FAST CARDIAC MAGNETIC RESONANCE IMAGING USING RADIAL TRAJECTORIES, TEMPORALLY CONSTRAINED RECONSTRUCTION, AND GRAPHICS PROCESSING UNIT CLUSTERS

by

Jordan P. Hulet

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biomedical Informatics

The University of Utah

December 2013

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Jordan P. Hulet**

has been approved by the following supervisory committee members:

Dennis Parker	, Chair	**05/30/2012** Date Approved
Julio Facelli	, Member	**05/30/2012** Date Approved
Gerald Treiman	, Member	**05/30/2012** Date Approved
Paul Clayton	, Member	**05/30/2012** Date Approved
Brian Chapman	, Member	**05/30/2012** Date Approved

and by **Julio Facelli** , Interim Chair of

the Department of **Biomedical Informatics**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Cine phase contrast (PC) magnetic resonance imaging (MRI) is a useful imaging technique that allows for the quantitative measurement of in-vivo blood velocities over the cardiac cycle. Velocity information can be used to diagnose and learn more about the mechanisms of cardio-vascular disease. Compared to other velocity measuring techniques, PC MRI provides high-resolution 2D and 3D spatial velocity information. Unfortunately, as with many other MRI techniques, PC MRI suffers from long acquisition times which places constraints on temporal and spatial resolution. This dissertation outlines the use of temporally constrained reconstruction (TCR) of radial PC data in order to significantly reduce the acquisition time so that higher temporal and spatial resolutions can be achieved. A golden angle-based acquisition scheme and a novel self-gating method were used in order to allow for flexible selection of temporal resolution and to ameliorate the difficulties associated with external electrocardiogram (ECG) gating. Finally, image reconstruction times for TCR are significantly reduced by implementation on a high-performance computer cluster. The TCR algorithm is executed in parallel across multiple GPUs achieving a 50 second reconstruction time for a very large cardiac perfusion data set.

For Caroline my love and Noah and Elijah my life.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my committee for their many contributions and sage advice. I would also like to thank the members of UCAIR who contributed time and effort in assistance of this research. Included among these are John Roberts, Seong-Eun Kim, Edward DiBella, and Ganesh Adluru.

This work was generously funded by several different sources: NLM training grant LM007124, the Ben B. and Iris M. Margolis Foundation, NIH grant HL 57990, and a Clinical Merit Review Grant from the VA Health Care System.

Finally, I would like to thank my family for their love, support, and encouragement; this work would not have been possible without them.

CHAPTER 1

INTRODUCTION

1.1 Magnetic Resonance Imaging

The principles behind magnetic resonance imaging (MRI) are elegant, to say the least. With essentially no moving parts, an MRI scanner is able to manipulate the protons in the object being imaged, arranging them in such a way that they can communicate anatomical and physiological information in great detail.

1.1.1 Nuclear Magnetic Resonance

All atomic nuclei have an intrinsic quantum mechanical property called spin. This spin is not equivalent to the classical notion of spin; however, MRI can be understood more simply if nuclei are pictured as positively charged spherical shells spinning around an axis of rotation in the classical sense (having an angular momentum Φ). For the remainder of this dissertation, classical analogies will be used in place of quantum mechanical descriptions of MRI. Nuclei with a non-zero spin have an angular momentum and a magnetic moment (μ), that is the product of the gyromagnetic ratio γ of the nucleus and its angular momentum [1, 2].

$$\mu = \gamma\Phi \tag{1.1}$$

Normally, individual magnetic moments are oriented randomly. However, in the presence of a strong external magnetic field, magnetic moments will have a tendency to align with the external field. In MRI, magnetic moments are placed in a strong magnetic field called the main field (B_0) that points longitudinally, or in the z direction. In the presence of the main field, magnetic moments will also exhibit resonance at a specific frequency called the Larmor frequency (ω) which depends on the gyromagnetic ratio of the nuclei and the strength of the external magnetic field [1, 2]: $\omega = B_0\gamma$.

Hydrogen nuclei (spin $\frac{1}{2}$) are the most commonly measured nuclei in MRI. Hydrogen has a gyromagnetic ratio of 42.58 MHz/T. Sodium, phosphorus, and fluorine are sometimes measured as well, but their concentrations in the body are much less than hydrogen and they produce much less measurable signal. At equilibrium, the orientation of individual magnetic moments is random, so the magnetic moments cancel each other out, resulting in no net magnetic moments of the spin system. However, in the presence of a strong magnetic external field, a small majority of magnetic moments align with the external magnetic field and generate a cumulative net magnetic moment of the spin system.

1.1.2 Excitation and Precession

Due to the Lorentz force, a magnetic moment will experience a torque from any external magnetic field with which it is not aligned. This torque will cause the magnetic moment to precess around the axis defined by the direction of the external field. Excitation is the process of applying an external magnetic field (B_1), rotating at the Larmor frequency, to magnetic moments aligned with B_0. The B_1 field is applied in the transverse plane and causes magnetic moments to precess away from B_0. Once magnetic moments are tipped away from B_0, they will experience a second torque that causes them to precess around B_0.

Using a rotating frame of reference (rotating at the Larmor frequency), the rotating B_1 field will appear stationary. As magnetic moments begin to precess around B_1, they will appear to tip directly away from B_0. In the stationary laboratory frame, this would look like a spiraling down instead of a simple tip due to the transverse precession caused by B_0. Ideally, after excitation, all magnetic moments (with the same gyromagnetic ratio) will precess at the same frequency with the same phase, resulting in a bulk precessing magnetic moment (M). The transverse component of this bulk precessing magnetic moment induces a voltage in a receive radio-frequency (RF) coil, creating a signal that can be measured.

The behavior of the bulk magnetic moment vector M in the presence of an external magnetic field (B) is described by the Bloch equation. B_0 contributes to B but, as will be described later, other external fields can also contribute to the net external field experience by M. A simplified version of the Bloch equation, only describing the effects of precession around B, is shown in Equation 1.2 where M and B are vectors [1].

$$\frac{dM}{dt} = M \times \gamma B \tag{1.2}$$

1.1.3 Gradients

As will be apparent in later sections on slice selection and image encoding, it is useful to expose magnetic moments to different field strengths as a function of their position. This is accomplished by applying an external magnetic field, called a gradient. Gradients add to or subtract from the z component of the total magnetic field. Gradients can be made to vary linearly in the x, y, and z directions. The strength of the magnetic field in the z direction (B_z) is described in Equation 1.3, where G is a vector representing the gradient that points in the direction that the main field, B_0, is changing and r is the position vector.

$$B_z(r) = B_0(r) + G \cdot r \tag{1.3}$$

1.1.4 Slice Selection

2D images are obtained by selectively exciting only the magnetic moments in a thin slice. This process is called slice selection. First, a gradient in the main field, in the direction normal to the slice, is induced. Next an RF pulse tuned to the range of Larmor frequencies for the desired slice width is applied. For 3D imaging, a similar process can be used to selectively excite a larger slab, the primary difference being a lower gradient strength, and/or a larger frequency band in the excitation RF pulse.

1.1.5 Relaxation

After excitation, signal generated by precessing magnetic moments will decay over time due to two primary effects called T_1 and T_2 relaxation. T_1 relaxation is directly related to a magnetic moment's tendency to realign itself with the main field and can be described by Equation 1.4, where M_0 is the equilibrium magnetization [1].

$$\frac{dM_z}{dt} = \frac{-(M_z - M_0)}{T_1} \tag{1.4}$$

Over time, the energy and angular momentum of a magnetic moment is lost to the lattice and it will eventually realign with the main field. T_2 decay is a separate effect caused by the dephasing of individual magnetic moments in close proximity to one another. The Larmor frequency depends on the strength of the local magnetic field. Tiny field inhomogeneities cause the various magnetic moments within close proximity to precess at different frequencies, resulting in dephasing of the magnetization over time. Over time,

T_1 and T_2 relaxation both contribute to a signal decay after excitation. T_2 decay can be described by Equation 1.5 [1].

$$\frac{dM_{xy}}{dt} = -\frac{M_{xy}}{T_2} \tag{1.5}$$

Adding the effects of signal decay from Equations 1.4 and 1.5 to the simplified version of the Bloch equation (Equation 1.2) yields the complete Bloch equation (1.6), where k is a unit vector in the z direction and i and j are orthogonal unit vectors in the transverse plane [1].

$$\frac{dM}{dt} = M \times \gamma B - \frac{(M_z - M_0)k}{T_1} - \frac{M_x i + M_y j}{T_2} \tag{1.6}$$

1.1.6 Signal Equation

Ignoring the effects of T_1 and T_2 decay, the solution to the Bloch equation is a phasor that rotates around the z axis at the Larmor frequency. This signal equation is shown in Equation 1.7 where $m(x, y)$ is the magnitude of the magnetic moment at position (x, y) [1].

$$S(t) = \int_x \int_y m(x, y) e^{-i\omega_0 t} dx dy \tag{1.7}$$

When a static magnetic gradient $G(x, y)$ is applied along with the main field, the precession frequency of magnetic moments becomes spatially dependent (r is the position vector) [1].

$$S(t) = \int_x \int_y m(x, y) e^{-i\omega_0 t} e^{-i\gamma G(x,y) \cdot r} dx dy \tag{1.8}$$

Before proceeding with the discussion of the imaging equation and image encoding, a brief discussion of the Fourier Transform is required.

1.1.7 The Fourier Transform

The concept of the Fourier Transform is essential to understanding how MRI images are encoded and decoded. The Fourier Transform transforms a signal to the inverse domain

(also called the Fourier domain) of the original signal. For example, a signal in the time domain (s) is transformed to the domain of $\frac{1}{s}$, or the *frequency domain*. Similarly, a signal in the spatial domain (x) is transformed to the domain of $\frac{1}{x}$ or the *spatial frequency domain*. The utility of the Fourier Transform lies in the fact that some problems are more easily solved when the domain is changed in such a way.

The Fourier Transform can be applied to both continuous and discrete signals. MRI makes discrete measurements and produces images with discrete voxels (small discrete volumes of magnetic moments that can be thought of as the 3D equivalent of pixel) so the discrete Fourier Transform (DFT) is used. The DFT and its inverse are shown in Equations 1.9 and 1.10. The coefficient k is used to represent Fourier domain data; this reflects the common practice in MRI of referring to Fourier domain data as k-space data. [3]

$$F(k) = \sum_{x=0}^{N-1} f(x) \cdot e^{-i2\pi k \frac{x}{N}} \tag{1.9}$$

$$f(x) = \frac{1}{N} \sum_{k=0}^{N-1} F(k) \cdot e^{i2\pi k \frac{x}{N}} \tag{1.10}$$

The DFT is separable and can be applied to any number of dimensions; in MRI, we are frequently interested in 2D images and 3D volumes. A 2D form of the DFT is shown in Equations 1.11 and 1.12, and can be similarly expanded to 3 dimensions. [3]

$$F(k,l) = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x,y) \cdot e^{-i2\pi (k \frac{x}{N} + l \frac{y}{M})} \tag{1.11}$$

$$f(x,y) = \frac{1}{NM} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} F(k,l) \cdot e^{i2\pi (k \frac{x}{N} + l \frac{y}{M})} \tag{1.12}$$

The DFT illustrates that any discrete signal can be expressed as a weighted sum of spatial harmonics. These harmonics show up in the 1D equations as $e^{i2\pi k \frac{x}{N}}$ where $k = [0, N-1]$. MRI hardware is responsible for creating spatial harmonics in the volume being imaged. Once the signals from these are measured, an image can be produced by applying the inverse DFT. A cartoon illustrating the basics of the DFT is shown in Figure 1.1.

It should be noted that a computationally optimized version of the DFT, called the Fast Fourier Transform (FFT, and IFFT for inverse FFT) is used in practice. The FFT takes less time to compute but produces the same output as the standard DFT.

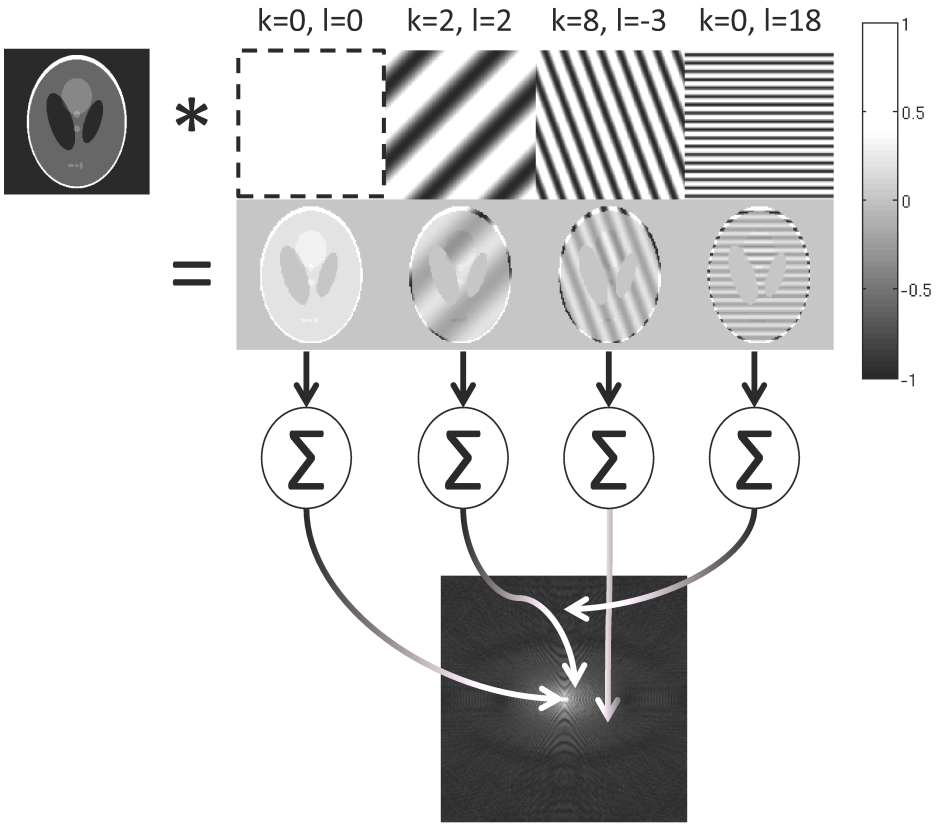

Figure 1.1. Illustration of the computation of 4 Fourier coefficients for a simple phantom. Only the real component for spatial frequencies and their product with the phantom are shown. A scaled version of the magnitude of k-space is shown.

1.1.8 Frequency and Phase Encoding and K-space

In MRI, relative frequencies and phases of magnetic moments are used to encode and decode images; as a result, the measured signal is demodulated, leaving only the base-band signal. Additionally, time-varying magnetic gradients are used to induce controlled frequency and phase differences. Another form of the signal equation, with the demodulated base-band signal and time-varying magnetic gradients, is shown in Equation 1.13 [1].

$$S(t) = \int_x \int_y m(x,y) e^{-i[x \int_0^t G_x(\tau)d\tau + y \int_0^t G_y(\tau)d\tau]} dxdy \tag{1.13}$$

Gradient waveforms can be designed in order to induce a specific linear phase difference at time t and G_x and G_y can be rewritten in terms of two new functions $k_x(t)$ and $k_y(t)$.

$$\gamma \int_0^t G_x(\tau)d\tau = 2\pi k_x(t) \tag{1.14}$$

$$\gamma \int_0^t G_y(\tau)d\tau = 2\pi k_y(t) \tag{1.15}$$

Finally, the signal equation can be written as in Equation 1.16 [1].

$$S(t) = \int_x \int_y m(x,y) e^{-i2\pi[k_x(t)x + k_y(t)y]} dxdy \tag{1.16}$$

This form of the signal equation illustrates that the MRI signal over time consists of coefficients in the spatial frequency domain, frequently called k-space. In the simplest case, reconstruction of images from MRI k-space data is performed by inverse FFT. Reconstruction using an inverse FFT requires that k-space be fully sampled. To accomplish this, magnetic moments must be properly encoded with spatial information.

One technique for measuring complete k-space can be thought of as two steps, namely phase and frequency encoding. Phase encoding is accomplished by applying a gradient field in the phase encoding direction before any signal is measured. This creates a spatially dependent phase difference in the phase encoding direction and determines the k_{PE}. Next, a gradient field in the readout direction is applied while the signal is measured simultaneously. This results in the measuring of multiple k_{RO} points as the phase difference in the readout direction evolves. This process is repeated with different phase encoding gradient pulses,

resulting in multiple lines of k-space in the RO direction, each with a difference phase encoding. Frequency encoding is used to encode one spatial dimension of an image and phase encoding can be used to encode one or two more spatial dimensions, allowing for the encoding of 2D and 3D images.

1.1.9 Spin Echoes

Signal lost to T_1 decay can only be recovered by tipping the magnetic moment back into the transverse plane. However, some of the effects of T_2 decay can be reversed by generating a spin echo. The principle of spin echoes relies on the fact that some field inhomogeneities are constant over a short period of time; these inhomogeneities cause dephasing of the magnetic moment, resulting in signal decay. Using an additional RF pulse, applied some time after the initial excitation pulse, magnetic moments can be rotated, or flipped, 180° around the x or y axis. Spins originally ahead in phase of precession are flipped behind, and magnetic moments originally behind in phase of precession are flipped ahead. This flip does not affect the local field strengths, so slower magnetic moments continue to precess slower and fast magnetic moments continue to precess faster. As a result, magnetic moments flipped behind in phase of precession will catch up to magnetic moments flipped ahead and they will be temporarily rephased causing a recovery of the signal lost due to the initial dephasing.

1.1.10 Gradient Echoes

Gradient echoes are another type of echo caused by the refocusing of magnetic moments that were initially dephased due to the application of the gradient fields. A gradient echo is equivalent to traversing the center of k-space since this is where all the magnetic moments across the volume are in phase. Frequently, spin echoes will be used in conjunction with gradient echoes and the timing will be arranged so that both occur at the same instant, producing the most signal at the center of k-space.

1.1.11 MRI Hardware

An MRI scanner is composed of three fundamental components: the main magnet, magnetic gradient coils, and transmit and receive radio-frequency coils (Figure 1.2). The main magnet is responsible for generating the strong main magnetic field used to align the magnetic moments of protons in the body. Gradient coils are used for spatial encoding, and the RF system is used to excite and measure the signal from precessing protons.

The main magnet produces a constant homogeneous magnetic field (B_0 that points in

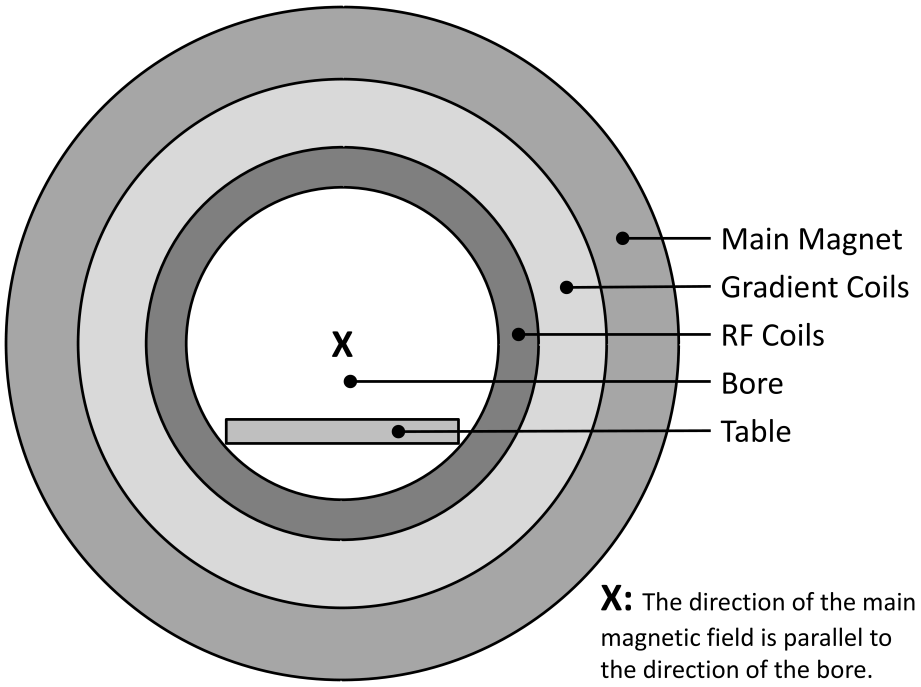

Figure 1.2. An axial cross section of a simplified MRI scanner is shown. The three main components, main magnet, gradient coils, and RF coils, are nested inside one another. The patient lays on the table that sits in the bore of the magnet.

the direction of the long axis of the bore) throughout the imaging volume that is responsible for aligning the bulk magnetic moment in the longitudinal direction, making precession of atomic nuclei possible. The main magnet of high-field MRI systems (generally 1.5T and above) is a super-conducting electromagnet that is super-cooled to less than 4° K using liquid helium. The super-conducting main magnet maintains its strong magnetic field as long as its low temperature is maintained.

Magnetic gradient coils are resistive electromagnets that create approximately linearly-varying magnetic fields that either oppose or contribute to the main magnetic field. Although the fields generated by the gradient coils contain components not parallel to the main field, only the components that are parallel to the main field have an effect on the frequency of precession and are relevant to imaging. Gradient coils produce much weaker magnetic fields than the main magnet, creating magnetic field gradients in the main field as large as 20 - 50 mT/m. Separate gradient coils are used to generate gradients in the x, y, and z directions. Gradient coils are used during imaging to manipulate magnetic moments for slice selection, spatial frequency encoding, velocity encoding, diffusion encoding, and other tasks performed by a given pulse sequence.

The RF system has two functions or operational modes, transmit and receive. When the RF system is transmitting, it generates a rotating magnetic field (perpendicular to the main field, rotating at the Larmor frequency) that is used to excite, flip, or otherwise influence the orientation of magnetic moments in the imaging volume. In receive mode, the RF system measures voltages induced by the rotating bulk magnetic moment of the imaging volume. These RF signals constitute the spatially encoded measurement data from the scanner and are reconstructed into images.

1.1.12 Computer Control

The MRI hardware is controlled by computer systems connected to the scanner. Generally, these systems consist of three separate machines: the host, scanner controller, and image reconstructor. The operator interacts with the host machine, using it to define protocols, perform scans, and view reconstructed images. The scanner controller controls the activity of the gradient and RF coils during an acquisition depending on the protocol that is prescribed by the host. The image reconstruction machine receives the raw data from the scan controller and reconstructs the image data that are sent back to the host for display and storage in the scanner image database.

1.1.13 Image Acquisition

MRI measurements consist of a series of spatially encoded free induction decays (FIDs) from magnetic moments that have been excited. This spatial encoding is usually described as two different steps: phase encoding and frequency encoding that will be outlined first, then a more general interpretation will be presented. When a gradient in the main magnetic field is applied in any spatial direction, it has the effect of inducing a linear phase in the spin system in the same direction. The longer the gradient is maintained or the higher the gradient strength, the steeper the linear phase change. In other words, magnetic gradients wind up the magnetic moments in the direction of the gradient; the stronger or longer the gradient, the more the magnetic moments are wound up.

1.1.14 Pulse Sequences

Pulse sequences are, in a sense, the programming language of MRI scanners. The type of pulse sequence used determines what kind of images are acquired. Pulse sequences define the operation of the RF, gradient, and ADC (analog to digital converter) systems. The timing and relative operation of each of these systems is very important. Pulse sequences define the RF waveforms that are used for excitation, including the frequency, phase, and envelope of the Rf pulse. Pulse sequences defined the current applied to the gradient fields. Each gradient (x, y, and z) can be operated independently. The ADC is used to convert the analog signal measured from the precessing magnetic moments, measured by the receive RF coils, to a digital signal that is recorded. The ADC "waveform" in pulse sequence diagrams represents the two states of the ADC, on and off. When the ADC is on, the signal is being recorded.

A diagram for a basic Cartesian spin echo pulse sequence is shown in Figure 1.3. This pulse sequence performs slice selective excitation, phase encoding, and a frequency encoding readout. In this diagram, the RF waveform only shows the envelope of the signal. The gradient waveforms in this diagram are box-shaped; however in reality, the gradients have limited rise times (current cannot be changed instantaneously) and the actual achieved waveforms are more similar to a trapezoidal shape. This pulse sequence diagram defines a single frame of a pulse sequence. This frame must be repeated many times to cover all of k-space. The multiple lines in the y gradient represent the fact that the magnitude of this waveform is changed for each frame. This is necessary as each frame consists of a single phase encoding.

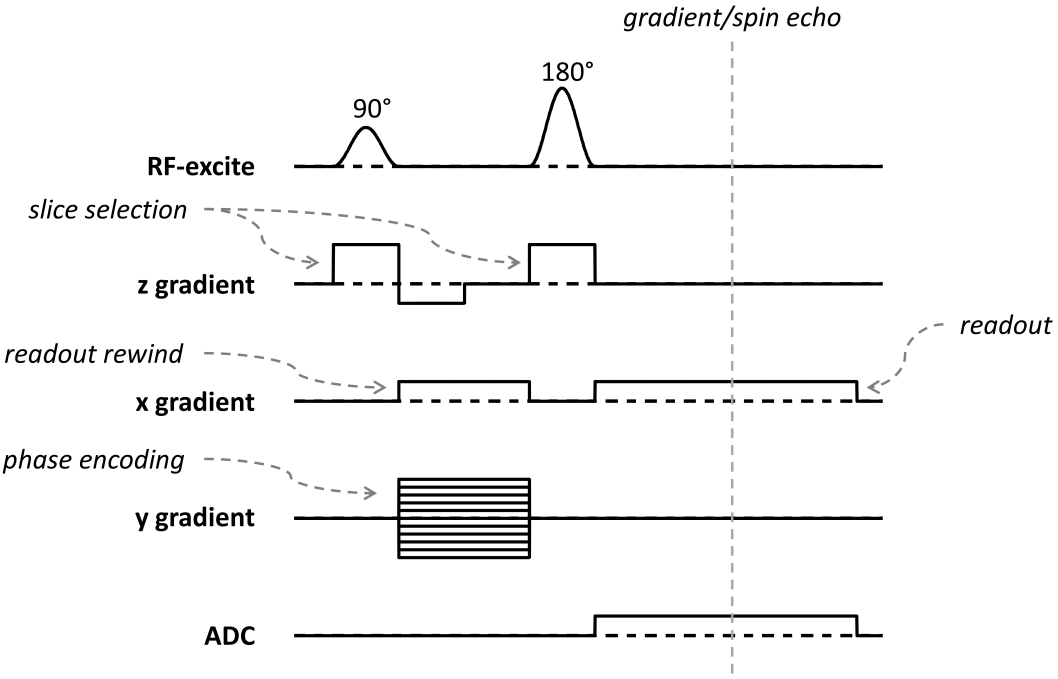

Figure 1.3. Diagram of a simple Cartesian spin echo pulse sequence. The amplitude of the y gradient is changed from repetition to repetition for different phase encodings.

1.1.15 Phase Contrast Imaging

Pulse sequences can be designed to measure the velocity of flowing blood. This is called phase-contrast (PC) imaging. [4, 5, 6]. PC imaging is possible due to the fact that magnetic moments moving through a magnetic gradient experience different magnetic field strengths at different times. When two gradient pulses, equal in magnitude and duration but opposite in polarity, are applied, they essentially wind and unwind the magnetic moment phases, respectively. Magnetizations that are in motion accumulate a different amount of phase during the first pulse than is lost during the second pulse, due to the change in magnetic field strength due to changing position in the gradient. This results in a net phase difference for moving magnetic moments and no phase difference for stationary magnetic moments and is called velocity encoding. When two acquisitions are performed, one with velocity encoding and one without, the relative phase of the two reconstructed acquisitions is proportional to the velocity of magnetic moments in each voxel, in the direction of the applied velocity encoding gradients. It is frequently useful to measure flow patterns over the period of a cardiac cycle.

1.1.16 Cine Imaging

Cine MRI is an imaging technique that results in a series, or loop, of images over time instead of a single, static image. Cine imaging is frequently applied to cardiac MRI where it produces multiple images spread across the cardiac cycle. These images can be shown as a movie of a single heart cycle. Cine imaging differs from real-time imaging because the resulting images represent an average cardiac cycle. This is due to the fact that multiple cardiac cycles must be measured in order to adequately sample the measurement space, allowing for the reconstruction of artifact-free images. This is accomplished by measuring a gating signal during image acquisition. This signal is frequently the output from electrocardiogram leads or a pulse-oximeter. The cardiac cycle is divided into temporal bins, the number of bins determines the total number of cine images that will be produced. Each measurement is compared to the gating signal and placed into the correct temporal bin. Once all of the bins are filled, they can be reconstructed as frames in the cine sequence.

1.1.17 K-Space Under-sampling

MRI acquisitions are limited by the amount of time that is required to measure k-space data. This is due to limitations of the scanner hardware and restrictions on the speed and magnitude of the gradient waveforms that can be safely used in-vivo. As a result, it is often desirable to only measure a subsample of the entire k-space. This results in reduced

acquisition times, but leads to an under-sampling artifact that can be understood as a convolution in image space by the sampling pattern in k-space. In order to utilize under-sampled acquisitions, special image reconstruction techniques are required to reconstruct images while minimizing the effect of under-sampling artifacts [7, 8].

1.1.18 Gridding

It is often desirable to measure MRI data using non-Cartesian trajectories. For example, using a radial trajectory to sample k-space results in an oversampling of the center of k-space due to the overlapping measurements of each radial line, or "view." Oversampling the center of k-space provides a degree of robustness to motion artifacts as well as being beneficial for under-sampled acquisitions.

In order to be reconstructed by the FFT, k-space data need to be on a Cartesian grid, as this is a requirement of the FFT operation. The process of transforming non-Cartesian k-space data to a Cartesian grid is called gridding. Gridding generally involves a convolution step that spreads the k-space data from neighboring non-Cartesian points to Cartesian grid locations. This convolution in k-space causes a roll-off artifact in image space that needs to be corrected using postapodization correction, or a multiplication in the image domain. If the roll-off correction is too aggressive, this can cause noise amplification at the edges of the reconstructed image [9].

1.1.19 Parallel Imaging

Parallel imaging is a general technique in which additional coil-sensitivity information is used in conjunction with measured k-space data to fill in un-measured k-space data and minimize under-sampling artifacts. As parallel imaging leverages the additional information provided by the coil sensitivities, the degree of under-sampling that it can handle (the acceleration factor) is limited by the number of receiver channels used to acquire the data. As a result, it is common practice for clinical systems to be designed with multiple receive channels to operate phased-array receiver coils, having as many as 32 or more independent receiver channels [10, 11, 12].

1.1.20 TCR

Temporally constrained reconstruction (TCR) is a reconstruction technique for under-sampled k-space data. It comes from the family of constrained reconstruction, or compressed sensing reconstruction algorithms. TCR estimates missing k-space data by placing a temporal constraint on MRI data containing a temporal dimension, such as cine or time resolved

data. The TCR algorithm converges to an image estimate by performing iterations that minimize a cost functional composed of the temporal constraint and a data fidelity term that constrains the solution to conform with the measurement data [13, 14, 15].

1.2 Parallel Computation

Reconstruction algorithms for under-sampled data can be much more complex than a simple IFFT required for reconstructing fully sampled Cartesian data. As a result, it is becoming an ever increasing trend in reconstruction computation to utilize parallel computing techniques to reduce computational time requirements. This is evidenced by the increasing number of cores found in commodity CPUs. Additionally, many academic centers have high-performance computational facilities consisting of clusters of multicore computers networked together. These facilities are frequently used by departments such as chemistry and physics to run complex simulations. With appropriately written software, large computational tasks can be distributed across many different nodes. In addition to computational clusters, massively parallel architectures such as graphics processing units (GPUs) and Intel's many integrated core (MIC) are becoming more and more accessible.

1.2.1 Computer Clusters

Computer clusters consist of many separate computers interconnected and outfitted with software that allows for communication between nodes and execution of parallelized jobs. In order to attain fast internode communication, clusters are frequently equipped with high throughput communication such as Infiniband (Mellanox Technologies) or fiber-channel connections. Compute clusters are generally a shared resource where access to compute power must be mediated in some way. As a result, a requested job may not execute immediately, and may wait in the queue for as long as several days while it waits for the appropriate resources to become available [16].

1.2.2 Graphics Processing Units

Graphics processing units (GPUs) are commodity hardware that were originally designed and developed to render graphics for 3D computer games. In order to be able to render high numbers of polygons quickly (30 - 60 frames per second), GPUs adopted a massively parallel architecture that performs computations using the single instruction multiple data (SIMD) model. Recently, GPUs have been used for general purpose computing (GPGPU) in addition to powering 3D graphics. Vendors such as NVIDIA and ATI have started manufacturing new GPUs with capabilities allowing for efficient GPGPU computing. Computing frameworks,

such as NVIDIA CUDA and open CL, are now available for GPGPU programming [17, 18, 19].

1.2.3 Difficulties of Parallelization

Software developed for serial execution on a CPU is not always easily parallelized. Computation must be able to execute in parallel in order to take advantage of parallel computing hardware. Usually, one cannot simply run serial software in a parellel environment and achieve any performance increase. Converting serial code to work in a parallel environment may require porting some or all of the code to a different language such as CUDA or OpenCL. Additional considerations, such as efficient data partitioning or communication as well as separate memory devices for CPUs and GPUs, need to be made.

1.3 Roadmap

The aim of this dissertation is to outline new work allowing for reduced acquisition and reconstruction times of cine and real-time cardiac imaging. By reducing both acquisition and reconstruction times, overall exam time can be reduced. Reduction in exam time can be directly translated to reduced costs and improved patient experience. The following is an outline of the following chapters of this dissertation.

1.3.1 Temporally Constrained Reconstruction of Radial Cine Cardiac Phase-Contrast Data

Cine phase-contrast (PC) imaging is an invaluable technique used to measure in-vivo blood velocities. PC imaging necessitates the inclusion of extra velocity encoding gradients which increases the total acquisition time. As a result, PC imaging sequences are even more lengthy than comparable magnitude only sequences. Due to long acquisition times, high temporal resolution cine sequences are not possible. Chapter 2 outlines an approach to using radial under-sampling along with temporally constrained reconstruction to reduce the acquisition time of a cine PC sequence while increasing the temporal resolution.

1.3.2 Under-Sampled Golden Ratio Cine Phase-Contrast MRI Using Automatic Distance Matrix Retrospective Self-Gating

Cine sequences must use a gating mechanism in order to label the cardiac phase of measurements, so that they can be appropriately binned. Traditionally, this is accomplished by measuring the signal from attached ECG leads or a pulse oximeter. These gating techniques can suffer from interference from the gradients, poor lead placement, or poor

pulse-ox signal due to a cold finger, and are just one more step required in an already lengthy, time-limited exam. Chapter 3 outlines a new approach, using golden-ratio radial under-sampled cine PC imaging, along with automatic distance matrix self-gating and temporally constrained reconstruction to allow for rapid cine PC acquisition without any external gating requirements. This new technique proved to be robust to missing and corrupted measurements while still allowing for a significantly reduced scan time.

1.3.3 Fast Temporally Constrained Reconstruction of a Large Cardiac Perfusion MRI Dataset Using a GPU Cluster

Solutions aimed at reducing MRI acquisition times require advanced image reconstruction techniques beyond a simple inverse FFT. This results in decreased acquisition times but increased reconstruction times. During an exam, the reconstructed images are frequently needed to determine if an acquisition must be repeated due to poor image quality or bad slice location. Unfortunately, during exam reconstruction, using methods such as TCR is not possible with current hardware, greatly decreasing the clinical utility of these reconstruction methods. Chapter 4 demonstrates the novel use of GPU clusters to perform TCR image reconstruction remotely. The reconstruction time for a large cardiac perfusion data set was reduced to the order of a minute, a reasonable amount of time to wait for image reconstruction on the scanner. Using this technique, image reconstruction could be performed remotely, allowing for exam-time reconstruction with methods such as TCR without the need for new hardware on the scanner.

1.3.4 Conclusion

Lastly, this dissertation will conclude with Chapter 5. This chapter will summarize the work outlined in this dissertation and propose future work.

1.4 References

[1] Dwight G. Nishimura. *Principles of Magnetic Resonance Imaging.* Stanford University, 2007.

[2] Jonathan Links Jerry L. Prince. *Medical Imaging Signals and Systems.* Prentice Hall, 2005.

[3] Richard E. Woods Rafael C. Gonzalez. *Digital Image Processing.* Prentice Hall, 2007.

[4] D. N. Firmin, G. L. Nayler, R. H. Klipstein, S. R. Underwood, R. S. Rees, and D. B. Longmore. In vivo validation of MR velocity imaging. *J Comput Assist Tomogr*, 11(5):751–756, 1987.

[5] D. Meier, S. Maier, and P. Bosiger. Quantitative flow measurements on phantoms and on blood vessels with MR. *Magn Reson Med*, 8(1):25–34, Sep 1988.

[6] P. G. Walker, S. Oyre, E. M. Pedersen, K. Houlind, F. S. Guenet, and A. P. Yoganathan. A new control volume method for calculating valvular regurgitation. *Circulation*, 92(3):579–586, Aug 1995.

[7] C. M. Tsai and D. G. Nishimura. Reduced aliasing artifacts using variable-density k-space sampling trajectories. *Magn Reson Med*, 43(3):452–458, Mar 2000.

[8] K. Arfanakis, A. A. Tamhane, J. G. Pipe, and M. A. Anastasio. k-space undersampling in PROPELLER imaging. *Magn Reson Med*, 53(3):675–683, Mar 2005.

[9] N. R. Zwart, K. O. Johnson, and J. G. Pipe. Efficient sample density estimation by combining gridding and an optimized kernel. *Magn Reson Med*, 67(3):701–710, Mar 2012.

[10] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn Reson Med*, 47(6):1202–1210, Jun 2002.

[11] D. K. Sodickson and W. J. Manning. Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays. *Magn Reson Med*, 38(4):591–603, Oct 1997.

[12] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger. SENSE: sensitivity encoding for fast MRI. *Magn Reson Med*, 42(5):952–962, Nov 1999.

[13] G. Adluru, S. P. Awate, T. Tasdizen, R. T. Whitaker, and E. V. Dibella. Temporally constrained reconstruction of dynamic cardiac perfusion MRI. *Magn Reson Med*, 57(6):1027–1036, Jun 2007.

[14] G. Adluru, C. McGann, P. Speier, E. G. Kholmovski, A. Shaaban, and E. V. Dibella. Acquisition and reconstruction of undersampled radial data for myocardial perfusion magnetic resonance imaging. *J Magn Reson Imaging*, 29(2):466–473, Feb 2009.

[15] N. Todd, G. Adluru, A. Payne, E. V. DiBella, and D. Parker. Temporally constrained reconstruction applied to MRI temperature data. *Magn Reson Med*, 62(2):406–419, Aug 2009.

[16] B. Kressler, P. Spincemaille, M. R. Prince, and Y. Wang. Reduction of reconstruction time for time-resolved spiral 3D contrast-enhanced magnetic resonance angiography using parallel computing. *Magn Reson Med*, 56(3):704–708, Sep 2006.

[17] M. S. Hansen, D. Atkinson, and T. S. Sørensen. Cartesian SENSE and k-t SENSE reconstruction using commodity graphics hardware. *Magn Reson Med*, 59(3):463–468, Mar 2008.

[18] S. S. Stone, J. P. Haldar, S. C. Tsao, W. M. Hwu, B. P. Sutton, and Z. P. Liang. Accelerating Advanced MRI Reconstructions on GPUs. *J Parallel Distrib Comput*, 68(10):1307–1318, Oct 2008.

[19] S. Roujol, B. D. de Senneville, E. Vahala, T. S. Sørensen, C. Moonen, and M. Ries. Online real-time reconstruction of adaptive TSENSE with commodity CPU/GPU hardware. *Magn Reson Med*, 62(6):1658–1664, Dec 2009.

CHAPTER 2

HIGHLY ACCELERATED CARDIAC CINE PC MRI USING AN UNDERSAMPLED RADIAL ACQUISITION AND TEMPORALLY CONSTRAINED RECONSTRUCTION

Jordan P. Hulet, Andreas Greiser, Jason K. Mendes, Chris McGann, Gerald Treiman, Dennis L. Parker. Highly Accelerated Cardiac Cine Phase-Contrast MRI Using an Undersampled Radial Acquisition and Temporally Constrained Reconstruction. *J. Magn. Reson. Imaging*, [in press].

JOURNAL OF MAGNETIC RESONANCE IMAGING 00:000–000 (2013)

Technical Note

Highly Accelerated Cardiac Cine Phase-Contrast MRI Using an Undersampled Radial Acquisition and Temporally Constrained Reconstruction

Jordan P. Hulet, BS,[1,2] Andreas Greiser, PhD,[3] Jason K. Mendes, PhD,[2] Chris McGann, MD,[4] Gerald Treiman, MD,[5] and Dennis L. Parker, PhD[1,2,6]*

Purpose: To evaluate a method to enable single-slice or multiple-slice cine phase contrast (cine-PC) acquisition during a single breath-hold using a highly sparsified radial acquisition ordering and temporally constrained image reconstruction with a spatially varying temporal constraint.

Materials and Methods: Simulated and in vivo cine-PC datasets of the proximal ascending aorta were obtained at different acceleration factors using a view projection acquisition order optimized for temporally constrained reconstruction (TCR). Reconstruction of the sparse cine-PC data performed with TCR was compared to reconstructions using zero-filled regridding and temporal interpolation.

Results: TCR resulted in more accurate velocity measurements than regridding or temporal interpolation. In one dataset, TCR of undersampled in vivo data (16 views per cardiac phase) resulted in a peak systolic velocity within 3.3% of the value measured by Doppler ultrasound while shortening the scan time to 13 seconds. High temporal-resolution undersampled TCR was also compared lower temporal-resolution, more highly sampled, regridding in three normal volunteers.

Conclusion: TCR proved to be an effective method for reconstructing undersampled radial PC data. Although TCR utilizes a temporal constraint, temporal blurring was minimized by using appropriate constraint weights in addition to a spatially varying temporal constraint. TCR allowed for the acquisition time to be reduced to the duration of a breath-hold, while still resulting in accurate velocity measurements.

Key Words: constrained reconstruction; iterative reconstruction; phase-contrast
J. Magn. Reson. Imaging 2013;00:000–000.

CINE PHASE-CONTRAST (cine-PC) imaging is an important magnetic resonance imaging (MRI) technique for measuring in vivo blood velocities (1–3). Unfortunately, PC scan times are intrinsically long because they require the acquisition $n+1$ datasets, where n is the number of velocity directions measured. Furthermore, cardiac PC imaging is affected by respiratory motion. Navigator echoes or breath-hold acquisitions are frequently used to avoid motion artifacts (4–6), but both of these techniques have negative consequences. Navigator echoes drastically increase scan time because data are only acquired during times when the heart position is sufficiently consistent. Breath-hold acquisitions can be performed rapidly, but are limited to the duration of a breath-hold, which places limits on the spatial and temporal resolution and total coverage of the acquisition, which are important for accurate velocity measurements. In order to allow for accelerated PC acquisitions with high temporal and spatial resolution, techniques that can reconstruct highly accelerated (sparsified) PC acquisitions, while still maintaining temporal resolution, are needed.

Accelerated PC

Several techniques have been previously used to accelerate cine-PC imaging. Nearly all of these techniques involve undersampling the measurements in each time frame of each image slice and then using one of a number of different techniques to reconstruct

[1]Biomedical Informatics, University of Utah, Salt Lake City, Utah, USA
[2]Utah Center for Advanced Imaging Research, University of Utah, Salt Lake City, Utah, USA
[3]Siemens Healthcare Sector, Erlangen, Germany
[4]Division of Cardiology, University of Utah, Salt Lake City, Utah, USA
[5]Veteran's Health Care Administration, Salt Lake City, Utah, USA
[6]Radiology Department, University of Utah, Salt Lake City, Utah, USA
Contract grant sponsor: National Institutes of Health (NIH); Contract grand numbers: NIH R01 HL48223; NIH R01 HL57990.
Contract grant sponsor: Siemens Medical Solutions.
Contract grant sponsor: Mark H. Huntsman Endowed Chair.
Contract grant sponsor: NLM training grant; Contract grand number: LM007124.
Contract grant sponsor: Ben B. and Iris M. Margolis Foundation..
*Address reprint requests to: D.L.P., Utah Center for Advanced Imaging Research, Radiology Department, 729 Arapeen Dr., Salt Lake City, UT 84108. E-mail: parker@ucair.med.utah.edu
Received July 19, 2011; Accepted March 6, 2013
DOI 10.1002/jmri.24160
View this article online at wileyonlinelibrary.com.

images from these undersampled datasets. Parallel image reconstruction methods, such as generalized autocalibrating partially parallel acquisition (GRAPPA) and sensitivity encoding (SENSE), rely on the assumed independence of the radiofrequency (RF) receiver coil sensitivities to provide extra information from which the missing measurements can be predicted. Because of the limited number of coils and the fact that the coil sensitivities are not completely independent, parallel image reconstruction methods are generally limited to low acceleration factors (R = 2 or R = 3) (7,8). By also taking advantage of correlations in time, various "k-t" methods, including k-t BLAST/SENSE and k-t GRAPPA, have been able to reconstruct sparse measurements with higher acceleration factors (9–11). These k-t methods can achieve higher acceleration factors, but they can suffer from temporal blurring and often require additional training data.

Image reconstruction methods which rely on correlations in time to fill in missing *k*-space data are based on the assumption that image space varies slowly in time. When this assumption is accurate, and only stationary or slowly varying tissues are imaged, high acceleration factors can be achieved without negative consequences. When, however, this assumption is not correct and quickly varying tissues are being imaged, high acceleration factors can lead to temporal blurring, even though much of the undersampling artifact (in x and y) may be removed.

TCR

Recently, the use of temporally constrained reconstruction (TCR) for reconstructing highly undersampled cine data has been demonstrated. TCR is an iterative technique that generates an image estimate based on the minimization of a cost functional containing a fidelity term and a temporal constraint term. The temporal constraint is tunable and can be adjusted to minimize temporal blurring. Additionally, the temporal constraint can be spatially dependent, allowing stationary regions of the image to be more temporally constrained than nonstationary regions, improving the results. Adluru et al (12,13) recently applied TCR to highly undersampled cardiac perfusion data. TCR is a promising technique for accelerating PC imaging, but its ability to reconstruct PC data has not yet been evaluated.

Objectives

The purpose of this study was to: 1) compare velocity measurements reconstructed with TCR to simple temporal interpolation of *k*-space, using retrospectively undersampled simulated data; 2) compare image quality and peak systolic velocity (PSV) measurements reconstructed with TCR to fully sampled in vivo data and Doppler ultrasound; and 3) determine if TCR can be used to reduce the acquisition time of a high-resolution cine-PC sequence enough to be performed during a short breath-hold, while still resulting in accurate velocity measurements.

MATERIALS AND METHODS

MRI Studies

Data Acquisition

All data were acquired using a cine radial spoiled gradient echo PC pulse sequence. Radial data were readout in a symmetric fashion with the center of each readout positioned at the center of *k*-space. Radial views were distributed equally over 180° for each cardiac phase. The acquisition pattern was offset between cardiac phases so that sampling patterns for neighboring phases overlapped as little as possible while still ensuring maximal combined *k*-space coverage across all cardiac phases. The acquisition pattern was rotated between cardiac phases according to Eq. (1) where $\varphi_{p,even}$ is the sampling pattern rotation for even-numbered cardiac phases and $\varphi_{p,odd}$ is the rotation for odd-numbered cardiac phases, p is the cardiac phase number, and n_v and n_p are the total number of radial views and cardiac phases, respectively:

$$\begin{aligned}\varphi_{p,even} &= \frac{p}{2} * \frac{\pi}{n_v n_p} \\ \varphi_{p,odd} &= \left[\frac{p-1}{2} + ceil\left(\frac{n_p}{2}\right)\right] * \frac{\pi}{n_v n_p}\end{aligned} \quad (1)$$

Three types of datasets were collected. First, three highly sampled datasets (128 views) were acquired from three different subjects and retrospectively undersampled. Peak systolic velocity (PSV) measurements from the original data were compared to PSV measurements from the undersampled data, reconstructed with temporal interpolation and TCR. Second, two in vivo datasets were acquired from the same subject, a highly sampled set with 192 views per cardiac phase (acquisition time: 70 sec) that was reconstructed with standard methods (called Standard-192), and a prospectively undersampled dataset with 16 views per cardiac phase (acquisition time: 13 sec) that was reconstructed with TCR (called TCR-16). PSV measurements from Standard-192 and TCR-16 were compared to the PSV measurement obtained by echocardiogram. Third, in five normal volunteers a high-temporal-resolution undersampled (16 radial views) TCR reconstruction was compared to a separate low-resolution more fully sampled (128 radial views) zero-filled regridded reconstruction. Volunteers were asked to hold their breath during each acquisition. The undersampled dataset obtained 16 radial views per cardiac phase, took between 15 and 19 seconds, depending on heart rate, and the temporal resolution was 19 msec. The more fully sampled dataset obtained 128 radial views per cardiac phase, took 24 to 27 seconds, depending on heart rate, and the temporal resolution was 79 msec. One subject was excluded due to a physiological finding during the exam that compromised the repeatability of the flow measurements and another subject was excluded due to poor image quality.

With informed consent and an Institutional Review Board (IRB)-approved protocol, all subjects were imaged on either a Siemens Magnetom Verio 3T or a Magnetom Trio Tim 3T MRI scanner (Siemens Healthcare, Erlangen, Germany), using a standard Siemens chest/spine coil, the scanner automatically selected the

optimal channels that were combined into five virtual channels. A single axial slice, positioned at the base of the ascending aorta, was acquired and velocity was measured in the through-plane direction. The acquisition matrix was 192 × 192 with a voxel size of 1.56 × 1.56 mm^2 and a slice thickness of 5 mm. Other parameters were TR/TE/α = 24.0msec/4.0msec/15° on the Verio and TR/TE/α = 36.2msec/6.7msec/15°, with velocity encoding = 150 cm/s being used on both. The sequence was gated to the cardiac cycle using either electrocardiogram (ECG) leads or a pulse oximeter placed on the subject's finger.

Simulated, Undersampled Data

To generate new radial measurements, for the retrospectively undersampled data, the fully sampled data were first interpolated to the Cartesian coordinate system and transformed to the image domain. Then a complex radon transform, followed by a transformation back to *k*-space, was used to generate new radial measurements that were then resampled to 192, 32, 16, and 8 radial views. This was done so that arbitrary view angles could be generated in order to conform to the undersampling scheme used.

TCR Implementation

Interpolation Onto Cartesian Grid

Initially, all radial data were regridded onto a double sized Cartesian grid that was subsequently convolved with bilinear interpolation gridding kernel and then re-sampled to the original matrix size. No post-apodization correction was performed as the phase-difference velocity images were of primary interest and they are not affected by roll-off like the magnitude images. However, the roll-off artifact in the magnitude images was not enough to be noticeable, especially after readout oversampling was removed. Data remained in the Cartesian coordinate system for all subsequent operations.

Cost Functional

The TCR cost functional, $C(\tilde{m})$, used in this article, is shown in Eq. (2):

$$C(\tilde{m}) = \| \sum_{c=1}^{coils} W \cdot \mathcal{F} S_c \cdot \tilde{m} - d_c \|_2^2 + \lambda \| \nabla_t \tilde{m} \|_1^1 \quad (2)$$

where W is the undersampling pattern, S is the relative coil sensitivity, $\mathcal{F}$ is the 2D Fourier transform, and $\tilde{m}$ is the image estimate, λ is a tunable parameter, which can also depend on position in the image, d_c is the original measurement data for receiver channel c, $\cdot$ is the piecewise multiplication operator, and ∇_t is the temporal gradient. An L_1 norm was used for the temporal constraint instead of an L_2 norm in order to preserve sharp changes in the temporal velocity profile, or, in other words, to reduce temporal blurring.

Coil Sensitivity Estimation

A single set of coil sensitivity estimates was calculated and used for reconstructing each individual cardiac phase. Assuming that signal variations through the cardiac cycle are relatively small, the *k*-space measurements in a single dataset were averaged across all cardiac phases to get a more complete set of *k*-space. Then, these data were inverse-Fourier transformed to the image domain and sensitivity estimates were calculated by dividing individual coil signals by the sum-of-squares combination of all coil signals, as described in Eq. (3), where $\bar{m}_c$ is the image data reconstructed from the time-averaged *k*-space data for the channel *c*:

$$S_c = \frac{\bar{m}_c}{\sqrt{\sum_{\hat{c}=1}^{colis} |\bar{m}\hat{c}|^2}} \quad (3)$$

Initial Image Estimate

As an intermediate step toward generating the initial image estimate ($\tilde{m}_0$), the measured *k*-space data (d) was temporally interpolated to fill in missing values, and then inverse-Fourier transformed to the image domain to form a vector of images, $\hat{m}$, one for each receiver coil. Due to the inclusion of the coil sensitivity term in the TCR reconstruction, $\tilde{m}_0$ could not have a channel dimension, so the channels in $\hat{m}$ needed to be combined. Image phase is essential to the PC reconstruction so the combination method chosen to form $\tilde{m}_0$ from $\hat{m}$ needed to preserve the phase difference between the reference and velocity-encoded (VE) sets. The phase difference ($\hat{\Delta\varphi}$) was calculated from $\hat{m}$ using the two-point method, where the velocity is determined to be proportional to the phase difference of the reference and VE sets, and the phase differences from all the input channels were combined using the weighted-sum method described in Eq. (6) to yield $\tilde{\Delta\varphi}$. Finally, the magnitudes of the reference and VE sets of $\tilde{m}_0$ were set to the sum-of-squares combination of the reference and VE sets of $\hat{m}$. The VE set of $\tilde{m}_0$ was multiplied by $e^{i2\pi\tilde{\Delta\varphi}}$ to restore the phase difference between the two sets. This resulted in an initial image estimate that had a reference set with zero phase and a VE set with a phase corresponding to the calculated phase difference of $\hat{m}$:

$$\tilde{m}_r = \sqrt{\sum_{c=1}^{colis} |\hat{m}_{r,c}|^2} \quad (4)$$

$$\tilde{m}_v = e^{i2\pi\tilde{\Delta\varphi}} \sqrt{\sum_{c=1}^{colis} |\hat{m}_{v,c}|^2} \quad (5)$$

PC Channel Combination

Multichannel phase-contrast data cannot be combined using the standard sum-of-squares approach due to the importance of the image phase. For this work, the weighted sum method was used (14). The phase-difference was computed using the two-point method for each channel individually, and then phase differences were combined as a weighted sum with weights corresponding to the square of the magnitude image (m_i) divided by the variance of the coil noise (σ_i):

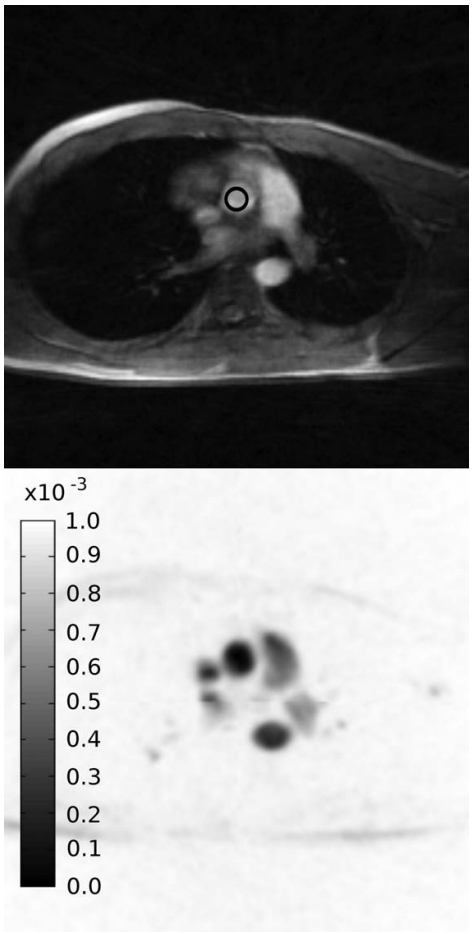

Figure 1. Magnitude image (top) and the corresponding spatial lambda map (bottom). Dark areas in the map are less temporally constrained. The ROI used to measure velocities is indicated by the black circle in the magnitude image.

$$\Delta\varphi_{comb} = \frac{\sum \Delta\varphi_i \cdot |m_i/\sigma_i|^2}{\sum |m_i/\sigma_i|^2} \tag{6}$$

Spatially Varying Temporal Constraint (λ)

A spatially varying temporal constraint weight was used to constrain areas with low or no flow more than areas with flow. An initial map (Λ) was derived from the initial image estimate $\tilde{m}_0$ by computing the magnitude of the complex difference of the VE and reference sets, of the initial image estimate, $\tilde{m}_0$, and summing in the cardiac phase dimension. This resulted in a map with high values in locations with high flow and low values where little or no flow existed. This map was then inverted and normalized to a prespecified λ_{max} to get λ. An example lambda map is shown in Fig. 1. The value of λ_{max} used in this work was determined empirically by reconstructing a retrospectively undersampled dataset with a range of different λ_{max} values and selecting the value that minimized the RMSE of the max velocity curve with respect to the fully sampled data:

$$\Lambda = \sum_{p=1}^{phases} |\tilde{m}_{0_{r,p}} - \tilde{m}_{0_{v,p}}| \tag{7}$$

$$\lambda = \lambda_{\max}\left(1 - \frac{\Lambda}{\max(\Lambda)}\right) \tag{8}$$

Nonlinear Gradient Descent Minimization

The cost functional (Eq. (2)) was minimized by an iterative gradient descent algorithm. The gradient of the cost functional (Eq. (3)) was calculated at each iteration and subtracted from the current image estimate; the result was used as the image estimate in the next iteration. The L_1 norm is not differentiable at zero so in order to calculate the gradient an approximation of the L_1 norm was used instead, that included a small offset (β^2) to make it differentiable at zero:

$$\nabla C(\tilde{m}) = 2\sum_{c=1}^{colis}(S_c * \cdot\mathcal{F}^{-1}\{W \cdot \mathcal{F}S_c \cdot \tilde{m} - d_c\}) + \lambda\nabla_t\left(\frac{\nabla_t\tilde{m}}{\sqrt{\nabla_t\tilde{m}^2 + \beta^2}}\right) \tag{9}$$

Analysis

For the simulated data, velocities were measured inside a circular region of interest (ROI) inside the proximal ascending aorta (Fig. 1). The position of the ROI was adjusted for each cardiac phase so that it remained at the center of the aorta. These ROIs were used to calculate PSV measurements, generate velocity curves, and correlation plots in a pixel-wise fashion over all cardiac time frames. For the in vivo data, images were reconstructed and visually inspected to assess the difference in image quality. Additionally, an ROI was used, as with the simulated data, to calculate PSV measurements that were compared to Doppler ultrasound.

RESULTS

Simulated Data

Max and mean velocity curves for one simulated dataset are shown in Fig. 2. At 32 radial views the velocity curves from both temporal interpolation and TCR match closely with the reference curve. Temporal blurring is evident in the temporal interpolation curves; this blurring gets worse as the number of radial views is decreased. Table 1 outlines the PSV values and percent errors computed for each of the simulated datasets with different undersampling factors. TCR resulted in PSV values with small error even at high undersampling factors while temporal interpolation resulted in more inaccurate PSV values

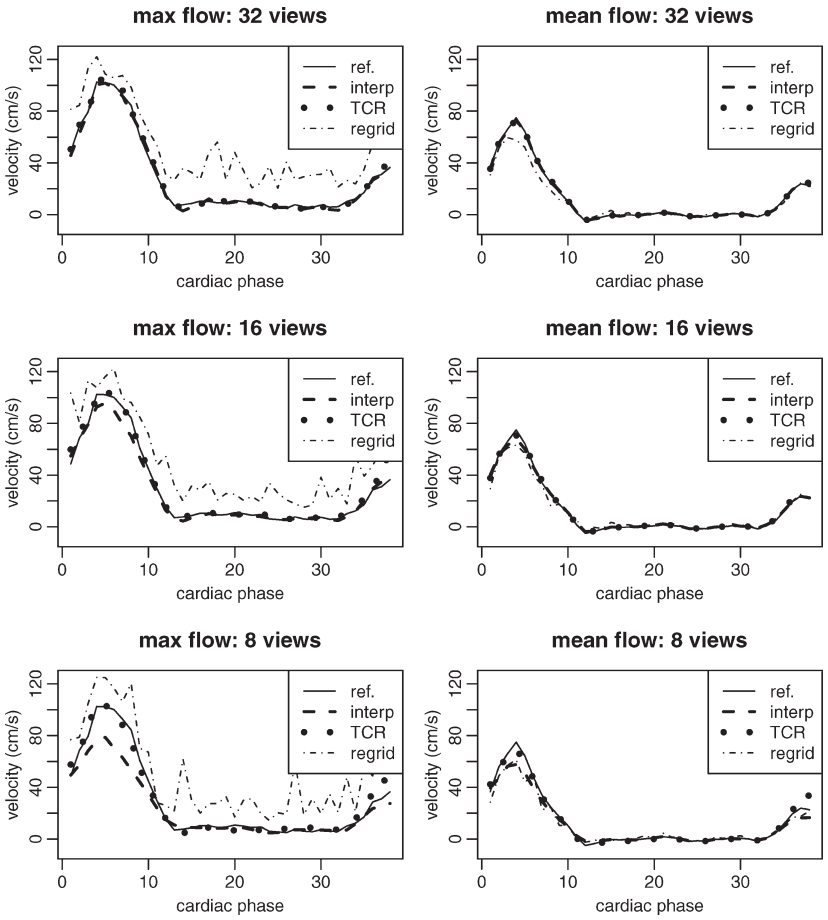

Figure 2. Velocity curves for one of the simulated datasets. Temporal interpolation results in significant temporal blurring, especially at higher acceleration factors.

as the number of radial views was decreased. Temporal interpolation, however, did produce PSV values with small errors at 32 radial views. Simple zero-filled regridding resulted in the lowest quality velocity curves due to significant streaking artifact and increased noise. A correlation plot from one of the simulated datasets, with different acceleration factors, is shown in Fig. 3 and correlation parameters are shown in Table 2. The plots suggest that at 32 radial views both TCR and temporal interpolation correlate well with the reference data. As the number of radial views is decreased, temporal interpolation shows signs of temporal blurring and high velocities are underestimated, while low velocities are overestimated.

In Vivo Patient Data

Magnitude and phase reconstructions for a single phase of TCR-16 are compared to the reconstructions from Standard-192 in Fig. 5. Even though the Standard-192 data were reconstructed from more views, the

Table 1
PSV Measurements (in cm/s) of the Simulated Data With Varying Levels of Undersampling Compared to the Fully Sampled Reference Measurements

	SIM3 (102.5)	SIM1 (90.7)	SIM2 (69.0)
Views	TCR (err%) / interp (err%)	TCR (err%) / interp (err%)	TCR (err%) / interp (err%)
32	106.0 (3.4%) / 101.7 (0.8%)	88.5 (2.5%) / 86.8 (4.4%)	71.7 (3.9%) / 66.9 (3.1%)
16	106.0 (3.4%) / 101.7 (0.8%)	87.6 (3.5%) / 78.8 (13.2%)	66.4 (3.9%) / 61.9 (10.4%)
8	109.5 (6.9%) / 78.6 (23.3%)	87.8 (3.2%) / 60.2 (33.7%)	66.6 (3.5%) / 52.4 (24.1%)

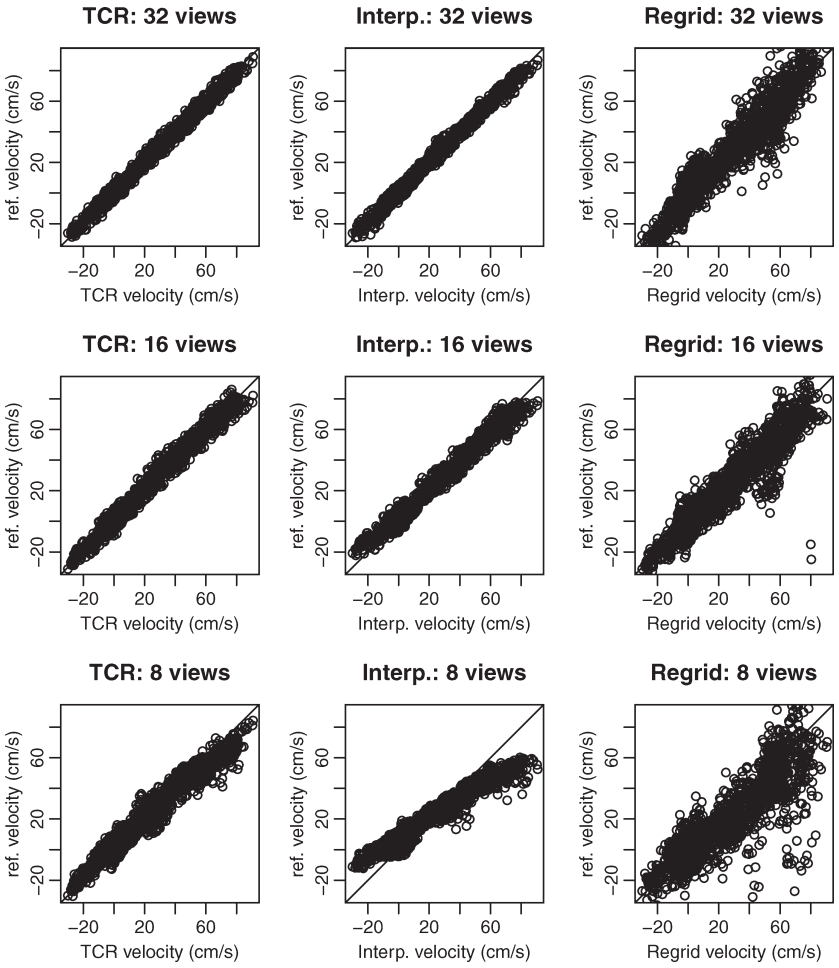

Figure 3. Correlation plots for one simulated dataset undersampled to 32, 16, and 8 radial views. Plots for TCR are shown on the left, plots for temporal interpolation are shown in the middle, and plots for zero-filled regridding are shown on the right. TCR maintains better correlation than temporal interpolation as the number of radial views is decreased, both show better correlation than simple regridding.

image quality is much poorer than the TCR-16 data. The subject imaged had difficulty holding still and holding his breath, which was the likely cause for the difference in image quality. Motion artifacts are evident in the Standard-192 data but the TCR-16 images are much clearer. The Standard-192 PC images show signs of velocity underestimation in the ascending and descending aorta. An echocardiogram indicated that the PSV in the ascending aorta was 125 cm/s. TCR-16 indicated a PSV of 121 cm/s (3.3% error) while Standard-192 indicated a PSV of 107 cm/s (17% error).

In Vivo Healthy Volunteer Data

Velocity histogram plots (Fig. 4) were used to compare the undersampled TCR data with the more fully sampled, regridded zero-filled data. The difference in temporal resolution is obvious, but both TCR and regridding indicate similar PSV values. TCR data were acquired with only 16 radial views as opposed to the regridded data, which were acquired with only 128 radial views, at a lower temporal resolution, which was necessary to allow for a breath-hold scan time around 30 seconds.

DISCUSSION

The purpose of this study was to assess the feasibility of using TCR to reconstruct highly undersampled cine-PC data. Using an appropriate radial view-order sampling pattern, TCR was applied to both simulated and in vivo data and the results were compared to a

Table 2
Correlation Parameters Correlation Coefficient R^2 for Correlation Plots Shown in Fig. 3

	Slope, intercept		R^2	
Dataset	Interp.	TCR	Interp.	TCR
SIM1: 32 v.	1.00, - 0.0013	0.99, 0.0023	0.9925	0.9884
16 v.	0.97, 0.0070	0.99, 0.0019	0.9783	0.9798
8 v.	0.79, 0.0455	0.94, 0.0169	0.9457	0.9541
SIM2: 32 v.	0.98, 0.0065	0.98, 0.0100	0.9916	0.9896
16 v.	0.94, 0.0205	0.99, 0.0063	0.9758	0.9808
8 v.	0.69, 0.0940	0.88, 0.0319	0.9384	0.9586
SIM3: 32 v.	0.98, 0.0043	0.99, 0.0050	0.9856	0.9831
16 v.	0.92, 0.0232	0.98, 0.0145	0.9736	0.9731
8 v.	0.73, 0.0299	0.93, 0.0144	0.9307	0.9428

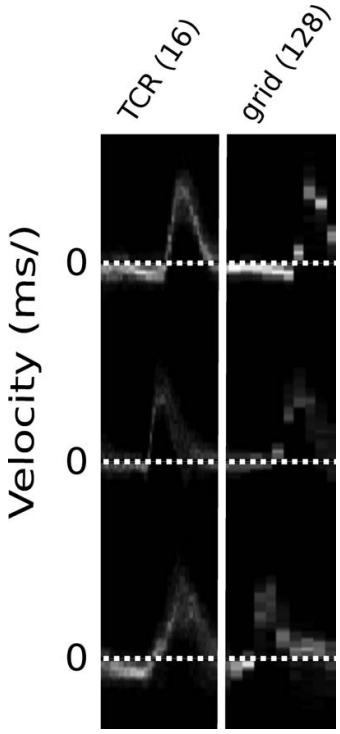

Figure 4. Velocity histogram plots for three healthy volunteers from a ROI at the center of the proximal ascending aorta. High-temporal-resolution TCR (left) is compared to low-temporal-resolution gridding (right).

simple temporal interpolation of *k*-space. PSV measurements from the in vivo MRI data were compared to Doppler ultrasound.

Results from the three simulated datasets showed that TCR produced more accurate velocity curves than temporal interpolation, which exhibited significant temporal blurring at high acceleration factors. However, with 32 views per cardiac phase, the performance of temporal interpolation was similar to TCR. This is most likely due to the high temporal resolution of the simulated data. Other experiments showed that temporal interpolation results in more significant temporal blurring when the temporal resolution of the measurement data is reduced. With 16 and 8 radial views per cardiac phase, TCR outperformed temporal interpolation, resulting in more accurate velocity curves and PSV measurements. With 8 radial views, PSV measurements from TCR remained within 3%–7% of the reference PSV, but temporal interpolation produced much larger PSV errors within the range of 23%–34%.

Results from the in vivo data suggest that TCR of undersampled data can produce better images and PSV measurements than standard reconstruction of more fully sampled data when subject motion, respiratory or otherwise, is a problem. TCR resulted in fewer motion artifacts and measured a PSV that was within 3.3% of the PSV measured by Doppler ultrasound. The highly sampled in vivo dataset was acquired with a large number of radial views (192) in order to avoid undersampling artifacts. This subject had difficulty holding still, and moved enough during the longer Standard-192 scan time to cause significant motion artifacts in the reconstructed images. However, the TCR-16 acquisition was short enough to be acquired without too much subject motion and the resulting images were free from motion artifacts. Image quality is highly dependent on subject compliance, and long scan times do not necessarily result in motion artifacts, but a decreased scan time lessens the risk of motion artifacts. In this instance, TCR allowed the scan time to be reduced to 13 seconds, which is within the duration of a breath-hold for most subjects.

Temporal blurring is a potential issue for any reconstruction technique that relies on signal correlations in time to reconstruct undersampled data. TCR has several advantages that allow it to reconstruct data without significant temporal blurring. Coil sensitivity information was incorporated into the TCR fidelity term, adding a parallel imaging aspect allowing for the filling of missing *k*-space values based on the coil sensitivities. With the inclusion of the coil sensitivity information in the fidelity term, TCR is essentially an iterative SENSE reconstruction (15) with an added temporal constraint. Additionally, the temporal constraint in TCR is tunable and spatially varying, with areas of flowing blood being less temporally constrained than static tissues.

A relatively small number of receiver channels was used in this work. The coils used in these experiments can be set to output uncombined data from each

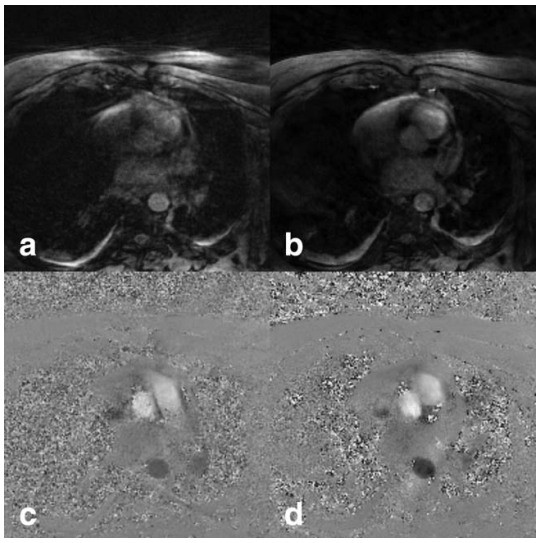

Figure 5. Magnitude and PC reconstructions of in vivo data. The 192 view data reconstructed using standard methods is shown in (**a**,**c**). The undersampled 16 view data reconstructed with TCR is shown in (**b**,**d**). Motion artifacts, most likely due to respiration, are evident in the 192 view data and the image quality of the 16 view data is much better.

channel, resulting in 3 times more receiver channels. Increasing the number of channels would likely improve TCR image reconstruction, but this has not yet been explored. Additionally, a relatively simple method for determining coil sensitivities was used. A different method may have produced more accurate sensitivity maps and improved the TCR reconstructions.

A radial acquisition scheme was used for two primary reasons. First, the PSV of a radial sampling pattern spreads artifact in multiple directions, creating less coherent artifacts. These less coherent artifacts can be more effectively resolved by TCR than the aliased image replicates created by Cartesian undersampling. Second, a radial acquisition oversamples the center of *k*-space, ensuring that the undersampled data has increased sampling density at the center of *k*-space. The sampling pattern selected was used so that aliasing patterns would be different between cardiac phases, making the data better conditioned for correction by a temporal constraint. Additionally, this ensured that there would be enough information to generate good coil sensitivity maps, as these were generated from the combined data of all cardiac phases. Radial data were transformed to the Cartesian coordinate system as a first step. An alternative to this approach would be to use a nonuniform fast Fourier transform (NUFFT) to transform the data from radial to Cartesian at each iteration. However, NUFFT is more computationally expensive than a traditional FFT because it performs a convolution step in addition to an FFT step. If NUFFT were used instead of pregridding of the initial data, an NUFFT and an inverse NUFFT would need to be performed at each iteration instead of an FFT and an IFFT and this would have added a significant amount of computation time.

In conclusion, TCR was able to reconstruct highly undersampled cine-PC data. Results from simulated and in vivo data indicate that TCR produced more accurate velocity measurements than temporal interpolation and zero-filled regridding and has the ability to improve image quality over more highly sampled data when subject compliance is an issue. TCR allowed for a single-slice single-velocity cine-PC acquisition in 13 seconds that resulted in a measured PSV similar to the PSV measured by Doppler ultrasound.

REFRENCES

1. Meier D, Maier S, Bosiger P. Quantitative flow measurements on phantoms and on blood vessels with MR. Magn Reson Med 1988;8:25–34.
2. Firmin DN, Nayler GL, Klipstein RH, Underwood SR, Rees RS, Longmore DB. In vivo validation of MR velocity imaging. J Comput Assist Tomogr 1987;11:751–756.
3. Walker PG, Oyre S, Pedersen EM, Houlind K, Guenet FS, Yoganathan AP. A new control volume method for calculating valvular regurgitation. Circulation 1995;92:579–586.
4. Sakuma H, Blake LM, Amidon TM, et al. Coronary flow reserve: noninvasive measurement in humans with breath-hold velocity-encoded cine MR imaging. Radiology 1996;198:745–750.
5. Ehman RL, Felmlee JP. Adaptive technique for high-definition MR imaging of moving structures. Radiology 1989;173:255–263.
6. Nagel E, Bornstedt A, Hug J, Schnackenburg B, Wellnhofer E, Fleck E. Noninvasive determination of coronary blood flow velocity with magnetic resonance imaging: comparison of breath-hold and navigator techniques with intravascular ultrasound. Magn Reson Med 1999;41:544–549.
7. Peng HH, Bauer S, Huang TY, et al. Optimized parallel imaging for dynamic PC-MRI with multidirectional velocity encoding. Magn Reson Med 2010;64:472–480.
8. Markl M, Kilner PJ, Ebbers T. Comprehensive 4D velocity mapping of the heart and great vessels by cardiovascular magnetic resonance. J Cardiovasc Magn Reson 2011;13:7.
9. Baltes C, Kozerke S, Hansen MS, Pruessmann KP, Tsao J, Boesiger P. Accelerating cine phase-contrast flow measurements using k-t BLAST and k-t SENSE. Magn Reson Med 2005;54:1430–1438.
10. Stadlbauer A, van der Riet W, Globits S, Crelier G, Salomonowitz E. Accelerated phase-contrast MR imaging: comparison of k-t BLAST with SENSE and Doppler ultrasound for velocity and flow measurements in the aorta. J Magn Reson Imaging 2009;29:817–824.
11. Jung B, Honal M, Ullmann P, Hennig J, Markl M. Highly k-t-space-accelerated phase-contrast MRI. Magn Reson Med 2008;60:1169–1177.
12. Adluru G, Awate SP, Tasdizen T, Whitaker RT, Dibella EV. Temporally constrained reconstruction of dynamic cardiac perfusion MRI. Magn Reson Med 2007;57:1027–1036.
13. Adluru G, McGann C, Speier P, Kholmovski EG, Shaaban A, Dibella EV. Acquisition and reconstruction of undersampled radial data for myocardial perfusion magnetic resonance imaging. J Magn Reson Imaging 2009;29:466–473.
14. Thunberg P, Karlsson M, Wigstrom L. Comparison of different methods for combining phase-contrast images obtained with multiple coils. Magn Reson Imaging 2005;23:795–799.
15. Pruessmann KP, Weiger M, Bornert P, Boesiger P. Advances in sensitivity encoding with arbitrary k-space trajectories. Magn Reson Med 2001;46:638–651.

CHAPTER 3

UNDER-SAMPLED GOLDEN RATIO CINE PHASE-CONTRAST MRI USING RETROSPECTIVE DISTANCE MATRIX SELF-GATING

3.1 Abstract

Cine phase-contrast imaging is a useful clinical technique for measuring in-vivo blood velocities. Cine acquisitions require a gating signal to determine which data belong to which cardiac phase bin. Traditional ECG gating is complicated by several common challenges, including poor ECG signal and signal interference caused by the MRI acquisition. In this work, a new distance matrix based self-gating technique for phase-contrast MRI is proposed. Three volunteers were imaged using an under-sampled radial phase-contrast sequence that was reconstructed with temporally constrained reconstruction. Traditional Cartesian ECG-gated phase-contrast data sets were also acquired for comparison. The self-gating signal was verified against the measured ECG signal and images for the new technique were compared to the standard Cartesian ECG-gated images. The new self-gating technique, in conjunction with the under-sampled radial acquisition and temporally constrained reconstruction, produced images as good as or better than traditional ECG-gated Cartesian phase-contrast while allowing for more flexible selection of temporal resolution and discarding of motion corrupted data.

3.2 Introduction

Cine phase-contrast (PC) MRI is a powerful clinical technique that allows accurate and noninvasive evaluation of cardiovascular blood flow without ionizing radiation or contrast. Cine PC MRI has been widely used in the assessment of congenital and valvular heart disease in particular [1, 2, 3]. Nonetheless, current techniques have limitations. Although breath-held cine PC MRIs minimize respiratory artifact, maintaining a short acquisition

time with current techniques constrains both spatial and temporal resolution. As a result, flow in small moving vessels and peak velocity in stenotic valves may be incorrectly assessed [3, 4, 5, 6]. As a result, cine PC techniques that allow for breath-held acquisitions without sacrificing temporal or spatial resolution are desirable.

Retrospective gating of cine PC data is traditionally accomplished by binning measurements collected over several heart cycles according to their temporal distance from the most recent ECG R-wave peak. Unfortunately, there are several challenges associated with ECG gating, such as signal interference from rapidly switching magnetic field gradients, signal interference due to the additional magnetic fields created from blood movement in the main magnetic field (the magnetohydrodynamic effect), and the increased exam time required for the proper application of ECG leads [7, 8, 9]. Gating based on a poor quality ECG signal can lead to image artifacts caused by individual measurements being binned into the incorrect cardiac phase.

In order to avoid the problems inherent in ECG gating, several self-gating techniques have been developed that derive the gating signal from the MR measurements instead of from an ECG waveform. Commonly, these techniques compute the self-gating signal from measurements regularly acquired at the center of k-space [10, 11, 12]. When a radial sampling pattern is used, no additional measurements need be acquired as the center of k-space is measured with each radial view [13]. For phase-contrast (PC) acquisitions, a complex difference of the reference and velocity encoded sets can be used for flow-gating cine data [14]. In addition to cardiac gating, self-gating techniques have been used for respiratory gating. In these cases, low frequency variations in the self-gating signal can be correlated with respiration.

Unfortunately, the nature of the self-gating signal can also vary significantly between exams and different anatomical locations, making robust, automated self-gating difficult. Deriving the self-gating signal from a smaller region of interest, selected through user intervention, and/or autocorrelation of the self-gating signal have been shown to improve reliability [14].

The objective of this work is to demonstrate a new PC imaging technique that is self-gated, can be acquired quickly, during a breath hold, and allows for detection of motion-corrupted data. A new self-gating technique utilizing distance matrices is presented using two different distance metrics that allow for flow gating of the PC signal and motion detection. Temporally constrained reconstruction is used to reconstruct under-sampled data. The technique is applied to three different anatomical locations: the proximal

ascending aorta, the pulmonary valve, and the coronary sinus. Standard Cartesian PC imaging is used as a gold standard for comparison.

3.3 Methods

3.3.1 Pulse Sequences and Acquisition Protocols

All PC imaging data were obtained using a retrospectively gated spoiled gradient echo cine phase-contrast pulse sequence. Velocity was measured in the through-plane direction; flow-compensated and velocity encoded measurements were acquired in an interleaved fashion. Separate Cartesian and radially-sampled data sets were acquired. Radial data were acquired by setting the view angle for each radial view to an ever increasing multiple of an angle based on the golden ratio [15] according to equation 3.1, where θ is the view angle and l is the sequential radial view number.

$$\theta = l\pi\frac{\sqrt{5}-1}{2} \tag{3.1}$$

The total number of lines acquired for the radial sequence is an adjustable parameter that determines total scan time. Cartesian data were acquired using the standard Siemens retrospective ECG-gated sequence with IPAT=2.

With informed consent and IRB approval, 3 volunteers were imaged on a Siemens Magnetom TIM Trio, or Siemens Magnetom Verio (Siemens, Erlangen Germany). In order to assess the new self-gating technique in different anatomies, three different slice positions were used: an axial slice through the ascending aorta at the level of the aortic valve, an oblique slice just above the pulmonary valve, and an oblique slice through the coronary sinus. The following sequence parameters were used for both the Cartesian and radial sequences: voxel size=0.83 x 0.83 mm^2 - 0.92 x 0.92 mm^2, slice thickness=5 mm, VENC=70 cm/s (coronary sinus) and 150 cm/s (pulmonary valve/ascending aorta). ECG leads were used to reconstruct the ECG-gated Cartesian data. Subjects were instructed to perform an exhaled breath hold during both the Cartesian and radial acquisitions. All radial acquisitions were set to measure 1200 radial views, resulting in an acquisition time of 14 seconds, Cartesian acquisitions acquired as many measurements as necessary to obtain Cartesian k-space for the set temporal resolution (40 ms), resulting in an acquisition time of 21 seconds.

3.3.2 Derivation of Self-Gating Signal

The new self-gating method described in this work is similar to the low-resolution ROI correlation method described by Larson et al. [13] because it relies on a comparison between

individual time frames of a low spatial resolution real-time reconstruction of the radial data. The self-gating signal is derived from a distance matrix calculated using a predefined distance metric between each pair of real-time frames.

The low temporal-resolution real-time image sequence was generated for each receiver channel by combining every 4 subsequent k-space radial views into a single temporal frame, cropping the high spatial frequency k-space components (in this work a square of 40 x 40 at the center k-space was retained), and filling in missing k-space data through linear temporal interpolation of adjacent temporal frames in a sliding-window fashion. The real-time image sequence was transformed to image space and the 2x oversampling from the measurement data was removed.

Distance matrices were computed for each receiver channel by measuring the distance, using one of two metrics, between each pair-wise combination of temporal frames from the real-time reconstruction. The first metric ($\Delta\phi_1$, equation 3.2) used was designed to identify differences between temporal frames of PC data and relies on the existence of separate reference and velocity encoded sets. This metric computes the complex difference between the reference and velocity encoded sets for each temporal frame, then computes the difference in this quantity between frames. This complex difference is not the same as computing the velocity data for the set which is done by taking the difference in the angle of the two complex sets. Taking the magnitude of complex difference eliminates signal from static tissue, improving correlation between like frames. The second metric ($\Delta\phi_2$, equation 3.3) does not rely on the velocity encoded set so it could potentially be used with non-PC data and was designed to detect respiratory motion in addition to providing the self-gating signal.

Individual receiver distance matrices were combined into a single matrix as the square root of the sum of the squared magnitudes of individual receiver matrices (equation 3.4). Metrics using only magnitude information could have been calculated on a channel-combined version of the real-time image sequence, but were computed this way to maintain consistency with other metrics requiring complex phase information. This combined matrix contains a periodic modulation due to image differences caused by the cardiac cycle, and may also contain variations due to respiration or other motion.

The self-gating signal was derived from the distance matrices in one of two ways. First it was derived by taking a projection of either the row or column direction; since the matrix is symmetric, both operations are equivalent. This method was used for matrices generated using metric 1 as it allowed for the identification of systolic and diastolic periods, as will

be discussed later. The second way in which self-gating signals were derived was using a template matching approach. The first 15 columns (enough columns to contain 1-2 heart cycles) of the distance matrix were convolved across the matrix in the column direction, resulting in a signal with peaks where the matrix correlated highly with the first 15 columns. This method was used for matrices generated using metric 2 because these matrices do not exhibit as regular a pattern as matrices using metric 1. Gating triggers were detected from the peaks in either self-gating signal by finding the local maxima using a sliding 10 element window.

$$\Delta\phi_1(\hat{m}_j, \hat{m}_k) = \|abs\left[(\hat{m}_{j,comp} - \hat{m}_{j,venc}) - (\hat{m}_{k,comp} - \hat{m}_{k,venc})\right]\|_2^2 \tag{3.2}$$

$$\Delta\phi_2(\hat{m}_j, \hat{m}_k) = \|abs\left[(\hat{m}_{j,comp} - \hat{m}_{k,comp})\right]\|_2^2 \tag{3.3}$$

$$D_{j,k} = \sqrt{\sum_{c=1}^{n-1} \Delta\theta(\hat{m}_{j,c}, \hat{m}_{k,c})^2} \tag{3.4}$$

3.3.3 Temporally Constrained Reconstruction

Using the aforementioned parameters (1200 radial views and a subset of 600 radial views and a temporal resolution of 10 ms) resulted in under-sampled k-space data in each cine bin, which was reconstructed using temporally constrained reconstruction (TCR) [16, 17, 18]. TCR reconstructs under-sampled k-space data by iteratively minimizing a cost functional involving a data fidelity term and a temporal constraint term. For this work, the following cost functional (equation 3.5), that includes an additional coil sensitivity constraint, was minimized using iterative gradient descent.

$$C(m) = \left\| \sum_{c=0}^{coils} W \cdot \mathcal{F} S_c \cdot \tilde{m} - d_c \right\|_2^2 + \lambda \left\| \nabla_t \tilde{m} \right\|_1^1 \tag{3.5}$$

W is the under-sampling mask, $\mathcal{F}$ is the 2D FFT, S_c is the complex receiver coil sensitivity for channel c, $\tilde{m}$ is the current image estimate (composed of both reference and velocity encoded sets, all receiver channels, and all cardiac phases), d is the original measurement data, and α is a tunable parameter for the temporal constraint. $\cdot$ denotes element-wise multiplication (not matrix multiplication), Δ_t is the temporal gradient operator, $\|\cdot\|_2$ is the l2 norm, and $\|\cdot\|_1$ is the l1 norm. The gradient of the l1 norm was approximated in

the same way as in [17]. Radial measurement data were pre-interpolated to a Cartesian grid and remained in the Cartesian coordinate system for the remainder of the algorithm. 1000 gradient descent iterations were performed in order to insure proper convergence of the image estimate. Radial measurement data were pre-interpolated to a Cartesian grid and remained in the Cartesian coordinate system for the remainder of the algorithm. 1000 gradient descent iterations were performed in order to insure proper convergence of the image estimate.

3.4 Results

3.4.1 Distance Matricies and Self-Gating Signal

A distance matrix computed from radial PC data obtained at the coronary sinus is shown in Figure 3.1 using metric 1 (equation 3.3). The diagonal identity line is zero, indicating that each time frame is no different from itself. Moving away from the diagonal, a periodic pattern of high and low difference measurements creates a checkerboard type pattern. Time frames during systole show low difference with a few temporally adjacent frames but a longer period of high difference as diastole occurs. Time frames during diastole show the reverse, with a longer period of low differences adjacent and a short period of high difference during systole. The ECG waveform was aligned with the self-gating signal by matching the timestamps associated with each ECG datum with the timestamps of individual k-space measurements. The self-gating signal is shown above the matrix along with the aligned ECG waveform. The self-gating signal correlates well with the ECG.

Distance matrices using both metrics and at three different anatomical locations (proximal ascending aorta, pulmonary valve, and coronary sinus) are shown in Figure 3.2. The same checkerboard pattern appears at each location, allowing for identification of systolic and diastolic cardiac phases using metric 1. The self-gated signal shown for distance matrices using metric 2 demonstrates that a self-gating signal can be obtained without any flow information. Figure 3.3 shows the same variations for the cardiac cycle as well as lower frequency variations caused by subject motion.

3.4.2 Reconstructed Image Comparison

Standard Cartesian PC images are compared to self-gated golden angle TCR images at the coronary sinus in Figures 3.4 and 3.5. In addition to magnitude and velocity reconstructions for a single cardiac phase bin, a temporal slice is shown. Image quality of the self-gated images is on par with the standard Cartesian images. The self-gated images appear to exhibit a slighter higher temporal resolution.

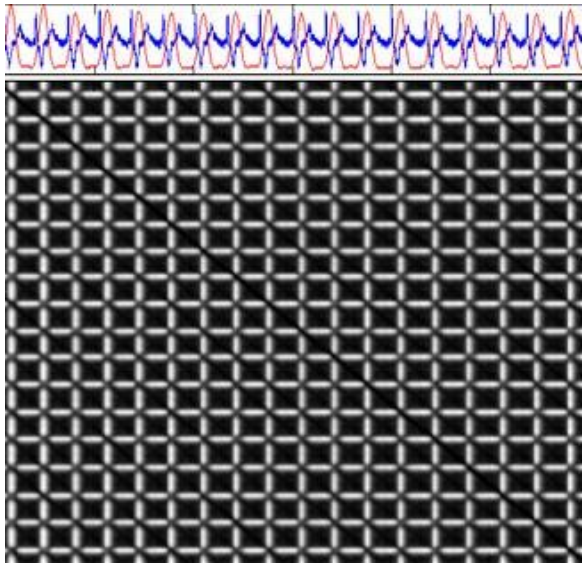

Figure 3.1. A difference matrix using metric 1 from data from the coronary sinus. The synchronized ECG signal is shown in blue, along with the self-gating signal in red. Periods of systole can be distinguished from diastole as they correspond to temporal frames that are dissimilar to more frames than they are similar.

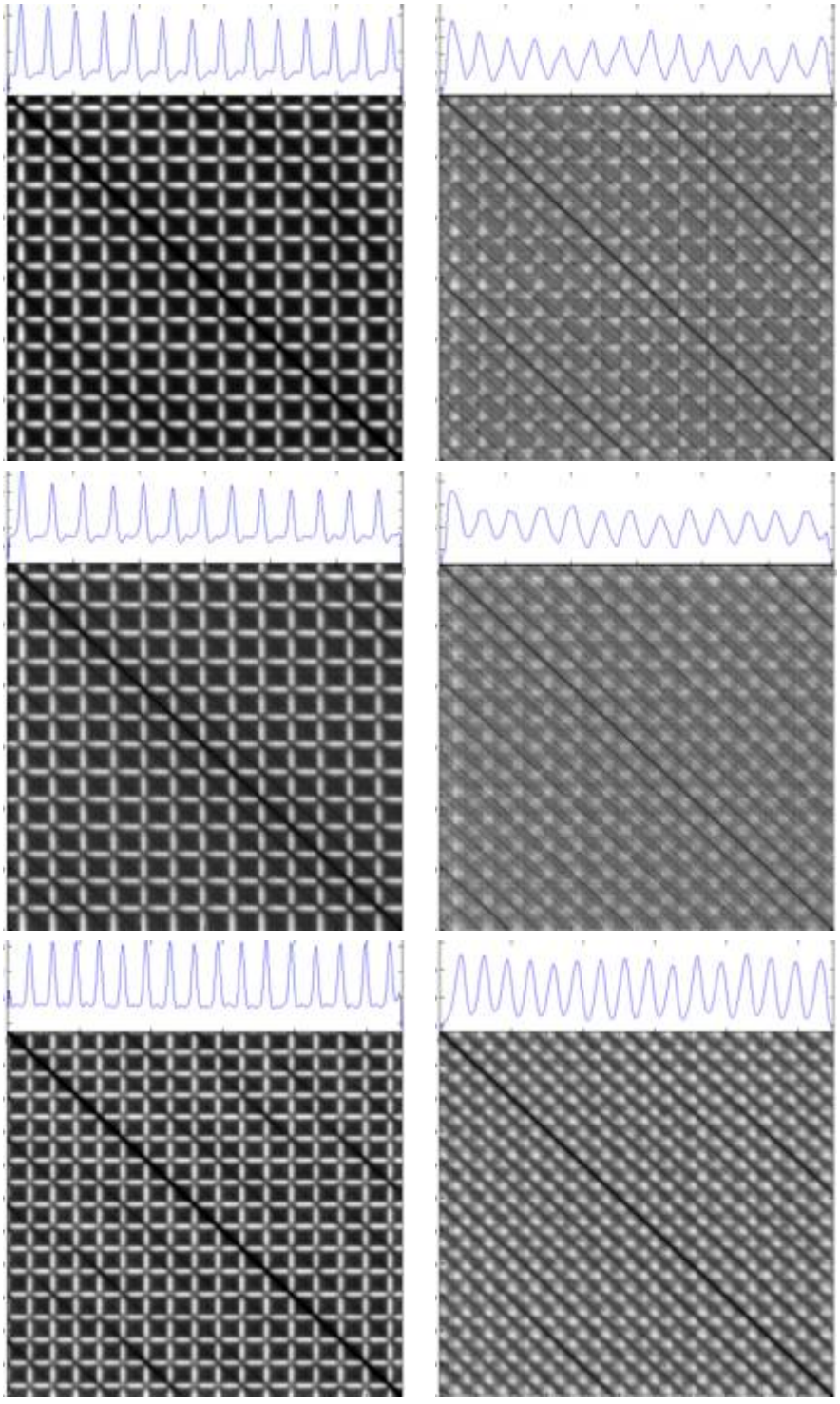

Figure 3.2. Difference matrices along with the derived self-gating signal, using metric 1 (left) and metric 2 (right) from 3 different data sets. (a) is located at the proximal ascending aorta, (b) is located at the pulmonary valve, and (c) is located at the coronary sinus.

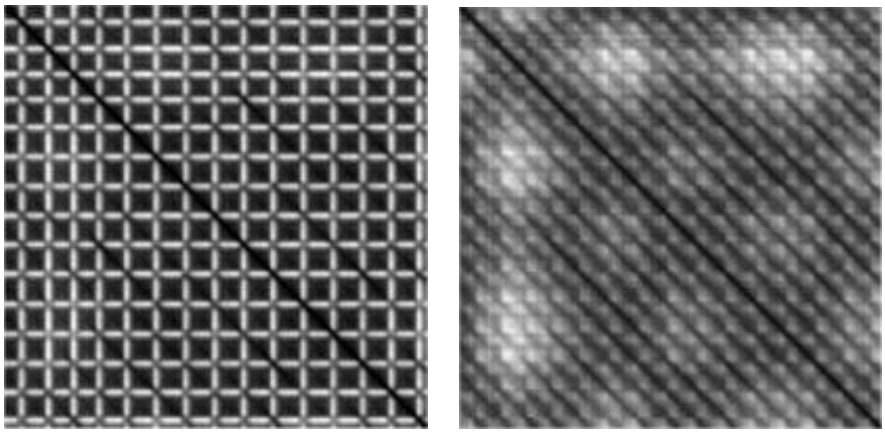

Figure 3.3. Distance matrices from the coronary sinus using metric 1 (left) and metric 2 (right). The effects of respiratory motion are evident in the difference matrix for metric 2 and show up as low frequency modulations in the difference signal

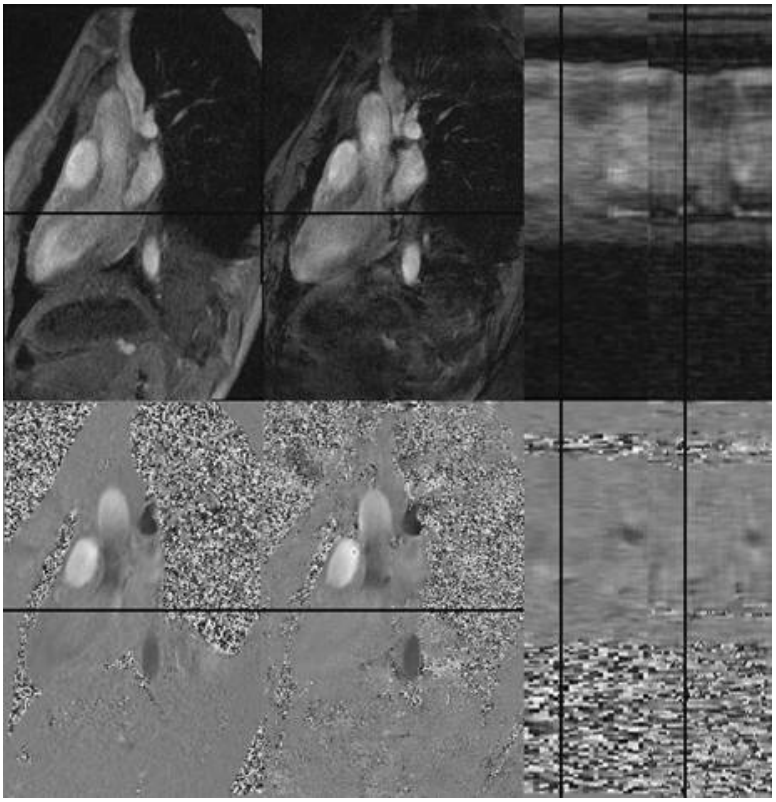

Figure 3.4. Magnitude and phase-difference velocity images of the standard reconstruction EGC-gated Cartesian data (a,c,e,g) and TCR golden angle self-gated radial data (b,d,f,g) are shown. A single cardiac phase is shown in (a,b,e,f) and image row x cardiac phase images are shown in (c,d,g,h). Black lines indicate the position of the corresponding images. Although the Cartesian images have fewer cardiac phases, the temporal images were resized to match the width of the self-gated images.

3.5 Discussion

The distance matrix self-gating technique for PC imaging proved to be robust over multiple acquisitions and multiple slice locations, providing a reliable self-gating signal as an alternative to ECG gating and its inherent difficulties. Distance metric 1 relies on a reference and a velocity encoding set so it could only be used with PC data sets. However, metric 2 uses only the reference data and the same technique could be used to self-gate purely anatomical magnitude cine images.

The use of the golden angle radial view pattern allows for flexible binning of cine data while still distributing measurements well over all the bins. As a result, an arbitrary temporal resolution, or multiple temporal resolutions, can be selected retrospectively in a manner that is tailored to the specific imaging application with no change to the imaging trajectory. A temporal resolution of 10 ms was chosen because experimentally, it produced

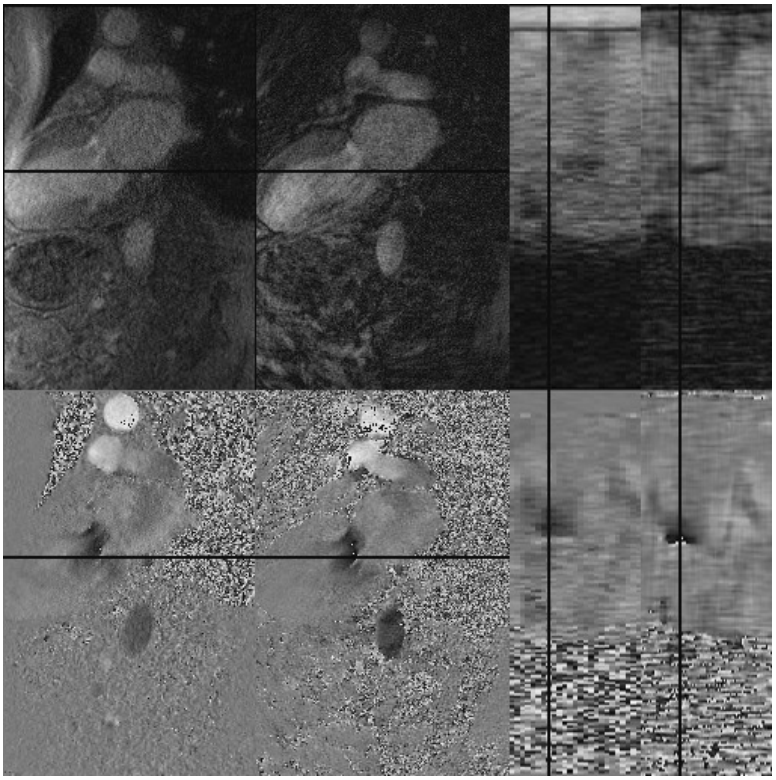

Figure 3.5. Magnitude and phase-difference velocity images of the standard reconstruction EGC-gated Cartesian data (a,c,e,g) and TCR golden angle self-gated radial data (b,d,f,g) are shown. A single cardiac phase is shown in (a,b,e,f) and image row x cardiac phase images are shown in (c,d,g,h). Black lines indicate the position of the corresponding images.

high-quality cine images without significant temporal blurring; however, this was not rigorously evaluated and selection of a different temporal resolution may have provided better results. Additionally, corrupted data from respiration, incongruent heart beats, or other motions can be discarded.

Subjects in this study had very regular heartbeats and produced ECG signals with clear waveforms. Patients with arrhythmias or poor ECG signals may benefit more from the self-gating method outlined here. Further work is needed to determine how effective the new self-gating method is when imaging patients with atrial fibrillation or ahhrythmias.

Reconstruction of the under-sampled data is not limited to TCR; non-Cartesian iterative SENSE, radial GRAPPA, or some other technique could be employed to fill in missing k-space values. However, in this work, TCR produced good results.

3.6 Conclusion

In conclusion, golden angle PC imaging using difference matrix self-gating and TCR is a flexible technique for obtaining high temporal resolution cine PC images in a fashion that is robust to under-sampling and corrupted data exclusion while still producing images on par with standard Cartesian PC.

3.7 References

[1] G. A. Varaprasathan, P. A. Araoz, C. B. Higgins, and G. P. Reddy. Quantification of flow dynamics in congenital heart disease: applications of velocity-encoded cine MR imaging. *Radiographics*, 22(4):895–905, 2002.

[2] J. F. Glockner, D. L. Johnston, and K. P. McGee. Evaluation of cardiac valvular disease with MR imaging: qualitative and quantitative techniques. *Radiographics*, 23(1):e9, 2003.

[3] P. D. Gatehouse, J. Keegan, L. A. Crowe, S. Masood, R. H. Mohiaddin, K. F. Kreitner, and D. N. Firmin. Applications of phase-contrast flow and velocity imaging in cardiovascular MRI. *Eur Radiol*, 15(10):2172–2184, Oct 2005.

[4] H. Arheden, M. Saeed, E. Tornqvist, G. Lund, M. F. Wendland, C. B. Higgins, and F. Stahlberg. Accuracy of segmented MR velocity mapping to measure small vessel pulsatile flow in a phantom simulating cardiac motion. *J Magn Reson Imaging*, 13(5):722–728, May 2001.

[5] P. Thunberg, K. Emilsson, P. Rask, and A. Kahari. Flow and peak velocity measurements in patients with aortic valve stenosis using phase contrast MR accelerated with k-t BLAST. *Eur J Radiol*, Jul 2011.

[6] P. J. Kilner, P. D. Gatehouse, and D. N. Firmin. Flow measurement by magnetic resonance: a unique asset worth optimising. *J Cardiovasc Magn Reson*, 9(4):723–728, 2007.

[7] M. J. Polson, A. T. Barker, and S. Gardiner. The effect of rapid rise-time magnetic fields on the ECG of the rat. *Clin Phys Physiol Meas*, 3(3):231–234, Aug 1982.

[8] R. E. Wendt, R. Rokey, G. W. Vick, and D. L. Johnston. Electrocardiographic gating and monitoring in NMR imaging. *Magn Reson Imaging*, 6(1):89–95, 1988.

[9] G. M. Nijm, S. Swiryn, A. C. Larson, and A. V. Sahakian. Extraction of the magnetohydrodynamic blood flow potential from the surface electrocardiogram in magnetic resonance imaging. *Med Biol Eng Comput*, 46(7):729–733, Jul 2008.

[10] T. A. Spraggins. Wireless retrospective gating: application to cine cardiac imaging. *Magn Reson Imaging*, 8(6):675–681, 1990.

[11] R. D. White, C. B. Paschal, M. E. Clampitt, T. A. Spraggins, and G. W. Lenz. Electrocardiograph-independent, "wireless" cardiovascular cine MR imaging. *J Magn Reson Imaging*, 1(3):347–355, 1991.

[12] M. E. Crowe, A. C. Larson, Q. Zhang, J. Carr, R. D. White, D. Li, and O. P. Simonetti. Automated rectilinear self-gated cardiac cine imaging. *Magn Reson Med*, 52(4):782–788, Oct 2004.

[13] A. C. Larson, R. D. White, G. Laub, E. R. McVeigh, D. Li, and O. P. Simonetti. Self-gated cardiac cine MRI. *Magn Reson Med*, 51(1):93–102, Jan 2004.

[14] R. B. Thompson and E. R. McVeigh. Flow-gated phase-contrast MRI using radial acquisitions. *Magn Reson Med*, 52(3):598–604, Sep 2004.

[15] S. Winkelmann, T. Schaeffter, T. Koehler, H. Eggers, and O. Doessel. An optimal radial profile order based on the Golden Ratio for time-resolved MRI. *IEEE Trans Med Imaging*, 26(1):68–76, Jan 2007.

[16] G. Adluru, S. P. Awate, T. Tasdizen, R. T. Whitaker, and E. V. Dibella. Temporally constrained reconstruction of dynamic cardiac perfusion MRI. *Magn Reson Med*, 57(6):1027–1036, Jun 2007.

[17] G. Adluru, C. McGann, P. Speier, E. G. Kholmovski, A. Shaaban, and E. V. Dibella. Acquisition and reconstruction of undersampled radial data for myocardial perfusion magnetic resonance imaging. *J Magn Reson Imaging*, 29(2):466–473, Feb 2009.

[18] N. Todd, G. Adluru, A. Payne, E. V. DiBella, and D. Parker. Temporally constrained reconstruction applied to MRI temperature data. *Magn Reson Med*, 62(2):406–419, Aug 2009.

CHAPTER 4

FAST TEMPORALLY CONSTRAINED RECONSTRUCTION OF A LARGE CARDIAC PERFUSION MRI DATASET USING A GPU CLUSTER

4.1 Abstract

Under-sampled MRI is a common technique used to reduce total acquisition times. Image reconstruction of under-sampled MR data requires advanced algorithms beyond the traditional Fourier Transform and a great variety of numerical methods have been developed in order to produce artifact-free images from under-sampled data. Unfortunately, many of these methods are computationally demanding and are not applicable in a clinical setting because of longer than acceptable reconstruction times due to scanner computer hardware limitations. In this paper, we demonstrate that using multiple modern graphics processing units (GPUs), it is possible to perform a fast temporally constrained reconstruction of a large cardiac perfusion MRI dataset within the clinical examination time constraints. This was accomplished using hardware that is marginally more complex and expensive than a typical workstation available in a MRI clinical center.

4.2 Introduction

Magnetic resonance imaging has significantly longer image acquisition times when compared to other medical imaging modalities. The speed at which the measurement space can be sampled is constrained by magnetic field gradient hardware limitations, the potential for peripheral nerve stimulation in the patient, and the need to fully sample the measurement space in order to avoid under-sampling artifacts. As a consequence, image properties such as signal to noise ratio (SNR) and temporal and spatial resolution are frequently sacrificed in favor of a shorter acquisition time.

Advances in MR image acquisition and reconstruction techniques allow for accurate reconstruction of under-sampled data when additional information, such as receiver coil sensitivity profiles, is taken into account. The use of under-sampled acquisitions can significantly decrease the total acquisition time. Techniques, such as GRAPPA and SENSE, work with MRI spatial frequency data (k-space data) acquired with a Cartesian imaging trajectory. Under good conditions, these techniques can reduce acquisition times by a factor of 2 to 4, depending on the receiver coil array configuration, with acceptable losses in image SNR. More advanced techniques, such as non-Cartesian iterative SENSE and constrained reconstruction, can utilize a non-Cartesian acquisition scheme and achieve much higher under-sampling. These techniques may also take advantage of additional redundant information beyond receiver sensitivity profiles to further reduce the amount of data that must be acquired, resulting in even shorter acquisition times.

Although under-sampling results in shorter acquisition times, the reconstruction methods required to obtain resolution equivalent to fully sampled data are computationally complex and result in lengthy image reconstruction times that are longer than the few minutes allowable for clinical utility. If image reconstruction cannot be performed during a clinical exam, it is impossible to assess image quality and determine if additional acquisitions are needed while the patient is still on the scanner. Even if image reconstruction is on the order of several minutes, and can be completed during an exam, longer reconstruction times may interfere with technician work-flow and can hold up subsequent image reconstruction tasks as the exam proceeds. As a result, fast reconstruction techniques are needed in addition to expedited acquisitions. In order to make advanced reconstruction techniques clinically relevant, it is necessary to achieve near subminute reconstruction times using modern hardware configurations that could be reasonably installed in most clinical MRI facilities.

4.2.1 Temporally Constrained Reconstruction (TCR)

Reconstruction of under-sampled data is an under-determined problem, and without the use of any additional information has an infinite number of possible solutions. TCR reconstructs under-sampled data by placing an additional temporal constraint on the reconstruction problem so that it can be solved. TCR is an iterative algorithm that identifies an ideal image estimate according to a cost functional that contains terms for both fidelity to the under-sampled data and the temporal constraint [1, 2, 3].

4.2.2 Multicore and GPU Systems

Computation using multicore commodity architectures is becoming more commonplace. Multicore CPUs are widely available on modern MRI systems, and computer systems using multisocket motherboards facilitating several multicore CPUs are relatively inexpensive. High performance computer clusters consisting of multiple computers with such configurations are also becoming more common in research institutions and universities, although they currently require a significant monetary investment to construct and are not widely available for MRI clinical applications. On the other hand, graphics processing units (GPUs) are another type of commodity hardware that can easily and inexpensively be added to multicore systems typical for MRI clinical facilities. Initially, GPUs were used solely to render computer graphics. However, now with the addition of double precision units, many of them are capable of general purpose computing using languages such as NVIDIA CUDA or OpenCL. GPUs are massively parallel architectures, with the current generation containing hundreds of compute cores that can deliver extraordinary performance under the right circumstances.

Several attempts have been made to reduce reconstruction times by utilizing parallel, multicore, or GPU computing architectures for MRI image reconstruction [4, 5, 6, 7, 8]. For example, Kressler et al. [9] computed spiral-sampled time-resolved 3D MR images using up to 64 CPU cores spread across 32 compute nodes and were able to reduce reconstruction time by a factor of 40 compared to reconstruction on a single CPU core. However, fast reconstruction techniques involving high-performance computer clusters are not amenable to clinical applications due to the large and specialized hardware requirements.

Construction and maintenance of computational clusters is expensive and requires specialized expertise. Often, such resources are shared among larger institutions such as universities. Although sharing helps to alleviate these problems, it also usually necessitates the use of processing queues, making around the clock on-demand access to the resources difficult if not impossible. One possible alternative to localized high-performance clusters is the use of "cloud computing" resources, which allow for on-demand allocation of high-performance computing clusters. With cloud computing, as many or as few compute nodes as necessary can be allocated and are usually paid for by the hour during use, and no cost is incurred when they are not used.

From the discussion above, it is apparent that developing algorithms that can be properly parallelized and executed with teraflop performance using hardware readily accessible to a clinical MRI facility can have a substantial clinical impact in allowing reconstruction

approaches for clinical practice. In this work, we demonstrate the use of multiple GPUs to outperform multicore systems and to significantly speed up temporally constrained reconstruction (TCR) of a very large, real-time cardiac perfusion data set. The performance of this configuration is compared to multinode CPU and single GPU implementations. Reconstructions were performed on both a local high-performance cluster and using cloud computing resources.

4.3 Methods

4.3.1 Data Acquisition

With informed consent and IRB approval, an under-sampled, radial cardiac perfusion dataset was obtained using an ungated radial saturation recovery turboFLASH sequence with the following parameters: TR=2.2 ms, TE=1.3 ms, pixel size=1.8x1.8 mm^2, radial views/repetition=24, repetitions=250, slices=4. A Siemens (Erlangen, Germany) 32 channel chest/spine receive coil was used. Radial views were continuously rotated by an increasing factor of the golden angle (θ in radians in equation 4.1) in order to ensure a well spaced angular distribution of radial views over time [10].

$$\theta = l\pi\frac{\sqrt{5}-1}{2} \tag{4.1}$$

4.3.2 Receiver Channel Dimensionality Reduction

Utilization of an ever increasing number of individual receiver channels is becoming common practice in MRI. This has the advantage of providing improved SNR over each of the smaller volumes of the individual coils. The drawback of a large number of receiver coils is that it effectively multiplies the size of the measurement data by the number of receiver coils used, resulting in large amounts of data and increasing the computational demand of image reconstruction. Consequently individual receiver coils are frequently combined into a smaller number of virtual receiver coils that are a linear combination of a subset of all the coils. In this work, principal component analysis (PCA) was used to reduce the number of individual receiver channels from 32 down to 5 virtual coils corresponding to the top 5 principal components of the 32 individual coils [11].

4.3.3 Data Partitioning

Parallelized image reconstruction requires some sort of data partitioning scheme so that work can be distributed between multiple independent jobs. In this work, two different

partitioning schemes were used.

The first sheme splits the data set into smaller temporally overlapping segments. An overlap of 3 temporal frames on each side were used, for a total of 6 redundant frames. A temporal overlap is necessary to deal with boundaries created by the data subdivision. After reconstruction, the overlapping segments are discarded before the data segments are rejoined. Each subset was reconstructed in parallel, and later assembled as depicted in Figure 4.1. One benefit of this partitioning scheme is that it has a high level of granularity when the number of total temporal frames is large. This means that data can be divided relatively evenly any number of jobs. The downside is that the overlapping temporal frames add to the total number of frames that need to be reconstructed, and if too small of an overlap is used, it can result in boundary artifacts in the reconsructed data.

The second partitioning scheme does not split the data in the temporal dimension, but splits the data into separate segments for each channel and slice, resulting in 20 data subsets (5 channels x 4 slices). Subsets are distributed as evenly as possible between jobs. The benefit of this approach is that no overlapping data are required and there is no potential for boundary artifacts. The downside is the low level of granularity. The data can only be divided into 20 smaller sets. If these sets can not be evenly divided into the number of jobs, not all jobs will process the same amount of data. Jobs with less data will finish early, but the whole reconstruction process will be held up until the jobs with more data finish.

4.3.4 Image reconstruction

Temporally constrained reconstruction (TCR) was used to reconstruct the under-sampled, partitioned, k-space data. A constraint on the coil sensitivity profiles was combined with the temporal constraint to reduce the solution space and improve the image reconstruction [12]. The cost functional in equation 4.2 was minimized using an iterative gradient descent algorithm with the following definitions, W: under-sampling mask, S: receiver coil sensitivities, $\mathcal{F}$: 2D FFT, $\tilde{m}$: image estimate, d: measurement data, λ: temporal constraint weight, and ∇_t: temporal gradient.

$$C(m) = \|W \cdot \mathcal{F}S \cdot \tilde{m} - d_c\|_2^2 + \lambda \|\nabla_t \tilde{m}\|_1^1 \tag{4.2}$$

Measurement data from the individual virtual coils generated using PCA was reconstructed independently and later recombined as the square root of the sum of squares to produce the final combined image estimate. Initially, the value of the cost functional was plotted against the number of iterations and it was determined that 100 iterations were

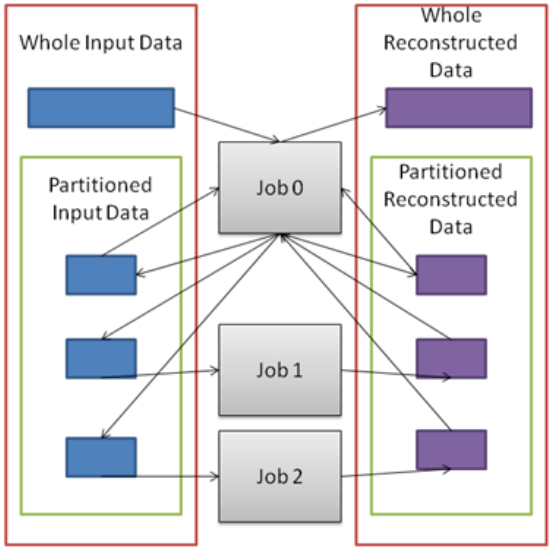

Figure 4.1. Illustration of how large data set is partitioned and processed by multiple jobs before being recombined. Job 0 is responsible for partitioning and recombining the data and also performs image reconstruction on one of the data subsets.

sufficient to ensure acceptable convergence of the image estimate. Calculation of the cost curve was only done to identify the appropriate number of iterations and was not performed during benchmarking.

4.3.5 Computation

Two versions of the TCR algorithm were implemented. GPU code was implemented using the NVIDIA CUDA framework and multithreaded CPU code was implemented using posix threads. Both the CPU and GPU code used the exact same compute kernels. Data communication across nodes was done using MPI. One master node was responsible for partitioning the data and sending the subsets out to the slave nodes. Once the slave nodes had finished reconstruting the data, they were sent back to the master node to be merged.

4.3.6 System Configuration

Image reconstruction was performed on one or more nodes of a 12 node computer cluster. Each node contained 2 Intel Xeon X5560 CPUs with 6 cores running 2.80 GHz, 48 gigabytes of RAM, and 2 NVIDIA Tesla M2090 GPUs with 6 gigabytes of on-card RAM and Mellanox (Sunnyvale, California) QDR Infiniband interconnect. Message passing interface (MPI) was used to distribute work between nodes. Each node ran a version of the Red Hat Linux operating system.

For alternate comparson, image reconstruction was also performed on remote Amazon.com EC2 compute nodes. These nodes are available for on-demand purchase by the hour and are available in several different configurations. In this work, several GPU compute nodes were used, each with 2 NVIDIA Tesla GPUs. The reconstruction process on these nodes was exactly the same as the other system except for an additional data transfer step that was necessary to shuttle the raw and reconstructed data back and forth.

TCR was performed on the same large dataset using several different software/hardware configurations on a local high-performance GPU cluster and on a cluster of remote EC2 nodes. On the local cluster, the CPU code was executed using 1, 2, 4, 8, and 12 cores on a single node and using a total of 60 cores across 5 nodes. The GPU code was executed on a single GPU using a single node, on 2 GPUs using a single node, and on 10 GPUs across 5 nodes. TCR was performed on the remote EC2 cluster using 10 GPU EC2 instances, providing a total of 20 GPUs.

The temporal data partitioning scheme was used for reconstructions on the local cluster. This scheme was chosen because of its granularity, which allowed data to be evenly split between 1, 2, 4, 8, and 12 tasks. The second channel/slice partitioning scheme was used for reconstruction on the EC2 instances because this resulted in 20 smaller data sets, the exact number of GPUs on the cluster.

4.4 Results

Local High-Performance Cluster

Reconstruction times for a single node are shown in Figure 4.2. Reconstruction using a single CPU core took over 2 hours and 35 minutes (155.0 minutes). Increasing the number of cores used to 12 on a single node (the total number of cores on the node) resulted in a reconstruction time of a little over 22 minutes (22.6 minutes). Speedup on a single node was nearly linear from 1 to 2 CPU cores, but speedup factors diminished when using more than 2 cores, presumably due to overhead incurred by shared memory and thread cost. Reconstruction times using GPUs are compared with some multinode CPU configurations in Figure 4.3. Reconstruction on a single node using a single GPU took 7.2 minutes, reconstruction on a single node using 2 GPUs took 3.9 minutes, and reconstruction across 5 nodes (each using 2 GPUs for a total of 10 GPUs) took only 1.0 minute. Reconstruction using 60 cores across all 5 nodes (12 cores on each node) took 3.8 minutes, a very similar amount of time compared to 2 GPUs on a single node. A log scale plot for all reconstruction configurations is shown in Figure 4.4 for comparison. For

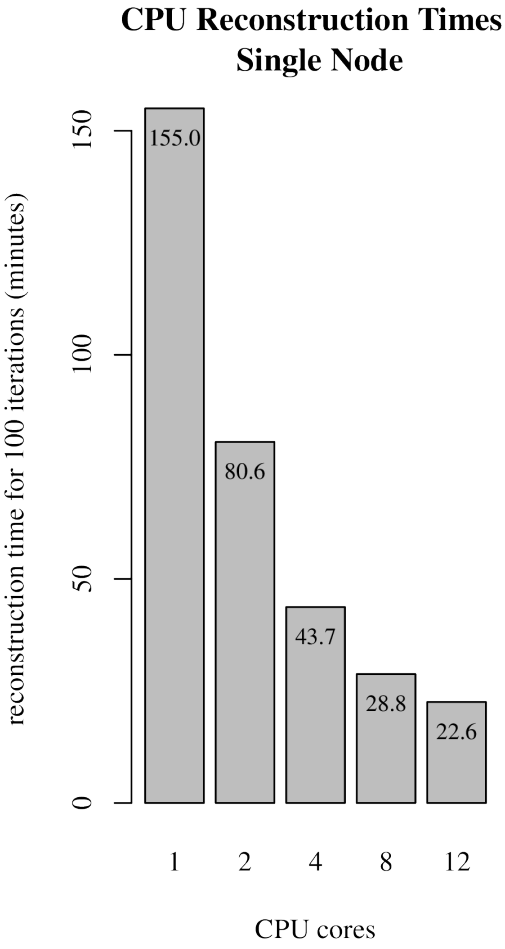

Figure 4.2. Total TCR reconstruction times (100 iterations) for the cardiac perfusion data set. Times shown are for a single node using 2, 4, 6, 8, and 12 cores.

illustration of the results, reconstructed images from several different time frames, for a single slice with channels combined, are shown in Figure 4.5.

Amazon EC2 Cluster

Five different reconstruction times for 10 Amazon EC2 GPU instances are outlined in Table 4.1. Reconstruction times include the amount of time that was required to transfer raw data to the nodes and to transfer reconstructed data back again. Each reconstruction was performed on exactly the same amount of data but there is a large difference in reconstruction times due to variability in data transfer speeds between nodes. The quickest reconstruction took 84 seconds and the slowest reconstruction took 174 seconds.

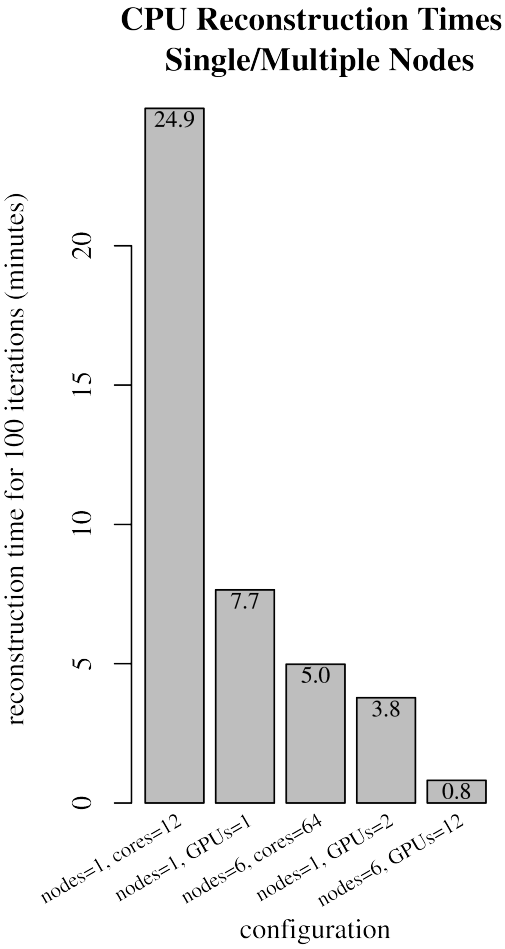

Figure 4.3. Total TCR reconstruction times (100 iterations) for the cardiac perfusion data set. Times shown are for a single node using 12 cores, a single node using a single GPU, 6 nodes using 12 cores each, a single node using 2 GPUs, and 6 nodes using 2 GPUs each.

4.5 Discussion

This work has shown that using multiple CPUs simultaneously for image reconstruction can drastically reduce the reconstruction time for TCR of a large dataset. Reconstruction time was reduced from 155.0 minutes using a single core to to 22.6 minutes using 12 cores. Even faster reconstruction times were achieved using 2 GPUs on a single node, in which case it was possible to perform image reconstruction in 3.9 minutes. This approaches the clinically desirable reconstruction time, using a hardware configuration easily obtainable in an MRI center setting. The marginal cost of adding two GPUs (as used in this work) to a standard server node as those used in this work is approximately $4,000. The fastest reconstruction times were achieved using multiple GPUs distributed across multiple compute

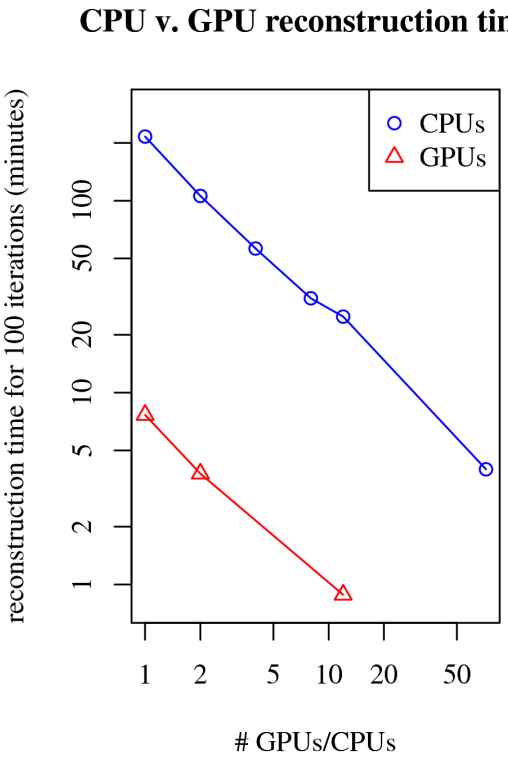

Figure 4.4. Log-log plots for all configurations.

nodes. A 1-minute reconstruction time was achieved using 10 GPUs distributed across 5 compute nodes. Unfortunately, this hardware configuration (5 GPU compute nodes) may not be readily obtainable in a clinical setting, due to cost.

As an alternative to a local GPU cluster, the performance of on-demand cloud cluster was also evaluated. Data transfer times dominated total reconstruction times in this case and added a great amount of variability. Further optimizations may be possible to reduce the data transfer overhead, such as network optimization and data compression.

Privacy is another potential concern when using public on-demand compute nodes. Image data must be transfered over the internet. In this work, data were transferred using an encrypted secure connection, but the fact that data are reconstructed on public hardware may be a concern. Additional measures to ensure that data privacy is maintained need to be investigated.

The TCR algorithm was exclusively computed using either the CPUs or GPUs. In principle, it would be possible to distribute the work across both types of processors simultaneously. However, due to the disparity between CPU and GPU performance, as

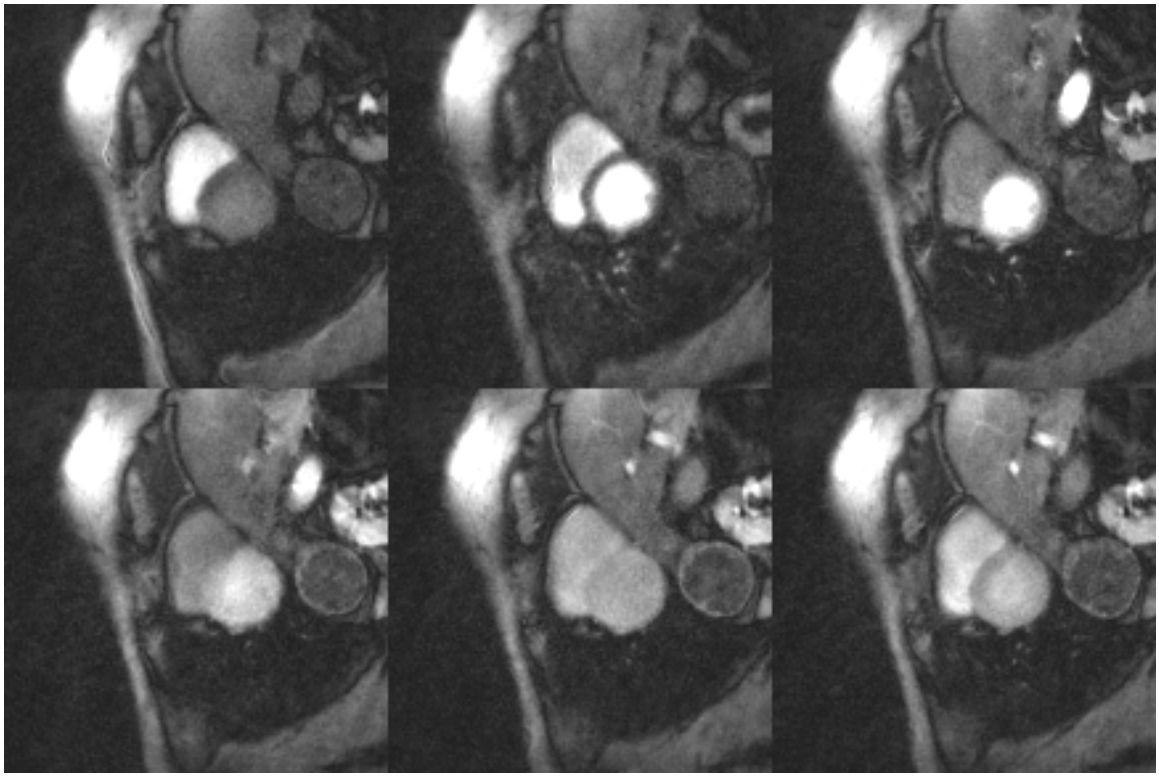

Figure 4.5. Six different time frames from the reconstructed cardiac perfusion data.

demonstrated by this work, it was determined that a dual CPU/GPU setup would not afford enough of a performance increase over a purely GPU setup to justify the additional programming complexity needed to use such a system.

4.6 Conclusions

In conclusion, the use of multi-GPU computer clusters for performing complex MRI image reconstructions of a large dataset can allow for exam-time image reconstruction (1 minute in this work). This allows for real-time assessment of image quality to determine if reacquisition is needed, eliminating the need to recall the patient for a repeat of the current study. Slower, but possibly acceptable, reconstruction times were achieved using a hardware configuration (single node, 2 GPUs) that should be relatively accessible in a typical clinical environment. Use of on demand computational resources could prove to be a low cost alternative to locally owned hardware if data transfer speeds can be reduced and data privacy can be ensured.

Table 4.1. Reconstruction times using 20 GPUs on 10 AWS EC2 nodes

Recon #	Reconstruction Time (s)
1	84
2	87
3	123
4	104
5	174

4.7 References

[1] G. Adluru, S. P. Awate, T. Tasdizen, R. T. Whitaker, and E. V. Dibella. Temporally constrained reconstruction of dynamic cardiac perfusion MRI. *Magn Reson Med*, 57(6):1027–1036, Jun 2007.

[2] G. Adluru, C. McGann, P. Speier, E. G. Kholmovski, A. Shaaban, and E. V. Dibella. Acquisition and reconstruction of undersampled radial data for myocardial perfusion magnetic resonance imaging. *J Magn Reson Imaging*, 29(2):466–473, Feb 2009.

[3] N. Todd, G. Adluru, A. Payne, E. V. DiBella, and D. Parker. Temporally constrained reconstruction applied to MRI temperature data. *Magn Reson Med*, 62(2):406–419, Aug 2009.

[4] M. S. Hansen, D. Atkinson, and T. S. Sørensen. Cartesian SENSE and k-t SENSE reconstruction using commodity graphics hardware. *Magn Reson Med*, 59(3):463–468, Mar 2008.

[5] S. S. Stone, J. P. Haldar, S. C. Tsao, W. M. Hwu, B. P. Sutton, and Z. P. Liang. Accelerating Advanced MRI Reconstructions on GPUs. *J Parallel Distrib Comput*, 68(10):1307–1318, Oct 2008.

[6] S. Roujol, B. D. de Senneville, E. Vahala, T. S. Sørensen, C. Moonen, and M. Ries. Online real-time reconstruction of adaptive TSENSE with commodity CPU/GPU hardware. *Magn Reson Med*, 62(6):1658–1664, Dec 2009.

[7] T. S. Sørensen, D. Atkinson, T. Schaeffter, and M. S. Hansen. Real-time reconstruction of sensitivity encoded radial magnetic resonance imaging using a graphics processing unit. *IEEE Trans Med Imaging*, 28(12):1974–1985, Dec 2009.

[8] D. Kim, J. Trzasko, M. Smelyanskiy, C. Haider, P. Dubey, and A. Manduca. High-Performance 3D Compressive Sensing MRI Reconstruction Using Many-Core Architectures. *Int J Biomed Imaging*, 2011:473128, 2011.

[9] B. Kressler, P. Spincemaille, M. R. Prince, and Y. Wang. Reduction of reconstruction time for time-resolved spiral 3D contrast-enhanced magnetic resonance angiography using parallel computing. *Magn Reson Med*, 56(3):704–708, Sep 2006.

[10] S. Winkelmann, T. Schaeffter, T. Koehler, H. Eggers, and O. Doessel. An optimal radial profile order based on the Golden Ratio for time-resolved MRI. *IEEE Trans Med Imaging*, 26(1):68–76, Jan 2007.

[11] M. Buehrer, K. P. Pruessmann, P. Boesiger, and S. Kozerke. Array compression for MRI with large coil arrays. *Magn Reson Med*, 57(6):1131–1139, Jun 2007.

[12] J. Mendes, D. L. Parker, J. Hulet, G. S. Treiman, and S. E. Kim. CINE turbo spin echo imaging. *Magn Reson Med*, 66(5):1286–1292, Nov 2011.

CHAPTER 5

CONCLUSIONS

MRI is a very rich imaging modality. Due to the mechanism by which measurement data are collected, pulse sequences can communicate a gamut of anatomical and physiological information. The MRI field is still growing rapidly with new techniques being developed and refined all the time. This work has contributed to the field of MRI in 3 different ways. First, this work has shown that TCR can be used to significantly reduce the acquisition time for PC imaging while maintaining good image quality and velocity measurements. Second, this work has demonstrated a new self-gating technique for PC imaging that uses distance matricies to robustly gate a cine PC acquisition. This work has shown that this self-gating method can be used in conjunction with TCR to produce good results in several imaging orientations in the heart. Third, this work has demonstrated the use of multiple GPUs for accelerating the reconstruction time of a large cardiac perfusion data set.

5.1 Summary of Work Complete

5.1.1 Temporally Constrained Reconstruction of Phase-Contrast MRI

Cine PC MRI is affected by respiratory motion; as a result, breath held acquisitions are frequently performed. Unfortunately, a breath hold constrains the total acquisiton time and temporal and/or spatial resolution must be sacrificed in order to allow for a short breath-held acquisition. The first paper in this dissertation demonstrated the use of TCR for reconstructing highly under-sampled radial cine PC data sets.

In order to allow for efficient under-sampling, a new radial sampling scheme was used that allowed for maximum k-space coverage between adjacent frames. The TCR cost functional was composed of a fidelity term that contained the coil sensitivity information in order to allow for higher acceleration. The temporal constraint term consisted of the L1 norm of the temporal gradient that was chosen over the L2 norm in order to preserve temporal resolution.

Three data sets were obtained from volunteers and one data set was obtained from a patient using the new radial PC sequence. Data sets from volunteers were retrospectively under-sampled to several different under-sampling factors; both fully sampled and prospectively under-sampled data sets were obtained from the patient. Under-sampled data were reconstructed with TCR and peak systolic velocity measurements from the under-sampled data were compared to the fully sampled reconstructions. Additionally, image quality of both sets was compared.

TCR proved effective in reconstructing under-sampled data, allowing for a reduction in scan time from 70 seconds to 13 seconds (a reasonable duration for a breath hold). PSV measurements from the under-sampled data compared well to the reference sets. Image quality was on par with the fully sampled reconstructions, except in the case of the patient where motion artifacts due to breathing were present in the fully sampled acquisition, the subsampled data were acquired much more quickly and did not suffer from the same artifacts.

5.1.2 Self-Gated Golden Ratio Radial Phase-Contrast

Traditional cine PC MRI requires the use of an external ECG signal in order to gate individual measurements into separate cardiac phase bins. ECG gating poses several problems, including poor ECG signal due to interference by the scanner and increased exam time needed for application of the ECG leads. Cine PC MRI has the potential to benefit from a new self-gating method and complementary acquisition scheme that allows for robust cine binning without the use of an ECG signal as well as flexibility in selection of temporal resolution and robustness to data exclusion in the presence of motion artifacts detected by the self-gating technique. The second paper in this thesis outlines a new self-gating technique based on distance matrices.

Extending on previous work, the aims of this work included a breath-hold duration acquisition and the use of an under-sampled radial acquisition. However, a new view ordering scheme, based on the golden angle, was used to allow for more robustness to data exclusion. The self-gating signal was generated from distance matrices derived from low spatial resolution real-time reconstructions of the under-sampled data. The distance, based on one of two distance metrics, was computed between each pair of real-time frames, resulting in an MxM matrix of distance values, where M is the number of real-time frames. Correlations between real-time frames cause patterns in the distance matrix that, using peak identification or template matching, can be used to identify similar frames. Similar frames were combined into cine bins and reconstructed using TCR.

Three volunteers were imaged using the new golden-angle PC sequence. A total of 1300 radial views were collected, resulting in an acquisition time of 13 seconds. Fully sampled ECG gated Cartesian acquisitions were also acquired and served as a gold standard for comparison. Slices were acquired in three different locations: the coronary sinus, proximal ascending aorta, and the pulmonary valve. Self-gating signals derived from the distance matrices were compared to acquired ECG signals.

The self-gating signal corresponded well with the ECG signal when the quality of the ECG signal was not too severely affected by interference from the scanner. Cine reconstructions using the golden-angle PC and distance matrix selfe-gating exhibited higher resolution than the standard Cartesian ECG gated images.

5.1.3 Accelerating TCR of Large Datasets Using Multiple GPUs

TCR has been shown to be effective in reconstructing highly under-sampled data, allowing for drastically reduced acquisition times. However, this reduction in acquisition time is necessarily traded for an increase in image reconstruction time due to the computational complexity of TCR. In order for TCR to be clinically relevant, reconstruction times must be reduced enough to allow for exam-time evaluation of the reconstructed images. The third paper in this thesis demonstrates the use of a cluster of GPU equipped computers for fast reconstruction of a large cardiac perfusion MRI data set.

The TCR algorithm was implemented to run on NVIDIA GPUs using the CUDA programming framework. An equivalent algorithm was also implemented using posix CPU threads in order to allow for comparison between CPU and GPU reconstruction times. Reconstruction was performed using one to five nodes of a high-performance GPU cluster, each node containing two NVIDIA Telsa GPUs and two 6-core Intel Xeon processors. Nodes were connected via Infiniband (Mellanox) interconnects. To allow for parallelization between nodes, the MRI data were divided by channel and slice resulting in 20 smaller data sets that were reconstructed independently. Reconstruction was performed using varying numbers of GPUs and CPU cores on a single node and across multiple nodes. When parallelized across 6 nodes using a total of 60 CPU cores, reconstruction time was reduced to 3.8 minutes. Finally, reconstruction using 10 GPUs across 5 compute nodes resulted in a reconstruction time of approximately 1 minute.

This work demonstrated that a GPU cluster can be used to reduce the reconstruction time for a large MRI data set to a reasonable amount of time (one minute) that would allow the images to be evaluated during the exam and potentially reacquired if any problems with

the acquisition are identified.

5.1.4 Impact on the Field of Medical Informatics

The field of biomedical informatics is growing out of the need to manage the recent inundation of biomedical data. Medical imaging, in particular MRI, has the ability to generate staggering amounts of data. As MRI technology continues to advance, the achievable image resolution and number of utilized receive coils will continue to grow, further increasing the amount of data generated. The natural question for biomedical informatics is, "How do we both manage the enormous amount of data we are presented, and how can we prevent waste by only collecting nonredundant information?" Without the answer to such a question, it will not be possible to reap the full benefit of the many advances that are being made. This dissertation has presented answers to these questions for a couple of different time-resolved cardiovascular MR imaging techniques. This dissertation has shown that TCR along with PCA coil compression techniques and the use of highly parallel systems can allow for the efficient sampling and processing of large data sets, allowing for higher temporal resolutions, less motion artifacts, and shorter acquisition times.

5.2 Future Work

5.2.1 Efficient and Robust TCR Parameter Selection

One of the primary difficulties in using TCR is the selection of optimal reconstruction parameters. With TCR, there is an intrinsic tradeoff between the visual quality of each temporal frame and the temporal resolution. The use of a large temporal constraint weight can result in undesired temporal blurring. Work in thesis has used several different approaches to decrease the amount of temporal blurring, including the use of a spatial temporal constraint weight, incorporation of coil sensitivities into the fidelity term, and use of an L1 norm instead of an L2 norm on the temporal gradient. However, even when these methods are used, a temporal constraint weight still needs to be chosen in an empirical fashion so that the reconstructed images are optimal.

In addition to the temporal constraint weight, an appropriate step size needs to be chosen for the gradient descent minimization. If the step size is too small, the reconstructions will require a larger number of iterations to converge. If the step size is too large, the image estimate can blow up instead of converge on an optimal solution. The use of variable step sizes of time could potentially be used to obtain the optimal minimization trajectory. Additionally, the cost of the image estimate could be computed at each iteration and compared to the previous iteration to ensure that the step size is not so large as to result in

an increased cost from one iteration to another; in that case, the step size could be reduced and the iteration could be computed again. However, one drawback of this approach is the increased computation required for the cost calculation at every iteration.

The total number of iterations also has an effect on the resulting image estimate. One could choose to iterate for a very large number of times in order to insure that the image estimate converges. Unfortunately, such an approach could unnecessarily increase the reconstruction time. Once again, the cost of the image estimate could be computed each iteration and the iterations halted once the cost converges; however, as mentioned before, this would only add to the computation time of each iteration.

In order for TCR to be usable in a clinical environment, these kinds of issues need to be resolved. Reconstruction parameters should be relatively consistent between acquisitions or some sort of optimization technique should be used to select the appropriate parameters each time.

5.2.2 Implementation on Clinical Systems

Another obstacle to clinical adoption of TCR is that no commercial implementation currently exists. This is most likely due in part to the previously discussed issues of parameter selection, but is also most likely due to lengthy reconstruction times on the hardware that is currently available on modern clinical scanners.

An image reconstruction framework that standardizes TCR and allows for the utilization of remote computation resources in order to perform TCR quickly could potentially pave the way for clinical adoption.