Compressed Sensing Reconstruction for Whole-Heart Imaging with 3D Radial Trajectories: A Graphics Processing Unit Implementation

Seunghoon Nam,1,2 Mehmet Akçakaya,1 Tamer Basha,1 Christian Stehning,3 Warren J. Manning,1,4 Vahid Tarokh,2 and Reza Nezafat1*

A disadvantage of three-dimensional (3D) isotropic acquisition in whole-heart coronary MRI is the prolonged data acquisition time. Isotropic 3D radial trajectories allow undersampling of k-space data in all three spatial dimensions, enabling accelerated acquisition of the volumetric data. Compressed sensing (CS) reconstruction can provide further acceleration in the acquisition by removing the incoherent artifacts due to undersampling and improving the image quality. However, the heavy computational overhead of the CS reconstruction has been a limiting factor for its application. In this article, a parallelized implementation of an iterative CS reconstruction method for 3D radial acquisitions using a commercial graphics processing unit is presented. The execution time of the graphics processing unit-implemented CS reconstruction was compared with that of the C++ implementation, and the efficacy of the undersampled 3D radial acquisition with CS reconstruction was investigated in both phantom and whole-heart coronary data sets. Subsequently, the efficacy of CS in supressing streaking artifacts in 3D whole-heart coronary MRI with 3D radial imaging and its convergence properties were studied. The CS reconstruction provides improved image quality (in terms of vessel sharpness and suppression of noise-like artifacts) compared with the conventional 3D gridding algorithm, and the graphics processing unit implementation greatly reduces the execution time of CS reconstruction yielding 34–54 times speed-up compared with C++ implementation. Magn Reson Med 000:000–000, 2012. © 2012 Wiley Periodicals, Inc.

Key words: compressed sensing; accelerated imaging; GPU implementation; 3D radial acquisition; cardiac MR

Cardiac MR data are typically acquired using multiple two-dimensional (2D) slices. Imaging using a single large three-dimensional (3D) slab covering the whole heart from the base to the apex can significantly simplify image prescription. Whole-heart coronary MRI, analogous to coronary multidetector computed tomography (CT), has replaced multiple small-slab targeted acquisitions for the individual coronary arteries (1–3). A single breath-hold accelerated 3D cine scan has been previously investigated for the evaluation of cardiac function (4–6). Free-breathing, 3D late gadolinium enhancement imaging has been used to identify fibrosis/scar with improved spatial resolution or coverage (7,8). Recently, 3D perfusion has also been applied to improve the spatial coverage (9,10). The advantages of a 3D acquisition include superior spatial resolution, especially through-plane, ease of image prescription, superior signal-to-noise ratio, and easy reformatting of the image in any desired plane. However, one major disadvantage of 3D imaging is the long data acquisition time. For coronary MRI, a longer scan time usually makes the scan more susceptible to respiratory motion. For late gadolinium enhancement, it results in imaging artifacts due to changes in optimal inversion time as the contrast washes out. For cine and perfusion, it typically results in lower temporal or spatial resolution. Therefore, methods to reduce scan time in 3D imaging could significantly improve the clinical utilization of 3D cardiac MR.

3D whole-heart data are commonly acquired using Cartesian k-space sampling; however, non-Cartesian sampling schemes, e.g., radial or spiral, have better data acquisition efficiency (11,12). Both 3D stack-of-radials and 3D radial (kooshball) acquisitions with isotropic spatial resolution have been previously used in 3D cardiac MR (11,13,14). In these sampling schemes, a Nyquist sampling rate is not necessary because undersampling does not yield distinct fold-over artifacts; instead, it typically results in streaking artifacts. This allows high undersampling rates with less pronounced imaging artifacts compared with Cartesian acquisitions at the same sampling density. These potential benefits have been previously exploited to achieve whole-heart coronary MRI with isotropic spatial resolution (15). It has also been extensively investigated to improve the dynamic imaging such as phase contrast, MR angiography, and cine imaging (16–18).

For single-phase anatomical imaging such as coronary MRI, a gridding algorithm is commonly used in the reconstruction of 3D radial acquisitions (19). Although the gridding algorithm can efficiently reconstruct data acquired using a 3D radial trajectory, its performance deteriorates significantly for highly undersampled data due to significant undersampling of outer k-space regions (20). Parallel imaging methods including...
sensitivity encoding (SENSE) (21) and generalized autocalibrating partially parallel acquisitions (GRAPPA) (22) have been previously applied for 2D radial acquisitions to reduce the streaking artifacts (23,24). Recently, compressed sensing (CS) has been applied to remove the streaking artifacts for 2D radial acquisitions (20,25). In this approach, additional constraints based on image properties are used to improve the image reconstruction. The CS reconstruction techniques are usually implemented with iterative procedures that solve the optimization problem with relatively computationally cheap matrix-vector multiplications (26). The CS reconstruction for 3D radial acquisition has been recently demonstrated for imaging of the hand using 512 radial profiles on a matrix size of $128^3$ (27). The computational overhead of the iterative CS reconstruction increases as the size of the 3D $k$-space increases, resulting in prolonged reconstruction time.

Recently, graphics processing units (GPUs) have become available for high computation-intensive applications. Hardware manufacturers provide parallel computing architectures (such as Compute Unified Device Architecture (CUDA) and FireStream) that enable researchers to implement GPU programs using high-level programming languages without knowledge of the GPU hardware structure. Recent studies have shown that GPU-accelerated reconstructions can be used to achieve reduced and low-latency reconstruction times for various MR applications (28–30). GPU-accelerated reconstructions for 2D radial acquisitions were demonstrated with ~6–32 times speed-up in reconstruction time compared with central processing unit implementations (28,31). GPU implementations have been shown to greatly accelerate CS reconstructions of 3D non-Cartesian trajectories such as stack-of-radials and stack-of-spirals (30,32), where the $k$-space samples are equidistantly spaced along one $k$-space dimension. However, GPU implementation for a true 3D non-Cartesian trajectory such as 3D radial sampling is not straightforward from those for stack-of-radials/spirals trajectories and has not been previously reported because of the large size of the 3D sampling data and GPU hardware limitations.

In this article, we propose to implement and evaluate the performance of a GPU-accelerated reconstruction for 3D radial reconstruction using the latest commercially available GPU hardware. Subsequently, we will investigate the efficacy of a 3D radial acquisition with CS reconstruction for whole-heart coronary MRI.

**MATERIALS AND METHODS**

All phantom and volunteer data were obtained using 1.5 T Achieva magnet (Philips Healthcare, Best, The Netherlands) with a five-channel phased-array coil. The acquired MR data were transferred to a stand-alone computer, and the image reconstruction was performed offline. All in vivo studies were approved by our institutional review board, and all subjects provided consent before participation in the study.

3D Radial Acquisition and Reconstruction

In this section, we will review and present the formulation for 3D radial image acquisition and reconstruction using CS. The 3D radial sampling trajectory consists of $N_i$ interleaves, where each interleaved has $N_p$ projection lines with $N_i$ sample points (15). Each interleaves is the rotated version of the first interleaved around the $k_z$-axis. The isotropy (or uniformity) of the sampling point distribution can be quantified by the standard deviation of the distance between adjacent sampling points on the $k$-space and is kept at <10% of the mean distance when the total number of projections $N_p \times N_i$ is between 100 and 10,000 (33). The sampling density of a 3D radial acquisition is defined as the ratio of the total number of $k$-space samples of the 3D radial acquisition over that of a Nyquist-sampled 3D Cartesian acquisition with the same resolution and the same field-of-view (FOV). A gridding algorithm (19) is commonly used to reconstruct 3D radial data. In the conventional gridding algorithm, each data point is compensated for its nonuniform sampling density by the density compensation function, which is calculated based on the sampling trajectory (34–36). The data point is convolved with a gridding kernel and resampled onto the Cartesian grid. The regridded $k$-space samples are then inverse Fourier transformed to obtain the desired image. Deapodization is performed after the inverse Fourier transform by dividing the image by the apodization function, which is given by the Fourier transform of the gridding kernel function (19). As all of the operations are linear, this procedure can be expressed in a matrix-vector format:

$$\hat{x} = D^* S^{-1} P^* F D x,$$  \[1\]

where $\hat{x}$ is the reconstructed image, $y$ is the measured 3D radial $k$-space data, $F$ is a diagonal matrix performing the density compensation, $S$ denotes the convolution matrix for the gridding operator, $F^*$ denotes the inverse fast Fourier transform (IFFT), and $D$ is a diagonal matrix performing the deapodization. We note that all the voxels of the 3D image are represented in a single column vector $x$ for mathematical convenience.

As an alternative approach, the acquired $k$-space signals can be formulated in an encoding matrix format as $y = A x$, where $A$ denotes the encoding matrix and $x$ denotes the actual image. $A$ can be considered as taking the reverse steps of the conventional gridding algorithm without the density compensation:

$$y = A x = S^* F D x,$$  \[2\]

where $D$ is a diagonal matrix performing the deapodization, $F$ denotes the fast Fourier transform (FFT) matrix, and $S^*$ denotes the convolution matrix from Cartesian to radial sample points. $x$ is deapodized and Fourier transformed into the $k$-space, and then the Cartesian $k$-space samples are regridded onto the 3D radial sample points using the gridding kernel. Unlike the conventional gridding algorithm, the density compensation is not required before the gridding because the density of the Cartesian grid is uniform (37). Equation 2 holds regardless of the Nyquist criterion, but the encoding matrix is not invertible for undersampled data, as Eq. 2 is underdetermined and there are multiple solutions that will satisfy the system equation. CS reconstruction uses the sparsity of the
where \( \lambda \) is a regularization parameter that determines the tradeoff between the data consistency and the sparsity level of the image. \( \| \cdot \|_p \) denotes the \( l_p \) norm of the vector, which is defined by \( \| s \|_p = (\sum_{i=1}^{n} |s_i|^p)^{1/p} \), and \( \Psi \) is a sparsifying transform matrix such as a wavelet transform or total variation operator.

To solve Eq. 3, we adopt an iterative method that alternately enforces the data consistency and sparsity of the image estimate at each iteration (38). The image update at the \((t+1)\)th iteration is given by solving the following two subproblems:

\[
\mathbf{u}^{t+1} = \mathbf{x}^{t+1} = \arg\min_{\mathbf{x}} \frac{1}{2} \| \mathbf{x} - \mathbf{u}^t \|_2^2 + \lambda \| \Psi \mathbf{x} \|_1.
\]

Equation 4 is called the data consistency step as the solution tends to decrease the \( l_2 \)-norm error between the measured data and the \( k \)-space of the image estimate. For any unitary sparsifying transform \( \Psi \), Eq. 5 can be reexpressed with respect to the transform domain vector \( \mathbf{z}^t = \Psi \mathbf{x}^t \) as

\[
\mathbf{z}^{t+1} = \arg\min_{\mathbf{z}} \frac{1}{2} \| \mathbf{z} - \mathbf{u}^t \|_2^2 + \lambda \| \mathbf{z} \|_1
\]

Equation 6 can be solved by a simple coefficient-wise thresholding function as follows:

\[
z_i^{t+1} = \begin{cases} 
w_i^t & \text{if } w_i^t \max \left( |w_i^t| - \frac{\lambda}{\alpha_t}, 0 \right) > 0 
\end{cases}
\]

where \( z_i^{t+1} \) and \( w_i^t \) denote the \( i \)th coefficient of the transform domain vector \( \mathbf{z}^t \) and \( \mathbf{w}^t = \Psi \mathbf{u}^t \) of the solution of the first subproblem in Eq. 4, respectively. The second subproblem is called the thresholding step. For \( \alpha_t \), we adopt the step size from (39), where \( \alpha_t \) is determined so that \( z_i^t \) approximates the Hessian of the data consistency term \( \nabla^2 \| \mathbf{Ax}^t - \mathbf{y} \|_2^2 \) as below:

\[
\alpha_t = \frac{\| \mathbf{Ax}^t - \mathbf{Ax}^{t-1} \|^2_2}{\| \mathbf{x}^t - \mathbf{x}^{t-1} \|^2_2}
\]

The overall iterative reconstruction procedure is summarized in Fig. 1. The reconstruction starts from an initial image estimate, which in our experiments was chosen to be the gridding reconstruction. The image is deapodized, Fourier transformed into \( k \)-space, and then regridded onto the radial sample points. The estimated radial samples are subtracted from the actual measurement data, convolved onto the Cartesian \( k \)-space grid, inverse Fourier transformed, and an image estimate is obtained after deapodization. The image estimate is combined with the intermediate image from the previous iteration. The combined image is then thresholded in the transform domain to produce a new image estimate, and the intermediate image is updated. The final image estimate is obtained as the result of the iterative procedures.

**GPU-Accelerated CS Reconstruction for 3D Radial Trajectory**

The computational burden of a 3D radial trajectory with CS reconstruction is a major drawback, and its feasibility has not been studied in the literature. In this section, we will present our implementation of a GPU-based reconstruction of a 3D radial acquisition that allows us to further explore the utility of this reconstruction for 3D whole-heart cardiac MR.

The reconstruction algorithm in this article was implemented using an NVIDIA (Santa Clara, CA) graphics card and parallel computing architecture, CUDA. The CUDA program consists of two parts: host code that is executed on the central processing unit and device code that is executed on the GPU. The code that has little or no parallelism in computation is written in host code using ANSI C language, and the code that has a large amount of parallelism in computation is written in device code.
using a slightly modified C-like language. The functions written in the device code are called kernels, and each kernel generates a large number of threads as a result of data parallelism once the kernel is invoked. All the threads generated by a kernel invocation are called a grid. The threads in a grid are grouped into blocks, which are the basic allocation unit for the execution resources on the hardware. All the blocks in the same grid must have the same number of threads.

The gridding and regridding operations are the most computationally intensive part of the iterative CS reconstruction. As the width of the convolution window is much smaller than the size of the entire k-space, the gridding/regridding can be performed in a parallel manner for each measured radial point and are well suited for CUDA implementation. In this article, we assigned each 3D radial data point to one CUDA thread. Each projection line corresponds to one block, which consists of \( N_s \) threads. The grid has a 2D block structure \((N_p, N_i)\) to represent all the projection lines and interleaves of the 3D radial trajectory. Figure 2 shows a simplified example of a grid hierarchy and thread assignment of our implementation, where we have eight sample points in one projection, three projection lines per interleaver, and two interleaves. In the gridding operation, contributions from adjacent radial samples are accumulated to a Cartesian sample point as illustrated in Fig. 3a, which results in cumulative memory writes during the parallelized execution. The cumulative memory writes can produce incorrect results if more than two threads try to access the same memory simultaneously. This is prevented using CUDA’s atomic operation, which is capable of reading and writing on a memory address without interruption by other threads, allowing concurrent threads to correctly perform the required memory access.

The performance of atomic operation in CUDA is greatly improved on recent “Fermi”-based GPUs offered by NVIDIA, which provide up to 20 times faster atomic operation compared with their former generation GPUs (40).

Besides the gridding/regridding operations, most of the CS reconstruction procedures including FFT/IFFT, wavelet/inverse-wavelet transforms, deapodization, and thresholding were parallelized and written in device code. cuFFT and cuBLAS packages were used for FFT/IFFT and other arithmetic operations. Because of the limited global memory size of current GPU hardware, we
could not parallelize the reconstruction for the multiple coil elements. The reconstruction was performed sequentially for each coil, and the final reconstructed image was obtained as the root-sum-square of the individual coil images. The CS reconstruction was also implemented in standard C++ environment for the comparison of the reconstruction time. The FFTw package (41) was used for FFT/IFFT operations. The GPU and C++ implementations of the CS reconstruction were based on single precision floating point arithmetic, and they were executed on a PC with Intel (Santa Clara, CA) Core2 Quad Q9400 central processing unit (2.66 GHz), 8.0 GB memory, and NVIDIA GeForce GTX 480 Graphics card (480 cores, 1.5 GB memory) running on a 64-bit Windows 7 operating system.

Phantom Study

Two experiments were performed for the phantom study. The first experiment is to demonstrate the capability of improving the reconstruction quality of 3D radial acquisitions using the CS reconstruction method. The second experiment is to investigate the convergence properties of the CS reconstruction method over different numbers of iterations.

For the first experiment, a high-resolution phantom was scanned with a steady-state free precession sequence using 3D radial trajectories with \( N_p = 344 \) and \( N_r = 10 \) for six different sampling densities 7.5, 10, 20, 30, 40, and 100%, which correspond to the number of projections per interleaf and 100%, which correspond to the number of projections per interleaf. The Kaiser-Bessel function with window size 4.0 was used for the convolution parameter, from 0.01 \( ||A^\gamma y||_\infty \) to 0.1 \( ||A^\gamma y||_\infty \) as in (39,43) and manually selected it to get the best image qualities; \( \lambda = 0.05 ||A^\gamma y||_\infty \) gave satisfactory results for most of the cases with both sparsity regularizations. The density compensation function for the gridding algorithm was calculated using the iterative procedure proposed in Ref. 34. For both the CS reconstruction and the gridding algorithm, a Kaiser-Bessel function with window size 4.0 was used for the convolution kernel (44).

For the second experiment on convergence properties of the reconstruction algorithm, iterative CS reconstructions with both image and wavelet domain regularizations were performed on the phantom data set with 7.5% sampling density, and the intermediate images were stored for different numbers of iterations.

Whole-Heart Coronary MRI

Whole-heart coronary MR images were acquired on nine healthy volunteers (two males, 26 ± 11 years). 3D free-breathing ECG-triggered steady-state free precession sequences were used for imaging the heart with 3D radial trajectories. A respiratory navigator with 7 mm gating window was used for gating and tracking the respiratory motion (43). The \( k \)-space data acquired within the gating window were accepted, and the \( k \)-space data acquired outside the gating window were rejected and reacquired until acquired within the gating window. Within the 7-mm gating window, the position of the imaging volume was adaptively adjusted using a tracking factor of 0.6. The data sets were acquired with \( N_v = 392 \) and \( N_t = 10 \) for various sampling densities: two data sets with 6.8, 12.1, 24.2, and 36.3%; seven data sets with 10, 20, 30, and 40%. The scan parameters were as follows: repetition time/echo time/\( \alpha \) = 3.9/1.9/60°, FOV = 256 × 256 × 256 mm³, and spatial resolution = 1.3 × 1.3 × 1.3 mm³. The nominal scan time for the data set with sampling density of 40% was reported to be 5 min 13 s assuming 100% navigator efficiency. For one volunteer, an additional scan with spatial resolution of 1.0 × 1.0 × 1.0 mm³ and sampling density of 40% was acquired. The acquired 3D radial data were reconstructed by the three reconstruction methods (i.e., gridding, CS with image domain regularization, and CS with wavelet domain regularization), and the reconstructed image quality was compared. We used \( \lambda = 0.05 ||A^\gamma y||_\infty \) as the regularization parameter for both sparsity regularizations. The density compensation function for the gridding algorithm was calculated by the same method used in the phantom study, and the Kaiser-Bessel function with window size 4.0 was used for the convolution kernel.

The empirical convergence properties of the CS reconstructions were also observed similar to the phantom study. The vessel sharpness and the vessel length of the right coronary artery (RCA) were measured using Soap-Bubble software (46) for quantitative assessment of the quality of the CS reconstruction method. The vessel sharpness is measured using a Deriche algorithm (47) as previously described (48), where vessel sharpness of 1.0 refers to a maximum signal intensity change at the vessel border. The sharpness and the length of the vessels with CS reconstruction were compared with the gridding algorithm using a paired \( t \)-test. A \( P \) value less than 0.05 was considered to be statistically significant.

RESULTS

GPU Implementation of the CS Reconstruction

Table 1 shows the average time required for the completion of one iteration of the iterative CS reconstruction with CUDA and C++ implementations. The reconstruction was performed on the in vivo data for four different sampling densities (10, 20, 30, and 40%), which correspond to the sampling parameters (\( N_v, N_p, N_t \) = (392, 396, 10), (392, 768, 10), (392, 1152, 10), and (392, 1536, 10), respectively. The measured time is averaged over 100 iterations. The most time-consuming parts of the C++ implementation are the gridding and regridding operations, amounting to 67.1, 79.5, 85.3, and 88.5% of the total reconstruction time for 10, 20, 30, and 40% sampling densities, respectively. The speed-up gains of the GPU implementation over the C++ implementation
are also the largest for the gridding and regridding operations: 56.5–58.8 times speed-up for the gridding operation and 111.5–111.8 speed-up for the regridding operation. As the proportion of the gridding and regridding operations in the total reconstruction time of the C++ implementation increases, the total speed-up gain of the GPU implementations increases as well. Overall, the speed-up of the CUDA implementation of the CS reconstruction with wavelet domain regularization was 34.3, 43.7, 50.2, and 53.9 for 10, 20, 30, and 40% sampling densities, respectively. The speed-up of the CS reconstruction with the wavelet domain regularization was 35.4, 42.7, 48.4, and 51.9 for 10, 20, 30, and 40% sampling densities, respectively. The execution time of the gridding operation was about twice as long as the execution time of the regridding operation in CUDA implementation for a given sampling density, whereas the execution time of the gridding and regridding operations was nearly the same for C++ implementation. The gridding operation in CUDA is hampered by the cumulative memory writes, which is not present in the regridding operation; this is due to the cumulative memory writes in the CUDA implementation. The images reconstructed with the wavelet domain regularization were visually identical to those reconstructed with the C++ implementation. The reference image from 100% sampled k-space data and the reconstructed image from the under-sampled k-space data. Both of the CS reconstructions show improved image quality compared with the conventional gridding reconstruction, and the improvement is more distinct with lower sampling densities. The streaking artifacts degrade the image quality of the conventional gridding reconstructions for lower sampling densities (20, 10, and 7.5%), whereas most of the streaking artifacts are removed on the CS reconstructed images for the same sampling densities. Overall, the CS reconstructions have less visible artifacts and improved image homogeneity compared with the gridding reconstructions. In particular, the image domain regularization provides better image quality at sharp edges, whereas the wavelet domain regularization is generally better at removing streaking artifacts. The CS reconstruction with image domain regularization provides the least normalized mean-squared error values at all sampling densities.

Figure 5 depicts the resulting images generated by the CS reconstruction with image domain regularization for different numbers of iterations. The streaking artifacts in the earlier iterations are gradually removed as the number of iterations increases, whereas the image loses the sharpness at the edges of the phantom object and becomes slightly more blurry up to 500 iterations. After additional 2500 iterations, the sharpness of the object is improved and the image looks more refined with preserved edges. A similar trend in the convergence of the CS algorithm is observed for the reconstructions with the wavelet transform as the regularization term. However, no improvement in the image quality was observed after 500 iterations in this case.

Table 1

<table>
<thead>
<tr>
<th>(N_i, N_p, N_j)</th>
<th>(392, 396, 10)</th>
<th>(392, 768, 10)</th>
<th>(392, 1152, 10)</th>
<th>(392, 1536, 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CUDA</td>
<td>C++</td>
<td>SU</td>
<td>CUDA</td>
</tr>
<tr>
<td>FFT</td>
<td>0.27 s</td>
<td>5.00 s</td>
<td>18.5</td>
<td>0.26 s</td>
</tr>
<tr>
<td>IFFT</td>
<td>0.27 s</td>
<td>5.04 s</td>
<td>18.6</td>
<td>0.26 s</td>
</tr>
<tr>
<td>Gridding</td>
<td>0.31 s</td>
<td>17.59 s</td>
<td>56.5</td>
<td>0.58 s</td>
</tr>
<tr>
<td>Re-gridding</td>
<td>0.15 s</td>
<td>17.64 s</td>
<td>111.5</td>
<td>0.30 s</td>
</tr>
<tr>
<td>Thresholding</td>
<td>0.01 s</td>
<td>1.10 s</td>
<td>69.1</td>
<td>0.01 s</td>
</tr>
<tr>
<td>Etc.</td>
<td>0.50 s</td>
<td>6.07 s</td>
<td>-</td>
<td>0.51 s</td>
</tr>
<tr>
<td>Total (cs-image)</td>
<td>1.52 s</td>
<td>52.48 s</td>
<td>34.3</td>
<td>1.96 s</td>
</tr>
<tr>
<td>DWT</td>
<td>0.24 s</td>
<td>8.51 s</td>
<td>35.5</td>
<td>0.24 s</td>
</tr>
<tr>
<td>IDWT</td>
<td>0.21 s</td>
<td>8.72 s</td>
<td>41.5</td>
<td>0.21 s</td>
</tr>
<tr>
<td>Total (cs-wavelet)</td>
<td>1.97 s</td>
<td>69.71 s</td>
<td>35.4</td>
<td>2.41 s</td>
</tr>
</tbody>
</table>

DWT, discrete wavelet transform; IDWT, inverse discrete wavelet transform.

Phantom Experiment

Figure 4 shows the reconstruction results of an example slice of the 3D radial acquisition using the aforementioned algorithms with different sampling densities of 7.5, 10, 20, 30, and 40%. At the bottom left of each image, a selected region of the phantom is shown at a larger scale. The normalized mean-squared error from the reference image with 100% sampling density is also included at the bottom right of each image, calculated as $MSE = ||x_{ref} - x_{under}||^2_2/||x_{ref}||^2_2$, where $x_{ref}$ denotes the reference image from 100% sampled k-space data and $x_{under}$ denotes the reconstructed image from the under-sampled k-space data. Both of the CS reconstructions show improved image quality compared with the conventional gridding reconstruction, and the improvement is more distinct with lower sampling densities. The streaking artifacts degrade the image quality of the conventional gridding reconstructions for lower sampling densities (20, 10, and 7.5%), whereas most of the streaking artifacts are removed on the CS reconstructed images for the same sampling densities. Overall, the CS reconstructions have less visible artifacts and improved image homogeneity compared with the gridding reconstructions. In particular, the image domain regularization provides better image quality at sharp edges, whereas the wavelet domain regularization is generally better at removing streaking artifacts. The CS reconstruction with image domain regularization provides the least normalized mean-squared error values at all sampling densities.

In Vivo Experiment

Figures 6 and 7 show the example slices of axial and reformatted sagittal views from 3D whole-heart views.
FIG. 4. Comparison of conventional 3D gridding reconstruction vs. 3D iterative CS reconstruction with different sparsity regularization (image domain and wavelet domain) for a 3D radial acquisition using five different sampling densities (40, 30, 20, 10, and 7.5%). The number of iterations was 3000 and 500 for CS with image domain sparsity and wavelet domain sparsity, respectively. For high sampling densities all three reconstruction methods yield comparable image qualities. For lower densities, both CS reconstructions provide superior image qualities compared with the gridding algorithm, whereas CS with image domain sparsity shows better results at sharp edges and CS with wavelet domain sparsity is better at smooth surfaces. The normalized mean-squared errors are also included at the right bottom of the images.

FIG. 5. CS reconstruction with image domain regularization for a phantom imaged with 3D radial with sampling density of 7.5% at different numbers of iterations, initiated with the conventional gridding reconstruction. The streaking artifacts are gradually removed with some blurring up to 500 iterations; however, with additional iterations the streaking artifacts are suppressed with improved sharpness.
acquisitions with isotropic 1.3-mm spatial resolution reconstructed with the gridding reconstruction as well as iterative CS reconstruction with image and wavelet domain regularizations for four different sampling densities (6.8, 12.1, 24.2, and 36.3%). The images reconstructed with gridding present streaking artifacts and high-frequency noise-like artifacts, especially at lower sampling densities. Both CS reconstructions were able to substantially suppress these artifacts at lower densities. Although the wavelet domain regularization provides cleaner and more homogeneous results in the blood pool, the image domain regularization provides more detailed and sharper edges. The wavelet domain regularization results in checkerboard-like artifacts in the reconstructed image with 6.8% sampling density.

Figure 8 illustrates the resulting images of the CS reconstruction with image domain regularization for different numbers of iterations. The artifacts associated with undersampling are gradually removed, and the image quality improves as the number of iterations increases. The blurring of the image during the iterations shown in the phantom (Fig. 5) was not observed in the in vivo result. Between 500 and 3000 iterations, there is a slight improvement in the image quality but it was less prominent than the phantom case. Similar trends were observed for the wavelet domain regularized CS reconstruction, but no visual improvement was observed after 500 iterations.

Figure 9 depicts the reformatted RCA images from 3D whole-heart data with a spatial resolution of 1.0 × 1.0 × 1.0 mm³ and sampling density of 40%, reconstructed by the iterative CS reconstruction with the image domain regularization. The data set is retrospectively undersampled to get 10 and 20% sampling densities, and the reconstructed images are shown. Because of the isotropic resolution of the 3D radial acquisition in all three dimensions, the image can be reformatted retrospectively in an arbitrary angle to obtain a desirable imaging plane for visualizing the vessels. Table 2 summarizes the quantitative results of the 3D whole-heart images from six complete data sets with sampling densities of 10, 20, 30, and 40%. The measured vessel lengths increase as the sampling density increases for all the reconstruction methods, but the vessel lengths are not significantly different among the three reconstruction methods. The CS reconstruction with image domain regularization provides higher vessel sharpness for all sampling densities, and the improvements are statistically significant for sampling densities of 10, 20, and 30% compared with the gridding reconstruction. The CS reconstruction with wavelet domain regularization, however, does not show significant improvement in the vessel sharpness over the gridding reconstruction for any of the sampling densities.
FIG. 7. Example slices of sagittal views from 3D whole-heart images reconstructed by conventional 3D gridding reconstruction and iterative CS reconstruction (with 1000 iterations for image domain regularization and 500 iterations for wavelet domain regularization) for different sampling densities. For all the sampling densities, CS reconstructions have less high-frequency streaking artifacts, and the improvement in the image quality is more distinct at lower sampling densities.

FIG. 8. An example slice from 3D data set (sampling density = 6.8%) of the coronary arteries reconstructed using CS with image domain regularization at different iterations. The high-frequency artifacts are gradually removed throughout the iterations up to 500 iterations. Slight improvement was observed after 500 iterations, but it was less prominent than the phantom case (Fig. 5).
DISCUSSION

In this study, we have evaluated the implementation of a GPU-accelerated CS reconstruction for 3D radial imaging. GPU allows substantial reduction of the reconstruction time for 3D radial imaging. Phantom and in vivo whole-heart coronary MRI studies demonstrated the efficacy of CS reconstruction in removing the streaking artifacts, especially at high acceleration rates.

For CUDA implementation of the CS reconstruction, the computations in gridding/regridding operations can be assigned to the device code either by dividing the 3D radial data points among threads (radial point driven) or by dividing the Cartesian grid points among threads (Cartesian point driven). The radial point driven assignment is a simple and intuitive approach and has a minimum number of memory reads (writes) in gridding (regridding), but results in a large amount of data sharing among threads and cumulative memory writes in gridding as illustrated in Fig. 3a. Each CUDA thread assigned to the radial data point will read the memory to get the measured k-space value for the sample point and distribute the value to the neighboring Cartesian grid points inside the convolution window. The Cartesian grid point will have different contributions from different radial sample points, resulting in cumulative memory access among different CUDA threads. In our experiment, the execution time of the gridding/regridding operation was only twice as long as the regridding operation despite the massive cumulative memory access and atomic operations. On the other hand, the Cartesian point driven assignment has minimum number of memory writes (reads) in gridding (regridding). However, one must compute the list of the radial points associated with the Cartesian grid point within the convolution window for every thread, which requires additional computations and/or additional memory usage. The Cartesian point driven assignment is illustrated in Fig. 3b. Each CUDA thread assigned to the Cartesian grid point will read the memory to get the measured k-space values from neighboring radial sample points inside the convolution window, combine the values, and write on the memory for the Cartesian point only once. The Cartesian point driven assignment has an uneven workload distribution over different threads and causes poor compute to global memory access ratio for outer k-space points, especially at low sampling densities. Each thread assignment strategy has its advantages and disadvantages, and it is not simple to determine which one is superior to the other. In this article, we used radial point driven assignment; more study and optimization on the thread allocation and memory management can be done in the future for further speed-up of the parallel implementation.

The proposed implementation of the 3D radial acquisition still takes a long time to be clinically feasible. For example, we have used 3000 iterations for the CS reconstruction of the phantom data (3443 voxels) and 1000 iterations for the iterative CS reconstruction with image domain regularization and 1000

![Reformatted images of the RCA with isotropic resolution of (1.0 mm)³ from whole-heart 3D radial data with three sampling densities (40, 20, and 10%) by the iterative CS reconstruction with image domain regularization and 1000 iterations on GPU. The actual scan time with sampling density of 40% was 7 min 28 s with the navigator gating efficiency of 54%. The RCA is clearly visualized with the CS reconstruction for all sampling densities, while slight blurring of the image and residual artifacts are observed at low sampling density (10%).](image)

<table>
<thead>
<tr>
<th>Sampling density</th>
<th>Reconstruction method</th>
<th>RCA sharpness</th>
<th>RCA length (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>CS-image</td>
<td>0.65 ± 0.05*,#</td>
<td>7.29 ± 3.02</td>
</tr>
<tr>
<td></td>
<td>CS-wavelet</td>
<td>0.52 ± 0.06</td>
<td>7.35 ± 2.95</td>
</tr>
<tr>
<td></td>
<td>Gridding</td>
<td>0.53 ± 0.03</td>
<td>6.99 ± 2.82</td>
</tr>
<tr>
<td>20%</td>
<td>CS-image</td>
<td>0.61 ± 0.03*,#</td>
<td>7.32 ± 4.10</td>
</tr>
<tr>
<td></td>
<td>CS-wavelet</td>
<td>0.54 ± 0.04</td>
<td>7.08 ± 3.91</td>
</tr>
<tr>
<td></td>
<td>Gridding</td>
<td>0.51 ± 0.04</td>
<td>6.89 ± 3.61</td>
</tr>
<tr>
<td>30%</td>
<td>CS-image</td>
<td>0.60 ± 0.05*,#</td>
<td>8.41 ± 2.82</td>
</tr>
<tr>
<td></td>
<td>CS-wavelet</td>
<td>0.54 ± 0.03</td>
<td>8.50 ± 2.76</td>
</tr>
<tr>
<td></td>
<td>Gridding</td>
<td>0.54 ± 0.02</td>
<td>8.52 ± 2.86</td>
</tr>
<tr>
<td>40%</td>
<td>CS-image</td>
<td>0.64 ± 0.04</td>
<td>9.07 ± 3.28</td>
</tr>
<tr>
<td></td>
<td>CS-wavelet</td>
<td>0.59 ± 0.06</td>
<td>9.13 ± 3.26</td>
</tr>
<tr>
<td></td>
<td>Gridding</td>
<td>0.58 ± 0.05</td>
<td>8.78 ± 3.40</td>
</tr>
</tbody>
</table>

CS reconstruction with image domain regularization improves the vessel sharpness for sampling densities 10, 20, and 30% compared with the gridding reconstruction.

*P < 0.05 compared with the gridding reconstruction.

#P < 0.05 compared with the CS reconstruction with wavelet domain regularization.
iterations for the in vivo data (392^3 voxels), and the final reconstruction times for the data with 20% sampling density were around 5 and 2.5 h, whereas the conventional gridding algorithm takes a few minutes with the C++ implementation for any case. However, CS reconstruction with fewer iterations (e.g., 100 iterations) still provides improved image quality compared with the gridding algorithm, and the reconstruction time in this case is around 10–16 min with 20% sampling density with the GPU implementation.

Multiple GPUs can be used for further speed-up as the reconstruction of the individual coil images can also be parallelized among the GPUs. In this article, the proposed CS algorithm only uses the sparsity property of the image for the reconstruction of the undersampled k-space data and does not exploit the coil sensitivity information from multiple coils, which may potentially enable further undersampling. Techniques aiming to combine parallel imaging and CS for even higher acceleration rates have been proposed (10,49–51), and GPU implementations were also proposed for some of the approaches (52,53). The 3D radial CS reconstruction may also be combined with such techniques for further acceleration. The main issue of combining parallel imaging for the reconstruction of the 3D radial acquisition is the huge amount of the data and the limited size of the GPU’s global memory. The data from multiple coils cannot be stored on the GPU’s global memory at the same time, and the reconstruction process needs to be divided into smaller jobs to fit on the GPU’s memory. This will result in frequent memory access between the main system memory and the GPU’s memory, serial execution of the divided processes, and additional handling of the shared data between the divided processes. As GPU hardware is fast developing for general purpose computing, it is expected that these limitations will be resolved, enabling efficient implementation of more advanced reconstruction methods without complicated designing and optimization for GPU programming.

We have used the identity transform and the Daubechies 4 wavelets as the sparsifying transforms for the CS reconstruction. The baseline assumption for successful CS reconstruction is that the MR images are sparse in these transform domains. Wavelets have been applied in many MR reconstruction studies (50,54), but the use of image domain sparsity has been limited to applications such as MR angiography (55,56). The 3D radial trajectories are generally oversampled in the read-out direction, and this results in an increased FOV larger than the prescribed FOV. The 3D image then contains redundant areas where there is not much signal, making the image sparse in the image domain itself. Both image domain and wavelet domain regularizations have provided improved image quality compared with the conventional gridding algorithm, but exhibited some issues that need to be improved. The CS reconstruction with image domain regularization has a slow convergence speed with the iterative algorithm described in this article. The CS reconstruction with wavelet domain regularization provides a better convergence speed than the image domain regularization, but shows checkerboard-like and blocky artifacts at low sampling densities. The two-step iterative CS reconstruction algorithm used in this article enables simple and efficient coefficient-wise thresholding for the thresholding step in Eq. 5 only when the sparsifying transform is given by a unitary matrix. The image domain (identity transform) and Daubechies wavelet transform satisfy this condition, whereas the well-known and commonly used total variation regularization does not. The use of other sparsifying transforms or more advanced techniques that can adaptively capture object-specific sparsity nature (57,58) can improve the CS reconstruction, which requires further investigation.

**CONCLUSION**

We have implemented a GPU-accelerated iterative CS reconstruction method for 3D radial acquisitions and evaluated its performance in 3D whole-heart coronary MRI. The CS reconstruction method improved the image quality of highly undersampled 3D radial data sets compared with the conventional gridding reconstruction, and the GPU implementation was able to substantially reduce the reconstruction time.

**ACKNOWLEDGMENT**

The authors thank Jaime Shaw for help with proofreading.

**REFERENCES**


