability and a failure to converge.

The advantage of these iterative techniques is that they are conceptually simple, computationally fast, and empirically give results as good as any current methods when the correct error profile p is used and the number of iterations is determined experimentally. A disadvantage is that, although they work conceptually and empirically, there is no theoretical basis for claiming that they converge to the optimal solution, and, in fact, experimental evidence indicates that the mean square error in $f_{\boldsymbol{w}}$ can actually rises if the algorithm is allowed to run too long. This may be due to numerical instability, or to a failure of the mathematical algorithm itself to technically converge.

An algebraic method of optimizing the PSF, which has more theoretical grounding than convolution-based methods, attempts to directly solve the inverse problem

$$\boldsymbol{G}\boldsymbol{G}^*\boldsymbol{w} = \boldsymbol{u},$$

where $\boldsymbol{u}$ is a vector of all ones. The direct solution to this problem via conjugate gradients using the NFFT was proposed in [21], but as with the algebraic solutions for $\boldsymbol{w}$ based on the least-squares method, this can result in a $\boldsymbol{w}$ with wide variations and sometimes even negative entries, which does not match our expectation for a density and empirically gives inferior results. A regularization of this method was proposed in [2] which instead attempts to solve

$$(\boldsymbol{G}\boldsymbol{G}^* + \sigma^2\boldsymbol{I})\boldsymbol{w} = \boldsymbol{u} + \sigma^2\boldsymbol{w}_1,$$

where $\boldsymbol{w}_1$ is an initial nonnegative and smoothly varying estimate of the density, say, Jackson's weight (10), or, more optimally, the result of one or two iterations of (11). This second approach ensures that the solution behaves as we would expect a DCF to behave, and empirically gives better results than the unregularized method. The algorithm given in [2] also incorporates Jacobi preconditioning to speed convergence of the conjugate gradient iterations.

Knowing that Pipe and Johnson's error profile p provides an optimal weight on errors in the point-spread function,

it might be preferable to modify the approach in [2] in two ways. The first is that, since we need to minimize PSF errors over twice the support of f, we replace the NDFT matrix $\boldsymbol{G}$ with $\boldsymbol{G}_1$, where the uniform grid has twice the radius of that used by $\boldsymbol{G}$. This avoids the risk that we might ignore PSF errors which, according to the convolution defining $\boldsymbol{f_w}$, appear in the Riemann reconstruction. The second is replacing $\boldsymbol{G}_1\boldsymbol{G}_1^*$, which treats all PSF errors as equally important, with $\boldsymbol{G}_1\boldsymbol{P}\boldsymbol{G}_1^*$, where $\boldsymbol{P}$ contains the values of the optimal error profile p. To the author's knowledge, this has never been tried, but in light of experiments reported by [13] indicating that the approaches taken in [2] and [13] both yield the some of the best results of methods proposed to date for arbitrary trajectories, combining their methods might produce the best results seen yet.

4. Acknowledgments

This work was part of an undergraduate research project under the supervision of Dr. Svetlana Roudenko. The author would also like to thank Ken Johnson and Dr. Jim Pipe for providing graphical illustrations and for helpful discussions of this content, as well as Dr. Doug Cochran for his assistance and input. The project was partially supported by NSF-DUE # 0633033 and NSF-DMS # 0652853.

References:

[1] Philip J. Beatty, Dwight G. Nishimura, and John M. Pauly. Rapid gridding reconstruction with a minimal oversampling ratio. *IEEE Trans. Med. Imag.*, 24(6):799–808, June 2005.

[2] Mark Bydder, Alexey A. Samsonov, and Jiang Du. Evaluation of optimal density weighting for regridding. *Magnetic Resonance Imaging*, 25:695–702, 2007.

[3] S. Dunis and D. Potts. Time and memory requirements of the nonequispaced fft. *Sampling Theory in Signal and Image Processing*, 7:77–100, 2008.

[4] A. Dutt and V. Rokhlin. Fast fourier transforms for nonequispaced data. *SIAM Journal of Scientific Computing*, 14(6):1368–1393, 1993.

[5] B. Elbel and G. Steidl. Fast fourier transform for nonequispaced data. In C. K. Chui and L. L. Schumaker, editors, *Approximation Theory IX*. Vanderbilt University Press, Nashville, 1998.

[6] Leslie Greengard and June-Yub Lee. Accelerating the nonuniform fast fourier transform. *SIAM Review*, 46(3):443–454, 2004.

[7] R.D Hodge, R.K.S. Kwan, and G.B. Pike. Density compensation functions for spiral MRI. *Magnetic Resonance in Medicine*, 38:117–128, 1997.

[8] J. Jackson, C. Meyer, D. Nishimura, and A. Macovski. Selection of a convolution function for fourier enversion using gridding. *IEEE Trans. Med. Imag.*, 10(3):473–478, Sep. 1991.

[9] Tobias Knopp, Stefan Kunis, and Daniel Potts. A note on the iterative mri reconstruction from nonuniform k-space data. *International Journal of Biomedical Imaging*, 2007.

[10] C. Meyer, B. S. Hu, D. Nishimura, and A. Macovski. Fast spiral coronary artery imaging. *Magnetic Resonance in Medicine*, 28:202–213, 1992.

[11] J. O'sullivan. A fast sinc function gridding algorithm for fourier inversion in computerized tomography. *IEEE Transactions on Medical Imaging*, MI-4:200–207, 1985.

[12] James G. Pipe.

[13] J.G. Pipe and Kenneth Johnson. Convolution kernel design and efficient algorithm for sampling density correction. *Preprint*, 2008.

[14] J.G. Pipe and P. Menon. Sampling density compensation in MRI: Rationale and an iterative numerical solution. *Magnetic Resonance in Medicine*, 41:799–808, June 2005.

[15] D. Potts. *Schnelle Fourier-Transformationen für nichtäquidistante Daten und Anwendungen*. Habilitation, Universität zu Lübeck, 2003.

[16] D. Potts, G. Steidl, and M. Tasche. Fast fourier transforms for nonequispaced data: A tutorial. In J. J. Benedetto and P. J. S. G. Ferreira, editors, *Modern Sampling Theory: Mathematics and Applications*, pages 247–270. Birkhäuser, Boston, 2001.

[17] D. Potts and M Tasche. Numerical stability of nonequispaced fast fourier transforms. *Journal of Computational Applied Mathematics*, 222:655–674, 2008.

[18] Y. Qian, J. Lin, and D. Jin. Reconstruction of mr images from data acquired on an arbitrary k-space trajectory using the same-image weight. *Magnetic Resonance in Medicine*, 48:306–311, 2002.

[19] V. Rasche, R. Proksa, R. Sinkus, P. Bornert, and H. Eggers. Resampling of data between arbitrary grids using convolution interpolation. *IEEE Trans. Med. Imag.*, 18:427–434, 1999.

[20] D. Rosenfeld. An optimal and efficient new gridding algorithm using singular value decomposition. *Magnetic Resonance in Medicine*, 40:14–23, 1998.

[21] A.A. Samsonov, E.G. Kholmovski, and C.R. Johnson. Determination of the sampling density compensation function using a point spread function modeling approach and gridding approximation. volume 11, 2003.

[22] Hossein Sedarat and Dwight G. Nishimura. On the optimality of the gridding reconstruction algorithm. *IEEE Transactions on Medical Imaging*, 19(4):306–317, 2000.

[23] G. Steidl. A note on fast fourier transforms for nonequispaced grids. *Advanced Computational Mathematics*, 9:337–353, 1998.