Blind Retrospective Motion Correction of MR Images

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Alexander Loktyushin
aus Wolgograd/Russland

Tübingen
2016

Tag der mündlichen Qualifikation:	**19.05.2015**
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Bernhard Schölkopf
2. Berichterstatter:	Prof. Dr. Andreas Schilling

Summary

Subject motion during MR image acquisition can cause severe degradation in image quality. The displacement of the imaged object even by a few millimeters is enough to generate motion artifacts, which usually appear as ghosts and blur, and make a scan unacceptable for medical analysis. Correcting the scans for the effects of motion still remains an open problem with no universal solution.

In this work, data-driven retrospective motion correction algorithms aimed at improving the quality of 3D MRI scans affected by both rigid and non-rigid motion are proposed. The crucial aspect of the proposed algorithms is that they require no information about the displacements of a patient in the scanner, i.e. guiding from the tracking cameras. Furthermore, the developed techniques use the raw data from standard imaging sequences and do not require any modifications in the scanning pipeline. Importantly, the proposed methods run on graphic cards in order to attain short computation times, which are in the order of seconds when rigid motion is assumed, and in the order of minutes when more complicated non-rigid motion needs to be corrected.

The rigid motion correction method is based on minimization of a cost function that characterizes the objective quality of the optimized image. The idea is to explore the space of possible motion trajectories in order to find a set of motion parameters that yields the minimum value of a cost function. As an estimator of the image quality, the entropy of the spatial gradients metric is used. In order to be able to correct for non-rigid motion, the forward model of the motion degradation is extended to approximate non-rigid motion by locally (patch-wise) rigid motions. To correct for the effects of such motion, an annealing-based optimization procedure based on alternations of objective function minimization with respect to the latent image and motion parameters has been developed. The proposed algorithms were evaluated on both simulated and real human subject data, and were able to significantly reduce ghosting and blurring artifacts in the images.

Zusammenfassung

Die Bewegung des Patienten während einer MRI Untersuchung kann die Bildqualität stark verringern. Eine Verschiebung des abzubildenden Objektes von nur ein paar Millimetern ist genug um Bewegungsartefakte zu erzeugen und der Scan unbrauchbar für die medizinische Diagnostik zu machen. Obwohl in den letzten 20 Jahren mehrere Verfahren entwickelt wurden, ist die Bewegungskorrektur immer noch ein ungelöstes Problem.

Wir schlagen einen neuen retrospektiven Bewegungskorrekturalgorithmus vor, mit dem man die Qualität von 3D MR Bildern verbessern kann. Mit diesem Verfahren ist es möglich sowohl starre als auch nicht starre Körperbewegungen zu korrigieren. Der wichtigste Aspekt unserer Algorithmen ist, dass keine Informationen über die Bewegungstrajektorie, z. B. von Kameras, nötig sind um die Bewegungskorrektur durchzuführen. Unsere Verfahren verwenden die RAW-Dateien von normalen MRT-Sequenzen und brauchen keinerlei Änderungen im Scanablauf. Wir benutzen Grafikprozessoren um die Bewegungskorrektur zu beschleunigen – im Fall von starren Körperbewegungen sind nur wenige Sekunden erforderlich, bei nicht starrer Körperbewegung nur einige Minuten

Unser Bewegungskorrekturalgorithmus für starre Körper basiert auf der Minimierung einer Kostenfunktion, die die objektive Qualität des korrigierten Bildes abschätzt. Die Hauptidee ist, durch Optimierung eine Bewegungstrajektorie zu finden, die den kleinsten Betrag der Kostenfunktion liefert. Wir verwenden die Entropie der Bildgradienten als Bildqualitätsfunktion. Um nicht starre Körperbewegungen zu korrigieren, erweitern wir unser mathematisches Modell von Bewegungseffekten. Wir approximieren nicht starre Körperbewegungen als mehrere lokale starre Körperbewegungen. Um solche Bewegungen zu korrigieren, entwickeln wir ein neues annealing-basiert Optimierungsverfahren. Während der Optimierung wechseln wir zwei Schritte ab - die Kostenfunktionsminimierung durch Bild- und Bewegungsparameter. Wir haben mehrere Simulationen sowie in vivo Versuche am Menschen durchgeführt – beide lieferten wesentliche Bildqualitätsverbesserungen.

Acknowledgments

First of all, I would like to thank Prof. Dr. Bernhard Schölkopf for giving me the opportunity to work on the very exciting project of motion correction of MR images. The very special atmosphere of intellectual freedom, friendship and openness that exists in the department of Prof. Schölkopf was a unique and important experience for me.

This dissertation would have been impossible without Dr. Hannes Nickisch, who shared with me an invaluable experience of doing a scientific research in a proper way, and showed how to structure things right. Hannes is also the person responsible for the very early steps in my project, the person who gave a kick start to it. We had numerous enlightening discussions over the time of my PhD, and wrote two journal papers together with Dr. Rolf Pohmann and Prof. Dr. Bernhard Schölkopf.

Special thanks go to Dr. Rolf Pohmann – together we made innumerable amount of MR experiments. Rolf introduced me into many of the fascinating aspects related to the Physics of MR image acquisition. I am particularly grateful to Dr. Stefan Harmeling, who was the leader of our image processing group. Stefan has a unique ability to illuminate even difficult concepts with ease, and I am deeply indebted to him for many amazing discussions we had.

I would also like to thank my office mates and colleagues Dr. Michael Hirsch, Dr. Christian Schuler, Dr. Christopher Burger, and Paul Joubert. We had really great times in a both busy and lovely atmosphere of genuine scientific research, and many wonderful discussions over lunch time about everything. Great thanks to Martin Eschelbach for helping me with translations of some parts of this thesis into German language.

I would also like to thank Prof. Dr. Andreas Schilling, who agreed to be my second official supervisor together with Prof. Dr. Bernhard Schölkopf. I am deeply indebted to Sabrina Rehbaum for her patience and evernice attitude when it came to resolving many of the general problems related to the formal side of doing PhD

and research. Sebastian Stark cannot be praised enough for his great help and deep expertise in resolving IT-related issues.

Finally, I want to thank my parents, and in particular, my grand father, who showed my the beauty of nature, and sparkled my interest in her mechanisms.

Nomenclature

CCD - Charge-Coupled Device

CSF - Cerebrospinal Fluid

CUDA - Compute Unified Device Architecture

DICOM - Digital Imaging and Communications in Medicine

DOF - Degree of Freedom

EPI - Echo Planar Imaging

FFT - Fast Fourier Transform

FID - Free Induction Decay

FLASH - Fast Low Angle Shot

FOV - Field of View

FSE - Fast Spin Echo

GLM - Generalized Linear Model

GPU - Graphics Processing Unit

GRAPPA - Generalized Autocalibrating Partially Parallel Acquisition

MIP - Maximum Intensity Projection

MPRAGE - Magnetization Prepared RApid Gradient Echo

MRA - Magnetic Resonance Angiogram

MSE - Mean Squared Error

MVM - Matrix Vector Multiplication

NMR - Nuclear Magnetic Resonance

NRMSE - Normalized Root Mean Squared Error

NUFFT - Non-Uniform Fast Fourier Transform

PDwI - Proton Density Weighted Imaging

PSF - Point Spread Function

SENSE - Sensitivity Encoding for Fast MRI

SNR - Signal-to-Noise Ratio

SPM - Statistical Parametric Mapping

TSE - Turbo Spin Echo

Contents

1 Introduction

1.1 Problem statement

Subject motion during MRI scans can cause severe degradation in image quality[1]. The problem of correcting for such motion artifacts is one of the most important problems in the field remaining to be solved. Even displacement of the imaged object by a few millimeters is enough to generate motion artifacts, which usually appear as ghosts and blur, and make a scan unacceptable for medical analysis. Some imaging modalities such as diffusion imaging are highly sensitive to even weaker motions. The acquisition time for high resolution scans can be in the order of minutes, which together with a requirement to keep motionless in millimeter scale is a challenge even for healthy and cooperative subjects. Patients with movement disorders, the elderly and children are particularly prone to motion during image acquisition, whilst at the same time being amongst the categories of patients who are likely to benefit significantly from MR diagnostics.

There exists a variety of approaches aimed at solving the problem of motion. An efficient way to avoid any bulk motion (and also minimize physiological motion such as breathing) is to sedate the patient. This, however, requires specialist medical equipment and the presence of an anesthesiologist, which greatly complicates the scanning procedure. Furthermore, sedation is associated with risks to the patient and is usually only used as a last resort. A straightforward way to limit the motion of the patient is to use physical restraints to fixate the scanned body part, i.e. soft head cushions can be placed between the head and receive coil. Being patient-friendly, such types of restraints offer only a limited reduction in motion, however. Using bite bars allows for a rigid fixation and minimal motion, but is associated with considerable discomfort to the patient.

In recent decades there has been significant interest in developing techniques that can correct for motion without causing discomfort to the patient. Two large branches

of such methods have emerged: prospective and retrospective motion correction approaches. Prospective motion correction methods adjust the encoding gradients in real-time to accommodate displacements of the scanned object due to motion. In order to operate they require the position of the scanned object at each time point to be known, which can be achieved by using tracking equipment, i.e. optical cameras. A strong aspect of prospective methods is that the k-space is always adequately sampled meaning that there is no missing spectral information, which is often the case whenever the scan is affected by strong movements by the subject. A drawback of prospective methods is that they require additional MR-compatible equipment to track the motion and modifications to the scanning sequences. Furthermore, they are restricted to correction of the rigid motion only since, by adjusting the encoding gradients, it is only possible to induce global transformations of the spatial alignment of the volume.

Retrospective methods impose corrections after the data is acquired. The idea behind them is to apply some transformation to the motion corrupted k-space data such that the effects of motion are undone. Importantly, and in contrast with prospective approaches, this allows for correction of non-rigid physiological motion. Retrospective methods can be further categorized into reference-based and autofocusing subclasses. Reference-based methods are similar to prospective approaches in that they also require motion tracking. Autofocusing methods are reference-free meaning that the motion is estimated directly from the acquired data.

Autofocusing-based methods have many advantages. They are patient-friendly requiring no patient involvement and discomfort. They do not require additional tracking equipment and modifications to the scanning sequences. And importantly, they make it possible to correct non-rigid motion. Yet, autofocusing-based methods are not in conflict with prospective methods, but can be used to complement them, i.e. to refine the image with residual artifacts due to motion tracking errors. This makes it advantageous to use autofocusing methods in clinical practice. The primary problem is that current autofocusing methods are computationally intense. It can take several hours to correct a realistic 3D volume for motion whereas, to be acceptable in clinical practice, the method is expected to operate on a scale of minutes or even seconds. In the case of non-rigid motion correction the computational demands for autofocusing methods are so large that it makes them impractical. The current non-rigid retrospective motion correction methods thus require guiding from external motion sensors to simplify the problem, but it still takes hours to correct

for motion artifacts.

1.2 Contribution

1. An autofocusing-based retrospective motion correction algorithm is proposed, which significantly reduces ghosting and blurring artifacts due to subject motion. The technique uses raw data from standard imaging sequences; no sequence modifications or additional equipment such as tracking devices are required. The method iteratively searches for the motion trajectory that yields the sharpest image as determined by minimization of the entropy of spatial gradients. The vast space of motion parameters is efficiently explored by gradient-based optimization with a convergence guarantee. The objective function that needs to be optimized is highly non-convex. To address the problems associated with local minima in the objective landscape, a multi-scale algorithm is proposed, which propagates the estimated motion parameters from coarse to fine scales. In this way, on each scale there are good initializations of motion parameters driving the optimization process into the vicinity of supposedly good local minima. The approach is evaluated on both synthetic and real data in 2D and 3D using standard imaging techniques. MR images are consistently improved over different kinds of motion trajectories. Using the GPU implementation, computation times are in the order of seconds for a full 3D volume, which makes the method practical in medical diagnostics. The presented technique can be an alternative to prospective motion correction methods. Compared to similar autofocusing approaches, the proposed technique allows for more efficient optimization, which leads to shorter computation times.

2. A fully retrospective non-rigid motion correction scheme that only needs raw data as input is proposed. The method is based on a forward model that describes the effects of non-rigid motion by partitioning the image into patches with locally rigid motion. Using this forward model, an objective function is constructed and optimized with respect to both unknown motion parameters defined for each patch and the underlying sharp image. The method is evaluated on both synthetic and real data in 2D and 3D. In vivo data was acquired using standard imaging sequences. The correction algorithm allows for significant improvement in terms of image quality. The CUDA-enabled GPU implementation ensures feasible computation times. The presented technique is the first computationally feasible retrospective method that allows for correction of non-rigid motion without guidance from external motion

sensors.

1.3 Thesis structure

Outlined below is the chapter structure of the current dissertation:

Chapter 2 gives a very basic introduction into the physics of MRI and the principles of image formation. Good understanding of these is essential in modelling the process of image distortion due to subject motion.

Chapter 3 gives an overview of the current motion correction methods. The connections between the algorithms developed in this research and related methods are discussed.

Chapter 4 introduces GradMC – a blind motion correction algorithm that does rigid body motion correction, does not require sequence modifications and operates on a scale of seconds. Chapter 5 deals with an advanced version of the GradMC algorithm that allows for correction of non-rigid motion and is capable of correcting a more complicated class of motions involving strong rotations.

Chapter 6 concludes the dissertation, and outlines possible future directions for the evolution of the proposed methods.

2 Physics of MRI

2.1 Brief history

Magnetic resonance imaging (MRI) is an imaging technique used primarily for medical diagnostics in order to obtain high quality images of the inner parts of the human body. MRI is based on the principles of nuclear magnetic resonance (NMR), a spectroscopic technique used by scientists to study chemical and physical information about molecules. The technique was initially called nuclear magnetic resonance imaging (NMRI) but was later renamed MRI because of the negative connotations associated with the word nuclear in the late 1970s. MRI began as a tomographic imaging technique generating an image from the NMR signal in a thin slice through a human body. Later MRI advanced to become a more generalized technique that allowed for volume imaging using principles different to tomographical ones. Before proceeding to a description of the physical principles of MRI, it is helpful to outline a brief history of this imaging modality. Felix Bloch and Edward Purcell, both of whom were awarded the Nobel Prize in 1952, discovered the magnetic resonance phenomenon independently in 1946. In the period between 1950 and 1970, NMR was developed and used for chemical and physical molecular analysis. In 1971, Raymond Damadian showed that the nuclear magnetic relaxation times of tissues and tumors differed, thus motivating scientists to consider magnetic resonance for the detection of cancer. Magnetic resonance imaging was first demonstrated on small test tube samples in 1973 by Paul Lauterbur when he used a back-projection technique similar to the one used in computational tomography. In 1975, Richard Ernst proposed the technique to become the basis of current MRI. This technique allowed for actual imaging and relied on the use of phase and frequency encoding together with the Fourier transforms. Just a few years later, in 1977, Peter Mansfield developed the echo-planar imaging (EPI) technique which was another breakthrough since it allowed for acquisition of videos instead of images with rates of 30 ms/frame. In 1980, Edelstein and coworkers demonstrated the imaging of the human body using Ernst's

Figure 2.1: Magnetic resonance scanner.

technique; it took them approximately five minutes to acquire a single image. By 1986, fast imaging techniques were developed allowing for a reduction of the acquisition time to a few seconds without a sacrifice in image quality. The same year NMR microscopy (10 µm resolution on one cm samples) was shown to be possible. In 1987, echo-planar imaging was used to perform the real-time video-rate visualization of a single cardiac cycle. In the same year, Charles Dumoulin used magnetic resonance angiography (MRA) to image blood flow without the use of contrast agents. In 1991, Richard Ernst was awarded the Nobel Prize in Chemistry for his achievements in pulsed Fourier Transform NMR and MRI. In 1992, functional MRI (fMRI) was developed, a technique that allowed the oxygenation levels in different regions of the human brain to be probed. This made it possible to study the function of the brain using a completely new methodology and led to many important discoveries. In 2003, Paul C. Lauterbur of the University of Illinois and Sir Peter Mansfield of the University of Nottingham were awarded the Nobel Prize in Medicine for their discoveries in the field of magnetic resonance imaging.

2.2 Spin physics

2.2.1 Spin property of nature

Macroscopic bodies are composed of molecules. These molecules can be small and consist of only a few atoms as is the case for Oxygen or Nitrogen molecules. They can also be very large – good examples are the (bio)polymers and proteins found in the

cells of living organisms. Each molecule is composed of the atoms of some distinct chemical element. The atoms are not the most fundamental entities of matter and are composed of even more basic particles such as protons, neutrons, and electrons. These, in turn, are composed of yet more elementary quarks. It is not currently clear whether quarks also possess an internal structure or if they are the most elementary form of matter.

Nature has very different properties on different scales. The most familiar scale to us as human beings is the one in which we dwell and to which our senses are adapted. This is the scale that has an order ranging from millimeters to kilometers. Beyond this range our imagination and intuition start to fail. The vastness of the observed universe is disturbing to us, and bizarre quantum mechanical effects puzzle us on the small scales. Indeed, most people find it difficult or even impossible to conceive with the use of mental imagery the quantum phenomena such as particle-wave duality or entanglement, and have to use the bridge of mathematics in order to study such phenomena.

The key to Magnetic Resonance is the spin property of the particles. Here again, human intuition can be misleading: it is common to think of a spinning particle as a fast rotating body, but if we calculate the velocity of some virtual point on the outer surface of the spinning particle we will find that it greatly exceeds the speed of light. Since it is not possible for matter to go faster than the speed of light, this suggests that understanding the spin as a fast rotation is not valid. Instead, the spin is a very special quantum mechanical property of nature that appears in micro scales, and does not have analogues in the macroscopic world. Still, it is common for MR-related textbooks to describe the particles as fast rotating spinning tops, although it is better to conceive of the spin of the particle as a fundamental property of nature such as charge or mass. Another curious fact about the spin is that it comes in discrete values, positive and negative multiples of 1/2. This discretization is yet another manifestation of the quantum mechanical flavor of the spin phenomenon.

Even though we cannot properly conceive of the particle spin in our minds, some familiar concepts from macro physics do hold. We can still find the angular momentum of a particle and its magnetic moment. It is actually this magnetic moment, which the particles with non-zero spin possess, that is of particular importance to MRI. Applying an external magnetic field we can make the spin particles move in a precessional motion, absorb electromagnetic radiation and re-emit it. Recording the re-emitted electromagnetic echoes, we can study the composition of matter. It

is only possible to do this, however, for the elements with non-zero spin. Electrons, protons and neutrons all have non-zero spins being equal to 1/2. In chemical elements that are made up of particles with spins of the opposite signs, it is possible for the spins to cancel out leaving a net spin equal to zero; this holds true for Helium to give one example. Of the utmost importance for MR imaging is the hydrogen element, which has the highest natural and biological abundance. Since the tissues of the human body have different densities of hydrogen, using MR imaging allows for observation of an internal structure of the body as a 3D image.

It is helpful to study the effect of magnetic resonance in greater detail. It is good to start with the following equation: $\nu = \gamma * B$. Here ν is the frequency of the absorbed photon, γ is the gyromagnetic ratio (for Hydrogen equal to 42.58 MHz / T), and B is the magnetic field. This equation means that when some particle is placed in a magnetic field, it can absorb the electromagnetic wave with the frequency proportional to the gyromagnetic ratio of the particle and the strength of the magnetic field. When an atom with non-zero spin is placed in an external magnetic field it tries to align with it. There are two configurations the spins can end up in: firstly, oriented in the direction of the field (low energy state) and secondly, anti-parallel to the field (high energy state). The particle with a spin can make a transition from a low to high energy state by absorbing a photon with energy equal to the difference between the two energy states (Figure 2.2). The transition energy is given by the formula $E = h * \nu$, where h is the Planck's constant and ν is the frequency of the absorbed photon, which is proportional to the strength of the magnetic field. In a similar process, the particle emits the photon with a frequency ν, while making a transition from a high to low energy state. A full equation relating energy to magnetic field is as follows: $E = h * \gamma * B$. Thus, the stronger the external magnetic field is the higher will be the energy of absorbed and emitted photons.

In MR research it is common to use magnetic fields with strengths in the order of a few Teslas, which are very strong fields – the magnetic field of the Earth is much weaker, having strength equal to just $3.1 * 10^{-5}$ T. The strong magnetic fields used in MR mean that absorbed and emitted electromagnetic radiation is in radio-frequency range. But why is it necessary to use such strong magnetic fields? Generating them requires the use of complex hardware, primarily superconducting magnets running huge currents which are cooled by liquid helium. The reason for such strong fields is explained in the following sub-section.

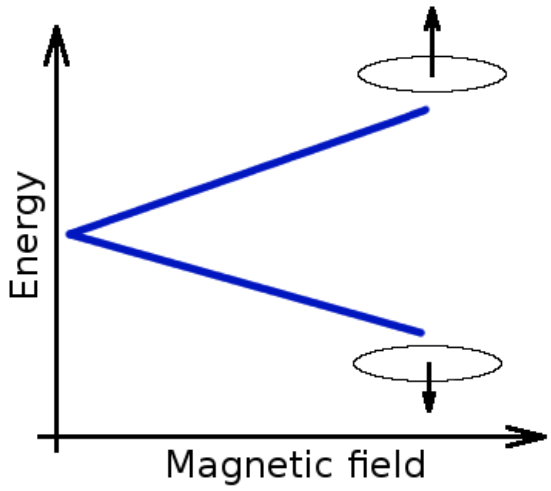

Figure 2.2: The energy required to make a transition from a low to a high state depends on magnetic field strength.

2.2.2 Boltzmann Statistics

When the spins are placed in the magnetic field, some of them end up in the low energy and some in the high energy states (Figure 2.3). The proportion of spins in the high energy state (N-) compared to the number of spins in the low energy state ($N+$) is shown by Boltzmann statistics: $N^-/N^+ = e^{-\frac{E}{kT}}$, where k is a Boltzmann constant and T is temperature. From this equation it is clear that if the temperature is high then there is an equal number of spins in the low and high energy states, since the ratio approaches one. When the temperature gets lower, spins in the low energy state start to dominate. Placing the spins into a magnetic field causes the energy of transition between two states E to become higher, thus increasing the number of spins in the low energy state compared to the number of spins in the high energy state. The Boltzmann statistics equation is important in terms of understanding what the MR signal relies on. The more spins make the transition from a high to a low energy state the stronger will be the recorded signal. Thus, in order to have stronger signals, high magnetic fields in the order of Teslas are used.

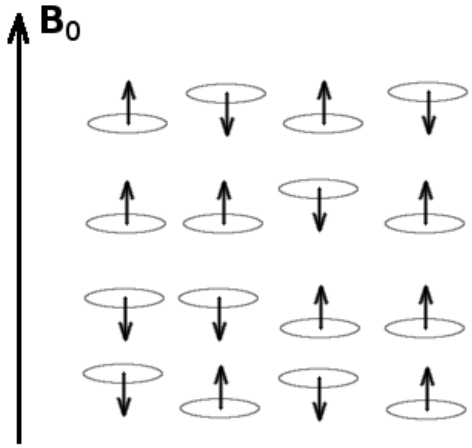

Figure 2.3: Alignment of several nuclei with spin in magnetic field.

2.2.3 Natural and biological abundance of nuclei with non-zero spin

To have a good signal not only is it necessary to have more spins in the low energy state than in the high energy state, but it is also important that the imaged object is composed of the nuclei with non-zero spin. Every chemical element has an isotope, which has a non-zero spin; however, the natural abundance of such isotopes can strongly vary for different chemical elements. For example, non-zero spin isotopes of Hydrogen (1H), Nitrogen (14N), Sodium (23Na) and Potassium (39K) are very abundant; on the other hand, MR-detectable isotopes of Carbon (13C) and Calcium (43Ca) are quite rare. Since MR imaging is extensively used to study biological organisms, it is important to take into account the biological abundance of non-zero isotopes in the chemical elements commonly found in living beings. The most important element is Hydrogen (1H), which has both high biological and natural abundances. It accumulates with varying density in different biological tissues, which allows a study of the internal structure of the organism due to density contrast between the tissues.

2.2.4 Spin packets and precession

Since MR is usually used to study the properties of matter on a scale far exceeding the scale of individual nuclei, it is convenient to consider the effects of magnetic field on spin on a macroscopic level. If a collection of spins experiences the same magnetic

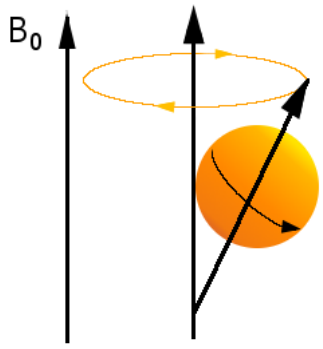

Figure 2.4: Nuclei with spin make a precessing motion around the axis parallel to the magnetic field.

field it can be treated as a single object characterized by a magnetization vector. Due to superposition effects, the individual magnetization vectors of each spin sum together, and the net magnetization vector describes the magnetization of the spin cluster well. The magnitude of this vector depends on the ratio of the number of spins in the low energy state to the number of spins in the high energy state. When the group of spins is excited by photons with a resonant frequency, it makes a transition from a low to a high energy state. It is interesting to consider what happens to the net magnetization vector in this case. Before the excitation, more spins are found in the low energy state. The spins try to align to the external field and stay parallel to it which makes the net magnetization vector at the equilibrium state point in the direction of the field $\boldsymbol{B}$.

Quite importantly, the spins do not stay completely parallel to the field, instead they make a precessing motion around the axis that is parallel to the field (Figure 2.4). This phenomenon is similar to what is happening to a spinning top in the gravitational field. When the spins are excited by the electromagnetic waves of the resonant frequency, they make a transition to the state anti-parallel to the field, so that they still perform the precessing motion and the net magnetization vector is pointing in the opposite direction. Importantly, the frequency of precession (called Larmor frequency) is equal to the resonant frequency. If multiple spins are excited by photons with the same frequency and phase, these spins will precess and the net magnetization vector will rotate around the axis parallel to the external magnetic field with Larmor frequency. It is convenient to describe the net magnetization vector in the frame of reference which is rotating with Larmor frequency around

the axis parallel to the magnetic field vector. This frame of reference is called the rotating frame of reference. By using strong pulsed electromagnetic radiation it is possible to saturate the net magnetization vector so that the number of spins in the low energy state equals the number of spins in the high energy state. In this case the net magnetization vector will have a zero component in the direction of the external field $\boldsymbol{B}$ (commonly denoted by Z axis), and so the magnetization vector will rotate in the X/Y plane. By applying even stronger excitation it is actually possible to flip the magnetization vector to the opposite direction. In this case there will be more spins in the high than low energy state. The precession of the net magnetization vector generates electromagnetic waves which can be detected by placing a loop of wire (Receive Coil) near the object with excited spins. These waves constitute the actual signal that is recorded in MR studies. However, special techniques are still needed in order to recover the spatial locations of the spin packets, and not just cumulative magnetization.

2.2.5 T1 and T2 processes

When the excitation ceases, the spins start to make spontaneous transitions to the low energy state by emitting photons. In a short time, spins return to an equilibrium, and the net magnetization vector has the same direction as the vector $\boldsymbol{B}$ again. The time that it takes the spins to reach equilibrium is called T1, and it is of great importance in MR imaging. T1 relaxation times depend on the temperature and viscosity of the tissues meaning that they can be used to generate a contrast for the acquired image. T1 relaxation affects the longitudinal (Z) component of the magnetization vector. Let M_0 be the magnetization (longitudinal) at the equilibrium, then behavior of the longitudinal component M_z of the magnetization vector is described by the following equation: $M_z = M_o(1 - e^{-t/T_1})$, where t is the time variable (see Figure 2.5 for illustration).

There is a second relaxation effect which is related to the transverse component (X/Y) of the magnetization vector. Since it is hard to make the external field $\boldsymbol{B}$ completely homogeneous, spins in different spatial regions experience fields of different strengths. This makes the spins in different spatial regions precess with slightly different frequencies (Figure 2.6). In a short time the spins are out of phase and the transverse component of the magnetization vector vanishes, since individual magnetizations of the out of phase spins cancel out. The time it takes for the

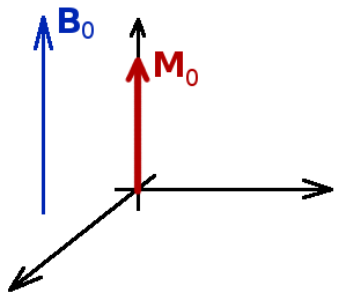

Figure 2.5: The net magnetization vector returns to the equilibrium M_0.

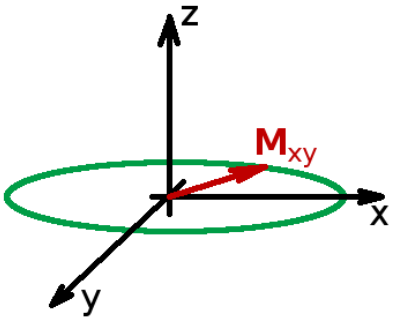

Figure 2.6: The projection of the net magnetization vector onto X/Y plane gradually becomes smaller after excitation due to the dephasing of spins.

transverse component to completely vanish due to dephasing is called T2, and it is a further important intrinsic variable that can be used to obtain contrast in the image. The time evolution of the transverse component M_{xy} of the magnetization vector is described by the following equation: $M_{XY} = M_{XY_o} e^{-t/T_2}$. T2 is always less than or equal to T1. Quite importantly, field inhomogeneities not only result from the fact that the main magnet is not perfect enough to make the field completely homogeneous but also from molecular interactions. The combination of these effects makes M_{xy} vanish in a characteristic time called T2*.

Both T1 and T2* processes occur simultaneously and can be described using the Bloch equations, which completely determine the behavior of the net magnetization vector of a spin system placed in an external magnetic field:

$$\frac{dM_x(t)}{dt} = \gamma(\mathbf{M}(t) \times \mathbf{B}(t))_x - \frac{M_x(t)}{T_2} \tag{2.1}$$

$$\frac{dM_y(t)}{dt} = \gamma(\mathbf{M}(t) \times \mathbf{B}(t))_y - \frac{M_y(t)}{T_2} \tag{2.2}$$

$$\frac{dM_z(t)}{dt} = \gamma(\mathbf{M}(t) \times \mathbf{B}(t))_z - \frac{M_z(t) - M_0}{T_1}. \tag{2.3}$$

Here $\times$ is the cross product, and $\mathbf{B}(t) = (B_{1x}(t), B_{1y}(t), B_0 + \Delta B_z(t))$. B_0 is the longitudinal component of the magnetic field generated by the primary scanner magnet.

B_1 is a pulsed magnetic field that generates electromagnetic waves of a resonant frequency which can be absorbed by the nuclei and causes the transitions of spins from low to high energy states. B_1 is generated by passing an alternating current with a Larmor frequency through a loop of wire called a transmit coil. The longer the time that B_1 is switched on the more spins make a transition. The pulse that makes M_z vanish and M_{xy} wind in the transverse plane is called a 90 degree pulse. To flip M_z to the opposite direction, a 180 degree pulse is used.

2.2.6 Pulsed Magnetic Fields

When 90 degree pulse is applied it causes the magnetization vector to rotate with the Larmor frequency in the transverse plane. The magnitude of M_{xy} slowly decreases due to both the effect of dephasing (T2) and the return of the spins to a low energy state. Using a receive coil it is possible to measure the direction and the strength of M_{xy}. The signal recorded by a receive coil over a short time is called a Free Induction Decay (FID) and has the shape of a damped sinusoid (Figure 2.7). The FID signal depends on the density of the spins, viscosity and the temperature of the imaged object, thus FID signals obtained from different spatial regions can be used to render an image using a suitable contrast. The amplitude of the FID signal S is given by $S = k\rho(1-e^{-TR/T_1})$, where k is a proportionality constant, ρ the density of spins and TR is a repetition time. To improve the SNR it is common to repeat the excite-read sequence multiple times, where the duration of each repetition is given by an instrumental variable called TR. Adjusting the repetition time it is possible to make the acquisition more sensitive to proton density, T1, or T2 effects, and thus

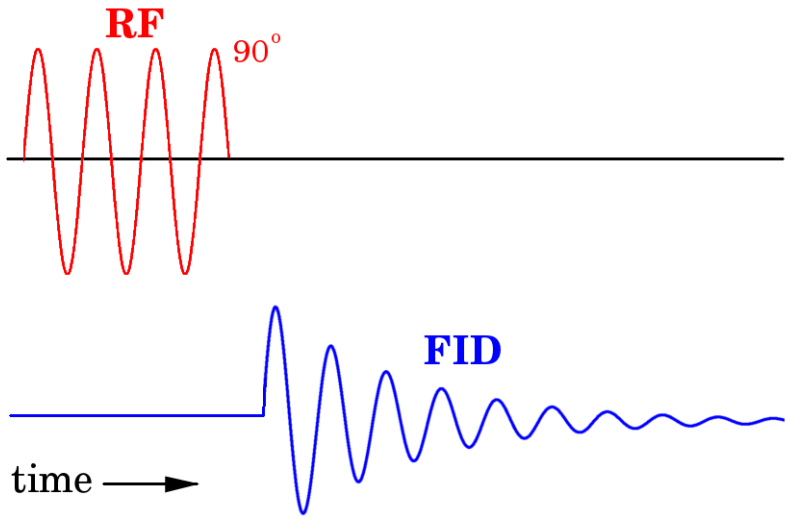

Figure 2.7: Applying an excitation pulse changes the spin states. Placing a radiofrequency coil around the spin system allows the evolution of the net magnetization vector, which is changing due to T1 and T2 effects to be recorded. The recorded signal is called the FID signal.

obtain different contrasts.

2.3 Spatial encoding

Just recording the FID signal is not enough for proper spatial imaging. Indeed, the recorded signal will reflect the average properties of the spin systems from the entire Field-Of-View (FOV). In medical imaging it is important to be able to study the structure of the biological tissues, and to do this it is necessary to carry out spatial imaging. There is a further important component of MR theory that deals with spatial imaging. The general idea is that making the spins precess with different Larmor frequencies, and given that they are susceptible to the photons of different characteristic frequencies, it is possible to tell apart the spatial locations of the spin packets. The resonant frequencies in different spatial regions are altered by superimposing an additional magnetic field that is linear in space onto the main magnetic field (Figure 2.8). To enable full 3D spatial imaging, three techniques are employed – frequency encoding, slice encoding and phase encoding.

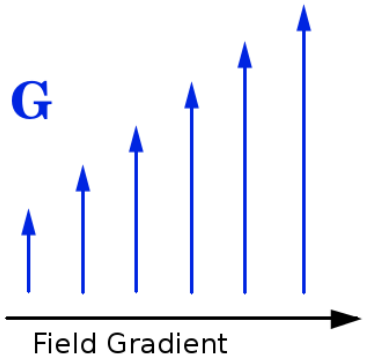

Figure 2.8: Applying an additional magnetic field gradient allows the spin systems in space to be localized.

2.3.1 Frequency encoding

The idea of frequency encoding is as follows. Exciting the spins within some volume with a pulsed magnetic field makes the magnetization vectors of the spin systems precess with the same frequency if the field is sufficiently homogeneous. Applying an additional linear field makes spins precess with different frequencies along the gradient of this field. The signal recorded by the receive coil will then show an entire spectrum of multiple (resonant) frequencies. Since these frequencies have a one-to-one correspondence with the spatial locations of the spin systems, the Fourier transform of the recorded signal allows for the magnetization vectors of the spin systems in different spatial regions to be told apart. By varying the repetition time (TR) it is possible to make the acquisition sensitive to T1 or T2 effects, and to have different contrasts to allow for visualization of the spatial arrangement of the matter with non-zero spins.

2.3.2 Slice encoding

With frequency encoding and a linear gradient it is possible to do one-dimensional imaging. Using a tomography principle of acquiring the image multiple times while changing the orientation of the applied linear field, it is possible to do a sampling of an entire 3D volume by rotating the gradient by an angle ranging from 0 to 360 degrees. The inverse Radon transform can then be used to reconstruct the image. This procedure is, however, quite cumbersome and is not used in practice. Instead, two other techniques are employed - slice and phase encoding.

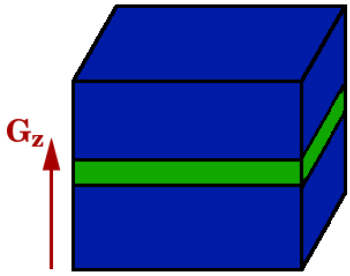

Figure 2.9: Applying the magnetic gradient in a slice encode direction together with a sinc pulse allows the spins in a thin slab of matter to be excited.

The principle of slice encoding is quite symmetric to the one of frequency encoding. In slice encoding it is not the receive stage, but rather a transmit stage that is modulated. Simultaneously, with the application of the excitation magnetic pulse, an additional linear magnetic field is generated in the direction that is orthogonal to the frequency encoding gradient. This additional field is called a slice encoding gradient, and it makes the spins in the respective direction receptive to excitations by the photons having an entire range of Larmor frequencies (Figure 2.9). That means that using a pulse with some characteristic frequency from this range allows the spins in a very thin slab, which is affected by the magnetic field associated with this frequency, to be excited. It is, however, more common to use pulses composed of superpositions of oscillations from some frequency band, which allows the spins in a slice thick enough to generate a good signal to be excited. The desired property of such a composite pulse is that it should select the spatial slice with precisely localized and sharp edges. This is achieved by shaping the excitation pulse to have the form of the sinc function, which has a Fourier pair that is a desirable rect function.

2.3.3 Phase encoding

The remaining problem to solve is to localize the signal in the third dimension, which is achieved by the technique called phase encoding. The idea is somewhat similar to frequency encoding, but is more complicated to understand. Before applying frequency encoding, an additional gradient that is called phase encoding is switched on. This gradient is applied in the direction both orthogonal to the frequency and slice encoding gradients. This causes the spins in this direction to precess with different frequencies. After a short time the gradient is shut down, which causes the spins to wind with the same frequency again but with different phases (Figure 2.10).

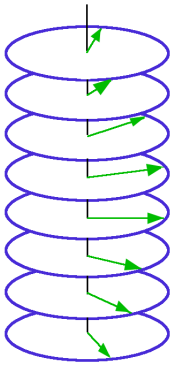

Figure 2.10: Applying the magnetic field gradient in phase encode direction for a short time makes the spins in different spatial locations precess with different frequencies. Turning off this gradient makes the spins precess with the same frequency again, however, with different phase. This allows spatial encoding to be carried out in yet another direction.

The unique phase allows the spatial localization in the remaining dimension, and thus makes the positions of the spin systems in 3D space unambiguous. The exact sequence of encoding steps is as follows: firstly, a slice encoding gradient together with excitation pulse is applied, which affects the spins in a small slab. Next, the phase encoding gradient is applied for a short time. Finally, the frequency encoding gradient together with a signal readout is performed. The recorded signal is then composed of superposition of oscillations of different frequencies and phases.

2.3.4 Volume imaging

To acquire a full 3D volume the encoding techniques described above are executed in a special order which is commonly called an acquisition sequence. A single frequency encoding step allows an entire spectrum of spatial frequencies to be acquired from the Field-Of-View (FOV) in respective direction. In contrast, phase encoding allows only one particular frequency component in the respective direction to be read at a time. Thus, acquisition of a 2D image of a single slice requires multiple phase encoding steps in which, at each step, the slope (steepness) of the phase encoding magnetic gradient is varied. This makes it possible to record the information from all frequency bands in the phase encode direction. It then remains to use a 2D inverse Fourier transform to recover the spatial positions of the spin systems characterized by the properties of interest (proton density, T1 or T2 effects) and store the result

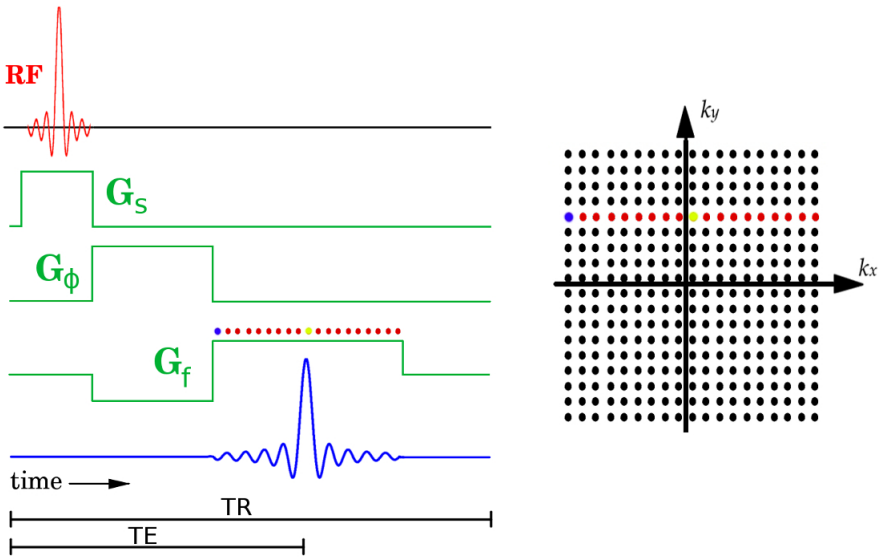

Figure 2.11: Illustration of a 3D imaging pulse sequence with frequency, phase and slice encoding gradients, which allows the acquisition of a 3D volume slice by slice.

in the form of a complex-valued image. An entire 3D volume can be then acquired slice by slice with a total of $N_p N_s$ repetitions, where N_p is the number of phase encode steps and N_s is the number of slice encode steps (Figure 2.11 illustrates a single repetition). Stacking the acquired slices together in spatial domain allows the final three-dimensional volume to be obtained.

This approach, however, does not allow the slice thickness to be made small (due to the vanishing signal intensity). To be able to acquire high-resolution images with isotropic voxel sizes another 3D imaging sequence is employed (see Figure 2.12 for the sequence diagram). The first step in the sequence now involves using an excitation pulse to flip the spins in the entire volume of interest rather than a single slice. Subsequently, two orthogonal phase encoding gradients are switched on at the same time instead of a single one. Adjusting the slopes of both gradients allows any particular frequency component in the directions orthogonal to the frequency encode direction to be probed. Thus, at each repetition, one line from the 3D cube of Fourier coefficients is acquired, until an entire 3D spectrum is sampled. Finally, an inverse 3D Fourier transform is used to recover the volume in spatial domain.

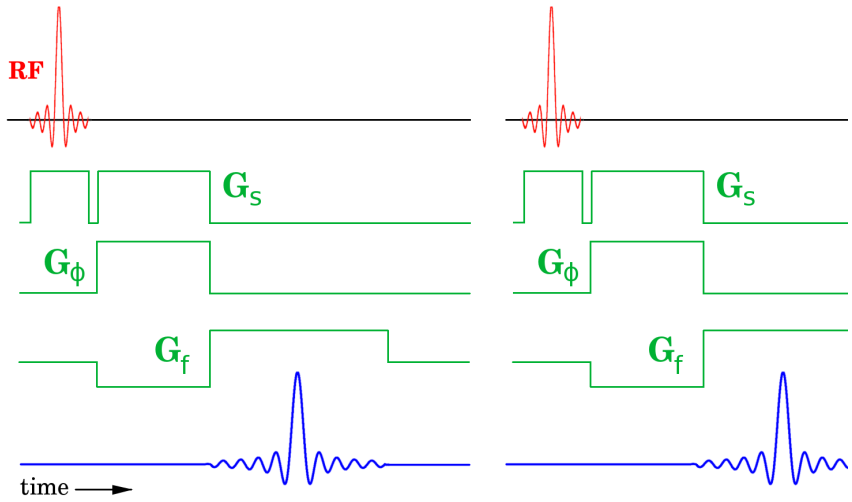

Figure 2.12: The diagram shows one and a half repetitions from a volume imaging sequence. At the beginning of each repetition, RF pulse is used to excite the spins in the entire volume. A combination of two phasing encoding and one frequency encoding gradient is then used to sample one line from a 3D spectrum of the volume.

2.3.5 Spin Echo imaging

So far we have considered imaging sequences, which involved different arrangements of encoding gradients. The flexibility of the sequence design in MRI goes far beyond this. In Spin Echo sequence it is the orchestration of RF pulses that allows interesting effects to be achieved. In this sequence, firstly, a 90 degree pulse is applied to flip the magnetization vector onto the transverse plane. T2 dephasing then starts to saturate the vector's magnitude, and while that happens, a second 180 degree pulse is applied. The effect of this pulse is to make the slow-precessing spins rotate faster, and the fast-rotating spins wind slower. A short time after this pulse, the net magnetization again has the full magnitude in the transverse plane. This refocusing of the net magnetization vector can be recorded by receive coil and generates a signal called echo, which has the shape of a growing and then dampening sinusoid. The time that elapses from the application of the 90 degree pulse to the moment when spins refocus is equal to TE (Echo Time), which is another important characteristic time. The interval between the 90 degree and 180 degree pulses is equal to TE/2. The equation describing the spin echo signal is as follows: $S = k\rho(1 - e^{-TR/T_1})e^{-TE/T_2}$, and it is only valid when TR >> TE. The spin sequence is useful because it allows pure T2 effects excluding the dephasing due to molecular effects (T2* effects) to be

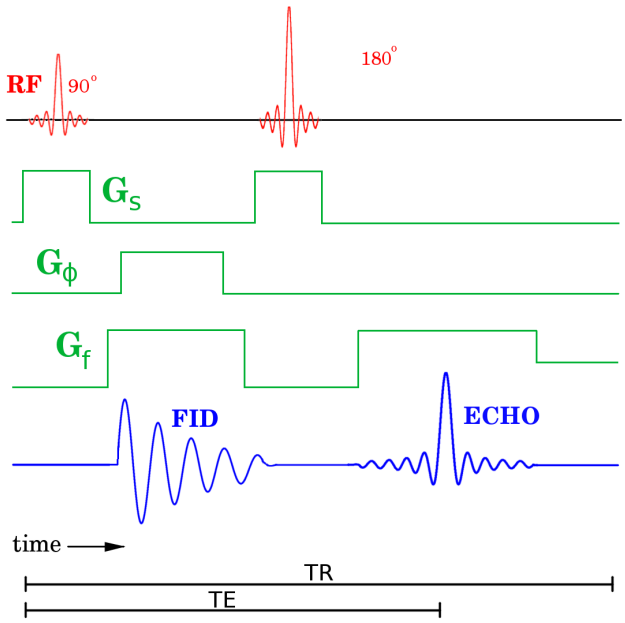

Figure 2.13: The diagram shows a single repetition from the Spin Echo sequence. A combination of 90 degree and 180 degree pulses is used to generate a signal called echo.

studied. The timing diagram that shows the arrangement of RF pulses and encoding gradients in the Spin Echo sequence is shown in Figure 2.13.

2.3.6 Gradient Echo imaging

The sequences described so far require the longitudinal component of the net magnetization vector to return to the equilibrium before the start of each subsequent repetition. If T1 is large this can lead to lengthy acquisition times. It is possible to trade off the acquisition times for the intensity of the acquired signal by using RF pulses that rotate magnetization vectors by the angle ϑ, which is less than 90 degrees. This makes the Z component of the magnetization vector return to the equilibrium much faster. This idea is implemented in a Gradient Echo sequence, which has a timing diagram shown in Figure 2.14. In this sequence, the slice selective RF pulse rotates the magnetization vector by an angle within a range of 10 to 90 degrees.

An important property of the Gradient Echo sequence is that it also produces an echo, which is, however, generated by a different mechanism than that used in Spin Echo imaging. In the Gradient Echo sequence the refocusing of spins is achieved by

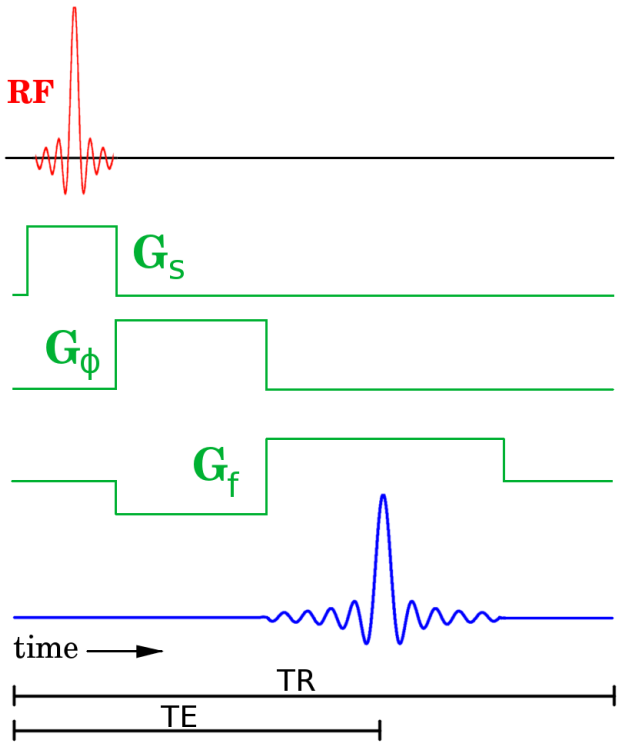

Figure 2.14: The diagram shows a single repetition of a Gradient Echo sequence. In this sequence an echo is generated by using the dephasing frequency encoding gradient prior to the read-out frequency encoding gradient.

simultaneous application of the phase encoding together with a special frequency encoding gradient that dephases the spins. This special dephasing frequency encoding gradient comes before the read-out and has a negative orientation compared to the primary read-out frequency encoding gradient. The echo is generated because the read-out frequency encoding gradient causes the spins, that were dephased by the first frequency encoding gradient, to refocus. Since the refocusing is now accomplished by the gradients, the sequence is called Gradient Echo. This together with the fact that RF pulse rotates the magnetization vector by a smaller angle, results in shorter acquisition times, which is useful when rapid imaging of 3D volumes is needed. Another useful property of the Gradient Echo sequence is that it is intrinsically sensitive to the field inhomogeneities, since it is based on the refocusing gradient.

2.4 Contrasts

MR imaging has the potential to produce a very broad repertoire of imaging sequences [2], which is one of the strongest characteristics of this imaging modality. By tuning various sequence parameters it is possible to trade off the spatial resolution for the contrast and SNR. In medical examinations it is important to discriminate different tissues from one another and from the pathological forms of tissues. In other words, it is necessary to have contrast between the tissues in terms of the spatial variations of acquired signal. This can be achieved by modifying the imaging sequence parameters so that the acquisition is sensitive to the intrinsic magnetization effects. Two tables below (courtesy of [3]) list such intrinsic parameters (or variables), and show their typical values for different tissues found in a human head:

Spin-Lattice Relaxation Time, T1
Spin-Spin Relaxation Time, T2
Spin Density, ρ
T2*

Tissue	T1 (s)	T2 (ms)	ρ*
CSF	0.8 - 20	110 - 2000	70-230
White	0.76 - 1.08	61-100	70-90
Gray	1.09 - 2.15	61 - 109	85 - 125
Meninges	0.5 - 2.2	50 - 165	5 - 44
Muscle	0.95 - 1.82	20 - 67	45 - 90
Adipose	0.2 - 0.75	53 - 94	50 - 100

The contrast between the two tissues can be formed by adjusting the scanning sequence parameters so that intrinsic variables have values such that two tissues of interest have non-overlapping ranges of values. Such sequence parameters are called the instrumental variables and are provided in the following table.

Repetition Time, TR
Echo Time, TE
Inversion Time, TI
Rotation Angle, ϑ
T2*

Setting the instrumental parameters according to the rules outlined in the following table makes it possible to make acquired images T1/T2/T2* or proton density weighted, and thus ensure a contrast between the tissues of interest.

Weighting	TR Value	TE Value
T1	smaller or equal to T1	much smaller than T2
T2	much larger than T1	larger or equal to T2
ρ	much larger than T1	much smaller than T2

3 Motion degradation in MRI

3.1 Introduction

In vivo studies usually deal with objects possessing a rich dynamical repertoire. Indeed, an entire spectrum of the imaging problems associated with varying levels of dynamical complexity exists – Figure 3.1. At one extreme there are problems in situations in which the motion is simple, i.e. the imaged object is a rigid body and its dynamics can be fully described by 6 parameters (three translational and three rotational parameters) at each time point. A good and practically relevant example of such a class of problems is the imaging of a human brain. In this case we are ultimately interested in the anatomical structure or the functional aspects of the brain, and bulk motion is a nuisance which hinders the imaging process. The effects of motion have a strong impact on the image quality, and moderately intense movement can render the imaging results practically useless. Luckily, the human brain can be approximated well as a rigid body meaning that the motion can easily be parameterized and described. In this case, the estimated motion parameters are secondary to the study goals but can be used to correct for motion or used as regressors in functional statistical studies to model the motion-related confounding effects.

At the other end of the spectrum there can be problems when we are interested in the dynamical aspects of the imaged object. In an extreme case we might even completely ignore the structural information and aim at estimating the displacement vector fields for each time point in order to completely describe the dynamics. An example of such a case is for a type of cardiac imaging when we are interested in the motion of blood cells inside the heart, which can be highly complex and turbulent. There are many other problems between these two extremes of differing levels of importance in terms of dynamical aspects for this study. In the middle of the spectrum, for example, is a kind of fetal imaging, the purpose of which is to

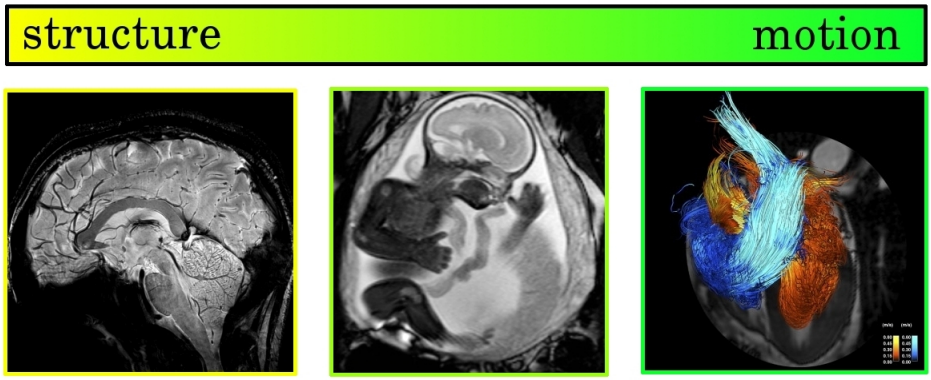

Figure 3.1: Depending on the conducted study we might be interested in anatomical structure (brain imaging), the dynamics (blood fluid motion in the heart) or both (early motor patterns of a human fetus).

observe the early motor patterns of human fetuses. In this case, knowing both the structure and dynamics is equally important.

The problem distinction outlined above is important when it comes to motion correction. A valuable insight is that we usually want to do motion correction when the imaged object is affected by a bulk rigid motion – see Figure 3.2. In this case, the problem of motion correction is greatly simplified, since efficient implementations of the forward models describing the effects of the global rigid motion exist. Furthermore, tracking the object with external devices in order to estimate the motion parameters is feasible in this case. On the other hand, if the imaged object is affected by a complex spatially-varying non-rigid motion we cannot usually simply discard the dynamical aspects of the problem. This means that instead of solving the motion correction problem, we would rather use imaging sequences that allow acquisitions with high temporal rates possibly sacrificing spatial resolution.

It is interesting to compare the problems of motion correction in MR imaging and conventional photography. It is helpful firstly to contrast these two acquisition modalities from the image processing perspective. The obvious difference is that MR scanners naturally allow 3D image acquisitions, which is very useful in in vivo studies, where we need to know the internal structure of the imaged object without compromising its integrity. The characteristic property of MR acquisitions is that the image is recorded in frequency domain, and an inverse Fourier transform is required to obtain a human-interpretable image. Even after doing an inverse FT the image is complex-valued, having a non-trivial non-uniform phase. The complex-

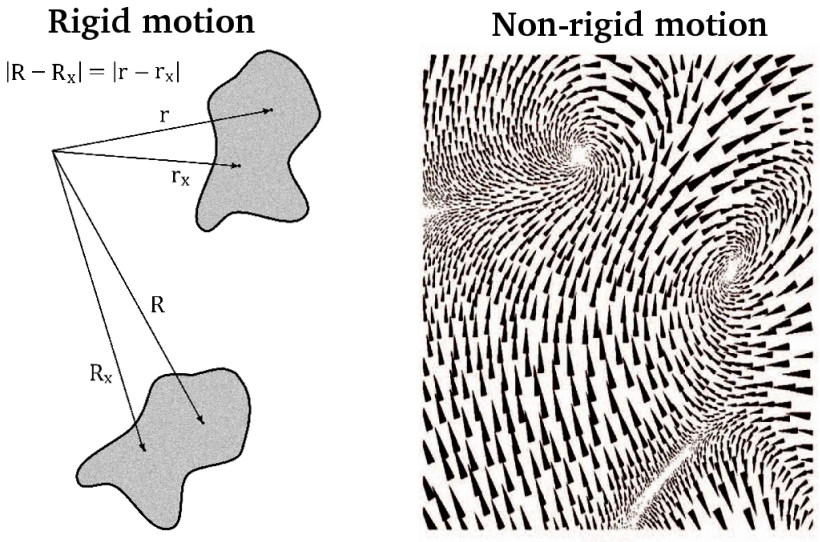

Figure 3.2: Rigid motion involves global translations and rotations and can be described by $6T$ parameters, where T is the number of time points. Non-rigid motion can be described by displacement vector fields specified for each time point.

valuedness of the image is the source of many challenges when it comes to motion correction as we will see later in this manuscript.

The most important difference that is relevant to solving the motion correction problem has to do with the image formation process. Digital cameras equipped with CCD sensors collect photons from the entire FOV at each time frame. Over the exposure the photon counts in each sensor element accumulate and the final image is composed of the time-stacked pixel intensities. Whenever, during the exposure, the imaged object or the camera is affected by translational motion this causes the stacking of the shifted copies of the imaged object, which makes the image appear blurry. Mathematically, the effects of motion can be described by a convolution of the latent motionless image with some point-spread-function (blurring kernel) that depends on the motion trajectory. In practice, object/camera motions usually also involve rotational components, and the motion degradation is described by convolutions with spatially-varying point-spread-functions.

As described in the previous section, MR acquisition is based on sequential read-outs of the k-space segments. Interestingly, if the imaged object is affected by pure translations, the image degradation due to motion can again be described as a convolution in spatial domain. The respective convolution kernel has certain peculiar properties – it is complex-valued, and its power spectrum is unit-valued in

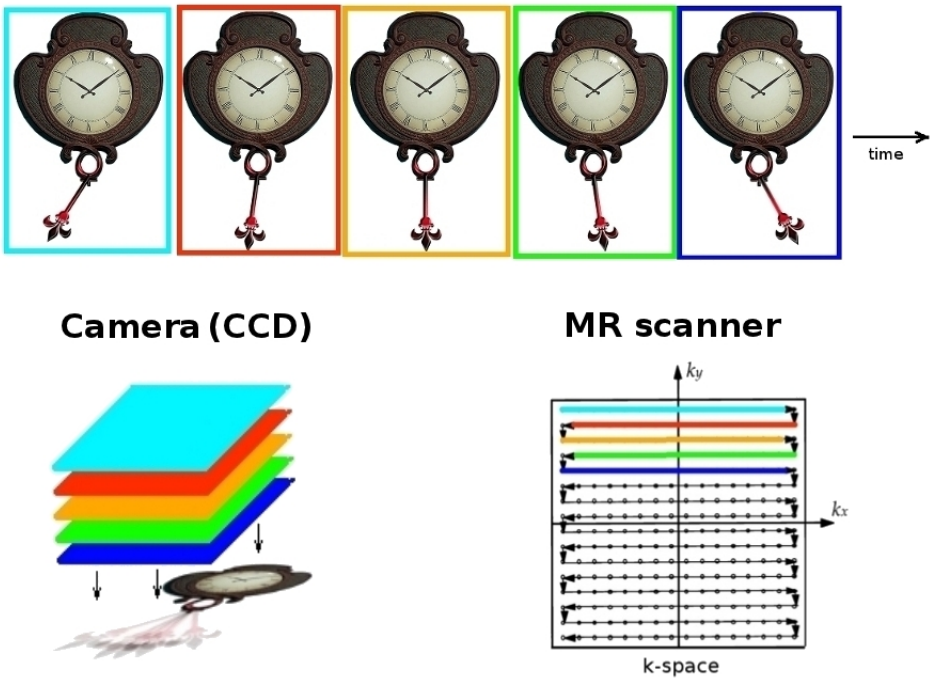

Figure 3.3: The temporal aspects of the image acquisition with CCD cameras and MR scanners: exposure over time with a conventional camera leads to intensity stacking in the spatial domain which, in case of motion, leads to blurring. MR imaging involves the acquisition of the different segments of the frequency domain at each time point. In this case, motion artifacts are due to incoherence between k-space segments, which leads to blurring and ghosting of the image data in spatial domain.

all frequencies. The Figure 3.3 illustrates the effects of motion in both acquisition regimes.

It is the combination of acquiring the image in frequency domain together with time-sequential readouts of k-space segments that makes the problem of MR motion correction unusual. Shown in Figure 3.4 are the typical effects of a human subject bulk motion on a recorded image of the brain.

3.1.1 Types of motion

It is good to start by giving an overview of the types of motion that are usually observed in human MR studies. There are two broad classes of motion: bulk involuntary movements, and organ and tissue deformations and displacements due to

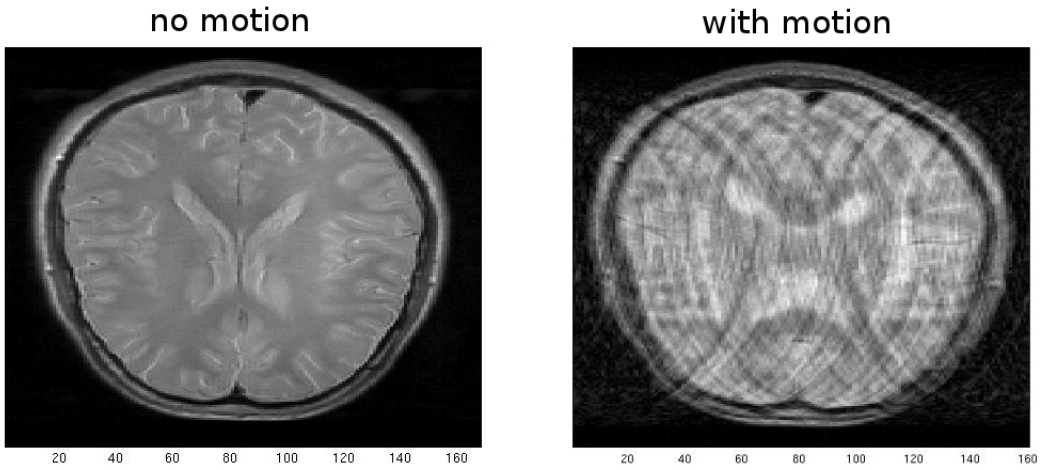

Figure 3.4: Typical motion artifacts due to subject motion during image acquisition.

the physiological function of the organism. The first class of motion is most commonly observed in subjects who are lacking control over their motor actions such as children, elderly people and patients with movement disorders. However, even healthy adult subjects can have trouble controlling their movements when it comes to lengthy acquisitions of high-resolution images. Furthermore, given how long it can take to perform a scan, it might require a large effort on the part of a subject to avoid semi-deliberate physiological acts such as coughing, sneezing and swallowing.

The second class of motions is even more problematic since it involves the autonomous body systems (cardiac, respiratory) which are beyond control. Such motions are most often observed in abdominal imaging where, if ignored, they can easily make the acquired image non-diagnostic. Compared to the head bulk motion, physiological abdominal motions are usually non-rigid, involving complex deformations of tissues, especially in cardiac muscles. Additionally, there exist hybrid types of motion, a good example being the motion of the fetus in the womb, which involves a combination of both mother's respiration and the bulk motion of the fetus.

3.2 Motion avoidance

Since motion is ubiquitous and the acquisition times for high-resolution/high SNR images are long, solving the problem of subject motion is practically very important. The most intuitive and straightforward way to avoid bulk motion is to physically

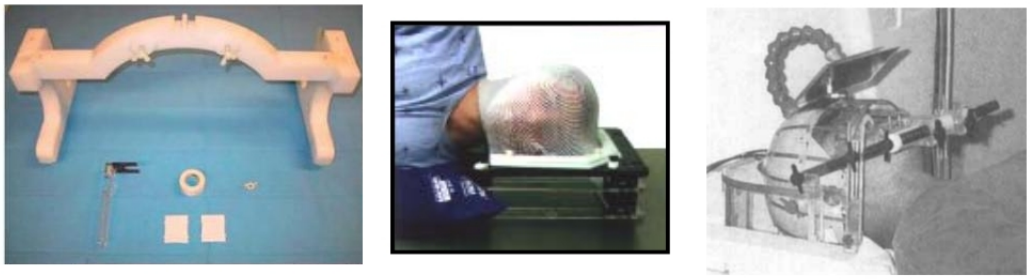

Figure 3.5: Physical constraints can be used to minimize the subject motion: bite bars, head nets, fixation helmets with a bite bar.

constrain the subject. Whilst doing head scans this can be achieved with the use of bite bars, head cushions, or fixation nets – see Figure 3.5. Head cushions are a gentle and subject-friendly type of motion restraint; however, they usually provide only limited aid against motion. They prevent gross motions, but still allow the subjects to make minor movements which can degrade the quality of high-resolution images.

Constraint devices such as bite bars (the subject is asked to hold them tightly with her teeth) and fixation helmets are much more efficient at preventing motion; however, they introduce significant discomfort to the human subject and can trigger claustrophobia leading to premature termination of the scan.

Even more extreme motion control methods involve using anesthetic agents and muscle relaxants. Such methods, however, tend to bring many complications both in terms of the scanning setup (due to the need to use additional life-support systems with anesthesia) and the health of the patient. They are usually employed in situations where a scan needs to be urgently performed on a subject (such as a pediatric patient) who cannot properly control his own movements.

3.3 Motion-tolerant sequences

When carrying out abdominal scans instead of correcting for motion artifacts, it is more common to use special sequences, which account for the regularity of respiratory and cardiac cycles. Such (gating-based) imaging techniques perform well-timed read-outs triggered by the phase of the physiological cycle, which can be obtained with the use of additional sensors such as ECG – see Figure 3.6. Using this tech-

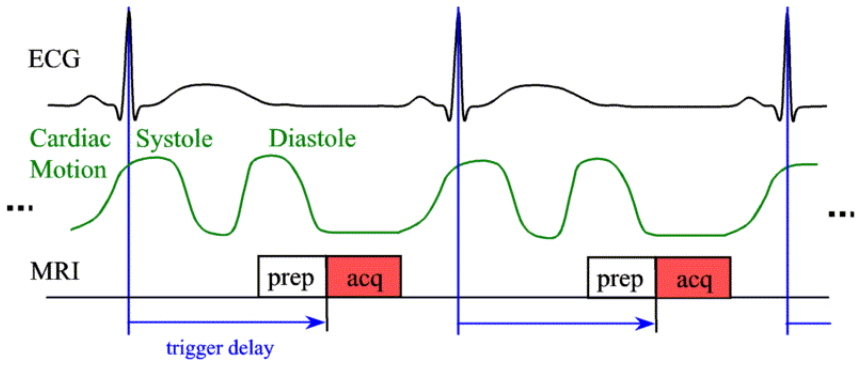

Figure 3.6: External physiological sensors (i.e. ECG-based) can be used to retrieve the phase of the physiological cycle. The k-space read-outs are then synchronized with the cycle, such that the imaged object appears motionless and in one phase. Image courtesy of [4].

nique the slices from the imaged volume are always acquired in the same physiological phase, which ideally leads to a scan free of motion artifacts. In practice, however, the phase estimation is not perfect and post-processing needs to be done to do a slice registration before the final volume is assembled.

In the case of involuntary bulk head motion the use of special acquisition sequences can also be helpful. There exists a trade-off between spatial and temporal resolution (Figure 3.7). In other words, a fixed scanning time budget can be used to acquire an image with high spatial resolution or to acquire multiple low-resolution frames for the same object. Whether to prioritize spatial or temporal resolution is dictated by the nature of the imaging problem, i.e. if the scanned object is affected by complex non-stationary deformations (non-rigid motion) it might be preferable to use an ultra-fast EPI-like sequence and treat the motion as an important dynamical aspect of acquired data rather than a confounder. The drawback of such acquisition schemas is that the fine details in the image might be indiscernible due to low spatial resolution. On the contrary, if the imaged object is a rigid body then the motion correction problem is easier to solve, and the scanning time budget can be dedicated to the acquisition of high-frequency regions of k-space resulting in high spatial resolution. There also exist hybrid approaches where some regions of k-space are acquired more often than others which allows an adjustable compromise between the spatial and temporal resolution. On the downside, such hybrid approaches usually require sophisticated post-processing.

It is also possible to design sequences which are intrinsically prone to motion. Such

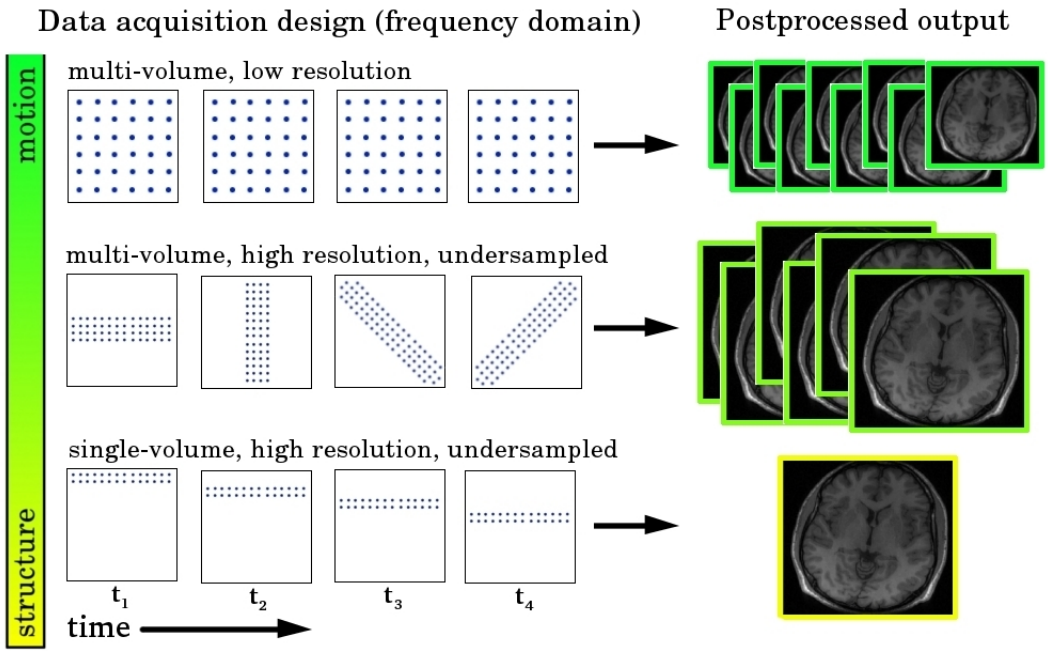

Figure 3.7: Depending on the complexity of anticipated motion, the operator of the scanner can choose different acquisition schemas. In the case of complicated non-rigid motion it might be preferable to acquire a sequence of low resolution images with high temporal fidelity. On the other hand, if the imaged object is a rigid body and the motion is temporally smooth, it is possible to acquire a single volume with a high resolution. Furthermore, there exist hybrid acquisition schemas, which involve the repetitive acquisitions of low-frequency regions of k-space together with unique (acquired only once) high-frequency k-space segments. This allows a trade-off between the spatial and temporal resolution in a flexible way.

sequences are usually based on special acquisition trajectories – see Figure 3.8. In a projection sequence, k-space lines are swept in a radial pattern around the center of k-space. The acquisition with radial trajectory is resistant to translational motion because the phases in each acquired line are always congruent. Similarly, the spiral trajectory has certain advantages when the scanned object is affected by rotational motion. There are further more sophisticated trajectories, which involve multiple overlaps and redundant data which can be used to infer the motion parameters. Reconstructing the final image in spatial domain is more complicated than just an inverse Fourier transform and usually requires more complicated post-processing.

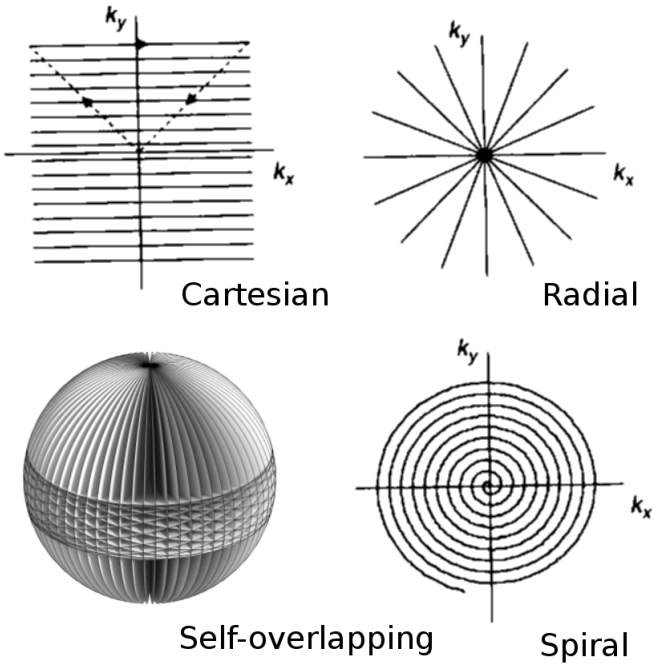

Figure 3.8: Shown in the figure are Cartesian, radial, spiral and self-overlapping trajectories, which can be used to traverse k-space. Acquisitions with radial and spiral trajectories are intrinsically resistant to certain types of motion. Using trajectories with self-overlapping segments allows the motion parameters from the regions of overlap to be inferred, and then the data for motion artifacts is retrospectively corrected.

3.4 Motion correction methods

Although there is a vast literature on motion correction, no universal solution to the problem yet exists. The available methods can broadly be classified into *prospective* and *retrospective* techniques. Prospective methods correct for motion during the scan by constantly adjusting the magnetic field gradients to follow the subject's pose (see Figure 3.9). The position of the imaged body part is measured by external sensors, like tracking cameras [5, 6]. Retrospective methods remove motion artifacts after the image is fully acquired. This can be done by estimating the point spread function due to translation, and then using its phase for correction [7, 8], by correlating the values in adjacent k-space lines to estimate the motion [9], by the method of generalized projections [10] in space-limited images with a known finite region of support, by special trajectories introducing redundancy into the acquired data that allows inference of the motion [11] or by navigator-based methods that measure the

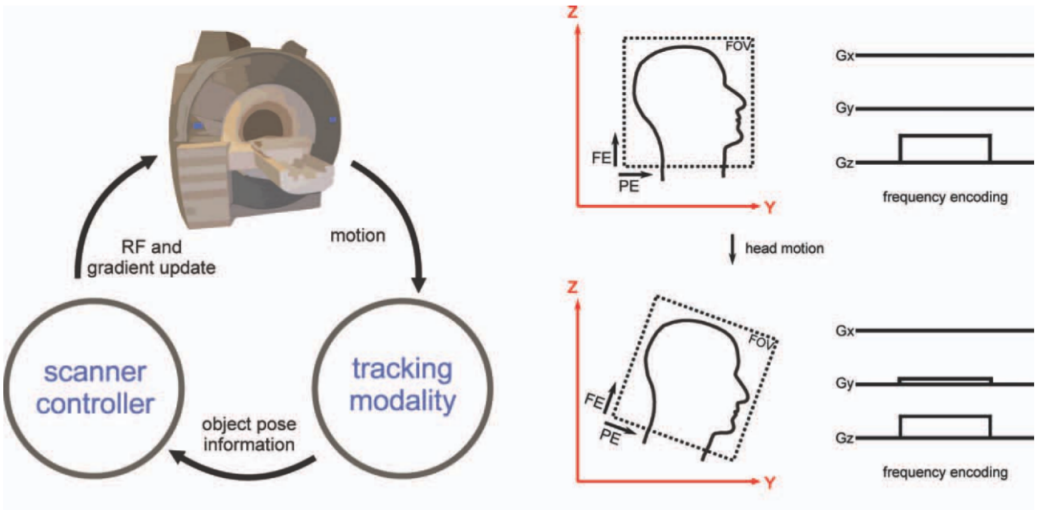

Figure 3.9: Prospective motion correction idea: adjust the gradients in real-time to keep the FOV stationary. Image courtesy of [25].

motion [12, 13] using additional echoes in the pulse sequence.

Autofocusing (AF) methods represent a large and important class of retrospective techniques. AF does not require the use of special trajectories, or any other additional data, and relies on optimization-based refining of the image w.r.t. some function evaluating the image quality. Such methods have been shown to be able to compete with navigator-based methods in terms of image quality [14]. The origin of AF methods can be traced back to early attempts to solve image denoising and deblurring problems [15, 16]. AF is used for motion compensation in radar measurements [17] where the acquisition, as in MR, takes place in the frequency domain. The seminal paper of Atkinson [18] adapts AF to MR motion correction. In recent years, AF methods for MR were improved substantially to accommodate complex motions involving rotations [19, 20], and to deal with 3D acquisitions [21]. Recently, retrospective approaches to correction of a non-rigid body motion were proposed [22, 23, 24]; however, to cope with the large number of unknown motion parameters, reference data (i.e. from navigator-based trackers) is needed.

3.4.1 Prospective motion correction

One advantage of prospective methods is that the image output by the scanner is already free of motion, thus no additional (potentially computationally intense) post-processing is required [26]. A further advantage is that k-space is always ad-

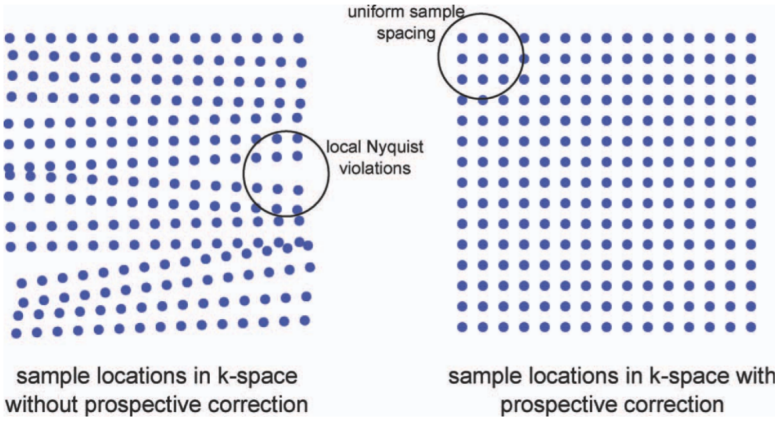

Figure 3.10: Rotational motion can cause an inconsistent sampling of k-space. Image courtesy of [25].

equately sampled, meaning that rotational motion does not lead to the missing k-space data, and the image reconstructed in spatial domain will not suffer from associated problems (see Figure 3.10 for illustration of this aspect). Additionally, prospective methods are naturally immune to the problems associated with events when the scanned object leaves the FOV. Prospective methods also avoid spin history effects which occur when spins in some slice that was excited by a multiband pulse change their positions due to motion. It is desirable to avoid such events since, when excited spins change their positions, they cease to be in a steady state which leads to complex interference from the previous spin states, and causes the image artifacts.

At present, prospective methods are mostly employed in research labs and are not routinely used in clinical practice. There are two reasons for this. Firstly, in order to do prospective correction, special imaging sequences need to be used. Secondly, MR-compatible hardware is required to track the motion and provide real-time estimates of motion parameters to the scanner.

3.4.1.1 Motion tracking

An essential prerequisite for the success of any prospective method is the availability of estimates of motion parameters. The approaches used to measure motion can be broadly classified into external hardware-based and scanner-based. External tracking devices include optical cameras, laser tracking devices and optical tracking tapes.

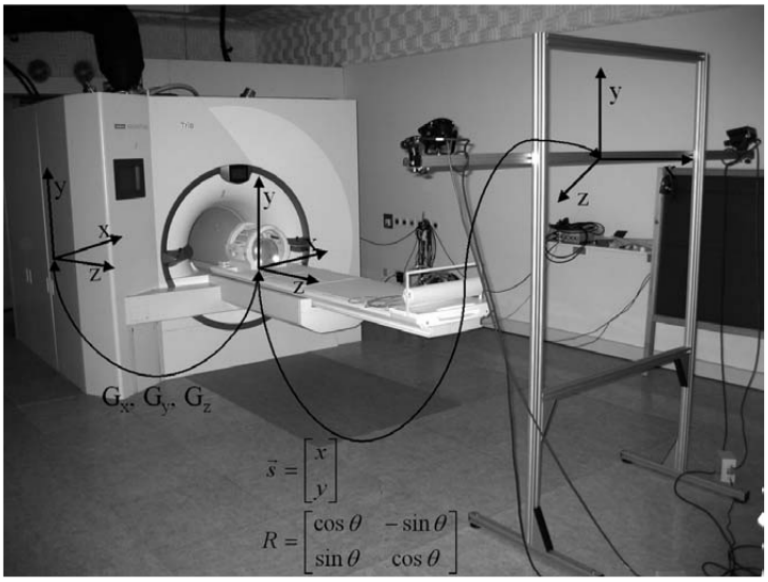

Figure 3.11: External tracking devices such as special MR-compatible cameras can be used to detect and measure subject motion. Image courtesy of [26].

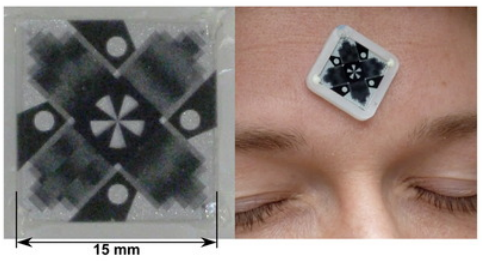

Figure 3.12: The tracking camera observes the marker with a special Moire-pattern that is attached to the forehead of the subject. Image courtesy of [27].

A commonly employed motion tracking method involves using optical cameras which are mounted near the opening to the bore of the scanner, or alternatively in the scanner room (Figure 3.11), such that a direct line of sight from the camera to the acquired object is maintained. The camera tracks the special marker, which is attached to the head of the human subject.

Moire-pattern markers (Figure 3.12) allow for high-precision estimation of both translational and rotational motions. In order for tracking to be accurate it is essential to attach the marker in such a way that it is rigidly connected to the subject.

An efficient way to achieve a rigid fixation is based on the use of bite bars that have

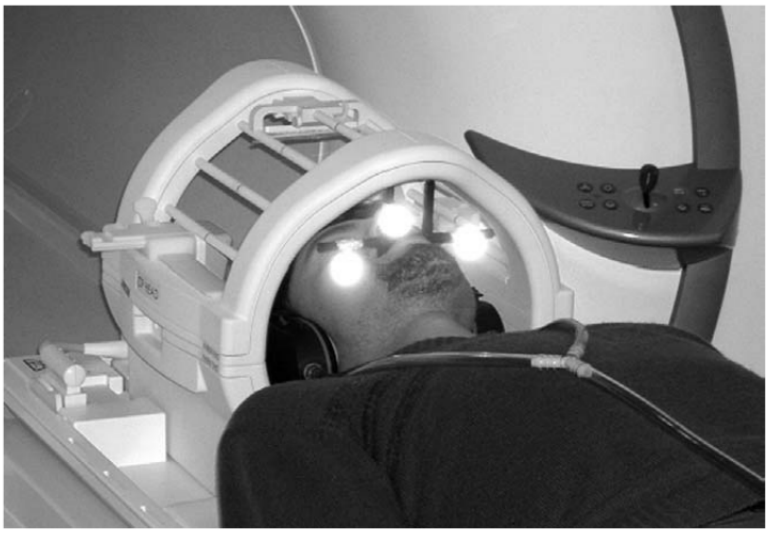

Figure 3.13: The tracking targets can be attached to a rigid mouth-bite piece for a better fixation. Image courtesy of [26].

markers attached to them (Figure 3.13). This is the method of choice in high-field (7T) ultra-high-resolution image acquisitions, which are sensitive to subject motion in sub-millimeter range. Bite bars, however, can lead to significant discomfort for the person in the scanner, and can be compared, in this respect, to rigid fixation methods that are used to physically constrain the movements of the patient.

Alternatively, a marker can be attached to the glasses which the subject wears during the scan. Although this method of marker fixation is associated with only a negligible discomfort it results in less accurate estimates of motion parameters because the glasses are not rigidly attached to the head. A marker can be also attached to the skin of the forehead. This is even better in terms of subject convenience but can lead to severe tracking errors if the subject contracts facial muscles.

Instead of using motion tracking equipment, the displacements of the imaged object can be measured by the scanner itself. The idea relies on modification of the acquisition sequence that adds additional navigator read-outs before each repetition [12, 28]. The navigator images are low-dimensional and can be acquired very quickly (within a second). Registering the navigator images against each other allows the displacements to be inferred and this information can then be used for online adjustment of encoding gradients and prospective correction [29]. This has the benefit of causing no discomfort to the patient; however, the accuracy of estimated motion parameters is lower compared to the accuracy of specialized tracking equipment. The

Figure 3.14: Active markers (field probes) can be used to determine position of the head without using tracking cameras. Image courtesy of [30].

accuracy can be made nearly perfect by attaching the active markers (field-probes) rigidly to the subject (mouth-piece), and then using the special sequence that tracks the position of the active markers [30] (Figure 3.14). Again, this is advantageous, since no extra tracking equipment is required; however, it is associated with all sorts of problems in terms of marker fixation as mentioned above.

While causing no discomfort to a patient, the navigator techniques involve severe modifications to the imaging sequences that usually lead to prolonged acquisitions. Furthermore, not all imaging sequences can easily be equipped with navigator readouts. For this reason such techniques have a limited scope in clinical practice.

Another technique that utilizes scanner capabilities to detect motion is based on Free Induction Decay navigators. The idea is to sample the FID signal with a high temporal resolution (at each repetition). Since the FID readout is made extremely fast, using this navigation technique does not lead to an increase in the image acquisition time. Importantly, implementing FID navigator readouts requires minimal modifications to the imaging sequence and is compatible with most sequences.

Although the FID signal from a single coil is not sufficient to estimate the motion trajectory it can be reliably used to find the time points, where large displacements have taken place. Acquiring the image with multiple coils (parallel imaging) allows several FID signals to be acquired which are localized to different spatial parts of the imaged object (as dictated by the position of each coil). This technique can be described as allowing for acquisition of very low spatial resolution representation of the object with high temporal resolution. It was recently shown that if ground-truth motion parameters are known by means of optical tracking, it is possible to

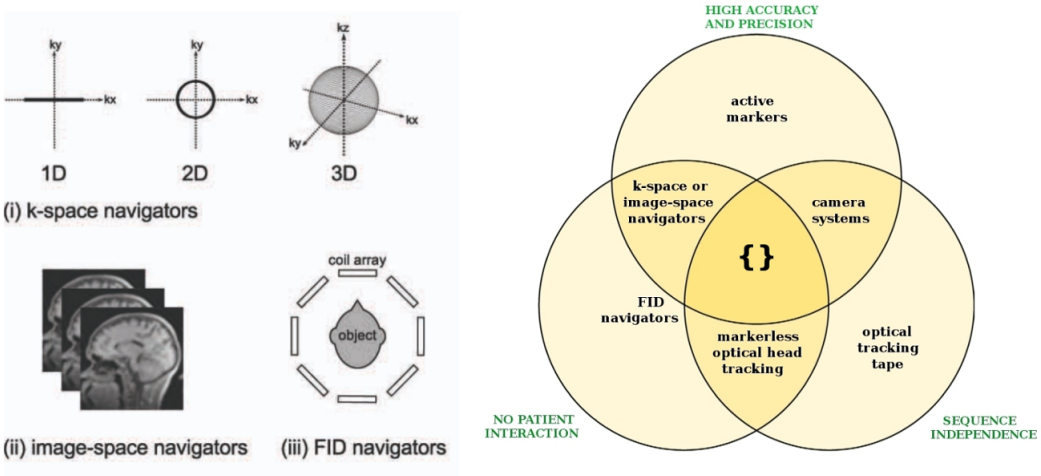

Figure 3.15: Advantages and disadvantages of different motion tracking methods. Image courtesy of [25].

find a linear map from FID signals to motion parameters. Thus FID signals from multiple channels can be used for the accurate prediction of motion parameters. The problem is that linear mappings are subject-dependent and need to be recalibrated in each scan. The linear maps depend on both the magnitudes and phases of acquired images. To further complicate things, even when imaging the same object at different times, it is possible that phase in the acquired image will change. This will require using the tracking cameras again and doing recalibration and linear map estimation. So at this stage the relevant technology requires significant improvement.

The overview of the different motion detection and tracking methods and their evaluation with respect to the comfort of the patient is shown in Figure 3.15.

3.4.2 Retrospective motion correction

Motion correction methods that are proposed in current work belong to the class of retrospective motion correction methods. There are two important limitations on prospective methods: firstly, they require additional tracking hardware to be used during image acquisition and secondly, they allow for correction of exclusively global affine motions. The second limitation is due to the fact that encoding gradients adjustment only allows for global transformations to be performed on the acquired volume meaning that non-rigid motion (with the exclusion of scaling and shearing) cannot be prospectively corrected. An example where prospective correction might

fail involves a jaw motion during a head scan. Since a jaw is not rigidly attached to the head, such motion can cause complicated non-local artifacts that cannot be removed by prospective correction systems. More complex and also more practically relevant scenarios comprise abdominal and fetal imaging when using prospective correction is problematic not only because of the global correction problem [25] but also because tracking the motion of the organs in an abdominal cavity with conventional methods such as optical cameras is not possible. To address such kind of problems the retrospective motion correction techniques are the methods of choice.

A common problem of retrospective methods is potentially long computation times whenever high-resolution images need to be corrected. Another problem already mentioned above has to do with the fact that certain types of motion cause inadequate sampling of k-space and missing spectral data. To a certain extent this can be compensated for by using prior knowledge on a typical medical image. A related way to address this problem is to use high capacity artificial systems (such as neural networks) trained on large databases of medical data in order to learn characteristic properties of medical images; this knowledge can then be used to aid image restoration. The second mechanism for data loss involves cases in which, due to pronounced motion, the imaged object partially leaves the FOV. This issue is of minor concern in practice though, since usually the FOV is configured in such a way that the imaged object is "insulated" by low proton density matter such as air, and motion in centimeter range is required to cause out-of-FOV motion induced spectral data corruption.

Finally, an important limitation of existing retrospective methods is the requirement to have raw data as input. Medical MRI data is commonly stored in DICOM format, where only the spatial magnitude of the image is preserved, and the phase is discarded. This has the effect on k-space that each line no longer experiences transformation characterized by just six motion parameters. Instead, motion at each time point of the acquisition now affects all k-space coefficients in a complicated way.

The problem of truncated input becomes even more severe in parallel imaging in which multiple coils are used to acquire the volume within a shorter time (acceleration) or with improved SNR. Thirty-two channel systems are associated with a large amount of raw data that requires gigabytes of storage space. For this reason, it is common to store just a single DICOM volume with images from each coil combined using the Sum-of-Squares method that also eliminates the spatial phase. Doing motion correction on combined multi-coil data is even more complicated since, in such

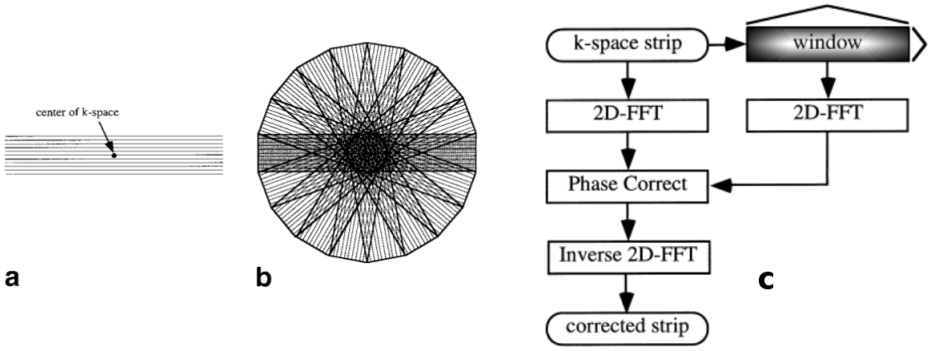

Figure 3.16: Self-navigating methods rely on using the navigator data for both motion estimation and image reconstruction. (a) A single strip which has only a few phase encode lines, thus a spatial domain representation of the strip has a very low resolution in the phase-encoded direction; (b) rotated strips, which constitute a full set of PROPELLER trajectories; (c) motion correction pipeline based on redundant PROPELLER trajectories. Image courtesy of [31].

circumstances, most of the raw data and not just a phase is missing. Ideally, the future approaches to retrospective motion correction need to be able to operate on such classes of inputs.

3.4.2.1 Self-navigation

Retrospective methods based on the self-navigation principle rely on the use of special imaging sequences to acquire the redundant data. Different chunks of such data can be registered against each other and motion parameters estimated - Figure 3.16. Importantly, the same reference data can also be used for other purposes, i.e. to improve the SNR of the image.

3.4.2.2 Combined prospective-retrospective correction

Tracking systems that are used for prospective correction usually never yield completely accurate estimates of motion parameters, and the acquired image has residual motion artifacts. The sources of the errors are tracking noise and marker attachment instability. The problem of residual artifacts can be solved by combining prospective and retrospective correction methods. The resulting combined method is very powerful since it is both tolerant to the problem of spectral data loss due to strong motion (due to adequate sampling enabled by prospective correction) and

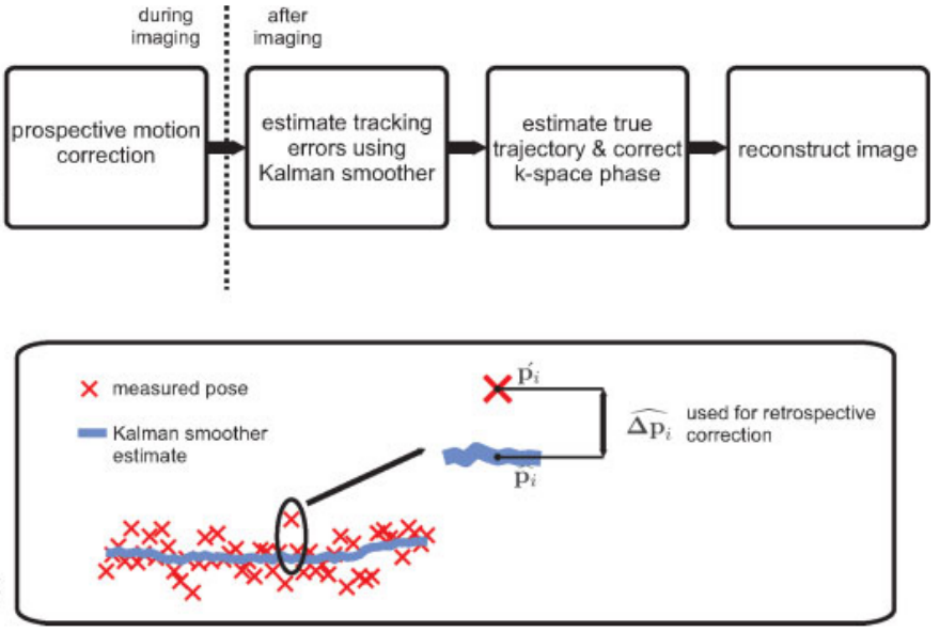

Figure 3.17: Combining prospective and retrospective motion correction methods allows for compensation in the motion tracking errors. Image courtesy of [32].

the problem of artifacts due to tracking errors, which can be efficiently removed by retrospective correction methods (Figure 3.17). Thus prospective and retrospective methods are in fact compatible and complimentary, and future methods of motion correction might involve a combination of both approaches.

3.4.3 Autofocusing-based approaches

An important subclass of retrospective motion correction methods that is of special relevance to the current work is the autofocusing methods. Such methods operate by applying some transformation to the k-space in a such a way that an appropriate image metric is optimized in the spatial domain. The seminal work was carried out by David Atkinson, who considered a closely-related problem in radar imaging (Figure 3.18), and suggested the use of a similar solution to solve MR motion correction problems retrospectively. It was assumed that in a valid motion-free medical image the entropy of the image intensities should be low. The reason for this is that medical images have a special structure, which involves strong edges and flat intensity regions. Motion has an effect of smearing the edges with a blur effect and producing global ghosts which cause a lot of pixel-per-pixel variation. It was empir-

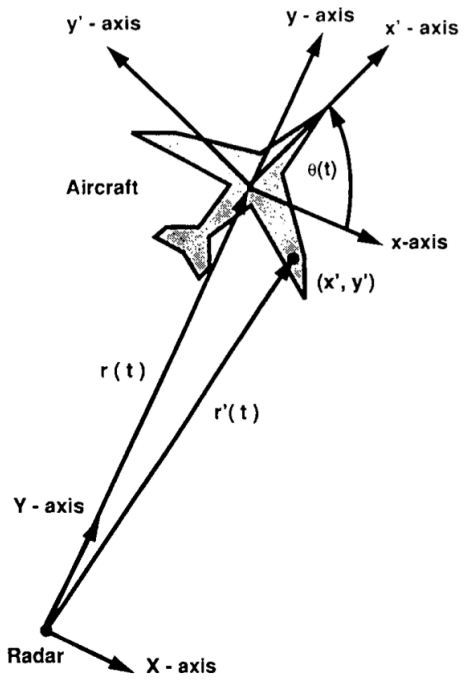

Figure 3.18: The autofocusing motion correction technique has a connection to the problem of radar imaging. Image courtesy of [17].

ically observed that the entropy of the image affected by motion artifacts is much larger than the entropy of the motion-free scans. The idea of auto-focusing methods is to try to find some motion parameters such that, if the image is corrected for the artifacts using the given motion parameters, the entropy in the spatial domain is minimized.

It is helpful firstly to consider a simpler problem of pure translational motion. In this case, the effect of motion on each k-space line can be described by multiplication with a phase ramp, whose slope is equal to the amount of spatial displacement. Thus, if the motion parameters for the entire trajectory are known, it is possible to construct the inverse phase ramps that, when applied to each line of k-space, will cancel the effects of motion. Thus, given purely translational motion, motion correction is a simple task, and the main problem is to actually find unknown motion parameters. In an early autofocusing approach by [18] this problem was solved by probing different motion parameters in an extensive combinatorial way. Firstly, the adjacent k-space lines are grouped together and assumed to be affected by the same motion (Figure 3.19). Then, for each group of these lines the method applies a set of trial displacements, which comprises discrete steps of [-5,-4,-3,-2,-1,0,1,2,3,4,5] pixels.

In the case of two-dimensional images, such displacements need to be performed in both spatial directions. For each trial displacement the image is converted to spatial domain by applying the inverse Fourier transform, and the entropy of the image is computed. The displacement that results in the lowest entropy is assumed to describe the actual movement of the imaged object. Once the motion parameters are estimated coarsely for the large blocks of k-space lines, the next step involves blocks with a smaller number of lines. Finally, each individual line is probed with trial displacements. Such a coarse-to-fine framework allows the use of motion parameters estimated in coarse grouping scales to constrain the search in finer groups.

In a similar way, in the case of rotational motion, an adaptation of the technique described above can be employed. It is no longer sufficient simply to apply the phase ramps to the frequency data in order to correct for motion. Since rotational motion causes displacements of each acquired k-space line around the DC component, in order to correct for rotations, it is necessary to do resampling of the spectral data using the interpolation. Given the rotational parameters, first the deformed k-space grid of sampled points is constructed. Then interpolation is used to estimate the coefficients on the deformed grid from the regular grid of k-space coefficients. This is a forward operation that simulates the effects of rotational motion. In order to actually correct for rotational motion the procedure should be inverted. This was done in [18], the follow-up work, for which not only the translational displacements, but also the different trial rotation parameters, were probed. For the different sets of motion parameters it is thus necessary to do the resampling in k-space, compute the inverse FT, and estimate the entropy metric in spatial domain.

The method described above has generated a plethora of follow-up work, and was used with different MR imaging sequences and in studies of various human body parts. The main problem with the method was found to be the computational complexity, which is due to the fact that the method is based on the combinatorial exploration of the space of motion parameters. For each motion parameter vector to be tested, it is necessary to do the Fourier transform and estimate the image metric. This is computationally prohibitive, especially in the case of high-resolution 3D volumes and complicated motions.

A crucial question when it comes to autofocusing methods is what image metric to use. There are two requirements for a good image metric: firstly, it should be sensitive to the image data. In other words, the image metric should capture the characteristic properties of the typical medical data. The second requirement is that

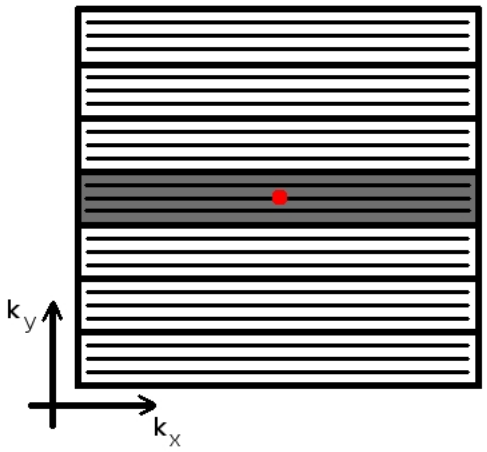

Figure 3.19: The autofocusing technique by [18] relies on the selection of a block in k-space and applying trial motion transformations to the lines within this block. The image metric is then calculated in spatial domain, providing an evaluation of how good the different trial motion parameters are.

the image metric should be a good detector of the artifacts due to motion. Even if the first requirement is not fully satisfied, and the image metric does not encode the medical image statistics well, if this metric is sensitive to the artifacts caused by motion it can still be used to do the motion correction.

An important study by [33] has actually addressed the question of what image metric to use for autofocusing retrospective correction of MR images (Figure 3.20). In this study, different metric functions were applied to medical images that were corrupted by varying degrees of motion. Additionally, the expert radiologists were asked to rate the motion corrupted and motion-free images. The ratings of the experts were then compared to the objective values of the metric functions and, in this way, it was decided what image metrics best capture the properties of medical images. The main result of this research was that the entropy of the image gradients, and a normalized gradient squared, are the metrics best congruent to the expert radiologists' ratings.

3.4.4 Non-rigid motion correction

An important class of motion correction problems that can be exclusively addressed by retrospective methods is in cases where the object is affected by non-rigid motion. Motion due to the physiological function of the organism is unavoidable even if

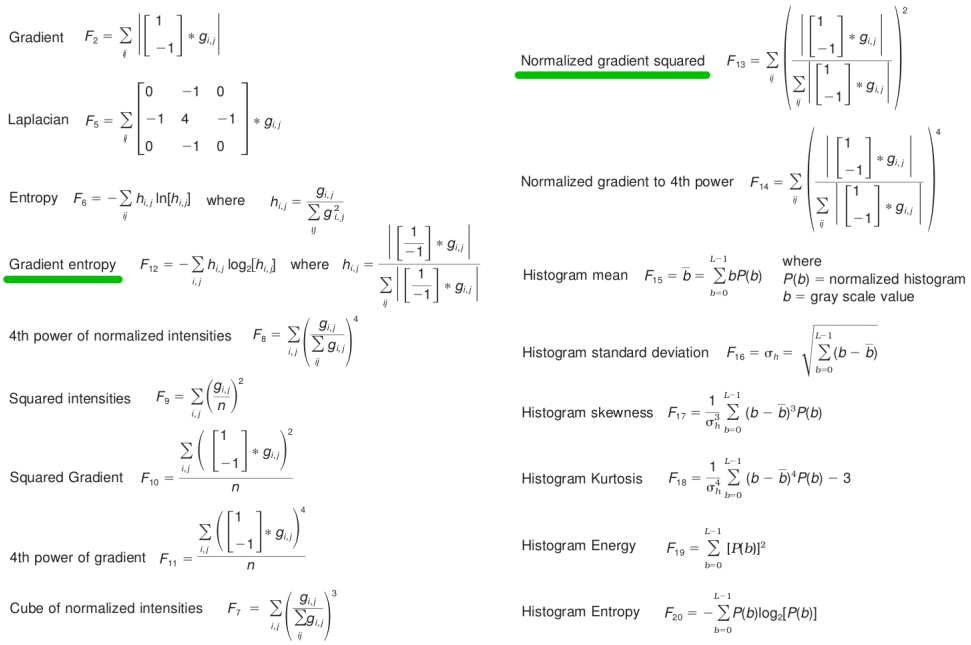

Figure 3.20: Different objective image metrics that were evaluated in [33].

the subject is sedated. In particular, respiration and heartbeat are of primary concern for thoracic and abdominal MR imaging. Motion artifacts due to respiration can be reduced with breath-held scans [34], but this imposes many limitations on acquisition, since scan times have to stay within tolerable breath-holding periods. Further examples of physiological motion include bowel movement, cardiovascular pulsation and even movement within the brain, characterized by non-rigid pulsations in the thalamus [35]. Usually, physiological motion leads to complex artifacts, which cannot be properly addressed by motion correction methods that assume the scanned object to be a rigid body. Prospective motion correction methods suffer from the global correction problem (see [25]), and are restricted to global affine motions, which include non-rigid shearing and scaling [36]. Most of the research on non-rigid motion correction has focused on developing retrospective methods, which correct for motion once the data is acquired.

The early research of [37] shows that it is possible to correct for non-rigid motion in kinetic joint studies. The motion estimation does not require the use of additional measurements and is driven by projections from a segmented radial acquisition. In their seminal paper, Batchelor et al. [22] derive a general description of effects of arbitrary (also non-rigid) motion during MR imaging applicable to standard sequences

and parallel imaging acquisitions. Motion degradation is modeled as a linear process described by matrix-vector multiplications. The authors show that it is possible to find an estimated inverse of the linear system, and thus to retrospectively recover the underlying sharp image. Their approach is of theoretical interest because they found it computationally prohibitive to estimate the unknown motion parameters. To solve this problem, external motion reference signals obtained from navigators were proposed [38]. This idea turned out to be fruitful and was later used for non-rigid motion correction in coronary MRI [39], cardiac CINE [40, 41, 42], and liver MRI [43].

A further important step was made by Odille et al. [24], who treat the image reconstruction and non-rigid motion parameter estimation as two coupled inverse problems. However, models for patient motion driven by external sensors such as ECG or navigators still need to be used to make the motion parameter estimation feasible. In follow-up research [44], how to reduce the number of unknown motion parameters with the use of adaptive meshes is shown, which involves grouping voxels with similar displacements and thus encoding local smoothness information. In spite of this, information from external motion sensors is still needed for reconstruction.

A recent work [23] proposes that non-rigid motion be approximated as simple linear translations in small spatial regions to reduce the problem complexity – Figure 3.21. Indeed, given that spatial patches are limited to the size of a single pixel, any arbitrary motion can be described by the optical flow or displacement vectors specified for each pixel. This allows for sufficiently fast implementation of a non-rigid motion correction algorithm, which uses butterfly navigators to estimate the unknown motion parameters. More recently [45], non-rigid motion correction was combined with compressed sensing to address the problem of respiratory motion during free-breathing cardiac acquisitions. Again, a navigator signal is required to estimate the motion.

3.5 Overview of related work

Presented below is a table which shows certain important aspects of existing retrospective methods in a structured way. Associated with each entry in the table is a short description of the core idea of the method.

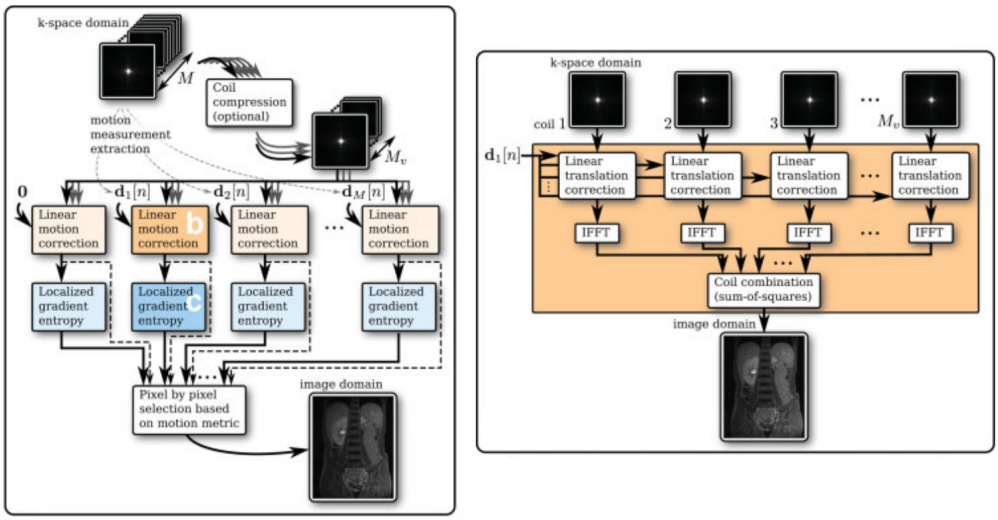

Figure 3.21: Non-rigid motion correction with the use of butterfly navigators. The complex non-rigid motion is locally approximated by the linear translational motions. Image courtesy of [23].

	REAL	NOREF	ANYSEQ	AUTOFOCUS	ROT	3D	ANYMOV	NONRIGID
Our method	yes	yes	yes	yes	yes	yes	yes	yes
1. Hedley et al. (1991)	yes	yes	yes	no	no	no	no	no
2. Atkinson et al. (1997)	yes	yes	yes	yes	no	no	no	no
3. Manduca et al. (1998)	yes	yes	yes	yes	no	no	no	no
4. Atkinson et al. (1999)	yes	yes	yes	yes	yes	no	no	no
5. Manduca et al. (2000)	yes	yes	yes	yes	no	no	no	no
6. McGee et al. (2000)	yes	yes	yes	yes	no	no	no	no
7. McGee et al. (2001)	yes	yes	yes	yes	yes	yes	no	no
8. Atkinson (2001)	yes	no	yes	no	yes	no	yes	no
9. Kholmovski et al. (2002)	yes	no	no	no	yes	no	no	no
10. Bourgeois et al. (2003)	yes	no	yes	yes	yes	no	yes	no
11. Atkinson (2003)	yes	no	yes	no	yes	no	no	no
12. Lin et al. (2005)	yes	yes	yes	no	yes	no	yes	no
13. Batchelor et al. (2005)	yes	no	yes	yes	yes	yes	yes	yes
14. Lin and Song (2006)	yes	yes	yes	yes	no	no	no	no
15. Lin et al. (2007)	yes	yes	yes	yes	yes	no	yes	no
16. Odille et al. (2008)	yes	no	yes	no	yes	yes	yes	yes
17. Lin and Song (2009)	yes	yes	yes	no	yes	yes	yes	no
18. Lee et al. (2009)	no	yes	yes	yes	no	no	no	no
19. Su et al. (2010)	yes	yes	yes	no	yes	yes	yes	no
20. Samsonov et al. (2010)	yes	yes	yes	no	yes	no	no	no
21. Johnson et al. (2011)	yes	yes	no	no	yes	yes	yes	no
22. Maclaren et al. (2011)	yes	no	yes	no	yes	no	yes	no
23. Kadah (2011)	yes	yes	no	no	yes	no	yes	no
24. Cheng et al. (2012)	yes	no	yes	yes	yes	yes	yes	yes
25. Menini et al. (2012)	yes	no	yes	no	yes	yes	yes	yes
26. Yang et al. (2013)	no	yes	yes	yes	no	no	no	no
28. Our method	yes	yes	yes	yes	yes	yes	yes	yes

The descriptions of the abbreviated column names are as follows:

REAL: was the method tested on real MR data? "No" means that the method was only tested on simulated data.

NOREF: can the method operate without motion reference, i.e. navigators or motion sensors? "Yes" means that the method is blind, and purely data-driven.

ANYSEQ: can the method be used with generic Cartesian sequences? "No" means that the method is based on the special acquisition sequence.

AUTOFOCUS: does the method rely on using an image quality metric?

ROT: does the method allow correction for rotational motion?

3D: does the method allow correction for motion in 3D volumes?

ANYMOV: can the method be applied to correct for any arbitrary motion? "No" means that the method relies on some assumptions on the motion trajectory.

NONRIGID: can the method be used to correct for non-rigid motion?

1) Motion Artifact Correction in MRI Using Generalized Projections [10]

To correct for the artifacts in 2D MR images due to translational motion the authors used an adaptation of the Gerchberg-Saxton algorithm which aims to recover the spatial phase from both image domain and spectrum magnitude information. The idea is as follows: since translational motion does not affect the power spectrum but only the phase of Fourier coefficients, Gerchberg-Saxton's method can be used to correct for spatial domain motion artifacts resulting from phase errors. Since the problem is ill-posed, prior information needs to be used. The prior information is based on the assumption that, since the Field-of-View is usually larger than the head of the subject, the clinically relevant part of the image must have a small and compact support. This assumption allows for use of a Hayes theorem that states that a two-dimensional sequence (such as an MR image) that has a finite support and irreducible z-transform can be recovered from the magnitude information. The recovered image in spatial domain is unique to within a sign change, translation and a coordinate system flip. The authors note that the z-transform of the acquired data is a two-dimensional polynomial, which is almost always irreducible and which allows for recovery of the phase from the magnitudes. Based on the idea described above, the authors came up with an algorithm that requires approximately 20 iterations to attenuate the motion artifacts. Furthermore, the algorithm was shown to be stable in the presence of noise.

2) Automatic Correction of Motion Artifacts in Magnetic Resonance Images Using an Entropy Focus Criterion [18]

The authors aimed at correcting 2D MR scans for artifacts due to translational rigid motion using an autofocusing technique. The idea was to use an image metric that prefers images that are free from motion artifacts. The authors proposed the use of an entropy function for this purpose, and noted that this favors the data with a larger number of low intensity pixels. Thus, such a metric penalizes the motion-corrupted data which is affected by ghosts and blurring that are known to reduce the number of non-dark pixels. Using the image quality criterion the authors came up with an optimization strategy that iteratively recovers the unknown motion trajectory. On each iteration, the motion artifacts are corrected by adjusting the spectral phase, and the entropy of the image in spatial domain is estimated in order to evaluate how good the recovered motion parameters are. The motion is assumed to be inter-view so that each k-space line is affected by a corresponding displacement which is given by the position of the patient during the acquisition of the view. The search of motion parameters is based on an extensive exploration by applying trial motion displacements for each k-space view. In order to make the search faster, the authors use an ad hoc technique for grouping k-space views firstly in large blocks, and then progressively making the blocks smaller up to the point of using just two views. The authors showed that this algorithm is efficient at correcting the translational motion artifacts in 2D images. The method requires a raw complex-valued image as input.

3) Autofocusing of Clinical Shoulder MR Images for Correction of Motion Artifacts [46]

In this work, the authors used the method of [18] (described just above) in order to correct motion artifacts in images of real clinical patients. The target data are high resolution 2D shoulder images. Instead of using the entropy focus criterion, the authors used the entropy of the image gradients (first order finite derivatives) asserting that this new metric leads to better results. The authors noted that the method operates remarkably well, even though it assumes no knowledge of the actual patient motion trajectory. Additionally, the authors compared the autofocusing motion correction results against the navigator echo based correction. They observed that, although autofocusing-based correction was 90% as efficient, it never degraded the image.

4) Automatic Compensation of Motion Artifacts in MRI (Atkinson et al., 1999)

This work is a follow-up to [18], an approach that presents an extension of the algorithm that is capable of correcting for in-plane rotations. The presented algorithm was shown to require substantially shorter computation times, which were in the order of the time required to do an additional scan of the patient. The algorithm features a new kind of multi-resolution search of motion parameters which is based on sequential 1D searches in each of the degrees of freedom. The rotational motion is corrected by re-gridding; however, this computationally expensive step only needs to be applied once for each rotational angle.

5) Autocorrection in MR Imaging: Adaptive Motion Correction without Navigator Echoes [14]

This work is an extension of previous research in [46], and also aims to further extend the autofocusing-based correction of coronal images of rotator cuff. The results of the study confirm the previous finding that motion correction by means of autofocusing is comparable to navigator echo based correction, although it is blind with respect to the underlying motion. The method still has some limitations such as inability to deal with through-plane motion as well as 3D volume data. Furthermore, the computation times were observed to be too long, taking around 20 minutes of the UNIX workstation operating time for each slice.

6) Image Metric-Based Correction (Autocorrection) of Motion Effects: Analysis of Image Metrics [33]

This study aims to evaluate and compare different image quality metrics that can be used for autofocusing-based correction. Four expert radiologists rated 164 clinical rotator cuff images with and without motion artifacts. These scores were then correlated with objective criteria of various image metrics. Both entropy of the gradients and an autocorrelation-based metric were found to be good indicators of the image quality which can be used to guide autofocusing-based correction in order to avoid solutions affected by motion artifacts.

7) Autocorrection of Three-Dimensional Time-of-Flight MR Angiography of the Circle of Willis [21]

The authors made two assumptions: the first assumption was that only abrupt (non-continuous) motion occurs during acquisition of 3D time-of-flight MR angiography data; and the second assumption was that motion along the Z axis is negligible for each k_y phase-encoding step. If the angle of rotation is small and the region of interest is far enough from the fulcrum of rotation then rotation of an object in this region can be broken into equivalent translations along the two in-plane axes. The authors applied this approximation to the motion corrupted data so that instead of regridding the k-space data, a translational motion model with linear phase correction factors was used to correct for 'equivalent' linear translations. The auto-correction of 3D time-of-flight data was performed by dividing k-space into blocks, in a similar manner to that used by autofocusing approaches described above. The authors used the entropy of the gradients image metric. The results of this study showed that auto-correction can improve the quality of 3D time-of-flight MR angiography MIP images, and that auto-correction does not adversely affect images without motion-induced artifacts. The results also indicate that the auto-correction algorithm is most effective at improving image quality when motion is in the axial plane.

8) Automatic Motion Correction Using Prior Knowledge [47]

In this study, the authors consider using multiple 2D images of the same subject in order to aid motion correction. Indeed, it is common to acquire multiple volumes in order to improve image SNR by averaging. Furthermore, multiple images with different contrast sequences (T1/T2 weighted) can be acquired. This means that some of these images that are less corrupted by motion can be used to guide the correction of the motion-affected ones. In order to choose the image least affected by motion artifacts the authors propose the use of the entropy criterion. The selected image is then used as a reference in order to compare each k-space view from the motion corrupted data against the reference ground-truth. This allows for estimation of translational and rotational motion parameters, and then the use of the shearing-based rotation method in order to remove the effects of motion in degraded data.

9) Motion Artifact Reduction Technique for Dual-contrast FSE Imaging [48]

The authors firstly assumed that motion events are sparse during scans, meaning that only a fraction of acquired MR data is motion corrupted. The second assumption was that the effects of motion manifest themselves in deviations of the scanned object position from the baseline. Given these assumptions, the authors considered the problem of motion correction of dual-contrast 2D FSE images, which generates PDw and T2w parts. The k-space data of both images needs to be similar and spatial phases consistent which allows for identification and recovery of k-space lines corrupted by motion. The identification is based on the correlated features in the image background which are caused by motion. In order to segment out the background, the authors rely on an image support mask, which needs to be known. This mask is applied in the spatial domain in order to zero out the object and leave only the background, which is then Fourier transformed. This has an effect on corrupted k-space lines which appear as peaks in k_x integrated power spectrum. This method allows convincing results to be achieved; however, it cannot be used to correct for motion that affects more than half of k-space.

10) Retrospective Intra-scan Motion Correction [49]

This method operates by first estimating the in-plane translations and rotations that have occurred between measurements of the data along the successive trajectories in k-space, correcting for these motions and eventually finding the motion-free image. Two important problems were addressed in this area: (1.) the automatic estimation of motion parameters from k-space data along individual trajectories and (2.) the accurate reconstruction of images from pseudo-randomly sampled k-spaces. Firstly, the intra-scan translation and rotation parameters for each dataset in a series are estimated by comparing the potentially corrupted dataset with a reference acquisition chosen within the series. Due to the rotation effects, the resulting "true" sampling grid is pseudo-random. After correcting for translational motion, the "motion-free" image is reconstructed using the true sampling grid. The image reconstruction uses Bayesian estimation, originally developed to deal with under-sampled and irregular MRI acquisitions. The present analysis has focused on 2D acquisition with Cartesian sampling (spin-warp imaging). To estimate the rotation angles the magnitudes of the distorted signal and of the reference signal are compared for each trajectory, which is possible since translations do not affect the magnitude of the distorted

signal. The comparison is done by maximizing the correlation coefficient between the distorted signal and the reference signal, re-gridded onto the rotation-deformed grid. The first step of the algorithm aims to optimize the quality of one particular acquisition in the series (for instance the first one) by correcting the artifacts due to the translations (this operation does not require any reference signal) and using the corrected data as reference data for the series. To correct the translation-induced linear phase differences of the signals along k-space trajectories, the authors use their own implementation of the algorithm of Gerchberg and Saxton. The accuracy of motion estimates assessed by the Monte Carlo studies is better than 0.1° for rotation, and 0.1 and 0.05 pixel respectively for translations along the read and phase directions. The algorithm reduces the distortions due to rotations down to the noise level.

11) Reconstruction after Rotational Motion [50]

The developed technique converts a 2D problem into a series of 1D re-griddings, each of which is solved using a matrix inversion. It then identifies under-sampled regions that prevent stable inversion, and inserts additional data into these regions to allow stable image reconstruction. Although the method is not a genuine pure data-driven autofocusing since it requires the knowledge of motion parameters, nevertheless, it can be seen as an important step towards correction of strong and temporally sophisticated rotational motion.

12) Correcting Bulk In-Plane Motion Artifacts in MRI Using the Point Spread Function [8]

Using 1-pixel-large markers to measure a spatial blurring kernel due to motion allows for correction of translational motion artifacts by applying a phase factor to the corrupted k-space data whose complex conjugate is the Fourier transform of the point spread function (PSF). Furthermore, by measuring the relative shift of two PSF markers, object rotation can be also estimated and data subsequently corrected. The proposed method relies on accurate measurement of the PSF and, therefore, robustness to noise is critical. As demonstrated in both phantom and in vivo examples, the proposed technique provides means to detect and subsequently correct general motion, including motion with no continuity between adjacent views.

13) Matrix Description of the General Motion Correction Applied To Multishot Images [22]

Complicated motion patterns, which involve non-rigid deformations, do not have a simple description in Fourier domain. Nevertheless, the authors showed, for the first time, that it is possible to correct such classes of motions retrospectively. The authors came up with a new mathematical formalism in which non-rigid motion effects are described using a large matrix that is applied to the image in spatial domain. Inversion of such matrix is hypothesized to result in attenuation of motion artifacts. The presented approach is of theoretical interest since it requires the motion parameters to be known. Given the motion parameters, it is possible to solve the huge, underlying linear system, but an efficient method is required to achieve this in a practically feasible time. The authors propose the use of the LSQR algorithm which is a robust implementation of the conjugate gradient of the normal equation. The results indicate that it is possible to correct for non-rigid types of motion that are commonly encountered in clinical practice, but long computation times pose an issue.

14) Improved Optimization Strategies for Autofocusing Motion Compensation in MRI Via the Analysis of Image Metric Maps [51]

One of the aims of this paper is to analyze the objective landscape (metric map) generated by the image metric when it is estimated for different translational motion parameters in the context of the autofocusing problem. In other words, evaluating the image metric for different trial displacements of two translational parameters generates a 2D image whose individual pixel intensities code for image metric values for any given translational displacement. The authors then analyze the patterns observed in the metric maps in order to develop an efficient optimization strategy to be used in the autofocusing algorithm with the final goal of achieving enhanced accuracy and shorter computation times. Two image metrics were considered: entropy and normalized gradient squared. Both metrics were observed to feature an interesting pattern: for the views located close to the zero frequency there was a higher absolute range of signal intensity (peak minus valley), which means that optimization of the motion parameters for the respective views is less susceptible to noise. A further unexpected observation was the existence of high-intensity maxima between the local minima in the metric maps. Importantly, an analysis indicated that local searches are prone to be trapped in local minima which are abundant in the metric maps.

15) Image Metric-Based Correction (Autofocusing) of Motion Artifacts in High-Resolution Trabecular Bone Imaging [20]

The authors adapted the method described in [18] in order to correct the motion artifacts in high resolution trabecular 2D bone scans. To correct for rotational motion, the shearing-based approach is used, implemented by three successive column-wise applications of the phase shifts. The authors used normalized gradient squared as an image metric. The optimization was carried out using a "base extension" approach which was based on progressive expansions of already corrected k-space views by two adjacent views until the entire k-space was corrected.

16) Generalized Reconstruction by Inversion of Coupled Systems (GRICS) Applied to Free-Breathing MRI [24]

In this work, the non-rigid motion correction framework described in [22] was extended to support the data input from multiple coil elements. The crucial requirement of [22] to have the initial knowledge of motion parameters was relaxed. The motion trajectory is described with the use of motion models governed by a reduced set of parameters which are obtained from the external motion sensors. The motion correction is then formulated in terms of two coupled inverse problems: image reconstruction (given motion) and model optimization (given image). In order to invert the coupled system, a fixed-point multi-resolution scheme was proposed. Major improvement of the image quality was observed; however, the optimization took 210 minutes for $256 \times 256 \times 32$ volume.

17) Extrapolation and Correlation (EXTRACT): a New Method for Motion Compensation in MRI [9]

The authors proposed an unusual way to carry out retrospective motion correction. The idea is to carry out a k-space extrapolation in order to generate a motion-free reference, which is then correlated with original k-space data in order to estimate motion parameters. The extrapolation is carried out by progressive expansion in a center-out fashion in k-space. Compared to the autofocusing-based approaches presented above, the proposed technique does not depend on a predetermined range of trial displacements. Furthermore, since image correlation can be performed very efficiently by doing pointwise multiplication in Fourier domain the EXTRACT ap-

proach is fast. Practically speaking, the technique was found to be stable against noise and capable of correcting both translational and rotational rigid motions in 3D images; however, the quality of reconstructed images was only moderately improved.

18) Compensation of Motion Artifacts in MRI via Graph-Based Optimization [7]

The authors proposed that the problem of correction of motion artifacts due to translational motion be treated as a blind deconvolution. Indeed, since translational motion has a multiplicative effect on the spectrum data, it means that it can be described as a convolution with a suitable kernel in spatial domain. This kernel, however, is complex-valued and is as large as a target image; this last property prohibits the use of state-of-the-art deconvolution or deblurring methods which become unpractical for large kernels. The proposed approach relies on the special structure of the translational motion PSF kernel whose magnitude contains multiple peaks corresponding to motion vectors. Strong edges in the image can be used to estimate such peak locations in the PSF kernel. In order to do an optimization, an auxiliary graph is constructed to indicate which k-space lines are affected by motion. The normalized gradient squared metric is used to define the energy of the graph, which allows the optimization problem to be solved and motion parameter estimates to be obtained. Practically speaking, the method was shown to be able to find closest integer motion vectors (approximating true motion, which is not integer-valued) leading to compensation of the motion artifacts.

19) MRI Motion Artifact Correction Based on Spectral Extrapolation with Generalized Series [52]

The EXTRACT method described in [9] was extended to allow for numerical optimization of the motion parameters. Additionally, it was proposed that the finite support with general series method be used instead of extrapolation. The improvement of the image reconstruction results over the original EXTRACT method was demonstrated.

20) POCS-Enhanced Correction of Motion Artifacts in Parallel MRI [53]

The idea of this proposed approach is to use the informational redundancy of the data acquired with parallel imaging in order to identify corrupted k-space segments.

This allows the motion affected data to be discarded and the remaining data to be used together with coil sensitivity profiles in order to recover clean images. The method showed reasonable performance in terms of image quality improvement in both simulated and real datasets.

21) Rigid Body Motion Compensation for Spiral Projection Imaging [11]

The authors proposed this method, which relies on a special trajectory involving a spiral-based acquisition of multiple 2D planes with different orientations. Since there are multiple overlaps of these planes it is possible to find the unknown motion parameters by looking at the intersections between the planes. Firstly, rotational parameters are estimated from the magnitude of the spectrum, which is known to be invariant to spatial translations. After the data is corrected for rotational motion translation, parameters are estimated. The experiments performed in simulation, phantom and real data environments demonstrated a high accuracy when using the method. Even with a relatively large range of motion (21.9 degrees and 12.6 mm from the baseline), the mean error was only 0.48 degrees/0.34 mm.

22) Combined Prospective and Retrospective Motion Correction to Relax Navigator Requirements [32]

Since prospective motion correction methods are prone to residual artifacts due to tracking errors, the authors proposed the use of an additional post-processing retrospective correction step. The Kalman filter is applied to the tracking data in order to approximate the true latent motion trajectory and find the error offsets in motion parameters. These extra offsets are then used to retrospectively correct the input data. The presented technique allows the requirement of having a high precision navigator to be dropped.

23) Overlapped *k*-Space Acquisition and Reconstruction Technique for Motion Artifact Reduction in Magnetic Resonance Imaging [54]

The idea of this method is to use a special imaging sequence in order to acquire multiple overlapping bands in k-space. The redundant spectrum data can then be used in order to estimate the motion parameters and correct for motion. The method was evaluated on both simulated and real data and showed promising results.

24) Non-rigid Motion Correction in 3D Using Autofocusing with Localized Linear Translations [23]

The authors noted that, in a sufficiently small spatial region, any kind of motion (rotation/non-rigid) can be approximated by simple translations. In order to correct for possibly non-rigid motion, an autofocusing algorithm based on a gradient-entropy metric was proposed. Since the search space of the motion parameters (given many local spatial regions) is enormous, the optimization needs to be guided by multi-channel navigator data. Experimental results showed an increase in sharpness of the images; however, after correction there were significant remaining residual motion artifacts.

25) Joint Reconstruction of Image and Motion in MRI: Implicit Regularization using an Adaptive 3D Mesh [44]

In this work, the method described in [24] was improved to enable mesh-based subject-specific smoothness constraints on the motion model. The idea was to solve for motion only in key points forming a mesh which can be seen as an implicit regularization. The mesh is automatically generated by an auxiliary algorithm proposed by the authors. Compared to the original work in this area, the method showed a faster convergence rate. The qual-ity of reconstructed images was shown to improve in both real and simulated data.

26) Sparse MRI for Motion Correction [55]

The authors proposed an optimization-based rigid motion correction method that seeks for the motion parameters that result in a corrected image to be sparse. Motion parameters and the latent image are jointly optimized using an iterative procedure. Promising results were obtained, but only on simulated data.

3.6 Some aspects of MR motion degradation

In this section, the results of experiments that simulated the process of motion degradation in MRI will be shown. Using the simulated environment allows for better understanding of certain aspects of MR image degradation due to motion and allows for evaluation of the image metrics for different types of motion trajectories. In the following discussion, only rigid motion is considered.

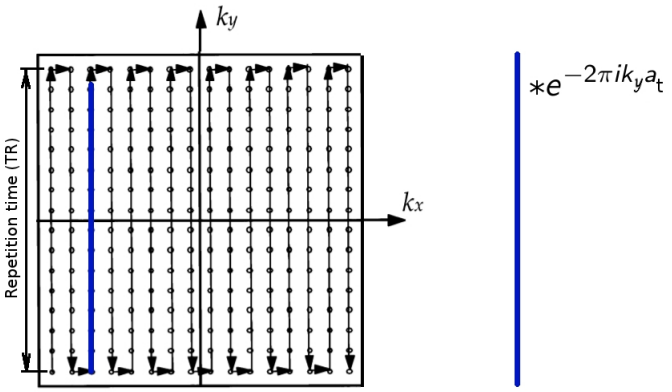

Figure 3.22: The translational motion only affects the phases in frequency domain. Mathematically, this corresponds to multiplying each acquired k-space line with a phase ramp whose slope is equal to the spatial displacement at the corresponding time point.

3.6.1 Translational motion

The effect of translational motion is described by the Fourier identity, which states that circular shifting of the signal in spatial domain corresponds to multiplication with a linear phase ramp in Fourier domain (see Figure 3.22). The slope of this ramp is proportional to the spatial displacement.

Alternatively, we can see the effects of translational motion as a convolution in spatial domain. This holds true since translations are described by pointwise multiplications in Fourier domain as is mentioned above. This means that a translational motion operator acting in spatial domain is a point spread function or a blurring kernel. Because of the nature of Cartesian MR acquisition (one k-space line at a time) this kernel has certain peculiar properties: firstly, it has a unit-valued power spectrum, which means that the image energy is preserved under translational transformation. Secondly, it is complex-valued in spatial domain, and thirdly, it has a large support in one dimension (phase encode direction), and small support in the other.

3.6.2 Rotational motion

Rotational motion is, conceptually, even simpler. If the zero frequency component is in k-space center, then rotation in spatial domain around the center of the image corresponds to the rotation in Fourier domain around the DC component. In

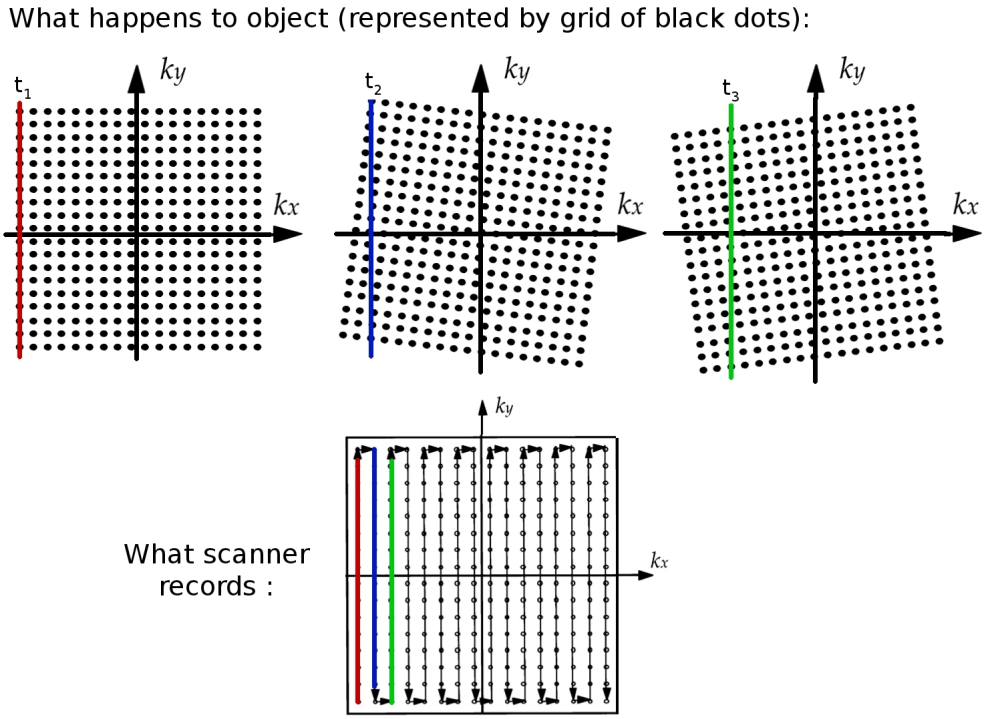

Figure 3.23: If the imaged object experiences rotational motion, this makes the scanner miss the intended spectral data. Rotation in spatial domain corresponds to rotation around the DC component in frequency domain. Thus, in the presence of rotational motion, the scanner acquires Fourier coefficients on the rotation distorted Cartesian grid.

Figure 3.23, the black dots correspond to the Fourier coefficients of the imaged object. The effects of rotation can be interpreted as non-Cartesian readouts dictated by a rotation deformed acquisition trajectory. Computation-wise, the implementation of the motion degradation due to rotational motion in the Fourier domain amounts to the gridding operation.

Dealing with a retrospective correction of rotational motion there are two basic problems to be solved. The first is the implementation of the forward process of the motion degradation involving rotations. Such implementation can be done both in spatial and Fourier domain, with the latter of the two being much faster to compute, especially given large 3D volumes. Computation-wise, the implementation of the forward process in the Fourier domain involves doing a Fourier transform followed by the gridding operation (resampling from a regular to a non-regular rotation-deformed grid). This is a well-studied operation, which is called a non-uniform Fourier transform of type II, see [56, 57]. The second problem relates to the inversion

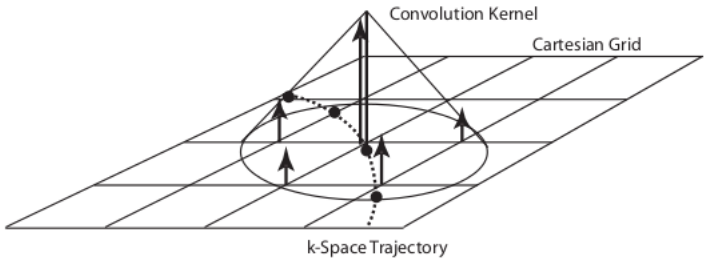

Figure 3.24: The interpolation can be seen as a convolution of the observed data with some convolution kernel. Image courtesy of [60].

of effects of rotational motion. This inversion can be done in closed form exclusively in the Fourier domain. Computationally, this amounts to a resampling from the non-regular to the regular grid followed by the inverse Fourier transform. This is a non-uniform Fourier transform of type I.

At the core of both the problems is the interpolation operation which needs to be done in order to estimate the values on either the regular or irregular grid depending on the type of the non-uniform Fourier transform. To carry out the interpolation, it is common to use a suitable function of the distance between the two points. Such an interpolating function returns a weighting that is applied to the value of some neighbor on a source grid in order to generate the interpolated estimate on some point of the target grid. This means that an interpolation can be represented as a convolution of the datapoints of the source grid with some interpolating kernel [58, 59] – Figure 3.24. The ideal kernel is given by a sinc function which causes no intensity attenuation in the spatial domain. Unfortunately, such a sinc kernel has a very large support meaning that the convolutions with such a kernel have a high computational cost. For this reason, it is common to use smaller kernels based on polynomials, Gaussians or Kaiser-Bessel functions. A common way to achieve higher precision in the gridding operation is to use the oversampled spectrum which can be achieved by zero padding in spatial domain. This results in a denser source grid of Fourier coefficients and higher precision in the interpolation operation. It also allows interpolating kernels with a smaller support to be used. The main drawback of oversampling is that it is more computationally demanding, requiring a factor of 2^d more data for processing, where d is the number of dimensions.

An additional complication, which is associated exclusively with a non-uniform

Fourier transform of type I, is due to the fact that the points on the source grid have a spatially-varying density. Some datapoints on the source grid can cluster together, while other points can be isolated with respect to their neighbors. Thus, an interpolation from the non-regular to the regular grid results in large values of the points on the target grid which are close to the clusters of points on the source grid. On the other hand, the points on the regular target grid, which are surrounded by only a few neighbors from the source grid, will have vanishing values after interpolation. This results in a non-uniform density of the interpolated target data, and can lead to all sorts of artifacts, which can appear as blurring or ringing in spatial domain. There are many approaches that can be used to address the problem of non-uniform density. A commonly employed method is based on using Voronoi tessellations [61] that are applied to the source grid. The idea is to split the grid into the cells that have areas of different size depending on the density of the resampled points in some region of the source grid. The areas of Voronoi cells can be used as correction factors to rescale the values on the target grid. The major problem with density compensation by Voronoi tessellations is a high computational cost especially in 3D data.

In practice, to do the gridding, Kaiser-Bessel window interpolation [56] is often used, but there is also fast Gaussian gridding that avoids the pre-computation of the interpolation weights [57]. Publicly available gridding software includes Jeff Fessler's `NUFFT` [1], Daniel Pott's `NUFFT` [2] and the Matlab toolkit by Matthew Ferrara [3]. Fessler's `NUFFT` is as fast as it can be once the underlying sparse matrix is constructed. If the matrix is needed only once, it is just as efficient to use Ferrara's code instead.

3.6.3 On complex-valued images and motion

When dealing with motion in MR, many challenges come from the fact that the images have a complex phase in spatial domain. Current retrospective approaches only allow for correction in cases where the raw data is preserved, meaning that, in spatial domain, complex phase is not zero. In practice, it is rather uncommon to store a spatial phase, and only the magnitude of the images is preserved in datafiles which are often saved in DICOM format. Since the blurring kernel (of

[1] `http://www.eecs.umich.edu/~fessler/irt/irt/nufft/`
[2] `http://www-user.tu-chemnitz.de/~potts/nfft/`
[3] `http://www.mathworks.com/matlabcentral/fileexchange/25135-nufft-nfft-usfft`

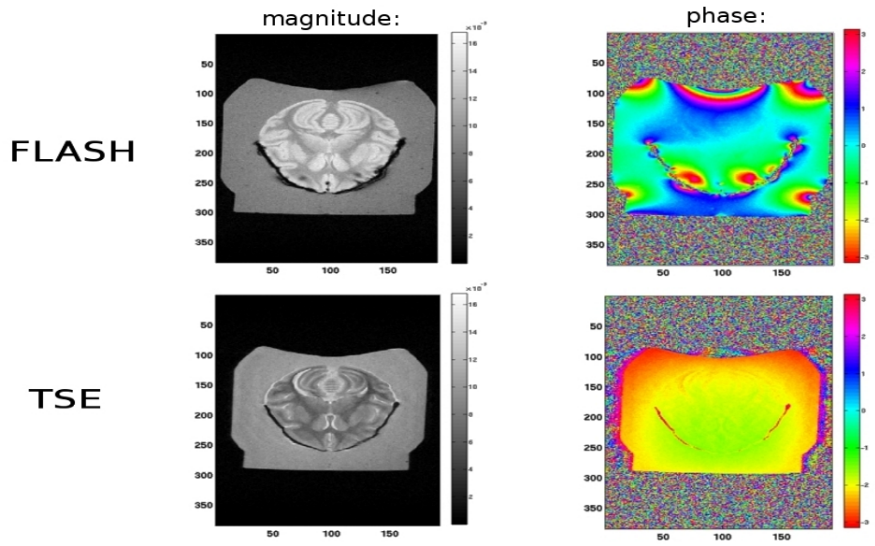

Figure 3.25: Due to the nature of the MR acquisition process, the image in spatial domain is complex-valued. The spatial phase depends on the sequence used and the magnetic field variations within the scanner.

the translational motion operator) is complex-valued, its interplay with intrinsic image phase results in many complicated image distortions and presents significant challenges for retrospective motion correction approaches. Usually, it is simpler to achieve better image quality improvement when the data is acquired with a sequence associated with a smooth spatial phase (such as Turbo-Spin-Echo) - see Figure 3.25.

3.6.4 Typical image artifacts in spatial domain

In the following experiment we restrict ourselves to the motion trajectory of a sinusoidal shape. Such motion trajectory causes regular and consistent effects on the Fourier coefficients in all frequencies, and is thus valuable for simulation-based experiments. Varying the amplitude of the sine it is possible to conveniently adjust the strength of motion. Shown in Figure 3.26 are the typical effects of the motion with sinusoidal trajectory on the image. The horizontal axis in the motion trajectory plot is the time or the repetition number, while the vertical axis shows the amount of displacement.

There follows a discussion of what happens to the image of the hash if the sinusoidal motion of a fixed amplitude is applied in different directions (Figure 3.27). For translations there are two possibilities: either, the object moves along k-space

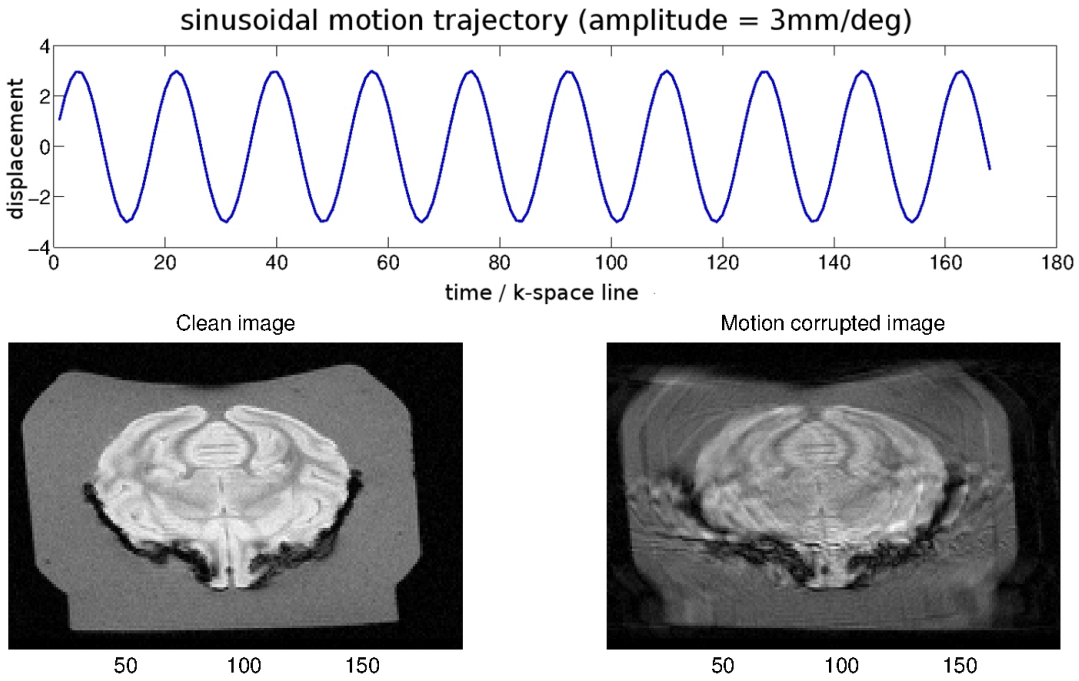

Figure 3.26: The effects of the translational motion of a sinusoidal shape on the image.

lines, or orthogonal to them. Irrespective of the direction of motion, the ghosts are always forming orthogonally to the acquisition lines. This is because each k-space line has coherent phases within it, and the incoherence builds up only between the different lines. Interestingly, the same holds for rotational motion – ghosting artifacts propagate orthogonally to acquisition lines. This is due to the fact that the phases in each rotated line are coherent in this case too.

It is the rotational motion that is most problematic for retrospective motion correction. As was mentioned above, it causes an irregular k-space sampling leading to the potential loss of the spectral data. In Figure 3.28, the effects of rotations on the image are shown. Rotations in the range of 36 degrees can render the image useless for medical analysis.

3.6.5 Rotation inversion with exact motion parameters

Shown in Figure 3.29 is a mechanism of the Fourier data loss due to rotational motion. Rotations cause some frequency information about the object to never be acquired, resulting in "holes" in k-space. A strong rotational motion in a range of

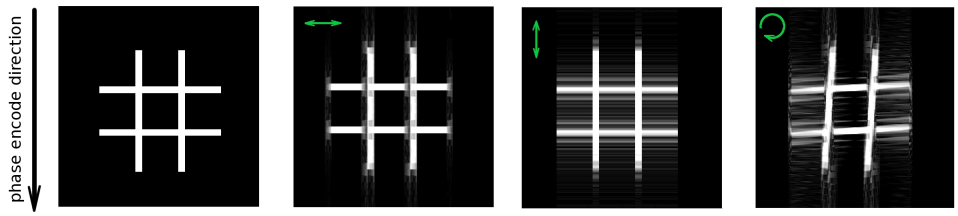

Figure 3.27: Effects of motion in different directions: translation in frequency encode direction causes blurring; translation in phase encode direction leads to ghosting; rotations cause both ghosting and blurring.

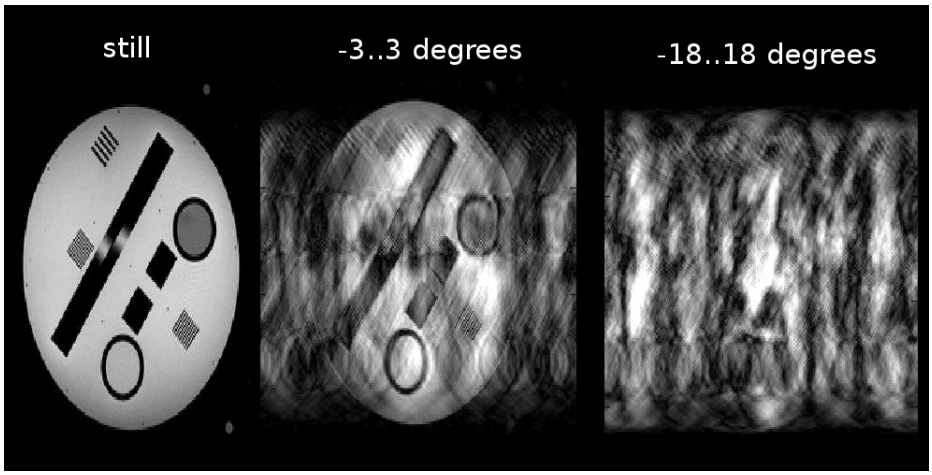

Figure 3.28: The effects of rotational motion of different strengths. Motion in the range of 36 degrees can be devastating.

several degrees can lead to a substantial loss of the spectral data, meaning that, in order to recover the artifact-free image, a highly underdetermined inverse problem needs to be solved.

A rotational motion can be described by a matrix, whose elements are the interpolation weights. In order to induce the rotational motion, this matrix is multiplied with a vectorized Fourier representation of the ground-truth image. A straightforward way to correct for motion artifacts due to rotational motion is to find the inverse of this matrix, and then multiply it with a motion degraded observation. Shown in Figure 3.30 are the results of motion correction by a matrix inversion, which were computed in the case of progressively increasing strength of rotational motion. It can be seen from the illustration that the rotational motion with a range of just 4 degrees (sine shape trajectory) leads to visible artifacts in the image obtained after matrix inversion-based correction. In contrast, the translational motion

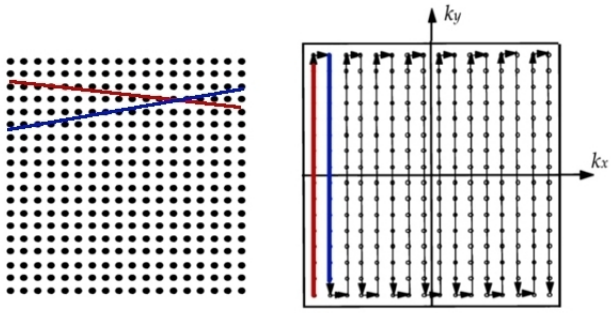

Figure 3.29: Strong rotational motion leads to missing k-space data due to a sampling inadequacy of mid-frequency components.

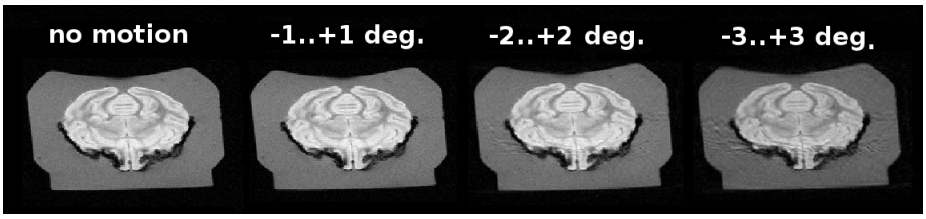

Figure 3.30: Example of correction of rotational motion using ground-truth motion parameters. Simple inversion of the motion operator results in visible artifacts in corrected results given a rotational motion of just 4 degrees.

is always perfectly invertible if the imaged object remains within a coil sensitivity region.

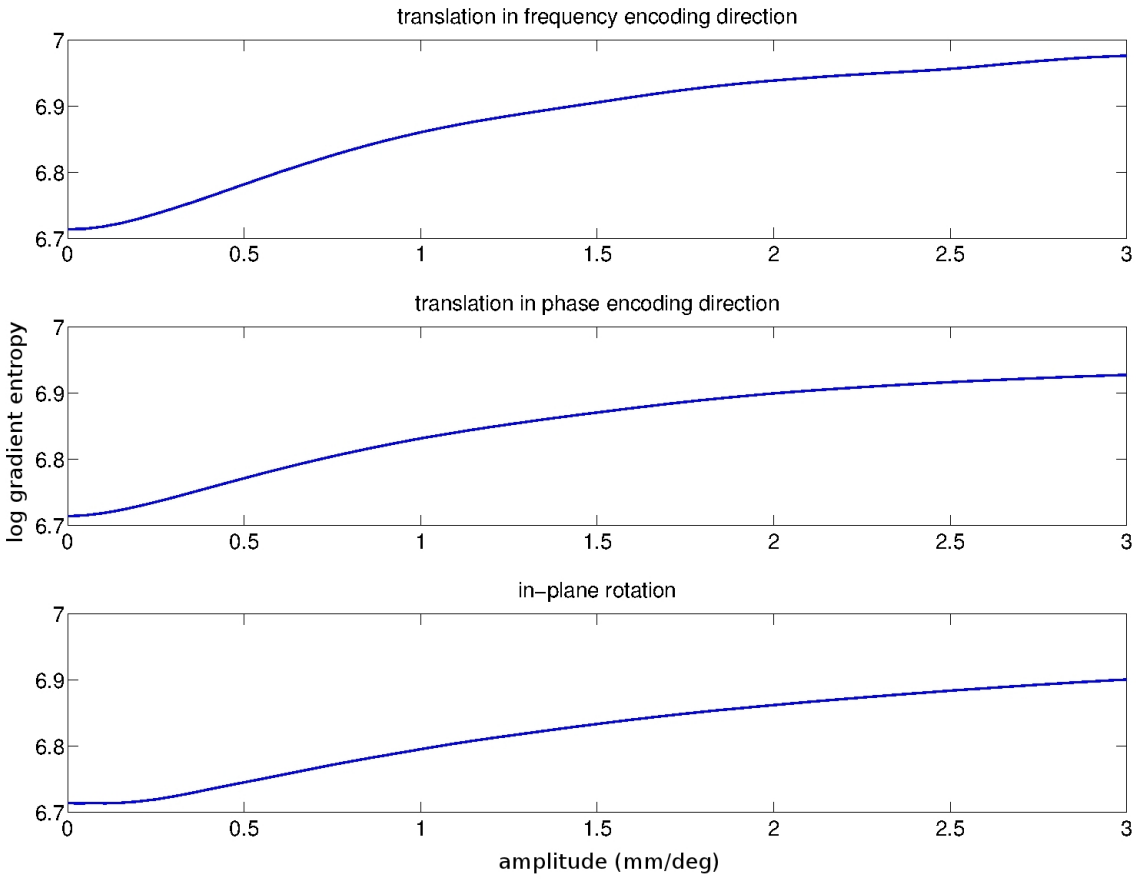

Figure 3.31: Entropy of the gradients image metric function is monotonically increasing when computed on the image corrupted by the motion of a sinusoidal shape with growing amplitude.

3.6.6 On image metric values

Autofocusing approaches rely on computing the image metric in order to do the motion correction. It is a desirable property of a metric function that its values are monotonically increasing given the progressive increase of motion intensity. Such metric function should also give a smallest value for the images not corrupted by motion. Shown in Figure 3.31 are three plots (2 for translations in phase/read directions, and one for in-plane rotation) of the values of the entropy of the gradients image metric calculated for the images corrupted by motion of progressively increasing strength (the amplitude of the sine motion trajectory was modulated). As the plots show, the gradient entropy is a monotonically increasing function of the motion amplitude.

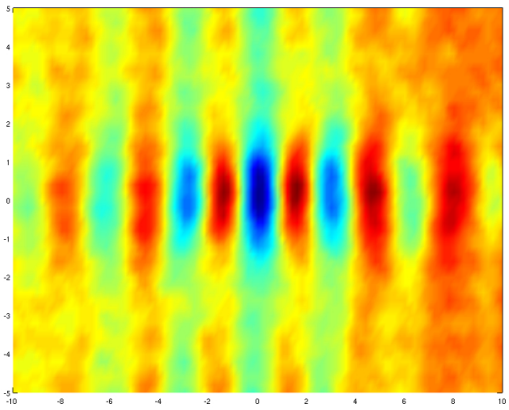

Figure 3.32: Reducing the number of degrees of freedom in the motion trajectory to just two (step motion in a fixed repetition with a variable translational offset in phase/read directions) allows an image metric landscape to be drawn. A complex non-linear structure with multiple local minima is obvious.

We can also go a step further and examine how the energy landscape of the metric function looks - see Figure 3.32. This is essentially the same plot of the image metric values, but now it is computed as a function of two parameters. Instead of the sine shape motion trajectory, we now use the trajectory with a single step, fixing the position of the step in time (repetition number) and varying the extent of spatial translations in both phase and read encode directions. The range of displacements is [-10 10] pixels. As shown in the figure, the resulting metric function landscape has a complicated structure with multiple local minima and maxima. Still, there is a global minimum, which corresponds to zero motion. Thus, in the case of a step-based motion trajectory, there is no longer a simple monotonic relationship between the displacement and the energy; instead the objective landscape has a complicated structure which it is important not to ignore when doing autofocusing-based correction.

4 Blind retrospective motion correction

In recent years, most successful motion correction techniques have been based on prospective motion correction (see [29, 62], and [25] for a comprehensive review). However, as discussed in the previous chapter, prospective methods have certain limitations, which make their use impossible in certain scenarios. In this chapter, I introduce a new retrospective motion correction algorithm that does not require tracking equipment and can be applied to already acquired data through post-processing. Quite importantly, in contrast with prospective techniques, the method preserves the original motion-corrupted image, which allows the motion correction results to be compared with the original image.

As was mentioned in the problem statement section of this dissertation, current autofocusing-based retrospective methods suffer from the problem of being too slow to be used in clinical practice. This is because the space of motion parameters is explored in a brute force, inefficient way, leading to highly suboptimal solutions or long computation times in high-resolution 3D scenarios due to the curse of dimensionality. Together with my colleagues, I propose an autofocusing method that is based on minimization of a cost function that characterizes the quality of the optimized image. The idea is to find the point in the space of possible motions such that the image corrected for motion yields the minimum value of a cost function. The challenge for our approach is the fact that for high-resolution 3D volumes, the optimization space is vast since there are 6 free parameters per phase encode step. We address this problem by using an analytic model for both motion degradation and its inverse, which allows the partial derivatives (gradients) of our cost function to be found, and an efficient optimization of the objective to be carried out. For this reason we call our algorithm GradMC (Gradient-based Motion Correction). GradMC starts with an initial estimate of the motion trajectory (which can be no-motion indicating zeros or some small random values), and iteratively corrects for

translation and rotation until no more progress in terms of image quality can be made. The image quality metric that we use in our cost function is the entropy of spatial image gradients (first finite derivatives).

The major difference between GradMC and other autofocusing approaches is the way in which the motion space is explored. Greedy forward selection (as in e.g. [18]) locally optimizes a quality metric by altering only a few motion parameters at a time. As mentioned above, we use the analytical gradient to drive the optimization process, and consistently explore the parameter space using quasi-Newton BFGS (Broyden-Fletcher-Goldfarb-Shanno). GradMC attains a provable convergence to a local optimum, which is a desirable property of an optimization algorithm. Interestingly, gradients can be computed quite cheaply by one single non-uniform fast Fourier transformation per degree of freedom (DOF). We conjecture that derivative-based optimization makes it possible to achieve higher image quality due to efficient exploration of the high-dimensional objective landscape.

We model the motion trajectory as a piecewise constant function and make the assumption that the motion time-scale is longer than the repetition time TR. This makes modeling of the motion degradation process much simpler, while still being a valid assumption due to the fact that strong displacement events are likely to be sparse over the course of image acquisition. Thus, only a few views are likely to be corrupted, and the source of the major image artifacts is the inconsistency between views. Similarly to other retrospective approaches, we also neglect second order effects such as the influence of motion on the magnetic field. We further assume that the scanned object is a rigid body (i.e. brain), but this assumption will be relaxed in the method extension presented in the next chapter. Given that these assumptions are satisfied, GradMC can correct for arbitrary motion trajectories in 6 degrees of freedom for both 2D and 3D acquisitions. For the convenience of the reader, and to promote the wide spread of autofocusing based methods, we make a Matlab implementation of GradMC along with 4 examples available online at: `http://mloss.org/software/view/430/`.

In this chapter I first introduce the theory required to understand our approach, and then explain the proposed method. The second part of the chapter presents the experimental results based on both simulated and in vivo data, and summarizes and discusses certain aspects of the proposed method.

4.1 Rigid motion transformation in Fourier domain

MR scanners acquire Fourier coefficients along smooth trajectories as dictated by the gradient shapes of the MRI sequence. We restrict ourselves to the Cartesian case where k-space is densely sampled line by line. In the following discussion, we describe a forward model of the measurement process that includes a moving subject.

Let $\mathbf{F} \in \mathbb{C}^{N\times N}$ be the orthonormal Fourier matrix, $\mathbf{u} \in \mathbb{C}^N$ be the unknown sharp image of size $N = n_1 \cdot n_2 \cdot n_3$ pixels, and $\mathbf{M} = \mathrm{dg}(\mathbf{m}) \in [0,1]^{N\times N}$, where $\mathbf{m} \in [0,1]^N$ is a diagonal masking matrix. Furthermore, let $\tilde{\mathbf{A}}_{\boldsymbol{\theta}}$ denote a general rigid motion transformation matrix parameterized by three translation shifts and three rotation angles contained in $\boldsymbol{\theta} \in \boldsymbol{\Theta} = \mathbb{R}^3 \times [0, 2\pi)^3$. Eventually, we want to think about rotations around a point which is different to the point of origin.

Assuming additive Gaussian noise $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$, the acquisition in k-space can be written as a noisy linear process

$$\mathbf{y} = \underbrace{\int_0^T \mathbf{M}_t\mathbf{F}\tilde{\mathbf{A}}_{\boldsymbol{\theta}_t}\mathrm{d}t}_{\mathbf{X}}\ \mathbf{u} + \boldsymbol{\varepsilon} \in \mathbb{C}^N,\ \mathbf{M}_t = \mathrm{dg}(\mathbf{m}_t),\ \mathbf{1} = \int_0^T \mathbf{m}_t\mathrm{d}t, \tag{4.1}$$

where $\tilde{\mathbf{A}}_{\boldsymbol{\theta}_t}$ denotes the translation or rotation of the imaged object at time t. Equivalently, one can think of all the matrices as general linear operators acting on continuous images instead of discretized ones. The formalism is the same and since we are ultimately interested in a pixel image, we base our explanations on the discretized case. The above model ignores second order effects such as the influence of motion on the magnetic field.

4.1.1 Pure translations case

If the subject makes purely translational movements, the matrix $\tilde{\mathbf{A}}_{\boldsymbol{\theta}_t}$ *is* a convolution $\tilde{\mathbf{A}}_{\boldsymbol{\theta}_t}\mathbf{u} = \tilde{\mathbf{a}}_{\boldsymbol{\theta}_t} \star \mathbf{u} = \mathbf{F}^{\mathsf{H}}\mathrm{dg}(\mathbf{F}\tilde{\mathbf{a}}_{\boldsymbol{\theta}_t})\mathbf{F}\mathbf{u}$ and hence the measurement matrix

$$\mathbf{X}\mathbf{u} = \int_0^T \mathbf{M}_t\mathbf{F}\tilde{\mathbf{A}}_{\boldsymbol{\theta}_t}\mathrm{d}t\ \mathbf{u} = \mathrm{dg}\left(\mathbf{F}\,\underbrace{\mathbf{F}^{\mathsf{H}}\int_0^T \mathbf{m}_t \odot (\mathbf{F}\tilde{\mathbf{a}}_{\boldsymbol{\theta}_t})\mathrm{d}t}_{\mathbf{k}}\right)\mathbf{F}\mathbf{u} = \mathbf{k} \star \mathbf{u} \tag{4.2}$$

turns out to be very simple. Consequently, the whole measurement process can be described by a *single* convolution with a filter $\mathbf{k}$.

4.1.2 Fourier domain representation

At every time point t, the subject is rigidly moving, i.e. (Equation 4.1) has $\tilde{\mathbf{A}}_{\boldsymbol{\theta}_t}\mathbf{u}$ instead of $\mathbf{u}$. As measurements are made in k-space, it is natural to represent $\tilde{\mathbf{A}}_{\boldsymbol{\theta}_t}$ in k-space. Mathematically speaking, since $\tilde{\mathbf{A}}_{\boldsymbol{\theta}_t}$ and $\mathbf{F}$ do not commute in general, we seek an operator $\mathbf{A}_{\boldsymbol{\theta}_t}$ such that $\mathbf{F}\tilde{\mathbf{A}}_{\boldsymbol{\theta}_t} = \mathbf{A}_{\boldsymbol{\theta}_t}\mathbf{F}$. Interestingly, both rotation and translation have an equivalent in k-space [63]. Rotation in image space corresponds to rotation in k-space and a translation corresponds to a pointwise rescaling. This simplifies our measurement equation (Equation 4.1) considerably because $\mathbf{F}$ can be moved outside the integral

$$\mathbf{y} = \int_0^T \mathrm{dg}(\mathbf{m}_t)\mathbf{A}_{\boldsymbol{\theta}_t}\mathrm{d}t\,\mathbf{F}\mathbf{u} + \boldsymbol{\varepsilon} \in \mathbb{C}^N,\ \mathbf{1} = \int_0^T \mathbf{m}_t \mathrm{d}t. \tag{4.3}$$

4.1.3 Discretizing time

In an MRI experiment, there are two characteristic times: the *echo time* TE and the *repetition time* TR. Our line-by-line setting is best described as a sequence of repetitions each of length TR. The acquisition of the signal inside each of the repetitions takes $2 \cdot TE$, hence $2 \cdot TE < TR$. In the following analysis, we assume that the movement is piecewise constant such that the rigid transform matrix $\mathbf{A}_{\boldsymbol{\theta}_t}$ does not change inside one repetition.

Thus, the measurement integral (Equation 4.3) becomes a sum and the masking becomes binary instead of continuous $\mathbf{M} = \mathrm{dg}(\mathbf{m}) \in \{0,1\}^{N\times N}$ where $\mathbf{m} \in \{0,1\}^N$

$$\mathbf{y} = \sum_{t=1}^{T} \mathrm{dg}(\mathbf{m}_t)\mathbf{A}_{\boldsymbol{\theta}_t}\mathbf{F}\mathbf{u} + \boldsymbol{\varepsilon} \in \mathbb{C}^N,\ \mathbf{1} = \sum_{t=1}^{T} \mathbf{m}_t. \tag{4.4}$$

Assume for now that the k-space lines of an image of size n_1 by n_2 pixels are measured from top to bottom, hence $T = n_1$ steps are needed. Then, in every repetition, a noisy version of $\mathrm{dg}(\mathbf{m}_t)\mathbf{A}_{\boldsymbol{\theta}_t}\mathbf{F}\mathbf{u} \in \mathbb{C}^N$ is measured. But since $\mathbf{m}_t$ is binary this is equivalent to saying that we acquire a noisy version of n_2 components instead of $n = n_1 n_2$, which we denote by $[\mathbf{A}_{\boldsymbol{\theta}_t}]_{\mathbf{m}_t}\mathbf{F}\mathbf{u} \in \mathbb{C}^{n_2}$ with the shortcut $[\mathbf{A}_{\boldsymbol{\theta}_t}]_{\mathbf{m}_t} \in \mathbb{C}^{n_2\times n}$

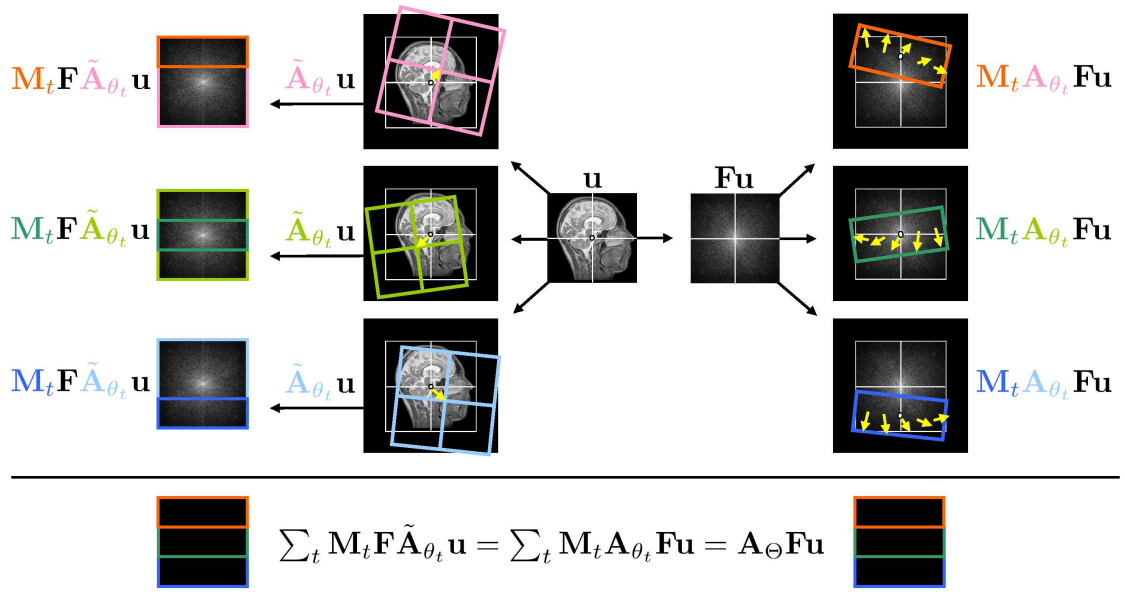

Figure 4.1: Illustration of the forward model. An unknown underlying image $\mathbf{u}$ experiences rotation and translation during every phase encode step represented by a masking matrix $\mathbf{M}_t$. Motion can be represented in the pixel domain (left) and in the Fourier domain (right). The latter allows for efficient computations.

$$\mathbf{y} = \mathbf{A}_{\boldsymbol{\theta}}\mathbf{Fu} + \boldsymbol{\varepsilon} \in \mathbb{C}^N, \ \mathbf{A}_{\boldsymbol{\theta}} := \sum_{t=1}^{T} \mathbf{M}_t\mathbf{A}_{\boldsymbol{\theta}_t} = \begin{bmatrix} [\mathbf{A}_{\boldsymbol{\theta}_1}]_{\mathbf{m}_1} \\ [\mathbf{A}_{\boldsymbol{\theta}_2}]_{\mathbf{m}_2} \\ \vdots \\ [\mathbf{A}_{\boldsymbol{\theta}_T}]_{\mathbf{m}_T} \end{bmatrix} \in \mathbb{C}^{N\times N}. \tag{4.5}$$

For an illustration of the Fourier forward model of motion, see Figure 4.1.

The bottom line is that a matrix vector multiplication (MVM) with $\mathbf{A}_{\boldsymbol{\theta}}$ can essentially be done in $\mathcal{O}(N)$ time since $\mathbf{A}_{\boldsymbol{\theta}}$ can be decomposed into blocks whose MVMs can be performed efficiently.

4.1.4 Field of view

In many 3D scans, the object to be acquired does not respect the circular boundary conditions. That is, a part of the object leaves the FOV due to motion. Formally, we do not see $\mathbf{u}$ but $\mathbf{v}_t \odot \mathbf{u}$, where $\mathbf{v} \in \{0,1\}^N$ is the field of view mask. Hence,

using the convolution theorem $\mathbf{F}(\mathbf{v} \odot \mathbf{u}) = (\mathbf{F}\mathbf{v}) \star (\mathbf{F}\mathbf{u}) = \mathbf{V}\mathbf{F}\mathbf{u}$, we have

$$\mathbf{y} = \sum_{t=1}^{T} \mathrm{dg}(\mathbf{m}_t)\mathbf{F}(\mathbf{v} \odot (\tilde{\mathbf{A}}_{\boldsymbol{\theta}_t}\mathbf{u})) = \sum_{t=1}^{T} \mathrm{dg}(\mathbf{m}_t)\mathbf{V}\mathbf{A}_{\boldsymbol{\theta}_t}\mathbf{F}\mathbf{u} + \boldsymbol{\varepsilon} \in \mathbb{C}^N,\ \mathbf{1} = \sum_{t=1}^{T} \mathbf{m}_t. \quad (4.6)$$

Here, $\mathbf{V}$ is a convolution matrix with a filter being the Fourier transform of the FOV mask $\mathbf{v}$. In our case, $\mathbf{v}$ is a (separable) box-shaped FOV which means $\mathbf{V}$ corresponds to a convolution with a (separable) sinc kernel.

The unfortunate bit in the above formula is that $\mathbf{y}$ does not simply consist of (disjoint) rotated and phase multiplied parts of $\mathbf{F}\mathbf{u}$ but that there is a subsequent (although local) convolution with $\mathbf{F}\mathbf{v}$ taking place, before $\mathbf{m}_t$ can clip out the corresponding Fourier coefficients. The computational load thus increases by a constant factor.

4.1.5 Vector space parameterizations of rigid body motion

A rigid body motion of $\mathbb{R}^d$ is a transformation $\phi : \mathbb{R}^d \to \mathbb{R}^d$ that leaves distances and orientations unchanged. Mathematically, all these transformations form a group $\mathrm{SE}(d)$, the special Euclidean group. As a subgroup of the affine group, every element $\phi \in \mathrm{SE}(d)$ can be represented by $\phi : \mathbf{z} \mapsto \mathbf{R}\mathbf{z} + \mathbf{b}$ with $d(d+1)/2$ degrees of freedom in $\mathbf{R}$ and $\mathbf{b}$. Furthermore, any rigid motion can be decomposed into a rotation with $d(d-1)/2$ degrees of freedom and a translation with d degrees of freedom by Mozzi-Chasles' theorem. Therefore, it suffices to consider those two transformations only. While a translation is easily represented by a vector $\mathbf{b} \in \mathbb{R}^d$, rotations $\mathbf{R} \in \mathrm{SO}(d)$ are harder to represent. Since $\mathrm{SO}(d)$ is a Lie group, one can use the associated Lie algebra consisting of skew symmetric matrices $\mathbf{X} = -\mathbf{X}^\top$ and the matrix exponential $e^{\mathbf{X}} := \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{X}^k$ to represent rotations in a vectorial way by $\mathbf{R} = e^{\mathbf{X}}$.
We restrict ourselves to $d = 3$ and obtain 6 degrees of freedom in total. The matrix exponential for $d = 3$ is simple:

$$e^{\mathbf{X}} = \mathbf{n}\mathbf{n}^\top + \cos(\alpha)\left(\mathbf{I} - \mathbf{n}\mathbf{n}^\top\right) + \sin(\alpha)\begin{bmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{bmatrix}, \quad (4.7)$$

$$\mathbf{X} = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \tag{4.8}$$

where $\mathbf{n} = \mathbf{x}/\left\|\mathbf{x}\right\|$ and $\alpha = \left\|\mathbf{x}\right\|$. Hence, $\mathbf{R} = e^{\mathbf{X}}$ can be interpreted as a rotation around the axis $\mathbf{n}$ with angle α. If we restrict ourselves to $\left\|\mathbf{x}\right\| < 2\pi$, the map $e : \mathbf{X} \mapsto \mathbf{R}$ is even invertible. Thus, we have a parameterization $p : \boldsymbol{\theta} \in \mathbb{R}^6 \mapsto (\mathbf{R}, \mathbf{b}) \in \mathrm{SE}(3)$ that we will use to compute $\mathbf{A}_{\boldsymbol{\theta}_t} \in \mathbb{R}^{n \times n}$, the linear map in pixel space.

Rotations around a point $\mathbf{c}$, i.e. $\phi_{\mathbf{c}} : \mathbf{x} \mapsto \mathbf{c} + \mathbf{R}(\mathbf{x} - \mathbf{c})$ such that $\mathbf{c}$ stays unchanged can be achieved by setting $\mathbf{b} = \mathbf{c} - \mathbf{R}\mathbf{c}$.

4.2 Matrix inverse, conjugate and motion inversion

In order to be able to invert the forward model, we study whether a rigid body transformation matrix $\mathbf{A}_{\boldsymbol{\theta}}$ can be inverted. For the case of pure translation, the matrix is diagonal in Fourier space, which allows for exact inversion. Since the diagonal elements have an absolute value of 1, we have $\mathbf{A}_{\boldsymbol{\theta}}^{-1} = \mathbf{A}_{\boldsymbol{\theta}}^{\mathsf{H}} = \mathbf{A}_{-\boldsymbol{\theta}}$.

In the case of rotation, the matrix cannot be inverted because some pixels can be mapped to a place outside the image. One can use the notion of invertability in a subspace (the so called Bott-Duffin-Inverse). Pixels inside the circular region in the image center form that subspace. For rotation matrices $\mathbf{A}_{\boldsymbol{\theta}}$, the objects $\mathbf{A}_{\boldsymbol{\theta}}^{\mathsf{H}}$ and $\mathbf{A}_{-\boldsymbol{\theta}}$ are distinctively different, but as we observed in our experiments $\mathbf{A}_{-\boldsymbol{\theta}}$ performs better.

4.3 Measures of image quality

The key ingredient of any AF method is an image quality metric. Ideally, the value of metric function is large for motion-corrupted images and small for motion-free ones. The second desirable property is that minima of such a metric function correspond to diagnostically useful images according to medical experts. We are not aware of a theoretically-justified optimal metric fulfilling these requirements. In an

autofocusing paper by David Atkinson [19], the entropy is used as objective

$$\mathcal{H}(\mathbf{u}) = \mathbf{v}^\top \ln \mathbf{v}, \mathbf{v} = \frac{|\mathbf{u}|}{\|\mathbf{u}\|} \in \mathbb{R}_+^n, |\mathbf{u}| = \sqrt{\Im^2(\mathbf{u}) + \Re^2(\mathbf{u})} \in \mathbb{R}_+^n, \|\mathbf{u}\| = \sqrt{\mathbf{u}^\mathsf{H}\mathbf{u}} \in \mathbb{R}_+. \tag{4.9}$$

In our experiments we use the gradient entropy

$$\phi(\mathbf{u}) = \mathcal{H}(\mathbf{D}_x\mathbf{u}) + \mathcal{H}(\mathbf{D}_y\mathbf{u}), \tag{4.10}$$

that was found to be superior to 24 estimators in an empirical study based on the expertise of radiologists [33].Here $\mathbf{D}_x$, $\mathbf{D}_y \in \{0, \pm 1\}^{N\times N}$ are horizontal and vertical finite difference matrices, and $\mathcal{H}(\cdot)$ is a pixel entropy defined as:

$$\mathcal{H}(\mathbf{u}) = -\mathbf{v}^\top \ln \mathbf{v},\ \mathbf{v} = \sqrt{\frac{\mathbf{u} \odot \bar{\mathbf{u}}}{\mathbf{u}^\mathsf{H}\mathbf{u}}} \in \mathbb{R}_+^N,\ \mathbf{u} \in \mathbb{C}^N. \tag{4.11}$$

Furthermore, $\bar{\mathbf{u}}$ denotes the complex conjugate and $\mathbf{a} \odot \mathbf{b}$ denotes the pointwise (Hadamard) product of two vectors $\mathbf{a}$ and $\mathbf{b}$. Note that the entropy is scale invariant $\mathcal{H}(\mathbf{u}) = \mathcal{H}(\alpha\mathbf{u})$, phase invariant $\mathcal{H}(\mathbf{u}) = \mathcal{H}(\exp(i\boldsymbol{\omega}) \odot \mathbf{u})$ and invariant w.r.t. a constant offset $\mathcal{H}(\mathbf{u}) = \mathcal{H}(\mathbf{u} + c\mathbf{1})$. The effect of such a focus criterion is to prefer images having uniform intensity regions separated by sharp edges. Indeed, motion corrupted images are known to be affected by blurring and ghosts, both leading to smooth intensity variations and thus high gradient entropy $\mathcal{H}$.

4.4 Blind rigid motion correction optimization problem

Our goal is to estimate the motion parameters $\hat{\boldsymbol{\theta}}$ which best describe the motion in the corrupted volume, and then to correct for motion. We consider a class of approximately invertible motions composed of translations and small angle rotations ($< 3^\circ$). For strong rotations, the information in high-frequency k-space regions is lost, and no artifact-free empirical inversion is possible. By empirical inversion we mean the application of $\mathbf{A}_{\boldsymbol{\theta}}$ directly to a motion corrupted observation $\mathbf{y}$

$$\mathbf{u}_{\boldsymbol{\theta}} = \mathbf{F}^\mathsf{H}\mathbf{A}_{\boldsymbol{\theta}}\mathbf{y}, \tag{4.12}$$

where $\mathbf{u}_{\boldsymbol{\theta}}$ is the resulting image in the spatial domain. We are interested in motion parameters $\hat{\boldsymbol{\theta}}$ that correspond to the sharpest image according to the minimum of our cost function

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\Theta}{\arg\min}\, \phi\left(\mathbf{u}_{\boldsymbol{\theta}}\right). \tag{4.13}$$

It is clear that the estimation of $6 \cdot T$ free motion parameters from a single image constitutes a highly underdetermined problem. Thus far, we have treated the motion parameters for all views independently. In practice, doing the motion correction with our objective, the recovered trajectories were found to contain regularly positioned strong spikes. The presence of these spikes was not found to cause a deterioration in image quality. However, such artifacts can be a nuisance, when the goal is to recover a realistic motion trajectory, thus we regularize the trajectory of the recovered motion parameters by putting a quadratic penalty on the differences between consecutive motion parameters

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\Theta}{\arg\min}\, \phi(\mathbf{F}^{\mathsf{H}}\mathbf{A}_{\boldsymbol{\theta}}\mathbf{y}) + \lambda \left\|\mathbf{D}\boldsymbol{\theta}\right\|^2, \tag{4.14}$$

where $\mathbf{D} \in \{0, \pm 1\}^{6\cdot T \times 6\cdot T}$ is a finite difference matrix (respecting the temporal order of the phase encodes), and $\lambda \in \mathbb{R}_+$ is a regularization parameter controlling the smoothness of the trajectory. This heavily penalizes the occurrence of spikes leading to smoother trajectories. In practice, the value $\lambda = 0.1$ was found to give best results. In order for a regularization to make sense, we assume that the order in which k-space lines are recorded by the scanner is known (in 3D there is a phase/slice direction ambiguity in this respect). Otherwise, the temporal structure of the motion trajectory recovered by the algorithm might not match the actual temporal sequence of acquisition in k-space. As an alternative to the difference penalty term, a blocking approach can be used, where the motion parameters in consecutive views are hard-constrained to be optimized together. This makes it possible to obtain reasonable and probable motion estimates for those regions in k-space where the accuracy of the algorithm is limited. While it may be worthwhile to take advantage of this complexity-reduction approach to further improve the motion correction results, the spirit of our current research is slightly different. We reduce complexity by means of an adjustable regularizer instead, while keeping the number of parameters fixed. This leaves the optimizer with more freedom to find the best setting among the

softly coupled parameters.

4.5 Implementing the Fourier motion operator

The practical implementation of the $\mathbf{A}_{\boldsymbol{\theta}}$ operator involves the treatment of translational and rotational motion represented by $6 \cdot T$ components in $\boldsymbol{\theta}$. The translation correction amounts to a multiplication of each k-space line with a linear phase ramp $\exp(-2\pi i k_x \theta_t)$, a function of translation θ_t in the spatial domain at time t, and the Fourier coefficient k_x of the affected view.

Rotation is more involved since it causes a rotation of one section of k-space relative to others. To compute its effect on the image, we first construct a deformed grid by rotating the points of each k-space line by their time-respective angles. In order to make gridding interpolation efficient, there are two strategies: One can do a local interpolation from many neighbors or one can resample the Fourier signal on a finer grid by zero padding in the pixel domain. Such resampling constitutes a fast and efficient (only expensive in terms of 2^d more memory). Then we carry out the interpolation in an oversampled (2X) k-space to estimate the values on the points of the rotated grid.

4.5.1 Derivatives

4.5.1.1 General representation

We have a parameterized matrix $\mathbf{A}_{\boldsymbol{\theta}}$ and a functional $\phi(\mathbf{u}, \boldsymbol{\theta})$ of the form

$$\phi(\mathbf{u}, \boldsymbol{\theta}) := \|\mathbf{A}_{\boldsymbol{\theta}}\mathbf{F}\mathbf{u} - \mathbf{y}\|_2^2 + \lambda\rho(\mathbf{B}\mathbf{u}) \tag{4.15}$$

and wish to compute $\frac{\partial\phi}{\partial\mathbf{u}}$ and $\frac{\partial\phi}{\partial\boldsymbol{\theta}}$. Derivatives w.r.t. the first variable $\mathbf{u}$ are simple

$$\frac{\partial}{\partial\mathbf{u}}\phi(\mathbf{u}, \boldsymbol{\theta}) = 2\mathbf{F}^{\mathsf{H}}\mathbf{A}_{\boldsymbol{\theta}}^{\mathsf{H}}(\mathbf{A}_{\boldsymbol{\theta}}\mathbf{F}\mathbf{u} - \mathbf{y}) + \lambda\mathbf{B}^{\top}\dot{\boldsymbol{\rho}}(\mathbf{B}\mathbf{u}) \tag{4.16}$$

while the second variable $\boldsymbol{\theta}$ is more challenging. Let's compute the derivative of a real-valued function $\xi : \mathbb{R}^{6\times T} \mapsto \mathbb{R}$ having the form

$$\xi(\boldsymbol{\theta}) = \mathbf{w}^{\mathsf{H}}\boldsymbol{\psi}\left(\mathbf{C}\mathbf{X}_{\boldsymbol{\theta}}\mathbf{v} - \mathbf{b}\right), \tag{4.17}$$

where $\mathbf{X} : \mathbb{R}^{6\times T} \mapsto \mathbb{C}^{N\times N}$, $\mathbf{X}_{\boldsymbol{\theta}} = \mathbf{A}_{\boldsymbol{\theta}}\mathbf{F}$, $\mathbf{C} \in \mathbb{R}^{\nu\times N}$, $\mathbf{v} \in \mathbb{C}^N$, $\mathbf{b}, \mathbf{w} \in \mathbb{C}^\nu$ and $\boldsymbol{\psi} : \mathbb{C}^\nu \mapsto \mathbb{C}^\nu$ a pointwise map. We further assume that every row of $\mathbf{X}_{\boldsymbol{\theta}}$ depends on the six components of one particular $\boldsymbol{\theta}_{1:6,t}$ only, which allows to decompose $\mathbf{X}_{\boldsymbol{\theta}}$ into $[\mathbf{X}_{\boldsymbol{\theta}_1}; \mathbf{X}_{\boldsymbol{\theta}_1}; ..; \mathbf{X}_{\boldsymbol{\theta}_T}]$. The derivative is given by

$$\mathrm{d}\xi = \mathbf{r}^{\mathsf{H}}(\mathrm{d}\mathbf{X}_{\boldsymbol{\theta}})\mathbf{v},\ \mathbf{r} := \mathbf{C}^\top\left(\mathbf{w}\odot\dot{\boldsymbol{\psi}}(\mathbf{C}\mathbf{X}_{\boldsymbol{\theta}}\mathbf{v}-\mathbf{b})\right),\quad \frac{\partial\xi}{\partial\theta_{i,t}} = \mathbf{r}^{\mathsf{H}}\frac{\partial\mathbf{X}_{\boldsymbol{\theta}}}{\partial\theta_{i,t}}\mathbf{v}. \tag{4.18}$$

Observe that only the t-th block in $\frac{\partial\mathbf{X}_{\boldsymbol{\theta}}}{\partial\theta_{i,t}}$ is nonzero, hence we can stack *all* derivatives $t = 1..T$ for a particular $i = 1..6$ into a single matrix

$$\frac{\partial\mathbf{X}_{\boldsymbol{\theta}}}{\partial\boldsymbol{\theta}_i} := \sum_{t=1}^{T}\frac{\partial\mathbf{X}_{\boldsymbol{\theta}}}{\partial\theta_{i,t}} = \begin{bmatrix} [\frac{\partial\mathbf{X}_{\boldsymbol{\theta}}}{\partial\theta_{i,1}}]_{\mathbf{m}_1} \\ [\frac{\partial\mathbf{X}_{\boldsymbol{\theta}}}{\partial\theta_{i,2}}]_{\mathbf{m}_2} \\ \vdots \\ [\frac{\partial\mathbf{X}_{\boldsymbol{\theta}}}{\partial\theta_{i,T}}]_{\mathbf{m}_T} \end{bmatrix} \in \mathbb{C}^{N\times N}. \tag{4.19}$$

One can even generalize further and establish the derivatives of the form

$$\frac{\partial\psi(\mathbf{A}_{\boldsymbol{\alpha}}\mathbf{u})}{\partial\boldsymbol{\alpha}} \in \mathbb{R}^N \tag{4.20}$$

where $\mathbf{A}_{\boldsymbol{\alpha}} \in \mathbb{C}^{N\times N}$ is a matrix with parameter $\boldsymbol{\alpha} \in \mathbb{R}^N$ such that every row i depends on α_i only, $\mathbf{u} \in \mathbb{C}^N$ a vector and $\psi : \mathbb{C}^N \to \mathbb{R}$ is a continuously differentiable function. We denote by $\mathbf{A}'_{\boldsymbol{\alpha}}$ the matrix such that $[\frac{\partial}{\partial\alpha_i}\mathbf{A}_{\boldsymbol{\alpha}}\mathbf{u}]_i = [\mathbf{A}'_{\boldsymbol{\alpha}}\mathbf{u}]_i$ i.e. $\mathbf{A}'_{\boldsymbol{\alpha}}$ contains all the information needed for the Jacobian

$$\frac{\partial\mathbf{A}_{\boldsymbol{\alpha}}\mathbf{u}}{\partial\boldsymbol{\alpha}^\top} = \mathrm{dg}(\mathbf{A}'_{\boldsymbol{\alpha}}\mathbf{u}), \tag{4.21}$$

which is diagonal because every row i depends on α_i only. Now, let $\frac{\partial\psi(\mathbf{v})}{\partial\mathbf{v}}$ denote the gradient of the objective function $\psi : \mathbb{C}^N \to \mathbb{R}$. Then, the desired derivative $\partial\psi(\mathbf{A}_{\boldsymbol{\alpha}}\mathbf{u})/\partial\boldsymbol{\alpha}$ can be obtained via the chain rule

$$\frac{\partial\psi(\mathbf{A}_{\boldsymbol{\alpha}}\mathbf{u})}{\partial\boldsymbol{\alpha}} = \frac{\partial(\mathbf{A}_{\boldsymbol{\alpha}}\mathbf{u})^\top}{\partial\boldsymbol{\alpha}}\frac{\partial\psi(\mathbf{v})}{\partial\mathbf{v}} = (\mathbf{A}'_{\boldsymbol{\alpha}}\mathbf{u})\odot\left(\frac{\partial\psi(\mathbf{v})}{\partial\mathbf{v}}\right). \tag{4.22}$$

To compute $\mathbf{A}'$ in general, we need to take a closer look at how the MVM $\mathbf{A}\mathbf{y}$ (we write $\mathbf{A}$ and $\mathbf{A}'_{\#}$ where $\# \in \{x, y, z, a, b, c\}$ in the following for short) is computed and how it depends on the motion parameters. First of all, $\mathbf{A}\mathbf{y}$ can be decomposed

into a translational part $\mathbf{t} = \mathbf{t}_x \odot \mathbf{t}_y \odot \mathbf{t}_z \in \mathbb{C}^N$ doing pointwise multiplication with a location specific phase shift, and a rotational part $\mathbf{R} \in \mathbb{R}^{N\times N}$ that is in fact a resampling matrix transforming the Cartesian grid $\mathbf{g} \in \mathbb{R}^{N\times 3}$ into the (phase-encode-wise) rotated grid $\mathbf{k} \in \mathbb{R}^{N\times 3}$ such that $\mathbf{Ay} = \mathbf{t}\odot(\mathbf{Wy})$. In particular, we have $\mathbf{t}_x = \exp(-2\pi i \mathbf{k}_x \odot \boldsymbol{\theta}_x)$, where the derivatives are given by $\mathbf{t}'_x = -2\pi i \boldsymbol{\theta}_x \odot \mathbf{t}_x$. Rotated grid vectors $\mathbf{k}^i \in \mathbb{R}^3$ are obtained from the Cartesian grid vectors by a local rotation $\mathbf{R}^i\mathbf{g}^i$ as achieved by a rotation matrix $\mathbf{R}^i \in \mathbb{R}^{3\times 3}$ depending on three rotation parameters a_i, b_i, c_i from $\boldsymbol{\theta}_i \in \mathbb{R}^6$. Furthermore, entries of the gridding matrix W_{ij} are obtained by application of a windowing function $w : \mathbb{R} \to \mathbb{R}$ (we use cubic resampling windows for ease of GPU implementation) to the new k-space location k_j and the old Cartesian location g_i such that $W^{ij} = w(k^i_x - g^j_x)w(k^i_y - g^j_y)w(k^i_z - g^j_z)$. Using the derivative of the window function w', we define the matrix $\mathbf{W}'_a$ (and similarly $\mathbf{W}'_b$ and $\mathbf{W}'_c$) containing the derivatives $[\mathbf{W}'_a]_{ij} = w'(k^i_x - g^j_x)w(k^i_y - g^j_y)w(k^i_z - g^j_z)$. Also, the derivative $\mathbf{k}'_\# \in \mathbb{R}^{N\times 3}$ of the rotated k-space grid w.r.t. rotation parameters $\# \in \{a, b, c\}$ can be written as $(\mathbf{k}^i{}_\#)' = (\mathbf{R}^i_\#)'\mathbf{g}^i$, where $\mathbf{g}^i$ is the Cartesian location and $(\mathbf{R}^i_\#)' = \partial\mathbf{R}^i_\#/\partial\#_i \in \mathbb{R}^{3\times 3}$ is the derivative of the rotation matrix in w.r.t. the rotation parameters. We use the axis-angle representation for the matrices $\mathbf{R}$, but other representations such as Euler angles are possible, too. Putting everything together, we obtain

$$\begin{aligned} \mathbf{A}'_x\mathbf{y} &= (\mathbf{t}'_x \odot \mathbf{t}_y \odot \mathbf{t}_z) \odot (\mathbf{Wy}), \text{ and} \\ \mathbf{A}'_a\mathbf{y} &= \mathbf{t} \odot \sum_{\#\in\{a,b,c\}} (\mathbf{W}'_\#\mathbf{y}) \odot \mathbf{k}'_a. \end{aligned} \tag{4.23}$$

Computation-wise, matrix vector multiplications with $\mathbf{A}'$ have a complexity of $\mathcal{O}(N)$ each so that the derivatives w.r.t. each of the six DOFs of the general 3D case can be computed efficiently.

4.5.1.2 Entropy criterion

Let us compute the derivative $\partial\phi(\mathbf{u})/\partial\mathbf{u}$ for $\phi(\mathbf{u}) = -\mathbf{v}^\top \ln \mathbf{v} \in \mathbb{R}$ with $\mathbf{v} = \sqrt{\mathbf{u}\odot\bar{\mathbf{u}}}/\sqrt{\mathbf{u}^\mathsf{H}\mathbf{u}} \in \mathbb{R}^n$ and $\mathbf{u} \in \mathbb{C}^N$.

$$\begin{aligned}\frac{\partial \phi}{\partial \mathbf{v}} &= -(\mathbf{1} + \ln \mathbf{v}) \\ \frac{\partial \phi}{\partial \mathbf{u}} &= -\left(\mathrm{dg}(\mathbf{v} \odot \bar{\mathbf{u}}^{-1}) - (\mathrm{sign}(\mathbf{u}) \odot \mathbf{v})\mathbf{v}^\top / \sqrt{\mathbf{u}^\mathsf{H}\mathbf{u}}\right)(\mathbf{1} + \ln \mathbf{v}) \\ &= (\mathrm{sign}(\mathbf{u}) \odot \mathbf{v})[\mathbf{v}^\top(\mathbf{1} + \ln \mathbf{v})]/\sqrt{\mathbf{u}^\mathsf{H}\mathbf{u}} - \mathbf{v} \odot \bar{\mathbf{u}}^{-1} \odot (\mathbf{1} + \ln \mathbf{v})\end{aligned} \tag{4.24}$$

Here $\mathrm{sign}(u) = \begin{cases} 0 & |u| = 0 \\ \frac{u}{|u|} & \text{else} \end{cases}$ and $|u| = \sqrt{(\Re u)^2 + (\Im u)^2}$.

We use a stabilized variant to avoid issues with zero norm $\mathbf{u}$, where we use $\mathbf{v}_\epsilon = \sqrt{\mathbf{u} \odot \bar{\mathbf{u}} + \epsilon^2 \mathbf{1}}/\sqrt{\mathbf{u}^\mathsf{H}\mathbf{u} + \epsilon^2},\ 0 < \epsilon \ll 1$ instead of $\mathbf{v}$.

$$\frac{\partial \phi_\epsilon}{\partial \mathbf{u}} = \mathbf{u} \odot \mathbf{v} \odot \mathbf{p}/(\mathbf{u} \odot \bar{\mathbf{u}} + \epsilon^2 \mathbf{1}) - \mathbf{u}\frac{\mathbf{v}_\epsilon^\top \mathbf{p}}{\mathbf{u}^\mathsf{H}\mathbf{u} + \epsilon^2} \tag{4.25}$$

4.5.1.3 In translational parameters

In order to compute $\mathrm{d}\mathbf{A}_{\boldsymbol{\theta}}$, we first compute the derivatives for translation. Since translation is represented by a diagonal in Fourier space, we only need $\mathrm{d}\exp(-2\pi i t) = -2\pi i \exp(-2\pi i t)\mathrm{d}t$.

4.5.1.4 In rotational parameters

Rotations are a bit more involved since

$$\mathbf{R}(\mathbf{x}) = \mathbf{n}\mathbf{n}^\top + \cos(\alpha)\left(\mathbf{I} - \mathbf{n}\mathbf{n}^\top\right) + \sin(\alpha)\begin{bmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{bmatrix}, \mathbf{n} = \mathbf{x}/\left\|\mathbf{x}\right\|, \tag{4.26}$$

$$\alpha = \|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2} = f_1(\alpha)\mathbf{x}\mathbf{x}^\top + \cos(\alpha)\mathbf{I} + f_2(\alpha)\begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}, \tag{4.27}$$

$$f_1(\alpha) = \frac{1-\cos(\alpha)}{\alpha^2},\ f_2(\alpha) = \frac{\sin(\alpha)}{\alpha}, \tag{4.28}$$

$$\frac{\partial \alpha}{\partial \mathbf{x}} = \frac{\mathbf{x}}{\alpha},\ \frac{\partial \mathbf{x}\mathbf{x}^\top}{\partial x_i} = \mathbf{e}_i\mathbf{x}^\top + \mathbf{x}\mathbf{e}_i^\top, \tag{4.29}$$

$$f_1'(\alpha) = \frac{\alpha^2 \cdot \sin\alpha - 2\alpha(1-\cos\alpha)}{\alpha^4},\ f_2'(\alpha) = \frac{\alpha\cdot\cos\alpha - \sin\alpha}{\alpha^2}, \tag{4.30}$$

$$\begin{aligned}\frac{\partial \mathbf{R}}{\partial x_i} &= \frac{x_i}{\alpha}\left(f_1'(\alpha)\mathbf{x}\mathbf{x}^\top - \sin(\alpha)\mathbf{I} + f_2'(\alpha)\begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}\right) \\ &\quad + f_1(\alpha)\left(\mathbf{e}_i\mathbf{x}^\top + \mathbf{x}\mathbf{e}_i^\top\right) + f_2(\alpha)\frac{\partial}{\partial x_i}\begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}.\end{aligned} \tag{4.31}$$

We can use the spectral representation $\mathbf{X} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\mathsf{H}$ to represent the directional derivative $\frac{\mathrm{d}}{\mathrm{d}t}e^{\mathbf{X}+t\mathbf{D}}$ [64, Theorem 4.5]:

$$\frac{\mathrm{d}}{\mathrm{d}t}e^{\mathbf{X}+t\mathbf{D}} = \mathbf{U}\left[(\mathbf{U}^\mathsf{H}\mathbf{D}\mathbf{U}) \odot \mathbf{\Phi}(t)\right]\mathbf{U}^\mathsf{H},\ \phi_{ij}(t) = \begin{cases} (e^{t\lambda_i} - e^{t\lambda_j})/(\lambda_i - \lambda_j) & \lambda_i \neq \lambda_j \\ te^{t\lambda_i} & \lambda_i = \lambda_j \end{cases}. \tag{4.32}$$

The gradient $\frac{\partial}{\partial X_{ij}}e^{\mathbf{X}}$ is its special form that comes up when using the matrix $\mathbf{E}_{ij}$ that is zero everywhere except for $[\mathbf{E}_{ij}]_{ij} = 1$:

$$\frac{\partial}{\partial X_{ij}}e^{\mathbf{X}} = \mathbf{U}\left[(\mathbf{U}^\mathsf{H}\mathbf{E}_{ij}\mathbf{U}) \odot \mathbf{\Phi}\right]\mathbf{U}^\mathsf{H},\ \phi_{ij} = \begin{cases} (e^{\lambda_i} - e^{\lambda_j})/(\lambda_i - \lambda_j) & \lambda_i \neq \lambda_j \\ e^{\lambda_i} & \lambda_i = \lambda_j \end{cases}. \tag{4.33}$$

As a side remark, we note that 2d rotations can be decomposed into a sequence of three shear operations[1] [65]

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} = \begin{bmatrix} 1 & -\tan\frac{\alpha}{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \sin\alpha & 1 \end{bmatrix} \begin{bmatrix} 1 & -\tan\frac{\alpha}{2} \\ 0 & 1 \end{bmatrix} \tag{4.34}$$

allowing an arbitrary 3D rotation to be represented by a sequence of shears. Shears can efficiently be accomplished in the Fourier domain by phase multiplications. However, these phase multiplications cannot be carried out jointly in 3D k-space.

4.5.1.5 Continuous coordinate transforms

Once the derivative $\frac{\partial \mathbf{R}}{\partial x_i}$ is known, the derivative $\frac{\partial \mathbf{A}_{\boldsymbol{\theta}}}{\partial x_i}$ can be obtained by the following reasoning of continuous valued objects: Let $u(\mathbf{p})$ denote the 3d image cube depending on the position $\mathbf{p}$. Further, let $\mathbf{f}_{\mathbf{x}} : \mathbf{p} \mapsto \mathbf{p}'$ denote a coordinate transform e.g. $\mathbf{f}_{\mathbf{x}}(\mathbf{p}) = \mathbf{R}_{\mathbf{x}}(\mathbf{p} - \mathbf{c}) + \mathbf{c}$ with $\frac{\partial \mathbf{f}_{\mathbf{x}}}{\partial x_i} = \frac{\partial \mathbf{R}}{\partial x_i}(\mathbf{p} - \mathbf{c})$. Then the derivative $\frac{\partial}{\partial x_i} u\left(\mathbf{f}_{\mathbf{x}}(\mathbf{p})\right)$ is given by

$$\frac{\partial}{\partial x_i} u\left(\mathbf{f}_{\mathbf{x}}(\mathbf{p})\right) = \left(\frac{\partial u}{\partial \mathbf{f}_{\mathbf{x}}}\right)^{\top} \frac{\partial \mathbf{f}_{\mathbf{x}}}{\partial x_i} = \left(\frac{\partial u}{\partial \mathbf{f}_{\mathbf{x}}}\right)^{\top} \frac{\partial \mathbf{R}}{\partial x_i}(\mathbf{p} - \mathbf{c}). \tag{4.35}$$

4.5.1.6 Interpolated discrete images

A simple way of approximating derivatives of a pixelized image $U_{i,j}$ involves the calculation of finite differences $U_{i+1,j} - U_{i,j}$ and $U_{i,j+1} - U_{i,j}$. However, if the image is subject to a coordinate transform $\mathbf{f}$, we have to interpolate the new pixels with reference to the old ones.

A pixel image based on values $U_{i,j}$ placed on a discrete lattice grid $(i, j) \in \mathbb{N}^2$ can be extended to the continuous domain $\mathbb{R}^2$ by interpolation. Denoting by $\lfloor x \rfloor$ the largest integer smaller or equal to x and by $\{x\}$ the fractional remainder $\{x\} = x - \lfloor x \rfloor \in [0, 1)$ we can write

$$u(x, y) = \sum_{(i,j) \in \mathcal{N}(\lfloor x \rfloor, \lfloor y \rfloor)} w(x - i) w(y - j) \cdot U_{i,j} \tag{4.36}$$

in full generality. The interpolation weights w do not depend on the exact position

[1] `http://www.ocf.berkeley.edu/~fricke/projects/israel/paeth/rotation_by_shearing.html`

(x, y) but on the distance $(x - i)$ to the neighbor only.

Another viewpoint is a convolutional one. Suppose that the sampled image $\underline{u_s(x, y)}$ is given by a sum of Dirac delta functions centered at the grid locations

$$u_s(x, y) = \sum_{(i,j)} \delta(x - 1)\delta(y - j) \cdot U_{i,j}. \tag{4.37}$$

Then the resampled image can be obtained as the convolution of a window function $w(x)$ with finite support and $\int w(x)\mathrm{d}x = 1$

$$u(x, y) = w(x) \star w(y) \star u_s(x, y) = (w(x)w(y)) \star u_s(x, y). \tag{4.38}$$

A widely used interpolation scheme is bilinear interpolation (or multilinear interpolation for higher dimensional grids):

$$\begin{aligned} u(x, y) = (1 - \{x\})(1 - \{y\}) \cdot U_{\lfloor x \rfloor, \lfloor y \rfloor} + (1 - \{x\})\{y\} \cdot U_{\lfloor x \rfloor, \lfloor y+1 \rfloor} \\ +\{x\}(1 - \{y\}) \cdot U_{\lfloor x+1 \rfloor, \lfloor y \rfloor} + \{x\}\{y\} \cdot U_{\lfloor x+1 \rfloor, \lfloor y+1 \rfloor}, \end{aligned} \tag{4.39}$$

where using the identity $\{\lfloor x \rfloor\} = 0$, we have, if evaluated at the grid

$$u(x, \lfloor y \rfloor) = (1 - \{x\}) \cdot U_{\lfloor x \rfloor, \lfloor y \rfloor} + \{x\} \cdot U_{\lfloor x+1 \rfloor, \lfloor y \rfloor} \text{ and } u(\lfloor x \rfloor, \lfloor y \rfloor) = U_{\lfloor x \rfloor, \lfloor y \rfloor}. \tag{4.40}$$

Using the derivatives $\mathrm{d}\lfloor x \rfloor / \mathrm{d}x = 0$, $\mathrm{d}\{x\}/\mathrm{d}x = 1$ and $\mathrm{d}(1 - \{x\})/\mathrm{d}x = -1$, we can derive

$$\begin{aligned} \frac{\partial u(x, y)}{\partial x} &= (1 - \{y\}) \cdot \left(U_{\lfloor x+1 \rfloor, \lfloor y \rfloor} - U_{\lfloor x \rfloor, \lfloor y \rfloor}\right) + \{y\} \cdot \left(U_{\lfloor x+1 \rfloor, \lfloor y+1 \rfloor} - U_{\lfloor x \rfloor, \lfloor y+1 \rfloor}\right), \\ \frac{\partial u(x, \lfloor y \rfloor)}{\partial x} &= U_{\lfloor x+1 \rfloor, \lfloor y \rfloor} - U_{\lfloor x \rfloor, \lfloor y \rfloor} =: D^x_{\lfloor x \rfloor, \lfloor y \rfloor}, \\ \Rightarrow \frac{\partial u(x, y)}{\partial x} &= (1 - \{y\}) \cdot D^x_{\lfloor x \rfloor, \lfloor y \rfloor} + \{y\} \cdot D^x_{\lfloor x \rfloor, \lfloor y+1 \rfloor} \\ \Rightarrow \frac{\partial u(x, y)}{\partial y} &= (1 - \{x\}) \cdot D^y_{\lfloor x \rfloor, \lfloor y \rfloor} + \{x\} \cdot D^y_{\lfloor x+1 \rfloor, \lfloor y \rfloor} \end{aligned} \tag{4.41}$$

hence finite differences coincide with derivatives at grid points and general deriva-

tives correspond to linearly interpolated finite differences. In general, we find

$$\frac{\partial u(x,y)}{\partial x} = \sum_{(i,j)\in\mathcal{N}(\lfloor x\rfloor,\lfloor y\rfloor)} \frac{\partial w_{i,j}(\{x\},\{y\})}{\partial x} \cdot U_{i,j}. \tag{4.42}$$

Nearest neighbor interpolation uses a box window function $w(x) = \begin{cases} 1 & |x| < \frac{1}{2} \\ 0 & \text{else} \end{cases}$, $w'(x) = 0$; linear interpolation uses the triangle window $w(x) = \max(1 - |x|, 0)$, $w'(x) = \frac{\text{sign}(x+1)}{2} + \frac{\text{sign}(x-1)}{2} - \text{sign}(x)$ and cubic interpolation can be used to approximate sinc interpolation. A practical kernel (continuously differentiable) [59] is given by

$$w(x) = \begin{cases} \frac{3}{2}x^3 - \frac{5}{2}x^2 + 1 & |x| \in [0,1] \\ -\frac{1}{2}x^3 + \frac{5}{2}x^2 - 4x + 2 & |x| \in (1,2] \\ 0 & \text{else} \end{cases} \tag{4.43}$$

$$w'(x) = \begin{cases} \frac{9}{2}x^2 - 5x & |x| \in [0,1] \\ -\frac{3}{2}x + 5x - 4 & |x| \in (1,2] \\ 0 & \text{else} \end{cases} \tag{4.44}$$

As a result, for discrete images $\mathbf{u}$ with rotation matrix $\mathbf{A}$ and derivative matrices $\mathbf{D}_1$, $\mathbf{D}_2$ and $\mathbf{D}_3$, this results in

$$\frac{\partial}{\partial x_i}\mathbf{A}_{\boldsymbol{\theta}}\mathbf{u} = \sum_{j=1}^{3}(\mathbf{D}_j\mathbf{A}_{\boldsymbol{\theta}}\mathbf{u}) \odot \mathbf{g}_j \Rightarrow \frac{\partial \mathbf{A}_{\boldsymbol{\theta}}}{\partial x_i} = \left(\sum_{j=1}^{3} \text{dg}(\mathbf{g}_j)\mathbf{D}_j\right)\mathbf{A}_{\boldsymbol{\theta}} \tag{4.45}$$

where $\mathbf{P} = [\mathbf{p}_1; \mathbf{p}_2; \mathbf{p}_3] \in \mathbb{N}^{3\times N}$ are the integer image coordinates and $\mathbf{G} = [\mathbf{g}_1; \mathbf{g}_2; \mathbf{g}_3] = \frac{\partial \mathbf{R}}{\partial x_i}(\mathbf{P} - \mathbf{c}\mathbf{1}^\top) \in \mathbb{R}^{3\times N}$ are the transformed image coordinates.

4.5.2 Motion trajectory parameterizations

The abstract objective function $\phi(\mathbf{u}, \boldsymbol{\theta})$ with $\mathbf{u} \in \mathbb{R}^N$ where $N = n_1 \cdot n_2 \cdot n_3$ and $\boldsymbol{\theta} \in \mathbb{R}^{6T}$ where $T = n_2 \cdot n_3$ depends on quite a few motion parameters in general. One has to reduce complexity in order to prevent undesired behavior, i.e. make the optimization problem ill-defined.

An obvious thing to do is to keep changes in position/orientation rather small, i.e.

to penalize $\|\mathbf{D}_t\boldsymbol{\theta}\|_p$ where $\mathbf{D}_t$ computes the finite differences in temporal direction. However, subjects do not need to rotate around the center but most often around a different point, e.g. the support point where the back of the head touches the pillow.

Another idea is to cut down the parameter space directly by choosing a parameterized version of the trajectory in the first place

$$(\boldsymbol{\alpha}, t) \mapsto \boldsymbol{\theta}_t \tag{4.46}$$

and to work with a much smaller set of parameters.

As an example, let us consider a situation in which the patient moves only once (or a few times) during the entire acquisition. We can model that by a sigmoid function

$$\mathrm{sig}(t) = \frac{1}{1+\exp(t)},\ \mathrm{sig}'(t) = \mathrm{sig}(t)\cdot(1-\mathrm{sig}(t)) \tag{4.47}$$

such that

$$\boldsymbol{\theta}_t = \mathrm{sig}(\beta(t-t_0))\cdot\boldsymbol{\theta}_T,\ \boldsymbol{\alpha} = \{\beta, t_0, \boldsymbol{\theta}_T\}. \tag{4.48}$$

In order to compute derivatives

$$\frac{\partial}{\partial\boldsymbol{\alpha}}\phi(\mathbf{u}, \boldsymbol{\theta}_{\boldsymbol{\alpha}}) = \frac{\partial\boldsymbol{\theta}_{\boldsymbol{\alpha}}^\top}{\partial\boldsymbol{\alpha}}\frac{\partial\phi(\mathbf{u}, \boldsymbol{\theta}_{\boldsymbol{\alpha}})}{\partial\boldsymbol{\theta}} \tag{4.49}$$

via the chain rule, we need to know the Jacobian $\frac{\partial\boldsymbol{\theta}_{\boldsymbol{\alpha}}^\top}{\partial\boldsymbol{\alpha}}$ of the parameterization $\boldsymbol{\theta}_{\boldsymbol{\alpha}}$.

4.6 Raw vs. magnitude-only data

In many practical cases, the raw data might not be available and only a magnitude phase-free image is left after scanning. It is not clear whether magnitude-only images can be corrected for motion. Lets treat the degradation as a convolution in spatial domain and see what is happening when phases are removed. For raw data we have $\mathbf{z} = \mathbf{k} \star \mathbf{u}$, where $\mathbf{k}$ is a convolution kernel due to motion (let's consider translation-only case) in spatial domain, and $\star$ is the convolution operator. We are now interested in what will happen if, instead of $\mathbf{z}$, we are given $\mathbf{z} \odot \bar{\mathbf{z}}$.

As an example which contains the properties of the problem, though rather trivial

in nature, we can consider a two-element image $\mathbf{u} \in \mathbb{C}^2$ and a two-element kernel $\mathbf{k} \in \mathbb{R}^2$, with entries equal to $\mathbf{u} = [r_1 e^{ip_1}, r_2 e^{ip_2}]$ and $\mathbf{k} = [k_1, k_2]$ respectively. We assume the kernel to be real-valued and the image to be complex-valued with magnitudes equal to $\mathbf{r}$ and phases to $\mathbf{p}$. After the convolution, the first element of $\mathbf{z}$, which is enough for the analysis, will be equal to $z_1 = k_1 r_1 e^{ip_1} + k_2 r_2 e^{ip_2}$. We are now interested in $z_1 \bar{z}_1$:

$$\begin{aligned} z_1 \bar{z}_1 &= (k_1 r_1 e^{ip_1} + k_2 r_2 e^{ip_2})(k_1 r_1 e^{-ip_1} + k_2 r_2 e^{-ip_2}) \\ &= k_1^2 r_1^2 + k_2^2 r_2^2 + k_1 k_2 r_1 r_2 \{e^{i(p_1 - p_2)} + e^{-i(p_1 - p_2)}\} \\ &= k_1^2 r_1^2 + k_2^2 r_2^2 + 2 k_1 k_2 r_1 r_2 \cos(p_1 - p_2) \\ &= (k_1 r_1 + k_2 r_2)^2 + 2 k_1 r_1 k_2 r_2 \{\cos(p_1 - p_2) - 1\} \end{aligned} \tag{4.50}$$

We see that the degraded result is not a convolution anymore as soon as $p_1 \neq p_2$, $r_1 > 0$ or $r_2 > 0$. Even worse it depends on the phase difference, which is lost. In a case where the phase variation is smooth over the image or one of the magnitudes r_1, r_2 is nonzero, we can drop the term with phases, but this is not likely to be the case for real images.

4.6.1 Magnitude-based data fidelity term

We have a parameterized matrix $\mathbf{A}_{\boldsymbol{\theta}}$ and a likelihood $\varphi(\mathbf{u}, \boldsymbol{\theta})$ of the form

$$\varphi(\mathbf{u}, \boldsymbol{\theta}) = \mathbf{1}^\top \boldsymbol{\psi}\left(\boldsymbol{\phi}(\mathbf{F}^{\mathsf{H}} \mathbf{A}_{\boldsymbol{\theta}} \mathbf{u}) - \boldsymbol{\phi}(\mathbf{F}^{\mathsf{H}} \mathbf{y})\right), \tag{4.51}$$

where $\mathbf{F} \in \mathbb{R}^{N \times N}$ is a Fourier transform matrix, $\boldsymbol{\phi}(\mathbf{x})_i = \phi_i(\mathbf{x}) = |x_i|$ and $\boldsymbol{\psi}(\mathbf{x})_i = \psi_i(\mathbf{x}) = |x_i|^2$ take into account that only magnitude of the image is known. Note that $\mathbf{u} \mapsto \varphi$ is convex if all $\mathbf{x} \mapsto \psi_i(\mathbf{x})$ are convex and non-decreasing in each x_i and all $\mathbf{x} \mapsto \phi_i(\mathbf{x})$ are convex. This is clearly the case for our choice from above. If the optimization turns out to be problematic, we can add another regularizer to penalize strong phase deviations

$$\xi(\mathbf{s}) = \mathbf{1}^\top \arg(\mathbf{s}) = \sum_{i=1}^{N} \arg(s_i), \ \mathbf{s} = \mathbf{B}\mathbf{u} \tag{4.52}$$

where we used a linear transformation $\mathbf{B}$ (e.g. a derivative matrix) and the arg : $\mathbb{C} \to \mathbb{R}$ function (highly non-convex but differentiable) defined by

$$\arg(z) = \theta,\ z = re^{i\theta}. \tag{4.53}$$

Note that the derivative $\mathrm{d}\arg(z)/\mathrm{d}z$ is very simple

$$\frac{\mathrm{d}\arg(z)}{\mathrm{d}z} = \partial\arg(z)/\partial a + i\partial\arg(z)/\partial b = i\frac{z}{|z|^2} = i\frac{\mathrm{sign}(z)}{|z|},\ z = a + bi. \tag{4.54}$$

To do the optimization we need to compute $\frac{\partial\varphi}{\partial\mathbf{u}}$ and $\frac{\partial\varphi}{\partial\boldsymbol{\theta}}$:

$$\frac{\partial\varphi}{\partial\mathbf{u}} = \mathbf{A}_{\boldsymbol{\theta}}^{\mathsf{H}}\mathbf{r}, \tag{4.55}$$

$$\frac{\partial\varphi}{\partial\boldsymbol{\theta}} = \mathbf{r}^{\mathsf{H}}\frac{\partial\mathbf{A}_{\boldsymbol{\theta}}}{\partial\theta_{i,t}}\mathbf{u}, \tag{4.56}$$

where $\mathbf{r} = \mathbf{F}\left(\dot{\boldsymbol{\phi}}(\mathbf{F}^{\mathsf{H}}\mathbf{A}_{\boldsymbol{\theta}}\mathbf{u}) \odot \dot{\boldsymbol{\psi}}\left(\boldsymbol{\phi}(\mathbf{F}^{\mathsf{H}}\mathbf{A}_{\boldsymbol{\theta}}\mathbf{u}) - \boldsymbol{\phi}(\mathbf{F}^{\mathsf{H}}\mathbf{y})\right)\right)$.

4.7 Avoiding bad local minima

Our cost function is highly nonlinear since it involves logarithms, complex exponentials and piecewise cubic interpolation polynomials. Minimizing such a function is difficult, due to the presence of local minima. We address this problem by using a multi-scale coarse-to-fine approach - see Algorithm 4.1 and Figure 4.2 for illustration. It is based on the observation that, in DC-centered k-space cubes containing low frequencies only, even strong motion will produce little offsets in the spatial domain. Furthermore, the signal intensity, and thus signal-to-noise ratio (SNR), is usually higher in low-frequency regions. This means that estimation and correction of motion is easier in coarsely-sampled versions of the image. Having estimated the motion parameters for the coarser scales, we proceed to the finer ones, for which we can use the motion parameters from coarser scales as an initialization. In this way, we drive the optimizer towards better minima in the objective landscape.

There is a fundamental difference between 2D and 3D multiscale optimization. In 2D, a temporal sequence of acquired views in the spectrum most often goes in

high-low-DC-low-high frequency order. In 3D, a sequence is more complex, and involves multiple alternations of high and low-frequency views (consider Cartesian line-by-line covering of the 3D k-space cube). In the first scale iteration, GradMC finds only the lowest frequency segments of the motion trajectory. These segments are surrounded by gaps corresponding to as yet unknown motion parameters of higher-frequency views. In each scale iteration the gaps shrink, until finally the whole trajectory is recovered. We exploit the alternation effect to our benefit – we initialize the motion parameters in the gaps by linearly interpolating from the values on their boundaries, where the algorithm has already determined the motion. This allows higher accuracy to be achieved in determining the motion in more problematic high-frequency views.

Algorithm 4.1 Multiscale optimization

Input: Corrupted volume $\mathbf{y}$ with $N = n_x \cdot n_y \cdot n_z$ k-space coefficients centered at $[c_x, c_y, c_z] = \left[\frac{n_x}{2}+1, \frac{n_y}{2}+1, \frac{n_z}{2}+1\right]$. Also, assume $n_y = n_z$.
Output: Restored volume $\mathbf{u}$ in spatial domain.
For $s \leftarrow \frac{64}{2}, ..., \frac{n_y}{2}$ **do**

- Use only k-space center of the raw data:
  $\hat{\mathbf{y}} \leftarrow \mathbf{y}(c_x,\ c_y{-}s \,..\, c_y{+}s,\ c_z{-}s \,..\, c_z{+}s)$.
- Compute best motion parameters:
  $\hat{\boldsymbol{\theta}} \leftarrow \arg\min_{\boldsymbol{\theta}}\ \phi(\mathbf{F}^{\mathsf{H}}\mathbf{A}_{\boldsymbol{\theta}}\hat{\mathbf{y}})$.
- Initialize central frequency part of $\hat{\boldsymbol{\theta}}$ on next finer scale:
  $\hat{\boldsymbol{\theta}} \leftarrow \boldsymbol{\theta}(c_y{-}s \,..\, c_y{+}s,\ c_z{-}s \,..\, c_z{+}s)$.

End
Finally, obtain the sharp image: $\hat{\mathbf{u}} \leftarrow \mathbf{F}^{\mathsf{H}}\mathbf{A}_{\hat{\boldsymbol{\theta}}}\mathbf{y}$.
See Figure 4.2 for an illustration.

4.8 Parallel imaging

The approach can be extended to images acquired with multiple coils. Since we optimize with respect to motion parameters, there is no need to know the sensitivity profiles of each coil. We seek the motion parameters $\hat{\boldsymbol{\theta}}$ that invert the motion on the data from each coil, such that the sum of the cost functions over all C coils is

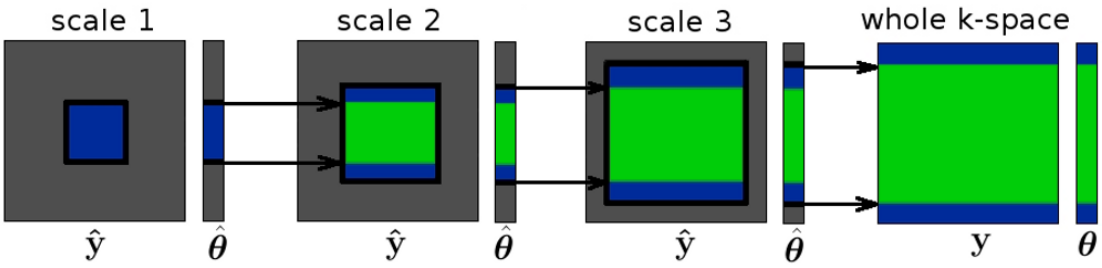

Figure 4.2: Multiscale optimization of motion parameters $\boldsymbol{\theta}$ is performed by progressively growing the processed data $\hat{\mathbf{y}}$ from the low-frequency center until the whole k-space is covered.

minimal

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\Theta}{\arg\min} \sum_{c=1}^{C} \phi\left(\mathbf{F}^{\mathsf{H}}\mathbf{A}_{\boldsymbol{\theta}}\mathbf{y}_c\right) + \lambda \left\|\mathbf{D}\boldsymbol{\theta}\right\|^2, \tag{4.57}$$

where $\mathbf{y}_c$ is the raw data from coil c.

Alternatively, if the coil sensitivities are known we can adjust our forward model such that

$$\mathbf{y}_c = \mathbf{A}_{\boldsymbol{\theta}}\mathbf{F}(\mathbf{s}_c \odot \mathbf{u}) + \boldsymbol{\varepsilon} \in \mathbb{C}^N \tag{4.58}$$

is pretty similar. Here, $\mathbf{s}_c \in [0,1]^N$ is a sensitivity map depending on the FOV of the coil and relative positioning of each. Assuming that the sensitivity maps $\mathbf{s}_c$ are known *a priori*, we can determine the motion parameters $\hat{\boldsymbol{\theta}}$ via

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_{c} \phi\left(\mathbf{F}^{\mathsf{H}}\mathbf{A}_{\boldsymbol{\theta}}\mathbf{y}_c\right) \tag{4.59}$$

and the underlying image via

$$\hat{\mathbf{u}} = \sum_{c} \mathbf{s}_{jc}^{-1} \odot (\mathbf{F}^{\mathsf{H}}\mathbf{A}_{\hat{\boldsymbol{\theta}}}\mathbf{y}_c) \text{ or } \hat{\mathbf{u}} = \mathbf{F}^{\mathsf{H}}\mathbf{A}_{\hat{\boldsymbol{\theta}}} \sum_{c} \mathbf{y}_c. \tag{4.60}$$

There is a trade-off between emphasizing noisy regions with $\mathbf{s}_c^{-1}$ and causing problems when optimizing the objective. Sensitivity profiles could be measured and made part of the optimization by means of a mask which penalizes all intensities outside itself.

4.9 Computational complexity

We use the limited memory BFGS nonlinear optimizer [66] with 50 iterations per scale. Note that the gradient of the cost function can be computed as fast as the objective function itself (see Appendix for details). The computational bottleneck are the fast Fourier transformations requiring an effort of $\mathcal{O}(N \cdot \log N)$ each. Figure 4.3 shows the computation times for various implementations against the total number of voxels in a 3D volume. Our pure Matlab implementation[2] is clearly impractical for large volumes, but has the advantage of clear structure with the entire pipeline represented by dedicated matrix classes [67] handling all major operations of the forward model. The first order speed-up was achieved by implementing the heavy resampling and gridding routines in plain C via Matlab's MEX interface. Finally, we took advantage of modern graphics cards (graphical processing unit, GPU), benefiting from the fact that most heavy operations in our pipeline are massively parallel. GPU implementation was done in CUDA language, and is also interfaced via MEX. For realistic volumes with 128^3 voxels, three minutes of processing are required. The current bottleneck of our GPU implementation is the result of a need to transfer the image from the main memory to the device memory of the GPU for each function evaluation. Thus, merging our GPU code with the code of the optimizer will make it possible to achieve a further speed-up. Experiments were done on an Intel(R) Core(TM)2 Duo CPU 2.66Ghz processor and a GeForce GTX 285 graphics card.

4.10 Experiments

This section contains the results of the experiments that we carried out to test our approach both in simulated and in vivo environments. The section progresses through several steps starting with simpler 2D simulated problems and ending with realistic 3D motion corrupted volumes acquired with an MR scanner.

4.10.1 Results: simulated data

Using simulated data is of benefit since it has ground truth for both the original clean image and the motion trajectory that was used to induce the motion transformation. This allows the stability and faithfulness of GradMC to be tested in recovering both

[2] Available from `http://mloss.org/software/view/430/`.

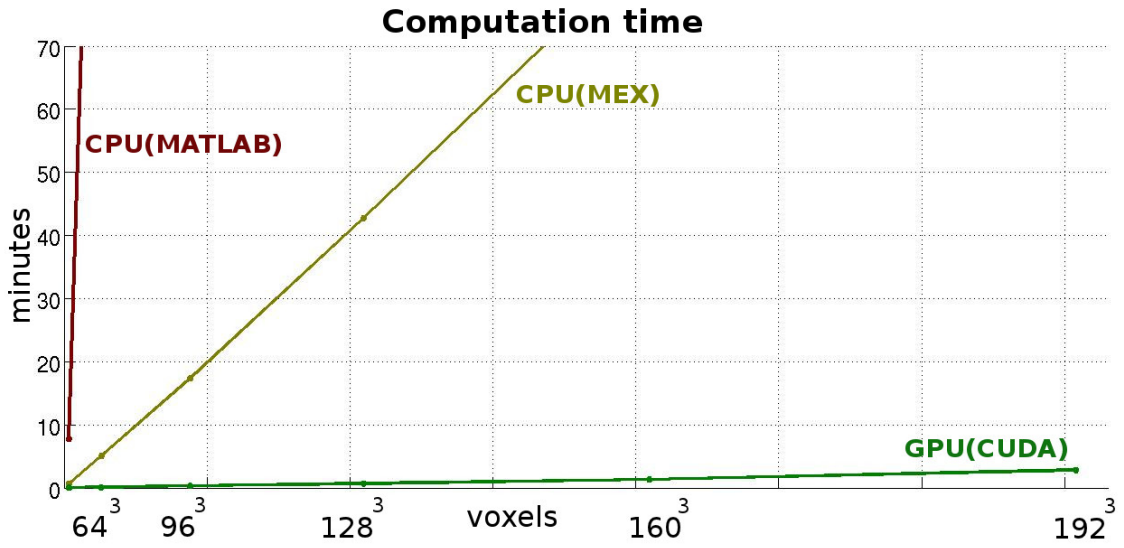

Figure 4.3: Comparison of the computation times for motion correction of 3D volumes of different sizes N for implementations in Matlab, plain C with MEX interface, and CUDA (GPU).

the image and motion parameters. We first carried out an experiment in a simpler 2D setting, where we used a motion-free image of a monkey's brain in order to generate motion corrupted data. We then compared the motion trajectory recovered by GradMC against the ground truth. The motion trajectory for each motion parameter was generated to be of a sinusoidal form (see Figure 4.4). The sine-form motion was chosen for convenient analysis of the recovered motion trajectories against the ground truth in all view/frequency ranges. In three separate experiments we simulated the translations in frequency encode direction, translation in phase encode direction and in-plane rotational motion.

In Figure 4.4 we show the results of the analysis of recovered trajectories. For a 2D image of the brain (matrix size N=192×192) we induced an out-of-phase sine-form motion simultaneously in all motion DOF (phase/frequency encode direction translations and in-plane rotation). In 100 successive trials, the optimizer was challenged with different random initializations which were in the range of the amplitude of the sinusoid. Plotted in solid are the ground truth sinusoidal trajectories. Dashed lines show the empirical mean of the recovered trajectories. The shaded gray tube around the empirical mean contains 95% of the probability mass computed by the respective quantiles. Analyzing the discrepancy of recovered mean trajectories to ground truth and the width of the error tubes, we observe that the accuracy is very high for trans-

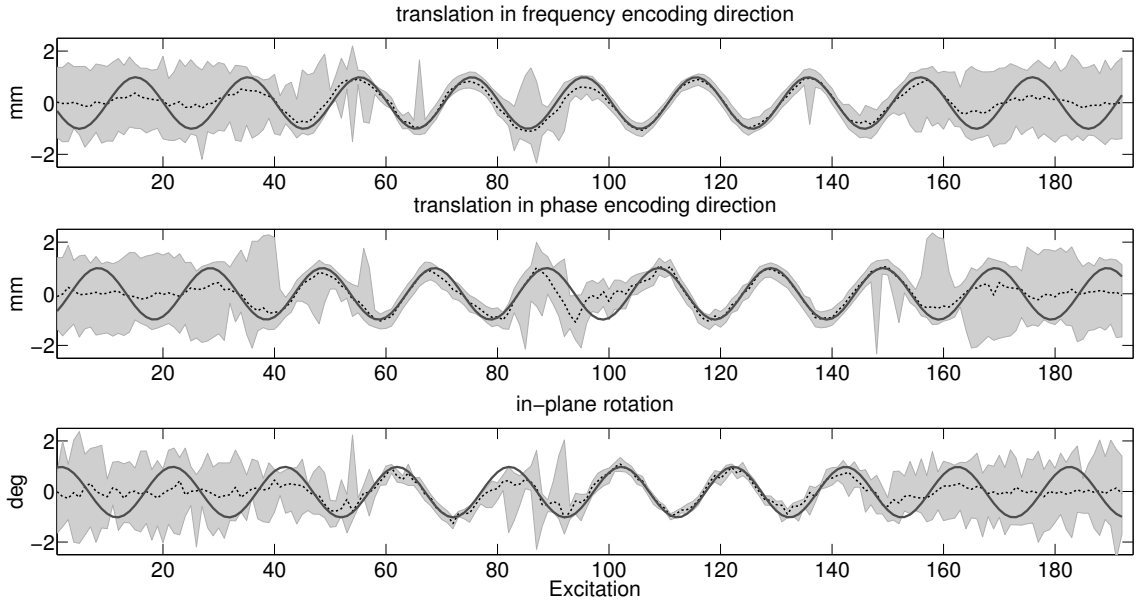

Figure 4.4: Recovery of underlying ground truth motion parameters in a 2D forward simulation with three DOFs starting from a random initialization. **Solid line:** underlying sinusoidal motion trajectory. **Dotted line:** the empirical mean over 100 recovered trajectories. **Shaded gray:** error tube over 100 recovered trajectories capturing 95% of the probability mass (corresponds to $\pm 2\sigma$ for Gaussian).

lations in low- and mid-frequency ranges (DC component is the center of abscissa), while moderate discrepancies occur in high-frequencies, likely due to low power of the signal in these regions of the spectrum. A large discrepancy also appears in the estimation of translational motion in phase direction around the DC component that is consistent over all trials, thus having a systematic origin. However, this discrepancy was not observed to cause deterioration in the restored image, probably because close-to-DC regions are not that sensitive to motion, with DC itself being completely invariant to it. Doing the restoration with motion parameters $\boldsymbol{\theta}$ initialized to zeros, the unitless image metric values computed on corrupted/restored/ground-truth images were equal to 1318/1114/1032, respectively.

In Figure 4.5, the same analysis is carried out in 3D (matrix size N=192×192×48). Since low- and high-frequency views are interleaved in 3D, the recovered trajectories have periodic large variance regions. These regions correspond to high-frequency views, and the intervals between their centers are equal to the matrix dimension in phase encode direction, so there are a total of 48 such regions for the given matrix size. One beneficial consequence of this is that if the true motion trajectory has

smooth variations, it is possible to actually constrain the variations in high-variance regions from the neighboring low-frequency supports.

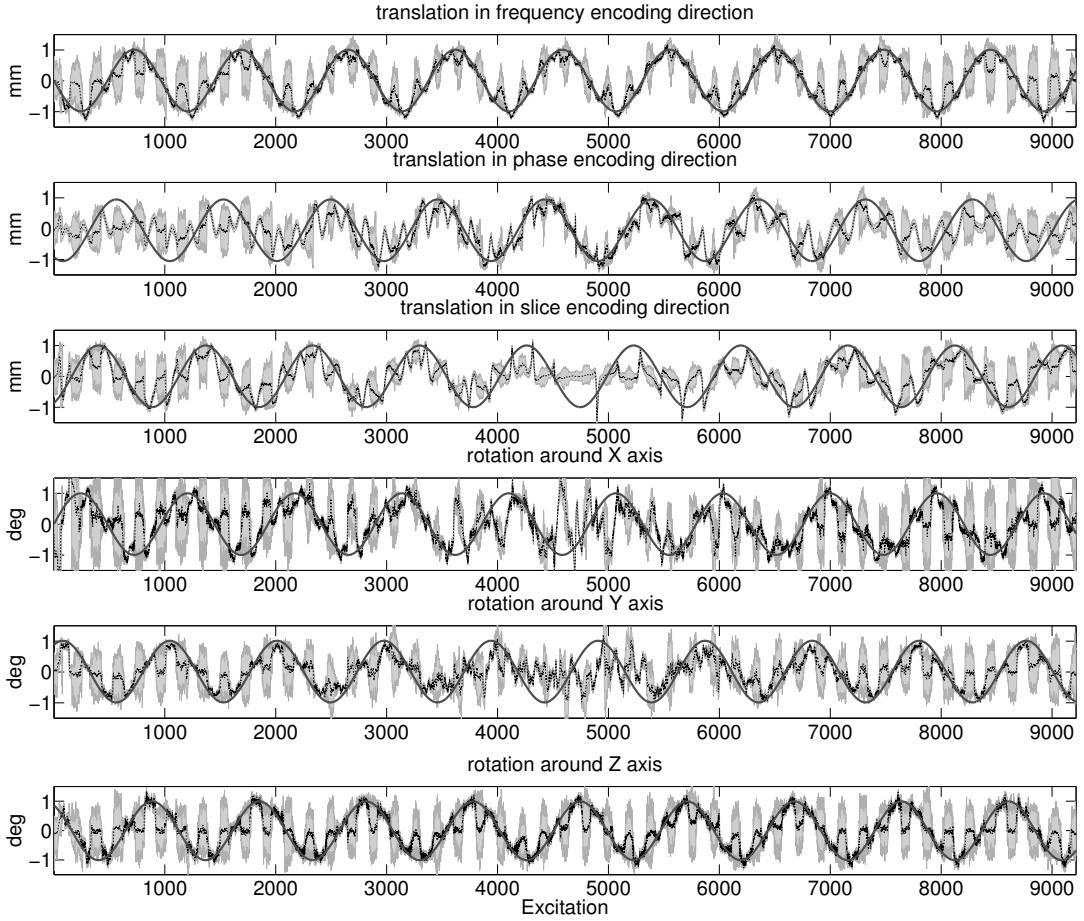

Figure 4.5: Recovery of underlying ground-truth motion parameters in a 3D forward simulation with six DOFs starting from random initializations. **Solid line:** underlying sinusoidal motion trajectory. **Dotted line:** the empirical mean over 100 recovered trajectories. **Shaded gray:** error tube over 100 recovered trajectories capturing 95% of the probability mass (corresponds to $\pm 2\sigma$ for Gaussian).

The last simulation experiment tests the stability of GradMC in an additive noise regime. In Figure 4.6, we show motion correction results together with recovered trajectories for the cases with and without additive Gaussian noise. As depicted in the right part of the figure, even adding substantial amount of noise does not break the ability of the algorithm to correct for the artifacts. However, the recovered motion trajectories seem to have larger deviations from the ground truth.

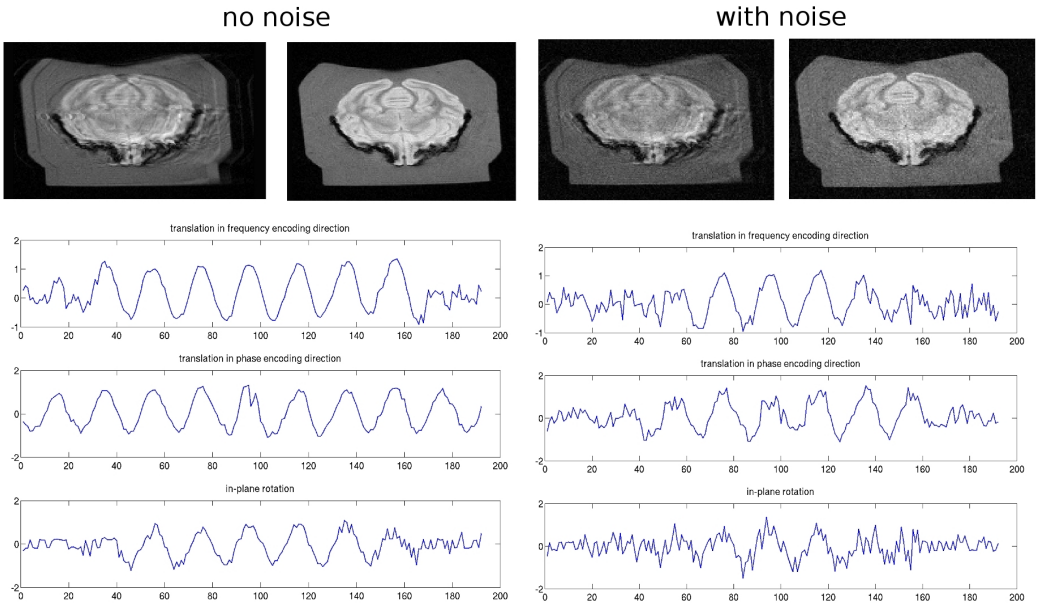

Figure 4.6: Left: Motion corrected image and recovered trajectory without additive noise. **Right:** Gaussian noise was added to the image with simulated motion. GradMC was still able to attenuate motion artifacts, although, the recovered motion trajectories were less accurate.

4.10.2 Results: real data

Imaging experiments were performed on a 3T scanner (Siemens Healthcare, Erlangen, Germany), using a single-channel birdcage coil for brain imaging. To obtain motion-free images and to inflict well-controlled motion and exclude non-rigid motion, a fixed monkey brain embedded in agarose was used for initial experiments. It was placed on a special, MR-compatible holder that allowed well-defined motion in 3D (two translational, one rotational DOF) during the scan.

In vivo human images were acquired on four subjects as approved by the local ethics board. The subjects, whose heads were loosely fixed inside the coil, were told to move during some of the experiments. Standard imaging protocols in two and three dimensions were used, with sequences and parameters similar to those used in medical routines and neuroscience research. For 2D images, a multi-slice RARE sequence was used. Acquiring eight echoes per echo train, four 4mm slices were imaged with a TE of 40ms and a TR of 2.5s within 75s. Including the scanner's automatic two-fold oversampling in read direction, a FOV of $400\times160\text{mm}^2$ was covered with a matrix size of 384×160. In addition, high-resolution data was obtained with a matrix size of 1024×666 over a FOV of $440\times286\text{mm}^2$. Those scans acquired

18 echoes per echo train with a TE of 61ms, a TR of 4s and a refocusing flip angle of 120° within 2.5 min. 3D in vivo data had a matrix size of 384×192×96 over a FOV of 440×220×115mm^3. FLASH images with a TE of 5.3ms, a TR of 16ms and a flip angle of 18° were acquired within 5min. For the monkey brain images, the FOV was reduced to 100×100×24mm^3 and only 16 slices were sampled with a flip angle of 35°. MP-RAGE scans had a similar duration (4.8min) and used an inversion time of 1.1s, a TE of 2.85ms, a TR of 1.5s and a flip angle of 9°.

Firstly, we show the results of applying our method to 2D images acquired by a RARE sequence. The motion corrupted and restored images (matrix size N=384×160) are shown in Figure 4.7. The subject made sideways head movements with rotation around the support point, where the head touched the table. The echo train length was equal to eight which, in principle, could allow us to reduce the number of motion parameters to be estimated by packing them into the train blocks. We did not exploit this advantage, and allowed the algorithm to determine the motion parameters for each view. The reconstructed result is of high quality compared to the observed image, where the anatomical details are almost fully occluded by ghosting artifacts. Additionally, the bottom part of Figure 4.7 shows the reconstruction of a high resolution (matrix size N=1024×640) RARE image, for which the method is also capable of improving the image quality.

We did additional experiments in 2D to test the multicoil procedure as described in the methods section. The results are shown in Figure 4.8, which compares the reconstruction using the data from each coil separately (middle) to simultaneous treatment (right). The reconstruction quality is slightly better if the coil data is motion-corrected jointly.

Next, we present the reconstruction results on more challenging 3D volumes. To completely avoid the effects of non-rigid motion, we first imaged the monkey brain. In Figure 4.9 we show the reconstruction results for 3D volume (matrix size N=384×192×16) acquired with FLASH sequence. From left to right two degraded/restored/no-motion slices are shown. GradMC was able to find an almost artifact-free solution.

The last set of experiments dealt with a realistic scenario of a freely moving human subject. We acquired 3D brain volumes of dimensions N=304×192×94 – the matrix size often used in clinical practice. The top rows of Figure 4.10 show the results from the FLASH sequence (matrix size N=384×192×96), while the bottom rows display

images acquired with an MPRAGE sequence.

For both sequences we have observed a consistent improvement in the image quality and removal of ghosting artifacts.

4.11 Summary

Given that the assumptions (which I have defined in the beginning of the chapter) are fulfilled, the GradMC method can be applied to any already acquired raw data. The strong advantage of the GradMC method (as well as most other autofocusing approaches) is that it does not require the use of any tracking equipment, or special imaging sequences. Our method is also general enough to be extendable to handle non-rigid body motion – a class of motions being out-of-scope of prospective correction approaches. I will show how our method can be extended to handle such a class of motions in the next chapter. On the downside, our algorithm is prone to the problem of irreversible loss of information due to large rotations and motion of the object out of the FOV – a common problem for retrospective methods (prospective approaches are immune to such issues). Luckily, the displacements involving large rotations are not likely to be a problem in clinical practice, since the head is usually softly constrained by cushions, physically limiting the range of possible movements.

As was shown in this chapter, gradient-based optimization is an efficient way to uncover motion parameters from motion-corrupted MRI scans. Even though the optimization is challenging due to nonlinearity, high dimensionality and heavy computational demand, the experimental results show that the proposed method can reconstruct undistorted images with substantial image quality improvement and reasonable computation time. In previous autofocusing approaches, motion parameters were obtained by application of trial translational/rotational displacements throughout the views in k-space and calculation of the image quality metric to find the best displacement for each view. The related follow-up works showed how, using intricate heuristics to select and group views to be optimized, it is possible to make the procedure faster. However, there is a price to pay in terms of complexity, accuracy and robustness of the algorithm. Challenged with high-resolution 3D volume inputs such methods run into trouble due to the curse of dimensionality and the associated enormous combinatorial search. Since the previous autofocusing approaches rely on trial sets of discrete displacements, it makes it hard to compare them with

our approach. Indeed, for any problem where the underlying motion trajectory has many deviations from the fixed offsets, trial-based reconstruction is fundamentally prone to error. In contrast, our method is immune to such problems, since any arbitrary continuous displacements are allowed. This makes a comparison of approaches unfair, and thus we do not carry out the respective experiments in this work.

There are two major challenges for our method: the first is due to the non-convex image quality metric that leads to the objective landscape with multiple local minima. We addressed this problem by using a multiscale algorithm. The idea is to initialize the motion parameters on each next-scale iteration with motion parameters from the previous coarser scale. This drives the optimization process into the vicinity of supposedly good local minima. It is much easier to accurately determine the motion parameters of low frequency regions in k-space because of higher signal power and less sensitivity to strong motion. The second problem is the computation time – an important limitation for many retrospective motion correction methods. We have developed a GPU version of GradMC featuring a highly parallel implementation of our forward model. For volumes of matrix sizes commonly encountered in clinical practice only a few minutes are needed for the reconstruction. There is still room for improvement with respect to computation time since, in the current implementation, it is necessary to move the data to the memory of the GPU and back for each function evaluation. We have used cubic interpolation to do the gridding, which might seem inferior to recent efficient gridding methods based on Kaiser-Bessel convolution kernels or jincs [68, 69]. The reason for using cubic interpolation is that it can be implemented simply and efficiently on GPUs. Better interpolation kernels require specialized mathematical functions like Bessel functions either not available on GPUs or slow to compute.

In our simulation experiments, the forward model of the motion degradation process was used to generate motion-corrupted data. This allowed for a simple test of stability of GradMC. The ground-truth motion trajectory was of sinusoidal form, which conveniently allows for control over the strength of motion (amplitude) and its time variation (frequency). Doing a motion correction on this simulated data, we started the optimization with random initializations of motion parameters. With respect to image quality, we invariably observed a substantial improvement for all initializations. Comparing recovered trajectories to ground truth, we observed that the mismatch in motion parameters was small for low- and mid-frequency views, and increasingly large towards high-frequency views. We expected to see such a

pattern which is due to an inverse power law drop of the signal strength from low to high frequencies. Additionally, the simulation experiment showed that multiple differing motion trajectories, when used to invert the motion, lead to similarly good reconstruction with respect to both image metric and visual quality. For 3D volumes, we also observed "cross-talk" effects between different DOFs, most strongly distinct between translational DOF in phase/slice encoding directions. This effect is most strongly observed if the data is simulated with motion in some DOF only. On reconstruction of such data, recovered motion trajectories show non-zero motion in "silent" DOF. Thus, artifacts produced by motion in one DOF can, to a certain extent, be mimicked by motion in another DOF. The bottom line is that the good match between the recovered trajectory and ground truth, although being desirable, is not always possible and this issue is particularly observable in high-frequency views. It seems, though, that the objective landscape allows for multiple local minima, which are also good in terms of image quality.

The acknowledged major limitation of our method is the inability to correct for motion that involves strong rotational components (angles larger than 3°). The visual quality of a reconstructed image never gets worse than that of a degraded image, however, the stronger the rotation, the less improvement can be achieved. For rotation angles larger than 10°, the corrected image looks essentially the same as the degraded one. This limitation stems from the reduced form of our objective, where instead of a data fidelity term (difference between the observation and the forward model applied to a motion-free estimate), we use an empirical inverse of the observed motion. An exact empirical inversion is not possible since, for strong rotations, the amount of missing k-space information is significant. Losses occur in high-frequency corners, and in k-space holes that appear due to the rotation of neighboring views against each other. The inversion-free treatment of the problem, capable of dealing with arbitrary rotations, involves only the forward model, taking the form of alternations between the estimation of motion parameters and recovery of the sharp image. Both alternating steps have data fidelity and regularization terms. Regularization terms are based on prior knowledge of the properties of motion parameters and the image. In the next chapter, I present more details on the outlined alternating procedure, and show how larger rotations and non-rigid motion can be corrected.

Instead of empirical inversion in a final image reconstruction step, it is possible to use a proper mathematical inverse, i.e. computed by conjugate gradients. In

respective experiments, we did not find the results obtained with a mathematical inverse to be better in terms of accuracy compared to the empirical inverse. The reason for this is that the objective function used to estimate motion parameters has an empirical inverse in its core resulting in inherently suboptimal estimates of motion parameters. A related effect, due to an empirical inversion in the objective, is that the pose of the object in the reconstructed image can be different compared to the ground truth. Indeed, translating/rotating the whole object does not influence the value of the quality metric. We do not consider this to be a flaw in the method.

We make a strong assumption that the imaged object behaves as a rigid body. Indeed, this is what allows us to carry out fast multiplications with the matrix $\mathbf{A}_{\boldsymbol{\theta}}$ in the Fourier domain. Slight deviations from non-rigidity are well-tolerated; however, gross effects (e.g. due to movement of the tongue during acquisition) make an artifact-free correction difficult, as we observed in a dedicated experiment. We also assume that the input to GradMC is the raw k-space data along with the order in which the k-space was sampled. In a clinical setting, such data is not always available, because often only magnitude images are preserved. Since real data from the scanner often has non-uniform spatial phase, taking the modulus of complex-valued pixels in the spatial domain results in distortions of the motion structure in the Fourier domain. It no longer holds that the same motion affects all coefficients in each k-space view. An interesting direction for future work would be to understand what clues about motion still survive in magnitude images, with the hope of exploiting them for corrections.

MR imaging is very flexible in the sense that it allows many different trajectories that can be used to traverse the k-space. We assume that the k-space is acquired in a Cartesian line-by-line fashion, which is commonly done with clinical sequences such as MPRAGE. Extending our approach to radial/spiral trajectories would be straightforward, i.e. for radial trajectories we would expect our method to achieve higher accuracy, since each k-space view would contain low-frequency coefficients, which are crucial for estimation of motion parameters.

In this chapter, I have touched upon the problem of correcting the data acquired with multiple coils. We have proposed a simple extension of the method that can handle such data. In our implementation, no knowledge of sensitivity profiles of coils is expected; however, if sensitivity profiles are known, GradMC can be modified to allow for an interesting variation of the image metric to be used to do the correction. The idea is to penalize the artifacts due to motion, which spread globally and occur

in the image regions, where the coil sensitivity is negligible. This would require using a mask, which can easily be obtained by subtracting the sensitivity profile from the unity vector. Such an image metric is interesting, since it does not make any assumption about the nature of the acquired data (i.e. no requirement for edges), and is only sensitive to the motion artifacts.

To summarize: in this chapter, I have presented a retrospective method aimed at blind estimation and correction of motion. The method uses the analytic formulation of the motion degradation process to search for motion parameters in a high-dimensional space being guided by the partial derivatives of the cost function based on the entropy of the gradients image metric. The results of the experiments substantiate the great potential of the proposed method.

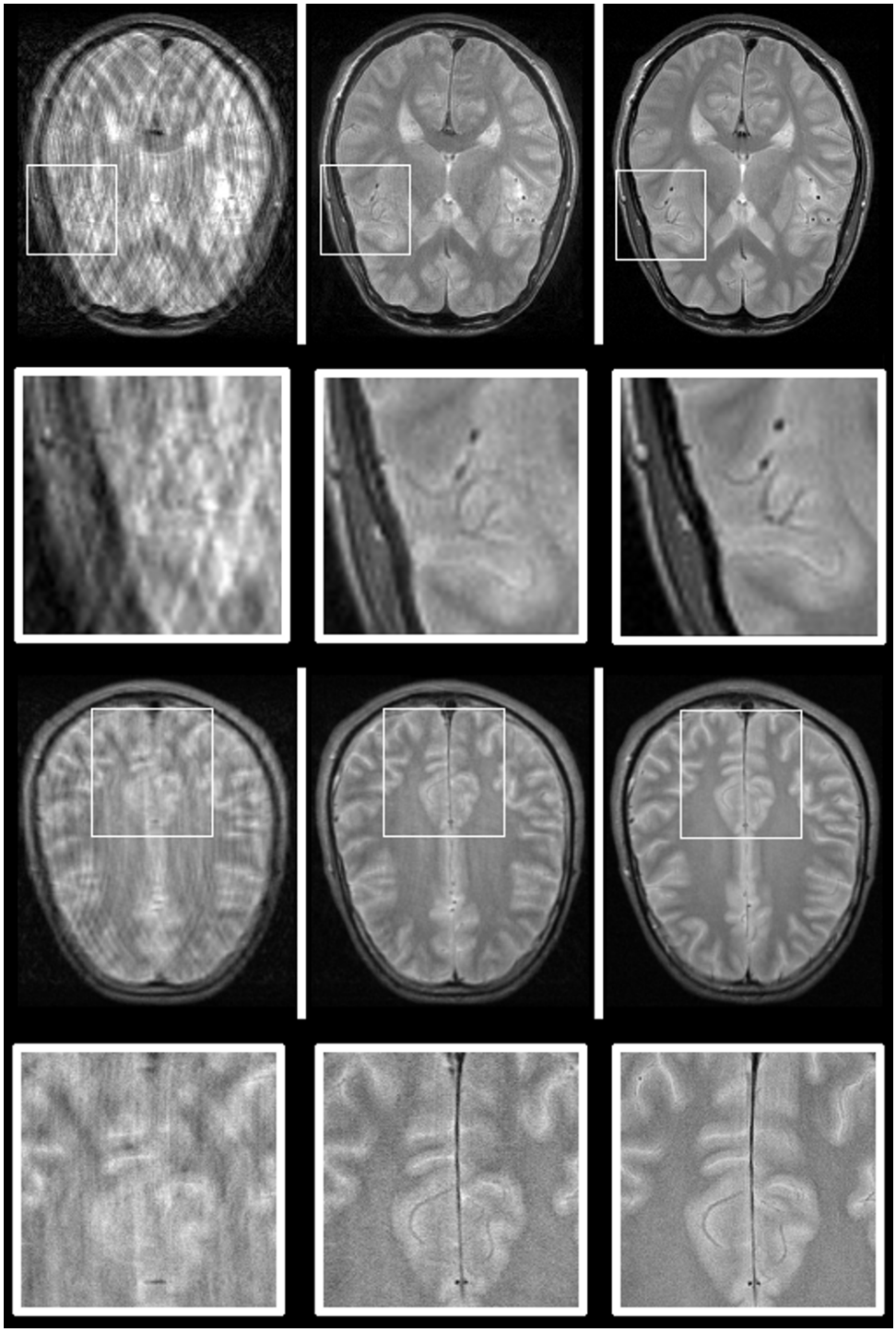

Figure 4.7: Motion correction of 2D RARE images in a freely moving human subject. **Left:** motion corrupted image. **Middle:** reconstruction. **Right:** no motion image. **Top:** low resolution. **Bottom:** high resolution (matrix size N=1024×640).

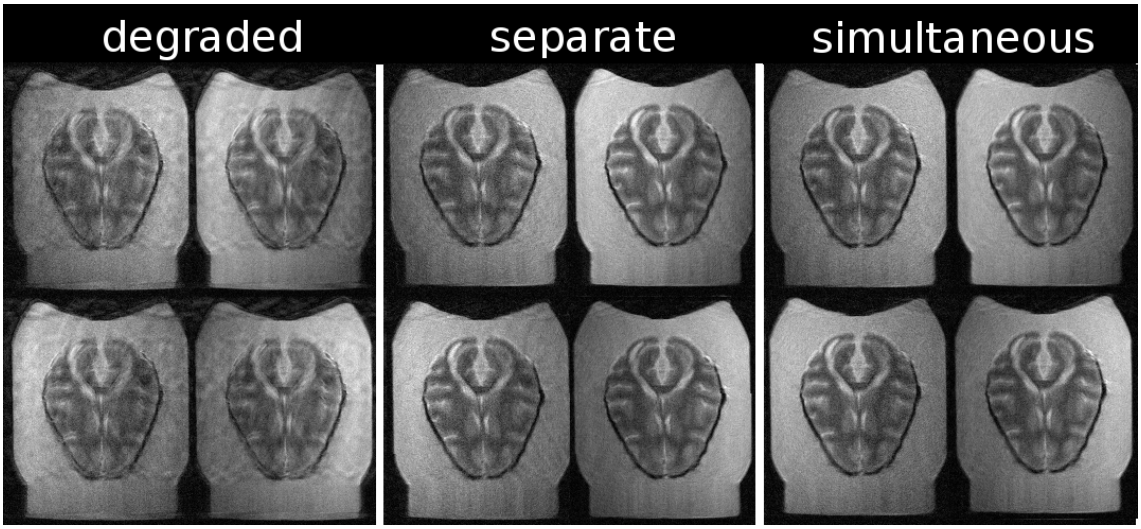

Figure 4.8: Motion correction of data recorded by 4 coils in a monkey brain in fixation gel. **Left:** motion corrupted data. **Middle:** reconstruction using the data from each coil separately. **Right:** reconstruction with simultaneous optimization of image metric over the data from all coils.

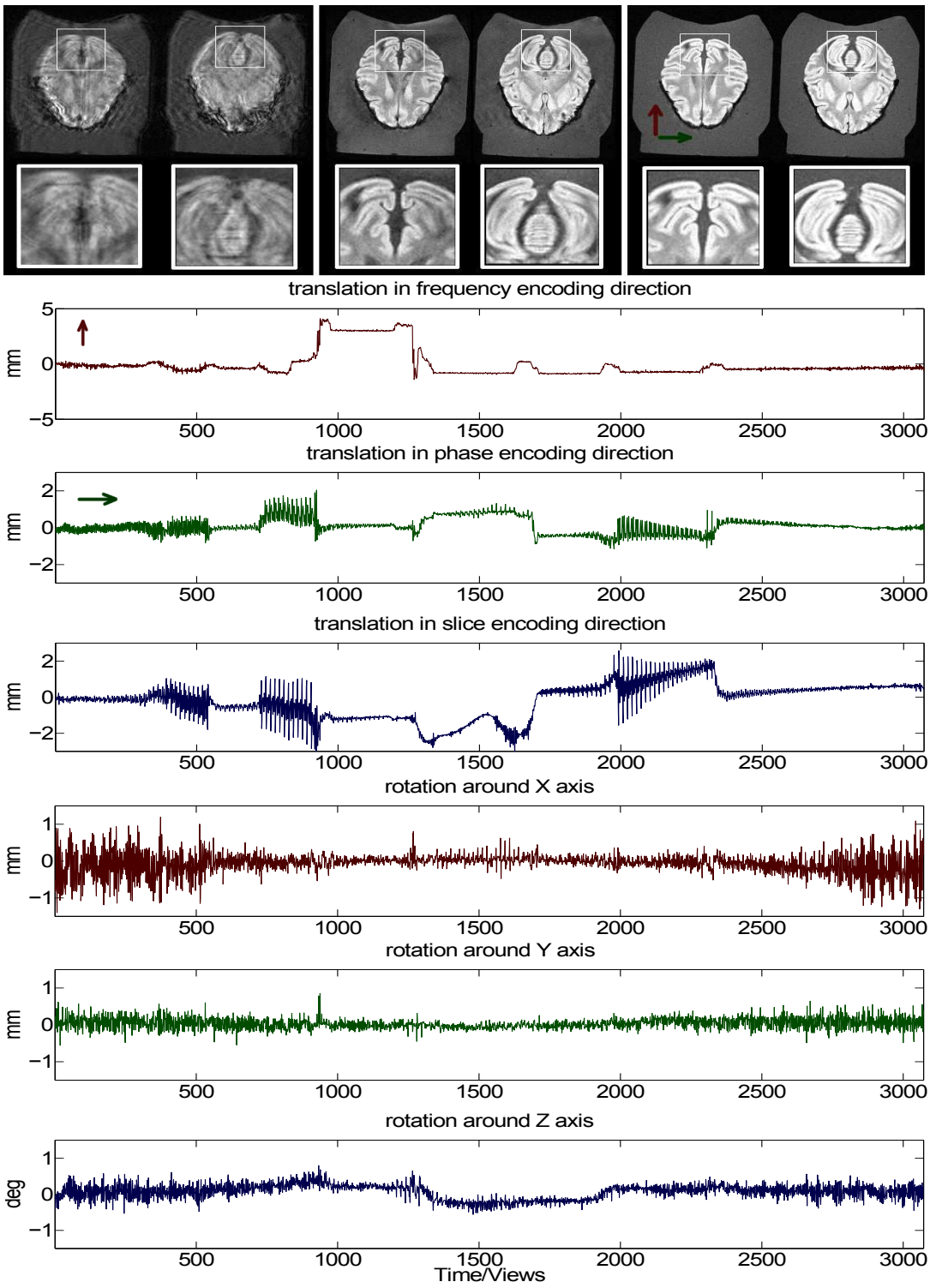

Figure 4.9: Motion correction of 3D FLASH images in a monkey brain in fixation gel. Motion was induced by manually moving the probe in the scanner by means of a specialized device. **Left:** motion corrupted slices. **Middle:** reconstruction. **Right:** no motion image. **Bottom:** recovered trajectories (from top to bottom - translations, then rotations)

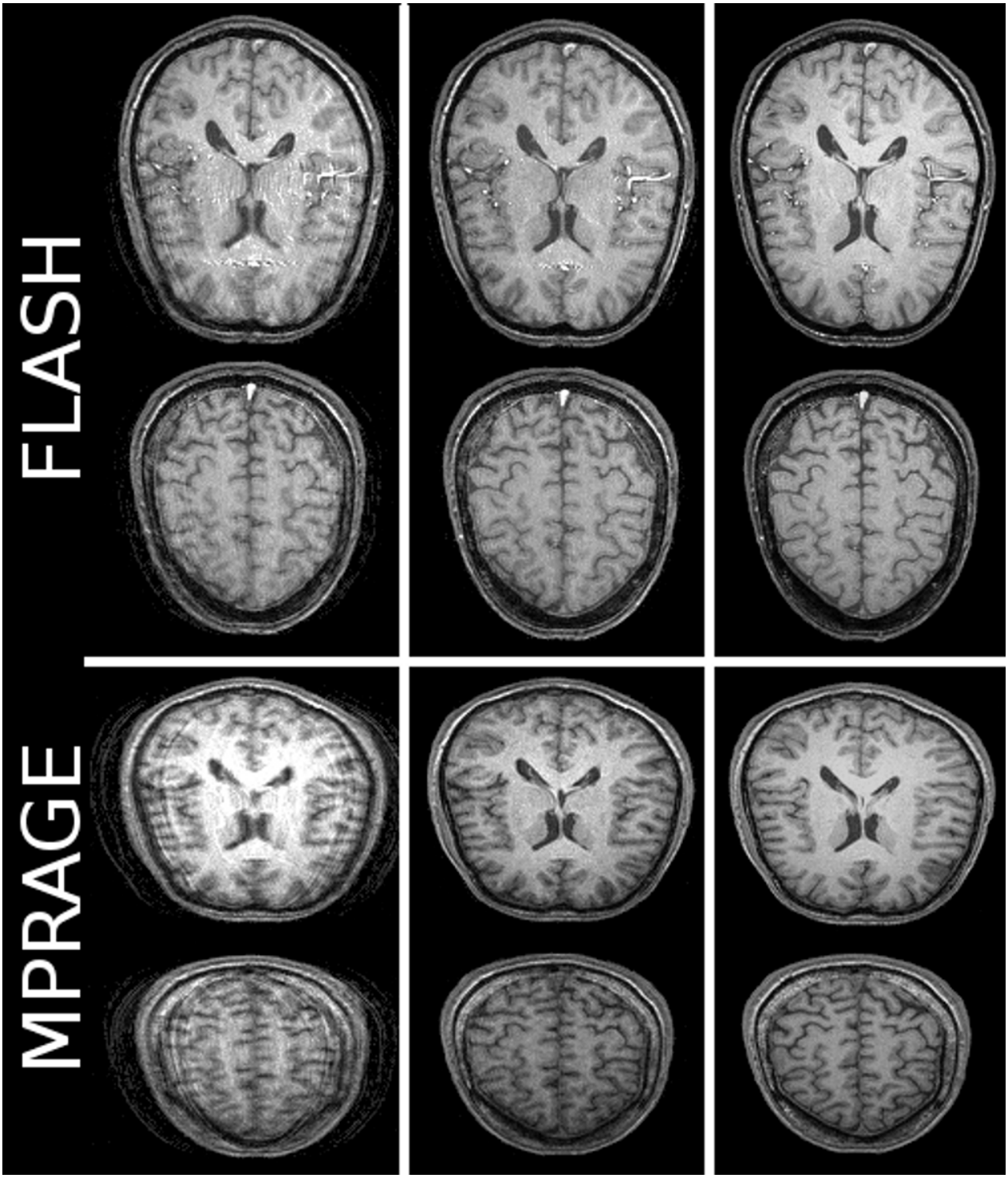

Figure 4.10: Motion correction of 3D FLASH and MPRAGE images in a freely moving human subject. **Left:** motion corrupted slices. **Middle:** reconstruction. **Right:** no motion image.

5 Multi-rigid motion correction

In this chapter, I present an extension of our blind retrospective motion correction method aimed at correction of non-rigid motion. Once again, the proposed method is purely image data-driven meaning that our correction algorithm (GradMC2) does not require information about motion from external sensors as input. Being able to carry out such reference-free correction in the case of complicated spatially-varying motion is the main contribution we make with the current work. Furthermore, the method can be used with generic acquisition sequences and is associated with a reasonable computation time. We approximate non-rigid motion by locally (patch-wise) rigid motions, and extend the forward model formulated in the previous chapter to accommodate this "multi-rigidness" assumption. The method is autofocusing-based, which means that, at the core, is a generic image quality functional. Given the high complexity of the problem to be solved, and the large number of unknowns, it is crucial to be able to come up with an efficient optimization algorithm to recover both the underlying sharp image and local displacements characterized by distinct sets of motion parameters specified for each patch. We come up with an annealing-based alternating optimization procedure, where we benefit from the closed form formulation of our forward model in order to do efficient optimization using derivatives of the quality functional with respect to both latent image and motion parameters.

As an ultimate goal, we aim for our method to be applied in routine clinical imaging, where the scans are affected by non-rigid physiological motion (e.g. in abdominal scans). In this study, we take a step towards this goal and, in our in vivo experiments, we correct for multi-rigid finger motion in wrist imaging, which can be seen as a reasonable simplification of the realistic non-rigid problem. The first part of this chapter details how to combine multiple locally rigid motions in image patches into a joint global non-rigid motion model and introduces the concept of image patch separability. Secondly, I formulate the trade-off between data fidelity and restored image quality measures and introduce our nested-loop optimization algorithm to recover the sharp image and the underlying motion parameters. At

the end of the chapter experimental results on both simulated and real data are presented followed by concluding remarks.

5.1 Non-rigid motion transformation in Fourier domain

Non-rigid body motion leads to image distortions that cannot be described by a single *global* pair of translation and rotation vectors. We consider (approximately rigid but) *local* motion inside an image patch, which we will call *multi-rigid* in the following analysis. We use the term "multi-rigid" instead of the more common "non-rigid" to emphasize that we view a general non-rigid motion as the extreme case, where each point of the body of interest possesses a unique motion vector. Note that any linear operation on the volume can be modeled by multi-rigid transformations if only the patches are made small enough, i.e. the number of patches P is in the order of the number of voxels N. We illustrate our model in Figure 5.1.

Formally, we partition the image volume into P spatial patches using windows $\mathbf{w}_p \in [0,1]^N$. This description is similar to the Efficient Filter Flow framework [70] that aims to solve a non-stationary deconvolution problem. We do not impose a particular shape nor do we require connectedness of the patches. The patches could, at least theoretically, be estimated from the measured k-space data $\mathbf{y}$; however, for now, we assume that they are given. Doing the normalization $\sum_{p=1}^{P} \mathbf{w}_p = \mathbf{1}$ we can interpret the windows $\mathbf{w}_p$ as the probabilities that the voxels will belong to the spatial patch p. Implementation-wise, we define the spatial windows using binary masks, which we convolve with a Gaussian kernel to enforce a smooth transition between the motion parameters of neighboring patches. Note that the width of the Gaussian kernel governs the influence region and the steepness of the motion parameter transition. The k-space equivalent of the pointwise multiplication with a spatial window $\mathbf{w}_p$ is a convolution matrix $\mathbf{W}_p$ such that $\mathbf{W}_p\mathbf{F}\mathbf{u} = (\mathbf{F}\mathbf{w}_p)\star(\mathbf{F}\mathbf{u}) = \mathbf{F}(\mathbf{w}_p \odot \mathbf{u})$, where $\star$ is the convolution operation and $\odot$ denotes pointwise multiplication. Each of the P patches has an associated motion trajectory parameterized by $\boldsymbol{\Theta}_p$. We can write the multi-rigid forward model as

$$\mathbf{y} = \sum_{p=1}^{P} \mathbf{A}_{\boldsymbol{\Theta}_p}\mathbf{F}(\mathbf{w}_p \odot \mathbf{u}) + \boldsymbol{\varepsilon} = \mathbf{A}_{\boldsymbol{\Theta}}\mathbf{F}\mathbf{u} + \boldsymbol{\varepsilon} \in \mathbb{C}^N, \tag{5.1}$$

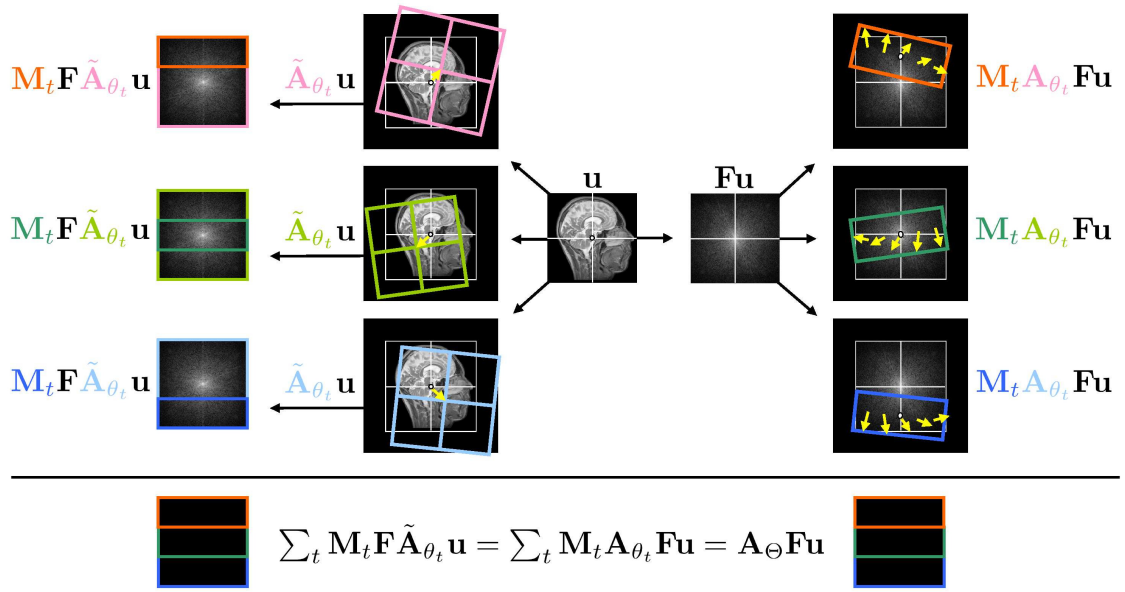

Figure 5.1: Illustration of the forward model. Left panel, rigid: an unknown underlying image $\mathbf{u}$ is affected by a linear transformation which is given by the operator $\tilde{\mathbf{A}}_{\boldsymbol{\theta}_t}$ in the spatial domain (left part of the panel), or by $\mathbf{A}_{\boldsymbol{\theta}_t}$ in the Fourier domain (right part) at every time point t. The observed image (bottom row) is obtained by summation (over the entire acquisition time $t = 1..T$) of k-space segments extracted by the trajectory matrix $\mathbf{M}_t$. Right panel, multi-rigid: the image is split into P patches (brain/jaw) by means of spatial windows $\mathbf{w}_p$. For each spatial window, a rigid motion (left panel) is applied with a distinct set of motion parameters $\boldsymbol{\Theta}_p$. The observed image (bottom row) is obtained by summing over spatial patches.

where we use P different rigid motion matrices $\mathbf{A}_{\boldsymbol{\Theta}_p}$ to construct the multi-rigid transformation $\mathbf{A}_{\boldsymbol{\Theta}} = \sum_{p=1}^{P} \mathbf{A}_{\boldsymbol{\Theta}_p}\mathbf{W}_p$ with an overall number of $6 \cdot T \cdot P$ free parameters. Note that for $P = 1$, we have $\mathbf{w}_p = \mathbf{1} \Rightarrow \mathbf{W}_p = \mathbf{I}$ that is we recover the rigid motion model introduced in the previous section as a special case.

It is well-known that the artifacts due to motion manifest as ghosts that propagate in phase encode direction and blur in the direction(s) of motion (see [1, 71]). In the case of pure translations, the motion degradation is described by a convolution with some point-spread-function, which has a large support (attributed to global ghosts) in phase encode direction, and much smaller support (resulting in local blur) in frequency encode direction.

This has strong implications when it comes to non-rigid motion correction. Consider a 2D multi-rigid problem, where the image is half-split into two rectangular

patches. If the image is split in frequency encode direction, we call the problem *separable*. This means that it is possible to segment the image into two patches in spatial domain by means of spatial windows, and then treat the two resulting images as two separate rigid motion correction problems, which can be addressed by existing retrospective rigid motion correction methods. In particular, the problem can be solved using our inverse-based approach (described in the previous chapter) in mere seconds of time. It is important to note that, since blurring also propagates in frequency encode direction, the problem is, strictly speaking, not completely separable. However, since the support of point-spread-function in frequency encode direction is small, the use of spatial windows with smooth cut-off/overlap allows the problem to be treated as approximately separable subject to small localized errors near the window boundary. In many cases, where the boundary between rigid parts has no diagnostic value these small errors might be tolerated. So far we were assuming that the image is affected by pure translational motion. It is important to note that a strong rotational motion (in a range of more than a couple of degrees) makes the problem non-separable, irrespective of whether the patches are arranged in phase or frequency encode directions.

In the non-separable case, the patches are arranged in phase encode direction, which leads to a considerable "artifact cross-talk" between the adjacent patches resulting in a genuine multi-rigid problem. We see the separable case as a "trivial" one, where a solution might be obtained using rigid motion correction methods. Our goal is to solve the non-separable problem and, in our experiments, we arrange the patches (and assume multiple motion trajectories) in phase encode direction. We explicitly test the property of separability in a dedicated experiment (see Results section).

5.1.1 Comparison with spatial domain formulation

Instead of doing the computations in the Fourier domain using P rigid motion transformations as depicted in the right panel of Figure 5.1, we can swap the summation over time t and patches p and equivalently evaluate the forward model in the spatial domain

$$\mathbf{y} = \sum_{t=1}^{T} \mathbf{F}\tilde{\mathbf{A}}_{\boldsymbol{\theta}_t}\mathbf{u} + \boldsymbol{\varepsilon},\ \tilde{\mathbf{A}}_{\boldsymbol{\theta}_t} = \sum_{p=1}^{P} \tilde{\mathbf{A}}_{\boldsymbol{\theta}_{t,p}}\mathrm{diag}(\mathbf{w}_p) \tag{5.2}$$

leaving the result unaltered but choosing a different computational route. Here, the main difference is the computational complexity of (a) the spatial approach and (b) the Fourier approach:

- Method (a) is dominated by T fast Fourier transforms $\mathcal{O}(\frac{2}{3}N \log N)$ following the application of the oversampled multi-rigid transformation $\mathcal{O}(2^D N)$ yielding $\mathcal{O}(T \cdot N \cdot (\frac{2}{3} \log N + 2^D))$.
- Method (b) is governed by the complexity of P rigid models where double oversampling is used for the Fourier transform and an additional double oversampling for the rotation interpolation leaving an overall $\mathcal{O}(P \cdot N \cdot (4^D + 2^D \log N))$.

Setting $n := n_x = n_y = n_z = 200$, $N = n^3$ as data size, $D = 3$ as data dimension, $T = n$ as the number of motion parameters (piecewise constant motion trajectory), and equating the complexities of (a) and (b), we find $P \approx 19$. For $n = 100$, we obtain $P \approx 10$ which illustrates how complex the non-rigid motion can become until spatial computations of method (a) are more efficient. Since modern processors and GPUs allow parallel computations, the resulting trade-off depends on the degree of parallelization. Both the P interpolated rotations in (b) and the summation over the T independent Fourier transformed motion transformations in (a) can be done in parallel. In our case, with a moderate number of patches $P \leq 6$ and sufficiently large images, the Fourier approach is computationally beneficial.

5.1.2 Possible computational savings

Let's firstly assume that we have a single global patch and $\mathbf{X}$ is a matrix describing the effect of the forward model, where $\mathbf{X} = \mathbf{A}_{\boldsymbol{\theta}}\mathbf{F}$ is the forward model. We often need to compute $\mathbf{v} = \mathbf{X}^{\mathsf{H}}\mathbf{X}\mathbf{u}$ e.g. in optimization steps, and thus it is important to understand the structure of $\mathbf{X}^{\mathsf{H}}\mathbf{X}$ since it corresponds to the effect of going forward and coming back. We think of $\mathbf{X}$ as being the product of a non-uniform Fourier matrix (for rotation) and a diagonal phase matrix (for translation). Interestingly, $\mathbf{X}^{\mathsf{H}}\mathbf{X}$ does not depend on the translation since $e^{-2\pi i \phi_{\mathbf{k}}}$ and $\overline{e^{-2\pi i \phi_{\mathbf{k}}}} = e^{2\pi i \phi_{\mathbf{k}}}$ cancel each other out. Following ideas in [72], where $\mathbf{k}$ contains arbitrary locations, and $\mathbf{r}$ and $\mathbf{s}$ are Cartesian grids, we have

$$\mathbf{y} = \mathbf{X}\mathbf{u},\ y_{\mathbf{k}} = e^{-2\pi i \phi_{\mathbf{k}}} \frac{1}{\sqrt{n}} \sum_{\mathbf{r}} e^{-2\pi i \mathbf{k}^{\top}\mathbf{r}} u_{\mathbf{r}},\ \mathbf{v} = \mathbf{X}^{\mathsf{H}}\mathbf{X}\mathbf{u}, \tag{5.3}$$

$$v_{\mathbf{s}} = \frac{1}{\sqrt{n}} \sum_{\mathbf{k},\mathbf{r}} e^{-2\pi i \mathbf{k}^\top (\mathbf{r}-\mathbf{s})} u_{\mathbf{r}} = \frac{1}{\sqrt{n}} \sum_{\mathbf{r}} w_{\mathbf{r}-\mathbf{s}} u_{\mathbf{r}}, \text{ where } w_{\boldsymbol{d}} = \frac{1}{\sqrt{n}} \sum_{\mathbf{k}} e^{-2\pi i \mathbf{k}^\top \mathbf{d}}. \quad (5.4)$$

Hence, we have a (non-circular) convolution $\mathbf{v} = \mathbf{w} \star \mathbf{u}$, where the convolution weights can be precomputed using the complex transpose of $\mathbf{X}$ i.e. $\mathbf{w} = \mathbf{X}^{\mathsf{H}}\mathbf{1}$. The convolution can be performed in the Fourier domain as usual by

$$\mathbf{F}\mathbf{v} = \mathbf{F}(\mathbf{w} \star \mathbf{u}) = (\mathbf{F}\mathbf{w}) \odot (\mathbf{F}\mathbf{u}). \quad (5.5)$$

However, the grid size of $\mathbf{w}$ is twice as large as the size of $\mathbf{u}$.

Let's now assume that we have several spatial windows p, and $\mathbf{X} = \sum_{p=1}^{P} \mathbf{A}_{\boldsymbol{\theta}_p} \mathbf{F} \mathrm{dg}(\tilde{\mathbf{w}}_p) = \sum_{p=1}^{P} \mathbf{A}_{\boldsymbol{\theta}_p} \mathbf{W}_p \mathbf{F}$ is the forward model. The question is whether we can benefit from an efficient implementation of an MVM with $\mathbf{Z}_{pq} = \mathbf{A}_{\boldsymbol{\theta}_p}^{\mathsf{H}} \mathbf{A}_{\boldsymbol{\theta}_q}$. Let us first expand the sum

$$\mathbf{X}^{\mathsf{H}}\mathbf{X} = \mathbf{F}^{\mathsf{H}} \left(\sum_{p,q=1}^{P} \mathbf{W}_p \mathbf{Z}_{pq} \mathbf{W}_q \right) \mathbf{F}. \quad (5.6)$$

We can approximate $\mathbf{X}^{\mathsf{H}}\mathbf{X}\mathbf{u} \approx \mathbf{F}^{\mathsf{H}} \left(\sum_{p=1}^{P} \mathbf{W}_p \mathbf{Z}_{pp} \mathbf{W}_p \right) \mathbf{F}\mathbf{u}$ i.e. drop the off-diagonal terms and ask when this expression is accurate. Intuitively speaking, this seems to be the case when $p \neq q$ and either the motion trajectories $\boldsymbol{\theta}_p$ and $\boldsymbol{\theta}_q$ do not let the patches overlap or the overlapping region has zero intensity in the true image $\mathbf{u}$.

The case where $\mathbf{A}_{\boldsymbol{\theta}_p}$ contains purely translational motion is very interesting since $\mathbf{A}_{\boldsymbol{\theta}_p} = \mathrm{dg}(\mathbf{a}_{\boldsymbol{\theta}_p})$ and $\mathbf{A}_{\boldsymbol{\theta}_p}^{\mathsf{H}} \mathbf{A}_{\boldsymbol{\theta}_p} = \mathbf{I}$ since

$$\mathbf{y} = \sum_{p=1}^{P} \mathbf{a}_{\boldsymbol{\theta}_p} \odot (\mathbf{F}(\tilde{\mathbf{w}}_p \odot \mathbf{u})) + \boldsymbol{\varepsilon} = \sum_{p=1}^{P} \mathbf{a}_{\boldsymbol{\theta}_p} \odot (\mathbf{W}_p \mathbf{v}) + \boldsymbol{\varepsilon} = \sum_{p=1}^{P} \mathbf{a}_{\boldsymbol{\theta}_p} \odot (\mathbf{w}_p \star \mathbf{v}) + \boldsymbol{\varepsilon}, \mathbf{v} = \mathbf{F}\mathbf{u}. \quad (5.7)$$

The question is whether we can compute the MVM $\mathbf{v} \mapsto \mathbf{A}_{\boldsymbol{\theta}}\mathbf{v}$, $\mathbf{A}_{\boldsymbol{\theta}} = \sum_{p=1}^{P} \mathrm{dg}(\mathbf{a}_{\boldsymbol{\theta}_p}) \mathbf{W}_p$ faster by evaluating the sum beforehand or as we go along.

5.2 Blind non-rigid motion correction optimization problem

In order to recover both unknown motion parameters $\mathbf{\Theta}$ and the underlying sharp image $\mathbf{u}$ from the motion corrupted k-space measurement $\mathbf{y}$, we need to formalize the trade-off between *data fidelity,* i.e. how well the data $\mathbf{y}$ is reproduced by the forward operator $\mathbf{A}_{\mathbf{\Theta}}$ and *regularization* of both $\mathbf{u}$ and $\mathbf{\Theta}$. Regularization here has two meanings: the regularizer for $\mathbf{u}$ measures the image quality, and the regularizer for $\mathbf{\Theta}$ softly constrains the vast space of motion parameters, so that only a tiny relevant subspace needs to be considered. Our objective function is

$$\psi(\mathbf{u}, \mathbf{\Theta}) = \phi(\mathbf{u}) + \lambda \cdot \|\mathbf{r}_{\mathbf{\Theta}}\|^2 + \mu \cdot \sum_{p=1}^{P} \mathrm{tr}(\mathbf{\Theta}_p \mathbf{D}_t^\top \mathbf{D}_t \mathbf{\Theta}_p^\top), \text{ with } \mathbf{r}_{\mathbf{\Theta}} = \mathbf{A}_{\mathbf{\Theta}}\mathbf{F}\mathbf{u} - \mathbf{y} \quad (5.8)$$

where $\lambda \in \mathbb{R}_+$ and $\mu \in \mathbb{R}_+$ are scalar weights for the data fidelity term and the motion trajectory regularizer, respectively. We measure data fidelity by the squared norm of the k-space residual $\mathbf{r}_{\mathbf{\Theta}}$, which corresponds to the squared error. Regularization of the motion parameters is achieved by requiring small temporal derivatives independently for each of the P motion trajectories where $\mathbf{D}_t \in \{0, \pm 1\}^{T \times T}$ is a discrete temporal derivative matrix and $\mathrm{tr}(\cdot)$ is the trace operator.

Using the full objective (Equation 5.8) we can correct for both strong rotations and multi-rigid motion. In this work, we use the sum of the gradient's absolute values $\phi(\mathbf{u}) = \|\mathbf{G}\mathbf{u}\|_1$, with $\mathbf{G} = [\mathbf{D}_x, \mathbf{D}_y, \mathbf{D}_z] \in \{0, \pm 1\}^{3N \times N}$, and $\|\mathbf{v}\|_1 = \sum_{i=1}^{N} |v_i|$ as an image regularizer. We use this l_1-norm of the gradient image metric for two reasons:

1. optimization of (Equation 5.8) w.r.t. $\mathbf{u}$ is a convex penalized least squares problem with a single global optimum.
2. the metric is widely used and compares well to other non-convex metrics in empirical studies [33].

For a discussion of the connection between the previously used inverse-based approach and Equation 5.8, see the Appendix.

5.2.1 Connection to inverse-based approach

Assuming $P = 1$, $\mu = 0$, $\mathbf{A}_{\mathbf{\Theta}}$ to be invertible and its inverse being a valid motion matrix $\mathbf{A}_{\mathbf{\Theta}'} := \mathbf{A}_{\mathbf{\Theta}}^{-1}$, and substituting $\mathbf{u} = \mathbf{F}^{\mathsf{H}}\mathbf{A}_{\mathbf{\Theta}'}\mathbf{z}$, we can rewrite the objective

$\psi(\mathbf{z}, \boldsymbol{\Theta}) = \phi(\mathbf{F}^{\mathsf{H}}\mathbf{A}_{\boldsymbol{\Theta}'}\mathbf{z}) + \lambda \cdot \|\mathbf{z} - \mathbf{y}\|^2$, where $\mathbf{z}$ is an auxiliary variable. For large[1] λ we obtain $\mathbf{y} = \mathbf{z}$ and hence $\psi(\boldsymbol{\Theta}) = \phi(\mathbf{F}^{\mathsf{H}}\mathbf{A}_{\boldsymbol{\Theta}'}\mathbf{y})$, which is the direct inverse objective function we used in our previous work – see Chapter 4. Although this allows for fast motion correction, the inverse-based objective clearly has a number of limitations, i.e. it does not allow for valid reconstruction in cases where rotational motion is stronger than a few degrees.

5.3 Solving non-rigid motion optimization problem

Using the objective function (Equation 5.8), we want to recover both the unknown sharp image $\mathbf{u}$ and the motion parameters $\boldsymbol{\Theta}$ simultaneously. Let us define the mapping ψ_* obtained by evaluating the objective ψ at the optimal image $\mathbf{u}_*$ by

$$\psi_*(\boldsymbol{\Theta}) = \psi(\mathbf{u}_*, \boldsymbol{\Theta}), \text{ with } \mathbf{u}_* = \arg\min_{\mathbf{u}} \psi(\mathbf{u}, \boldsymbol{\Theta}). \tag{5.9}$$

Empirically, small changes in $\boldsymbol{\Theta}$ can lead to big changes in the optimal value $\mathbf{u}_*$. Also, optimization w.r.t. $\mathbf{u}$ is convex and simple, whereas ψ is a non-convex function in $\boldsymbol{\Theta}$. This suggests the special structure of an optimization scheme, which we detail in Algorithm 5.1. The purpose of the *annealing loop* is to progressively relax the importance of the regularization terms in both the image $\mathbf{u}$ and motion parameters $\boldsymbol{\Theta}$ by increasing the weight of the data fidelity term. In the *motion loop,* ψ_* is optimized w.r.t. motion parameters $\boldsymbol{\Theta}$; here we use the fact that the forward model allows the motion parameter gradient to be computed in closed form. The most inner *image loop* involves computation of the current best estimate of the image $\mathbf{u}_*$ obtained by optimizing the convex objective function (Equation 5.8). Additionally, for strong motions, a multiscale approach is used, which is implemented by the outermost *multiscale loop*. The idea is to first solve a simpler low-resolution problem and then use the estimated motion parameters as initializations for higher resolution scales. This has an effect of avoiding bad local minima, which the optimizer is prone to become stuck in given the non-linear objective function. To carry out the optimization, we use the LBFGS package [66].

The literature on blind deconvolution [73, 74] suggests that it is beneficial to compute the data fidelity term in the gradient domain for blur kernel estimation

[1] From the theory of regularization paths we know, that we do not need to drive λ to infinity but there exists a threshold beyond which the optimizer will always obey $\mathbf{z} = \mathbf{y}$.

Algorithm 5.1 Nested-loop optimization algorithm.

Input: $\mathbf{y} \in \mathbb{C}^N$, $N{=}n_x \cdot n_y \cdot n_z$ **// Corrupted volume**
$\mathbf{s} = [s_1, s_2, ..., s_K]$ **// Multiscale schedule**
$\boldsymbol{\lambda} = [\lambda_1, \lambda_2, ..., \lambda_L]$ **// Annealing schedule**
$N_{\boldsymbol{\Theta}}$ **// Number of motion estimation steps**
$N_{\mathbf{u}}$ **// Number of image estimation steps**
Output: $\mathbf{u}_*$ **// Restored image**

Start from zero motion $\boldsymbol{\Theta} \leftarrow \mathbf{0}$.
For $\mathbf{s} = [s_1, s_2, ..., s_K]$ **do** **// Multiscale loop**
  Use only k-space center of the raw data.
  For $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, ..., \lambda_L]$ **do** **// Data fidelity annealing loop**
    For $\ell = 1, ..., N_{\boldsymbol{\Theta}}$ **do** **// Motion estimation loop**
      Start from zero image $\mathbf{u} \leftarrow \mathbf{0}$. **// Image estimation loop**
      Find $\mathbf{u}_* \leftarrow \arg\min_{\mathbf{u}} \psi(\mathbf{u}, \boldsymbol{\Theta})$ using $N_{\mathbf{u}}$ LBFGS steps.
      Compute the motion gradient $\nabla\psi_* \leftarrow \partial\psi(\mathbf{u}_*, \boldsymbol{\Theta})/\partial\boldsymbol{\Theta}$.
      Make a conjugate gradient step along $\nabla\psi_*$.
    End
  End
  Initialize central frequency part of $\boldsymbol{\Theta}$ on next finer scale.
End

(which corresponds to our motion estimation loop in Algorithm 5.1). We also adopt this approach and use a data fidelity of the form $\left\|\mathbf{G}\mathbf{F}^{\mathsf{H}}\mathbf{r}_{\boldsymbol{\Theta}}\right\|^2$ in the motion estimation loop of Algorithm 5.1. Here $\mathbf{r}_{\boldsymbol{\Theta}} = \mathbf{A}_{\boldsymbol{\Theta}}\mathbf{F}\mathbf{u} - \mathbf{y}$ is the k-space residual and $\mathbf{G} = [\mathbf{D}_x, \mathbf{D}_y, \mathbf{D}_z] \in \{0, \pm 1\}^{3N \times N}$ computes the gradient using finite differences. Since the gradient computation operation is diagonal in k-space and the Fourier matrix $\mathbf{F}$ is orthonormal (leaving the norm invariant), the data fidelity term becomes the weighted sum $\|\mathbf{d}_x \odot \mathbf{r}_{\boldsymbol{\Theta}}\|^2 + \|\mathbf{d}_y \odot \mathbf{r}_{\boldsymbol{\Theta}}\|^2 + \|\mathbf{d}_z \odot \mathbf{r}_{\boldsymbol{\Theta}}\|^2$ instead of the plain norm $\|\mathbf{r}_{\boldsymbol{\Theta}}\|^2$ in Equation 5.8. The effect of the reweighting is similar to a preconditioner in a linear system: high frequencies are magnified and low frequencies receive less emphasis. This is in line with an intuition that motion can be best estimated from the localized high-frequency structures such as edges and cusps. Without the reweighting, due to the second order structure of natural (and also medical) images, - the norm $\|\mathbf{r}_{\boldsymbol{\Theta}}\|^2$ would be dominated by the low frequencies. The reweighting is not a prerequisite for our method to work, but the reconstruction results for 3D volumes are better and the optimization is faster.

In case of pure translations, the matrix $\mathbf{A}_{\boldsymbol{\Theta}_p}$ is diagonal with the elements given by the Fourier translation identity $\mathcal{F}[g(x+a)] = \int_{-\infty}^{+\infty} g(x)e^{-2\pi ik(x+a)}dx = \mathcal{F}[g(x)] \cdot e^{-2\pi ika}$. For rotational motion, the matrix $\mathbf{A}_{\boldsymbol{\Theta}_p}$ has non-zero off-diagonal elements, which correspond to interpolation coefficients used for regridding. The number of non-zero elements in each row depends on the number of neighbors used for interpolation. The interpolation quality can be increased by using more neighbors, and also by using an oversampled grid, which can be obtained by zero-padding in the spatial domain. In our implementation, we are using 2-fold oversampling and 4^D neighboring points, where D is the number of dimensions.

5.4 Computational complexity

Figure 5.2 compares computation times for different setups. The simplest class of problems involves rigid bodies affected by weak (translation $<$ 3mm, rotation $<$ 4°) but commonly encountered motion. Such problems can be addressed with our fast inverse-based approach which only requires a single loop for estimation of motion parameters. Going beyond this regime (multi-rigid mode and/or strong motion) requires additional computational work. The Fourier transforms still constitute the bottleneck; however, we now perform optimization in several nested loops, which increases the computation time by a factor depending on the number of outer loop iterations. Thus, there is a trade-off between the computation time and an achievable level of complexity of the motion correction problem to be solved. Still, using GPUs the computation times even for complicated classes of problems are tolerable (significantly below one hour), and will improve due to the rapid advances in modern GPU technology.

It is important to note that the matrix $\mathbf{A}_{\boldsymbol{\Theta}_p}$ never needs to be constructed explicitly, an operation that can easily become memory prohibitive when dealing with large data volumes. It is, however, necessary to implement the matrix vector operations: multiplications with $\mathbf{A}_{\boldsymbol{\Theta}_p}$, its transpose and derivative with respect to motion parameters. Computationally, these operations amount to point-wise multiplications with translational phase ramps and interpolations to estimate the values on rotation-deformed grids. Since these operations involve local multiplications and additions in frequency domain they can be efficiently handled by computational architectures capable of massively parallel computations. We use graphic processing units (GPUs) for this purpose, and implement our method in CUDA C.

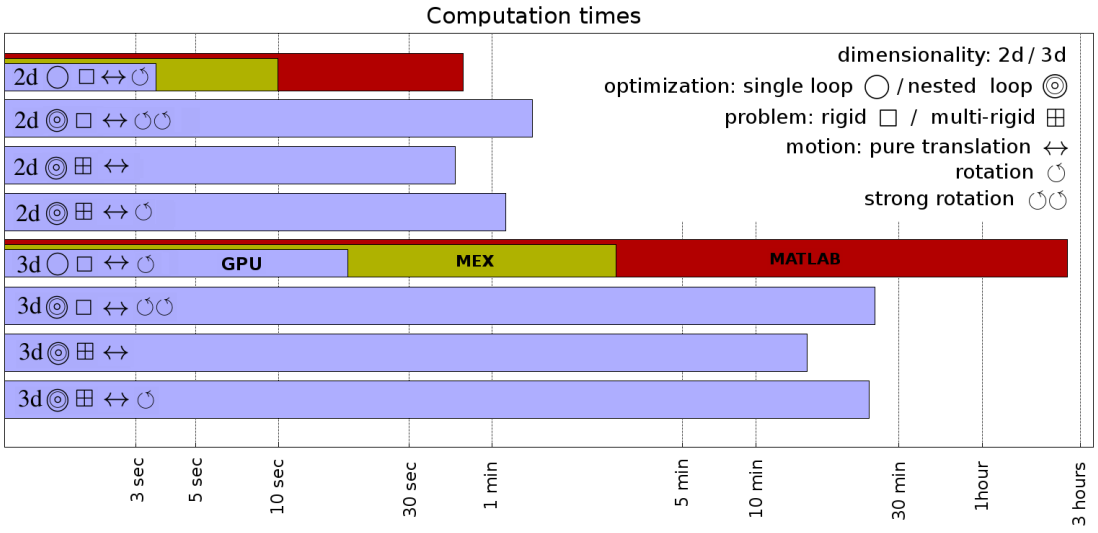

Figure 5.2: Computation times (in log scale) for different motion correction problem classes. The problems differ in the dimensionality (2D, 192×192 versus 3D, $192 \times 192 \times 16$) of the raw data, the kind of motion (global rigid motion $\square$ versus multi-rigid motion $\boxplus$) and the type of motion (translation $\leftrightarrow$, rotation $\circlearrowleft$, strong rotation $\circlearrowleft\circlearrowleft$). The different problem classes require different flavors of the algorithm (single-loop $\bigcirc$ direct inverse objective versus nested-loop $\circledcirc$ full objective optimization). For some (feasible) problems we also show CPU computation times for C(MEX), and MATLAB implementations.

It is not only the operations with $\mathbf{A}_{\boldsymbol{\Theta}_p}$ matrix which can benefit from a parallel computation. In fact, the objective function (Equation 5.8) involves further ingredients that are subject to parallelization such as computation of spatial gradients, fast Fourier transforms and image metric evaluation. To have the maximum computational gain, we implement and run the entire pipeline of Algorithm 5.1 on GPU. In particular, this allows for minimization of the memory transactions from the graphic card to the main memory, which can be a bottleneck when transferring large data volumes.

The most computationally active part of the algorithm is the inner image estimation block which is iterated $LN_{\boldsymbol{\Theta}}$ times, and itself involves repetitive application of the motion operator. Since the motion parameters do not change within this block, this allows both translational phase coefficients and interpolation weights to be precomputed before starting the block; they can then be stored in GPU memory. This makes the image estimation loop computationally cheap, and the entire algorithm operable in practical time.

5.5 Experiments

Imaging experiments were performed on a 3 T scanner (Magnetom Trio, Siemens Healthcare, Erlangen, Germany), using single-channel volume coils designed for wrist examinations. In a non-human set of experiments, we imaged kiwi fruits, which feature high-contrast fine-detailed regular internal structure, and are thus very helpful in assessing the ability of the correction algorithm to improve visual quality. To induce the motion, we used a special MR-compatible actuator. 2D kiwi data had a matrix size of 384x384 over a Field-of-view (FOV) of 150x150 mm^2 with a slice thickness of 3 mm. FLASH images with an echo time (TE) of 5.9 ms, a repetition time (TR) of 250 ms and a flip angle of 50° were acquired within 1.5 min.

In vivo images were acquired with the approval of the local ethics board. The hand of a human subject was imaged; the subject was told to move an index finger while trying to keep the rest of the hand stationary. A standard FLASH sequence was used in both two and three dimensions. 2D images with a matrix size of 384x384 over a FOV of 180x180 mm^2, a TE of 5.9 ms, a TR of 250 ms and a flip angle of 50° were acquired within one minute. 2D images with a multi-finger motion were acquired within two minutes, with a matrix size of 320x320 over a FOV of 256x256 mm^2, a TE of 7.1 ms, a TR of 500 ms,a flip angle of 60°, and a slice thickness of 3 mm. 3D wrist data had a matrix size of 256x256x24 over a FOV of 180x180x36 mm^3, a TE of 3.1 ms, a TR of 20 ms and a flip angle of 20° and were acquired within 2 min.

5.5.1 Results: simulated data

We applied our method to a set of simulated datasets to study the properties of the algorithm. Firstly, we considered a rigid motion problem with a single global patch – see Figure 5.3. Strong rotational motion (in the range of 20°) of sinusoidal form is applied to the 2D clean image of an isolated monkey brain. An analysis of the reconstruction results leads to the following conclusions: firstly, our new method passes the sanity check by being able to correct for basic rigid motion. Secondly, since strong rotation leads to severe undersampling of the spectrum, it cannot be corrected by empirical inversion. Thus, our GradMC method described in the previous chapter, and even inversion with ground-truth parameters, fails. However, our new method, based on the forward model and the image regularizer, is able to significantly improve image quality. The reconstructed image is oversmoothed, which

happens because of severe loss of high-frequency information due to rotations. Finally, it can be observed that the pose of the reconstructed object is different from ground truth, which is due to the invariance of our objective function to global (all-view) rotations and translations.

Next, we simulate motion-corrupted datasets, where we know both the ground-truth image and motion trajectory. In the first experiment (see Figure 5.4), we examined the convergence properties of the nested-loop Algorithm 5.1. We evaluated the correction progress by tracking the error in the estimated image and trajectory on each outer loop iteration. As a ground truth we used a 2D image (size of 192x192) of the monkey brain and set the motion trajectory to a global (rigid, single patch) rotation of pure sine form with an amplitude of 2 degrees. In the center of Figure 5.4 we show the error (absolute value of the difference) in the estimated rotational motion parameters computed w.r.t. ground truth for each iteration of the algorithm. The optimization time arrow is pointing from the top to the bottom of Figure 5.4, and the improvement (color-coded from red to blue) towards the end of optimization can be seen. Periodically occurring low-error motion estimates (best seen in the very first iterations) are due to the fact that motion parameters were initialized to zeros; thus for zero-crossings of the sine trajectory they will be initially valid. We used two annealing steps $\lambda = [2.5, 17]$, which split the optimization process into two epochs (each epoch takes 50 iterations). In the first epoch (top half of Figure 5.4), the algorithm gradually improved the motion parameter estimates corresponding to low-frequency excitations. This is because the heavily weighted regularization term eliminates high-frequency information in the image. During the second epoch, less weight was put into the regularization term, and the algorithm was able to recover high-frequency motion parameters.

The effects of the two epochs are easily seen in the recovered images. In the left column of Figure 5.4, from top to bottom, we show the motion corrupted image, the image estimated after the first epoch (notice oversmoothing due to strong regularization) and the image recovered after the second epoch. Additionally, we computed the normalized root mean-squared error (NRMSE) between estimated images and ground truth, and plotted it in the right part of Figure 5.4.

Our next experiment (see Figure 5.5) probes the separability property that we introduced in the Methods section. Firstly we studied a non-separable case and split the image into two patches (blue, green) in phase encode direction. We used our multi-rigid forward model to apply the translational motion (uniformly random bi-

directional motion in a range of four pixels) to the blue patch. Since ghosts propagate globally (from the nearby blue patch) one can see the decrease of image quality in the green patch, although it is not directly affected by motion. We then used spatial windows to extract the contents from each patch and correct them separately using the rigid motion inverse-based approach described in the previous chapter. No improvement in image quality can be observed. In the separable case, the patches were arranged in frequency encode direction. The same motion trajectory as used in the non-separable experiment was applied to the blue patch. The green patch was not affected by motion. In contrast to the non-separable case, a great improvement of image quality is now evident.

Finally, we tested the multi-rigid motion correction capabilities of our method. We simulated the global rotational motion in the image of a monkey brain. We split the image into multiple patches of the same size, and attempted to carry out motion correction by assuming pure translational motion in each patch (see Figure 5.6). We varied the number of patches from one single global patch to a dense splitting of 8 by 8 patches. The MSE of the reconstruction w.r.t. the ground-truth image is computed for each splitting level. The results show that the optimum w.r.t. MSE corresponds to a moderate level of splitting density (4 by 4 patches). Our interpretation for this is that, while too few patches are clearly not enough to approximate the global rotation by multiple translations, too many patches may excessively increase the complexity of the problem.

5.5.2 Results: real data

In a similar manner to that used in the simulated data experiments, we firstly considered the global single-patch problem involving strong rotational motion. Figure 5.7 demonstrates the ability of our method to correct real data affected by strong rigid motion, and its superiority over our previous inverse-based method. We used $N_{\boldsymbol{\Theta}} = N_{\mathbf{u}} = 50$ for motion/image estimation loops and the annealing schedule of $\boldsymbol{\lambda} = [e^1, e^2, e^3]$.

For our next experiments, we used a wrist coil to acquire images of kiwi fruits and a hand of a human subject. Note that the spatial windows (indicated by color-coded overlays) were arranged in phase-encode direction leading to hard, non-separable problems. For the first real data experiment (see Figure 5.8) we acquired the image of two kiwi fruits. During the acquisition we used an MR-neutral actuator to displace

the left kiwi fruit. While carrying out the motion correction, we split the image into two patches, and used our multi-rigid nested-loop algorithm. To evaluate the quality of reconstructed images we additionally acquired a motion-free ground-truth image. As can be seen in Figure 5.8, the algorithm was able to remove strong artifacts, and reconstructed results look very similar to ground truth.

In the next experiment we acquired in vivo data of the human hand (see Figure 5.9). The subject was asked to move the index finger and keep the wrist stationary. In this case as well, with a more biologically plausible motion, our algorithm was able to significantly improve image quality. Additionally, in the bottom section of the figures we plot the motion trajectories recovered by the algorithm. The motion parameters estimated in the left kiwi and index finger patches indicate a strong motion. Please note, however, that the motion trajectories recovered in non-motion patches also contain significant variation. This is due to the fact that blind autofocusing motion correction approaches suffer from an ambiguity in the recovered motion parameters which does not influence the image quality of the recovered image. Additionally, if strong regularization of the image is used, this has an effect of penalizing high-frequency information and means that the recovered image is tolerant to variation in high-frequency excitation motion parameters.

Next, we tested our algorithm on a more complicated problem, which involved multiple patches (Figure 5.10). The subject was moving all fingers simultaneously trying to displace them in a disorganized manner, so that each finger has a unique motion trajectory. Also in this case, major improvements in image quality were possible. Carrying out the reconstructions we used the following algorithm parameters: $N_{\boldsymbol{\Theta}} = N_{\mathbf{u}} = 50$, $\boldsymbol{\lambda} = [e^1, e^2]$.

Finally, we tested our method on a 3D dataset, which was also corrupted by multi-rigid motion. In Figure 5.11 we show four slices from the motion-corrupted, corrected and motion-free volumes. A significant improvement in image quality can be observed in the reconstructed result. We used the following algorithm parameters: $N_{\boldsymbol{\Theta}} = 100$, $N_{\mathbf{u}} = 50$, $\boldsymbol{\lambda} = [e^1, e^2]$, and three multiscale iterations.

5.6 Summary

We have, thus, experimentally demonstrated the possibility of correction for multi-rigid motion using a fully blind retrospective method. We assume the patch masks

to be given as input (although in theory they could be obtained by discrete optimization) either by the user or an automatic process like a coarse segmentation of the 3D image. Of course, this means that before using our algorithm the patches need to be provided (either by their coordinates, or by a graphical interface). This is a difficult problem in itself if the multi-rigid motion is complex, or if the observation is covered by strong artifacts. In a practical setting, e.g. knee or hand imaging, the patch splitting can be obtained by segmentation or atlas registration so that the degrees of freedom of the imaged body parts are reflected in the parts of the segmentation. Another way to avoid the need for user-created masks is through the use of regular patch-splitting of the image. With increasing patch density, any non-rigid motion can be described, an extreme case being the vector field over all voxels. At this extreme the problem is, of course, computationally very hard to solve; furthermore, it is severely underconstrained given the number of parameters to optimize. Further research will go into the direction of coupling the motion parameters over patches in order to reduce the effective number of parameters to optimize. In particular, upgrading our algorithm to use an adaptive patch splitting [44] looks like a reasonable next step.

Compared to the inverse-based approach described in the previous chapter, the method described above is computationally more demanding. Modern graphics cards now render retrospective motion correction problems, that were considered almost intractable years ago, solvable in a reasonable time. Although, in our current implementation we are already experiencing massive time savings due to parallelization, there is still room for improvement by code optimization. For example, when performing line searches in the objective landscape, one could avoid computing the gradient direction at each step.

In contrast with the inverse-based motion correction framework described in the previous chapter, our new algorithm has more parameters to tune, and cannot simply be used out of the box. The tuning is a price to pay for being able to solve the more complex multi-rigid problem. In our current research, we faithfully report all the parameters that we were using in our experiments. Future research might allow for simplification of the choice of parameters.

We acknowledge that, in this study, the validity of our approach was demonstrated through somewhat artificial instances of multi-rigid motion. Our future goal is to address the problem of motion correction in abdomen data sets, i.e. images of livers or kidneys, which are affected by breathing. This would be challenging for many

reasons: firstly, motion due to breathing is rather strong, and secondly, the task of splitting the abdomen data into patches is non-trivial. Usually there are many different tissues in the field of view, and they are likely to possess different modes of motion, i.e. spine and walls of the abdomen are likely not to be moving, while the internal organs are affected by translations and possibly small rotations.

While in the current study we assumed that the data was acquired with a single coil, this may be a more difficult scenario compared to multi-rigid motion correction of multi-coil data. This is due to the fact that spatial sensitivity profiles of the coils naturally disentangle the effects (artifacts) of non-rigid motion over the spatial domain that otherwise propagate globally in phase encode direction. Recent non-rigid motion correction methods (see Introduction) operate on data from multiple coils and should strongly benefit from this property. The bigger challenge is that, in practice, the input to the reconstruction algorithm would likely be a magnitude-only DICOM image. Thus, not only the information from each coil is not accessible, but the spatial phase in the image is gone. To our knowledge, there is no method for doing retrospective motion correction of both rigid and non-rigid motion based on pure magnitude data.

Our main contribution is the retrospective method capable of correcting a multi-rigid motion that does not require external information on motion as a reference. At the core of our new method is an analytic non-linear optimization of objective function of both the image and motion subject to fidelity and regularization constraints. We have demonstrated the ability of the method to correct for motion on both simulated and real data sets.

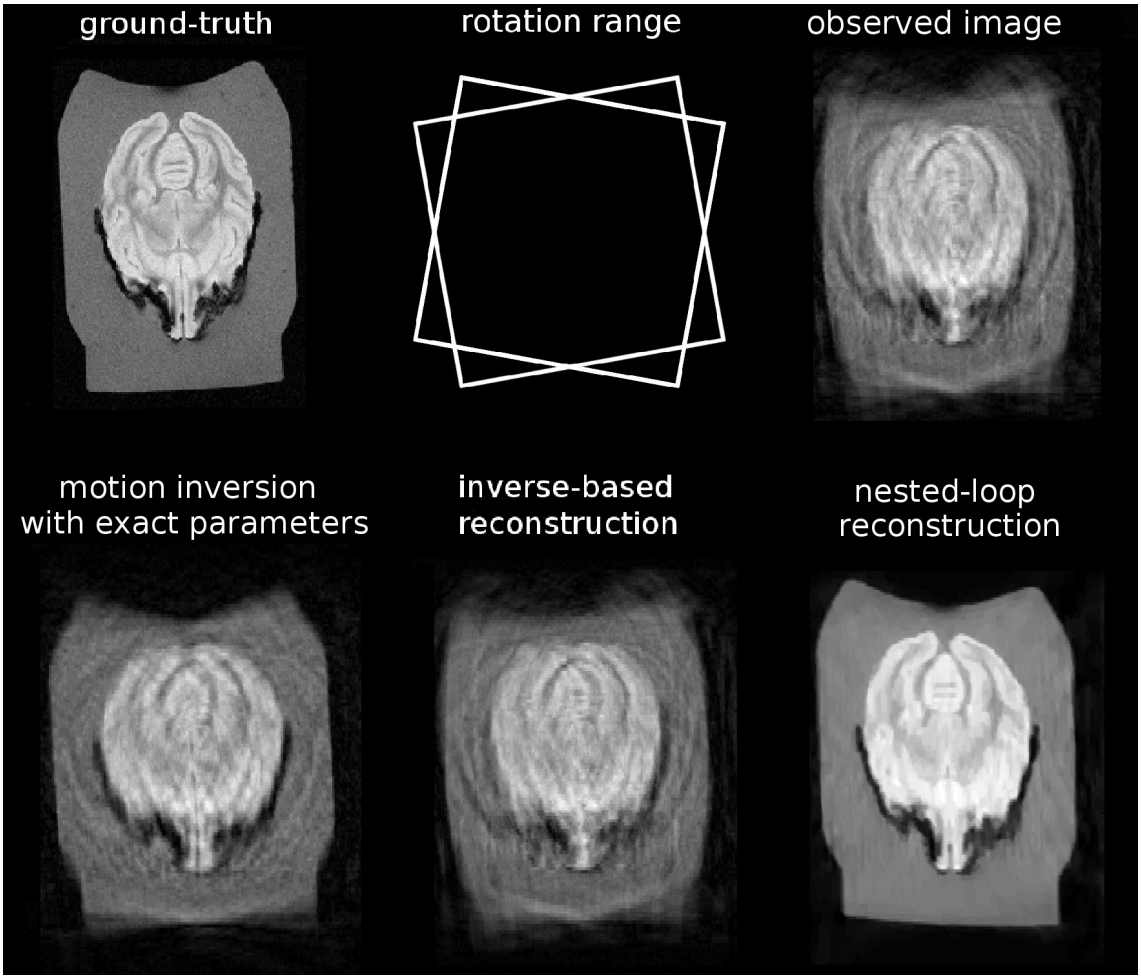

Figure 5.3: Correction of strong rotational motion (in the range of 20°). We compare the reconstruction with our new nested-loop algorithm against our previous inverse-based approach GradMC. In the top row, we show the original image together with the motion-corrupted observation. The amplitude of periodic rotations from -10° to + 10° is indicated by white frames. In the bottom row, we show (left) the empirical inversion with ground truth rotational parameters, (middle) the restoration with the use of our previous inverse-based method, and (right) the reconstruction results obtained using our new method based on the forward model.

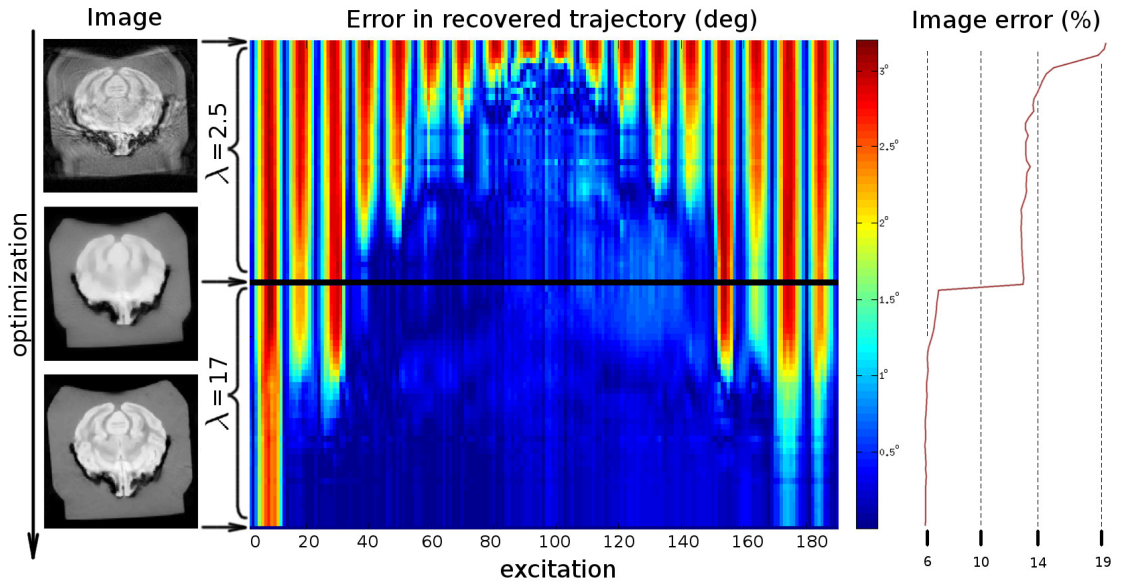

Figure 5.4: Analysis of the motion correction process w.r.t. the error decrease in the reconstructed image and motion trajectory. The optimization time line is going from top to bottom. The regularization parameter annealing scheme consists of two epochs delimited with a black line. Left: the images before, in the middle (after the first annealing epoch) and after correction. Center: the change in the absolute value of the difference between estimated motion parameters and ground truth; the color-coding (red to blue) indicates the error decrease. Right: normalized RMSE in reconstructed image computed w.r.t. ground truth.

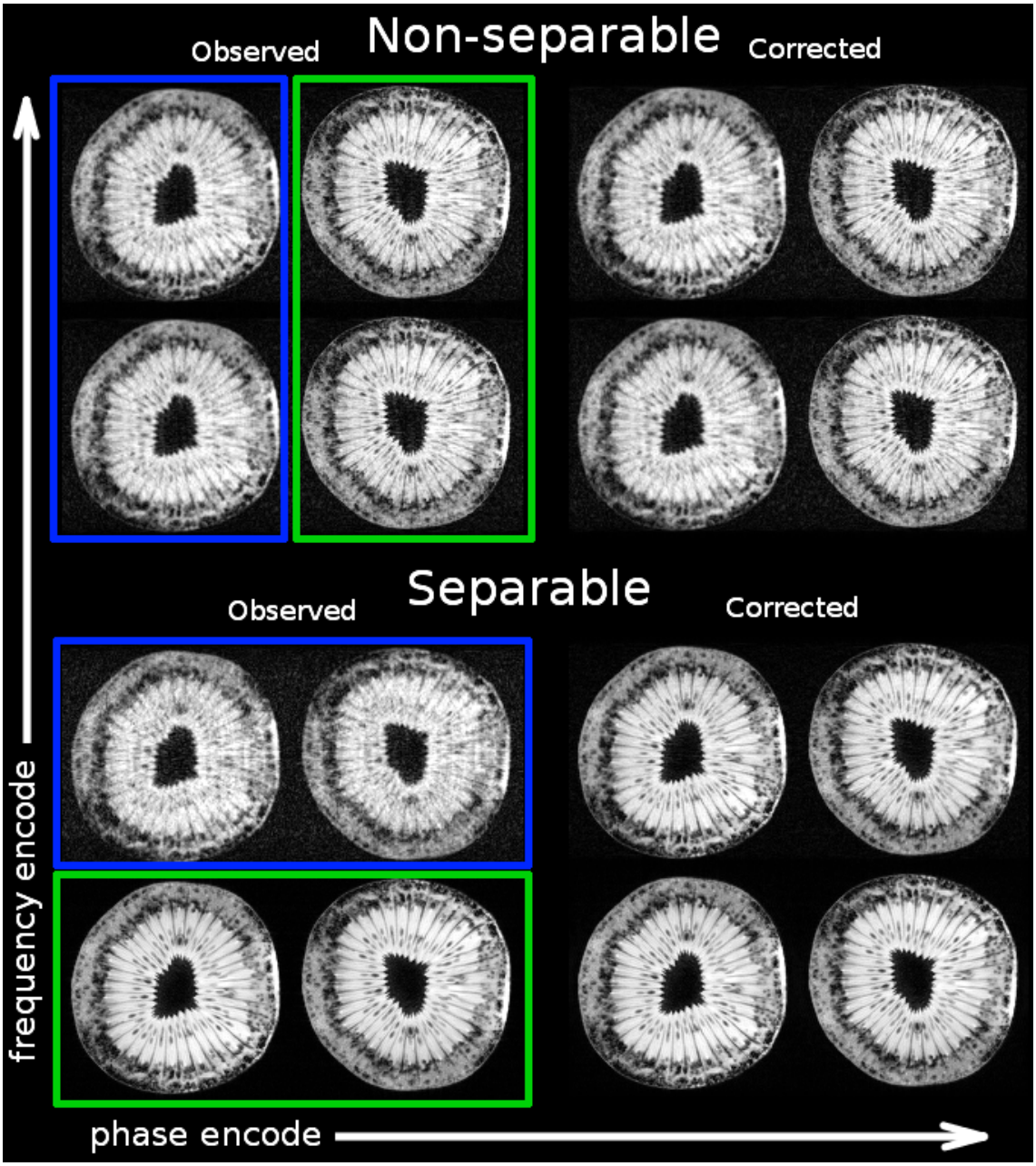

Figure 5.5: Rigid motion correction in the case of non-separable (top) and separable (bottom) problems. In both cases the same random and bi-directional translational motion was applied to patch enclosed in a blue frame. No motion was applied to green-frame patch. The left part of the figure shows the simulated motion-affected images, and the right part shows the results of motion correction.

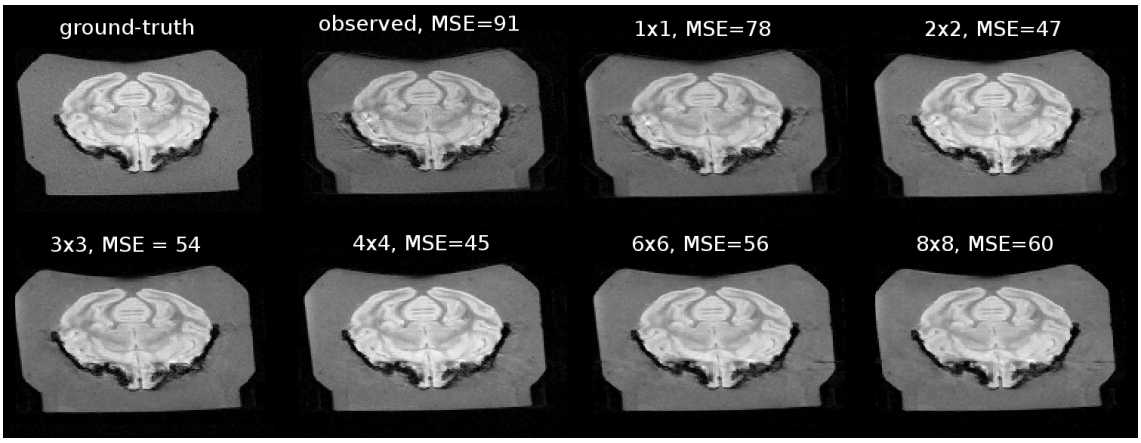

Figure 5.6: Correction of the image affected by global rotational motion by pure translations. Reconstruction is carried out on different patch splitting levels ranging from one global patch (rigid motion) to 8x8 patches splitting. Mean squared error (MSE) is calculated for reconstructions corresponding to different splitting levels.

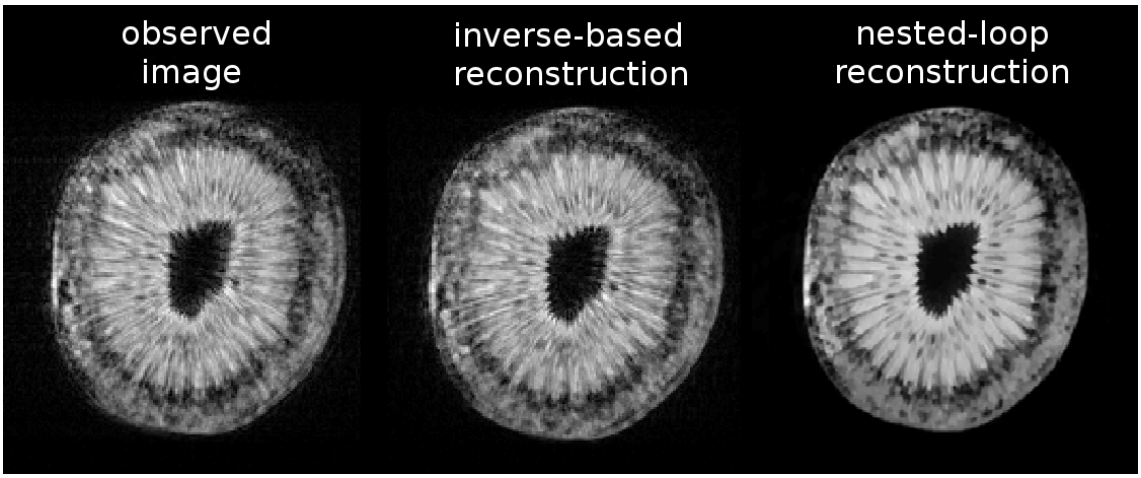

Figure 5.7: Real data. Left to right: image of a kiwi corrupted by rotational motion; restoration using the inverse-based objective; correction using the nested-loop algorithm

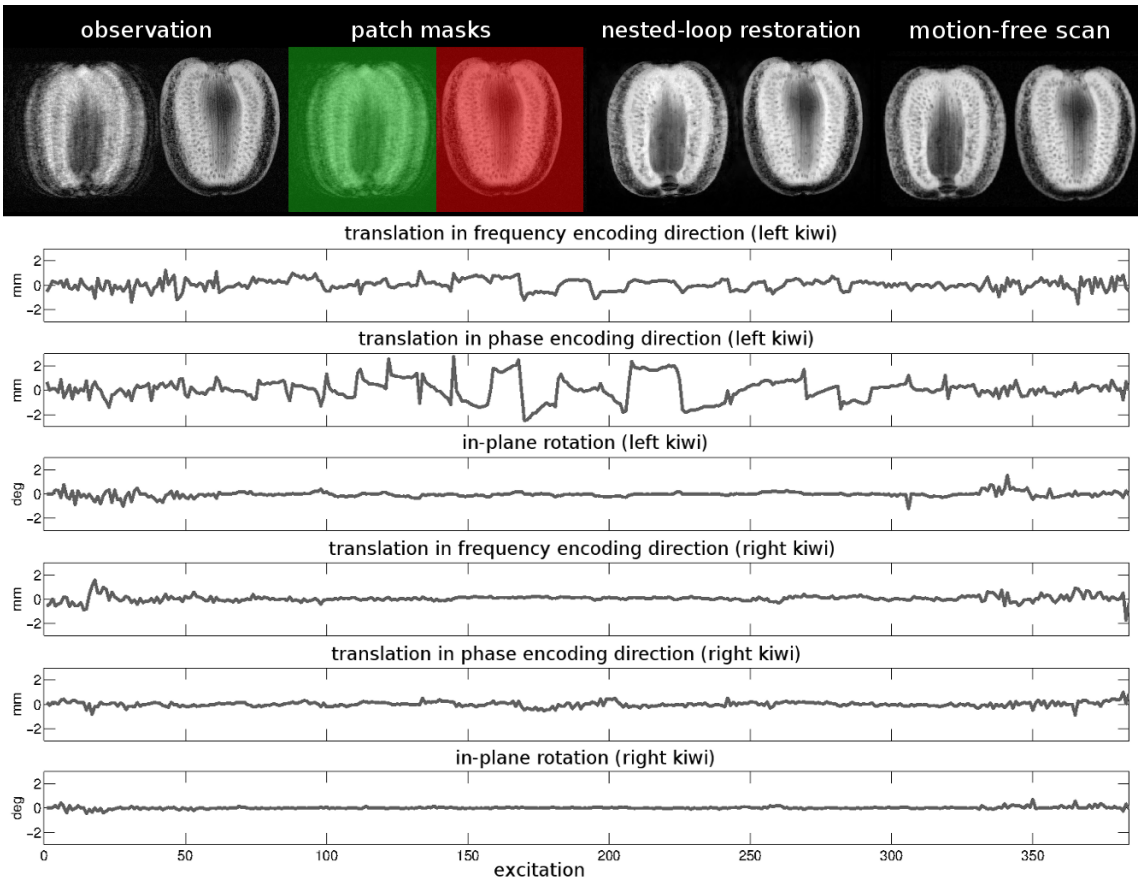

Figure 5.8: Real data. Multi-rigid motion correction was performed on the image split into two patches (indicated by red/green overlays). Top: from left to right we show the observed image (of kiwi fruits), the spatial masks used, motion correction result and ground-truth image from the motion-free scan. Bottom: the trajectories for both patches recovered by the algorithm.

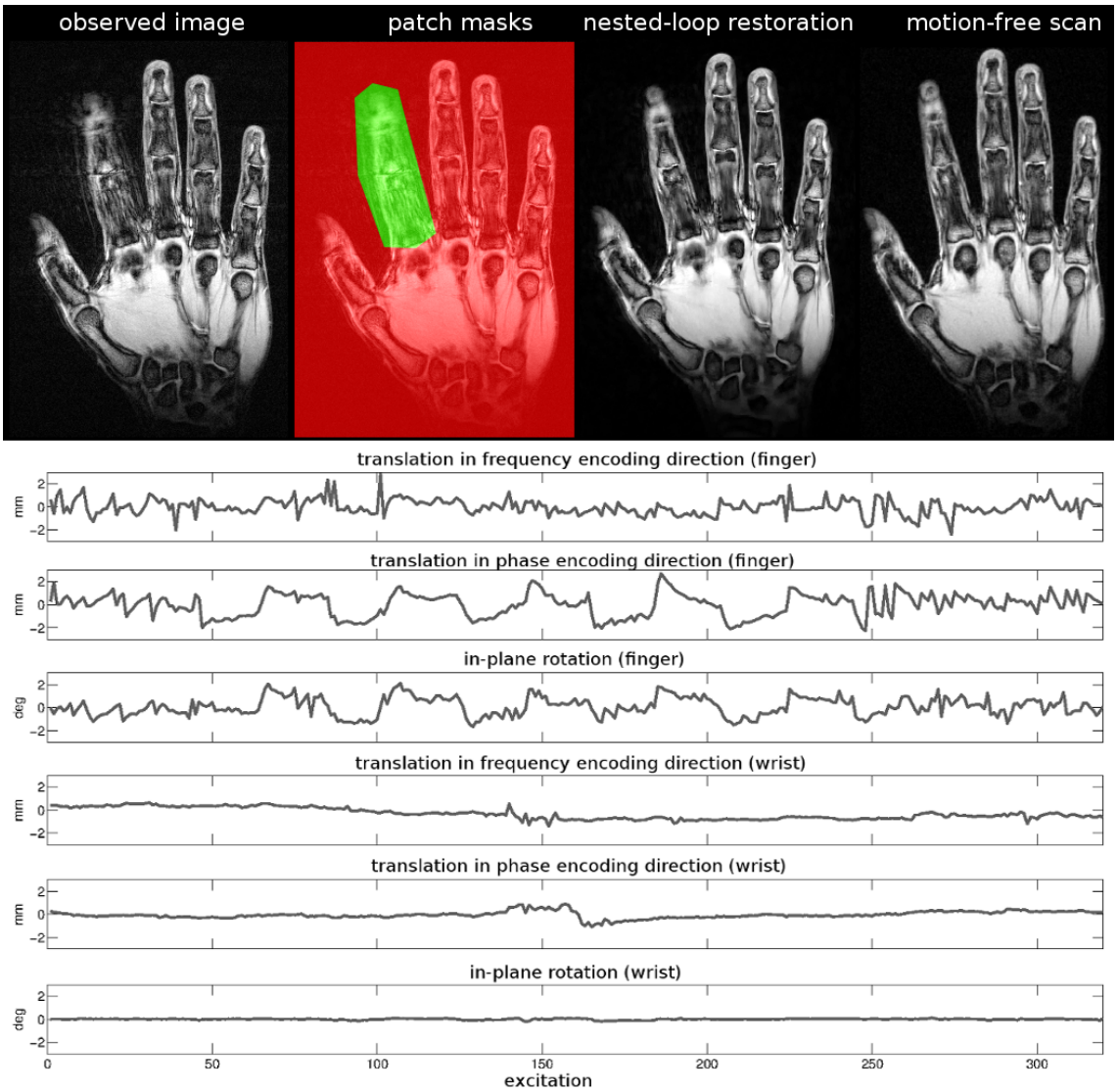

Figure 5.9: Real data. Multi-rigid motion correction was performed on in vivo data, where the image of the human wrist was acquired. During the experiment the subject was moving the index finger and was trying to keep the hand stationary. Top: from left to right, we show the observed image, the spatial masks used, motion correction result and ground-truth image from the motion free scan. Bottom: the trajectories for both patches recovered by the algorithm.

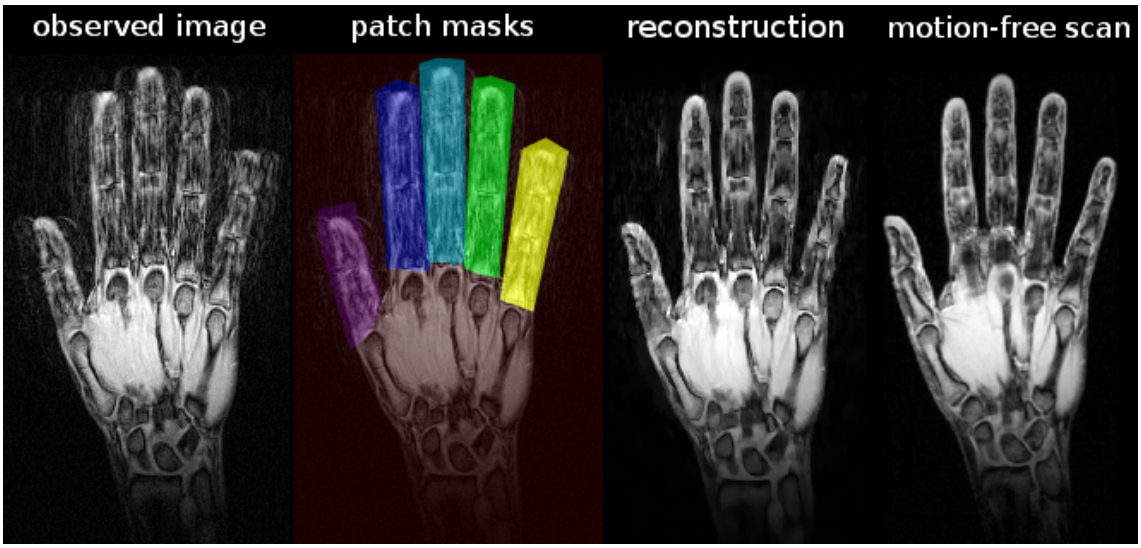

Figure 5.10: Real data. Multi-rigid motion correction was performed on the image split into $P = 6$ patches. The human subject was moving all six fingers during the experiment.

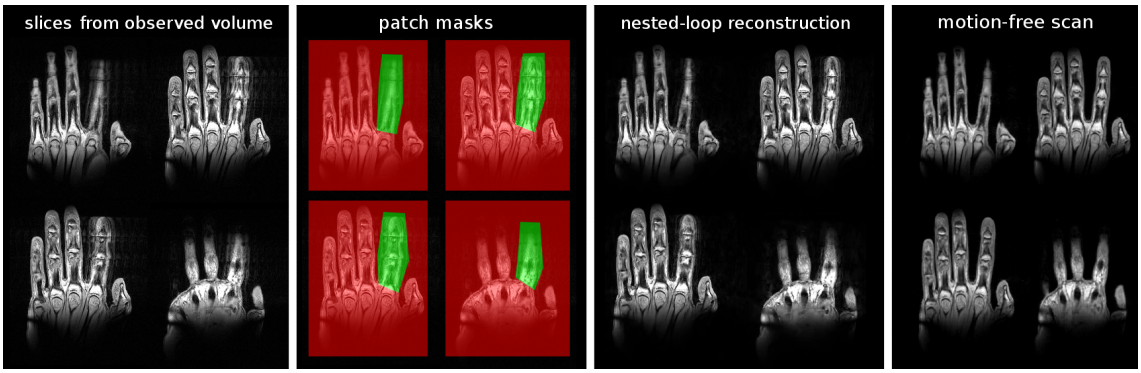

Figure 5.11: Real data. Multi-rigid motion correction was performed on the 3D image of the human hand. During the experiment the subject was moving the index finger. Left: four slices from the observed 3D dataset (containing 32 slices in total). Middle: spatial windows. Right: multi-rigid motion correction result, and the slices from the reference volume not affected by motion.

6 Conclusion

6.1 Overview

The main hypothesis of this thesis is that analytic gradient-based optimization improves the performance of autofocusing methods both with respect to the quality of reconstructed images and computation time. The existing autofocusing methods suffer from the problem of a curse of dimensionality, because the search of motion parameters is formulated in a combinatorial way, where exhaustive trial and error selection of the motion parameters that result in a best image metric is performed. Given the volume sizes of realistic medical data, such trial and error parameter selection leads to a combinatorial explosion making the search for good motion parameters computationally prohibitive. Our hypothesis was that, using the analytic model of the motion degradation in an objective function that combines the model of the motion degradation together with the image quality metric, it is possible to carry out a search of motion parameters efficiently. Importantly, our objective functions are differentiable with respect to both the image and motion parameters. Even though the resulting optimization problem is highly nonlinear, by using the multi-scale approach, we could achieve reconstruction results that were superior to those generated by existing methods.

Our first method, GradMC, was limited in not being applicable to the data corrupted with strong rotational and non-rigid motions. To address this limitation we developed an extension to our rigid motion method called GradMC2. Our hypothesis was that it is possible to use a fully blind motion correction method to address the problem of non-rigid motion in a way that does not also require sequence modifications and external tracking hardware. To our knowledge, prior to this work, there existed no algorithm that could correct for non-rigid motion without additional information on the motion trajectory. Importantly, our GradMC2 method is also capable of doing reconstruction carried out in the order of minutes making

it a practical tool. Low computation cost was achieved by using a fast GPU implementation of our a nested-loop algorithm featuring hierarchical annealing-type optimization with respect to the image and motion parameters. As a positive side effect of using a forward model-based nested-loop approach, we were also able to correct for strong rotational rigid motion, which was not possible in our original GradMC method. We have evaluated GradMC2 on both realistic and simulated data and observed reasonable reconstruction results.

To conclude, it should be emphasized that different existing methods of motion correction have distinct strengths and weaknesses, and there does not seem to exist an approach that is universally good. There is a trade off in computation time, requirements for extra (tracking) hardware, scanning time durations, and robustness to motions of varying strengths. The greatest strength of the methods that we have developed is that they are blind with respect to imaged objects' displacements, meaning that no information about motion parameters is required as input. Furthermore, our methods can be applied to the scans obtained with the use of generic sequences. Additionally, our method complements prospective motion correction methods. In high-resolution imaging, prospective methods can suffer from errors in estimated motion due to tracking noise or marker fixation instability. Applying our GradMC method to the data corrected with prospective approaches can further improve the image quality.

Active research in high-field MR hints that future scans will be acquired with even higher resolutions. At such resolutions, in vivo acquisitions are likely to suffer from the ubiquitous physiologic and small bulk motion. This means that motion correction will remain a fruitful area of research in the foreseeable future.

6.2 Future directions

6.2.1 Correction of strong motion

In our experiments we were mostly dealing with motion of low-to-mid intensity, which is a reasonable choice, since in clinical practice it is common to use cushions to constrain the movements of the subject, preventing large displacements from being physically possible. Studying the behavior of our algorithm in simulated problems of varying complexity we observed that strong motion (especially rotations) make the algorithm get stuck in bad local minima which leads to low quality of reconstructed

images. Future research should give a clear answer with regard to the bounds of the reconstruction capacity of the algorithm when it comes to strong motion.

6.2.2 Using the method in clinical practice

The approaches presented in this work were tested on the data acquired in research facilities. In order to facilitate the use of the method on a routine basis in hospitals, it is necessary to make the method easy and fast to use by medical personnel. Ideally, the method needs to be implemented in Siemens Image Calculation Environment (ICE), which will make it possible to do a correction just after the scan is acquired using the same graphical interface that is used to set up the scan. Additionally, the method can be incorporated into the commonly used Matlab SPM toolbox, and other similar environments such as FSL and AFNI.

6.2.3 Intra-view correction

One of the major assumptions that we make is that there is no substantial intra-view motion during image acquisition. This assumption is justified by the fact that a readout is made in a very short time interval. Still, it is quite possible that the subject moves during the readout making the data within a view inconsistent. It is an open question whether the retrospective correction of this kind of degradation is viable, and to what extent it is possible to improve the image quality.

6.2.4 Motion out of the coil sensitivity region

Another assumption that we make is that the outer parts of the scanned object never leave the sensitivity region of the receive coil. This assumption is fulfilled in most practical cases, since there is usually a zero intensity "insulator" (air) around the imaged object whenever the coil sensitivity profile is sufficiently broad. Still, it is quite possible that some parts of the scanned object leave the sensitivity region due to strong motion. This will result in a loss of spectral information on the object, and inadequate Fourier encoding. If the amount of missing information is substantial, the inverse-based approach will likely fail. A possible solution to this problem is to use the forward model that incorporates the effects of object motion out of the coil sensitivity region.

6.2.5 Learning the typical patterns of subject motion

An interesting direction to explore is suggested by the following question: what is the space of possible motion trajectories admitted by some particular part of a human body? It seems to be a valid guess that the classes of motorically plausible motions occupy a low-dimensional manifold within a space of all possible motions. For example, observing the spectra of typical motion trajectories might lead to a conclusion that human subjects are unlikely to make periodic movements with certain frequencies. Furthermore, there are likely to be constraints in degrees of freedom of a given body part, due to the obvious property that any body part is connected to the rest of the body. To illustrate this, one can consider the movements of the head, which is attached to the rest of the body by a neck, meaning that the amplitude of motion in the dorsoventral axis is limited. Additionally, the effects of gravity and weight might make certain movements less probable. For example, a human subject is likely to turn his head around a point, which is located at the back of her head, where the head touches the pillow.

It is also natural to assume certain kinds of regularities in the motion patterns of human subjects. It is reasonable to expect, for example, that the motion of the subject at some time point t_n will depend on the motion at time t_{n-1}. In order to properly address the question of typical motor patterns, any future study could involve the acquisition of a large database of motion trajectories of different parts of the human body from many human subjects. This would require the use of optical tracking cameras. Using better models of human motion would allow a higher quality of motion correction to be achieved, including robustness with respect to strong and complicated motions and shorter reconstruction times.

6.2.6 Motion trajectory parameterizations

The previous subsection naturally leads to the issue of motion trajectory parameterizations. In our experiments we used an identity trajectory parameterization (3 translation and 3 rotation parameters at each time point coding for position), which makes no assumptions about possible couplings of motion parameters in time and DOF. The identity parameterization is robust making the algorithm applicable to arbitrary classes of motions. It also has certain strengths when it comes to optimization, since having more parameters to optimize allows more flexible exploration of the objective landscape and opens up routes of escape from the bad local minima.

In a few experiments we also tried using a difference parameterization. In this parameterization it is not the position of the object, but its velocity at each time point that is described by parameters. Thus, this parameterization essentially involves first order finite derivatives of the motion parameters from identity parameterization. Such parameterization has certain nice properties, i.e. in the case of a piece-wise constant motion, most of the parameters are zero and thus, the underlying vector of velocities in each DOF is sparse. Although, theoretically this can lead to certain optimization benefits, in practice while using this kind of parameterization we have not observed any substantial difference when compared to naive identity parameterization. This might be due to the fact that real motion trajectories are far from being piece-wise constant.

Still, there is a plenty of room for potential research when it comes to different types of motion trajectory parameterizations. A good initial choice might be a kind of parameterization which takes into account the fact that human subjects are likely to turn their heads around a point, where the head touches the pillow. Dealing with non-rigid motions in the human abdomen might require even more elaborate motion models that should reflect the dynamic properties of internal organs.

6.2.7 Image metric functions

Another direction for possible future research has to do with objective image quality estimator functions. In current work we mostly employed the entropy of the image gradients. Recently, there has been a lot of progress in the field of computational photography related to the use of robust image metrics in certain inverse problems such as blind and non-blind deconvolution [73]. One finding was that, since the distribution of the gradients of natural images is heavy-tailed, it is beneficial to use the hyper-laplacian priors. Along the same lines, efficient approximation to L0 norm was successfully used by [75]. The transfer of knowledge from the field of computational photography to the field of image reconstruction in MRI will require a dedicated study of the statistical properties of medical images, and an understanding of what their similarities to the natural images acquired with optical modalities are. It is rather likely that certain similarities, such as power law decrease of energy of the spectrum as a function of frequency, exist, meaning that many of the findings from the field of inverse problems can be applicable in autofocusing motion correction of MR images. Another interesting idea is to use large capacity learners such as neural

networks in order to learn the characteristic features of medical images and then use them to aid reconstruction.

6.2.8 Confidence on the reconstruction results

When it comes to retrospective motion correction, often the question is raised about whether the reconstruction results can be trusted by radiologists. It is possible to imagine, for example, that given a strong motion the reconstructed result can still have the artifacts that mimic the anatomical or pathological image data such as false lesions. Problems of such a kind might become even more severe in a case where there is a strong emphasis on using medical image priors by retrospective motion correction method. Reasonable treatment of such problems involves providing the medical expert with an estimate of the confidence of the algorithm in the corrected result. Thus, given a strong motion, the confidence is expected to be low. Having an objective measure of how trustworthy the reconstructed image is, a radiologist can decide on the next action, which may involve re-acquisition of an extra scan.

6.2.9 Magnitude image correction

In clinical practice, it is common to store medical images in DICOM format, which usually only preserves the spatial magnitude of the images. As was explained in Chapter 4, retrospective motion correction of magnitude only data is very hard to do. In particular, it is not possible to use the inverse-based approach to do the correction, because the assumption that each line in k-space is only affected by motion from one particular time point is violated. Instead, the motion distortion of each line now depends on the motion parameters from an entire motion trajectory in a complicated way. Additionally, since the spatial phase is discarded, there is a lot of missing information which makes the naive inversion invalid. In our experiments, we tried using the forward model, which incorporated the effects of applying a modulus operator in spatial domain to simulate the phase elimination. We then used an alternating procedure introduced in Chapter 5 to carry out the correction, but we did not observe any improvement in the quality of reconstructed images. This is likely due to the fact that the objective involving modulus operator is highly non-convex and even non-differentiable at origin. The problem of retrospective correction of the magnitude data is still unsolved, while being of a high practical value. It would

be interesting to try using machine learning approaches to tackle this problem, i.e. neural networks can be trained to learn mapping from the motion corrupted magnitude data to the motion-free outputs. Such an approach to the problem would require access to a sufficiently large database of medical images in order to avoid the problem of overfitting.

6.2.10 Parallel imaging: SENSE/GRAPPA

It is common to use parallel imaging in order to acquire data in an accelerated manner, where some of the k-space lines are not sampled. This allows acquisition times to be shorter, making the scan less likely to be affected by motion. Still, even with accelerated acquisitions, the scanning usually takes a few minutes, which makes it likely that the subject will move. In order to reconstruct the image acquired with acceleration, often the SENSE or GRAPPA method is used. The SENSE method relies on the knowledge of coil sensitivity profiles, and carries out the reconstruction in spatial domain. The GRAPPA method is a Fourier domain method which requires full sampling of some small low-frequency region of the spectrum called the autocalibration region. This region is used to learn the kernel that is then applied to the undersampled data in order to infer the missing spectral k-space lines. An interesting question is whether our Fourier domain based method can be combined with GRAPPA reconstruction in order to correct the accelerated data for motion. One plausible thing to try is to alternate the GRAPPA reconstruction and inverse-based motion correction steps which, after several iterations, might result in an attenuation of artifacts due to motion.

6.2.11 Radial/Spiral acquisitions

In this work we were dealing with line-by-line Cartesian k-space sampling trajectories. There exist other possible acquisition patterns, some of which have certain useful properties when it comes to motion correction. For example, radial acquisitions are prone to artifacts due to translational motion. This is because in each acquired view the phases are congruent, which means that the phases are consistent in all directions. Our framework is general enough to be easily extended to allow correction of images acquired with arbitrary k-space traverse trajectories. The extension of the method would involve a different process of the grid construction to

be used in the resampling step. This would allow the strengths of a fully blind retrospective approach to be combined with motion-resistant properties of certain non-Cartesian acquisition trajectories.

6.2.12 Automatic mask extraction

Our multi-rigid GradMC2 algorithm expects segmentation masks to be provided as input. Mask creation is a potentially cumbersome process, and automatic means to generate the spatial windows would be useful. There are two possible ways to address this problem. Firstly, a dense regular patch splitting could be used instead of arbitrarily shaped masks. At sufficiently fine splitting level there is a guarantee that any non-rigid motion can be modeled. The drawback of such an approach is high computational complexity, which is due to the fact that each of the small patches has an associated set of motion parameters. The second possible solution would be to use automatic segmentation which can split the image into regions where the motion is locally rigid. This could be done, for example, by automatic analysis of the motion artifacts in different spatial regions, but it is not clear how accurate such masks would be, especially given strong motion.

6.2.13 Optimizing GPU implementation

The fast computation times attained by our methods are due to the fact that we run our algorithms on modern graphics cards. Profiling our algorithms we observed that a bottleneck arises in the Fast Fourier transform meaning that our implementation is already quite efficient. In order to avoid the common problem of low memory transfer bandwidth associated with the use of GPUs, we minimized the host to device memory transactions by storing all persistent data in device memory without the need to move it on each objective function evaluation. Still, many computational savings are possible; for example, not recomputing the objective function gradients on each line search could lead to a substantial decrease in running time. Furthermore, in current implementations we have been using 2X overgridding to avoid large errors in the resampling step. It might be the case that, in order to estimate the motion parameters, a high precision in resampling is not necessary which would allow resampling to be done without overgridding. This means that data of a smaller size would need to be Fourier transformed leading to significant computational savings.

6.2.14 Using FID-navigators for motion detection/estimation

An interesting direction for possible future research would be to relax the requirement for the method to be completely blind, and use weak guidance from some external source of knowledge about motion. A particularly good candidate would be the Free Induction Decay navigator [76], which could be acquired for each repetition with a negligible time cost of only a few milliseconds. Thus, it would be possible to integrate FID navigators into most sequences without increasing acquisition time. The signal from FID navigators alone is not sufficient to recover motion parameters, but it can be used to detect repetitions where strong motion has occurred. This would allow for development of a special motion trajectory parameterization to reduce the number of unknowns for use in our autofocusing method.

Bibliography

[1] M. Wood and R. Henkelman, "MR image artifacts from periodic motion," *Medical Physics*, vol. 12, p. 143, 1985.

[2] M. A. Bernstein, K. F. King, and X. J. Zhou, *Handbook of MRI Pulse Sequences*. Elseviewer, 2004.

[3] J. P. Hornak, "The Basics of MRI.".

[4] D. Manke, K. Nehrke, P. Rösch, and O. Dössel, "Study of Respiratory Motion of the Heart in Coronary MR Angiography," *International Journal of Bioelectromagnetism*, vol. 2, no. 2, 2000.

[5] M. Zaitsev, C. Dold, G. Sakas, J. Hennig, and O. Speck, "Magnetic resonance imaging of freely moving objects: prospective real-time motion correction using an external optical motion tracking system," *Neuroimage*, vol. 31, pp. 1038–1050, 2006.

[6] L. Qin, P. van Gelderen, J. A. Derbyshire, F. Jin, J. Lee, J. A. de Zwart, Y. Tao, and J. H. Duyn, "Prospective head-movement correction for high-resolution MRI using an in-bore optical tracking system," *Magnetic Resonance in Medicine*, vol. 62, pp. 924–934, 2009.

[7] T. Lee, H. Su, S. Lai, and T. Chang, "Compensation of motion artifacts in mri via graph-based optimization," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pp. 2192–2199, june 2009.

[8] W. Lin, F. Wehrli, and H. Song, "Correcting bulk in-plane motion artifacts in MRI using the point spread function," *IEEE Transactions on Medical Imaging*, vol. 24, no. 9, pp. 1170–1176, 2005.

[9] W. Lin and H. Song, "Extrapolation and correlation (EXTRACT): a new method for motion compensation in MRI," *IEEE Transactions on Medical Imaging*, vol. 28, no. 1, pp. 82–93, 2009.

[10] M. Hedley, H. Yan, and D. Rosenfeld, "Motion artifact correction in MRI using

generalized projections," *IEEE Transactions on Medical Imaging*, vol. 10, no. 1, pp. 40–46, 1991.

[11] K. O. Johnson, R. K. Robison, and J. G. Pipe, "Rigid body motion compensation for spiral projection imaging," *IEEE Transactions on Medical Imaging*, vol. 30, no. 3, pp. 655–665, 2011.

[12] Z. Fu, Y. Wang, R. Grimm, P. Rossman, J. Felmlee, S. Riederer, and R. Ehman, "Orbital navigator echoes for motion measurements in magnetic resonance imaging," *Magnetic Resonance in Medicine*, vol. 34, no. 5, pp. 746–753, 1995.

[13] E. Welch, A. Manduca, R. Grimm, H. Ward, and C. Jack, "Spherical navigator echoes for full 3-d rigid body motion measurement in MRI," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001*, 2001.

[14] A. Manduca, K. McGee, E. Welch, J. Felmlee, R. Grimm, and R. Ehman, "Autocorrection in MR imaging: Adaptive motion correction without navigator echoes," *Radiology*, vol. 215, no. 3, pp. 904–909, 2000.

[15] B. Frieden, "Restoring with maximum likelihood and maximum entropy," *Journal of the Optical Society of America*, vol. 62, no. 4, pp. 511–518, 1972.

[16] R. Muller and A. Buffington, "Real-time correction of atmospherically degraded telescope images through image sharpening," *Journal of the Optical Society of America*, vol. 64, no. 9, pp. 1200–1210, 1974.

[17] R. Bocker and S. Jones, "ISAR motion compensation using the burst derivative measure as a focal quality indicator," *International Journal of Imaging Systems and Technology*, vol. 4, no. 4, pp. 285–297, 1992.

[18] D. Atkinson, D. Hill, P. Stoyle, P. Summers, and S. Keevil, "Automatic correction of motion artifacts in magnetic resonance images using an entropy focus criterion," *IEEE Transactions on Medical Imaging*, vol. 16, no. 6, pp. 903–910, 1997.

[19] D. Atkinson, D. L. G. Hill, P. N. R. Stoyle, P. E. Summers, S. Clare, R. Bowtell, and S. F. Keevil, "Automatic compensation of motion artifacts in MRI," *Magnetic Resonance in Medicine*, vol. 41, pp. 163–170, 1999.

[20] W. Lin, G. Ladinsky, F. Wehrli, and H. Song, "Image metric-based correction (autofocusing) of motion artifacts in high-resolution trabecular bone imaging," *Journal of Magnetic Resonance Imaging*, vol. 26, no. 1, pp. 191–197, 2007.

[21] K. McGee, J. Felmlee, C. Jack, A. Manduca, S. Riederer, and R. Ehman, "Au-

tocorrection of three-dimensional time-of-flight MR angiography of the circle of willis," *American Journal of Roentgenology*, vol. 176, no. 2, pp. 513–518, 2001.

[22] P. Batchelor, D. Atkinson, P. Irarrazaval, D. Hill, J. Hajnal, and D. Larkman, "Matrix description of general motion correction applied to multishot images," *Magnetic Resonance in Medicine*, vol. 54, no. 5, pp. 1273–1280, 2005.

[23] J. Y. Cheng, M. T. Alley, C. H. Cunningham, S. S. Vasanawala, J. M. Pauly, and M. Lustig, "Nonrigid motion correction in 3d using autofocusing with localized linear translations," *Magnetic Resonance in Medicine*, vol. 68, no. 6, pp. 1785–1997, 2012.

[24] F. Odille, P. Vuissoz, P. Marie, and J. Felblinger, "Generalized reconstruction by inversion of coupled systems (GRICS) applied to free-breathing MRI," *Magnetic Resonance in Medicine*, vol. 60, no. 1, pp. 146–157, 2008.

[25] J. Maclaren, M. Herbst, O. Speck, and M. Zaitsev, "Prospective motion correction in brain imaging: A review," *Magnetic Resonance in Medicine*, vol. 69, no. 3, pp. 621–636, 2012.

[26] C. Dold, M. Zaitsev, O. Speck, E. A. Firle, J. Hennig, and G. Sakas, "Advantages and limitations of prospective head motion compensation for MRI using an optical motion tracking device," *Academic Radiology*, vol. 13, pp. 1093–1103, 2006.

[27] J. Maclaren, B. S. R. Armstrong, R. T. Barrows, K. A. Danishad, T. Ernst, C. L. Foster, K. Gumus, M. Herbst, I. Y. Kadashevich, T. P. Kusik, Q. Li, C. Lovell-Smith, T. Prieto, P. Schulze, O. Speck, D. Stucht, and M. Zaitsev, "Measurement and correction of microscopic head motion during magnetic resonance imaging of the brain," *PLoS ONE*, vol. 7, p. e48088, 11 2012.

[28] A. J. W. van der Kouwe, T. Benner, and A. M. Dale, "Real-time rigid body motion correction and shimming using cloverleaf navigators," *Magnetic Resonance in Medicine*, vol. 56, pp. 1019–1032, 2006.

[29] N. White, C. Roddey, A. Shankaranarayanan, E. Han, D. Rettmann, J. Santos, J. Kuperman, and A. Dale, "Promo: Real-time prospective motion correction in mri using image-based tracking," *Magnetic Resonance in Medicine*, vol. 63, no. 1, pp. 91–105, 2010.

[30] M. B. Ooi, S. Krueger, W. J. Thomas, S. V. Swaminathan, and T. R. Brown, "Prospective real-time correction for arbitrary head motion using active markers," *Magnetic Resonance in Medicine*, vol. 62, pp. 943–954, 2009.

[31] J. G. Pipe, "Motion correction with PROPELLER MRI: Application to head motion and free-breathing cardiac imaging," *Magn. Reson. Med.*, vol. 42, no. 5, pp. 963–969, 1999.

[32] J. Maclaren, K. Lee, C. Luengviriya, O. Speck, and M. Zaitsev, "Combined prospective and retrospective motion correction to relax navigator requirements," *Magnetic Resonance in Medicine*, vol. 65, no. 6, pp. 1724–1732, 2011.

[33] K. McGee, A. Manduca, J. Felmlee, S. Riederer, and R. Ehman, "Image metric-based correction (autocorrection) of motion effects: Analysis of image metrics," *Journal of Magnetic Resonance Imaging*, vol. 11, no. 2, pp. 174–181, 2000.

[34] C. Plathow, S. Ley, J. Zaporozhan, M. Schöbinger, E. Gruenig, M. Puderbach, M. Eichinger, H.-P. Meinzer, I. Zuna, and H.-U. Kauczor, "Assessment of reproducibility and stability of different breath-hold maneuvres by dynamic mri: comparison between healthy adults and patients with pulmonary hypertension.," *Eur Radiol*, vol. 16, no. 1, pp. 173–9, 2006.

[35] M. Soellinger, A. K. Rutz, S. Kozerke, and P. Boesiger, "3d cine displacement-encoded mri of pulsatile brain motion.," *Magn Reson Med*, vol. 61, no. 1, pp. 153–62, 2009.

[36] D. Manke, P. Rösch, K. Nehrke, P. Börnert, and O. Dössel, "Model evaluation and calibration for prospective respiratory motion correction in coronary MR angiography based on 3-d image registration.," *IEEE Transactions on Medical Imaging*, vol. 21, no. 9, pp. 1132–41, 2002.

[37] T. Schäffter, V. Rasche, and I. C. Carlsen, "Motion compensated projection reconstruction," *Magnetic Resonance in Medicine*, vol. 41, no. 5, pp. 954–963, 1999.

[38] F. Odille, N. Cîndea, D. Mandry, C. Pasquier, P. Vuissoz, and J. Felblinger, "Generalized MRI reconstruction including elastic physiological motion and coil sensitivity encoding," *Magnetic Resonance in Medicine*, vol. 59, no. 6, pp. 1401–1411, 2008.

[39] J. Schmidt, M. Buehrer, P. Boesiger, and S. Kozerke, "Nonrigid retrospective respiratory motion correction in whole-heart coronary MRA.," *Magnetic Resonance in Medicine*, vol. 66, no. 6, pp. 1541–1549, 2011.

[40] F. Odille, S. Uribe, P. G. Batchelor, C. Prieto, T. Schäffter, and D. Atkinson, "Model-based reconstruction for cardiac cine MRI without ECG or breath holding," *Magnetic Resonance in Medicine*, vol. 63, no. 5, pp. 1247–1257, 2010.

[41] P. Vuissoz, F. Odille, B. Fernandez, M. Lohezic, A. Benhadid, D. Mandry, and J. Felblinger, "Free-breathing imaging of the heart using 2d cine-GRICS (generalized reconstruction by inversion of coupled systems) with assessment of ventricular volumes and function," *Journal of Magnetic Resonance Imaging*, vol. 35, no. 2, pp. 340–351, 2012.

[42] M. Hansen, T. Sørensen, A. Arai, and P. Kellman, "Retrospective reconstruction of high temporal resolution cine images from real-time MRI using iterative motion correction," *Magnetic Resonance in Medicine*, vol. 68, no. 3, pp. 741–50, 2012.

[43] M. J. White, D. J. Hawkes, A. Melbourne, D. J. Collins, C. Coolens, M. Hawkins, M. O. Leach, and D. Atkinson, "Motion artifact correction in free-breathing abdominal MRI using overlapping partial samples to recover image deformations," *Magnetic Resonance in Medicine*, vol. 62, pp. 440–449, Aug. 2009.

[44] A. Menini, P. Vuissoz, J. Felblinger, and F. Odille, "Joint reconstruction of image and motion in MRI: Implicit regularization using an adaptive 3D mesh," in *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, vol. 15, pp. 264–271, 2012.

[45] M. Usman, D. Atkinson, F. Odille, C. Kolbitsch, G. Vaillant, T. Schaeffter, P. G. Batchelor, and C. Prieto, "Motion corrected compressed sensing for free-breathing dynamic cardiac MRI," *Magnetic Resonance in Medicine*, vol. doi, p. 10.1002/mrm.24463. [Epub ahead of print], 2012.

[46] A. Manduca, K. McGee, E. Welch, J. Felmlee, and R. Ehman, "Autofocusing of clinical shoulder MR images for correction of motion artifacts," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 1998*.

[47] D. Atkinson and D. Hill, "Automatic motion correction using prior knowledge," in *Proc 9th Annual Meeting ISMRM, Glasgow, 2001*, 2001.

[48] E. Kholmovski, A. Samsonov, and D. Parker, "Motion artifact reduction technique for dual-contrast FSE imaging," *Magnetic Resonance Imaging*, vol. 20, no. 6, pp. 455–462, 2002.

[49] M. E. Bourgeois, F. T. Wajer, M. Roth, A. Briguet, M. Décorps, D. van Ormondt, C. Segebarth, and D. Graveron-Demilly, "Retrospective intra-scan motion correction," *Journal of Magnetic Resonance*, vol. 163, pp. 277–287, 2003.

[50] D. Atkinson and D. Hill, "Reconstruction after rotational motion," *Magnetic*

Resonance in Medicine, vol. 49, no. 1, pp. 183–187, 2003.

[51] W. Lin and H. Song, "Improved optimization strategies for autofocusing motion compensation in MRI via the analysis of image metric maps," *Magnetic Resonance Imaging*, vol. 24, no. 6, pp. 751–760, 2006.

[52] H. Su, T. Lee, S. Lai, and T. Chang, "MRI motion artifact correction based on spectral extrapolation with generalized series," in *IEEE International Conference on Image Processing*, 2010.

[53] A. Samsonov, J. Velikina, Y. Jung, E. Kholmovski, C. Johnson, and W. Block, "POCS-enhanced correction of motion artifacts in parallel MRI," *Magnetic Resonance in Medicine*, vol. 63, no. 4, pp. 1104–1110, 2010.

[54] Y. Kadah, "Overlapped k-space acquisition and reconstruction technique for motion artifact reduction in magnetic resonance imaging," *Journal of Medical Imaging and Health Informatics*, vol. 1, no. 1, pp. 61–65, 2011.

[55] Z. Yang, C. Zhang, and L. Xie, "Sparse mri for motion correction," *CoRR*, vol. abs/1302.0077, 2013.

[56] J. Fessler and B. P. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Transactions on Signal Processing*, vol. 51, pp. 560–574, 2003.

[57] L. Greengard and J. Y. Lee, "Accelerating the nonuniform fast fourier transform," *SIAM Review*, vol. 46, no. 3, pp. 443–454, 2004.

[58] V. Rasche, R. Proksa, R. Sinkus, P. Börnert, and H. Eggers, "Resampling of data between arbitrary grids using convolution interpolation," *IEEE Transactions on Medical Imaging*, vol. 18, no. 5, pp. 385–392, 1999.

[59] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, pp. 1153–1160, 1981.

[60] J. Pauly, "Non-cartesian reconstruction," `http://www.stanford.edu/class/ee369c/notes/non_cart_rec_07.pdf` 2007.

[61] R. D. Hoge, R. K. S. Kwan, and G. B. Pike, "Density compensation functions for spiral MRI," *Magnetic Resonance in Medicine*, vol. 38, pp. 117–128, 1997.

[62] M. D. Tisdall, A. T. Hess, M. Reuter, E. M. Meintjes, B. Fischl, and A. J. W. van der Kouwe, "Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical mri," *Magnetic Resonance in*

Medicine, vol. 68, no. 2, pp. 389–399, 2012.

[63] R. N. Bracewell, *The Fourier Transform and Its Applications*. Mc Graw Hill, 1999.

[64] I. Najfeld and T. F. Havel, "Derivatives of the matrix exponential and their computation," *Advances in Applied Mathematics*, vol. 16, pp. 321–375, 1995.

[65] A. W. Paeth, "A fast algorithm for general raster rotation," *Proceedings Graphics Interface*, pp. 77–81, 1986.

[66] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM Journal on Scientific and Statistical Computing*, vol. 16, pp. 1190–1208, 1995.

[67] H. Nickisch, "glm-ie: Generalised linear models inference & estimation toolbox," *Journal of Machine Learning Research*, vol. 13, pp. 1699–1703, 2012.

[68] P. Beatty, D. Nishimura, and J. Pauly, "Rapid gridding reconstruction with a minimal oversampling ratio," *Medical Imaging, IEEE Transactions on*, vol. 24, pp. 799 –808, june 2005.

[69] K. O. Johnson and J. G. Pipe, "Convolution kernel design and efficient algorithm for sampling density correction," *Magnetic Resonance in Medicine*, vol. 61, no. 2, pp. 439–447, 2009.

[70] M. Hirsch, S. Sra, B. Schölkopf, and S. Harmeling, "Efficient filter flow for space-variant multiframe blind deconvolution.," in *CVPR*, pp. 607–614, IEEE, 2010.

[71] J. Morelli, V. Runge, F. Ai, U. Attenberger, L. Vu, S. Schmeets, W. Nitz, and J. Kirsch, "An image-based approach to understanding the physics of mr artifacts.," *Radiographics*, vol. 31, no. 3, pp. 849–66, 2011.

[72] F. T. A. W. Wajer and K. P. Pruessmann, "Major speedup of reconstruction for sensitivity encoding with arbitrary trajectories," in *ISMRM*, p. 767, 2001.

[73] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," in *ACM Transactions on Graphics (TOG)*, vol. 25, pp. 787–794, ACM, 2006.

[74] S. Cho and S. Lee, "Fast motion deblurring," in *ACM Transactions on Graphics (TOG)*, vol. 28, p. 145, ACM, 2009.

[75] L. Xu, S. Zheng, and J. Jia, "Unnatural l0 sparse representation for natural image deblurring.," in *CVPR*, pp. 1107–1114, IEEE, 2013.

[76] T. Kober, J. P. Marques, R. Gruetter, and G. Krueger, "Head motion detection using fid navigators," *Magnetic Resonance in Medicine*, vol. 66, no. 1, pp. 135–143, 2011.