

# Motion-Corrected Reconstruction in Cone-Beam Computed Tomography for Knees under Weight-Bearing Condition

Bewegungskorrigierte Bildrekonstruktion für die  
Kegelstrahl-Computertomographie von Knien unter Belastung

Der Technischen Fakultät  
der Friedrich-Alexander-Universität  
Erlangen-Nürnberg

zur

**Erlangung des Doktorgrades Dr.-Ing.**

vorgelegt von

Martin Berger  
aus  
Günzburg

Als Dissertation genehmigt  
von der Technischen Fakultät  
der Friedrich-Alexander-Universität Erlangen-Nürnberg

Tag der mündlichen Prüfung:	28.09.2016
Vorsitzende des Promotionsorgans:	Prof. Dr. P. Greil
Gutachter:	Prof. Dr.-Ing. habil. A. Maier Prof. Dr.-Ing. M. Stamminger

## Abstract

Medical imaging plays an important role in diagnosis and grading of knee conditions such as osteoarthritis. In current clinical practice, 2-D radiography is regularly applied under weight-bearing conditions, which is known to improve diagnostic accuracy. However, 2-D images cannot fully cover the complexity of a knee joint, whereas current 3-D imaging modalities are inherently limited to a supine, unloaded patient position. Recently, cone-beam computed tomography (CBCT) scanners for 3-D weight-bearing imaging have been developed. Their specialized acquisition trajectory poses several challenges for image reconstruction. Patient motion caused by standing or squatting positions can substantially deteriorate image quality, such that the motion has to be corrected during reconstruction. Initial work on motion correction is based on fiducial markers, yet, the approach prolonged image acquisition and required a large amount of manual interaction. The goal of this thesis was to further develop innovative motion correction methods for weight-bearing imaging of knees.

Within the course of this thesis, the marker-based motion correction was steadily enhanced. Manual annotation of markers has been replaced by a robust, fully automatic detection of markers and their correspondences. A clear disadvantage of markers is the often tedious attachment, which decreases patient comfort and interferes with the acquisition protocol. Also, the method is limited to rigid motion and an extension to nonrigid deformations is nontrivial. To alleviate these drawbacks, we introduce a novel motion estimation approach that makes use of a prior, motion-free reference reconstruction. The motion of femur and tibia is determined individually by rigid 2-D/3-D registration of bone segmentations from the prior scan, to each of the acquired weight-bearing projection images. Reliability of the registration is greatly influenced by the large amount of overlapping structures, especially for lateral view directions. We compare two different similarity measures used for 2-D/3-D registration and also introduce a temporal smoothness regularizer to improve registration accuracy. A common evaluation of marker- and registration-based approach yields superior image quality using 2-D/3-D registration, particularly in presence of severe, nonrigid motion. Further reduction of the algorithm's runtime and an automation of bone segmentations could allow for a complete replacement of marker-based motion correction in future applications.

In case the clinical setup prohibits acquisition of a prior scan, motion correction relies solely on the acquired projection images. We derived a new motion correction method based on Fourier consistency conditions (FCC) which is independent of surrogates or prior acquisitions. So far, FCC have not been used for motion correction and were typically limited to fan-beam geometries. We first introduced the motion estimation for the fan-beam geometry, followed by a practical extension to CBCT. Numerical phantom simulations revealed a particularly accurate estimation of high-frequency motion and of motion collinear to the scanner's rotation axis. FCC are currently limited to nontruncated, full-scan projection data, and thus, not yet applicable to real weight-bearing acquisitions. However, a dynamic apodization technique is introduced to account for axial truncation, allowing application to a squatting knee phantom with realistic motion. Given the large improvements in image quality, we are confident that FCC is a future candidate for a completely self-contained motion correction approach in CBCT weight-bearing imaging of knees.

## Kurzfassung

Die medizinische Bildgebung übernimmt eine wichtige Rolle bei der Diagnose von Kniegelenkserkrankungen wie z.B. Arthrose. Im klinischen Umfeld wird hierbei oft die 2-D Röntgenbildgebung im Belastungszustand angewandt, welche die hohe Komplexität des Kniegelenks allerdings nicht vollständig abdecken kann. Aktuelle 3-D Bildgebungsverfahren sind wiederum nur sehr bedingt unter Belastung anwendbar. Erst vor Kurzem wurden spezielle Kegelstrahl-Computertomographen (KSCT) für die 3-D Belastungsbildgebung entwickelt. Allerdings stellen deren Aufnahmetrajektorien auch erhöhte Anforderungen an die Bildrekonstruktion. So kann z.B. Patientenbewegung die Bildqualität deutlich verschlechtern, so dass eine bewegungskorrigiert Rekonstruktion nötig ist. Erste Ansätze basieren auf metallischen Markern, welche jedoch die Bildaufnahmen verzögern und ein hohes Maß an manueller Interaktion verlangen. Ziel dieser Dissertation war es, die Entwicklung bestehender Bewegungskorrekturen voranzutreiben und um hochinnovative Methoden zu erweitern.

Teil dieser Arbeit war eine stetige Weiterentwicklung der marker-basierten Korrektur, wobei die manuelle Annotation durch eine robuste, vollautomatische Markerdetektion ersetzt wurde. Einen Nachteil stellt das mühsame Anbringen der Marker dar, welches nicht nur den Ablauf der Bildaufnahmen, sondern auch den Patienten negativ beeinflusst. Außerdem ist die Methode derzeit auf rigide Bewegungen beschränkt und nur schwer auf elastische Deformationen erweiterbar. Abhilfe schaffte eine neu entwickelte Bewegungskorrektur, basierend auf einer bewegungsfreien Referenzrekonstruktion. Eine 2-D/3-D Bildregistrierung von segmentierten Knochen aus der Referenzrekonstruktion, zu jedem Projektionsbild der Belastungsbildgebung, ermittelt Bewegungen von Femur und Tibia. Die Robustheit der Registrierung wird hierbei stark durch überlappende Strukturen beeinflusst. Um die Genauigkeit zu erhöhen, verglichen wir verschiedene Ähnlichkeitsmaße und integrierten einen Regularisierer, welcher eine zeitliche Glattheit der Bewegung erzwingt. Die gemeinsame Auswertung von marker- und registrierungs-basiertem Ansatz zeigt eine überlegene Bildqualität bei Einsatz der 2-D/3-D Registrierung, insbesondere bei starker nicht-rigider Bewegung. Mittels einer automatisierten Knochensegmentierung und einer beschleunigten Registrierung könnte der marker-basierte Ansatz somit komplett ersetzt werden.

Sollte keine bewegungsfreie Referenzrekonstruktion möglich sein, beruht die Bewegungskorrektur einzig auf den aufgenommenen Bildern. Basierend auf Fourier-Konsistenzbedingungen (FKB) entwickelten wir eine Bewegungskorrektur, welche unabhängig von zusätzlichen Informationen oder Aufnahmen arbeitet. Bisher wurden FKB nicht für die Bewegungskorrektur eingesetzt und waren auf die Fächerstrahlgeometrie (FSG) begrenzt. Wir definierten zunächst eine Bewegungsschätzung für die FSG, welche im Folgenden für KSCT erweitert wurde. Numerische Simulationen zeigten eine besonders genaue Bewegungsschätzung für hochfrequente Bewegung, aber auch für Bewegungen die kollinear zur Rotationsachse des Scanners erfolgten. Derzeit sind FKB nicht auf realen Daten anwendbar, da sie nur für einen trunkierungsfreien, vollständigen Scan gültig sind. Eine eigens entwickelte dynamische Apodisation erlaubt den Umgang mit axialer Trunkierung, was die Evaluation auf numerischen Kniedaten mit realer Bewegung ermöglichte. Eine starke Verbesserung der Bildqualität zeigt, dass eine komplett unabhängige Bewegungskorrektur in der Belastungsbildgebung von Knien möglich ist.

## Acknowledgment

This thesis would not have been possible without the support and trust of so many people I met during my time as a PhD student. In the following I would like to take the time to thank at least a subset of those people.

First and foremost, I want to thank Prof. Dr.-Ing. habil. Andreas Maier for his outstanding supervision, his constant encouragement, and many fruitful technical discussions. He was always available in urgent cases, even at unusual business hours. I had the chance to learn plenty of details about CT reconstruction from him and will never forget the numerous discussions about our reconstruction framework CONRAD.

I also want to express my gratitude to Prof. Dr.-Ing. Joachim Hornegger. With his fascinating lectures on medical image processing, he laid the foundation for my interest in the field. When I had to decide on my future career, he provided me with the opportunity to pursue my PhD at his lab, in such an uncomplicated fashion.

No less of my appreciation goes to Prof. Rebecca Fahrig, PhD for making me part of the weight-bearing project and for an exciting stay at Stanford, including excellent supervision. Despite her busy schedule, she was there for me in scientific (as well as medical) emergencies and never got tired of proof reading my papers.

Many thanks also go to Prof. Dr.-Ing. Marc Stamminger, for a flawless lead of our Research Training Group, a great internship at his lab that led to many new ideas on 2-D/3-D registration, and for providing me a quiet desk to write up this thesis.

A big, big “Thank You” goes to all the colleagues and friends at the Pattern Recognition Lab, who surely built one of the main pillars of this work. In particular I would like to thank: Dr.-Ing. Yan Xia, for a great and fun time in our office and your invaluable thoughts on Fourier Consistency Conditions. Dr.-Ing. Kerstin Müller, for an uncountable number of new ideas for motion-corrected CT and memorable conference trips. Mathias Unberath, for our top-notch discussions on motion correction, your thorough reviews of my papers, and glorious micro table tennis matches. André Aichert, for your infinite knowledge on projection geometries and many curious technical discussions about consistency conditions. Dr.-Ing. Christian Riess, for an unforgettable time in Stanford and your inspiring “the show must go on” attitude. Finally, I want to thank Johannes Jordan, Chris Schwemmer, Eva Eibenberg, Christoph Forman, Peter Fischer, Jakob Wasza, Dr.-Ing. Sebastian Bauer, Dominik Neumann, Felix Breunig-Lugauer, Christian Jaremenko, and Bastian Bier, for being a part of my PhD.

I would like to acknowledge funding of the Research Training Group 1773 Heterogeneous Image Systems by the German Research Foundation (DFG). I also acknowledge support from the NIH Shared Instrument Grant S10 RR026714, the NIH Weight-bearing imaging Grant R01 AR065248, and Siemens AT.

Last but most certainly not least, I want to thank my family and friends for their moral support and encouragement. In particular, I would like to thank my wife Stefanie, for being there for me in all the demanding phases of this doctorate. This thesis would have never been possible without you.

Martin Berger

To My Little Son

# Contents

<b>Contents</b>	<b>i</b>
<b>Chapter 1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	2
1.2 Contributions . . . . .	2
1.2.1 Fully Automatic, Marker-Based Motion Correction . . . . .	2
1.2.2 Motion Correction Using 2-D/3-D Image Registration . . . . .	3
1.2.3 FCC for Motion Correction in Fan- and Cone-Beam CT . . . . .	4
1.2.4 Other Contributions to CBCT and Weight-Bearing Imaging . . . . .	4
1.3 Organization of this Thesis . . . . .	5
<b>I Background</b>	<b>9</b>
<b>Chapter 2 Imaging of the Knee Under Weight-Bearing Conditions</b>	<b>11</b>
2.1 Medical Background . . . . .	11
2.1.1 Knee Anatomy . . . . .	11
2.1.2 Common Disorders of the Knee Joint . . . . .	12
2.1.3 Imaging for OA and PFPS . . . . .	13
2.2 CBCT Weight-Bearing Imaging . . . . .	14
2.2.1 Motivation . . . . .	14
2.2.2 CBCT Systems . . . . .	15
2.3 Clinical Studies Using Weight-Bearing CBCT . . . . .	16
2.3.1 Feasibility Evaluation . . . . .	16
2.3.2 Cartilage Deformation Under Weight-Bearing Conditions . . . . .	17
2.3.3 Verification of Treatment Methods for PFPS . . . . .	17
2.4 Challenges for Image Reconstruction . . . . .	18
2.4.1 Lateral Truncation / Limited FOV . . . . .	18
2.4.2 Detector Saturation . . . . .	19
2.4.3 Patient Motion . . . . .	20
2.5 Conclusion . . . . .	21
<b>Chapter 3 Image Reconstruction and Motion Correction in CBCT</b>	<b>23</b>
3.1 Mathematical Definitions . . . . .	23
3.2 CBCT Imaging . . . . .	24
3.2.1 X-ray Physics . . . . .	24
3.2.2 Fan-Beam CT . . . . .	25
3.2.3 CBCT . . . . .	27
3.2.4 Iterative Reconstruction . . . . .	29
3.2.5 Discretization and Projection Matrices . . . . .	30

3.3	Motion Correction . . . . .	32
3.4	Motion Estimation Methods. . . . .	33
3.4.1	Motion Models . . . . .	33
3.4.2	Based on Surrogate Signals (I) . . . . .	35
3.4.3	Based on Reference Image (II) . . . . .	37
3.4.4	Without Reference and Surrogate Data (III) . . . . .	39
3.5	Motion Compensation Methods . . . . .	42
3.5.1	Updated Backprojection for Affine Motion . . . . .	43
3.5.2	Updated Backprojection for Nonrigid Motion . . . . .	43
3.6	Conclusion. . . . .	44
<b>II</b>	<b>Theory</b>	<b>47</b>
	<b>Chapter 4 Marker-Based Motion Correction</b>	<b>49</b>
4.1	Robust, Fully Automatic Marker Detection . . . . .	49
4.1.1	2-D Detection Using Fast Radial Symmetry Transform. . . . .	50
4.1.2	Automatic Extraction of 3-D Reference Positions . . . . .	52
4.1.3	Determining 2-D/3-D Point Correspondences . . . . .	53
4.2	Optimization and Outlier Detection . . . . .	54
4.2.1	Gradient of Cost Function . . . . .	55
4.2.2	Iterative Detection and Removal of Outliers . . . . .	55
4.3	Removal of Markers . . . . .	56
4.3.1	Comparison of Extrapolation Methods . . . . .	57
	<b>Chapter 5 Image-Based Motion Correction Using 2-D/3-D Registration</b>	<b>59</b>
5.1	Method Overview . . . . .	60
5.2	Digitally Rendered Radiographs . . . . .	61
5.3	Similarity Measures. . . . .	62
5.3.1	Gradient Correlation . . . . .	63
5.3.2	Normalized Gradient Information. . . . .	63
5.4	Temporal Regularization and Optimization. . . . .	64
5.5	Nonrigid Motion Field Generation. . . . .	65
5.6	Noise Reduction in DRR Images. . . . .	66
	<b>Chapter 6 Image-Based Motion Correction Using Consistency Conditions</b>	<b>69</b>
6.1	FCC and their Extension to CBCT . . . . .	70
6.2	FCC With Arbitrary Motion Models . . . . .	71
6.3	Efficient Implementation for Detector Translations . . . . .	72
6.3.1	Gradient of Cost Function . . . . .	73
6.3.2	Invariance to Static Detector Offsets . . . . .	74
6.4	Regularization, Optimization and Motion Compensation . . . . .	75
6.5	Practical Considerations . . . . .	76
6.5.1	Discretization and Adjustment of the Mask . . . . .	76
6.5.2	Axial Truncation: Problems and Solutions . . . . .	78
6.5.3	Implementation Details . . . . .	81

<b>III Experiments and Results</b>	<b>85</b>
<b>Chapter 7 Data Acquisition and Common Metrics</b>	<b>87</b>
7.1 Simulated Phantom Data . . . . .	87
7.1.1 Bead Removal Phantom. . . . .	89
7.1.2 FORBILD Head Phantom. . . . .	89
7.1.3 XCAT Dynamic Squat Phantom . . . . .	92
7.2 Real CBCT Acquisitions . . . . .	93
7.2.1 Subject 1 . . . . .	94
7.2.2 Subject 2 . . . . .	94
7.2.3 Subject 3 . . . . .	95
7.3 Common Image Quality Metrics . . . . .	96
7.3.1 MSE, RMSE and rRMSE . . . . .	97
7.3.2 SSIM and UQI . . . . .	97
<b>Chapter 8 Joint-Evaluation of Registration- and Marker-Based Method</b>	<b>99</b>
8.1 Materials and Methods . . . . .	100
8.1.1 Parameter Selection And Image Reconstruction. . . . .	100
8.2 Reconstruction Results . . . . .	102
8.3 Quantitative Assessment of Image Quality . . . . .	108
8.3.1 Quantitative Evaluation Pipeline . . . . .	108
8.3.2 UQI Results . . . . .	110
8.4 Cross-Evaluation at Markers and Bone Surfaces . . . . .	111
8.4.1 Bone Surface Profiles . . . . .	112
8.4.2 Fiducial Registration Error (FRE) . . . . .	112
8.5 Removal of Marker-Induced Metal Artifacts . . . . .	115
8.5.1 Results Based on Numeric Phantom . . . . .	115
8.5.2 Results for Real Data. . . . .	116
8.6 Discussion and Conclusion . . . . .	117
<b>Chapter 9 Evaluation of FCC-Based Motion Correction</b>	<b>123</b>
9.1 Materials and Methods . . . . .	123
9.1.1 Fan-Beam Study . . . . .	124
9.1.2 Cone-Beam Study . . . . .	124
9.1.3 Parameter Selection And Image Reconstruction. . . . .	125
9.2 Initial Evaluation for Fan-Beam CT. . . . .	126
9.3 FCC-based Motion Correction for CBCT . . . . .	129
9.3.1 Reconstruction Results. . . . .	129
9.3.2 Quantitative Assessment of Image Quality . . . . .	134
9.3.3 Accuracy of Motion Estimation . . . . .	136
9.4 Results for Simulated Knee Phantom . . . . .	138
9.4.1 Parameter Selection and Image Reconstruction . . . . .	139
9.4.2 Image Quality of Reconstructed Images . . . . .	139
9.4.3 Improvements in Projection and Fourier Domain . . . . .	140
9.5 Discussion and Conclusion . . . . .	141

<b>Chapter 10 Outlook</b>	<b>147</b>
<b>Chapter 11 Summary</b>	<b>151</b>
<b>Chapter A Appendix</b>	<b>155</b>
A.1 Definition and Derivative of Rigid Motion . . . . .	156
<b>List of Acronyms</b>	<b>159</b>
<b>List of Symbols</b>	<b>163</b>
<b>List of Figures</b>	<b>171</b>
<b>List of Tables</b>	<b>173</b>
<b>List of Algorithms</b>	<b>175</b>
<b>Bibliography</b>	<b>177</b>

# Introduction

1.1 Motivation.....	2
1.2 Contributions.....	2
1.3 Organization of this Thesis.....	5

Since its discovery in 1895 by Wilhelm Conrad Röntgen, X-ray imaging has steadily evolved, making it to one of the leading modalities in modern radiology. For a long period of time, X-ray imaging was limited to radiography or fluoroscopy, showing single or a sequence of 2-D projection images of the imaged anatomies. These projection images offered the first possibilities to look inside a human body without surgical intervention. They were especially helpful for conditions of the skeletal system as bones yielded a suitable contrast due to their high absorption coefficient. However, a spatially resolved representation of the imaged object is not possible with 2-D projection imaging as the images correspond to line integrals through the object. All materials and objects between X-ray source and the detector are superimposed and the order of transmission cannot be resolved from a single image.

The connection between line integrals acquired on a circular orbit and the object function was the starting point of X-ray **computed tomography (CT)**. The theoretical foundation was given by Radon [Rado 17] and Cormack [Corm 63] who solved this connection, allowing computation of the function describing the imaged object. This process is known as Radon inversion and generally referred to as **CT image reconstruction**. Based on findings from Cormack and Radon, the first **CT scanner** applied in the medical context was built by Hounsfield in the early 1970's [Houn 73]. Since that time **CT** has become an invaluable tool for diagnosis in the field of radiology, with application areas ranging over almost the entire spectrum of medical fields.

Current systems can be separated into diagnostic and interventional **CT**. Diagnostic **CT** consists of a single or multiple rotating X-ray sources and circular detectors. The plane of rotation and the **source-detector-distance (SDD)** as well as **source-isocenter-distance (SID)** is fixed. A linear motion of the patient table as well as a fast rotation of X-ray source and detector creates a helical source trajectory around the patient. Interventional **CT** is usually done by a C-arm **cone-beam computed tomography (CBCT)** device using flat panel or image intensifier detectors. The C-arm device can be used inside the interventional suite and allows for 2-D fluoroscopic but also 3-D **CT** imaging. In contrast diagnostic **CT**, which allows for a full 360° rotation, C-arm systems are usually limited to a short-scan rotation angle of about 200°. Their X-ray source trajectory moves in a plane while acquiring high-resolution projection images. A comparison of diagnostic and interventional **CTs** may be obtained from [Mull 14a].

## 1.1 Motivation

The baseline of this thesis is work by Choi et al. [Choi 14a], who describe a novel scanning setup allowing acquisitions of the lower extremities under weight-bearing conditions. An interventional C-arm CBCT device is used, where the X-ray source trajectory lies in a horizontal plane, such that the patient is able to stand or squat during the imaging process. In general, image reconstruction for CT as well as CBCT assumes a static object with an ideal geometry of the scanner, e. g., an X-ray source moving on a perfect circle. An implication of weight-bearing imaging is an increased level of patient motion, leading to severe motion artifacts. Among other sources of artifacts, patient motion causes the highest degradation of image quality when using the proposed scanning setup and needs to be corrected during image reconstruction.

As an initial motion correction method, Choi et al. proposed to use metallic fiducial markers that can be tracked in the individual projection images acquired during a weight-bearing scan [Choi 12, Choi 13, Choi 14c]. However, the method requires a high amount of manual interaction which limited its practical applicability and thus, its clinical value. Furthermore, the attachment of markers might be impractical for the clinical workflow and causes discomfort to the patient. Also, the markers may not accurately reflect motion at the knee joint, as they are attached to the skin at the periphery of the joint.

The goal of this thesis is to further develop motion management methods for weight-bearing imaging of the knees. In particular, the thesis focuses on the design and validation of motion-corrected image reconstruction algorithms that do not require dedicated surrogate signals, such as fiducial markers. Purely image-based motion correction methods are derived for knee imaging, exploiting information of motion-free supine scans or data consistency conditions (CC). In addition, robustness and autonomy of the marker-based approach are further optimized.

## 1.2 Contributions

In the course of this thesis, several contributions to the state-of-the-art in motion correction and CT reconstruction have been made. Many of these contributions are part of publications in national and international conferences but also in well known journals. The main achievements are introduced in the following, accompanied with the corresponding reference to the literature.

### 1.2.1 Fully Automatic, Marker-Based Motion Correction

An important part of marker-based motion estimation is the detection of the markers in the projection image and the identification of their correspondences between different images. In prior work [Choi 14c] this process required a large amount of manual interaction. We propose a **fully automatic detection of 2-D marker locations and their correspondences**. The method is based on an intermediate reconstruction of projection images showing segmented markers locations. For increased robustness we propose to use the **fast radial symmetry transform (FRST) algorithm**

for segmentation. In addition, a method for the **removal of marker-induced metallic artifacts** is introduced. To further improve robustness of the motion estimation, two **methods for detection of outliers** and a the use of a **gradient-based optimizer** is proposed, where the gradient of the cost function is computed analytically. The methods are detailed in Chap. 4 and were presented in three conference and a journal publication:

- [Berg 14a]
M. Berger, C. Forman, C. Schwemmer, J. H. Choi, K. Müller, A. Maier, J. Hornegger, and R. Fahrig. “Automatic Removal of Externally Attached Fiducial Markers in Cone Beam C-arm CT”. In: T. Deserno, H. Handels, H.-P. Meinzer, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2014*, pp. 318–323, Mar 2014
- [Mull 15a]
K. Müller, M. Berger, J.-H. Choi, S. Datta, S. Gehrisch, T. Moore, M. P. Marks, A. Maier, and R. Fahrig. “Fully Automatic Head Motion Correction for Interventional C-arm Systems using Fiducial Markers”. In: S. King and M. Glick, Eds., *Proc. Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, pp. 534–537, Jun 2015
- [Mull 15b]
K. Müller, M. Berger, J.-H. Choi, A. Maier, and R. Fahrig. “Automatic Motion Estimation and Compensation Framework for Weight-bearing C-arm CT scans using Fiducial Markers”. In: D. A. Jaffray, Ed., *IFMBE Proceedings*, pp. 58–61, Jun 2015
- [Berg 16a]
M. Berger, K. Müller, A. Aichert, M. Unberath, J. Thies, J.-H. Choi, R. Fahrig, and A. Maier. “Marker-free motion correction in weight-bearing cone-beam CT of the knee joint”. *Medical Physics*, Vol. 43, No. 3, pp. 1235–1248, Mar 2016

### 1.2.2 Motion Correction Using 2-D/3-D Image Registration

A marker-free approach for nonrigid motion estimation of the knee joint was developed, based on a bone-wise **2-D/3-D registration of an artifact-free prior volume to the individual projection images**. The bones are segmented in a reconstruction of a motion-free supine scan and then registered to the projection images, yielding the individual bone motion over time. Evaluation included a comparison with marker-based motion correction and of different similarity measures used for 2-D/3-D registration. Further details are presented in Chap. 5 and have been published in a conference and a journal article:

- [Berg 15]
M. Berger, K. Müller, J.-H. Choi, A. Aichert, A. Maier, and R. Fahrig. “2D/3D Registration for Motion Compensated Reconstruction in Cone-Beam CT of Knees Under Weight-Bearing Condition”. In: D. A. Jaffray, Ed., *IFMBE Proceedings*, pp. 54–57, Jun 2015

[Berg 16a] M. Berger, K. Müller, A. Aichert, M. Unberath, J. Thies, J.-H. Choi, R. Fahrig, and A. Maier. “Marker-free motion correction in weight-bearing cone-beam CT of the knee joint”. *Medical Physics*, Vol. 43, No. 3, pp. 1235–1248, Mar 2016

### 1.2.3 FCC for Motion Correction in Fan- and Cone-Beam CT

Fourier Consistency Conditions (FCC) define locations in the Fourier domain of the projection images that are required to have an absolute value close to zero. We show that **FCC can be violated in case of object motion**. Additionally, a **novel approach for motion estimation based on FCC** has been proposed. The method seeks to minimize the energy of the zero regions and requires no additional acquisition or surrogate signal. We initially introduced the method for the 2-D fan-beam geometry, followed by an extension to 3-D cone-beam geometries. Details may be obtained from Chap. 6 and have been presented in a conference and a journal publication:

[Berg 14b] M. Berger, A. Maier, Y. Xia, J. Hornegger, and R. Fahrig. “Motion Compensated Fan-Beam CT by Enforcing Fourier Properties of the Sinogram”. In: F. Noo, Ed., *Proc. The third international conference on image formation in x-ray computed tomography*, pp. 329–332, Jun 2014

[Berg 16b] M. Berger, Y. Xia, W. Aichinger, K. Mentl, M. Unberath, A. Aichert, C. Riess, J. Hornegger, R. Fahrig, and A. Maier. “Translational Motion Compensation for Cone-Beam CT Using Fourier Consistency Conditions”. *Physics in Medicine and Biology*, 2016. (under revision)

### 1.2.4 Other Contributions to CBCT and Weight-Bearing Imaging

A series of contributions to the current state of CT reconstruction literature have been made, that are in close connection to this thesis. Many of them are also dedicated to challenges in image reconstruction of weight-bearing acquisitions of knees (cf. Sec. 2.4 for more details). The main developments include:

- a novel method to increase the **field of view (FOV)** of the scanner, using modified source trajectories and detector offsets [Herb 14, Herb 15]
- a method for correction of detector saturation artifacts based on fitting simple geometric objects to the acquired images [Preu 15]
- an investigation of hardware-based correction of saturation artifacts [Shi 16]
- a removal of artifacts originating from objects outside the scanner’s FOV [Bier 16a]
- an extrapolation of missing data using a cost function based on FCC [Pohl 14]

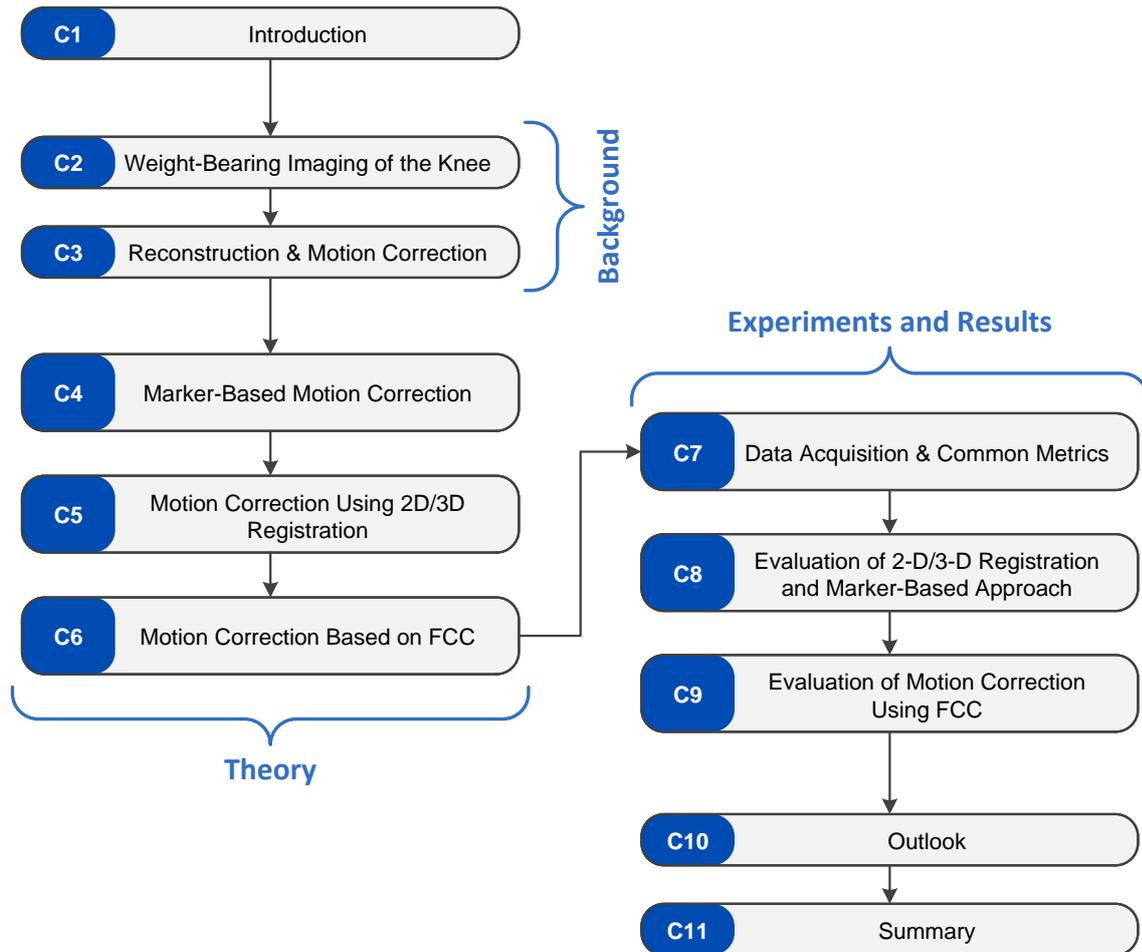
- nonlinear denoising techniques in projection and reconstruction domain [Lorc 15]
- a preliminary approach for motion estimation based on 2-D/2-D registration of maximum intensity projections (MIP) and acquired projections [Unbe 15]
- improvements of image quality for marker-based motion estimation [Choi 14b]
- a method to correct for artifacts caused by an insufficient scan range [Ries 13]
- advancements for truncation correction using a dynamic collimator in full-, short-, and super-short-scans [Xia 13, Xia 14c, Xia 14b]
- a truncation correction that incorporates the patient outline [Xia 14a, Xia 15]
- a scatter correction using a spatial modulation of the X-ray source [Bier 16b]
- a learning-based material decomposition in poly-energetic projection images [Lu 16]

## 1.3 Organization of this Thesis

Let us summarize the structure and contents of this thesis. A graphical overview of parts and chapters is given by Fig. 1.1. In general, the thesis consists of three parts, i. e., background, theory, along with experiments and results.

The thesis starts with an introduction in Chap. 1, providing an overview of the topic and an outline of the thesis' main contributions to the state-of-the-art. The background part begins with Chap. 2, which elaborates on the medical background of knee joint imaging, introduces CBCT weight-bearing imaging, and presents its main challenges for image reconstruction. This is followed by a revision of basic principles for image reconstruction in fan- and cone-beam CT presented in Chap. 3. In addition, the chapter includes a thorough review of literature associated to motion correction and introduces fundamental algorithms and concepts needed throughout this thesis.

The following theory part details the proposed methods and consists of three chapters. Chap. 4 presents all modifications and improvements made to the original formulation for marker-based motion correction [Choi 14c] in a unified mathematical framework. This includes a fully automatic detection of marker, an outlier removal method, and an analytic derivative of the cost function to allow for gradient-based optimization. Motion correction based on 2-D/3-D registration is outlined in Chap. 5. We present a novel approach for 2-D/3-D registration of multiple segmented, artifact-free bone volumes to the individual projection images of a weight-bearing scan. Aside from the method's description, this chapter outlines the principle of 2-D/3-D registration in general and provides details to similarity measures and nonrigid extrapolation of motion fields using thin plate splines (TPS). A novel method for motion correction based on FCC is presented in Chap. 6. We show the derivation of a cost function for motion estimation based on the 2-D and 3-D Fourier transform of fan- and cone-beam projection data, respectively. Additionally, a particularly efficient implementation of the cost function for optimizing detector translations is introduced. The chapter



**Figure 1.1:** Overview of this thesis, separated into background (top, Part I), theory (left center, Part II), and experiments and results (right center, Part III).

concludes with practical considerations, providing an efficient strategy for the implementation of the cost function and its gradient. Note that all methods in the theory part are based on one or more cost functions used for motion estimation. Optimization and regularization of these cost functions is outlined separately in each chapter.

The experiments and results part begins with Chap. 7, where we introduce acquisition details on real and simulated data, as well as common metrics used for the experiments of this thesis. In Chap. 8 we conduct a common evaluation of the marker-based motion correction and the approach based on 2-D/3-D registration using two different similarity measures. The evaluation includes a qualitative comparison of reconstructed images, an image-based quantitative comparison w. r. t. the motion-free supine reconstruction, and an evaluation of the reprojection error using the attached fiducial markers. In addition, a comparison of different removal techniques for marker-induced metal artifacts is shown. The evaluation of the FCC-based method is presented separately in Chap. 9, as it is not yet applicable to real weight-bearing acquisitions. The assessment is performed on a challenging numerical phantom for a series of simulated 3-D translational motion patterns. It contains a

qualitative but also quantitative analysis of the image quality w. r. t. a motion-free reconstruction. Additionally, deviations of true and estimated motion parameters are obtained. Finally, a first evaluation of FCC-based motion correction for simulated knee data is shown.

Both result chapters end with a discussion and conclusion providing strengths and weaknesses of the evaluated methods and a short summary of the main findings.

The thesis concludes with an outlook in Chap. 10, which provides ideas for future research directions, followed by a summary of the presented contents in Chap. 11.



PART I

# Background



# Imaging of the Knee Under Weight-Bearing Conditions

2.1	Medical Background .....	11
2.2	CBCT Weight-Bearing Imaging .....	14
2.3	Clinical Studies Using Weight-Bearing CBCT .....	16
2.4	Challenges for Image Reconstruction .....	18
2.5	Conclusion .....	21

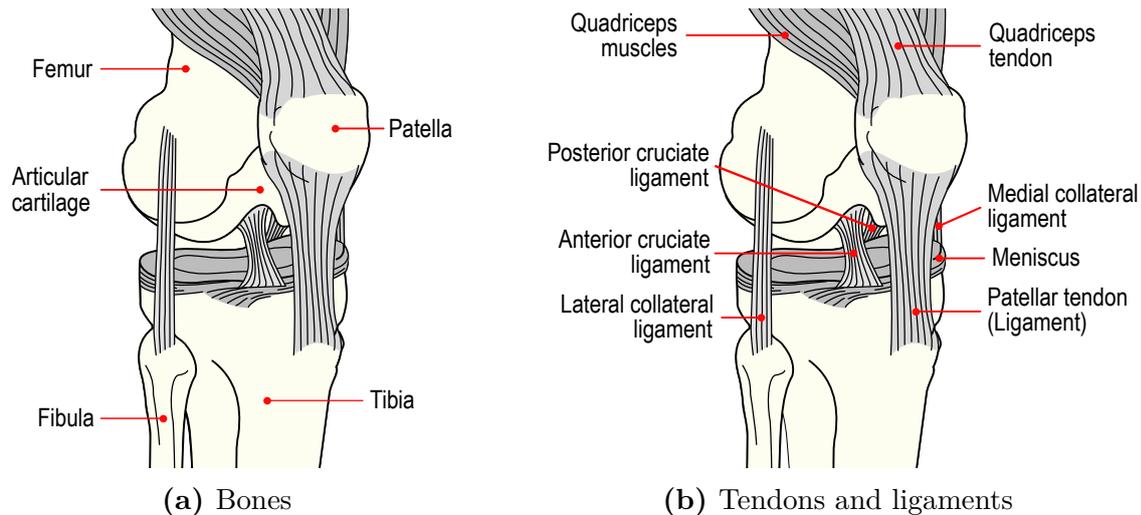
The first part of this chapter contains the clinical motivation (see Sec. 2.1) including a description of the knee’s anatomy, common knee joint disorders and the importance of weight-bearing conditions for diagnosis. This is followed by an overview of CBCT weight-bearing imaging and a more detailed description of the clinical studies relevant for this thesis (see Sections 2.2 and 2.3). Thereafter, Sec. 2.4 describes the implications of an acquisition under weight-bearing conditions for 3-D image reconstruction. Finally, a conclusion on the clinical relevance and its implications on weight-bearing imaging is given in Sec. 2.5.

## 2.1 Medical Background

### 2.1.1 Knee Anatomy

The knee is the largest and most complex joint of the human body allowing rotation in vertical axis but also flexion and extension [Kulo 32]. As can be seen in Fig. 2.1(a), the knee joint is comprised by three different bones, i. e., femur (thigh), tibia (shin) and patella (knee cap). Fig. 2.1(b) shows ligaments and tendons. Ligaments consist of fibrous tissue and ensure stability of the knee by holding the individual bones together. The lateral collateral ligament connects the fibula (calf-bone) to the femur, whereas the posterior and anterior cruciate ligament as well as the medial collateral ligament connect femur and tibia. The patella is attached to the quadriceps muscle group and to the tibia by the quadriceps tendon and the patellar tendon, respectively.

Articular cartilage provides a low-friction surface at the distal femur and proximal tibia, but also at the posterior side of the patella. Cartilage surfaces ensure a smooth sliding motion during flexion, extension and rotation of the knee. Between tibial and femoral cartilage are the menisci, i. e., two wedge-shaped layers of fibrocartilaginous tissue that disperse friction and act as a shock absorber between femur and tibia [Choi 14a].



**Figure 2.1:** Anatomy of the knee joint with focus on bones (Fig. 2.1(a)) as well as tendons and ligaments. (Images in public domain).

During flexion of the knee the collateral ligaments are relaxed, whereas the cruciate ligaments are tense and effectively limit the amount of flexion to approximately  $120^\circ$  to  $150^\circ$ . The cruciate ligaments are twisted and unwind during medial and lateral rotation, respectively. During large medial rotation, the medial collateral ligament becomes tensed. Together with the maximum twist of the cruciate ligaments the medial rotation is limited to about  $45^\circ$  to  $60^\circ$  [Plat 04, pp. 206-213].

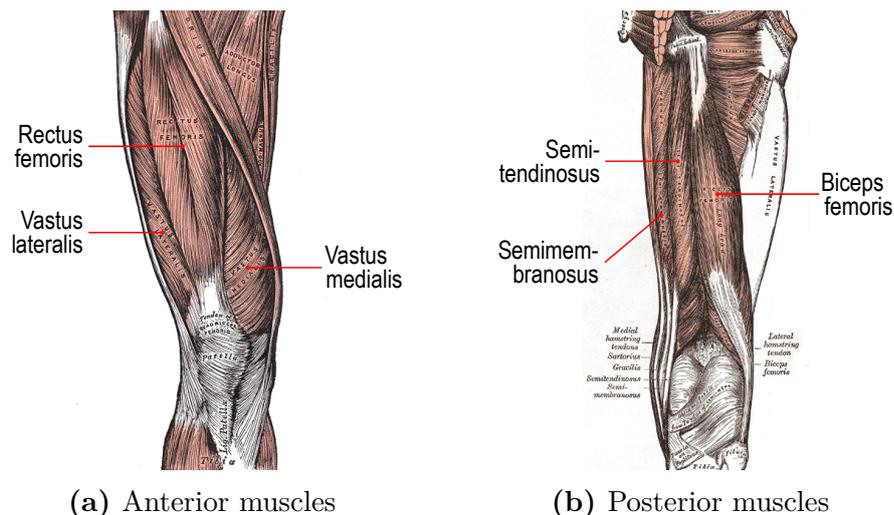
The muscles of the knee joint are illustrated in Fig. 2.2. They can be categorized into two main groups. The quadriceps group consists of the rectus femoris, the vastus lateralis, the vastus medialis and the vastus intermedius, where the latter is located underneath the rectus femoris and is therefore not visible in Fig. 2.2(a). All quadriceps muscles are located at the anterior side of the femur and function mainly as an extensor of the knee joint. The hamstring group located at the posterior side of the femur consists of the semitendinosus, the semimembranosus, and the biceps femoris muscle. Hamstring muscles build the counterpart to the quadriceps muscles and function as flexor but also decelerate the lower leg.

### 2.1.2 Common Disorders of the Knee Joint

The functionality of the knee joint can be affected by a variety of conditions. The focus of this thesis is on the two most commonly observed disorders, i. e., knee osteoarthritis (OA) and patellofemoral pain syndrome (PFPS).

#### Osteoarthritis

OA is a degenerative joint disease and the most frequent joint disorder in the United States, with knee OA showing the highest incident rates among all joints [Zhan 10]. The risk of suffering from knee OA during lifetime has been reported 45 % and even 60 % for a body mass index  $\geq 30$  [Murp 08]. OA affects the whole knee joint, typically causing a degeneration of the cartilage tissue and possibly also damage the synovium



**Figure 2.2:** Muscles of the knee joint. Labels in (a) depict the quadriceps group and labels in (b) the hamstring group. (Images in public domain)

or the joint capsule. Additionally, ligaments can be thickened, the menisci may wear off and osteophytes are built at affected bone surfaces [Engl 12]. Typical symptoms include knee pain, a limited range of joint motion, stiffness and joint swelling, which all affect regular daily activities and work life.

### Patellofemoral Pain Syndrome

PFPS is a common knee condition describing knee pain that originates from the contact area of patella and femur. Recently an incidence rate of 2.2% per year has been reported in a US-based study involving 1525 participants [Boli 10]. Physically active people performing sports such as running, basketball or other disciplines involving load bearing of the knee, are known to have a higher risk factor. People suffering from PFPS experience discomfort at activities involving knee flexion, such as prolonged sitting, squatting, or climbing up- and downstairs. Besides several other reasons, patellar maltracking is suggested to cause PFPS [Fred 06], yet the exact pathophysiological background is still an active field of research [Dye 05]. Patellar maltracking refers to an abnormal position or movement trajectory of the patella w. r. t. the trochlea (patellar groove) of the femur, increasing stress on the patellar cartilage. An imbalance between the vasti muscles, i. e., vastus intermedius, vastus lateralis and vastus medialis, is thought to cause patellar maltracking [Fulk 02]. More specifically, weakness or an activation delay of the vastus medialis was shown to correlate with patellar maltracking [Wern 95, Saka 00, Pal 11].

#### 2.1.3 Imaging for OA and PFPS

A standard clinical diagnosis and classification tool for knee OA is the Kellgren-Lawrence scale (KLS) [Kell 57]. An essential part of the KLS is 2-D radiography to measure narrowing of the joint-space width (JSW), development of osteophytes and bone sclerosis. This information acts as a surrogate signal for the degeneration of

cartilage and menisci. It has been shown that weight-bearing radiographs are more sensitive when grading knee OA than non-weight-bearing acquisitions, where the best results are achieved in anteroposterior rather than lateral radiographs [Altm 87]. Grading accuracy could be further improved using weight-bearing anteroposterior radiographs acquired with a certain knee flexion angle [Rose 88]. A limitation of radiographic imaging is that the cartilage is not directly visible in X-ray-based modalities and it is also restricted to 2-D projection images. In contrast, magnetic resonance imaging (MRI) offers great soft tissue contrast and is regularly used to analyze the course of cartilage and meniscus consistency [Ecks 06]. However, its application under weight-bearing conditions remains difficult.

Several imaging techniques have been proposed for diagnosis of PFPS, where the imagery is mostly used to identify irregular positions of the patella. For example, misalignment can be measured by a patella tilt angle, as well as a lateral or medial shift of the patella w.r.t. the trochlea [Pal 11, Stan 88]. Fulkerson et al. [Fulk 02] proposed weight-bearing anteroposterior radiographs with a 45° flexion angle, identical to imaging of knee OA used by Rosenberg et al. [Rose 88]. 3-D imaging methods for patellar pose measurements include MRI [Powe 99] but also weight-bearing, open-bore MRI [Powe 03, Besi 05, Drap 11].

## 2.2 CBCT Weight-Bearing Imaging

### 2.2.1 Motivation

It has been shown that weight-bearing imaging can provide an important tool to better understand knee joint kinematics essential for diagnosis and analysis of knee joint disorders, such as OA and PFPS. Regular daily activities causing pain in PFPS patients are often performed under weight-bearing conditions, whereas the majority of prior work is based on non-weight-bearing imaging methods [Souz 10]. Furthermore, studies showed substantial differences in kinematic properties of the knee under weight-bearing and non-weight-bearing conditions [Powe 03, Drap 11]. This is further supported by Thawait et al. [Thaw 15] who reports improved diagnosis of OA using upright weight-bearing imaging. Single- or dual-plane weight-bearing radiography was successfully applied for diagnosis of OA patients [Leac 70, Rose 88] but also for analysis of knee kinematics [Tash 03, Li 08]. A drawback of radiographic imaging is that it is inherently limited to 2-D images which may not fully cover the complex 3-D structure of the knee joint [Choi 14a].

Previous work in 3-D weight-bearing imaging include the usage of MRI in supine scanning position with specialized weight-application devices [Pate 03, Coto 11]. Further, open-bore MRI [Powe 03, Besi 05, Drap 09, Pal 11] has been used that allowed standing and squatting on a slanted weight-bearing platform. However, conventional but also open-bore MRI is still limited to supine or semi-upright postures due to the scanner design. Fully upright standing is possible with CBCT scanners dedicated to scanning extremities [Zbij 11, Tuom 13, Carr 14], yet, these devices offer only a limited FOV, which might not allow for squatting positions or simultaneous imaging of both knees. Choi et al. proposed the use of an interventional C-arm CBCT



(a) First prototype



(b) New robotic C-arm device

**Figure 2.3:** Typical weight-bearing scanning setups with the first prototype system and the newer, robotic C-arm device.

with a high-resolution flat panel (FP) detector and a large SDD, thus substantially increasing the FOV [Choi 14c].

### 2.2.2 CBCT Systems

The underlying imaging systems for this thesis are two interventional C-arm CBCT systems as described in [Choi 14c] and [Choi 14a]. In contrast to a standard interventional use case, their X-ray source trajectory can be positioned in the axial plane which allows scanning in fully upright standing and even squatting patient position. Maier et al. [Maie 11a] investigated the feasibility of horizontal weight-bearing trajectories for interventional C-arm devices. They conclude that horizontal trajectories suffer from similar levels of C-arm “wobble” than vertical trajectories and that proper trajectory calibration yields a stable and reproducible reconstruction quality over multiple acquisitions.

A typical scanning setup for the prototype C-arm weight-bearing system (Axiom Artis dTA, Siemens Healthcare GmbH, Forchheim, Germany) can be seen in Fig. 2.3(a). It shows an interventional C-arm CBCT which was adjusted to enable scanning with horizontal weight-bearing trajectories. The subject was positioned on a stand to align the FOV with the center of the knee joint, as a lower trajectory was not possible due to mechanical restrictions of the scanner. To fit both knees into the FOV both legs have been wrapped using two hook-and-loop straps. Additionally a layer of plasticine has been attached around the legs to avoid detector saturation artifacts (see Sec. 2.4.2).

A newer weight-bearing imaging system can be seen in Fig. 2.3(b). It shows a C-arm CBCT device that is mounted on an industrial robot system (Artis Zeego, Siemens Healthcare GmbH, Forchheim, Germany) allowing for a wide variety of X-ray source trajectories [Herb 15, Maie 15b, Yu 15, Stro 16a, Stro 16b, Gou 16]. In contrast to the prototype system, no height adjustment of the subject is necessary allowing for free, upright standing or squatting. The image also shows the patient position-

ing platform that ensures a standardized pose throughout studies involving multiple subjects or follow up scans. The plasticine wraps used for the prototype system are replaced by two synthetic arc-shaped beam blocking devices that increase X-ray absorption in anteroposterior view direction [Choi 14a].

Both prototype as well as robotic C-arm system are also capable of regular supine acquisitions required in most studies to obtain a non-weight-bearing reference scan. They cover a scan range of about  $200^\circ$  acquiring 100 to 500 projection images over a scan time of 3s to 20s. The corresponding detector resolution equals  $1240 \times 960$  pixels, with an isotropic pixel size of 0.308 mm. Further, the SID is 785 mm with an SDD of 1198 mm, leading to a maximum isotropic reconstruction resolution at the isocenter of  $\sim 200 \mu\text{m}$ .

## 2.3 Clinical Studies Using Weight-Bearing CBCT

In the following, an overview on current clinical applications of weight-bearing CBCT is given. The studies make use of the C-arm systems introduced in Sec. 2.2.2 and can be grouped into an initial feasibility evaluation, a study related to OA investigating dynamic cartilage properties, and an investigation of a novel treatment method for PFPS patients. This section is mostly based on the work of Choi, who presented a detailed description and motivation of the cartilage and PFPS study in [Choi 14a]. Note that the prototype C-arm system has only been used for the feasibility evaluation in Sec. 2.3.1. All other studies (see Sections 2.3.2 and 2.3.3) are based on acquisitions using the robotic C-arm device.

### 2.3.1 Feasibility Evaluation

The first step of validating the feasibility of weight-bearing acquisitions was based on numerical simulations [Choi 13]. Motion measurements based on an optical tracking system were acquired from nine volunteers, holding a squat for 20s. Besides an analysis of the motion itself, the measurements have been incorporated into a numerical dynamic knee phantom, which could then be used as part of a proof of concept study to show feasibility of weight-bearing imaging [Choi 13]. As a next step in the weight-bearing CBCT project, images from healthy volunteers were acquired. With the application of improved OA diagnosis in mind, the goal was to verify whether the new imaging modality could be used for 3-D measurements of the JSW, that serves as biomarker for OA grading. Derived from typical 2-D radiography weight-bearing protocols, all acquisitions included a supine non-weight-bearing reference scan, followed by multiple weight-bearing scans with different flexion angles, e. g.,  $0^\circ$ ,  $30^\circ$  and  $60^\circ$ . Additionally, a calibration phantom has been scanned to allow for a subject-specific geometric calibration of the weight-bearing trajectory. This early project phase confirmed expected challenges for 3-D image reconstruction, most importantly patient motion as explained in Sec. 2.4. In addition to the verification of the weight-bearing imaging mode, the acquired data helped to develop initial motion correction algorithms based on fiducial-markers [Choi 12, Choi 13, Choi 14c].

### 2.3.2 Cartilage Deformation Under Weight-Bearing Conditions

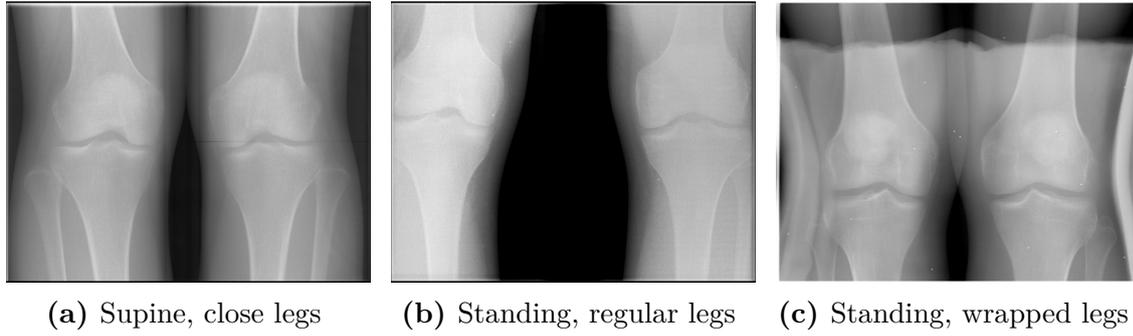
As an extension to JSW measurements, this study aims to investigate the cartilage deformation over time. First results indicate that this deformation can directly be associated to cartilage health [Choi 16a]. However, cartilage tissue is usually not visible in X-ray imaging modalities due to its low contrast w. r. t. the surrounding tissue. To visualize the cartilage surface, direct arthrography is applied, i. e., a contrast medium is injected between femoral and tibial cartilage [Llop 12]. Cartilage thickness can be determined by measuring the distance of the contrast medium's surface to those of tibia and femur. This measurement is even more accurate than a comparable MRI-based method [El K 04]. In addition, the high spatial resolution of  $\sim 200\ \mu\text{m}$  and the relatively fast acquisition time of  $\sim 8\ \text{s}$  is well suited for in vivo weight-bearing measurement of cartilage thickness.

The study protocol is as follows: (1) The volunteers are asked to avoid heavy weight-bearing activities a day prior to the acquisitions. (2) No load is applied to the leg of interest for at least 30 min prior to examination. (3) During this time, the contrast agent is injected into the knee joint and a supine reference scan with high angular resolution is acquired. (4) The subject is guided to the acquisition platform (cf. Fig. 2.3(b)) and weight-bearing scanning starts at the moment load is applied to the leg. (5) Five to six scans are acquired over a period of 15 min to 25 min with a  $0^\circ$  flexion angle. Guided by the reading of a force plate, the subject continuously holds  $\sim 50\%$  of his body weight on the leg of interest. The scanning time points are more densely sampled at the beginning of weight-bearing imaging to account for faster deformations. (6) The study concludes with a supine scan to obtain another high-resolution reference after prolonged and consistent application of load.

Subsequently, the supine and weight-bearing data is reconstructed and femur and tibia are segmented. Inter-scan 3-D/3-D registrations of femur and tibia aligns them to a common coordinate system. The contrast surface between femur and tibia is manually segmented. Finally, the cartilage thickness for a given point on the bone's surface is calculated by the Euclidean distance to its closest point on the contrast agent's surface. The result is a spatially and temporally resolved cartilage strain which could provide further insight into cartilage changes caused by OA.

### 2.3.3 Verification of Treatment Methods for PFPS

As explained in Sec. 2.1.2, one reason for PFPS can be caused by an imbalance of activation times in the vasti muscles. Recent studies report a significant reduction of pain after weakening the vastus lateralis using injections of Botulinum toxin type A [Sing 11, Silb 12]. Chronic PFPS patients with known patellar maltracking are recruited. Treatment includes the injection of Botulinum into 8 locations of the vastus lateralis, followed by six weeks of guided home exercise program to strengthen the quadriceps muscles. Two common pain scores are collected prior to treatment and 6, 12, 18 and 24 weeks after intervention. In this study, the volumetric changes of the vastus medialis and vastus lateralis are examined before and after treatment using supine CBCT imaging. Additionally, changes of the patellar position w. r. t. the femur are evaluated under weight-bearing conditions to investigate the treatments effect on



**Figure 2.4:** Posteroanterior projection images for supine and standing positions. Note that the FOV is too small to cover both legs in a regular stance (see Figure (b)).

patellar maltracking. CBCT imaging is conducted before treatment and six weeks after Botulinium injection.

The imaging pipeline for each session is as follows: (1) A high-resolution supine scan of the thigh is acquired to validate muscle volumes. (2) Then the patellofemoral joint is scanned with a high-resolution supine protocol, followed by (3) two weight-bearing scans with  $0^\circ$  and  $30^\circ$  flexion angle and a specified foot positioning [Choi 14a].

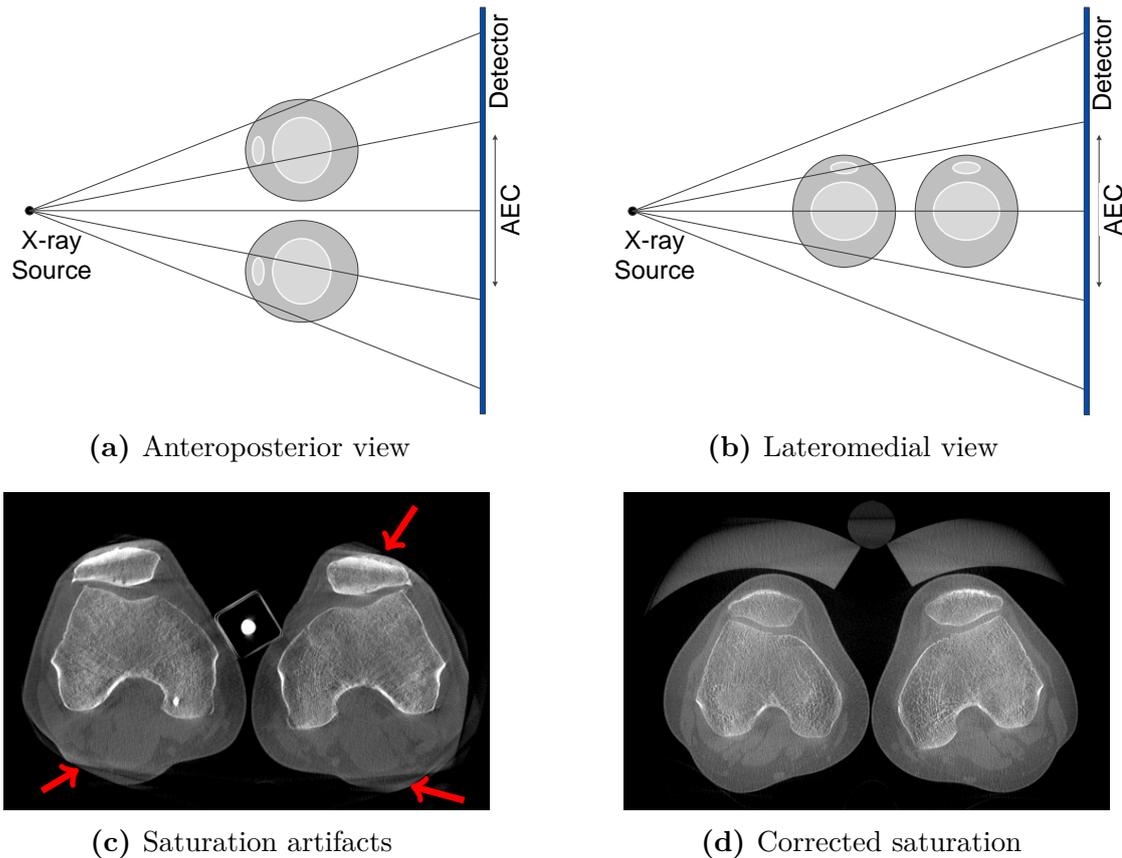
After reconstructing the acquired data, the vastus lateralis and medialis are segmented to compare the ratio of their volumes before and after treatment. Additionally, a 3-D joint coordinate system is extracted from the patellofemoral scans to identify the position of the patella w. r. t. the femur. This yields a total of six parameters that can be used to track post-intervention changes of the patellar position.

## 2.4 Challenges for Image Reconstruction

Weight-bearing imaging of the knee joint poses a series of challenges for 3-D image reconstruction. In the following, the sources and implications of individual problems are outlined in more detail.

### 2.4.1 Lateral Truncation / Limited FOV

Even though C-arm scanners have a considerably higher FOV than dedicated weight-bearing scanners it may not be large enough to cover both legs in an posteroanterior view direction. As a result lateral truncation at posteroanterior projection images occurs that leads to truncation artifacts in the reconstruction domain, e. g., intensity cupping. Fig. 2.4(b) shows a projection image of a subject with a regular standing position where lateral truncation occurred at both legs. Adjustments to the scanning setup can reduce truncation artifacts. For example in supine scanning, truncation is less apparent as the legs are naturally closer together (cf. Fig. 2.4(a)). One possibility is to fasten the two knees together for imaging in a standing position, yet, this comes with the cost of an unnatural stance, which may limit the diagnostic quality (cf. Fig. 2.4(c)). Additionally, changing the system's focus to only one knee of interest could improve image quality as truncation artifacts, such as cupping, have usually a higher influence on the periphery of the FOV. Other methods aim to cor-



**Figure 2.5:** Top: Illustration of erroneous exposure measurements for an AEC in case of knees and a centered measurement region. Bottom: Reconstructions suffering from saturation artifacts and a corrected version using a beam blocking device.

rect truncation algorithmically by extrapolation of the missing data using truncation correction algorithms [Ohne 00, Hsie 04, Denn 13, Maie 12b, Xia 15]. Connected to this thesis, Herbst et al. proposed to extend the FOV in lateral direction using a dynamic detector offset in combination with an extended scan range [Herb 15]. An analytic description between detector offset and necessary scan range is given for elliptical FOVs suitable for knee imaging. However, such type of trajectories are not supported by the robotic C-arm device used for this thesis and image reconstruction requires specialized algorithms that are still under development. Hence, it is not yet possible to apply this promising new technique during data acquisition.

### 2.4.2 Detector Saturation

Modern X-ray systems use an **automatic exposure control (AEC)** system to adjust the acceleration voltage (kVp), exposure times (s), and the tube current (mA) to keep the image quality consistent for all projection images. First, the actual exposure is measured at a central region on the detector, which is then used to adjust kVp and mAs settings accordingly. This procedure works generally well for dense, cylindrical anatomies, yet, it causes problems for acquisitions of knees. Fig. 2.5 shows an illustration of the acquisition for an anteroposterior and a lateromedial view. In case

of an anteroposterior view a large number of X-ray photons will travel almost unaffected from the X-ray source through the legs to the central detector region, leading to a high signal provided to the AEC. In contrast, photons need to pass both legs to reach the central detector region in a lateromedial view. The latter, causes the AEC to increase exposure such that enough photons arrive at the central detector region. Due to the finite number of bits available for discretization, the pixels showing the periphery of the legs, where only low-density tissue is located, are saturated. A global reduction of the requested peak tube voltage could reduce saturation artifacts, yet, this may lead to photon starvation in case of the anteroposterior views.

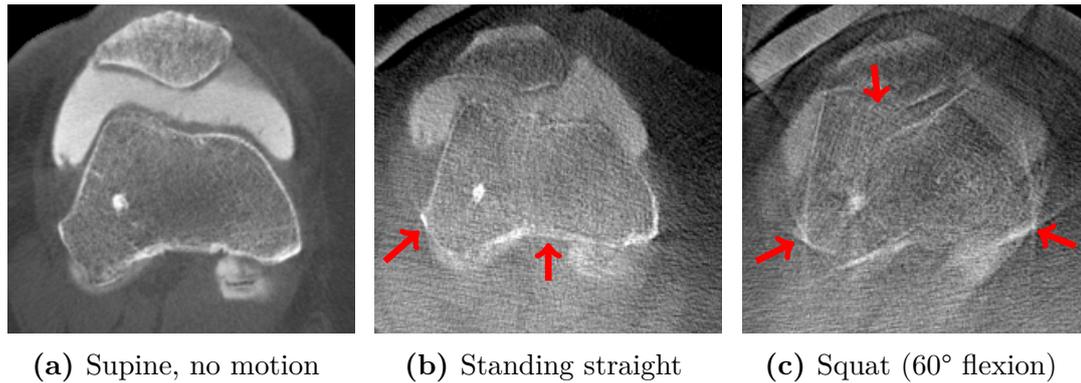
A possible solution is an artificial increase of X-ray absorption at the periphery of the legs, i. e., at the skin. For the first prototype scanning setup this was achieved by wrapping the legs with a layer of plasticine which successfully restored the skin outline in the reconstructed images [Choi 14a]. As the plasticine prohibits a natural standing position, later acquisitions were conducted with specifically designed arc-shaped beam blockers placed at the anterior of the knees (cf. Fig. 2.3(b)). An example for saturation artifacts is shown in Fig. 2.5(c) along with a reconstruction using the beam blocking device in Fig. 2.5(d).

Connected to this thesis, a study was conducted to correct for saturation artifacts as a postprocessing step prior to reconstruction. Therefore, saturated areas are detected and treated as missing data, which are then extrapolated by using circular and ellipsoidal shapes representing leg, femur, and patella [Preu 15]. The ellipsoids' radii, densities and positions are estimated using a cost function based on Helgason-Ludwig consistency conditions (HLCC) [Helg 80, Ludw 66]. Promising results have been reported on a numerical phantom study, however, the method is currently limited to the fan-beam geometry and has not yet been applied to real data.

Recently, we also proposed a saturation correction based on two very low energy projection images used to determine the legs' outline [Choi 16b]. This prior knowledge is then incorporated into an extrapolation method [Xia 15], to recover the overexposed projection data. Further, it has been shown that the exact patient outline can also be determined using an external range imaging camera, thus, avoiding the additional acquisitions entirely [Raus 16].

### 2.4.3 Patient Motion

It has been shown that calibration of the acquisition geometry is necessary to account for C-arm "wobble" and micro motion that deviates from an ideal trajectory [Maie 11a]. However, much larger motion artifacts can be expected from patients rather than unusual scanner motion. At an early phase of the weight-bearing project associated to this thesis, an analysis of the lower body motion was conducted when holding a squat of 60° flexion angle for 20 s [Choi 13]. Translations and flexion angles of nine healthy subjects were recorded using a high-accuracy optical tracking system. Results showed a mean translation of the knee center of 2.22 mm and a mean flexion of 0.49°, where the subject with the largest motion showed a mean translation of 3.43 mm and a flexion of 0.63° as well as a maximum translation of 12.41 mm and flexion of 1.97°. Considering the C-arm systems' maximum isotropic image resolution of ~200 μm, a considerable amount of motion artifacts were expected given the results



**Figure 2.6:** Comparison of reconstructions with supine, standing and squatting patient pose. Motion artifacts are present in the weight-bearing scans and need to be corrected, especially for higher flexion angles.

of the motion analysis. First acquisitions on the prototype system supported these findings, showing clear motion-induced streaking and blurring artifacts especially for squatting positions. Fig. 2.6 shows axial reconstructions of the femur and patella acquired from the same subject in supine, straight standing and squatting position. No visible motion artifacts are present in the supine acquisition seen in Fig. 2.6(a). Note the slight increase of motion artifacts in case of straight standing (cf. Fig. 2.6(b)), compared to the severe artifacts for a squatting pose (cf. Fig. 2.6(c)). For both standing and squatting acquisitions, the diagnostic value is heavily reduced due to blurring and streaking artifacts at bone outlines. To limit weight-bearing acquisition time and thus patient motion, the weight-bearing scans contained only half projection images than the supine scans. This leads to an increased noise level and a lower spatial resolution, as can be seen in Figures 2.6(b) and 2.6(c).

Motion artifacts can be reduced by faster acquisitions or a fixation of the patient, however, faster acquisitions would further reduce image quality and fixation limits the diagnostic value. Hence, motion needs to be corrected programmatically in the course of image reconstruction. A wide variety of motion correction approaches has been introduced in the field of cardiac imaging [Laur 06, Prum 06, Prum 09b, Mull 13, Mull 14c, Schw 13, Unbe 16a], yet, they all require a periodic motion pattern and projection gating, which is not feasible for weight-bearing imaging of knees.

In prior work, knee motion could be successfully corrected using metallic fiducial markers attached to the surface of the knees [Choi 13, Choi 14c]. Markers can be detected in the projection images due to their high X-ray attenuation. These detections can be tracked over time, i. e., over the projection images, to estimate patient motion. A more detailed explanation of the marker-based approach can be found in Sec. 3.4.2.

## 2.5 Conclusion

This chapter gives an introduction into the anatomy of the knee joint with focus on the bones, ligaments, and tendons as well as the muscles involved in the locomotor system of the lower body. OA and PFPS are the most common knee joint disorders

with generally high incident rates. Sec. 2.1.3 shows that imaging under weight-bearing conditions is crucial for a more precise diagnosis of OA, but also to accurately identify patellar maltracking known to be a common cause of PFPS.

2-D radiography is regularly applied under weight-bearing conditions, yet, the complex 3-D structure of the knee may not be accurately covered by 2-D projection images only [Choi 14a]. Current 3-D weight-bearing imaging of the knee includes MRI using special weight-application devices but also open-bore MRI. Whereas the first is still limited to an unnatural supine patient position, the latter is not commonly available in clinics. Further, given the long acquisition times of MRI, patient motion causes a substantial problem. This thesis is based on data acquired from CBCT systems that allow 3-D weight-bearing imaging in an almost unsupported position with reasonably fast acquisition times and high spatial resolution (see Sec. 2.2.2).

As shown in Sec. 2.4, CBCT weight-bearing imaging poses several challenges for image reconstruction, including data truncation, detector saturation but foremost patient motion. Correction of patient motion during image reconstruction is mandatory due to the large decline in image quality. In prior work, metallic fiducial markers have been used to estimate motion parameters, yet, the approach required manual interaction and lead to an increased acquisition time [Choi 13, Choi 14c].

# Image Reconstruction and Motion Correction in CBCT

3.1	Mathematical Definitions .....	23
3.2	CBCT Imaging .....	24
3.3	Motion Correction .....	32
3.4	Motion Estimation Methods .....	33
3.5	Motion Compensation Methods .....	42
3.6	Conclusion .....	44

This chapter provides the theoretical foundation on image reconstruction and motion correction algorithms in CBCT. Starting from the related fan-beam geometry, Sec. 3.2 gives basic knowledge about CBCT imaging, its acquisition geometry and associated reconstruction algorithms. The definition and classification of motion correction algorithms can be obtained from Sec. 3.3. Subsequently, Sec. 3.4 provides a detailed review of state-of-the-art motion estimation methods, followed by an overview of motion compensation algorithms in Sec. 3.5. Sec. 3.6 concludes this chapter with a summary of methods and their effects to CBCT weight-bearing imaging of the knee joint.

## 3.1 Mathematical Definitions

Some important functions and variables in this section are generically defined in the  $N$  dimensional space, to allow later use with fan- ( $N = 2$ ) and cone-beam ( $N = 3$ ) geometries. If not mentioned otherwise, the following mathematical notation style is used in this thesis. We use upper-case, bold letters (e.g.,  $\mathbf{A} \in \mathbb{R}^{N \times N}$ ) to define matrices and lower-case, bold letters (e.g.,  $\mathbf{a} \in \mathbb{R}^N$ ) to denote vectors. Elements of matrices and vectors are denoted by the same letter, but in lower-case and regular font with a subscript describing the element's location. For example,  $a_i$  is the  $i$ -th element of  $\mathbf{a}$  and  $a_{ij}$  is the element of matrix  $\mathbf{A}$  corresponding to the  $i$ -th row and  $j$ -th column. Lower-case, regular font letters (e.g.,  $a \in \mathbb{R}$ ) can also refer to scalar values or constants. Upper-case, nonbold letters (e.g.,  $\Omega$ ) are used for mathematical sets or for geometrical parameters of the acquisition geometries (e.g.,  $L$ ). Integer indices, e.g., subscripts for vectors, are lower-case, nonbold letters that may also be defined by a range (e.g.,  $i \in [1, \dots, N]$ ). The range of indices starts with 1 and ends with the maximum integer number also given by an upper-case, nonbold letter. Functions may be introduced by definition of their input and output space,

e. g.,  $\mathbf{s}(\mathbf{q}) : \mathbb{R}^3 \rightarrow \mathbb{C}^2$  represents a function  $\mathbf{s}$  that maps a vector  $\mathbf{q} \in \mathbb{R}^3$  to a complex valued vector of length two, i. e.,  $\mathbb{C}^2$ . Additional superscripts and subscripts can be used, e. g., to denote dependencies or indices of 3-D structures, but will be introduced accordingly. Throughout this thesis, rotation matrices are defined by rational numbers, i. e.,  $\mathbf{R} \in \mathbb{R}^{N \times N}$ . However, all rotation matrices are required to fulfill  $\mathbf{R}\mathbf{R}^\top = \mathbf{I}_N$  and that  $|\mathbf{R}| = 1$ , where  $\mathbf{I}_N$  is an  $N$ -dimensional identity matrix and  $|\mathbf{R}|$  is the determinant.

## 3.2 CBCT Imaging

This section introduces essential information about CBCT imaging, including the imaging geometry, the imaging process and basic reconstruction algorithms. Many of the presented concepts build on the books of Buzug [Buzu08], Kak and Slaney [Kak01] and Zeng [Zeng09].

### 3.2.1 X-ray Physics

Fundamental X-ray physics can be separated into three categories, i. e., X-ray generation, matter interaction and detection. For a detailed review of all categories we refer to the book of Buzug [Buzu08, pp. 15-72].

X-rays belong to the class of electromagnetic waves and can be classified by their energy  $E_p$  or wavelength  $\lambda_p$ . X-ray photons are produced inside the X-ray tube, by collision of accelerated electrons with an anode material. A filament is used as electron source, followed by an electric field of high intensity to accelerate the electrons. The interaction of electrons with the anode material releases X-ray photons with a certain energy  $E_p$ . In medical CBCT scanners the individual X-ray energies differ, thus, building a spectrum of energies. The shape of the X-ray spectrum is based on the acceleration voltage, as well as the anode material. The number of emitted X-ray photons  $N_0$  depends on the exposure time and the electric current produced by the electron beam.

When an X-ray photon interacts with matter it may get absorbed, scattered or it passes unaffected. The total amount of attenuation depends on object properties and  $E_p$ . In the monochromatic case, i. e., for a fixed energy, this may be summarized by the object function  $f(\mathbf{x}) : \mathbb{R}^N \rightarrow \mathbb{R}$ , which provides the object's linear attenuation function given the fixed photon energy and location  $\mathbf{x} \in \mathbb{R}^N$ . We define the location vector to be  $\mathbf{x} = (x \ y)^\top$  for  $N = 2$  and  $\mathbf{x} = (x \ y \ z)^\top$  for  $N = 3$ . Let us assume that the photons propagate on a straight line, parametrized by the X-ray source position  $\mathbf{c} \in \mathbb{R}^N$  and a unit vector  $\mathbf{r} \in \mathbb{S}^N$ , corresponding to the direction of the line. Then, the number of X-ray photons  $N_p$ , passing unaffected through the object is given by Beer-Lambert's law

$$N_p = N_0 e^{-\int_0^\infty f(\mathbf{c} + \beta \mathbf{r}) d\beta} , \quad (3.1)$$

where the integral is also known as the X-ray transform. Considering that  $E_p \in [0, E_{max}]$  and that the distribution  $N_0(E_p)$  is known, we can specify Beer-Lambert's law for polychromatic X-ray energies

$$N_p = \int_0^{E_{max}} N_0(E_p) \left( e^{-\int_0^\infty f(\mathbf{c} + \beta \mathbf{r}; E_p) d\beta} \right) dE_p . \quad (3.2)$$

Note, that in this case the function  $f$  also depends on the energy. However, most reconstruction algorithms described in literature are based on the assumption of a monochromatic attenuation. Polychromatic attenuation is also not of main interest in this thesis. Thus, simulations have been done with a fixed monochromatic energy and real data has been converted to a comparable monochromatic energy.

The goal of image reconstruction is to recover  $f(\mathbf{x})$  from a set of line integral measurements. Line integrals are obtained from Eq. (3.1) by

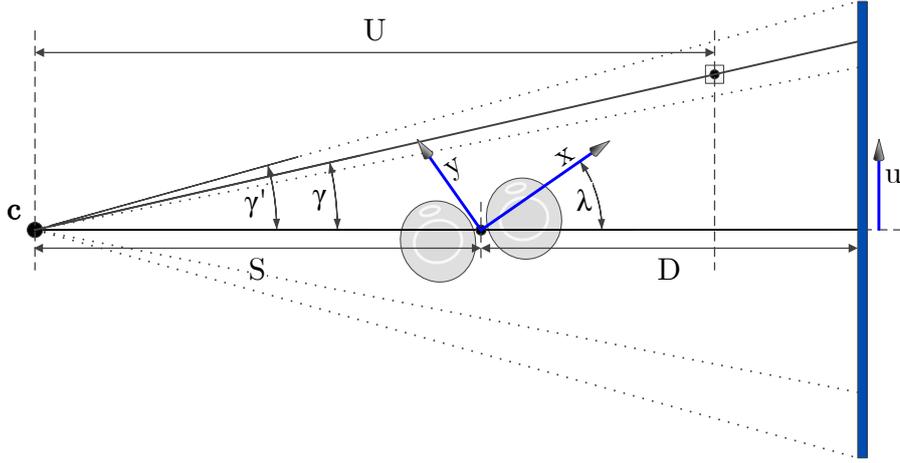
$$-\log \frac{N_p}{N_0} = \int_0^\infty f(\mathbf{c} + \beta \mathbf{r}) d\beta . \quad (3.3)$$

Unless explicitly stated otherwise, in the remainder of this thesis all references made to projection and sinogram domain refer to line integral data.

### 3.2.2 Fan-Beam CT

In the following the fan-beam geometry and algorithms are described in more detail. The dimensionality of previously introduced functions and variables is  $N = 2$ . This thesis contains data and experiments based on fan-beam as well as CBCT geometry. Note, that both approaches find their roots in the practically less relevant parallel-beam geometry and its associated Radon transform [Helg 11]. Yet, a detailed description of the parallel-beam geometry is omitted in this work. More information on the connection and equalities from fan- and cone- to the parallel-beam geometry may be obtained from Buzug [Buzu 08].

The 2-D fan-beam acquisition geometry can be considered an extension to the conventional parallel-beam geometry and has a close connection to 3-D CBCT. It is directly connected to parallel-beam imaging. Thus, line integral values can be converted between fan and parallel geometries, given data completeness. The imaging geometry is shown in Fig. 3.1. It depicts the X-ray source position  $\mathbf{c}$  on the left and the detector on the right. The imaging process describes the transform of the 2-D image domain spanned by the  $x$  and  $y$  coordinate system, to the 2-D sinogram domain described by the rotation angle  $\lambda$  and the detector position  $u$ . The source  $\mathbf{c}$  rotates around the isocenter with an SID of  $S$  and a rotation angle  $\lambda$ . The detector rotates opposite to the source with a detector-isocenter-distance (DID) given by  $D$ . Let us also define the SDD to be  $F = S + D$ . X-rays emitted from  $\mathbf{c}$  travel on a linear path to the detector and interact with the object based on Eq. (3.1). The line directions  $\mathbf{r}$  may be described by the fan angle  $\gamma = \arctan\left(\frac{u}{F}\right) \in [-\gamma', \gamma']$  w. r. t. the central ray direction of the system. In theory the maximum fan angle is bounded by  $\gamma' = \frac{\pi}{2}$ , however, in a practical application  $\gamma'$  is either limited by the finite size of the object



**Figure 3.1:** Fan-beam scanning geometry. The source  $\mathbf{c}$  and the detector rotate around the isocenter and acquire line integrals through the object.

or the detector. Using the X-ray transform from Eq. (3.1), the fan-beam sinogram  $g(\lambda, u) : \mathbb{R}^2 \rightarrow \mathbb{R}$  may be defined by

$$g(\lambda, u) = \int_0^\infty f(\mathbf{c}(\lambda) + \beta \mathbf{r}(\lambda, u)) d\beta \quad (3.4)$$

$$\mathbf{c}(\lambda) = \mathbf{R}(\lambda) \begin{pmatrix} -S \\ 0 \end{pmatrix} \quad (3.5)$$

$$\mathbf{r}(\lambda, u) = \mathbf{R}\left(\lambda + \arctan\left(\frac{u}{F}\right)\right) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (3.6)$$

$$\mathbf{R}(\phi) = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix}, \quad (3.7)$$

where  $\mathbf{c}(\lambda) : \mathbb{R} \rightarrow \mathbb{R}^N$  and  $\mathbf{r}(\lambda, u) : \mathbb{R}^2 \rightarrow \mathbb{S}^N$  compute the source location and X-ray direction w. r. t.  $\lambda$  and  $u$ . Further,  $\mathbf{R}(\phi) : \mathbb{R} \rightarrow \mathbb{R}^{2 \times 2}$  provides a 2-D rotation matrix based on the angle  $\phi$ .

A scan range of  $\lambda \in [0, 2\pi[$  corresponds to a full scan which measures all ray paths twice. To obtain the minimum set of complete data the trajectory needs to satisfy Tuy's conditions [Tuy 83]. It states that all possible lines through a point  $\mathbf{x}$  belonging to the object, need to intersect the source trajectory. In case of the fan-beam geometry it can be shown that this is the case for  $\lambda \in [0, > (\pi + 2\gamma')[$  also known as a short scan [Park 82].

## Reconstruction

The most common image reconstruction algorithm for fan-beam CT is **filtered back-projection (FBP)**. It is derived from the parallel-beam geometry by exploiting identities between parallel and fan geometry [Kak 01]. It is given by three individual steps, namely cosine weighting (cf. Eq. (3.8)), (2) ramp filtering (cf. Eq. (3.9)) and

(3) backprojection with distance weighting (cf. Eq. (3.10)). The following derivation result is obtained from [Kak 01] and given by

$$\tilde{g}(\lambda, u) = g(\lambda, u) \cos(\gamma) = g(\lambda, u) \frac{F}{\sqrt{F^2 + u^2}} \quad (3.8)$$

$$\hat{g}(\lambda, u) = \tilde{g}(\lambda, u) * h_{\text{ramp}}(u) , \quad (3.9)$$

where  $h_{\text{ramp}}(u)$  is the ramp filter kernel and  $\tilde{g}$  and  $\hat{g}$  correspond to the sinogram after cosine weighting and ramp filtering, respectively. Further,  $[*]$  is the convolution along  $u$  direction. The backprojection step projects the sinogram values back to the image domain and is given by

$$f(\mathbf{x}) = \frac{1}{2} \int_0^{2\pi} \frac{1}{U(\mathbf{x}, \lambda)^2} \hat{g}(\lambda, F \cos(\gamma(\mathbf{x}, \lambda))) d\lambda \quad (3.10)$$

$$U(\mathbf{x}, \lambda) = \left( \frac{\mathbf{c}(\lambda)^\top}{-S} \right) (\mathbf{x} - \mathbf{c}(\lambda)) \quad (3.11)$$

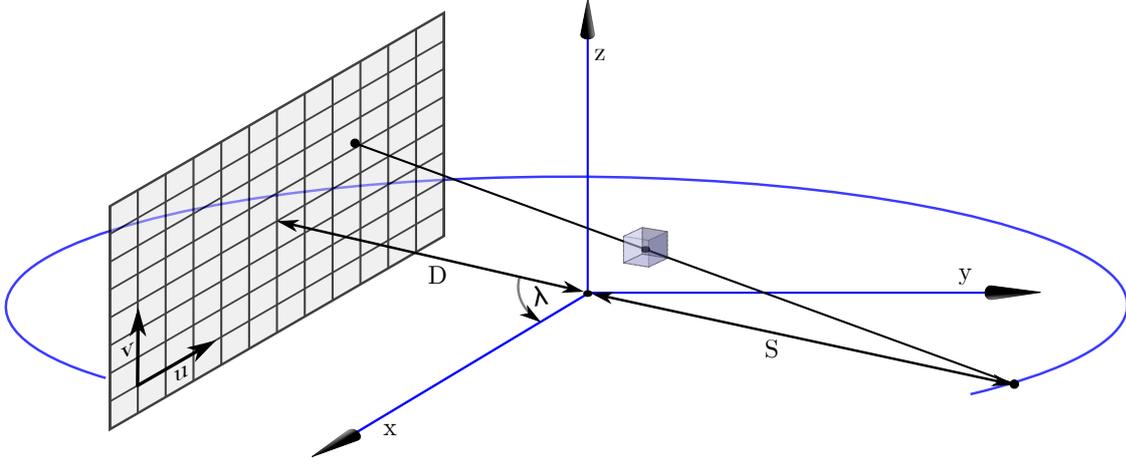
$$\cos(\gamma(\mathbf{x}, \lambda)) = \frac{U(\mathbf{x}, \lambda)}{\|(\mathbf{x} - \mathbf{c}(\lambda))\|_2} , \quad (3.12)$$

where  $U(\mathbf{x}, \lambda) : \mathbb{R}^{N+1} \rightarrow \mathbb{R}$  is the projection of the vector connecting source and world point  $(\mathbf{x} - \mathbf{c}(\lambda))$  onto the central ray and is also known as depth of  $\mathbf{x}$ . Further,  $\gamma(\mathbf{x}, \lambda) : \mathbb{R}^{N+1} \rightarrow \mathbb{R}$  is the angle between the principle ray and vector  $(\mathbf{x} - \mathbf{c}(\lambda))$ . Its cosine value can be computed by the ratio of the depth, to the length of vector  $(\mathbf{x} - \mathbf{c}(\lambda))$ . The division by  $U(\mathbf{x}, \lambda)^2$  is called distance weighting and is specific to fan-beam reconstruction.

Eq. (3.10) describes the backprojection formula for a full scan. The redundancy in the data can easily be corrected with a global multiplication by  $\frac{1}{2}$ , as all rays are sampled twice. In case of a short scan only some rays are sampled twice whereas others are sampled only once. Parker weighting can be used to reduce the weight for oversampled data and is applied prior to the FBP algorithm [Park 82]. Additionally, the range of the integral in Eq. (3.10) reduces from  $2\pi$  to the short-scan range  $(\pi + 2\gamma')$ . In practice, an even smaller angular range (i. e., super-short-scan) is possible, yet, requires specialized reconstruction algorithms [Noo 02]. The recently developed Riess weighting accounts for missing data by using specialized redundancy weights in angular direction [Ries 13].

### 3.2.3 CBCT

The cone-beam geometry with an FP detector is shown in Fig. 3.2. It has a close relation to the fan-beam geometry, but uses a 2-D instead of a 1-D detector. Whenever, it is referred to a cone-beam geometry we assume a dimension of  $N = 3$ . The additional detector direction is denoted as  $v$  direction, whereas the remaining geometric parameters remain unchanged w. r. t. the fan-beam geometry explained in Sec. 3.2.2. The world coordinate system is spanned by the  $x$ -,  $y$ - and  $z$ -axis, i. e.,  $\mathbf{e}_x = (1 \ 0 \ 0)^\top$ ,  $\mathbf{e}_y = (0 \ 1 \ 0)^\top$  and  $\mathbf{e}_z = (0 \ 0 \ 1)^\top$ . Without loss of generality the rotation axis is defined to be  $\mathbf{e}_z$ . Let us further define that the principle ray intersects the detector at  $v = 0$  and  $u = 0$ , corresponding to a centered detector. Thus,



**Figure 3.2:** CBCT scanning geometry. The X-ray source orbits around the object on a circular trajectory acquiring 2-D projection images for rotation angles  $\lambda$ .

for  $v = 0$  the cone-beam geometry degenerates to the fan-beam geometry which is also known as central slice.

Similar to fan-beam CT, the image formation process of the cone-beam geometry can be defined by using the X-ray transform

$$p(\lambda, u, v) = \int_0^{\infty} f(\mathbf{c}(\lambda) + \beta \mathbf{r}(\lambda, u, v)) d\beta \quad (3.13)$$

$$\mathbf{c}(\lambda) = -S \mathbf{R}_z(\lambda) \mathbf{e}_x \quad (3.14)$$

$$\mathbf{r}(\lambda, u, v) = \mathbf{R}_z(\lambda + \gamma) \mathbf{R}_y(\kappa) \mathbf{e}_x \quad (3.15)$$

$$\mathbf{R}_z(\phi^z) = \begin{pmatrix} \cos(\phi^z) & -\sin(\phi^z) & 0 \\ \sin(\phi^z) & \cos(\phi^z) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.16)$$

$$\mathbf{R}_y(\phi^y) = \begin{pmatrix} \cos(\phi^y) & 0 & \sin(\phi^y) \\ 0 & 1 & 0 \\ -\sin(\phi^y) & 0 & \cos(\phi^y) \end{pmatrix}, \quad (3.17)$$

where  $\mathbf{R}_z(\phi^z) : \mathbb{R} \rightarrow \mathbb{R}^{3 \times 3}$  and  $\mathbf{R}_y(\phi^y) : \mathbb{R} \rightarrow \mathbb{R}^{3 \times 3}$  are 3-D rotation matrices around  $z$ - and  $y$ -axis, respectively. Further,  $\kappa = \arctan\left(\frac{v}{F}\right) \in [-\kappa', \kappa']$  is the cone-angle in  $v$  direction,  $\mathbf{r}(\lambda, u, v) : \mathbb{R}^3 \rightarrow \mathbb{S}^N$  is the ray direction in 3-D and  $p(\lambda, u, v) : \mathbb{R}^3 \rightarrow \mathbb{R}$  is the function that represents the acquired projection images. Note that in theory  $\kappa' = \frac{\pi}{2}$ , however, it is usually limited by the object or detector extent.

It is known that the circular source trajectory does only fulfill Tuy's condition for data completeness at the central slice, but not for off-center slices, causing cone-beam artifacts [Buzu 08, p. 458]. Full sampling is possible using modified trajectories, such as circle-plus-arc, circle-plus-line or reverse-helix trajectories [Hsie 13, Yu 13], yet, this comes with the price of an increased complexity of the scanner and image reconstruction algorithms.

### Reconstruction

The Feldkamp-Davis-Kress (FDK) algorithm is most commonly used for image reconstruction in current clinical CBCT [Feld 84]. FDK is an intuitive but heuristic extension to the fan-beam FBP and can be implemented efficiently leading to very fast reconstruction results [Wu 16]. All CBCT reconstruction algorithms derived in this thesis build on the FDK algorithm. The algorithm is given by the same three steps as for the fan-beam FBP, i. e., 3-D cosine weighting (cf. Eq. (3.18)), row-wise ramp filtering (cf. Eq. (3.19)) and a 3-D backprojection (cf. Eq. (3.20)). The formulas for cosine weighting and ramp filtering are given by

$$\tilde{p}(\lambda, u, v) = p(\lambda, u, v) \cos(\gamma) \cos(\kappa) = p(\lambda, u, v) \frac{F}{\sqrt{F^2 + u^2 + v^2}} \quad (3.18)$$

$$\hat{p}(\lambda, u, v) = \tilde{p}(\lambda, u, v) * h_{\text{ramp}}(u) \quad , \quad (3.19)$$

where  $h_{\text{ramp}}(u)$  is the ramp filter kernel and  $[*]$  denotes a convolution over  $u$  direction. Further,  $\tilde{p}$  and  $\hat{p}$  correspond to the projection images after cosine weighting and ramp filtering, respectively. The backprojection is given by

$$f(\mathbf{x}) = \frac{1}{2} \int_0^{2\pi} \frac{1}{U(\mathbf{x}, \lambda)^2} \hat{p}(\lambda, F \cos(\gamma(\mathbf{x}, \lambda)), F \cos(\kappa(\mathbf{x}, \lambda))) d\lambda \quad (3.20)$$

$$U(\mathbf{x}, \lambda) = \mathbf{e}_x^\top \mathbf{R}_z(\lambda)^\top (\mathbf{x} - \mathbf{c}(\lambda)) \quad (3.21)$$

$$\cos(\gamma(\mathbf{x}, \lambda)) = \frac{U(\mathbf{x}, \lambda)}{\left\| \begin{pmatrix} \mathbf{e}_x & \mathbf{e}_y \end{pmatrix}^\top \mathbf{R}_z(\lambda)^\top (\mathbf{x} - \mathbf{c}(\lambda)) \right\|_2} \quad (3.22)$$

$$\cos(\kappa(\mathbf{x}, \lambda)) = \frac{U(\mathbf{x}, \lambda)}{\left\| \begin{pmatrix} \mathbf{e}_x & \mathbf{e}_z \end{pmatrix}^\top \mathbf{R}_z(\lambda)^\top (\mathbf{x} - \mathbf{c}(\lambda)) \right\|_2} \quad , \quad (3.23)$$

where  $U(\mathbf{x}, \lambda)$  is again the depth of  $\mathbf{x}$  based on  $\mathbf{c}$  and the principle ray direction. Further,  $\gamma(\mathbf{x}, \lambda)$  and  $\kappa(\mathbf{x}, \lambda) : \mathbb{R}^{N+1} \rightarrow \mathbb{R}$  are the angles between the principle ray and the projection of  $(\mathbf{x} - \mathbf{c}(\lambda))$  onto the rotated  $x$ - $y$  and  $x$ - $z$  plane, respectively. The FDK algorithm may as well be extended to a short-scan acquisition by a row-wise application of Parker weights [Park 82] prior to the reconstruction. Note, that FDK is an approximate algorithm for off-center planes, yet, reconstruction artifacts are usually dominated by the missing data due to the circular trajectory [Feld 84]. Theoretically, exact reconstruction algorithms have been presented by Grangeat [Gran 91] and Katsevich [Kats 02]. However, their application to circular CBCT is of little benefit, as the acquired data does generally not meet Tuy's completeness condition.

### 3.2.4 Iterative Reconstruction

FBP as well as FDK are the most common reconstruction algorithms and also the basis for the reconstruction algorithms derived in this thesis. Other, reconstruction methods include algebraic and iterative reconstruction methods, where the inversion is solved by iterative methods. This usually requires a regularization term based on assumptions on the object function  $f(\mathbf{x})$ , such as total variation

[Sidk 06, Wu 12, Huan 16]. Iterative reconstruction techniques are not used in this thesis and therefore omitted here, yet, the interested reader is referred to the following prior work [Chen 08, Sidk 08, Pan 09, Hara 09, Tang 09].

### 3.2.5 Discretization and Projection Matrices

The projection and reconstruction algorithms in Sections 3.2.2 and 3.2.3 are based on several geometric properties, e. g., perfect alignment of rotation axis to the world coordinate's  $z$ -axis, perfect orthogonality of detector  $u$ - and  $v$ -axis or an ideal intersection of the principle ray in the center of the detector. Also, the presented algorithms are defined in the continuous domain. However, detector and angular sampling are finite for real world CBCT scanners. Also, a more generic geometric description is required to model imperfect detector coordinate systems, but also deviations in the X-ray source trajectory. In this thesis we used the concepts of projective geometry, i. e., projection matrices and homogeneous coordinates, to model the mapping of 3-D world coordinates to 2-D detector coordinates. Many of the following principles are based on the book of Hartley and Zisserman [Hart 04].

Let us first define  $\Delta\lambda$  to be the angular spacing associated to rotation angle  $\lambda$  as well as  $\Delta u$  and  $\Delta v$  to be the discrete pixel spacing in detector  $u$  and  $v$  direction, respectively. Further, the dimensions of the detector are given by  $I$  and  $J$  in  $u$ - and  $v$ -axis, respectively, while the number of projections is given by  $K$ .

We also define that the homogeneous representation of a vector  $\mathbf{a} \in \mathbb{R}^N$  is given by  $\underline{\mathbf{a}} = \begin{pmatrix} \mathbf{a}^\top & 1 \end{pmatrix}^\top$ . To revert the homogeneous mapping we introduce the function  $h(\underline{\mathbf{a}}) : \mathbb{P}^N \rightarrow \mathbb{R}^N$  that performs a division by the last component

$$h(\underline{\mathbf{a}}) = \frac{1}{\underline{a}_{N+1}} \begin{pmatrix} \underline{a}_1 & \cdots & \underline{a}_n \end{pmatrix}^\top. \quad (3.24)$$

Note that we use  $h$  to map a point from projective space to the Euclidean space. The presented notation is not defined in case  $\underline{a}_{N+1} = 0$ , which would be equal to a point in infinity. However, it can be ensured that for all practical applications used in this thesis it holds that  $\underline{a}_{N+1} \neq 0$ .

The projection matrix  $\mathbf{P} \in \mathbb{R}^{3 \times 4}$  encodes six extrinsic parameters, that define rotation  $\mathbf{R} \in \mathbb{R}^{3 \times 3}$  and translation  $\mathbf{t} \in \mathbb{R}^3$  of the world coordinate system. In addition, it holds five intrinsic parameters in the matrix  $\mathbf{K} \in \mathbb{R}^{3 \times 3}$ , that determine the mapping of world to detector coordinate system. Eq. (3.25) defines how  $\mathbf{P}$  is built given  $\mathbf{K}$ ,  $\mathbf{R}$  and  $\mathbf{t}$ .

$$\mathbf{P} = \mathbf{K} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{pmatrix} \quad (3.25)$$

The  $\mathbf{K}$  matrix may be defined by

$$\mathbf{K} = \begin{pmatrix} \frac{F}{\Delta u} & k_s & c_u \\ 0 & \frac{F}{\Delta v} & c_v \\ 0 & 0 & 1 \end{pmatrix}, \quad (3.26)$$

where  $k_s$  is a parameter to model skewness between detector  $u$ - and  $v$ -axis and  $c_u$  and  $c_v$  denote the location of the principle point, i. e., the intersection between principle ray and detector plane given, in pixel coordinates.

After applying  $\mathbf{R}$  and  $\mathbf{t}$  to the X-ray source  $\mathbf{c}$ , it is by definition located at the world coordinate origin and points in direction  $\mathbf{e}_z$  [Hart 04]. However, the circular trajectory described in Sec. 3.2.3 is defined by the rotation axis  $\mathbf{e}_z$ , an initial source location  $-S \mathbf{e}_x$  and a view direction  $\mathbf{e}_x$ . Thus, a rotation matrix  $\mathbf{A} \in \mathbb{R}^{3 \times 3}$  is introduced that maps the initial location of the x-ray source from  $x$ -axis to the  $z$ -axis, the view direction from  $\mathbf{e}_x$  to  $\mathbf{e}_z$ , and that aligns the detector  $v$ -axis with the rotation axis. Hence, the rotations and translations for our circular trajectory can be created by

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad (3.27)$$

$$\mathbf{R} = \mathbf{A} \mathbf{R}_z(\lambda)^\top \quad (3.28)$$

$$\mathbf{t} = \mathbf{A} S \mathbf{e}_x = S \mathbf{e}_z . \quad (3.29)$$

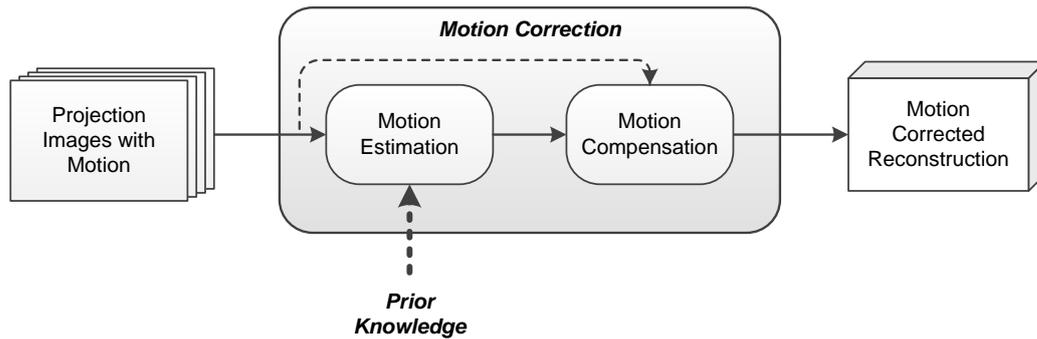
Note, that the null space of  $\mathbf{P}$  is related to the X-ray source by  $\ker(\mathbf{P}) = \underline{\mathbf{c}}$ , where  $\ker(\cdot)$  denotes the null space computation. Further, the rows of  $\mathbf{P}$  may be interpreted as planes, where the first two rows correspond to planes that contain  $\mathbf{c}$  and also the detector  $v$ - and  $u$ -axis, respectively. The plane described by the last row contains  $\mathbf{c}$  and is parallel to the detector plane [Hart 04, pp. 158ff.].

In real world systems, projection matrices are computed by system calibration, using a calibration phantom that allows identification of 3-D positions from 2-D measurements [Maie 11a]. For the remainder of this thesis the projection matrix corresponding to the  $k$ -th projection image is referred by  $\mathbf{P}_k$ , with its elements  $\mathbf{K}_k$ ,  $\mathbf{R}_k$  and  $\mathbf{t}_k$ , where  $k \in [1, \dots, K]$ . Note that projection matrices map 3-D points in world coordinates to 2-D pixel coordinates  $\mathbf{u} = \begin{pmatrix} u_1 & u_2 \end{pmatrix} \in \mathbb{R}^2$ . In contrast, the previously introduced coordinates  $u$  and  $v$  are defined in mm. Let us define projection image functions  $p_k(\mathbf{u}) : \mathbb{R}^2 \rightarrow \mathbb{R}$  for the  $k$ -th projection image. Their connection to the continuous projection data  $p(\lambda, u, v)$  can be defined by

$$p_k(\mathbf{u}) = p((k-1)\Delta\lambda, u, v) \quad (3.30)$$

$$\mathbf{u} = h \left( \mathbf{K}_k \begin{pmatrix} u & v & F \end{pmatrix}^\top \right) . \quad (3.31)$$

The images  $p_k(\mathbf{u})$  are usually discretized on a pixel grid with spacing  $\Delta u$  and  $\Delta v$ . We will use indices  $i \in [1, \dots, I]$  and  $j \in [1, \dots, J]$  to identify pixels in  $u_1$  and  $u_2$  direction, respectively. Despite discretization of the images, the continuous representation  $p_k(\mathbf{u})$  of the projection images is used in this thesis. Data readout at subpixel locations can be done using an interpolation method. If not mentioned otherwise, bilinear interpolation is used as interpolation method [Keck 14].



**Figure 3.3:** Motion correction consists of a motion estimation and motion compensation step. These steps can appear in a single algorithm or separated. They may also be executed repeatedly in an iterative manner.

Using projection matrices, we can now compute the discrete version of the continuous backprojection shown in Eq. (3.20). It is given by

$$\underline{\mathbf{u}} = \mathbf{P}_k \mathbf{x} \quad (3.32)$$

$$f(\mathbf{x}) = \frac{1}{2} \sum_{k=1}^K \frac{1}{(\underline{u}_3)^2} \hat{p}_k(h(\underline{\mathbf{u}})) , \quad (3.33)$$

where we use the convenient property that the third component of  $\underline{\mathbf{u}}$  equals the depth of  $\mathbf{x}$  w.r.t. the X-ray source and the principle ray. Note, that this property holds for our definition of  $\mathbf{P}$ , but depends generally on the scaling of the projection matrix and the encoded view direction [Hart 04, pp. 162f.]. Similar to the continuous case,  $\hat{p}_k$  refers to the projection images  $p_k$ , after application of cosine weighting and ramp filtering. For discretization of cosine weighting and ramp filtering we refer to the book of Kak & Slaney [Kak 01].

### 3.3 Motion Correction

Motion correction, motion estimation and motion compensation are often used indiscriminately in literature. In the scope of this thesis, **motion correction** consist of two building blocks, i. e., motion estimation and motion compensation. It refers to a reconstruction algorithm that exploits information about the underlying object or scanner motion to reduce motion artifacts in the reconstructed image. Fig. 3.3 depicts the key components of motion correction.

During **motion estimation**, parametrized motion models or even nonparametric motion fields are estimated. This is usually based on a cost function that takes the projection images containing the irregular motion, but also prior knowledge as an input. Motion estimation typically results in 3-D rigid or nonrigid deformations defined for individual or sets of projection images, yielding temporally resolved, 4-D motion data. The output of the motion estimation step is a parameter vector  $\boldsymbol{\alpha} \in \mathbb{R}^L$  that holds  $L$  parameters used to define the motion models. Alternatively, it encodes the deformation directly containing 3-D translation vectors for each reconstructed voxel and time point.

**Motion compensation** refers to the process of incorporating the estimated motion into the reconstruction process. The goal of motion compensation can be a static 3-D reconstruction for a specific point in time or a time frame. Alternatively, 3-D reconstructions can be conducted for a series of time points, yielding spatially and temporally resolved data, also known as 4-D CT reconstruction.

Motion correction is essentially very similar to geometric calibration of the system parameters. It can be shown that motion of the X-ray source, e. g., due to scanner “wobble”, may be modeled by object motion [Hart 04]. By definition, calibration and motion correction are treated separately in this thesis. It is assumed that the calibration of the system is performed prior to the scan, yielding a set of calibrated projection matrices  $P_i$ .

## 3.4 Motion Estimation Methods

We separate between three groups of “prior knowledge” for motion estimation algorithms. First, **surrogate signals** that require simultaneous acquisition of the surrogate (e. g., an **Electrocardiogram (ECG)** signal for coronary imaging) or dedicated alterations to the image content (e. g., fiducial-markers for motion tracking). Second, the usage of a **prior reference image** in the projection (2-D), reconstruction (3-D) or temporally resolved reconstruction (4-D) domain. And third, **without reference and surrogate signals**, but possibly with assumptions imposed on the motion model or the reconstructed image. The main differences are:

- Algorithms using **surrogate signals** usually interfere with the acquisition process, e. g., by attaching fiducials or a surrogate acquisition.
- Algorithms based on prior image data do not interfere with the acquisition of motion-corrupted data. However, depending on the acquisition protocol and patient history, additional acquisitions may be required.
- Algorithms that work **without reference and surrogate signals** do not require additional image data and do not cause alterations in the acquisition process.

In consequence, the amount of data decreases from group one to three, yet, this may cause limited accuracy or requires a reduction of complexity in the estimated motion field. A higher flexibility during clinical acquisition may especially be given in the third group.

### 3.4.1 Motion Models

Before the object motion can be estimated, a mathematical description that models the motion needs to be defined. Rigid and nonrigid, but also parametric and nonparametric motion models exist. Parametric models exist in both rigid and nonrigid domain, whereas nonparametric motion is usually only applied in the field of nonrigid image registration.

### Rigid Motion in 2-D and 3-D

We define rigid motion to be limited to translations and rotations, resulting in three degrees of freedom for 2-D images and six degrees of freedom for 3-D volumes. In that sense, rigid motion can be conveniently integrated to CBCT imaging when using the concept of projection matrices introduced in Sec. 3.2.5. Rigid motion of the object can be modeled by right multiplication of a projective transformation matrix to the projection matrix, whereas rigid detector motion may be modeled by a left multiplication. Let us define

$$\mathbf{T}_k(\boldsymbol{\alpha}) = \begin{pmatrix} \mathbf{R}_k(\boldsymbol{\alpha}) & \mathbf{t}_k(\boldsymbol{\alpha}) \\ \mathbf{0}^\top & 1 \end{pmatrix}, \quad (3.34)$$

such that  $\mathbf{T}_k(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{R}^{4 \times 4}$  computes a rigid object motion that consists of a rotation  $\mathbf{R}_k(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{R}^{3 \times 3}$  and translation  $\mathbf{t}_k(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{R}^3$  for the  $k$ -th projection image, given the parameter vector  $\boldsymbol{\alpha}$ . Further details on the structure of  $\boldsymbol{\alpha}$  and its mapping to the rigid motion matrices may be obtained from appendix A.1. Similarly, 2-D rigid detector motion may be defined by

$$\mathbf{Q}_k(\boldsymbol{\alpha}) = \begin{pmatrix} \cos(\phi_k(\boldsymbol{\alpha})) & -\sin(\phi_k(\boldsymbol{\alpha})) & s_k(\boldsymbol{\alpha}) \\ \sin(\phi_k(\boldsymbol{\alpha})) & \cos(\phi_k(\boldsymbol{\alpha})) & t_k(\boldsymbol{\alpha}) \\ 0 & 0 & 1 \end{pmatrix}, \quad (3.35)$$

where  $\phi_k(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{R}$  is the 2-D rotation angle for the  $k$ -th projection image. Further,  $s_k(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{R}$  and  $t_k(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{R}$  are the corresponding detector translations in detector  $u$  and  $v$  directions, respectively, given in pixel coordinates. The motion may be incorporated into the projection matrices by

$$\mathbf{P}_k(\boldsymbol{\alpha}) = \mathbf{Q}_k(\boldsymbol{\alpha}) \cdot \mathbf{P}_k \cdot \mathbf{T}_k(\boldsymbol{\alpha}), \quad (3.36)$$

such that  $\mathbf{P}_k(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{R}^{3 \times 4}$  are the updated projection matrices based on motion parameters  $\boldsymbol{\alpha}$ .

It is worth noting, that this type of linear model is not only applicable for rigid, but to a much large class of motions, defined in 2-D or 3-D. An extension is straightforward by incorporating more parameters, e. g., for scaling or shear, into matrices  $\mathbf{Q}_k(\boldsymbol{\alpha})$  or  $\mathbf{T}_k(\boldsymbol{\alpha})$ .

### Nonrigid Motion

The scope of this thesis is restricted to parametric models. For more information on nonparametric models it is referred to [Mode03, Clar06, Daum11]. The most common parametric nonrigid deformation models use B-splines [Ruec99] or TPS [Davi97, Mull13]. Both consist of a global rigid transformation and an additional local transformation yielding a nonrigid motion field  $\mathbf{d}_k(\mathbf{x}, \boldsymbol{\alpha}) : \mathbb{R}^{N+L} \rightarrow \mathbb{R}^N$ . B-splines and TPS use control points  $\mathbf{v}_{ak} \in \mathbb{R}^N$  with  $a \in [1, \dots, A]$  and  $A$  being the number of control points used. According to Davis et al. [Davi97] the TPS deformation function for dimension  $N = 3$  is given by

$$\mathbf{d}_k(\mathbf{x}, \boldsymbol{\alpha}) = \sum_{a=1}^A \mathbf{G}(\mathbf{x} - \mathbf{v}_{ak}) \mathbf{b}_{ak}(\boldsymbol{\alpha}) + h(\mathbf{T}_k(\boldsymbol{\alpha}) \mathbf{x}), \quad (3.37)$$

where  $\mathbf{T}_k(\boldsymbol{\alpha})$  contains the rigid motion parameters and  $\mathbf{G}(\mathbf{x} - \mathbf{v}_{ak}) : \mathbb{R}^N \rightarrow \mathbb{R}^{N \times N}$  is the kernel function that provides the location specific weights for the spline coefficients  $\mathbf{b}_{ak}(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{R}^N$ . It is given by

$$\mathbf{G}(\mathbf{x} - \mathbf{v}_{ak}) = \|\mathbf{x} - \mathbf{v}_{ak}\|_2 \mathbf{I}_3 , \quad (3.38)$$

where  $\|\cdot\|_2$  is the L2 norm and  $\mathbf{I}_N \in \mathbb{R}^{N \times N}$  is an  $N$ -dimensional identity matrix. Optimization of this type of motion field can be done in two ways. First, the unknowns in  $\mathbf{T}_k(\boldsymbol{\alpha})$  and  $\mathbf{b}_{ak}(\boldsymbol{\alpha})$  are contained in  $\boldsymbol{\alpha}$  and optimized directly, thus, the total number of parameters is  $(NA + 6)K$ . Another method for motion estimation is to estimate the translations for each control point  $\mathbf{v}_{ak}$ . In this case, inserting the estimated deformations into Eq. (3.37) allows for a closed form solution of  $\mathbf{T}_k(\boldsymbol{\alpha})$  and  $\mathbf{b}_{ak}(\boldsymbol{\alpha})$ , thus, the number of parameters is  $NAK$ . The second variant may also be used for extrapolation of sparse motion fields [Mull 13, Maie 14, Berg 16a]. In either case we define that  $\boldsymbol{\alpha}$  contains the necessary information to compute translations at arbitrary locations  $\mathbf{x}$ . A similar approach has been shown for B-spline motion fields [Ruec 99].

### 3.4.2 Based on Surrogate Signals (I)

This section provides a literature review of motion estimation methods using surrogate signals or dedicated alterations to the image content. We start with a detailed description of the prior work on motion estimation of knees under weight-bearing condition, which also built the baseline at the beginning of this thesis. This is followed by a more general overview of motion estimation methods used in other fields of CBCT imaging.

#### CBCT Weight-Bearing Imaging of Knees Using Fiducial Markers

At an initial phase of the weight-bearing imaging project associated to this thesis a fiducial marker-based solution was used to estimate for knee motion during acquisition [Choi 12, Choi 13, Choi 14c]. Tantalum beads with a 1 mm diameter were attached to the knee and absorbed most X-ray photons due to their high attenuation coefficient. This enables identification of markers in the individual projection images. Choi et al. investigated performance of three different motion models, including translations of projection images, nonrigid deformation of projection images using TPS and 2-D/3-D registration of average 3-D marker locations to the individual projection images. Results showed a superior image quality for the latter approach when evaluated with a numerical XCAT phantom [Sega 10] but also on real acquisitions [Choi 13, Choi 14c]. In the following we introduce the motion estimation algorithm described by [Choi 13] in more detail. Let the number of parameters be  $L = 6K$  such that  $\boldsymbol{\alpha}$  contains three rotations and translations for each of  $K$  projection images. Then the cost function is given by

$$\underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \sum_{k=1}^K \sum_{m=1}^M \|\mathbf{m}_{km} - \mathbf{u}_{km}\|_2^2 \quad (3.39)$$

$$\mathbf{m}_{km} = h(\mathbf{P}_k \cdot \mathbf{T}_k(\boldsymbol{\alpha}) \cdot \mathbf{x}_m) , \quad (3.40)$$

where  $\mathbf{x}_m \in \mathbb{R}^3$  is the average 3-D position of the  $m$ -th marker and  $\mathbf{u}_{km} \in \mathbb{R}^2$  are the detected bead locations of the  $m$ -th marker on the  $k$ -th projection. Further,  $\mathbf{m}_{km} \in \mathbb{R}^2$  is the projection of  $\mathbf{x}_m$  to the  $k$ -th projection image. Note, that  $\mathbf{m}_{km}$  and  $\mathbf{u}_{km}$  are defined in pixel coordinates.

To evaluate Eq. (3.39), one has to identify potential candidates for  $\mathbf{u}_{km}$ . Choi et al. proposed to first enhance marker locations by building the gradient of the projection images. After thresholding the gradient images with a heuristic lower bound, the Hough transform was used to identify 2-D bead locations  $\mathbf{u}_{km}$  [Choi 14c].

Moreover, the average 3-D positions  $\mathbf{x}_m$  of the markers need to be identified and also their correspondence to the detected 2-D candidate points. This is done in a semi-manual approach. First, the position of identical markers  $\tilde{\mathbf{u}}_{km}$  is annotated manually in a small set ( $\sim 4$ ) of projection images, distributed over the full-scan range. The average 3-D position of the annotations  $\tilde{\mathbf{x}}_m$  can be computed in a closed form solution using singular value decomposition (SVD) as shown in [Choi 14c]. This is equal to finding the point, that is closest to all lines backprojected from the annotated points on the detector to the X-ray source w.r.t. the L2 norm. The initially computed  $\tilde{\mathbf{x}}_m$  are then forward projected to all projection images using Eq. (3.40), yielding first estimates for  $\mathbf{m}_{km}$ . The correspondence problem is now solved by assigning detections  $\mathbf{u}_{km}$  to those  $\mathbf{m}_{km}$ , that are closest w.r.t. the Euclidean distance. The estimation of the final  $\mathbf{x}_m$  may be computed using all assigned points  $\mathbf{u}_{km}$ . Repeating this process multiple times leads to an iterative refinement of  $\mathbf{x}_m$  and their assignments to the detected 2-D points.

Drawbacks of the approach presented by Choi et al. include the time consuming manual annotations  $\tilde{\mathbf{u}}_{km}$ . Assuming that  $M \geq 6$  which is the lower bound to avoid an underdetermined optimization problem, the number of annotations would be  $\geq 6 \times 4 = 24$ . Also, the robustness of the Hough transform for small objects like beads may be reduced in noisy environments, either limiting the number of detected markers (high threshold) or increasing the number of wrong detections (lower threshold). As markers manipulate the image data, they also cause high intensity streaking artifacts in the reconstruction domain which reduces clinical usability [Berg 14a].

## General Overview of Surrogate-based Motion Estimation

In addition to motion estimation of knees, fiducial markers have been used regularly in the field of radiotherapy but also for estimating skeletal motion. During radiotherapy it is essential to track tumor motion precisely to avoid harmful irradiation of healthy tissue. Fiducial markers have been implanted close to the tumor allowing real-time tracking of tumor motion during treatment, based on simultaneously acquired projection images [Schw 00, Shir 03, Russ 05]. The estimated motion can then be used to readjust or trigger the high energy radiotherapy beam.

Especially breathing motion causes problems during radiotherapy as an external fixation is hardly possible. A multitude of motion estimation methods propose the usage of infrared markers, externally attached to the chest of the patient. Motion of the markers can be correlated to the internal motion based on a previously acquired 4-D CT. An overview of methods may be found in the review of Rit et al. [Rit 13]. Others proposed to use external 3-D range imaging sensors that are independent

of markers, to track breathing motion during treatment [Baue 12, Baue 13, Wasz 16, Geim 16].

A typical surrogate signal in the field of cardiac CBCT imaging is the ECG signal simultaneously acquired with the projection images. Motion-corrected 4-D imaging is possible for certain heart phases using ECG-gating. During gating the acquired projections are assigned to a set of heart phases identified from the ECG signal. Those projections can then be used for 3-D reconstruction of specified heart phases [Desj 04]. The position and size of gating-windows is known to be crucial, where a reduced weight to periphery of the gating window has been shown to be beneficial [Rohk 11]. An extensions to regular ECG-gating are registration-based approaches that aim to find the motion between individual heart phases [Schw 13, Mull 14b, Taub 16]. A different solution is to model the entire 4-D reconstruction as a single optimization problem, solved by iterative reconstruction methods, e.g., using a spatial and an ECG-driven, temporal total variation regularizer [Wu 12, Haas 16].

### 3.4.3 Based on Reference Image (II)

The majority of motion estimation methods are based on artifact-free reference images that may be defined in the projection, reconstruction or 4-D reconstruction domain. These reference images can be acquired before or after the motion-corrupted scan. They can also be derived from the motion-corrupted data directly, which usually requires prior knowledge on the imaged object or surrogate signals. If reference data is present, motion estimation is possible by (1) 3-D/3-D registration in the volume domain, (2) 2-D/2-D registration of projection images and (3) 2-D/3-D registration that estimates 3-D motion of a reference volume to 2-D projection images [Prum 09a, Mull 14a]. In this context, all registration methods refer to image-based measures, i. e., the cost functions used for registration are based on the image intensities of the projection images, reconstructed volumes, or both.

3-D/3-D registration-based motion estimation requires that static 3-D volumes can be reconstructed for individual time points with sufficient image quality. These volumes can be pairwise registered using 3-D/3-D registration to determine the time-dependent motion field between the individual motion states [Mull 14b, Taub 16]. A key problem is the generation of static 3-D reconstructions with sufficient accuracy. Projection gating can be used to determine a subset of projection images that belong to the same motion state. However, due to the limited and nonuniform angular sampling, gated reconstructions typically suffer from view aliasing artifacts which in turn reduces registration accuracy [Prum 09b, Mull 14b]. In the field of radiotherapy, gating information may be extracted directly from the projection images. For example the Amsterdam shroud method [Sonk 05] exploits the high contrast diaphragm which is tracked in superiorinferior direction. It is assumed that the diaphragm's motion has a strong correlation with the respiratory motion. However, this method is not applicable to knee imaging as it is restricted to motion mostly parallel to the rotation axis of the C-arm. In general, knee motion in straight standing position is not periodic and no intuitive surrogate signal is given.

Treatment planning in radiotherapy often involves a prior four-dimensional (4-D) CT acquisition which is first partitioned into a small number of breathing phases, each

corresponding to a static 3-D volume [Li 10, Li 11, Li 06, Rit 09a, Rit 09b, Wang 14, Schw 05]. Time dependent deformations are obtained by nonrigid 3-D/3-D registration to a reference breathing phase. A typical assumption is that the prior motion is similar to the respiratory motion during treatment. The acquired projections can be assigned to one of the 3-D volumes, based on a similarity measure computed between digitally reconstructed radiographs (DRRs) and the projection images. Thus, the current motion estimate can simply be used from the precomputed deformations. Li et al. extended this approach by training a statistical motion model on the prior 4-D CT data, which is then optimized by 2-D/3-D registration [Li 10, Li 11] during treatment. If the prior 4-D CT is not directly available it can be replaced by a set of gated 3-D reconstructions [Li 07] or generated numerically by interpolation between two 3-D reconstructions of maximum inhale and exhale [Schw 05].

Other work has investigated 2-D/2-D registration to estimate motion. Typically, the motion-corrupted projection images are registered to DRRs. A prior gated reconstruction was used to compute the DRRs in coronary imaging [Schw 13, Hans 08]. Otherwise, a prior acquisition may be used to render DRRs. Often segmentations of high contrast objects are used to enhance anatomies of interest in the projection domain prior to registration. For example, centerline extraction of vessels [Hans 08], MIPs [Hans 08, Unbe 15] or top-hat filtering [Schw 13].

2-D/3-D registration algorithms estimate the 3-D pose of a prior volume such that it fits best to the projection image. Most algorithms are based on the “projection strategy” [Mark 12]. First, the estimated motion is incorporated into the DRR process which yields an updated, synthetic projection image. This projection image is then compared to the acquired image by a similarity measure. The similarity measure is optimized by an iterative refinement of the estimated motion. 2-D/3-D registration has been widely used in the field of respiratory motion estimation for radiotherapy. We refer to Rit et al. for a comprehensive review of the state-of-the-art respiratory motion estimation approaches [Rit 13]. Additionally, methods have been developed for vascular [Blon 06, Groh 08, Mitr 13], cardiac [Prum 09a] or skeletal motion [Russ 05, Bifu 09, Otak 13, Otak 15, Tsai 10, Zhu 12a, Zhu 12b, Lin 13]. Zeng et al. have used nonrigid 2-D/3-D registration of a prior CT scan of the thorax to the current on-board fluoroscopy image for respiratory motion estimation during radiotherapy [Zeng 05, Zeng 07]. A parametric 3-D B-spline deformation was estimated by optimization an sum-of-squared-differences (SSD) and logarithmic cross correlation similarity measure. Gendrin et al. focused on a run-time optimized rigid 2-D/3-D registration where a single registration of 5 rigid deformation parameters (excluded depth information) was done in less than 0.5 seconds using graphics processing unit (GPU) programming and a normalized mutual information (MI) similarity measure [Gend 12]. Fu et al. separate the rigid registration problem into in-plane and out-of-plane rotations which are estimated using a combined pattern intensity pattern intensity (PI) and SSD similarity measure. For a detailed reviews of similarity measures we refer to [Penn 98] and [Gend 11].

In the field of skeletal 2-D/3-D registration, Russakoff et al. used an MI similarity measure to register the rigid deformation of a prior 3-D spine CT to 2-D radiographs, where a single fiducial marker is tracked and used as a regularizer to improve registration robustness [Russ 05]. Normalized cross correlation was used to register a

vertebra volume to fluoroscopic images [Bifu02, Bifu09]. Otake et al. claim that **normalized gradient information (NGI)** is a particularly robust similarity measure for overlapping structures, when registering vertebrae to projection images [Otak13]. NGI evaluates the similarity of gradient directions and weights them with the lower gradient magnitude. More recently, however, they decided to use the **gradient correlation (GC)** similarity measure [Otak15], which was originally described by Penney et al. [Penn98]. Whereas, NGI weighs with the minimum gradient magnitude to provide a more robust registration result in the presence of overlapping edges, GC is known to be robust to linear transformations of the intensities between DRR and projection image.

2-D/3-D registration for the knee anatomy includes work by Tsai et al., who introduce the **weighted edge matching score (WEMS)** similarity measure that compares edges extracted by a Canny edge detector, where longer edges are assigned a higher weight [Tsai10]. They reported a superior estimation of out-of-plane translations compared to PI and **gradient differences (GD)**. Lin et al. have also used WEMS to register a motion-free knee MRI volume to an MRI slice acquired when using a weight-bearing apparatus [Lin13]. Similar methods for registration of bones in knee and shoulder joints are presented by Zhu et al. [Zhu12a, Zhu12b]. First, the bone meshes are extracted in the prior volume, then the bone outline on the projection images is computed based on a direct projection of the bone meshes. Registration is performed between the projected outline and the outline of the acquired bone, obtained by segmentation in the projection domain.

A vast majority of related work uses a single or biplane imaging system, thus, the number of unknowns for rigid 2-D/3-D registration is limited to six or 12, respectively. Very little work was done when registering 3-D volumes to a full stack of projections. Recently, Ouadah et al. proposed a geometric calibration for CBCT systems based on the NGI similarity measure and a statistical optimizer [Ouah16]. They registered a prior CT scan to each acquired 2-D projection image by optimizing up to nine geometrical parameters per projection for a maximum of 496 projection images, yielding a total of  $9 \times 496 = 4464$  parameters.

### 3.4.4 Without Reference and Surrogate Data (III)

Motion estimation becomes more challenging in case no prior imagery is available and also no suitable surrogates can be acquired. Thus, less information is available for the optimization of a motion model. Approaches, that are independent of prior images and surrogates exist and can be assigned two one of two groups. First, assumptions or prior knowledge about the shape, intensity distribution and other **motion artifact metrics (MAMs)** can be optimized in the reconstruction domain. Second, mathematical CC that describe redundancies in the projection domain may be exploited for motion estimation.

#### MAMs for Motion Estimation

The term MAM was defined by Rohkohl et al. [Rohk13] and refers to image features defined in the reconstruction domain that are ideally sensitive to different types of motion artifacts. Optimization of such features could then be used to estimate a

motion model. Image features used in literature are widespread. Wicklein et al. compared a multitude of image features w.r.t. their sensitivity to motion-induced artifacts in the field of neuroradiology [Wick12]. They found that the image entropy based on the histogram of the reconstruction provided best sensitivity over a set of different motion types. The goal is to enhance sharpness of the image by minimization of the gray level entropy, usually measured in a certain region of interest (ROI).

Entropy minimization has also been used for motion estimation by Kyriakou et al. [Kyri08], who optimized motion parameters for circular CBCT on a thorax phantom and in a rat study. However, the complexity was limited to a small set of geometric parameters. In contrast, Kingston et al. use the sum of gradient magnitudes, also known as total variation, to measure image sharpness. The sharpness is maximized on a subset of reconstructed slices in volume space, using a multiresolution approach. Only a global alignment was optimized estimating a total of four parameters.

In the field of cardiac CT imaging, Rohkohl et al. optimize entropy but also positivity to enhance image quality in an ROI extracted from a segmentation of coronary arteries [Rohk13]. Positivity aims to keep gray values in the ROI above a user defined threshold. Based on a nonrigid spline-based motion model with control points defined in spatial and temporal domain, they report promising results for reconstructing coronary arteries.

Recently, Sisniega et al. proposed an MAM-based algorithm to correct for motion of a knee acquired under weight-bearing conditions [Sisn16]. They estimated six rigid parameters per projection image, using total variation, entropy and a regularizer that penalizes high-frequency changes of motion parameters. Large improvements in image quality were reported when evaluating the method on a cadaveric hand. Image quality also increased when applied to an in-vivo weight-bearing scan of a knee joint, however, streaking artifacts were still visible after correction.

### What are Consistency Conditions?

Another strategy for motion estimation is based on CC, which describe redundancies in the projection or sinogram domain. These redundancies originate from repeated scanning of the same physical object from different view points.

HLCC were derived by Helgason [Helg80] and Ludwig [Ludw66] and are best known and most often used in literature. The HLCC build a set of relations between the Radon transform (i. e., a parallel-beam sinogram) and the imaged object in the world coordinate system. They are defined by

$$0 = \int_0^{2\pi} \underbrace{\left( \int_{-\infty}^{\infty} (u_p)^\nu q(u_p, \theta) du_p \right)}_{\text{Moments along } u_p} e^{-in\theta} d\theta \quad (3.41)$$

$$\forall n, \nu \in \mathbb{N} \quad \bigwedge \quad n > \nu \geq 0 ,$$

where  $q(u_p, \theta) : \mathbb{R}^2 \rightarrow \mathbb{R}$  is the parallel-beam sinogram with detector position  $u_p$  and rotation angle  $\theta$ . Also,  $n$  is the order of the Fourier series expansion along the rotation angle and  $\nu$  is the order of the image moment. More specifically, HLCC state that image moments of order  $\nu$  cannot cause contributions to Fourier series

coefficients of order  $n > \nu$ . The theorem holds for ideal monochromatic line integrals and the continuous Radon transform. In case of physics-, geometry- or sampling-based deviations the HLCC are no longer fulfilled which allows their usage as cost function to correct for these effects. It is referred to Welch et al. [Welc98] for a helpful and intuitive interpretation of image moments and their relations to the world coordinate system.

Multiple extensions to fan-beam CT have been made over the last decade [Chen05, Yu06, Leng07, Clac15]. Recently, Clackdoyle et al. proposed an extension to CBCT geometries [Clac13]. However, their theorem requires that all X-ray sources are positioned on a plane which must not intersect with the object, thus, excluding circular cone-beam trajectories.

Applications of the HLCC and their fan-beam extensions are widespread and include attenuation correction of positron emission tomography (PET) and single photon emission computed tomography (SPECT) acquisitions [Hawk88, Natt93, Welc98, Welc03], extrapolation of missing or truncated projection data [Erla00, Chen05, Star05, Van06, Xu10] and correction of beam-hardening artifacts [Mou06, Tang11]. In work associated to this thesis, we recently proposed a fan-beam extension of HLCC used to estimate location, density and shape parameters of simple geometric objects with the application to overexposure correction [Preu15].

Kudo et al. presented a different form of the HLCC where the moments are converted to Chebyshev polynomials [Kudo91]. They also propose an efficient evaluation of the properties using the sine transform along the detector and the Fourier transform over the angular direction. This is already very similar to another type of CC which we call FCC. Edholm et al. [Edho86] and Natterer [Natt86] showed that the 2-D Fourier transform of a parallel-beam sinogram must have triangular regions with absolute values close to zero. An extension to fan-beam geometries has been made by Hawkins et al. [Hawk88] and Mazin et al. [Mazi10] who show that

$$\left| \int \int g(\lambda, u) e^{-i2\pi(\lambda\omega + u\xi)} d\lambda du \right| \approx 0, \quad (3.42)$$

where the regions can be defined by the following properties of  $\xi$  and  $\omega$

$$\left| \frac{\omega}{\omega - \xi(S + D)} \right| > \frac{r_p}{S}. \quad (3.43)$$

$\xi$  and  $\omega$  are axes in frequency domain that correspond to the detector direction  $u$  and the rotation angle  $\lambda$ , respectively. The size of the regions depends on the maximum extent of the object  $r_p$  w.r.t. the rotation center but also on geometric parameters of the scanner geometry as introduced Sec. 3.2.2. FCC have been applied for data extrapolation [Karp88, Pohl14] and attenuation correction in PET and SPECT [Xia95, Abel09]. Extensions of the FCC have been made by Desbat et al. [Desb04] and Brokish et al. [Brok06] who derived sampling properties for cone-beam geometries using the 3-D Fourier transform of cone-beam projection data.

Both HLCC and FCC were defined for parallel-beam sinograms and subsequently extended to fan- and cone-beam data. In contrast, John [John38] presented CC that define a system of ultrahyperbolic partial differential equations directly for cone-beam projection images. Practical extensions of John's theoretic concept have been

made by Patch et al. with the application of estimating missing projection data [Pate02b, Pate02a].

### Consistency Conditions for Motion Estimation

All of the previously introduced CC are derived on ideal assumption. The projection images are assumed to have an infinite extent such that no truncation is possible and the physical model is restricted to monochromatic absorption as described in Sec. 3.2. Furthermore, the CC do often not incorporate sampling information and are presented for continuous signals. Also, no irregular patient or scanner motion is allowed. The presented CC can be violated if any of these assumptions are not fulfilled. This effect can be exploited in a practical scenario, e. g., if patient motion violates a CC we can try to estimate the motion such that the consistency is best restored.

Motion estimation based on HLCC and its extensions has been presented for fan-beam CT by Yu et al. [Yu06], where they derive a closed form solution to estimate 2-D object translations directly from the sinogram, using polynomials as motion model. In a later study they extended their motion class to 2-D translations and rotations [Yu07]. In general a closed form solution to motion estimation is of great interest, however, their methods are currently limited to fan-beam geometries and have not been applied to any real data which reduces their applicability to real CBCT acquisitions.

Only a small number of algorithms exist that use CC for motion estimation in 3-D cone-beam geometries. Leng et al. shows a heuristic extension of an existing fan-beam CC [Chen05] with application to motion estimation, yet, the extension is limited to only small cone-angles. Aichert et al. proposed epipolar consistency conditions (ECC) based on the equivalence of line integrals over epipolar line pairs [Aich14, Aich15]. ECC showed promising results for estimating rigid motion using six X-ray source positions distributed in 3-D space around the object. ECC have also been applied for motion estimation on real short-scan CBCT acquisitions of a head [Frys15] and for correction of improper geometric calibration of CBCT systems [Debb13, Maas14]. However, a circular CBCT geometry does not impose ideal conditions for their algorithm due to unfortunate orientations of epipolar lines in case that all X-ray sources are located in the same plane [Aich15].

## 3.5 Motion Compensation Methods

The final step of motion correction is the reconstruction given the estimated motion models. Even though motion estimation is often more challenging and time consuming, it is not necessarily given that reconstruction algorithms exist that are able to provide exact solutions. In fact, it has been shown that algebraically exact reconstruction is only possible for a certain class of motion models. Roux et al. derived an exact algorithm for reconstruction of parallel- and fan-beam geometries with 2-D affine motion of the object [Roux04]. They also show admissibility conditions that need to be fulfilled during data acquisition given the affine motion. Desbat et al. showed that an even larger class than affine transformations is possible, with the only

limitations that rays in parallel-beam geometry remain parallel after the mapping and rays in a fan-beam geometry also remain a fan-shaped [Desb07].

More flexibility w. r. t. the motion model may be provided by approximate algorithms. For many applications the limited accuracy of motion estimation or insufficient motion models build the bottleneck of the motion correction's performance, with only little improvement when using exact compensation methods. The scope of this thesis is restricted to approximate motion compensation methods, that are able to incorporate arbitrary rigid and nonrigid deformations in projection- or reconstruction domain.

### 3.5.1 Updated Backprojection for Affine Motion

A well known compensation method has been introduced by Schäfer et al. [Scha06] and Prümmer et al. [Prum06] for the field of cardiac imaging. The idea is easily integrated into the standard FDK algorithm shown in Sec. 3.2.3. Let us first present its application for rigid deformations in projection, as well as reconstruction domain. The results of motion estimation is a parameter vector  $\boldsymbol{\alpha}$ , that contains the coefficients, describing the affine parameters in 2-D, 3-D or both. Looking at the affine mapping for a rigid motion model in Eq. (3.36), we can obtain that all kinds of affine motions can be incorporated into projection matrices. The idea presented in [Scha06, Prum06] is to only manipulate the backprojection step from the conventional FDK (cf. Eq. (3.33)). Its mapping from 3-D to 2-D coordinates is adjusted by

$$\underline{\mathbf{u}} = \mathbf{P}_k(\boldsymbol{\alpha}) \underline{\mathbf{x}} . \quad (3.44)$$

Note, that the subsequent backprojection given by Eq. (3.33) remains equal. Only, the static mapping  $\mathbf{P}_k$ , that maps  $\underline{\mathbf{x}}$  to  $\underline{\mathbf{u}}$  has been replaced by the updated projection matrices  $\mathbf{P}_k(\boldsymbol{\alpha})$  according to Eq. (3.36). The algorithm is approximate as no update of ramp filtering or cosine weighting is performed [Scha06]. However, this also ensures that the algorithm has no additional complexity compared to a static FDK reconstruction, except the computationally cheap matrix multiplications of Eq. (3.36).

### 3.5.2 Updated Backprojection for Nonrigid Motion

As for rigid motion only the mapping function needs to be replaced when using a nonrigid deformation field  $\mathbf{d}_k(\mathbf{x}, \boldsymbol{\alpha})$  in reconstruction domain, i. e., for dimensionality  $N = 3$ . The mapping of 3-D to 2-D coordinates now evaluates to

$$\underline{\mathbf{u}} = \mathbf{P}_k \begin{pmatrix} \mathbf{x} + \mathbf{d}_k(\mathbf{x}, \boldsymbol{\alpha}) \\ 1 \end{pmatrix} . \quad (3.45)$$

Similarly for 2-D deformations, i. e.,  $N = 2$ , we may rewrite the mapping function to

$$\begin{aligned} \hat{\underline{\mathbf{u}}} &= \mathbf{P}_k \mathbf{x} \\ \underline{\mathbf{u}} &= \hat{\underline{\mathbf{u}}} + \hat{\underline{u}}_3 \begin{pmatrix} \mathbf{d}_k(h(\hat{\underline{\mathbf{u}}}), \boldsymbol{\alpha}) \\ 0 \end{pmatrix} . \end{aligned} \quad (3.46)$$

Note, that the remainder of the backprojection, including distance weighting and summation, is identical for all motion compensated reconstructions and also for the static FDK reconstruction. The main benefit of this type of motion compensation is its simplicity but also its flexibility. The theoretic framework allows all possible motion models and combinations thereof. However, the approximate nature of the algorithm requires the deformations to be small to ensure sufficient image quality [Tagu08].

## 3.6 Conclusion

This chapter introduces image reconstruction algorithms and corresponding acquisition geometries. In addition, the mathematical foundation of the thesis is given by introducing important concepts such as projection matrices, rigid and nonrigid motion models and motion compensated reconstruction. In an extensive literature review, a variety of different motion estimation methods is presented including a detailed introduction of the state-of-the-art motion correction method for weight-bearing imaging of the knee, based on fiducial markers.

Unfortunately, the majority of the presented approaches for motion estimation are not directly applicable to weight-bearing imaging of the knee joint. For example, an acquisition of a prior 4-D CT data, as often done in radiotherapy, is currently not possible under weight-bearing conditions. Moreover, even if prior 4-D CT data is available, the motion is not periodic and therefore likely not similar to the motion from the prior scan. Many approaches for motion estimation in cardiac imaging rely on an ECG signal. Such surrogate signals do also exploit periodicity which is rarely given for knee motion. The focus for weight-bearing knee joint imaging lies on an imaging protocol that adds little to none complexity to the acquisitions, compared to static scans in supine position. Hence, surrogate signals like external tracking devices or range imaging cameras are not a desired option.

However, easily accessible prior information exists also for weight-bearing imaging. For example, all our clinical applications of weight-bearing CBCT of the knee include a prior, motion-free acquisition in supine position (cf. Sec. 2.3). This data may be used for motion estimation based on 2-D/3-D registration. Multiple work was presented for motion estimation of the knee, yet, almost all of this work uses a single or biplane imaging system. In this case the projection images can usually be positioned such that no overlap with high-contrast background structures occurs. In weight-bearing CBCT the knees are imaged using a horizontal trajectory and a regular standing position, i.e., overlap of Tibia, Femur, Fibula or Patella from left and right leg but also within each leg is inevitable. This occlusion makes registration methods that require a segmentation in the projection image domain such as WEMS [Tsai 10, Lin 13, Zhu 12a, Zhu 12b] hardly applicable and also increase the complexity for regular, intensity-based algorithms.

Finally, the chapter discusses motion estimation approaches that work without surrogate signals and prior images. Entropy or total variation have been shown to be robust, image-based features used for iterative motion estimation. However, the corresponding measures aim to increase image sharpness and are therefore more sensitive to geometric blurring or smearing of structures, which is usually related to

low-frequency motion. High-frequency motion causes streaking artifacts which do not drastically influence entropy and total variation constraints as they may cause sharp edges (cf. Fig. 2.6(c)). Sisniega et al. recently proposed an image-based method for weight-bearing imaging of the knee, based on Entropy and total variation [Sisn 16]. Their results show clear reduction of motion blur, yet, sharp streaking artifacts due to high-frequency motion could not be removed entirely.

A detailed literature review is provided for **CC** which are defined in the projection domain. **CC** have also been used for motion estimation, where previous work showed promising results for using **ECC** or extensions of **HLCC** for fan- and cone-beam geometries. However, the majority of **CC** is still limited to 2-D parallel- or fan-beam imaging.



PART II

# Theory



# Marker-Based Motion Correction

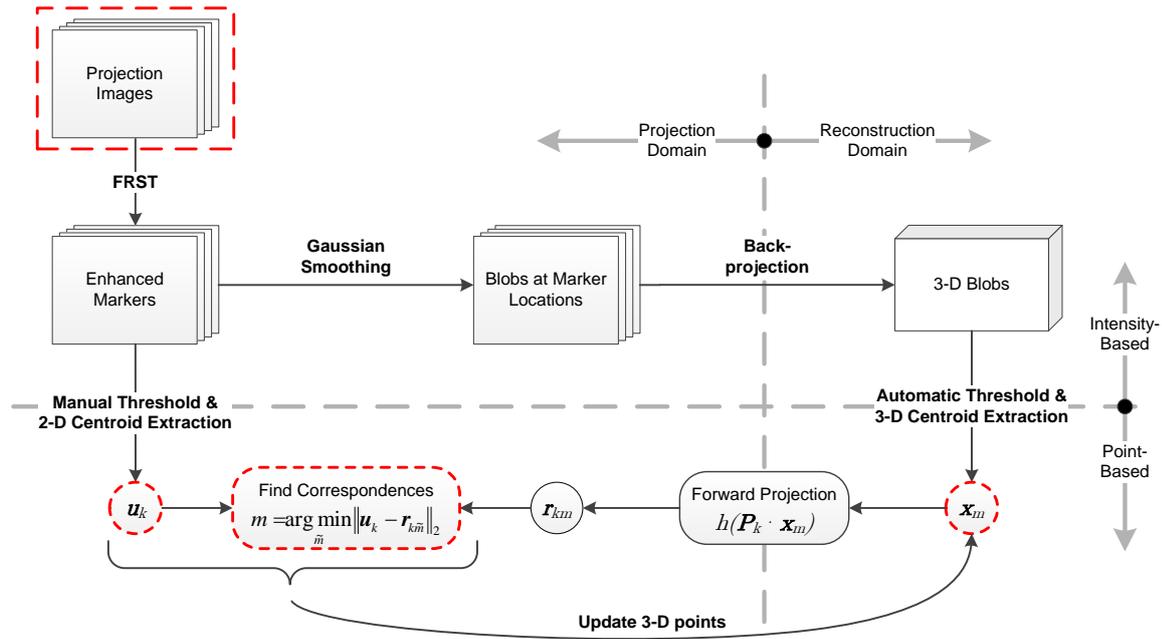
4.1 Robust, Fully Automatic Marker Detection.....	49
4.2 Optimization and Outlier Detection.....	54
4.3 Removal of Markers.....	56

The baseline of this thesis was prior work from Choi et al. who proposed motion-corrected reconstruction using externally attached fiducial markers [Choi 14a]. Before acquisition in weight-bearing position, metallic markers are attached to the skin. They can be tracked in the projection images allowing for motion estimation using a point-based 2-D/3-D registration process. The method is introduced in more detail in the literature review of surrogate-based motion estimation methods in Sec. 3.4.2. Within the scope of this thesis, the proposed method has been steadily enhanced. One goal was to develop a more robust but also fully automatic detection of markers, which is shown in Sec. 4.1. The focus in Sec. 4.2 is on improved accuracy but also robustness of the method. Therefore, an outlier detection method and an advanced gradient-based optimization scheme is presented. Finally, in Sec. 4.3 we present multiple methods that allow removal of markers from the reconstruction images, to avoid high intensity streaking artifacts. The majority of presented approaches are based on two conference [Berg 14a, Mull 15b] and a journal publication [Berg 16a].

## 4.1 Robust, Fully Automatic Marker Detection

One drawback of the previously introduced motion correction method [Choi 14c] is, that it required an initial set of manual annotations for each marker. In a typical weight-bearing imaging protocol up to 16 markers are distributed around both legs. Assuming an average of four annotations per marker as reported in [Choi 14c] and up to six weight-bearing scans, e. g., for the cartilage study (cf. Sec. 2.3.2), a total of  $6 \times 4 \times 16 = 384$  manual annotations may be required for evaluation of a new subject. Therefore, a fully automatic approach was developed with the goal to reduce the manual interaction to a minimum, thus, minimizing the overall time needed to perform motion-corrected reconstructions.

The complete pipeline is shown in Fig. 4.1. The acquired projection images are given as input indicated by the red, dashed frame. First, the FRST is applied to enhance the markers and subtract the background information, followed by a Gaussian smoothing and a 3-D backprojection. The FRST filtered projections, but also



**Figure 4.1:** Schematic representation of the marker detection and the automatic detection of their average position in 3-D. As input we have the acquired projection images denoted by the red, dashed frame. The outputs are 2-D locations  $\mathbf{u}_k$ , mean marker positions  $\mathbf{x}_m$  in 3-D, and a list of correspondences between  $\mathbf{x}_m$  and  $\mathbf{u}_k$ .

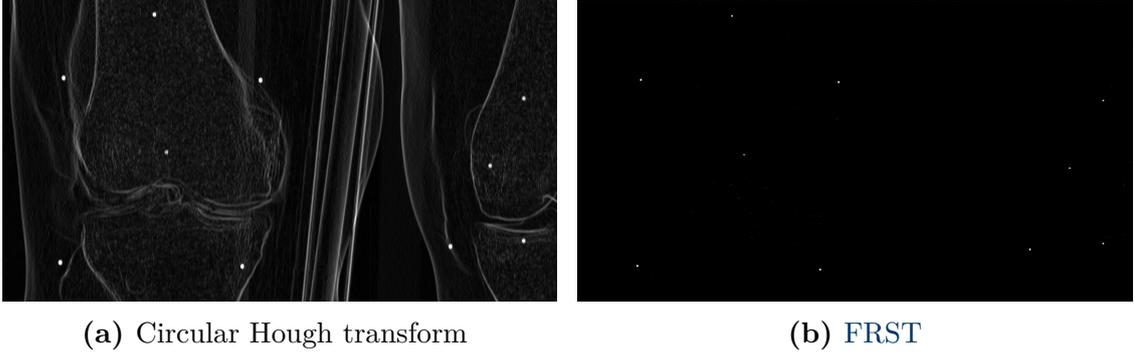
the backprojection result, are then discretized followed by a center point extraction in 2-D and 3-D, respectively. This yields detected positions  $\mathbf{u}_k$  in projection domain and initial estimates for the mean marker positions in reconstruction domain  $\mathbf{x}_m$ . An additional result are the correspondences between a 3-D position and its detections in projection domain.

In the following sections the individual contributions for a robust, fully automatic marker-based motion correction are shown. A more robust detection in projection domain is presented in Sec. 4.1.1. The extraction process of mean marker positions in reconstruction domain is outlined in Sec. 4.1.2, followed by an explanation of the method used to determine point correspondences in Sec. 4.1.3.

### 4.1.1 2-D Detection Using Fast Radial Symmetry Transform

The initial marker detection presented in [Choi 14c] applies the circular Hough transform [Gonz 08] to detect potential positions  $\mathbf{u}_k$ . The transform is based on an edge extraction step, using the gradient magnitude. Given a specified radius, all pixels belonging to an edge are assigned a confidence value, which indicates if the pixel belongs to a circle.

Part of this thesis was a replacement of the circular Hough transform with the FRST, which has been defined by Loy et al. [Loy 03] to detect radially symmetric points of interest in images. In contrast to the Hough transform, the FRST does also take the gradient directions into account.



**Figure 4.2:** Comparison of circular Hough transform and FRST w. r. t. their performance in separating markers from the background. The visualization windows were set from zero to the threshold where all beads are visible.

As a first step the gradient images are calculated. Let  $\nabla p_k(\mathbf{u}) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  be the gradient of the  $k$ -th projection image w. r. t. detector  $u_1$  and  $u_2$  direction, defined by

$$\nabla p_k(\mathbf{u}) = \begin{pmatrix} \frac{\partial p_k(\mathbf{u})}{\partial u_1} \\ \frac{\partial p_k(\mathbf{u})}{\partial u_2} \end{pmatrix}. \quad (4.1)$$

The result of the FRST is a pixel-wise weight for the presence of a symmetric feature with certain radii, where negative values correspond to dark features and positive values to bright features. In the following we introduce the main principle using only a single radius  $r_f$ , defined in pixel coordinates. First, two images  $\text{inc}_k(\mathbf{u}) : \mathbb{R}^2 \rightarrow \mathbb{R}$  and  $\text{mag}_k(\mathbf{u}) : \mathbb{R}^2 \rightarrow \mathbb{R}$ , of same size as the input image are initialized with zeros, where one holds the orientation information and the other one the gradient magnitude information.

Let us first define positive and negative locations  $\mathbf{u}_{pos} \in \mathbb{R}^2$  and  $\mathbf{u}_{neg} \in \mathbb{R}^2$  by

$$\mathbf{u}_{pos} = \mathbf{u} + \text{round} \left( \frac{\nabla p_k(\mathbf{u})}{\|\nabla p_k(\mathbf{u})\|_2} r_f \right) \quad (4.2)$$

$$\mathbf{u}_{neg} = \mathbf{u} - \text{round} \left( \frac{\nabla p_k(\mathbf{u})}{\|\nabla p_k(\mathbf{u})\|_2} r_f \right), \quad (4.3)$$

where  $\text{round}(\cdot)$  rounds to the nearest integer at each vector element and  $\frac{\nabla p_k(\mathbf{u})}{\|\nabla p_k(\mathbf{u})\|_2} r_f$  is the gradient direction scaled to length  $r_f$ . For each pixel location  $\mathbf{u}$ , orientation and magnitude images are updated by

$$\text{inc}_k(\mathbf{u}_{pos}) = \text{inc}_k(\mathbf{u}_{pos}) + 1 \quad (4.4)$$

$$\text{inc}_k(\mathbf{u}_{neg}) = \text{inc}_k(\mathbf{u}_{neg}) - 1 \quad (4.5)$$

$$\text{mag}_k(\mathbf{u}_{pos}) = \text{mag}_k(\mathbf{u}_{pos}) + \|\nabla p_k(\mathbf{u})\|_2 \quad (4.6)$$

$$\text{mag}_k(\mathbf{u}_{neg}) = \text{mag}_k(\mathbf{u}_{neg}) - \|\nabla p_k(\mathbf{u})\|_2 \quad (4.7)$$

After all pixel locations are processed the final result is built by

$$\widehat{\text{frst}}_k(\mathbf{u}) = \frac{\text{mag}_k(\mathbf{u})}{s_f} \left( \frac{|\widehat{\text{inc}}_k(\mathbf{u})|}{s_f} \right)^{c_{\text{rad}}}, \quad (4.8)$$

where

$$\widehat{\text{inc}}_k(\mathbf{u}) = \begin{cases} \text{inc}_k(\mathbf{u}) & \text{if } |\text{inc}_k(\mathbf{u})| < s_f \\ s_f & \text{otherwise} \end{cases} . \quad (4.9)$$

Note that  $c_{\text{rad}}$  is a parameter that enforces radial strictness and  $s_f$  is a scaling parameter.

Fig. 4.2 shows a comparison of the confidence value using the circular Hough transform (cf. Fig. 4.2(a)) and the FRST (cf. Fig. 4.2(b)). For both algorithms the visualization window was adjusted from zero to the minimum threshold such that all beads are detected. Both algorithms use the same radius  $r_f$  as input. It can be seen that the Hough transform is more sensitive to edges of noncircular structures, e. g., the bone outlines. In contrast, the FRST could accurately separate the markers from the background structures. Furthermore, the markers appear more localized in case of FRST.

After application of the FRST to each projection image, negative and small contributions are removed from  $\widehat{\text{frst}}_k(\mathbf{u})$  by

$$\text{frst}_k(\mathbf{u}) = \begin{cases} \widehat{\text{frst}}_k(\mathbf{u}) - t_{\text{proj}} & \text{if } \widehat{\text{frst}}_k(\mathbf{u}) > t_{\text{proj}} \\ 0 & \text{otherwise} \end{cases} , \quad (4.10)$$

where  $t_{\text{proj}} \in \mathbb{R}$  is a manually selected thresholding value that separates the marker contributions from residual background and noise. After thresholding the method employs a 2-D connected components analysis [Gonz 08] to identify pixel clusters. A set of potential marker locations  $\widehat{U}_k$  is extracted from each projection image by

$$\widehat{U}_k = \{\mathbf{u}_k \mid \mathbf{u}_k = \text{cluster centroid in } k\text{-th projection image}\} , \quad (4.11)$$

where  $\#\widehat{U}_k$  equals the number of identified clusters. Note that  $\#\widehat{U}_k$  may deviate from the actual number of markers, due to wrong detections or markers that are not in the FOV.

### 4.1.2 Automatic Extraction of 3-D Reference Positions

A necessary step for marker-based motion correction is to determine initial marker location  $\mathbf{x}_m$  in reconstruction domain. In [Berg 14a] we proposed an automatic extraction step of these locations without the need of manual annotations.

As a first step the result of the FRST algorithm is convolved by

$$\widetilde{\text{frst}}_k(\mathbf{u}) = \text{frst}_k(\mathbf{u}) * g_{\sigma_f}^2(\mathbf{u}) \quad (4.12)$$

where  $g_{\sigma_f}^2(\mathbf{u}) : \mathbb{R}^2 \rightarrow \mathbb{R}$  is an isotropic 2-D Gaussian low-pass filter,  $\sigma_f$  denotes its standard deviation, and  $[*]$  is the 2-D convolution operator. After low-pass filtering we can observe 2-D blob-like structures centered at marker locations.

Instead of performing an FDK reconstruction on the originally acquired projections,  $\widetilde{\text{frst}}_k(\mathbf{u})$  is directly backprojected to obtain a volume representation. Using Eq. (3.33) we can write the backprojection as

$$\begin{aligned} \underline{\mathbf{u}} &= \mathbf{P}_k \mathbf{x} \\ f_{\text{frst}}(\mathbf{x}) &= \frac{1}{2} \sum_{k=1}^K \frac{1}{(\underline{u}_3)^2} \widetilde{\text{frst}}_k(h(\underline{\mathbf{u}})) , \end{aligned} \quad (4.13)$$

where  $f_{\text{first}}(\mathbf{x}) : \mathbb{R}^3 \rightarrow \mathbb{R}$  is the backprojection result. High-intensity contributions, belonging to the same marker, should intersect in a single 3-D point during backprojection, assuming accurate separation of markers and no object motion. If we relax these constraints to moderate object motion and deviations in marker detection accuracy, an ideal intersection is no longer given. However, with the application of the Gaussian low-pass filter  $g_{\sigma_t}^2(\mathbf{u})$  we can artificially widen the marker detection to enforce an overlapping of intensities during backprojection. Thus,  $f_{\text{first}}(\mathbf{x})$  will contain 3-D blob structures. We assume that these blobs will cover the majority of the moving marker locations and that the centroid of the  $m$ -th blob represents the average marker location  $\mathbf{x}_m$ .

Calculation of connected components requires a binarized image. The threshold  $t_{\text{rec}}$  is determined automatically by using the maximum entropy method [Saho 88], which aims to maximize the sum of entropies, given by the gray level distributions of the segmented region and the residual region. The thresholding and binarization operation is given by

$$\widehat{f_{\text{first}}}(\mathbf{x}) = \begin{cases} 1 & \text{if } f_{\text{first}}(\mathbf{x}) > t_{\text{rec}} \\ 0 & \text{otherwise} \end{cases}, \quad (4.14)$$

where  $\widehat{f_{\text{first}}}(\mathbf{x}) : \mathbb{R}^3 \rightarrow \{0, 1\}$  is the binarized object function. A 3-D connected component analysis is performed on  $\widehat{f_{\text{first}}}(\mathbf{x})$ , yielding a centroid for each detected blob. As an output we obtain a set of average marker locations

$$X = \{\mathbf{x}_m \mid \mathbf{x}_m \in (\text{detected centroids of } M \text{ largest blobs})\}. \quad (4.15)$$

Because the total number  $M$  of markers is known, only the  $M$  largest blobs are assigned to marker locations.

### 4.1.3 Determining 2-D/3-D Point Correspondences

Note that the entries in  $X$  are already indexed with the marker number  $m$ , whereas the 2-D candidate points in  $\widehat{U}_k$  are not yet labeled. The correspondences between these sets are still to be defined. We adopted the method proposed in [Choi 14c] to extract correspondences, it is summarized by Algorithm 4.1.

Labeling can be executed separately for each projection image. In each image all 3-D marker positions  $\mathbf{x}_m$  are forward projected to the projection domain, followed by an exhaustive search for the closest Euclidean distance between projected and detected points. Let us also define a maximum valid distance  $d_u$  in projection domain. The label  $m$ , corresponding to the pair with the smallest distance, is assigned to the 2-D measurements. If the distance is already larger than  $d_u$ , the assignment is discarded. In any case, both values are removed from the search space before repeating the search for the next closest pair. Finally, we obtain sets  $U_k$  containing all assigned  $\mathbf{u}_{km}$  for a specific projection image  $k$ .

Note that the number of detections  $\#U_k$  is not necessarily equal to  $M$ , as some markers may not have been detected. It is recommended to adjust detection parameters such that  $6 \leq \#U_k \leq M$  for all projection images. Thus, it is ensured that the subsequent optimization is not ill-posed, using a rigid motion model with six parameters per projection image. Adjustments to increase  $\#U_k$  include an increase of the

---

**Algorithm 4.1:** Algorithm to extract point correspondences between 3-D marker locations and 2-D marker detections.

---

```

input :  $X, \hat{U}_k, \mathbf{P}_k$ 
output:  $U_k$ 

for  $k \leftarrow 1$  to  $K$  do
    // Initialize output and temporal set for 3-D locations
     $U_k \leftarrow \{\}$ ;
     $\hat{X} \leftarrow X$ ;
    while  $\#\hat{X} > 0 \wedge \#\hat{U}_k > 0$  do
        // Min. distance of projected points in  $\hat{X}$  to points in  $\hat{U}_k$ 
         $\tilde{\mathbf{u}}, \tilde{\mathbf{x}}_m \leftarrow \underset{\mathbf{u}, \mathbf{x}_m}{\operatorname{argmin}} \|\mathbf{u} - h(\mathbf{P}_k \cdot \mathbf{x}_m)\|_2 \quad \forall \mathbf{u} \in \hat{U}_k \wedge \mathbf{x}_m \in \hat{X}$ 
        if  $\|\tilde{\mathbf{u}} - h(\mathbf{P}_k \cdot \tilde{\mathbf{x}}_m)\|_2 < d_u$  then // Is distance small enough?
            // Assign marker and proj. index to closest 2-D point
             $\mathbf{u}_{km} \leftarrow \tilde{\mathbf{u}}$ ;
            // Add labeled point to output set
             $U_k \leftarrow U_k \cup \{\mathbf{u}_{km}\}$ ;
        end
        // Points are assigned  $\rightarrow$  Remove points from candidate sets
         $\hat{U}_k \leftarrow \hat{U}_k / \{\tilde{\mathbf{u}}\}$ ;
         $\hat{X} \leftarrow \hat{X} / \{\tilde{\mathbf{x}}_m\}$ ;
    end
end

```

---

maximum allowed distance  $d_u$ , but also a decrease of the binarization threshold  $t_{\text{proj}}$  to obtain additional candidate points. Note that both adjustments may also lead to an increase of wrong detections or false assignments of point correspondences.

## 4.2 Optimization and Outlier Detection

The rigid motion model, as well as most parts of the optimization are adopted from [Choi 14c] and have been introduced in Sec. 3.4.2. Thus, a parameter vector  $\boldsymbol{\alpha}$  of length  $L = 6K$  is optimized, given the 2-D points in sets  $U_k$  and the mean 3-D marker positions in  $X$ . The cost function given in Eq. (3.39) can be adjusted accordingly to

$$\underset{\boldsymbol{\alpha}}{\operatorname{argmin}} e(\boldsymbol{\alpha}) = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \sum_{k=1}^K \sum_{m=1}^{\#U_k} \left( \frac{1}{2K (\#U_k)} \|h(\underline{\mathbf{m}}_{km}) - \mathbf{u}_{km}\|_2^2 \right), \quad (4.16)$$

where  $\underline{\mathbf{m}}_{km}$  is the projection of the moved 3-D marker position, given by

$$\underline{\mathbf{m}}_{km} = \mathbf{P}_k \cdot \mathbf{T}_k(\boldsymbol{\alpha}) \cdot \underline{\mathbf{x}}_m \quad .$$

Note that the motion is encoded in the rigid motion matrices  $\mathbf{T}_k(\boldsymbol{\alpha})$ . In contrast to the previously presented cost function in Eq. (3.39), only the variable number of detections  $\#U_k$  and two normalization terms have been incorporated.

Motion compensated reconstruction after optimization is performed using the regular FDK algorithm described in Sec. 3.2.3 with an updated backprojection approach given by Equations (3.33) and (3.44) as described in Sec. 3.5.

### 4.2.1 Gradient of Cost Function

Choi et al. used a Nelder-Mead simplex algorithm to optimize the motion parameters separately for each projection image. It has been shown that if a gradient can be obtained from the cost function, it is beneficial to use a gradient-based optimizer, which is usually more accurate and converges faster [Noce 06]. As part of this thesis the optimization method was enhanced to a gradient-based method in [Mull 15b] where the gradient was numerically estimated by building forward-differences. This was further enhanced to an analytic gradient computation in [Berg 16a]. The gradient of Eq. (4.16) is given by  $\nabla e(\boldsymbol{\alpha}) = \left( \frac{\partial e(\boldsymbol{\alpha})}{\partial \alpha_1} \quad \dots \quad \frac{\partial e(\boldsymbol{\alpha})}{\partial \alpha_{6K}} \right)^\top$ . With help of the chain rule for multivariate functions we can compute the partial derivatives with respect to the individual elements of the parameter vector by

$$\frac{\partial e(\boldsymbol{\alpha})}{\partial \alpha_l} = \sum_{k=1}^K \sum_{m=1}^{\#U_k} \left( \frac{1}{K (\#U_k)} \left( h(\mathbf{m}_{km}) - \mathbf{u}_{km} \right)^\top \cdot J_h(\mathbf{m}_{km}) \cdot \mathbf{P}_k \frac{\partial \mathbf{T}_k(\boldsymbol{\alpha})}{\partial \alpha_l} \mathbf{x} \right), \quad (4.17)$$

where  $J_h(\mathbf{a}) : \mathbb{R}^3 \rightarrow \mathbb{R}^{2 \times 3}$  denotes the Jacobian of function  $h(\mathbf{a})$  and is given by

$$J_h(\mathbf{a}) = \begin{pmatrix} \frac{1}{a_3} & 0 & -\frac{a_1}{a_3^2} \\ 0 & \frac{1}{a_3} & -\frac{a_2}{a_3^2} \end{pmatrix}. \quad (4.18)$$

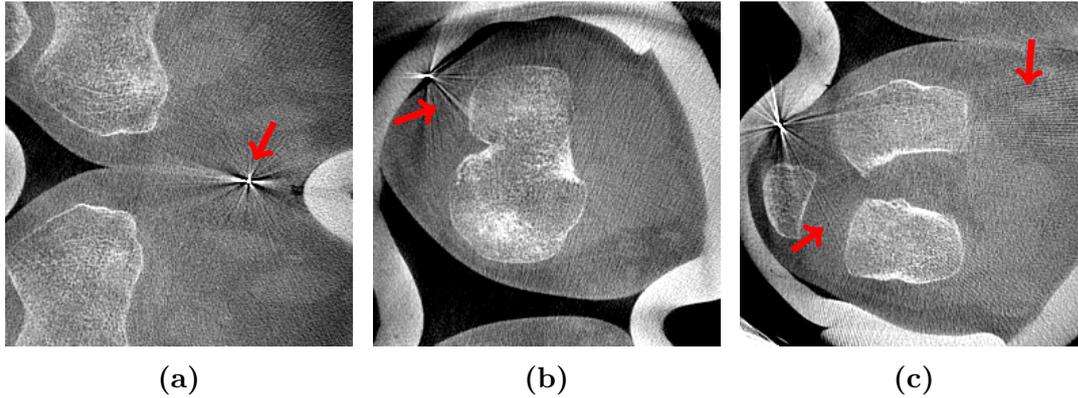
A detailed definition of the construction of  $\mathbf{T}_k(\boldsymbol{\alpha})$ , including its derivatives  $\frac{\partial \mathbf{T}_k(\boldsymbol{\alpha})}{\partial \alpha_l}$ , can be obtained from appendix A.1.

Optimization is done simultaneously for all projection images, using a nonconstrained gradient-based minimization method. The gradient is computed analytically and the Hessian is approximated using the BFGS method [Noce 06]. The step's direction is then computed by attempting a Newton step, based on the approximated Hessian. The step-size is calculated by a line-search method. The rotational parameters are optimized in degrees instead of radians to ensure that rotation and translation parameters are in a similar range. For more information it is referred to the optimizer's documentation [Maie 10].

### 4.2.2 Iterative Detection and Removal of Outliers

In addition to a gradient-based optimization, a method is proposed to remove outliers. Outlier removal is necessary to handle false or missing detections in 2-D and wrong point correspondences. Within the scope of this thesis we presented an outlier removal technique based on 3-D spline fitting of the marker trajectory in projection space [Mull 15a, Mull 15b]. However, the focus here is on the iterative outlier removal presented by Berger et al. [Berg 16a].

The procedure is based on an iterative removal of the worst contributions, with repeated optimizations of the cost function in Eq. (4.16). After each optimization, those  $\mathbf{u}_{km}$  are identified that belong to the  $\delta\%$  of pairs that show the largest Euclidean



**Figure 4.3:** Example of streaking artifacts caused by metallic tantalum beads used for marker-based motion correction. The streaks also propagate through the objects, reducing image quality at diagnostically important regions (W: 2025 HU, C: 145 HU).

distances w. r. t. their forward projected reference point  $\mathbf{m}_{km}$ . Subsequently, those points  $\mathbf{u}_{km}$  are removed from the corresponding measurement set  $U_k$ . However, each time a maximum of one detection per projection image is removed, i. e., the one showing the largest distance. In addition, the point is only removed if the number of residual measurements for that projection image is higher than a threshold, i. e.,  $\#U_k > M_{\min}$ . The number of iterative repetitions of the optimization is fixed to  $N_{\text{opt}}$ .

### 4.3 Removal of Markers

One of the main drawbacks of a marker-based motion correction is, that the markers have a direct influence on the projection images. The attached markers for weight-bearing imaging are metallic tantalum beads with a diameter of 1 mm. The linear attenuation coefficient for tantalum is substantially higher than that of regular tissue or bone. On the one hand, this allows a relatively straightforward detection of marker positions as described in Sec. 4.1.2.

On the other hand, such high attenuation coefficients can cause photon starvation, i. e., none or only very little amount of energy is measured at pixel locations covered by a marker. This causes high or even infinite values, after conversion to the line integral domain using Eq. (3.3). The ramp filtering step of a reconstruction algorithm yields a high response, due to the small size of the markers, their high absorption values, and the filter's high-pass characteristics. In combination with the limited detector, but also angular resolution, these effects cause severe streaking artifacts in the reconstruction domain. Examples of typical metal artifacts are presented in Fig. 4.3, where we can clearly see the star-like appearance of the markers. Note that the centers of the markers do not substantially influence the diagnosis, as most of the interesting anatomies are located inside the knee. However, the streaks also propagate through the internal structures of the knee, reducing the diagnostic value of the images, as can be seen in Figures 4.3(b) and 4.3(c).

To avoid a reduction in image quality the metal artifacts need to be corrected. Large regions are covered by the artifacts in the reconstruction domain, whereas the

markers are well localized in the projection domain. Thus, the idea is to remove them directly in the projection images. We can use the detected marker positions  $\mathbf{u}_{km}$  after outlier removal to localize the markers, however, due to missing detections and outlier removal not all markers will be captured. Therefore, the missing locations are filled with the forward projections of  $\mathbf{x}_m$  using the estimated motion parameters. Let us define

$$\tilde{U}_k = U_k \cup \{\mathbf{m}_{km} \mid \mathbf{u}_{km} \notin U_k \wedge \mathbf{m}_{km} = h(\mathbf{P}_k \cdot \mathbf{T}_k(\boldsymbol{\alpha}) \cdot \underline{\mathbf{x}}_k)\} , \quad (4.19)$$

to be sets which contain 2-D locations for all markers, thus,  $\#\tilde{U}_k = M$ . Removal of markers is now modeled as a missing data problem. It is assumed that the absorption values at marker locations, including a certain neighborhood, are defective and need to be extrapolated. In the process of this thesis a comparison was conducted to evaluate different extrapolation techniques for their ability to remove metal artifacts caused by markers [Berg 14a].

The sets  $\tilde{U}_k$  are used as an input to extract sets

$$W_{km} = \{\mathbf{u} \mid (\|\mathbf{u} - \mathbf{u}_{km}\|_2 < r_w) \wedge \mathbf{u}_{km} \in \tilde{U}_k\} \quad (4.20)$$

$$W_k = \{\mathbf{u} \mid \mathbf{u} \in W_{km} \wedge m \in [1, \dots, M]\} , \quad (4.21)$$

where  $W_{km}$  includes defective pixels for a single marker location and  $W_k$  all defects inside the  $k$ -th projection image. Further,  $r_w$  provides the invalid circular neighborhood. The extrapolation is separately executed for each marker based on the intensities inside a squared neighborhood, centered at the location of the marker. These regions can be conveniently defined by a set

$$\Omega_{km} = \{\mathbf{u} \mid (\|\mathbf{u} - \mathbf{u}_{km}\|_\infty < N_w/2) \wedge \mathbf{u}_{km} \in \tilde{U}_k\} , \quad (4.22)$$

where the regions' side length  $N_w$  is incorporated by the infinity norm  $\|\cdot\|_\infty$ .

The task of an extrapolation method is to estimate the defective values given by  $(\Omega_{km} \cap W_{km})$  based on the known intensities provided by  $(\Omega_{km} \setminus W_{km})$ . Note that only missing values  $W_{km}$  of one marker are extrapolated, based on all valid pixel intensities in  $\Omega_{km}$ . This is important in case multiple defects are located within a squared neighborhood.

### 4.3.1 Comparison of Extrapolation Methods

Evaluation included six extrapolation methods for estimation of the missing data. As an intuitive approach, [bilinear interpolation \(LinInt\)](#) was applied, which estimates the missing pixel by summation of closest existing pixels, weighted according to their distance. Another method, which was proposed by Mitrovic et al. [Mitr 13], is based on the fitting of 1-D cubic [B-splines \(BSpl\)](#) through the known intensities over row and column directions. Intensities at defect pixels are computed by the average of the corresponding row and column spline. A natural extension to this concept are [thin plate smoothing splines \(TPSS\)](#), which are similar to regular [TPS](#) introduced in Sec. 3.4.1, but do not enforce that the spline fits the values at control points exactly. Control points are the locations of valid pixels, with its values being the

pixel intensities. Extrapolation is then simply a readout of the fitted spline model at defect pixel locations.

Two methods are used that are extensions to a Gaussian low-pass filtering. First, **normalized convolution (NConv)** is applied as previously proposed by Knutsson et al. [Knut 93]. It represents a pixel-wise division of the low-pass filtered projection image and the low-pass filtered defect mask

$$p_k^{\text{corr}}(\mathbf{u}) = \frac{g_{\sigma_{\text{nc}}}^2(\mathbf{u}) * (p_k(\mathbf{u}) w_k(\mathbf{u}))}{g_{\sigma_{\text{nc}}}^2(\mathbf{u}) * w_k(\mathbf{u})} \quad (4.23)$$

$$w_k(\mathbf{u}) = \begin{cases} 0 & \text{if } \mathbf{u} \in W_{km} \\ 1 & \text{otherwise} \end{cases}, \quad (4.24)$$

where  $(p_k(\mathbf{u}) w_k(\mathbf{u}))$  sets all defect pixels to 0 and  $g_{\sigma_{\text{nc}}}^2(\mathbf{u})$  represents a 2-D Gaussian low-pass filter with standard deviation  $\sigma_{\text{nc}}$ . In addition the **subtract-and-shift (SAS)** approach proposed by Schwemmer et al. [Schw 10] is applied. It aims to remove high-density objects such as catheters while preserving occluded edge information. First, the projection images are filtered with a high-pass filter. The resulting intensities are used to replace defect pixel regions. As a final step the mean of the defect regions is shifted such that it aligns with the neighboring, valid pixels.

The last method is **spectral interpolation (SpecInt)** proposed by Aach et al. [Aach 01]. It is an iterative deconvolution method defined in frequency domain that estimates the entire corrected projection image based on the spectrum of  $p_k(\mathbf{u})$  and  $w_k(\mathbf{u})$ . It has been shown that **SpecInt** minimizes the **SSD** between corrected and uncorrected image over all valid positions in  $(\Omega_{km} \setminus W_k)$ . Finally, only the defect pixels are copied from the resulting corrected image.

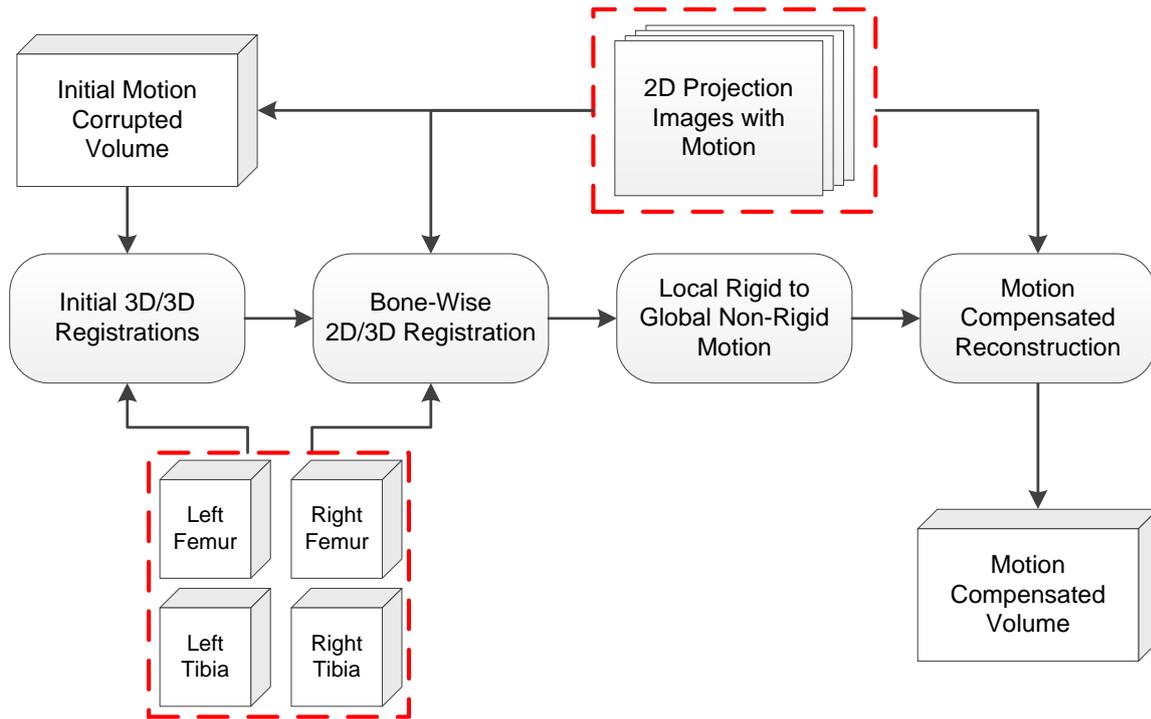
# Image-Based Motion Correction Using 2-D/3-D Registration

5.1 Method Overview .....	60
5.2 Digitally Rendered Radiographs .....	61
5.3 Similarity Measures .....	62
5.4 Temporal Regularization and Optimization .....	64
5.5 Nonrigid Motion Field Generation .....	65
5.6 Noise Reduction in DRR Images .....	66

A marker-based motion correction method for weight-bearing CBCT imaging of knees, has been presented in Sec. 3.4.2. As part of this thesis, several extensions to the marker-based methods have been made which are explained in more detail in Chap. 4. Even though this method showed promising results for motion correction, several aspects reduce its practical applicability.

First, the time needed for attaching the markers is not to be underestimated. Markers need to be attached carefully around the knee, such that as little as possible overlaps in projection domain occur. Also their arrangement should be well distributed in 3-D to achieve a well conditioned cost function. This additional time interferes directly with the acquisition protocol which may cause patient discomfort and inconveniences in clinical routine. It also remains uncertain whether the motion estimated by markers attached to the skin precisely reflects the internal motion at the joint which may be nonrigid due to joint rotations. Finally, the markers modify the projection images, causing streaking artifacts of high intensity which require removal of the markers prior to reconstruction, as shown in Sec. 4.3.

As a consequence a motion correction approach that can be applied without alterations to the acquisition protocol is desirable. Thus, the total acquisition time can be substantially reduced which is of great benefit for patients as well as clinics. A multitude of motion estimation algorithms based on an additional artifact-free acquisition are presented in Sec. 3.4.3, usually requiring some sort of image registration techniques. All clinical studies performed in context of this thesis include a prior, motion-free acquisition in supine position, as shown in Sec. 2.3. This chapter describes how this data may be used for nonrigid motion estimation based on 2-D/3-D registration. It starts with an overview in Sec. 5.1 that provides an outline of the proposed method and includes a short summary of 2-D/3-D registration. Methodologies

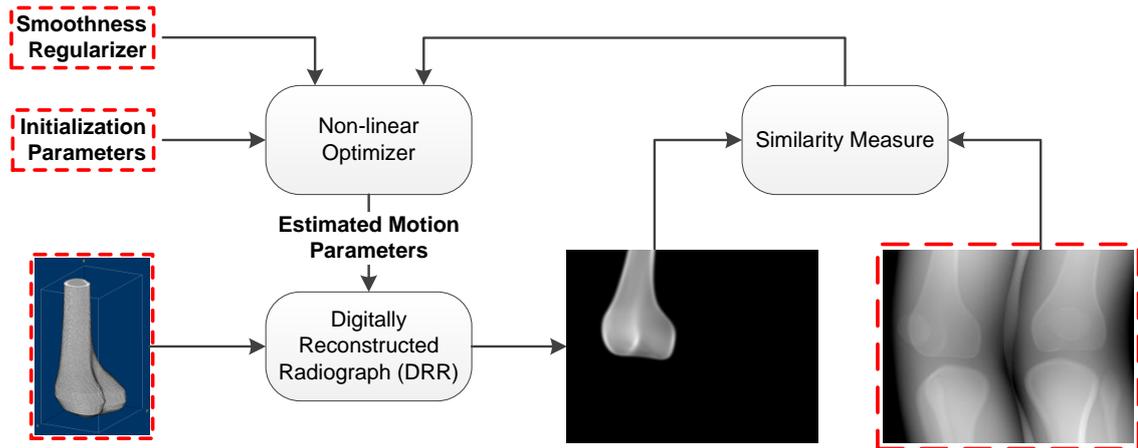


**Figure 5.1:** Outline of the registration-based motion correction approach. Inputs are segmented volumes of femur and tibia and the projection images acquired under weight-bearing conditions.

for the generation of DRR images are presented in Sec. 5.2. Two different similarity measures are used for image registration. They are introduced in Sec. 5.3, followed by the details on cost function optimization and temporal regularization in Sec. 5.4. The algorithm includes the extrapolation of nonrigid motion which is outlined in Sec. 5.5. Finally, Sec. 5.6 discusses a method to improve registration accuracy in case of real acquisitions. This chapter builds on a conference publication [Berg 15] where the method was applied to simulated data and a journal publication [Berg 16a] that provides evaluations on real data and several other advancements.

## 5.1 Method Overview

The important steps of the algorithm are outline in Fig. 5.1. The main idea is to apply 2-D/3-D registration between each projection image and each segmented bone volume. It can be assumed that bones move only rigidly over time, thus the registration for each bone is limited to a rigid motion model. A total of eight bones is present in two knees (i.e., femur, tibia, patella, and fibula). However, to reduce complexity of the registration approach we focus only on femur and tibia, which build the knee joint and are also the largest bones. Thus, the inputs to the method are the acquired weight-bearing projection image stack, as well as segmented volumes of the motion-free scan in supine position, showing tibia and femur from left and right leg. For improved clarity the input data is emphasized by a dashed, red frame in Fig. 5.1. As a first step the segmented volumes need to be roughly aligned to the weight-bearing



**Figure 5.2:** Overview of a 2-D/3-D registration approach. DRR images are rendered from the volume based on the current motion estimate. These are compared by a similarity measure, which is iteratively optimized by adjusting the motion estimates.

scan, as they originate from an entirely different coordinate system, i. e., in supine and standing position. This is achieved by an initial 3-D/3-D registration between the segmented bones and a reconstruction of the weight-bearing acquisition without motion correction. As the pose of the bones among one other may be different in supine and standing position, this step is required for each of the segmentations. It is worth noting that the 3-D/3-D registration does not correct for any patient motion but merely aligns the coordinate systems of the segmentations and the weight-bearing acquisition to provide a reasonable initialization for the subsequent 2-D/3-D registration. Patient motion is now estimated by registration of each bone to each of the acquired projection images, yielding the individual rigid motion parameters for each bone over time. As a final step the nonrigid motion between the bones, e.g., at muscle or skin tissue, is estimated by a TPS extrapolation as explained in Sec. 3.4.1, resulting in individual 3-D deformation fields for each projection image.

2-D/3-D registration is the most essential part of the algorithm as its accuracy has a direct impact on the image quality of the final reconstruction. The overview of the proposed 2-D/3-D registration method is shown in Fig. 5.2 and refers to the “projection strategy” as explained in [Mark 12]. Ray casting is used to build DRR projection images of the segmented volumes, given the currently estimated motion parameters. These images are then compared to the acquired projections using a suitable similarity measure. A nonlinear optimizer aims to iteratively minimize (or maximize) the similarity measure by adjustment of the motion parameters. To incorporate prior information on the parameters a regularizer may be provided to the optimizer, e. g., to enforce smooth variations of motion parameters over time.

## 5.2 Digitally Rendered Radiographs

The task of creating a DRR, given motion parameters and the volume data, is often done by ray casting [Goit 83]. For each detector pixel discrete line integrals through the volume are computed on rays connecting the virtual X-ray source and the pixel

location. Essentially, it is a modified version of the X-ray transform for CBCT geometries described in Eq. (3.13), where the source location  $\mathbf{c}_k(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{R}^3$  and the ray direction  $\mathbf{r}_k(\mathbf{u}, \boldsymbol{\alpha}) : \mathbb{R}^{L+2} \rightarrow \mathbb{S}^3$  can be extracted from the corresponding projection matrix  $\mathbf{P}_k(\boldsymbol{\alpha})$  [Hart 04, pp. 161f.]. Rigid motion may be incorporated by  $\boldsymbol{\alpha}$  according to Eq. (3.36). Using Eq. (3.13) we can define the motion dependent DRR images  $d_k(\mathbf{u}, \boldsymbol{\alpha}) : \mathbb{R}^{L+2} \rightarrow \mathbb{R}$  by

$$d_k(\mathbf{u}, \boldsymbol{\alpha}) = \int_0^\infty f\left(\mathbf{c}_k(\boldsymbol{\alpha}) + \beta \mathbf{r}_k(\mathbf{u}, \boldsymbol{\alpha})\right) d\beta, \quad (5.1)$$

where  $\beta$  is the differential of the line integral which usually has its upper limit when the ray intersects with the detector. A majority of the similarity measures used in skeletal 2-D/3-D registration are based on the gradient images. Similar to the gradient of the acquired projections defined in Eq. (4.1), we can obtain the gradient of  $d_k(\mathbf{u}, \boldsymbol{\alpha})$  by

$$\nabla d_k(\mathbf{u}, \boldsymbol{\alpha}) = \begin{pmatrix} \frac{\partial d_k(\mathbf{u}, \boldsymbol{\alpha})}{\partial u_1} \\ \frac{\partial d_k(\mathbf{u}, \boldsymbol{\alpha})}{\partial u_2} \end{pmatrix}. \quad (5.2)$$

An estimation of  $\nabla d_k(\mathbf{u}, \boldsymbol{\alpha}) : \mathbb{R}^{L+2} \rightarrow \mathbb{R}^2$  can be done by convolutions with the Sobel filter kernels for horizontal and vertical direction. In addition to the computational complexity, this may also cause artificial edges in DRRs in case the segmented volume is truncated.

In the presented method, we avoid these effects by generating the gradient DRRs directly, based on the precomputed gradient  $\nabla f(\mathbf{x}) : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , defined in volume domain. It was shown by Livyatan et al. and Wein et al. [Livy 03, Wein 05], that

$$\nabla d_k(\mathbf{u}, \boldsymbol{\alpha}) = \mathbf{D}_k^{uv}(\boldsymbol{\alpha}) \int_0^\infty \beta \cdot \nabla f\left(\mathbf{c}_k(\boldsymbol{\alpha}) + \beta \mathbf{r}_k(\mathbf{u}, \boldsymbol{\alpha})\right) d\beta, \quad (5.3)$$

where the volume gradient is weighted by the length of the line integral  $\beta$  during integration. Further, the projection of the volume gradient to the detector domain is done by building the inner product with directions of the detector  $u$ - and  $v$ -axis, in the world coordinate system. They are represented by a function  $\mathbf{D}_k^{uv}(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{R}^{2 \times 3}$  that returns a matrix which has the directions of  $u$ - and  $v$ -axis on its rows. Note that  $\mathbf{D}_k^{uv}(\boldsymbol{\alpha})$  can be extracted from the projection matrices  $\mathbf{P}_k(\boldsymbol{\alpha})$ .

### 5.3 Similarity Measures

A challenging part of 2-D/3-D registration is to find an appropriate similarity function to compare DRRs with the acquired data. Two different measures are compared in this work w. r. t. the final image quality of the reconstruction. The well known GC [Penn 98] is used, which computes the normalized cross correlation between DRR and vertical as well as horizontal gradient images. Additionally, the NGI is applied, that has been reported to provide increased stability in presence of outlier intensities and shall be a better choice in case of overlapping structures [Otak 13]. Both measures compare projection images in the gradient domain and therefore have their focus on aligning edges.

### 5.3.1 Gradient Correlation

The most popular similarity measure for 2-D/3-D registration of high density structures is **GC**. It was initially proposed by Penney et al. in context of a comparison of similarity measures [Penn 98]. In the following we present the formal definition of the **GC** measure. The gradient of the **DRR**  $\nabla d_k(\mathbf{u}, \boldsymbol{\alpha})$  is directly rendered using Eq. (5.3), whereas the gradient of the acquired projection images  $\nabla p_k(\mathbf{u})$  is computed prior to optimization using convolution-based Sobel filtering. The **GC** is defined to be the mean of normalized cross correlations between the individual partial derivatives in  $\nabla p_k(\mathbf{u})$  and  $\nabla d_k(\mathbf{u}, \boldsymbol{\alpha})$ . It may be formulated by

$$GC(p, d, \boldsymbol{\alpha}) = \frac{1}{2} \sum_{k=1}^K \sum_{\mathbf{u} \in \Pi_k} \left( (\nabla d_k(\mathbf{u}, \boldsymbol{\alpha}))^\top \cdot \mathbf{Y}^{-1} \cdot \nabla p_k(\mathbf{u}) \right) . \quad (5.4)$$

The set  $\Pi_k$  defines **ROIs** on the projection images used for computation of the **GC** measure. The regions may differ for different view angles which is indicated by the subscript  $k$ . Normalization is important to account for intensity differences of acquired and **DRR** images. The normalization part of the individual cross correlations are implemented by a division with the standard deviations of gradient intensities in regions  $\Pi_k$ , which is incorporated by the weighting matrix  $\mathbf{Y}$ . We define the set  $\Pi_k$  by

$$\Pi_k = \{ \mathbf{u} \mid \|\nabla d_k(\mathbf{u}, \boldsymbol{\alpha})\|_2 > 0 \} , \quad (5.5)$$

such that only those areas are included where the **DRR** images have nonzero values. The weighting matrix is given by

$$\mathbf{Y} = \begin{pmatrix} \sqrt{\sum_{\mathbf{u}} \left( \frac{\partial d_k(\mathbf{u}, \boldsymbol{\alpha})}{\partial u_1} \right)^2} & \sqrt{\sum_{\mathbf{u}} \left( \frac{\partial p_k(\mathbf{u})}{\partial u_1} \right)^2} & 0 \\ 0 & 0 & \sqrt{\sum_{\mathbf{u}} \left( \frac{\partial d_k(\mathbf{u}, \boldsymbol{\alpha})}{\partial u_2} \right)^2} \sqrt{\sum_{\mathbf{u}} \left( \frac{\partial p_k(\mathbf{u})}{\partial u_2} \right)^2} \end{pmatrix} . \quad (5.6)$$

Using cross correlations on the gradient images, rather than on the pixel intensities leads to an increased robustness in presence of background materials such as tissue, muscle or skin.

### 5.3.2 Normalized Gradient Information

**NGI** has been developed with the goal to achieve an increased robustness in presence of overlapping structures that can cause misleading edges. Gradient directions are compared by computing the dot product for each location in  $\Pi_k$ . The result is then weighted pixel-wise by the gradient magnitude. An improve robustness is achieved by using only the minimum of the two gradient magnitudes as a weighting factor, corresponding to either **DRR** or acquired image. It was shown that this weighting scheme is more robust in presence of overlapping structures causing interfering edges [Otak 13]. However, compared to the **GC** approach, no intensity normalization can be applied. As a consequence we propose a manual scaling of the **DRRs** to roughly adjust their gray levels to those of the acquired data. The equation for the **NGI** can be given by

$$NGI(p, d, \boldsymbol{\alpha}) = \frac{GI(p, d, \boldsymbol{\alpha})}{GI(p, p)}, \quad (5.7)$$

where the numerator is defined by

$$GI(p, d, \boldsymbol{\alpha}) = \sum_{k=1}^K \sum_{\mathbf{u} \in \Pi_k} \left( \frac{1}{2} \frac{(\nabla d_k(\mathbf{u}, \boldsymbol{\alpha}))^\top \nabla p_k(\mathbf{u})}{\|\nabla d_k(\mathbf{u}, \boldsymbol{\alpha})\|_2 \|\nabla p_k(\mathbf{u})\|_2} + \frac{1}{2} \right) \times \min(\|\nabla d_k(\mathbf{u}, \boldsymbol{\alpha})\|_2, \|\nabla p_k(\mathbf{u})\|_2) \quad (5.8)$$

and the constant denominator by

$$GI(p, p) = \sum_{k=1}^K \sum_{\mathbf{u} \in \Pi_k} \|\nabla p_k(\mathbf{u})\|_2. \quad (5.9)$$

In case of the *NGI* measure, the *ROIs* defined by  $\Pi_k$  can be adjusted such that they cover all pixel locations. This is because the gradient magnitude of  $\nabla d_k(\mathbf{u}, \boldsymbol{\alpha})$  is zero at locations which are not covered by the projection of the volume. These contributions have no influence on the cost function as they are canceled out by the minimum operator in Eq. (5.8).

## 5.4 Temporal Regularization and Optimization

In contrast to the marker-based approach presented in Chap. 4, the parameter vector  $\boldsymbol{\alpha}$  has a length of  $L = 6KB$  and consists of six rigid parameters for each of  $K$  projections but also for each of the  $B$  segmented bone volumes. The arrangement of parameters inside vector  $\boldsymbol{\alpha}$  is given by

$$\boldsymbol{\alpha} = \left( \boldsymbol{\zeta}_{11}^\top, \dots, \boldsymbol{\zeta}_{K1}^\top, \boldsymbol{\zeta}_{12}^\top, \dots, \boldsymbol{\zeta}_{KB}^\top \right)^\top, \quad (5.10)$$

where  $\boldsymbol{\zeta}_{kb} \in \mathbb{R}^6$  are vectors with rotation and translation parameters for the  $k$ -th projection and the  $b$ -th segmented volume. They are defined by

$$\begin{aligned} \boldsymbol{\zeta}_{kb} &= \left( \phi_{kb}^x \quad \phi_{kb}^y \quad \phi_{kb}^z \quad t_{kb}^x \quad t_{kb}^y \quad t_{kb}^z \right)_{kb}^\top \\ &= \left( \alpha_{i+1} \quad \alpha_{i+2} \quad \alpha_{i+3} \quad \alpha_{i+4} \quad \alpha_{i+5} \quad \alpha_{i+6} \right) \\ \hat{l} &= (6K(b-1) + 6(k-1)), \end{aligned} \quad (5.11)$$

where  $\hat{l}$  determines the position in  $\boldsymbol{\alpha}$ , given  $k$  and  $b$ . The dimensionality of the optimization problem can be very large, where the number of unknowns are computed by  $6KB$ , which equals up to 24 parameters per projection image in our case. Thus, a regularization of the parameters may be beneficial to improve convergence speed and robustness of the optimization. We assume that the variation between motion states is physically limited by the anatomy of the knee joint. Therefore, a temporal smoothness regularization is added to the original cost function. The goal is to limit high temporal variations, computed by subtraction of the estimated parameters by

their Gaussian filtered version. This represents a minimization of high temporal frequencies contained in the motion parameters.

$$r_{\text{tmp}}(\boldsymbol{\alpha}) = \sum_{k=1}^K \sum_{b=1}^B \left\| \boldsymbol{\zeta}_{kb} - \left( \boldsymbol{\zeta} * g_{\sigma_t}(k) \right)_{kb} \right\|_2^2 \quad (5.12)$$

Eq. (5.12) shows the temporal regularizer  $r_{\text{tmp}}(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{R}$ , where  $g_{\sigma_t}(k)$  is a 1-D Gaussian filter kernel, where  $\sigma_t$  defines the standard deviation over the temporal direction  $k$ . Further,  $(\boldsymbol{\zeta} * g_{\sigma_t}(k))$  denotes the convolution that applies the filter in direction of  $k$ .

The combined cost function is then built using the method of Lagrangian multipliers

$$\underset{\boldsymbol{\alpha}}{\text{argmin}} \quad -GC(p, d, \boldsymbol{\alpha}) + \eta r_{\text{tmp}}(\boldsymbol{\alpha}) \quad , \quad (5.13)$$

where  $GC(p, d, \boldsymbol{\alpha})$  may be replaced by  $NGI(p, d, \boldsymbol{\alpha})$  if desired. Note that **GC** but also **NGI** need to be maximized. This is usually done by minimization of the negative similarity measure as indicated in Eq. (5.13).

The same nonlinear optimizer as for the marker-based cost function is used (cf. Sec. 4.2). All parameters are estimated simultaneously as the regularization prohibits sequential optimization. The gradient is estimated by forward-differences and the Hessian is approximated using the BFGS method [Noce 06]. The step's direction is then computed by attempting a Newton step based on the approximated Hessian. The step-size is calculated by a line-search method. The rotational parameters are optimized in degrees instead of radians to ensure that rotation and translation parameters are in a similar range. For more information it is referred to the optimizer's documentation [Maie 10].

## 5.5 Nonrigid Motion Field Generation

Before a motion compensated reconstruction can be performed the individual bone motions estimated by the 2-D/3-D registration are combined to a global, nonrigid motion. This is done by the **TPS** extrapolation introduced in Sec. 3.4.1. The goal is to obtain a separate motion field  $\mathbf{d}_k(\mathbf{x}, \boldsymbol{\alpha})$  for each projection image. During segmentation of the bones in the artifact-free, supine scan, not only volumes are obtained but also the bones' surfaces represented by meshes of triangles. The meshes consist of surface normals and vertices, where the latter can be used directly as control points for the **TPS**. Let  $\mathbf{T}_b^{\text{init}} \in \mathbb{R}^{4 \times 4}$  be the initial rigid motion that was determined by manual 3-D/3-D registration of the segmented bones and a weight-bearing reconstruction without motion correction. Also,  $\hat{\mathbf{T}}_{kb} \in \mathbb{R}^{4 \times 4}$  are the results of the 2-D/3-D registration for each bone and projection image. Note that the dependency of motion matrices on parameter vector  $\boldsymbol{\alpha}$  is omitted for better visualization.

A common problem when performing motion correction is that no fixed coordinate system exists due to the patient motion, where all motion states represent a valid coordinate system. Therefore, a time point  $\hat{k}$  needs to be selected as a reference, such that the common coordinate system may be presented by  $\hat{\mathbf{T}}_{\hat{k}b}$ . It is recommended to choose  $\hat{k}$  such that it corresponds to a view direction that guarantees a

high registration accuracy, e. g., a projection image with good image quality and no occlusions or overlapping structures. For knees, a posteroanterior viewing direction is recommended to avoid overlapping bones. After selection of  $\hat{k}$ , all estimated transformations need to be adjusted to make sure that they also operate in the coordinate system of the reference. The adjusted motion matrices  $\mathbf{T}_{kb}$  can be obtained by

$$\mathbf{T}_{kb} = \hat{\mathbf{T}}_{kb} \hat{\mathbf{T}}_{\hat{k}b}^{-1} . \quad (5.14)$$

Let us now define  $\hat{\boldsymbol{\vartheta}}_b^a \in \mathbb{R}^3$  to be the  $a$ -th vertex located on the surface of the  $b$ -th bone. Note that the index  $a$  is omitted in the following derivations to preserve clarity. All vertices are defined in the supine coordinate system and are now transformed to the reference coordinate system of the weight-bearing scan. We define  $\boldsymbol{\vartheta}_b \in \mathbb{R}^3$  to be the static vertex location for time point  $\hat{k}$ . It can be computed by

$$\boldsymbol{\vartheta}_b = \left( \hat{\mathbf{T}}_{\hat{k}b} \mathbf{T}_b^{\text{init}} \right) \hat{\boldsymbol{\vartheta}}_b , \quad (5.15)$$

where the corresponding moving vertices are given by

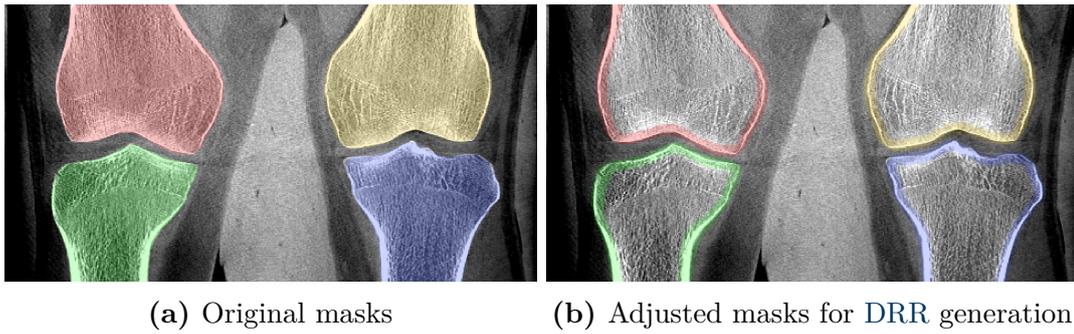
$$\boldsymbol{\vartheta}_{kb} = \mathbf{T}_{kb} \boldsymbol{\vartheta}_b . \quad (5.16)$$

We can now compute known translations for each vertex and time point by  $(\boldsymbol{\vartheta}_{kb} - \boldsymbol{\vartheta}_b) \in \mathbb{R}^3$ . These translations can be plugged into the TPS equation presented in Eq. (3.37), which yields a system of linear equations that can be solved for the splines' coefficients by a closed form solution [Davi97]. To constrain the TPS model at the periphery of the FOV the corners of the reconstructed volume are added as control points. The translations at corner locations are computed using the rigid motion of the geometrically closest bone, e.g., for the left-superior corners the motion of the left femur is used.

The resulting deformation function  $\mathbf{d}_k(\mathbf{x}, \boldsymbol{\alpha})$  can now be incorporated into a motion compensated reconstruction. The regular FDK algorithm is used as described in Sec. 3.2.3, yet, the updated backprojection approach given by Equations (3.33) and (3.45) is needed (cf. Sec. 3.5). Note, that  $\mathbf{d}_k(\mathbf{x}, \boldsymbol{\alpha})$  needs to be evaluated during backprojection for each voxel location  $\mathbf{x}$  and each projection image, which increases computational complexity.

## 5.6 Noise Reduction in DRR Images

When applied to real data, we could observed a high noise level in the gradient of the DRRs which influenced the accuracy of the 2-D/3-D registration method. The noise originates from the segmented bones in the supine scan. The following steps are suggested to reduce the effect of noise. First, edge-preserving noise filtering [Maie11b] can be applied in the supine reconstruction, where we suggest to apply a 3-D bilateral filter [Lorc15, Maie15a] prior to the calculation of the volume gradient. Further noise reduction can be achieved by a manipulation of the segmentation masks. As the trabecular bone as well as the bone marrow are rather homogeneous they contain only few structures useful for the 2-D/3-D registration when using gradient-based similarity measures. Hence, major parts of these regions can be removed during



**Figure 5.3:** Variation to 3-D segmentations to reduce the amount of noise in DRRs.

segmentation such that only cortical bone is left which corresponds to the outline of the bones.

This removal step is done with 3-D morphological operators. First, an erosion applied to the original segmentation masks is applied and in another step a dilation. Subtracting eroded from the dilated mask gives a segmentation that has its focus on the bone outline. The mask adjustment is illustrated in Fig. 5.3.

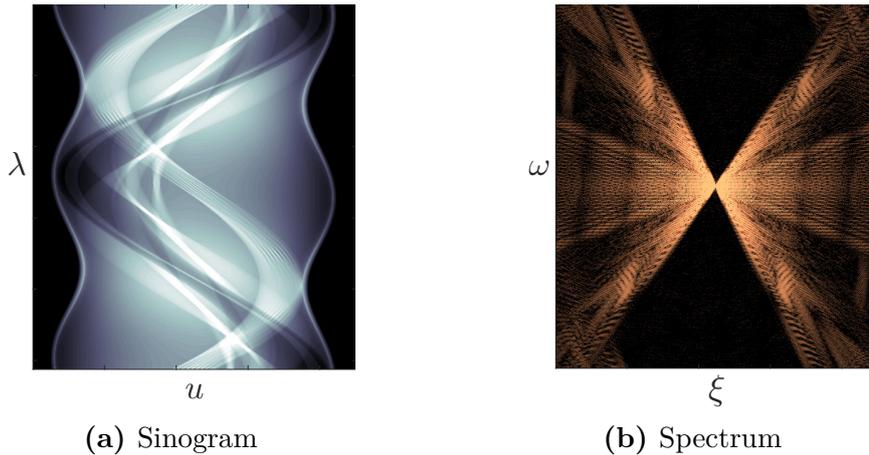


# Image-Based Motion Correction Using Consistency Conditions

6.1	FCC and their Extension to CBCT .....	70
6.2	FCC With Arbitrary Motion Models .....	71
6.3	Efficient Implementation for Detector Translations .....	72
6.4	Regularization, Optimization and Motion Compensation .....	75
6.5	Practical Considerations .....	76

It was shown that motion correction for weight-bearing CBCT of knees may be achieved by a marker-based approach. However, markers cause several problems during acquisition, but also for the image reconstruction process, as outlined in Chap. 4. As part of this thesis a surrogate-free motion correction method is proposed that aims to alleviate the drawbacks caused by a marker-based approach. It is based on 2-D/3-D registration of a prior, artifact-free reconstruction to the projection images of a weight-bearing scan and is presented in Chap. 5. Current research protocols for weight-bearing imaging of the knee joint involve this type of prior acquisition to obtain a non weight-bearing reference volume, required for associated clinical studies. However, in a potential clinical application this prior scan may not always be vital for diagnosis, therefore, causing unnecessary dose for the patient. In case no markers are used and also no additional motion-free scan is available, motion correction needs to be done using only the acquired projection images. Methods that fall in this group are introduced in Sec. 3.4.4 and are either based on assumptions of the image's intensities in reconstruction domain or on CC defined in the projection domain. A benefit of using CC is, that they do not require a reconstruction to evaluate the cost function and are therefore computationally efficient.

Part of this thesis is a method based on FCC defined in the projection domain, as introduced in Sec. 3.4.4. FCC have been extended to the cone-beam geometry to find ideal sampling conditions for CBCT [Desb04, Brok06]. Up to now they have not been used for artifact correction. This chapter is structured as follows: First, a practical extension of FCC from fan- to cone-beam geometry is given in Sec. 6.1, based on findings of Brokish et al. [Brok06]. Using the derived FCC, a cost function for motion estimation based on arbitrary motion models is defined in Sec. 6.2. An extension to a particularly efficient implementation of the cost function is given in Sec. 6.3. It is based on a motion model that estimates horizontal but also vertical



**Figure 6.1:** Example of a sinogram and the triangular regions of the FCC.

detector shifts with the goal of correcting for 3-D translational object motion. In Sec. 6.3.2 it is shown that the cost function is invariant to constant detector offsets. A regularizer to cope with this invariance is proposed in Sec. 6.4, together with details on optimization and motion compensated reconstruction. Finally, the discretization issues as well as implementation details are discussed in Sec. 6.5. The concepts presented in this chapter are based on a conference contribution proposed for the fan-beam geometry [Berg 14b] and a journal publication dedicated to an application to CBCT [Berg 16b].

## 6.1 FCC and their Extension to CBCT

The basic concept of FCC have been introduced in Sec. 3.4.4, where a detailed review of various CC is given. In Fig. 6.1 an illustration of the property is provided. Fig. 6.1(a) shows a fan-beam sinogram of a head phantom and Fig. 6.1(b) depicts the logarithmic representation of the absolute value of the sinogram's spectrum, where we can clearly see the triangular regions defined by the FCC. As shown in Sec. 3.4.4, the FCC for fan-beam CT can be defined by the inequality given in Eq. (3.43) which is an approximate definition proposed by [Mazi 10]. Thus, the shape and orientation of the triangles in Fig. 6.1(b) may be defined by

$$\left| \frac{\omega}{\omega - \xi(S + D)} \right| > \frac{r_p}{S}, \quad (6.1)$$

where  $\omega$  is the frequency axis corresponding to the rotation angle  $\lambda$  and  $\xi$  is the frequency axis associated with the detector  $u$ -axis. Their units are given by  $\text{rad}^{-1}$  and  $\text{mm}^{-1}$ , respectively. As explained in Sections 3.2.2 and 3.2.3 the variables  $S$  and  $D$  describe the acquisition geometry and correspond to SID and DID, respectively. Further,  $r_p$  defines the maximum extent of the object w.r.t. the rotation center of the acquisition system. In general, CC describe mathematical properties of the projection data, that need to be fulfilled by an ideal acquisition. The FCC is valid for continuous projection images, using a full rotation of the X-ray source on a perfect circular geometry. The projection images are also assumed to have an infinite extent,

i. e., no truncation is possible and the physical model is restricted to monochromatic absorption as described in Sec. 3.2. If these assumptions are not met, this can lead to an increase of energy in the triangular regions, which was exploited several times for artifact correction in fan-beam CT [Karp 88, Xia 95, Abel 09, Pohl 14].

In a publication associated with this thesis, Berger et al. [Berg 14b] showed that also unusual motion of the scanner or patient can violate the requirements of FCC. The total energy inside the triangular region increases when periodical translating the object during a scan. This observation allows us to use FCC as a cost function, such that the energy is minimized by optimization of a motion model defined in projection domain. Initial work for motion compensation in fan-beam CT has been presented in [Berg 14b].

For an extension to the cone-beam geometry we first introduce the frequency variable  $\psi$  for the vertical detector directions  $v$ . The idea is to apply the 2-D fan-beam properties of Eq. (6.1) directly in the 3-D Fourier domain of the projection images. Brokish et al. and Desbat et al. investigated the essential support of this Fourier domain, where they derived sampling conditions to find maximum sampling distances in detector and angular directions to allow for a reconstruction without aliasing artifacts [Brok 06, Desb 04]. The essential support can be interpreted as the opposite of the triangular regions, describing spectral areas that possibly contain information about the object. An important finding of Brokish et al. was, that for the majority of practical cone-beam devices, it is sufficient to determine the support of  $\omega$  for a fixed value of  $\psi = 0$ , which is equal to the 2-D fan-beam case. The variations of the support region along  $\psi$ -axis are especially small in presence of a small cone-angle [Brok 06]. This is the case for most C-arm CBCT devices, where the systems described in Chap. 2 have a cone-angle of  $(2\kappa') \approx 14^\circ$  and a fan-angle of  $(2\gamma') \approx 18^\circ$ . Based on a numerical simulation they show that the support's variation along  $\psi$  direction is negligible. Thus, we propose a straightforward extension of the fan-beam properties presented in Eq. (6.1) such that the triangular regions are simply extended in direction of  $\psi$ . This means that the zero regions in the 3-D Fourier transform of the projection stack are assumed to be independent of  $\psi$ . Thus, Eq. (6.1) also holds for the cone-beam geometry without further modification.

## 6.2 FCC With Arbitrary Motion Models

In the following, we present a cost function formulated for sampled data. Thus it can be applied directly to discretized projection images. The formulation is based on  $K$  projection images containing  $I$  and  $J$  pixels in  $u$  and  $v$  direction, respectively. This results in a total of  $C = I \times J \times K$  values that are measured during an acquisition. The cost function for CBCT may be defined by

$$e(\boldsymbol{\alpha}) = \frac{1}{2} \left\| \mathbf{W}(\mathbf{F} \mathbf{p}(\boldsymbol{\alpha})) \right\|_2^2, \quad (6.2)$$

where  $\mathbf{F} \in \mathbb{C}^{C \times C}$  is a symmetric discrete Fourier transform (DFT) matrix that performs the Fourier transform in all three dimensions and  $\mathbf{p}(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{R}^C$  computes a vector that contains all pixel intensities of the measured projection images. It is

based on a motion model defined in the projection domain given the motion parameters in  $\boldsymbol{\alpha}$ . The discretized mask which determines the triangular region is given by  $\mathbf{W} = \text{diag}(\mathbf{w}) \in \mathbb{R}^{C \times C}$ , which is the diagonal matrix of vector  $\mathbf{w} \in \mathbb{R}^C$  that holds the zero regions. The function  $\text{diag}(\cdot)$  creates a diagonal matrix where the diagonal elements are taken from the input vector.

Let us also define discrete values  $\xi_i$ ,  $\psi_j$  and  $\omega_k$  for the axes in Fourier domain corresponding to the detector directions  $u$  and  $v$  as well as to the rotation angles  $\lambda$ , respectively. Using the sampling intervals  $\Delta u$ ,  $\Delta v$ , and  $\Delta \lambda$ , defined in Sec. 3.2.5, they can be computed by

$$\begin{aligned}\xi_i &= \text{freq}(i, \Delta u, I) \\ \psi_j &= \text{freq}(j, \Delta v, J) \\ \omega_k &= \text{freq}(k, \Delta \lambda, K) ,\end{aligned}\tag{6.3}$$

where function  $\text{freq}$  is defined by

$$\text{freq}(idx, \Delta, Z) = \frac{idx - 1 - \lfloor \frac{Z}{2} \rfloor}{Z\Delta} .$$

$\text{freq}$  computes the discrete values of the frequency axis for a signal of length  $Z$ , a discrete index  $idx \in [1, \dots, Z]$ , and a sampling interval of  $\Delta$ . The mask vector  $\mathbf{w}$  can be defined by

$$\mathbf{w} = \left( w_{1,1,1}, \dots, w_{i,j,k}, \dots, w_{I,J,K} \right)^\top ,\tag{6.4}$$

where the indices refer to the pixel location and the projection index. As determined at the end of Sec. 6.1 we can use Eq. (6.1) without any alterations, to compute the entries of  $\mathbf{w}$  by

$$w_{i,j,k} = \begin{cases} 1, & \text{if } \left| \frac{\omega_k}{\omega_k - \xi_i(L+D)} \right| > \frac{r_p}{L} \\ 0, & \text{otherwise} \end{cases} .\tag{6.5}$$

### 6.3 Efficient Implementation for Detector Translations

In the following we restrict our motion model to 2-D translations of the detector. In contrast, many applications seek to compensate for 3-D movement of a patient. Still, translation parallel to the detector can be explained very well by detector shifts. The image coordinate system is aligned with the plane of rotation for CBCT acquisition. In that case, shifts in  $v$  direction directly account for translations along the  $z$ -axis in the world coordinate system, except that cone-beam artifacts may be slightly amplified. Accordingly, 3-D motion parallel to detector  $u$ -axis can be covered by shifts in detector  $u$ . Finally, a 3-D translation towards and away from the detector makes the object or patient appear bigger and smaller, which could be modeled by image scaling. The distance an object has to be moved towards the detector to introduce an error of about one pixel depends on fan and cone angles. In conventional C-arm CBCT data it is considerably larger than the amount of motion parallel to the detector, required to introduce an error of similar magnitude. In consequence, for practical

applications detector shifts can recover 3-D movements in very good approximation. In addition, translations of the detector can be applied in the 2-D Fourier domain of the individual projection images with help of the shift theorem. Thus computational complexity is reduced as the 2-D Fourier transform of the projection images may be precomputed. The parameter vector  $\boldsymbol{\alpha}$  has a length of  $L = 2K$  and is given by

$$\boldsymbol{\alpha} = (s_1, t_1, s_2, t_2, \dots, s_K, t_K)^\top, \quad (6.6)$$

where  $s_k$  and  $t_k$  are translations in  $u$  and  $v$  direction for projection index  $k$ . As the 2-D fast Fourier transform (FFT) of the projection images can be precomputed, an evaluation of the cost function uses only the 1-D Fourier transforms over  $\lambda$ . The 3-D FFT has a complexity of  $\mathcal{O}(C \log C)$ , which is now reduced to  $\mathcal{O}(C \log K)$ . The cost function in Eq. (6.2) can be reformulated to

$$e(\boldsymbol{\alpha}) = \frac{1}{2} \left\| \mathbf{o}(\boldsymbol{\alpha}) \right\|_2^2 = \frac{1}{2} \left\| \mathbf{W} (\mathbf{F}_\lambda (\mathbf{H}(\boldsymbol{\alpha}) \mathbf{F}_{uv} \mathbf{p})) \right\|_2^2, \quad (6.7)$$

where the function  $\mathbf{o}(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{C}^{C \times C}$  represents the inner part of the cost function, i. e., the argument of the squared L2 norm. Further,  $\mathbf{F}_\lambda \in \mathbb{C}^{C \times C}$  and  $\mathbf{F}_{uv} \in \mathbb{C}^{C \times C}$  are the DFTs over the angles and the projection images, respectively, and can also be denoted by matrix multiplications. Further,  $\mathbf{p} \in \mathbb{R}^C$  contains the acquired projection data without application of any motion motion model. The motion is now incorporated by a function  $\mathbf{H}(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{C}^{C \times C}$ , which builds a diagonal matrix of phase factors, encoding the translations provided by  $\boldsymbol{\alpha}$ . We define that first the rows  $u$ , then the columns  $v$ , and then the angles  $\lambda$  are mapped to the linear vector  $\mathbf{p}$ . Hence, the structure of  $\mathbf{H}(\boldsymbol{\alpha})$  is given by

$$\mathbf{H}(\boldsymbol{\alpha}) = \text{diag} \left( \left( e^{-i2\pi(\xi_1 s_1 + \psi_1 t_1)}, \dots, e^{-i2\pi(\xi_I s_1 + \psi_I t_1)}, \right. \right. \\ \left. \left. e^{-i2\pi(\xi_1 s_2 + \psi_1 t_2)}, \dots, e^{-i2\pi(\xi_I s_K + \psi_I t_K)} \right)^\top \right). \quad (6.8)$$

### 6.3.1 Gradient of Cost Function

Gradient-based optimizers can potentially yield faster convergence and more accurate results than non-gradient-based optimizers [Noce 06]. The duration of the optimization can be substantially faster if the gradient or Hessian can be provided analytically and if their computation can be implemented efficiently. The gradient of the cost function in Eq. (6.7) is defined by

$$\nabla e(\boldsymbol{\alpha}) = \left( \frac{\partial e(\boldsymbol{\alpha})}{\partial s_1}, \frac{\partial e(\boldsymbol{\alpha})}{\partial t_1}, \dots, \frac{\partial e(\boldsymbol{\alpha})}{\partial s_K}, \frac{\partial e(\boldsymbol{\alpha})}{\partial t_K} \right)^\top, \quad (6.9)$$

where the individual partial derivatives w. r. t. the detector shifts are computed by

$$\begin{aligned} \frac{\partial e(\boldsymbol{\alpha})}{\partial \alpha_l} &= \frac{\partial}{\partial \alpha_l} \frac{1}{2} \left\| \mathbf{o}(\boldsymbol{\alpha}) \right\|_2^2 \\ &= \frac{\partial}{\partial \alpha_l} \frac{1}{2} \left( \mathbf{o}(\boldsymbol{\alpha})^H \mathbf{o}(\boldsymbol{\alpha}) \right) \end{aligned} \quad (6.10)$$

$$= \mathbf{o}(\boldsymbol{\alpha})^H \left( \frac{\partial}{\partial \alpha_l} \mathbf{o}(\boldsymbol{\alpha}) \right) \quad (6.11)$$

$$= \left( \mathbf{p}^H \mathbf{F}_{uv}^H \mathbf{H}(\boldsymbol{\alpha})^H \mathbf{F}_\lambda^H \mathbf{W}^H \right) \mathbf{W} \mathbf{F}_\lambda \left( \frac{\partial}{\partial \alpha_l} \mathbf{H}(\boldsymbol{\alpha}) \right) \mathbf{F}_{uv} \mathbf{p} \quad (6.12)$$

$$= \mathbf{p}^H \mathbf{F}_{uv}^H \mathbf{H}(\boldsymbol{\alpha})^H \mathbf{F}_\lambda^H \mathbf{W} \mathbf{F}_\lambda \left( \frac{\partial}{\partial \alpha_l} \mathbf{H}(\boldsymbol{\alpha}) \right) \mathbf{F}_{uv} \mathbf{p} . \quad (6.13)$$

$(\cdot)^H$  denotes the Hermitian of a vector or matrix, i. e., the transposed, complex conjugate of the matrix. Further, we define  $\alpha_l$  to be the  $l$ -th element of  $\boldsymbol{\alpha}$ , which can correspond to either  $s_k$  or  $t_k$ . That means it is a translation in either  $u$  or  $v$  direction associated with the  $k$ -th projection image. Given the structure of  $\boldsymbol{\alpha}$  introduced in Eq. (6.6), the indices are connected by  $k = \lceil l/2 \rceil$ , such that

$$\alpha_l = \begin{cases} s_k & \text{if } \text{mod}(l, 2) \neq 0 \\ t_k, & \text{otherwise} \end{cases} . \quad (6.14)$$

In Eq. (6.10) the squared L2 norm for complex numbers is decomposed into a dot product, which can be easily derived as shown in Eq. (6.11). In Eq. (6.12)  $\mathbf{o}(\boldsymbol{\alpha})$  was replaced by its components according to Eq. (6.7). It can be shown that  $\mathbf{W}^H = \mathbf{W}$  and  $\mathbf{W}^H \mathbf{W} = \mathbf{W}$ , due to the diagonal and also binary structure of  $\mathbf{W}$ . Thus, we obtain Eq. (6.13) where  $\mathbf{W}$  appears only once.

In Eq. (6.15) we introduce  $\boldsymbol{\psi} \in \mathbb{C}^C$ , containing the partial derivatives of the arguments given in the exponential functions of  $\mathbf{H}(\boldsymbol{\alpha})$  (cf. Eq. (6.8)).

$$\boldsymbol{\psi} = \begin{cases} (0, \dots, 0, -i2\pi\xi_1, \dots, -i2\pi\xi_I, 0, \dots, 0)^\top & \text{if } \alpha_l = s_k \\ (0, \dots, 0, -i2\pi\psi_1, \dots, -i2\pi\psi_J, 0, \dots, 0)^\top & \text{if } \alpha_l = t_k \end{cases} \quad (6.15)$$

To conclude gradient computation, we obtain the partial derivatives of the motion matrix  $\mathbf{H}(\boldsymbol{\alpha})$  by

$$\frac{\partial}{\partial \alpha_l} \mathbf{H}(\boldsymbol{\alpha}) = \text{diag}(\boldsymbol{\psi}) \mathbf{H}(\boldsymbol{\alpha}) . \quad (6.16)$$

It is worth noting, that  $s_k$  or  $t_k$  causes only variations to the 2-D DFT of the  $k$ -th projection image. As a consequence,  $\boldsymbol{\psi}$  is sparse and has only  $I \times J$  nonzero entries in the range of  $[(k-1)IJ+1, (k)IJ]$ . It is well known that the multiplication by  $\text{diag}(\boldsymbol{\psi})$  in Fourier domain, is associated to a spatial derivative of the  $k$ -th projection image in either  $u$  or  $v$  direction.

### 6.3.2 Invariance to Static Detector Offsets

In the following we show that the FCC are invariant to temporally constant shifts, e. g., a fixed detector offset. Let us define  $\varphi_{i,j,n} \in \mathbb{C}$  to be a coefficient of the 2-D DFT

of the  $n$ -th projection image. Also,  $c_1 \in \mathbb{R}$  and  $c_2 \in \mathbb{R}$  are two constants representing a static detector offset in  $u$  and  $v$  direction, respectively. Further,  $o_{i,j,k}(\boldsymbol{\alpha}) : \mathbb{R}^L \rightarrow \mathbb{C}$  is a function that computes the coefficient of the 3-D DFT, located at frequency bin  $(i, j, k)$ . It may be computed by

$$o_{i,j,k}(\boldsymbol{\alpha}) = w_{i,j,k} \sum_{n=1}^K \varphi_{i,j,n} e^{-i2\pi(\xi_i s_n + \psi_j t_n)} e^{-i2\pi kn/K} . \quad (6.17)$$

Let us now set all shift values to the introduced constants, i. e.,  $s_n = c_1$  and  $t_n = c_2$  for all  $n \in [1, \dots, K]$ , such that  $o_{i,j,k}$  is independent of  $\boldsymbol{\alpha}$ . Then the phase factor may be factorized out of the DFT's summation by

$$o_{i,j,k} = w_{i,j,k} e^{-i2\pi(\xi_i c_1 + \psi_j c_2)} \sum_{n=1}^K \varphi_{i,j,n} e^{-i2\pi kn/K} \quad (6.18)$$

$$o_{i,j,k} = w_{i,j,k} e^{-i2\pi(\xi_i c_1 + \psi_j c_2)} \Phi_{i,j,k} , \quad (6.19)$$

where  $\Phi_{i,j,k} \in \mathbb{C}$  is a coefficient of the 3-D DFT of projection data  $\mathbf{p}$ , at frequency bin  $(i, j, k)$ . FCC considers only the magnitude of the Fourier coefficients, which we realized by the complex form of the squared L2 norm. Eq. (6.7) may be reformulated by

$$e(\boldsymbol{\alpha}) = \frac{1}{2} \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K w_{i,j,k} o_{i,j,k} \overline{o_{i,j,k}} . \quad (6.20)$$

Applying Eq. (6.19) to Eq. (6.20) yields

$$e(\boldsymbol{\alpha}) = \frac{1}{2} \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K w_{i,j,k} \Phi_{i,j,k} \overline{\Phi_{i,j,k}} \underbrace{e^{i2\pi(\xi_i c_1 + \psi_j c_2)} e^{-i2\pi(\xi_i c_1 + \psi_j c_2)}}_{=1} , \quad (6.21)$$

which shows that the cost function is indeed invariant to any constant translations  $c_1$  or  $c_2$ .

## 6.4 Regularization, Optimization and Motion Compensation

The implications of the cost function's invariance to static detector shifts are twofold. Given an alignment of detector  $v$ -axis with the rotation axis, a constant offset in  $v$  direction is similar to an offset of the object in  $z$  direction. However, a detector offset orthogonal to the CBCT's rotation axis, e. g., in  $u$  direction, causes distortions in the reconstructed image. A possible solution to this limitation is to enforce fixed values for the shifts in a selected projection image. In our method we fixed the motion parameters of the first projection image to  $s_1 = \tilde{s}$  and  $t_1 = \tilde{t}$ .

The fixed parameters can either be incorporated directly, reducing the parameter space by two elements. Alternatively,  $\tilde{s}$  and  $\tilde{t}$  may be incorporated as regularizer. In this case, the final cost function is obtained by

$$\tilde{e}(\boldsymbol{\alpha}) = e(\boldsymbol{\alpha}) + \tilde{\eta} \frac{1}{2} \left\| \begin{pmatrix} s_1 - \tilde{s} \\ t_1 - \tilde{t} \end{pmatrix} \right\|_2^2 , \quad (6.22)$$

where we enforce that  $s_1$  and  $t_1$  are close to the provided constants using the squared L2 norm of their difference, weighted with the Lagrangian multiplier  $\tilde{\eta}$ . The gradient w. r. t. the motion parameters may now be obtained by

$$\nabla \tilde{e}(\boldsymbol{\alpha}) = \nabla e(\boldsymbol{\alpha}) + \tilde{\eta} \begin{pmatrix} (s_1 - \tilde{s}) & (t_1 - \tilde{t}) & 0 & \cdots & 0 \end{pmatrix}^\top . \quad (6.23)$$

The same gradient-based nonlinear optimizer is used as presented for the marker-based and the registration-based methods in Chapters 4 and 5. The optimizer is unconstrained, thus  $\tilde{\eta}$  needs to be defined before optimization to allow optimization of Eq. (6.22). The gradient of the cost function w. r. t. the motion parameters is computed analytically following Eq. (6.23), whereas the Hessian is approximated during optimization using the BFGS method based on the computed gradient. An attempted Newton step solves for the step direction given the approximated Hessian. The step size is computed using a line-search approach. Additional information on the optimizer may be obtained from [Maie10].

After optimization, the parameters in  $\boldsymbol{\alpha}$  are directly incorporated into the projection matrices to update the mapping of 3-D to 2-D coordinates using Eq. (3.44). This can be achieved using Equations (3.35) and (3.36). However, the formulations in this chapter lead to estimated translations in mm. They, need to be converted to pixel coordinates before integration into the projection matrices  $\mathbf{P}_k$ . The updated projection matrices may be obtained by

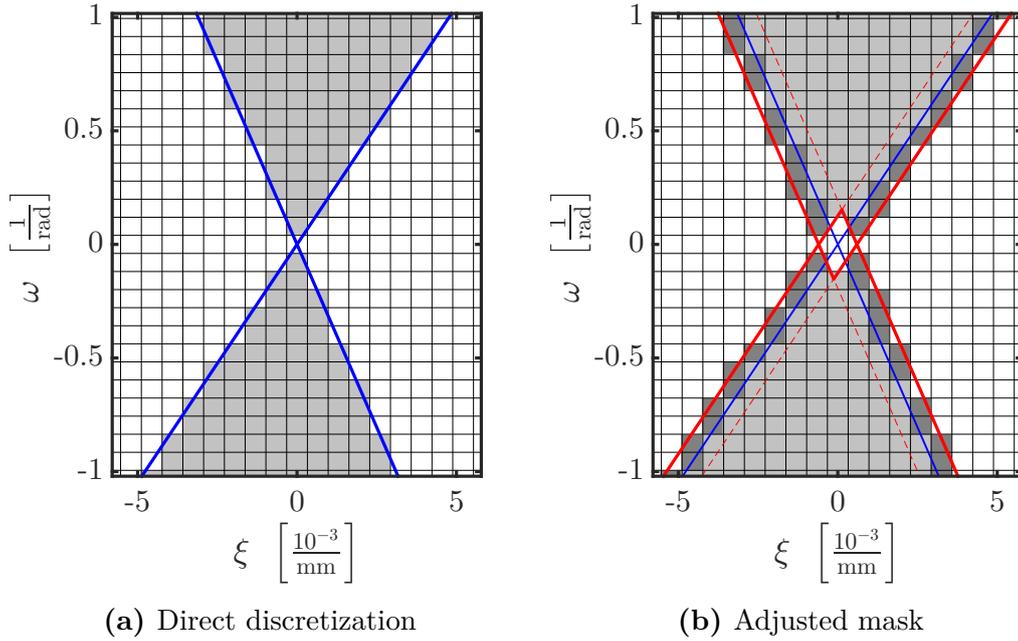
$$\mathbf{P}_k(\boldsymbol{\alpha}) = \begin{pmatrix} 1 & 0 & -\frac{s_k}{\Delta u} \\ 0 & 1 & -\frac{t_k}{\Delta v} \\ 0 & 0 & 1 \end{pmatrix} \mathbf{P}_k . \quad (6.24)$$

Note that Eq. (6.24) essentially updates the principle point of  $\mathbf{P}_k$ . If the detector is shifted by  $s_k$  and  $t_k$ , then the principle point moves in the opposite direction which causes the negative signs in the equations. Standard FDK can be applied for image reconstruction as described in Sec. 3.2.3, however, using the updated, instead of the original projection matrices.

## 6.5 Practical Considerations

### 6.5.1 Discretization and Adjustment of the Mask

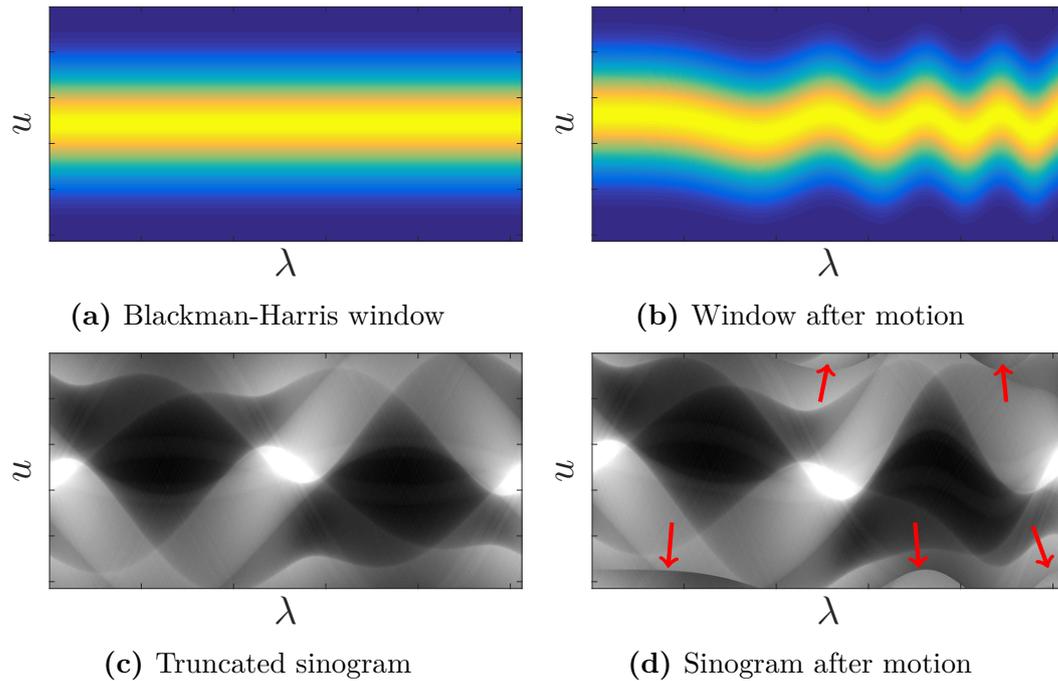
Using Eq. (6.5) to implement the mask that encodes the triangular regions can be interpreted as a discretization of Eq. (6.1). Essentially, Eq. (6.1) can be interpreted as a description of two continuous straight lines that build the triangular regions. Discretization now requires a representation of these lines using the discretized grid in Fourier domain which makes correct identification of the mask's boundaries challenging. The process of discretization is illustrated in Fig. 6.2(a). It shows the sampling grid, determined by Eq. (6.3), but also the discretized mask and the continuous representation of the mask given by Eq. (6.1). It can be seen that in some cases the discretized mask extends over the continuous boundary whereas other regions have almost no connection to the continuous lines.



**Figure 6.2:** Unmodified discretization in (a) and discretization after heuristic adjustment of regions to cover all motion related energies in (b).

This may cause several problems in a practical application. In case the discretized mask is too large, energies that are shifted to the triangular regions due to motion are not allowed to redistribute close to the support region. We assume implicitly that these energies are zero which may not be the case, even in an acquisition without motion. In contrast, if the discretized mask is too small, energy can be distributed to frequency bins close to the ideal boundary, but not fully covered by the mask. In consequence, these frequency bins will have only little effect on the cost function and are also not important for the consistency measure. Based on heuristic experiments, a mask that is too small appeared less robust for motion estimation as it often lead to an interference of the estimated motion with a sinusoidal signal of low-frequency but with high amplitude.

To account for these discretization effects a heuristic adjustment of Eq. (6.5) is proposed, which extends the mask regions similarly in all directions. The adjusted mask and its discretization can be seen in Fig. 6.2(b), where the discretization result is visualized as a solid gray area. First, the original continuous boundaries (dashed, red lines) are shifted outwards in a direction orthogonal to themselves, for each triangle boundary. Thus, four new lines (solid, red lines) are created which form two overlapping triangles, that contain the adjusted mask regions. However, the frequencies covered by both of the triangles are excluded from the mask to ensure that the center and very low frequencies are not contained. An advantage of this approach is that it only requires a single parameter  $\epsilon$ , which defines the distance of ideal boundaries to the shifted, parallel lines.

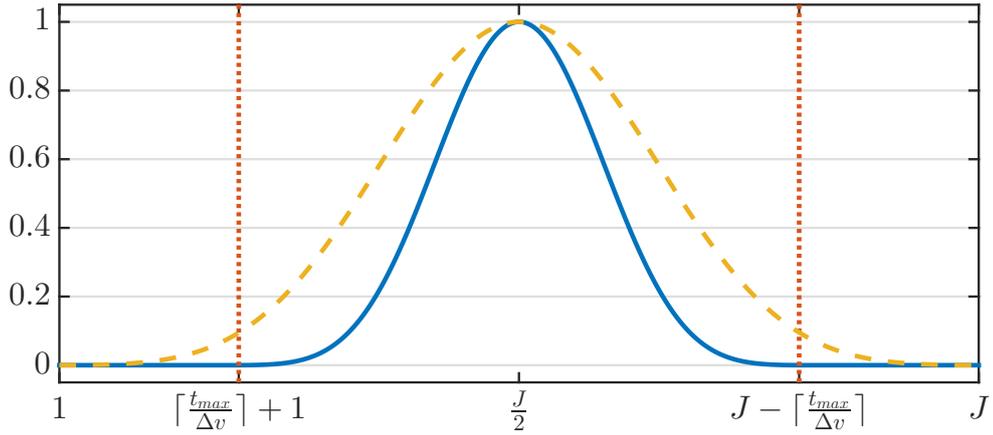


**Figure 6.3:** (a) Example of a static Blackman-Harris apodization window for the detector axis in fan-beam geometry. (b) Shape of window after application of artificial translational motion. (c) Example for a truncated sinogram. (d) Reappearance of sinogram content at opposite boundary after applying shifts in Fourier domain. For both cases, i. e., for (b) and (d), the erroneously introduced intensity variations over  $\lambda$  have a negative effect on FCC.

### 6.5.2 Axial Truncation: Problems and Solutions

Truncation refers to the problem of missing data, located outside the FOV covered by the detector. In case of the presented approach based on FCC, the missing data becomes apparent after application of detector translations. In Sec. 6.3 we describe how detector translations may be implemented in the Fourier domain. The Fourier transform assumes a periodic extension of the signal. Thus, a translation in  $u$  or  $v$  direction, using phase factors, causes content at the boundary of the detector to reappear at the opposite side. This process is visualized in Figures 6.3(c) and 6.3(d), where we show an example of a truncated fan-beam sinogram and its shifted version using an artificial motion pattern. Even if we would apply the translation in spatial domain, we do not know which values originate from the periphery. In case no truncation is present, values at the boundary will be close to zero as they correspond to the absorption of air. If the “empty” boundary in projection images is larger than the maximum translation, no reappearance of the object at image boundaries will occur. Even, if the object fits the detector exactly we could use zero-padding in  $u$  or  $v$  direction, prior to the 2-D Fourier transform, to accurately solve the truncation problem.

In addition to the missing information, the derivations of FCC are based on a limited object extent in axial, lateral, and also longitudinal direction [Mazi 10, Brok 06]. Within this section we focus on axial truncation problems as it would appear in

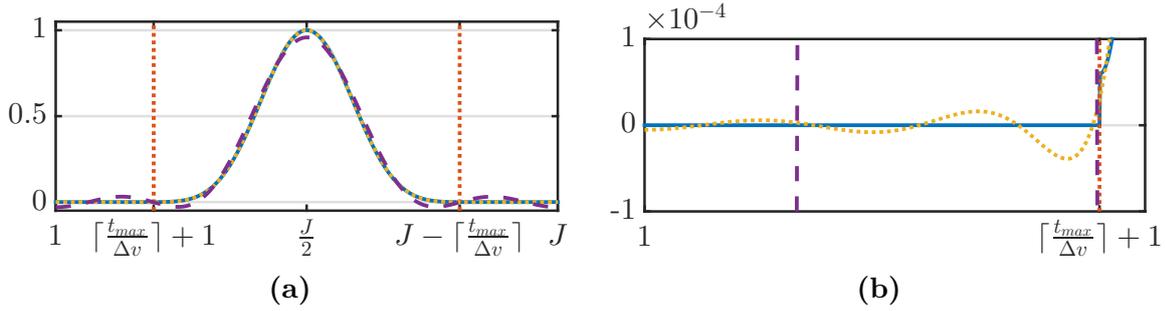


**Figure 6.4:** Adjusted apodization window to allow for translations of length  $t_k \in [-t_{max}, t_{max}]$ . Yellow, dashed line: Original Blackman-Harris window extending from  $j \in [1, J]$ . Blue, solid line: Adjusted Blackman-Harris window extending from  $j \in [\lceil \frac{t_{max}}{\Delta v} \rceil + 1, J - \lceil \frac{t_{max}}{\Delta v} \rceil]$ .

knee imaging, assuming the detector is wide enough to cover both legs. To ensure that the object is of limited extent, we propose to multiply each projection image with an apodization function that smoothly fades out to zero at the vertical detector boundaries. Apodization emulates a maximum object size which corresponds to the detector height. It is only used for motion estimation and neglected during reconstruction. Yet, apodization needs to be applied dynamically, after the application of estimated translations, to avoid additional influence of intensity variations due to the apodization window. Initial experiments using a static apodization applied to axially truncated data prior to optimization, causes the estimated translations  $t_k$  to be limited to  $[-\frac{\Delta v}{2}, \frac{\Delta v}{2}]$ , i. e., within the size of a single pixel. The reason for this is twofold. As described above, any further shift in  $v$  direction causes the data to reappear the opposite boundary. Because translations can be different for each projection, this reappearance causes irregular intensity variations in angular direction, increasing the energy in the triangular regions. Similar artificial intensity variations are obtained from a static apodization after application of different translations. For a better understanding we visualize a static apodization window and the effect of motion in Figures 6.3(a) and 6.3(b), using a fan-beam example with detector truncation.

### Proposed Solution

We propose to use an adjusted Blackman-Harris window for apodization as shown in Fig. 6.4. The window is multiplied with each column of all projection images, after the translations have been applied. Instead of using a regular Blackman-Harris window with a full extent over detector  $v$  direction, we limit its range and clamp it to zero. The start and end point of the window are determined as shown in Fig. 6.4, where  $t_{max}$  is the zero margin added at the boundaries. In consequence, any vertical translation  $t_k \in [-t_{max}, t_{max}]$  will not lead to a periodic reappearance at bottom or top of the detector.



**Figure 6.5:** Approximation of adjusted Blackman-Harris window by retaining spectral coefficients. Solid, blue line: Theoretically computed window function. Dashed, purple line: Approximation retaining only four coefficients. Dotted, orange line: Accurate approximation retaining eight coefficients.

Integration of apodization into the cost functions in Equations (6.2) and (6.7) is straightforward. Let  $\mathbf{A} = \text{diag}(\mathbf{a}) \in \mathbb{R}^{C \times C}$  be a diagonal matrix that holds the weights of the apodization window defined by vector  $\mathbf{a} \in \mathbb{R}^C$ . Following Harris [Harr 78], we can formalize this vector by

$$a_{i,j,k} = \begin{cases} 0 & \text{if } j \notin \left[ \lceil \frac{t_{max}}{\Delta v} \rceil + 1, J - \lceil \frac{t_{max}}{\Delta v} \rceil \right] \\ c_0 - c_1 \cos\left(\frac{2\pi j'}{J'}\right) + c_2 \cos\left(\frac{4\pi j'}{J'}\right) - c_3 \cos\left(\frac{6\pi j'}{J'}\right) & \text{otherwise} \end{cases} \quad (6.25)$$

$$c_0 = 0.35875, \quad c_1 = 0.48829, \quad c_2 = 0.14128, \quad c_3 = 0.01168,$$

where  $j' = j - \lceil \frac{t_{max}}{\Delta v} \rceil - 1$  and  $J' = J - 2\lceil \frac{t_{max}}{\Delta v} \rceil$  represent the adjusted index and window width.

The cost function for arbitrary motion in Eq. (6.2) can now be adjusted to

$$e(\boldsymbol{\alpha}) = \frac{1}{2} \left\| \mathbf{W} (\mathbf{F} \mathbf{A} \mathbf{p}(\boldsymbol{\alpha})) \right\|_2^2. \quad (6.26)$$

However, incorporating the apodization into the efficient description of the cost function given by Eq. (6.7), requires an additional inverse (i. e.,  $\mathbf{F}_{uv}^H$ ) and forward (i. e.,  $\mathbf{F}_{uv}$ ) 2-D FFT. The equation is then given by

$$e(\boldsymbol{\alpha}) = \frac{1}{2} \left\| \mathbf{W} \left( \mathbf{F}_\lambda \underbrace{(\mathbf{F}_{uv} \mathbf{A} \mathbf{F}_{uv}^H)}_{\text{apodization}} (\mathbf{H}(\boldsymbol{\alpha}) \mathbf{F}_{uv} \mathbf{p}) \right) \right\|_2^2. \quad (6.27)$$

As a consequence the computational advantage of the efficient description vanishes. Fortunately, it is known that a multiplication by apodization windows in spatial domain equals a convolution with the Fourier transformed window in spectral domain. In fact, applying apodization in the spectral domain is particularly easy when using generalized cosine windows, like the Blackman-Harris window, as they have only a small number of unique, nonzero Fourier coefficients [Harr 78]. For example, for  $t_{max} = 0$  it holds that the window described in Eq. (6.25) has only four unique spectral coefficients, requiring seven multiplications and six summations at each Fourier coefficient to perform apodization.

Yet, for  $t_{max} > 0$  this is no longer true and in theory the convolution would be global, requiring  $J$  multiplications and  $J - 1$  summations for each Fourier coefficient. However, in practice using only a certain number of most significant coefficients, reduces complexity and still leads to good approximations of the window. In Fig. 6.5 we show two approximations when retaining four (dashed, purple line) and eight (dotted, orange line) coefficients alongside the ground-truth window function (solid, blue line). Only very little difference can be seen between real and approximated function when using eight spectral coefficients.

Note that the computation of the cost function's gradient described in Eq. (6.13) is still valid, but now includes  $(\mathbf{F}_{uv} \mathbf{A} \mathbf{F}_{uv}^H)$  as introduced in Eq. (6.27). Thus, obtaining partial derivatives now also requires described convolution in spectral domain.

### L2 regularizer

A side effect of dynamic apodization is that the total energy of the spectrum is no longer constant. Thus, the energy in the triangular regions may also be minimized by translations that shift high intensity objects, e. g., bones, upwards or downwards such that they observe a lower weight during apodization. To prevent this behavior, we propose to regularize the amplitude of the estimated motion, using its L2 norm. The updated cost function, including the previously defined regularizer of Eq. (6.22), may be formalized by

$$e'(\boldsymbol{\alpha}) = \tilde{e}(\boldsymbol{\alpha}) + \eta' \left( \frac{1}{2L} \boldsymbol{\alpha}^\top \boldsymbol{\alpha} \right) . \quad (6.28)$$

Consequently, the overall gradient of the cost function may be computed by

$$\nabla e'(\boldsymbol{\alpha}) = \nabla \tilde{e}(\boldsymbol{\alpha}) + \eta' \left( \frac{1}{L} \boldsymbol{\alpha} \right) . \quad (6.29)$$

Similar to the derivation in Sec. 6.4, we have introduced a Lagrangian multiplier  $\eta'$  to create a single combined cost function used for optimization.

### 6.5.3 Implementation Details

In this section several optimization steps for an improved, computationally efficient implementation are proposed. The overall runtime of an optimization is important for its practical relevance. It often depends on the runtime of an individual evaluation of the cost function and its gradient. Based on an initial, single-threaded CPU implementation a calculation time of several minutes for one evaluation of the cost function was measured. The reason was repeated access to the whole projection data  $\mathbf{p}$  and its 3-D Fourier transform. As a first optimization step, we incorporated the detector shifts by phase multiplications according to Eq. (6.7). This allows a precomputation of  $\mathbf{F}_{uv} \mathbf{p}$ , such that it needs to be computed only once during optimization. The 1-D FFT over the rotation angles is executed  $I \times J$  times for each cost function evaluation. An additional speed up of this FFT was achieved by a rearrangement of the data layout, such that the projection angles are stored sequentially in memory, facilitating faster memory access by caching. In addition, most of the operations needed to evaluate the cost function are highly parallelizable such that we implemented them on the GPU using the OpenCL programming language.

---

**Algorithm 6.1:** Outline of a cost function evaluation for the  $n_{it}$ -th iteration.

---

```

input :  $\alpha^{n_{it}}, \alpha^{n_{it}-1}, (\mathbf{H}(\alpha^{n_{it}-1})\mathbf{F}_{uv}\mathbf{p})$ 
output:  $e(\alpha^{n_{it}}), (\mathbf{H}(\alpha^{n_{it}})\mathbf{F}_{uv}\mathbf{p})$ 

// Compute relative parameter vector
 $\Delta\alpha^{n_{it}} \leftarrow \alpha^{n_{it}} - \alpha^{n_{it}-1};$ 
// Apply shifts by multiplication with the phase factors
 $\mathbf{H}(\Delta\alpha^{n_{it}})$ . Two vectors holding all  $\psi$  and  $\xi$  values are
precomputed. Only  $\Delta\alpha^{n_{it}}$  needs to be transferred to the GPU.
 $(\mathbf{H}(\alpha^{n_{it}})\mathbf{F}_{uv}\mathbf{p}) \leftarrow \mathbf{H}(\Delta\alpha^{n_{it}}) \cdot (\mathbf{H}(\alpha^{n_{it}-1})\mathbf{F}_{uv}\mathbf{p});$ 
// Angular 1-D FFT. A GPU-based library was used for the FFT. The
implementation does not support in-place transforms, which
required an additional storage of the 3-D FFT result.
 $\mathbf{F}_\lambda \cdot (\mathbf{H}(\alpha^{n_{it}})\mathbf{F}_{uv}\mathbf{p});$ 
// L2 norm and mask: Sum of squared magnitudes, contained in the
mask, by a GPU-based parallel reduction. The mask is
precomputed before optimization, thus, only the summation
result needs to be transferred back to the host.
 $e(\alpha^{n_{it}}) \leftarrow \frac{1}{2} (\mathbf{o}(\alpha^{n_{it}}))^H \mathbf{W} \mathbf{o}(\alpha^{n_{it}});$ 

```

---

### Evaluation of the Cost Function

A typical data size for a complex representation of  $\mathbf{p}$  in our studies could be  $1240 \times 960 \times 248 \times 2 \times 4\text{Byte}/(1024)^2 \approx 2250\text{MB}$ . Thus, storing multiple copies of this data on the GPU's memory might not be feasible as they are still limited to only few GB's of data. In consequence, an implementation needed to be found that requires as few as possible copies of the projection data. One solution is to maintain only the shifted projection data in GPU memory. Let  $n_{it}$  be the iteration count and  $\alpha^{n_{it}}$  the associated parameters. As a first step in each iteration the differential parameter vector is built given by  $\Delta\alpha^{n_{it}} = \alpha^{n_{it}} - \alpha^{n_{it}-1}$ , with  $\alpha^{n_{it}-1}$  being the parameter vector from the previous iteration. The trick is now to apply the phase factors of the relative shifts  $\Delta\alpha^{n_{it}}$ , to the Fourier transform of the already translated projection data, instead of applying  $\alpha^{n_{it}}$  to the original data. Hence,  $\mathbf{p}$  is not required to evaluate the cost function nor its gradient. The process of a single evaluation of the cost function can be summarized by Algorithm 6.1.

### Gradient Computation

Equations (6.9) and (6.11) show that the gradient computation requires evaluation of the partial derivatives  $\mathbf{o}(\alpha) \frac{\partial}{\partial \alpha_i} \mathbf{o}(\alpha)$  w. r. t. each of  $L$  parameters. Certain conditions need to be fulfilled to make an optimization algorithm based on an analytic gradient faster than a numerical, forward differences gradient estimation. First, the average computation time of a single partial derivative needs to be lower than the average computation time of a cost function evaluation. Second, a potentially faster convergence of the analytic gradient approach could cause a faster optimization even if the first condition is not met. In this method, a fast gradient computation is achieved by

the following principles.  $\mathbf{o}(\boldsymbol{\alpha})$  is already computed during cost function evaluation and can be reused as is. Further,  $\frac{\partial}{\partial \alpha_l} \mathbf{o}(\boldsymbol{\alpha})$  can be decomposed into a multiplication of the precomputed 2-D DFT  $\mathbf{F}_{uv} \mathbf{p}$ , with the derivative of the shifting matrix  $\frac{\partial}{\partial \alpha_l} \mathbf{H}(\boldsymbol{\alpha})$  and finally by the angular DFT  $\mathbf{F}_\lambda$ . Hence, a straightforward implementation needs to compute  $\mathbf{F}_\lambda$  by 1-D FFTs for each partial derivative. However, using well known properties of the Fourier transform we can avoid computation of the angular FFTs completely.

First, we use the fact that  $\frac{\partial}{\partial \alpha_l} \mathbf{H}(\boldsymbol{\alpha})$  is only nonzero for a single projection image with index  $m = \lceil l/2 \rceil$ . Consequently,  $\frac{\partial}{\partial \alpha_l} \mathbf{H}(\boldsymbol{\alpha}) \mathbf{F}_{uv} \mathbf{p}$  has the same property. If we adopt the notation from Sec. 6.4 the angular DFT may be computed by

$$\frac{\partial}{\partial \alpha_l} o_{i,j,k}(\boldsymbol{\alpha}) = w_{i,j,k} \sum_{n=1}^K \varphi_{i,j,n} \underbrace{\left( \frac{\partial}{\partial \alpha_l} e^{-i2\pi(\xi_i s_n + \psi_j t_n)} \right)}_{=0 \quad \forall n \neq m} e^{-i2\pi kn/K} , \quad (6.30)$$

where the derivative of the phase factors is only nonzero for  $n = m$ . Thus, the summation of the DFT degenerates to

$$\frac{\partial}{\partial \alpha_l} o_{i,j,k}(\boldsymbol{\alpha}) = w_{i,j,k} \varphi_{i,j,l} \left( \frac{\partial}{\partial \alpha_l} e^{-i2\pi(\xi_i s_m + \psi_j t_m)} \right) e^{-i2\pi km/K} . \quad (6.31)$$

As a result, only a simple multiplication by  $e^{-i2\pi km/K}$  is required to compute the angular Fourier transform, substantially reducing the computational complexity. As a last step, it is ensured that partial derivatives  $\frac{\partial}{\partial s_m} e(\boldsymbol{\alpha})$  and  $\frac{\partial}{\partial t_m} e(\boldsymbol{\alpha})$  are computed in a single execution on the GPU. Thus, the overall runtime can be almost halved, as the data transfer causes the majority of computation time.



PART III

# Experiments and Results



# Data Acquisition and Common Metrics

7.1 Simulated Phantom Data .....	87
7.2 Real CBCCT Acquisitions .....	93
7.3 Common Image Quality Metrics .....	96

The evaluation part of this thesis is based on a variety of real as well as simulated CT datasets. This chapter contains information on the individual acquisition and simulation parameters, provided in Sec. 7.1 and Sec. 7.2. Aside from the acquisition protocols in case of real data, we provide the geometric parameters used for data generation, such as image resolution or SDD. Essential parameters will be summarized in a dedicated table for each dataset. Note that sizes and spacing of the reconstruction volumes are not part of the acquired data as they can be freely adjusted during back-projection. Thus, reconstruction related parameters and settings will be introduced in the corresponding sections of the evaluation chapters. An overview of all acquired datasets is provided in Tab. 7.1, which also introduces abbreviations used to refer to an individual dataset.

The chapter ends with an introduction of common image metrics, used in several of the following experiments to obtain quantitative measures of image quality. In detail, Sec. 7.3 shows descriptions of the mean squared error (MSE), root mean squared error (RMSE), relative root mean squared error (rRMSE), universal quality index (UQI), and structural similarity index (SSIM).

## 7.1 Simulated Phantom Data

Numerically simulated data has been used during a comparison of different extrapolation techniques used for marker removal and for the evaluation of the FCC-based motion correction method. Additionally, realistic simulations of the knee anatomy are generated for a qualitative assessment of the FCC method, using the dynamic XCAT phantom [Sega 10]. In the following we introduce geometric and physical parameters of the individual datasets.

Note, that all numerically simulated phantoms consist of analytically described shapes, e. g., spheres, ellipsoids, or spline surfaces. Simulation consists of a simulated X-ray transform, as described in Eq. (3.13), where line integrals are built by computing intersection lengths with individual geometric shapes that represent the phantom.

## Numerical phantoms and simulations

Phantom	Description	Dataset	Abbreviation
Bead removal phantom	Metallic markers attached to artificial knee model	Ground truth	MR_GT
		With eight markers	MR_8M
FORBILD head phantom	Static and dynamic fan-beam FORBILD head phantom	Fan-beam ground truth	FFB_GT_HQ
		Fan-beam <i>Oscil</i> motion	FFB.OSC_HQ
FORBILD head phantom	Static and dynamic cone-beam FORBILD head phantom	Cone-beam ground truth	FCB_GT_HQ
		Cone-beam <i>Oscil</i> motion	FCB.OSC_HQ
		Cone-beam <i>Chirp</i> motion	FCB.CHP_HQ
		Cone-beam <i>Rect</i> motion	FCB.RCT_HQ
		Cone-beam <i>LF-1</i> motion	FCB.LF1_HQ
		Cone-beam <i>LF-2</i> motion	FCB.LF2_HQ
XCAT knee phantom	Realistic numeric phantom of knees, including a real motion pattern	Ground truth	XCAT_GT
		Severe motion	XCAT_MOT

## Real acquisitions

Subject	Description	Dataset	Abbreviation
Subject 1	Left leg for marker removal evaluation	Supine scan with eight markers	S1_SUP
Subject 2	Supine and three weight-bearing scans of both legs at different flexion angles	Supine scan	S2_SUP
		Weight-bearing, 0° flexion	S2.WB0
		Weight-bearing, 35° flexion	S2.WB35
		Weight-bearing, 60° flexion	S2.WB60
Subject 3	Supine and six weight-bearing scans of left leg with injected contrast agent (cartilage study)	Supine scan	S3_SUP
		Weight-bearing, 0 s	S3.WB1
		Weight-bearing, 10 s	S3.WB2
		Weight-bearing, 20 s	S3.WB3
		Weight-bearing, 3 min	S3.WB4
		Weight-bearing, 10 min	S3.WB5
		Weight-bearing, 25 min	S3.WB6

**Table 7.1:** Overview of acquired data and abbreviations used for individual datasets. HQ for FORBILD head phantom refers to high-quality, noise-free simulations and may be replaced by LQ for noisy data, which are not shown here.

Parameter	Symbol	Unit	Value
SID	$S$	mm	600
DID	$D$	mm	600
Detector size	$I \times J$	-	$1240 \times 960$
Pixel size	$\Delta u \times \Delta v$	$\text{mm}^2$	$0.308 \times 0.308$
#Projections	$K$	-	494
Angular spacing	$\Delta\lambda$	deg	0.4049

**Table 7.2:** Simulation parameters for bead removal phantom.

Each shape is assigned to a physical material, e. g., water or bone. Computed intersection lengths are weighted with the materials attenuation coefficient, according to a selected physical model (cf. Sec. 3.2.1). In case geometric shapes overlap, the simulation uses a priority scheme to ensure that only a single material is evaluated at a specific position along the X-ray path. Physical properties of the individual materials are obtained from the well known [National Institute of Standards and Technology \(NIST\)](#) database [[Hubb 95](#)]. All simulations are done with the software framework [CONE-beam in RADIology \(CONRAD\)](#) [[Maie 13](#)]. Further details on the numerical simulation of projection images may be obtained from [[Maie 12a](#)]. If not stated otherwise, the simulated X-ray trajectory describes an ideal circle with the rotation axis being identical to the  $z$ -axis, thus, all source positions lie in the plane spanned by the  $x$ - and  $y$ -axis. Furthermore, the principle ray intersects the detector in its center. This setting is identical to the description of [CBCT](#) in Sec. 3.2.3.

### 7.1.1 Bead Removal Phantom

A phantom was developed which consists of three nested cylinder shapes, representing a simplified model of the knee. The cylinders represent leg tissue as well as cortical bone and bone marrow of a femur. Their materials are set to water, cortical bone, and bone marrow, respectively. Their radii are adjusted to 80 mm, 35 mm and 31.5 mm and their axis of rotation is aligned with the  $z$ -axis. Additionally, a total of eight metallic beads (stainless steel type 304) with a diameter of 1 mm are attached to the periphery of the largest cylinder. They are distributed on a helix and overlap with surface of the cylinder by 0.1 mm. For simplicity, we avoided a physical model based on the energy of X-ray photons and used the materials' density as linear attenuation coefficient. This equals a mass attenuation coefficient of  $1 \text{ cm}^2 \text{ g}^{-1}$  for all materials.

Two sets of projection images are generated, one without the metallic markers, representing the ground-truth data, and one including the markers, which is then used for extrapolation. All necessary parameters can be obtained from [Tab. 7.2](#).

### 7.1.2 FORBILD Head Phantom

We use the FORBILD head phantom [[Laur 01](#)] to evaluate the [FCC](#)-based motion correction. The phantom represents a head and consists of bone, several low-contrast

structures, a resolution pattern, and a challenging ear part with fast transitions between bone and air. A total of nine physical materials are assigned to a large number of geometrical shapes, eventually building up the phantom.

### Definition of Translational Motion

To use the phantom for an evaluation of motion correction capabilities, simulated motion patterns need to be incorporated into the rendering process of projection images. In the following, we introduce all simulated motion patterns of the object. The motion was limited to 3-D object translations, where three high-frequency and two low-frequency motion functions were simulated. Selection of these functions has been done w. r. t. their visual effect on the triangular zero regions, used for optimization in the FCC method. This way we could provoke different amounts of energy shifting into the zero regions.

As the algorithm for rendering projection images is based on projection matrices, the motion patterns can be easily incorporated using Equations (3.34) and (3.36), where  $\alpha \in \mathbb{R}^{6K}$  contains all simulated translations and the rotation parameters are set to zero. Please, see appendix A.1 for further information on the structuring of  $\alpha$  and its relation to rigid motion parameters. The updated projection matrices used for rendering of projection images are given by

$$\mathbf{P}_k(\alpha) = \mathbf{P}_k \begin{pmatrix} 1 & 0 & 0 & t_k^x \\ 0 & 1 & 0 & t_k^y \\ 0 & 0 & 1 & t_k^z \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (7.1)$$

We now introduce the individual motion functions. Note, that Equations (7.2) to (7.6) are presented with a normalized temporal parameter given by  $\tau_k = \frac{k-1}{K-1} \in [0, 1]$ .

**Oscillating Motion (*Oscil*)** The first motion is an accelerated, oscillating translation along the  $x$ -,  $y$ - and  $z$ -axis. It has a zero mean and consists of a cosine oscillation with only little low-frequency contributions, due to a limited acceleration setting. The translation function is computed by

$$t_k^x = t_k^y = t_k^z = m_{\text{oscil}}(a, b, f) = a \left( \frac{2}{1 + \exp(b \cos(2\pi f \tau_k))} - 1 \right), \quad (7.2)$$

where the amplitude, acceleration and frequency are given by  $a$ ,  $b$  and  $f$ , respectively. The parameters were set to  $a = 3$  mm,  $b = 4$  and  $f = 16$  Hz.

**Frequency Increases Linearly (*Chirp*)** This motion pattern is also an oscillating cosine motion with zero mean. Yet, the frequency increases linearly over time in the limits of  $[0, f_{\text{max}}]$ . We expect that the *Chirp* motion can be used to identify the behavior of motion correction algorithms for a combined low- and high-frequency motion. It can be defined by

$$t_k^x = t_k^y = t_k^z = a \cos(2\pi f_k \tau_k), \quad (7.3)$$

where the frequency is given by  $f_k = (f_{\text{max}} \tau_k)$  with a frequency pitch of  $f_{\text{max}} = 64$  s<sup>-2</sup> and a moderate amplitude of  $a = 1.5$  mm.

**Rectangular Steps (*Rect*)** Sharp transitions of motion states are modeled by this motion, where we used two interfering rectangular functions with different amplitude and frequency. The rectangular function is created with help of Eq. (7.2), by simply setting a very high value for the acceleration factor, thus

$$t_k^x = t_k^y = t_k^z = m_{\text{oscil}}(a_1, b, f_1) + m_{\text{oscil}}(a_2, b, f_2) , \quad (7.4)$$

where amplitudes and frequencies are defined by  $a_1$  and  $a_2$  as well as  $f_1$  and  $f_2$ , respectively. The parameters were set to  $b = 128$ ,  $a_1 = 1.5$  mm,  $a_2 = 1.0$  mm,  $f_1 = 16$  Hz, and  $f_2 = 4$  Hz.

**Low-Frequency Motion 1 (*LF-1*)** This motion originates from [Yu06] who used the  $t_k^x$  and  $t_k^y$  translations for their simulations when evaluating a motion correction method based on fan-beam CC. It is simply a linear object motion using three different slopes. The motion can be considered low frequent and has a mean not equal to zero. The starting and end position of the object are not identical. In consequence the periodic extension of the projection data in direction of  $\lambda$ , assumed by the Fourier transform, is no longer given. We can define the motion by

$$t_k^x = q_1 \tau_k, \quad t_k^y = q_2 \tau_k, \quad t_k^z = q_3 \tau_k , \quad (7.5)$$

where  $q_1$ ,  $q_2$  and  $q_3$  are the individual slopes. They have been fixed to  $q_1 = 6$  mm s<sup>-1</sup>,  $q_2 = 4$  mm s<sup>-1</sup> and  $q_3 = 3$  mm s<sup>-1</sup>.

**Low-Frequency Motion 2 (*LF-2*)** The second type of low-frequency motion does have the same start and end position and executes a low-frequent forth-and-back transition. Artifacts caused by this motion are limited to motion blur and slowly varying intensity gradients. The motion may be formalized by

$$t_k^x = t_k^y = t_k^z = a \left( \frac{\exp(-\cos(2\pi f \tau_k) + 1) - 1}{\exp(2) - 1} \right) , \quad (7.6)$$

where the frequency was set to  $f = 1$  Hz and the amplitude to  $a = \sqrt{\frac{25}{2}}$  mm.

## Projection Image Generation

Experiments included data for fan- and cone-beam CT. Let us first introduce the simulation parameters for the fan-beam geometry, presented in [Berg14b], which uses only the central axial plane of the phantom. In fact, we used the 3-D description of the phantom and incorporated motion according to Eq. (7.1), however, the detector height is limited to  $J = 1$ . Simulations are conducted for a monochromatic absorption model with a certain X-ray energy  $E_p$ , as presented in Sec. 3.2.1. Generated sinograms include noise-free but also noisy data. Poisson distributed noise was incorporated, specified by an initial number of X-ray photons  $N_0$ , emitted for each simulated ray path. We first introduced the *Oscil* motion for the fan-beam case [Berg14b], where the translation in  $z$  direction was set to  $t_k^z = 0$ . Thus, fan-beam evaluations are based on a total of four generated sinograms, i. e., the motion-free, ground-truth and

Parameter	Symbol	Unit	Ideal	Noisy
SID	$S$	mm	600	600
DID	$D$	mm	0	0
Detector size	$I$	-	1240	620
Pixel size	$\Delta u$	mm	0.25	0.5
#Projections	$K$	-	892	240
Angular spacing	$\Delta\lambda$	deg	0.404	1.5
Absorption model	-	-	monochromatic	monochromatic
Photon energy	$E_p$	keV	80	50
#Photons per pixel	$N_0$	-	-	30 000

**Table 7.3:** Fan-beam simulation parameters for FORBILD phantom.

Parameter	Symbol	Unit	Ideal / Noisy
SID	$S$	mm	600
DID	$D$	mm	600
Detector size	$I \times J$	-	$640 \times 480$
Pixel size	$\Delta u \times \Delta v$	mm <sup>2</sup>	$1.2 \times 1.2$
#Projections	$K$	-	512
Angular spacing	$\Delta\lambda$	deg	0.703
Absorption model	-	-	monochromatic
Photon energy	$E_p$	keV	80
#Photons per pixel	$N_0$	-	0 / 5000

**Table 7.4:** Cone-beam simulation parameters for FORBILD phantom.

motion-corrupted data using the modified *Oscil* motion. Both were simulated with high-quality, noise-free and low-quality, noisy settings. Simulation parameters are summarized in Tab. 7.3. The detector was simulated to be in the rotation center for convenience.

CBCT simulations are performed very similar to those in the fan-beam case. We also use a monochromatic model, and render noise-free and noisy data. However, a total of six motion types were simulated, including five motion patterns presented in Equations (7.2) to (7.6) and a ground-truth, motion-free rendering. All motion patterns were rendered with noise-free and noisy settings yielding 12 sets of projection images. Simulation parameters for the cone-beam geometry are given in Tab. 7.4.

### 7.1.3 XCAT Dynamic Squat Phantom

The 4-D XCAT phantom is based on segmentations of real anatomies of a full-body CT scan [Sega 10]. A large number of anatomies was segmented using surfaces, represented by B-splines. Rigid and nonrigid motion can be incorporated, modeled by a

Parameter	Symbol	Unit	Ideal
SID	$S$	mm	780
DID	$D$	mm	418
Detector size	$I \times J$	-	$620 \times 480$
Pixel size	$\Delta u \times \Delta v$	mm <sup>2</sup>	$1.2 \times 1.2$
#Projections	$K$	-	256
Angular spacing	$\Delta\lambda$	deg	1.406
Absorption model	-	-	monochromatic
Photon energy	$E_p$	keV	80

**Table 7.5:** Simulation parameters for XCAT dynamic squat phantom.

deformation of the spline surfaces over time. Choi et al. developed an XCAT-based numerical phantom of knees for weight-bearing conditions [Choi 13]. Motion of nine healthy subjects, holding a squat at 60° flexion angle for 20 s, was recorded using an optical tracking system. These motion parameters were subsequently transferred to the XCAT phantom, creating a numerical knee phantom based on realistic motion patterns. In this thesis we made use of this knee phantom using the most severe motion pattern acquired during the study, showing an average translational variation of 3.43 mm and an average flexion angle variation of 0.63°. The simulated scene includes leg tissue as well as cortical bone and bone marrow of femur, tibia, patella, and fibula. The transition of tracking system coordinates to XCAT parameters is detailed in [Choi 13].

Within this thesis, the XCAT dynamic squat phantom was used for a qualitative evaluation of a motion correction based on FCC. Simulation parameters differed to [Choi 13], as the FCC method is limited to full 360° rotation angles and projection images without lateral truncation. The phantom was rendered with an increased pixel spacing, such that no lateral truncation was present. We chose to use the same monochromatic physical model as for the cone-beam simulations of the FORBILD phantom. The parameters are presented in Tab. 7.5. Projection images were generated with and without motion, yielding two sets of projection images.

## 7.2 Real CBCT Acquisitions

Our evaluation also involves real weight-bearing acquisitions of healthy volunteers. At this point we refer to Sec. 2.2.2 where we have introduced the weight-bearing CBCT scanners in more detail. We had access to data from three volunteers, consisting of multiple acquisitions based on the prototype system, but also on the robotic C-arm device. Overall 12 real scans have been used for evaluation. In the following we describe the acquisition protocols and parameters for each of the three volunteers.

Parameter	Symbol	Unit	Supine
SID	$S$	mm	780
DID	$D$	mm	418
Detector size	$I \times J$	-	$1240 \times 960$
Pixel size	$\Delta u \times \Delta v$	mm <sup>2</sup>	$0.308 \times 0.308$
#Projections	$K$	-	494
Angular spacing	$\Delta\lambda$	deg	$\sim 0.403$
Acquisition time	-	s	$\sim 20$
Tube voltage	-	kVp	70
Dose at detector (per projection)	-	$\mu\text{Gy}$	1.2

**Table 7.6:** Acquisition parameters for the supine scan of Subject 1.

### 7.2.1 Subject 1

Subject 1 refers to a female volunteer with a weight of 102 kg and a height of 165 cm. The subject was part of the cartilage study described in Sec. 2.3.2, however, in this thesis only a single supine scan is used to validate the correction method for marker-induced metallic artifacts. The FOV of the supine scan was centered to the knee joint of the left leg, while the right leg was not part of the FOV. Scanning was performed with the prototype system (Axiom Artis dTA, Siemens AG, Forchheim, Germany) in a supine position after injection of contrast agent, used for improved cartilage visibility. Projection matrices for each projection image were obtained from the system’s calibration database. The system used an AEC which regulates tube voltage and tube current, to ensure a certain dose level at the detector. To correct for AEC-based detector saturation artifacts, explained in Sec. 2.4.2, a layer of plasticine was used to cover the leg at its periphery. Eight tantalum markers with a diameter of 1 mm are attached at the skin in proximity to the knee joint. Further acquisition parameters can be obtained from Tab. 7.6. After acquisition, the projection images underwent preprocessing to convert their intensities to the line integral domain, which included an AEC compensation [Schw 10] and low contrast water correction [Zell 05].

### 7.2.2 Subject 2

Subject 2 is a male volunteer who was part of the feasibility study based on the prototype system (Axiom Artis dTA, Siemens AG, Forchheim, Germany). We use a supine and three weight-bearing scans for a joint evaluation of the marker- and registration-based approach. The legs have been adjacent to each other and were wrapped with a plasticine layer to avoid saturation artifacts (cf. Sec. 2.4.2). 16 metallic tantalum markers with 1 mm diameter were attached to both legs for motion correction purposes.

Acquisition began with a high-resolution supine scan, followed by three weight-bearing acquisitions. For all acquisitions the FOV was adjusted to the center of the

Parameter	Symbol	Unit	Supine	Weight-bearing
SID	$S$	mm	780	780
DID	$D$	mm	418	418
Detector size	$I \times J$	-	$1240 \times 960$	$1240 \times 960$
Pixel size	$\Delta u \times \Delta v$	mm <sup>2</sup>	$0.308 \times 0.308$	$0.308 \times 0.308$
#Projections	$K$	-	495	248
Angular spacing	$\Delta\lambda$	deg	$\sim 0.403$	$\sim 0.806$
Acquisition time	-	s	$\sim 20$	$\sim 8$
Tube voltage	-	kVp	70	70
Dose at detector (per projection)	-	$\mu\text{Gy}$	1.2	1.2

**Table 7.7:** Parameters used for acquisitions of Subject 2.

two knee joints. Only a small amount of lateral truncation occurred at the periphery of the legs. Bones and the joint were not affected from lateral truncation. The weight-bearing scans were acquired in a straight standing ( $0^\circ$  flexion angle) and in squatting positions, for flexion angles of  $35^\circ$  and  $60^\circ$ . The amount of motion increased with the flexion angle, where the  $35^\circ$  and  $60^\circ$  scans showed the highest amount of motion throughout all subjects acquired for the feasibility study. As for Subject 1, an AEC compensation [Schw 10] and low contrast water correction [Zell 05] was applied, to convert projection images to the line integral domain. The acquisition parameters are provided by Tab. 7.7.

The weight-bearing trajectory required a system calibration to obtain projection matrices suitable for image reconstruction. Calibration was performed prior to acquisition of the subject, by scanning a calibration phantom which provided ground-truth locations, allowing the estimation of projection matrices by solving a linear system of equations. For more details we refer to Hoppe [Hopp 09, pp. 36ff.] and Choi et al. [Choi 14c].

### 7.2.3 Subject 3

Subject 3 corresponds to a female volunteer with a weight of 70 kg, a height of 163 cm, and an age of 58. The subject participated in the cartilage study described in Sec. 2.3.2. Acquisitions were conducted using the robotic C-arm device (Artis Zeego, Siemens Healthcare GmbH, Forchheim, Germany). The subject represents the most recently scanned weight-bearing data used within this thesis (Exam date: June 2015). The acquisition protocol for the cartilage study is well defined and was previously explained in Sec. 2.3.2. Aside from a high-resolution supine scan, six weight-bearing scans have been acquired over a period of 25 min. Supine and weight-bearing acquisitions had an FOV that focused on the left leg. The subject was asked to pull up the right leg during supine scanning, thus only a single leg is visible in the motion-free reconstructions. Contrast agent was injected into the joint prior to the acquisitions

Parameter	Symbol	Unit	Supine	Weight-bearing
SID	$S$	mm	780	780
DID	$D$	mm	418	418
Detector size	$I \times J$	-	$1240 \times 960$	$1240 \times 960$
Pixel size	$\Delta u \times \Delta v$	mm <sup>2</sup>	$0.308 \times 0.308$	$0.308 \times 0.308$
#Projections	$K$	-	496	248
Angular spacing	$\Delta\lambda$	deg	$\sim 0.404$	$\sim 0.806$
Acquisition time	-	s	$\sim 20$	$\sim 8$
Tube voltage	-	kVp	81	81
Dose at detector (per projection)	-	$\mu\text{Gy}$	1.2	1.2

**Table 7.8:** Parameters used for acquisitions of Subject 3.

and is visible in all acquisitions. Six metallic tantalum markers of 1 mm diameter were attached to the leg of interest and four to the right leg.

A multi-sweep trajectory, performing three sequential acquisitions, was used at the beginning of weight-bearing scanning to improve the temporal resolution at the initial weight-bearing phase. Calibration of the trajectory was performed as explained for Subject 2, but with an additional calibration scan for the multi-sweep trajectory. In consequence, the first three weight-bearing scans are performed within the first 30 s after application of load, whereas the remaining scans are acquired after 3 min, 10 min and 25 min, respectively. In contrast to Subject 1 and Subject 2, preprocessing of projection images, i. e., conversion to line integrals, was performed using the vendor’s software.

The subject was selected for use within this thesis, because the weight-bearing scans have a large spectrum w. r. t. the amount of motion. The first and second acquisition contain severe motion artifacts, the third acquisition suffers from moderate motion and acquisitions 4 to 6 contain only little motion artifacts. For more detailed acquisition parameters we refer to Tab. 7.8.

### 7.3 Common Image Quality Metrics

A quantitative measure of image quality is important to allow for an objective comparison of methods and validate the performance of an algorithm. Within the scope of this thesis we use a set of quantitative measures which are explained in the following. All of the presented measures are able to compare two image functions and can be used for functions of arbitrary dimension, e. g., 2-D fan-beam or 3-D cone-beam reconstructions. Without loss of generality, we define the measures to compare similarity of an object function  $f(\mathbf{x}) : \mathbb{R}^N \rightarrow \mathbb{R}$  with a reference function  $v(\mathbf{x}) : \mathbb{R}^N \rightarrow \mathbb{R}$ .

### 7.3.1 MSE, RMSE and rRMSE

A well known measure for image similarity of two functions is the **RMSE**, which is the square root of the **MSE**. Let us define the **MSE** between  $f(\mathbf{x})$  and  $v(\mathbf{x})$  by

$$\text{MSE}(f, v) = \frac{1}{(\#V)} \sum_{\mathbf{x} \in V} (f(\mathbf{x}) - v(\mathbf{x}))^2, \quad (7.7)$$

which is simply the average of the squared differences between the two functions. Further,  $V$  is the set that holds all locations  $\mathbf{x}$  which are used for evaluation. Consequently, we can define the **RMSE** but also the **rRMSE** by

$$\text{RMSE}(f, v) = \sqrt{\text{MSE}(f, v)} \quad (7.8)$$

$$\text{rRMSE}(f, v) = \frac{\text{RMSE}(f, v)}{I_v(v)}, \quad (7.9)$$

where  $I_v(v)$  is the intensity range of the reference function  $v$  and may be defined by

$$I_v(v) = \max_{\mathbf{x} \in V} (v(\mathbf{x})) - \min_{\mathbf{x} \in V} (v(\mathbf{x})). \quad (7.10)$$

A benefit of using **rRMSE** is its invariance to intensity scaling and that it can be presented as percentage, relative to the reference image.

### 7.3.2 SSIM and UQI

Wang et al. showed that the **RMSE** might not reflect the visual perception of differences between two images [Wang 04]. In contrast, **SSIM** and **UQI** are both measures that have been designed to incorporate perceptive information. They combine information on intensity, contrast and structural differences of two images into a single measure. The **SSIM** is defined by

$$\text{SSIM}(f, v) = \frac{(2\mu_f\mu_v + C_1)(2\sigma_{fv} + C_2)}{(\mu_f^2 + \mu_v^2 + C_1)(\sigma_f^2 + \sigma_v^2 + C_2)}, \quad (7.11)$$

where  $\mu_f$  and  $\mu_v$  are the mean intensities of object and reference function over all positions  $\mathbf{x} \in V$ . Similarly,  $\sigma_f$  and  $\sigma_v$  are the standard deviations and  $\sigma_{fv}$  is the correlation coefficient. The **SSIM** contains two constants,  $C_1$  and  $C_2$ , which assure stability of the measure in case the evaluated functions have a zero mean ( $\mu_f^2 + \mu_v^2 \approx 0$ ) or are without significant variation ( $\sigma_f^2 + \sigma_v^2 \approx 0$ ).

The **UQI** measure [Wang 02] can be interpreted as the predecessor of **SSIM**. They are identical when setting the constants associated with the **SSIM** to zero. Thus, when we refer to the **UQI**, we computed the **SSIM** for  $C_1 = C_2 = 0$ . For all experiments based on the **SSIM**, the constants are adjusted as proposed in the original formulation by Wang et al. [Wang 04]:  $C_1 = (0.01 I_v(v))^2$  and  $C_2 = (0.03 I_v(v))^2$ .

#### Evaluation Strategies for SSIM and UQI

We differentiate between three different types of evaluating **SSIM** and **UQI**.

1. The set  $V$  may be adjusted to a manually selected ROI that is to be compared.
2. Complete images are compared element-wise, where for each element an SSIM or UQI value over a certain neighborhood is obtained. The final value is the mean of all local SSIM or UQI measurements. Note that this method represents the original formulation for comparing 2-D images [Wang 04]. For 3-D reconstructions we used a cubical neighborhood of  $9 \times 9 \times 9$ .
3. Variant 2 can be further adjusted by limiting the voxel or pixel locations to a certain ROI, e. g., a segmented bone.

# Joint-Evaluation of Registration- and Marker-Based Method

8.1	Materials and Methods .....	100
8.2	Reconstruction Results.....	102
8.3	Quantitative Assessment of Image Quality .....	108
8.4	Cross-Evaluation at Markers and Bone Surfaces.....	111
8.5	Removal of Marker-Induced Metal Artifacts.....	115
8.6	Discussion and Conclusion .....	117

The main goal of this evaluation is to verify if the results of the marker-based method can be achieved by a registration-based method and which of the methods yields the best image quality at the knee joint. Specifically, we investigate if the bone outline is accurately restored. Also, the quality of the contrast boundary for data of Subject 3 is analyzed, which is an important factor for the cartilage study introduced in Sec. 2.3.2. Additionally, we want to identify if the marker-based method is able to restore the shape of structures close to the knee joint. In contrast, the accuracy of the registration-based method is evaluated w. r. t. the legs periphery, where the deformation fields are based on TPS extrapolations of internal bone motion.

Comparative qualitative and quantitative evaluations of the 2-D/3-D registration approach and the marker-based correction are shown in this chapter. The majority of evaluations have been done for both marker-based and 2-D/3-D registration, providing a good overview of advantages and disadvantages for each of the methods. In Sec. 8.1 we summarize all evaluation approaches and describe parameter settings used for image reconstruction as well as the individual methods. A qualitative and quantitative comparison of image quality is shown in Sections 8.2 and 8.3. In Sec. 8.4 we analyze the result of the marker-based method w. r. t. the internal bone structures and the performance of registration approach w. r. t. the periphery. This is followed by a comparison of different correction methods used for removal of marker-induced metal artifacts in Sec. 8.5. The chapter ends with a discussion and conclusion presented in Sec. 8.6.

## 8.1 Materials and Methods

Motion-corrected reconstructions contain inaccuracies based on all individual steps of a motion correction method, e. g., parameter settings, registration accuracy, or wrong point correspondences. Therefore, the focus of our assessment is image quality in reconstruction domain, which allows for a fair and complete comparison of methods and provides an initial validation of clinical applicability. An important field for future applications of weight-bearing imaging is the improved diagnosis of cartilage diseases. For example, in the cartilage study introduced in Sec. 2.3.2, we are interested in deformation and strain properties measured between tibia and femur to the contrast agent's surface. Other recently introduced measures are also based on segmentations of tibia and femur [Cao 15], thus, requiring a suitable image quality of these bones. As a consequence, the evaluation pipeline used within this chapter is tailored to focus on image quality of bone structures.

Evaluations include the marker-based approach (cf. Chap. 4), two variants of the motion correction method based on 2-D/3-D registrations (cf. Chap. 5), but also reconstructions without motion correction. We separate between registrations using GC and NGI as similarity measure. Within this chapter we refer to the individual methods by abbreviations MB, GC, NGI, and NoCorr. The evaluations are based on real weight-bearing acquisitions, i. e., the data of Subject 1 to Subject 3 as introduced in Sec. 7.2. We use the labels defined in Tab. 7.1 to refer to individual datasets. One part of the evaluation is a comparison between 2-D/3-D registration using GC and NGI. Results based on data from Subject 2 revealed a superior performance of NGI, hence, evaluation for Subject 3 is limited to NGI, MB, and NoCorr.

### 8.1.1 Parameter Selection And Image Reconstruction

Let us first introduce the software tools and parameter settings for the individual methods and datasets, as well as the reconstruction pipeline used during motion compensated reconstruction. The registration-based method and some of our evaluation methods require segmentations of tibia and femur. Segmentation masks and surface meshes were obtained semi-automatically using ITK-Snap [Yush 06], based on reconstructions of supine datasets S2\_SUP and S3\_SUP. Note, that S2\_SUP included left and right leg, whereas S3\_SUP showed only the left leg, allowing segmentation of four and two bones, respectively. Subsequent dilation and erosion of segmentation masks, as described in Sec. 5.6, was done with the software 3-D Slicer [Piep 04]. The same software was used for an initial, rigid 3-D/3-D registration of segmented bone volumes to motion-corrupted reconstructions. The rigid transform parameters were manually adjusted within 3-D Slicer, based on a superimposed visualization of supine and initial weight-bearing reconstructions.

#### Parameters for MB

Subject 2 had 16 and Subject 3 had 10 markers attached, which were distributed over both legs. Not necessarily all markers were in the FOV during the weight-bearing scans. We counted the number of visible markers in the FRST result and adjusted

	S2_WB0	S2_WB35	S2_WB60	Subject 3 (all scans)
$M$	16	15	14	10
$t_{\text{proj}}$	0.1	0.17	0.17	0.3

**Table 8.1:** Parameter settings for marker-based motion correction.

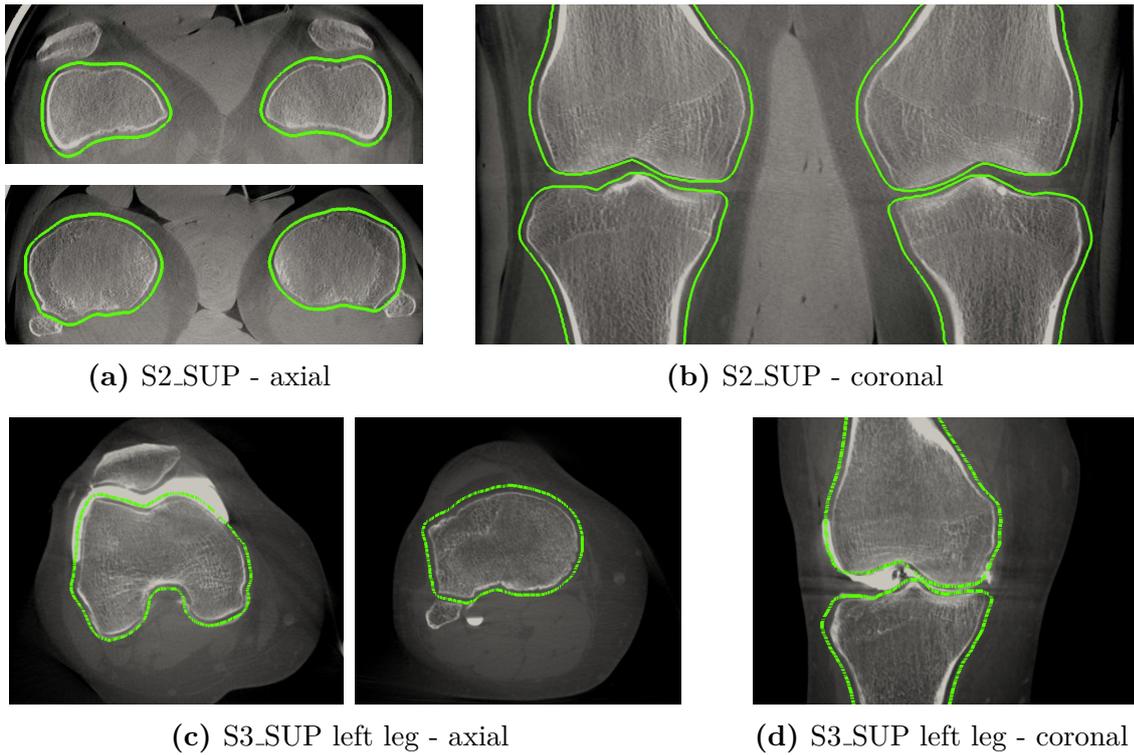
the number of markers  $M$  accordingly. The binarization threshold  $t_{\text{proj}}$  of the **FRST** is manually adjusted. An **MIP** of the **FRST** result is used to adjust  $t_{\text{proj}}$  as it allows a clear separation of noise and marker intensities. Further, we set intensities at the boundary of projection images to zero in the **FRST** result, as the systems' collimator edge led to wrong detections. We used a detection radius of  $r_f = 3.0$  px and a radial strictness of  $c_{\text{rad}} = 3$  for all datasets and fixed the scaling parameter  $s_f$  according to [Loy 03]. The maximum search distance for point correspondences is set to  $d_u = 70$  px and  $d_u = 150$  px for Subject 2 and Subject 3, respectively. Selected values for the number of beads  $M$  and the binarization threshold  $t_{\text{proj}}$  are provided in Tab. 8.1 for each dataset. Outlier removal was applied iteratively for  $N_{\text{opt}} = 6$  times, detecting  $\delta = 0.5\%$  of worst contributions for each iteration, where the minimum number of detections to retain per projection image was given by  $M_{\text{min}} = 6$ .

### Parameters for **GC** and **NGI**

A multiresolution approach was used for the 2-D/3-D registration, where the resolution of the projection images and the ray tracing step size for **DRR** generation was adjusted accordingly. The first optimization was done with a downsampled projection image size of  $310 \times 240$ . For Subject 3 we used a quadrupled step size for ray tracing while blurring the gradient volume using a 3-D Gaussian with  $\sigma = 1$ . This was mainly done to reduce the total runtime of the registration as twice as many datasets were present compared to Subject 2. In a second optimization the parameters were initialized with results from the first optimization and an image size of  $620 \times 480$  was used. The Lagrangian for the temporal smoothness regularizer was adjusted to  $\eta = 5$  and  $\eta = 0.5$  for Subject 2 and Subject 3, respectively. It was further reduced to  $\eta = 0.1$  for S3\_WB1 to account for the severe and rapidly changing type of motion. The standard deviation of the Gaussian was set to  $\sigma_t = 2$  for all cases. If not mentioned otherwise, parameters were manually determined on a single dataset and kept constant for all optimizations within a subject. We made sure that the initial values of **GC** and **NGI** had a similar range, using a normalization factor. This is important to ensure a similar influence of the smoothness regularizer for both **GC** and **NGI**. The reference frame needed for generation of **TPS** motion fields was set to  $\hat{k} = 122$ , as this corresponded to an posteroanterior view in all datasets with little overlap of bones, ensuring an accurate 2-D/3-D registration.

### Reconstruction of Real Data

Reconstructions were obtained on a grid of  $512 \times 512 \times 256$  voxels with an isotropic resolution of 0.5 mm for all datasets and methods. The reconstruction of S3\_SUP was



**Figure 8.1:** Reconstructions of supine data S2\_SUP and S3\_SUP. Only the left leg was scanned for S3\_SUP. (a) and (c): Axial slices inferior and superior of the knee joint. (b) and (d): coronal slices showing tibia and femur. The ROIs depicted by the solid, green line are used for qualitative evaluation.

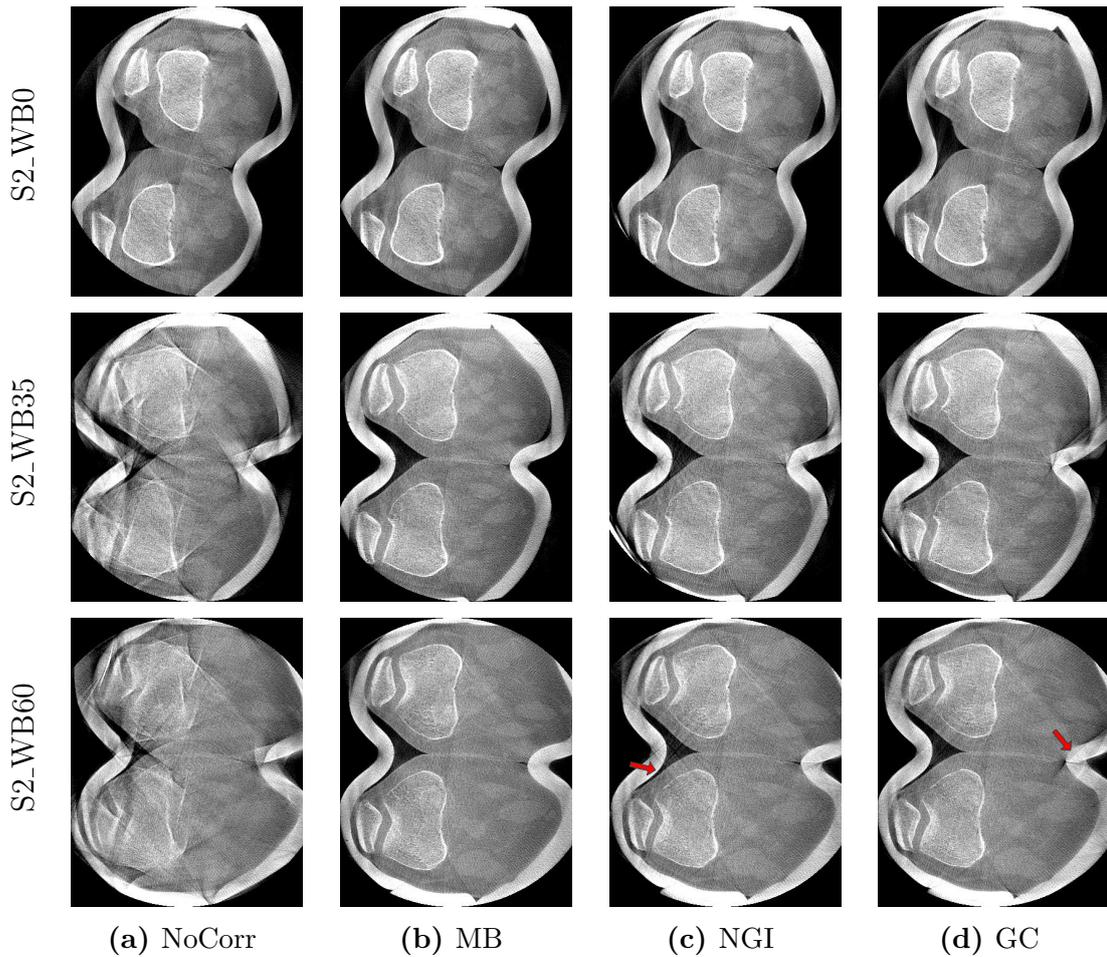
done on a voxel grid of size  $512 \times 512 \times 512$  to improve visibility of the tibia. Intermediate backprojections of the FRST, used to detect 3-D marker locations, required a larger reconstruction volume to allow detection of laterally truncated markers. The volume size was adjusted to  $512 \times 512 \times 512$  with an isotropic voxel size of 0.7 mm.

The following steps are included for all reconstructions: 1) a simplified beam-scatter-kernel scatter estimation [Ruh11] assuming that the object consists only of water and that the water-equivalent-thickness is uniform, 2) cosine weighting, 3) Parker redundancy weighting [Park82], 4) a simple truncation correction [Ohne00], 5) ramp filtering with a smooth Shepp-Logan kernel [Kak01] and 7) a motion-compensated GPU backprojector [Sche07].

Rigid motion for MB and NoCorr is directly incorporated into projection matrices followed by a regular GPU-based backprojector, whereas, GC and NGI used a backprojector that evaluates the generated nonrigid TPS deformations. Please see Sec. 4.2 and Sec. 5.5 for more information.

## 8.2 Reconstruction Results

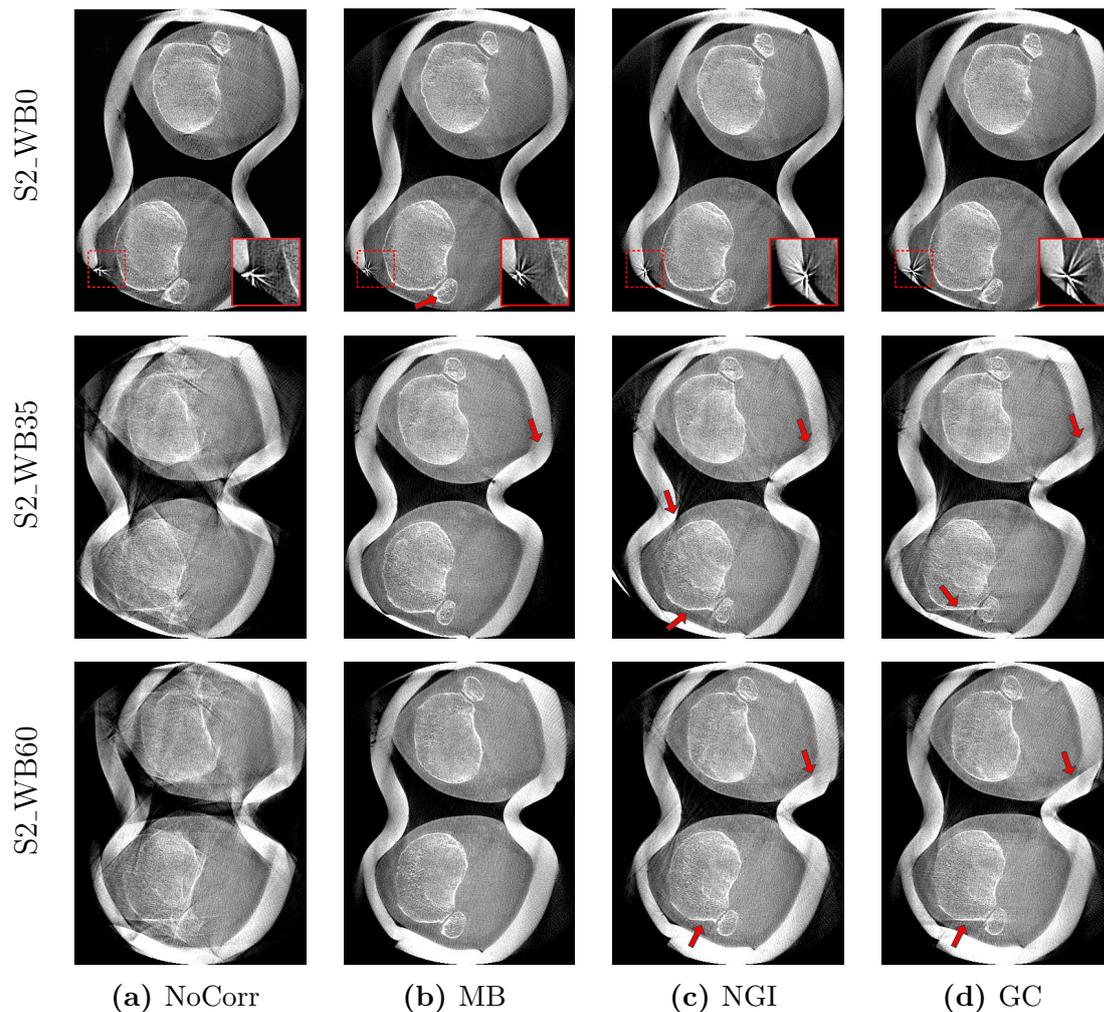
Let us first present reconstructions of the supine, motion-free data shown in Fig. 8.1. Figures 8.1(a) and 8.1(b) and Figures 8.1(c) and 8.1(d) show axial and coronal slices for reconstructions of S2\_SUP and S3\_SUP, respectively. For both reconstructions we



**Figure 8.2:** Axial slices through femur and tibia of weight-bearing reconstructions from Subject 2. From (a) to (d): Without correction, the MB method, registration-based methods using NGI and GC. The rows correspond to S2\_WB0 ( $0^\circ$  flexion), S2\_WB35 ( $35^\circ$  flexion), and S2\_WB60 ( $60^\circ$  flexion). (W: 2025 HU, C: 145 HU).

added an overlay depicting the ROIs used for quantitative analysis of image quality presented in Sec. 8.3. Note the plasticine surrounding the legs in S2\_SUP, which was used to prevent detector saturation explained in Sec. 2.4.2. Saturation was avoided in S3\_SUP by imaging only the left leg. The injected contrast agent for cartilage visualization can be clearly seen between femur and patella in Fig. 8.1(c) and between femur and tibia in Fig. 8.1(d). Due to the supine position, no motion related artifacts are visible. In general, the double amount of angular resolution led to a better image quality, compared to the weight-bearing reconstructions.

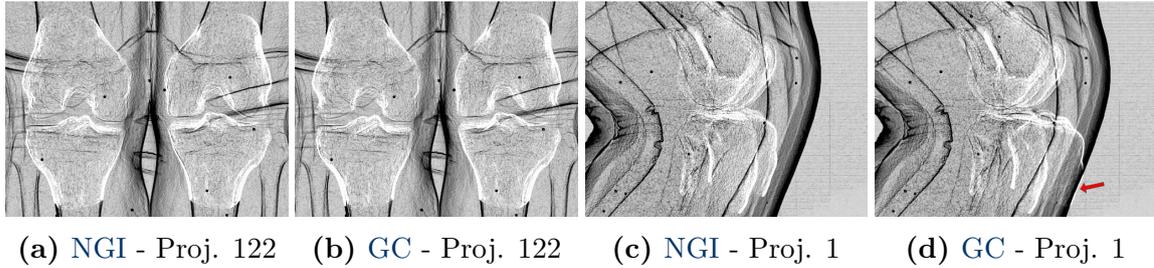
Axial slices of three weight-bearing reconstructions from Subject 2 are visualized in Figures 8.2 and 8.3, where the first shows superior slices through femur and patella and the latter inferior slices through tibia and fibula. Only little motion artifacts are present for S2\_WB0, with slight streaking at the bones' outlines in case of NoCorr. All methods restored the bone structures and show similar image quality as can be seen in the top row of Figures 8.2 and 8.3. However, the MB method shows a slightly sharper appearance of the fibula indicated by a red arrow. We have embedded zoomed views



**Figure 8.3:** Axial slices through tibia and fibula of weight-bearing reconstructions from Subject 2. Arrangement is identical to Fig. 8.2. (W: 2025 HU, C: 145 HU).

to the reconstructed marker in the inferior images of S2\_WB0 (cf. top row Fig. 8.3). The zooms reveal an accurate reconstruction of the marker for MB and NGI, whereas GC shows a slight increase of streaking artifacts.

NoCorr reconstructions of S2\_WB35 and S2\_WB60 show severe motion artifacts. All motion correction methods reduced streaking artifacts and improved image quality drastically. Within the correction methods only minor differences are seen at bone outlines of femur and patella in Fig. 8.2. As expected, a better restoration of the transition between skin and plasticine was achieved by MB, whereas slight streaking artifacts originate from the misaligned plasticine in NGI and GC. For the inferior slices in Fig. 8.3, MB and NGI show similar image quality with slightly more residual streaking of NGI, especially for S2\_WB35. The GC method was not able to accurately restore the left tibia and shows increased streaking artifacts, particularly for the left tibia. This is in line with misregistrations that occurred for the GC similarity measure. Fig. 8.4 depicts registration results for S2\_WB60 in projection domain, using an overlay of gradient magnitudes derived from the registered DRR images (positive intensities) and the acquired projection images (negative intensities).

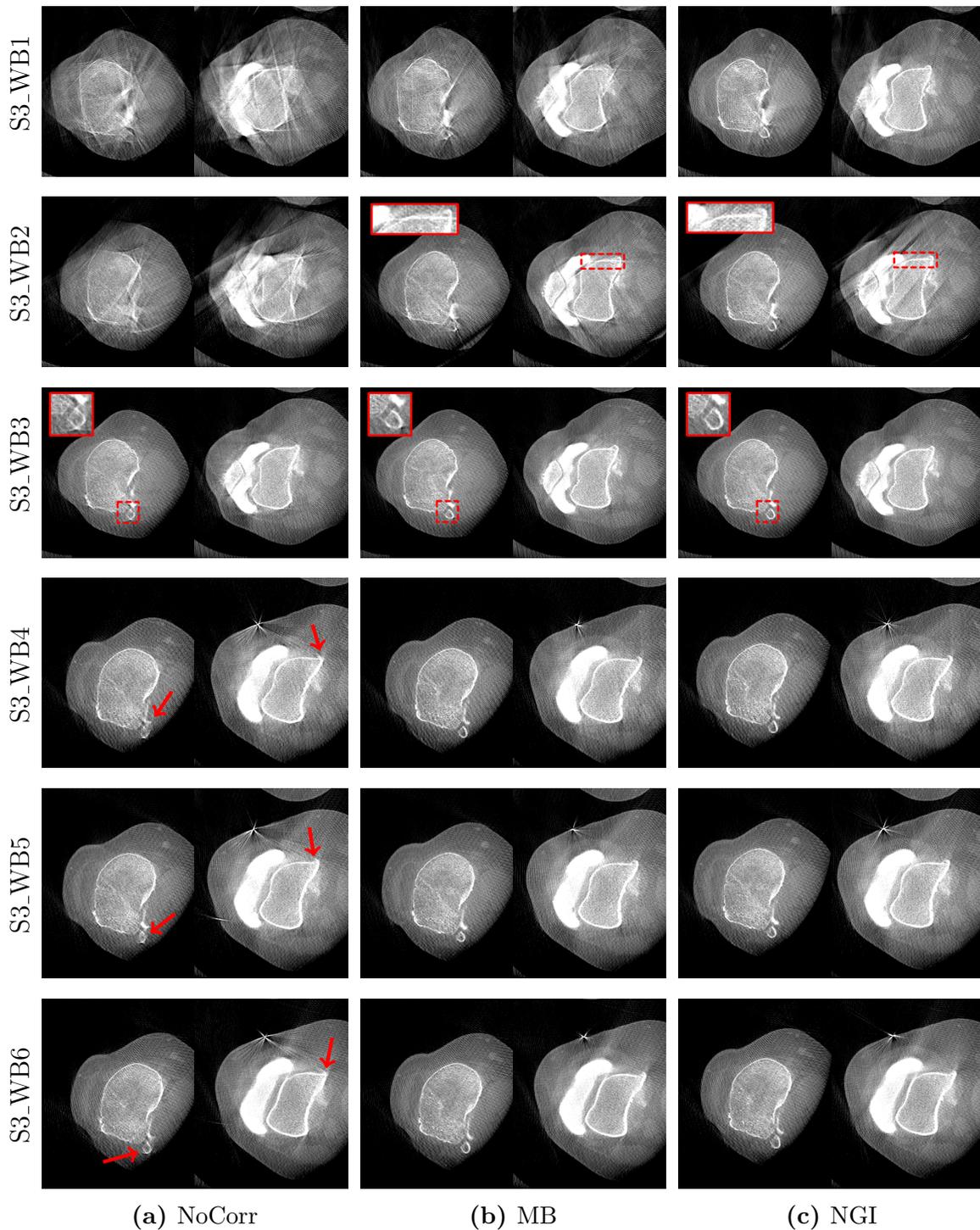


**Figure 8.4:** Difference image of gradient magnitudes extracted from registered DRRs and the acquired projection images from dataset S2\_WB60. Left: Projection images corresponding to the reference coordinate system. Right: Lateral projections with multiple occlusions, which led to a wrong registration of the left tibia for GC.

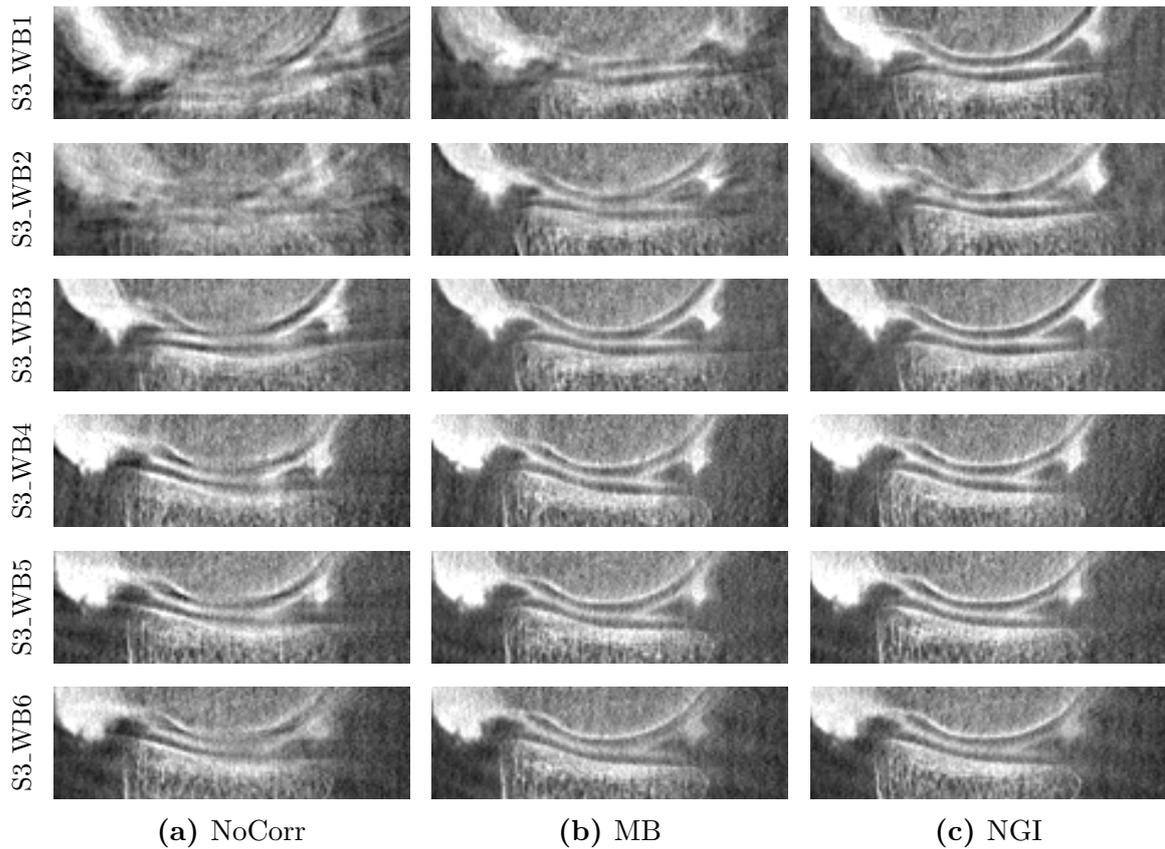
Accurate alignments are shown for a posteroanterior view direction for NGI and GC in Figures 8.4(a) and 8.4(b). Note that this view was also selected for extraction of the reference coordinate system. Registration was more problematic for lateral views, as shown in Figures 8.4(c) and 8.4(d). Individual bones are overlapping and skin- as well as plasticine boundaries create additional distractions for gradient-based similarity measures. Despite these effects, NGI was able to align tibia and femur accurately, whereas the left tibia was registered to the plasticine in case of GC, causing streaking artifacts in reconstruction domain.

Reconstructed images for Subject 3, i.e., for the cartilage study, are shown in Fig. 8.5 for all six acquired weight-bearing scans. To reduce overall complexity, we excluded GC from the evaluation of Subject 3, because NGI showed a superior performance for all cases of Subject 2. The individual datasets are shown on the rows, where S3\_WB1 to S3\_WB3 originate from a single acquisition with three sweeps acquired at the time load is applied to the leg. S3\_WB4 to S3\_WB6 are single sweep scans, acquired 3 min, 10 min and 25 min after application of load. The columns correspond to reconstruction results for NoCorr, MB, and NGI, respectively. Each image is a concatenation of an inferior axial slice through tibia and proximal fibula (left) as well as a superior axial slice through the femur (right). Visual observations of the projection images revealed very severe motion for S3\_WB1 and S3\_WB2, moderate motion for S3\_WB3 and only little motion for S3\_WB4 to S3\_WB6. This is supported by the reconstruction results of NoCorr in Fig. 8.5(a), showing a high amount of streaking and clear structural loss of the bones, soft tissue and contrast agent for S3\_WB1 and S3\_WB2. Mild streaking artifacts are also present at tibia, fibula, femur, and contrast agent of S3\_WB3. Motion artifacts for S3\_WB4 to S3\_WB6 are small and manifest mostly as blurring, which can be seen well at the proximal fibula and at the small extension of the femur’s medial surface. The artifacts can be corrected well for the last four acquisitions by both NGI and MB, yielding a sharper outline of fibula and femur. Note the metallic marker in the last three rows of Fig. 8.5. Both correction methods could improve the marker’s locality after motion correction, yet, with superior performance of MB. In contrast, NGI yielded a sharper reconstruction of the fibula for S3\_WB3.

Even though both methods can drastically improve image quality in case of severe motion (S3\_WB1 and S3\_WB2), a fair amount of streaking artifacts remain after



**Figure 8.5:** Each image shows an axial slice through tibia (left) and femur (right) of weight-bearing reconstructions from Subject 3. From (a) to (c): Without correction, the MB method, and NGI. The rows correspond to individual weight-bearing acquisitions, scanned at beginning of load application (first three rows) and after 3 min, 10 min and 25 min (last three rows). (W: 2025 HU, C: 145 HU).



**Figure 8.6:** Registered sagittal reconstructions of medial condyles from Subject 3. Note the contrast agent’s surface which separates femoral and tibial cartilage. Contrast visibility is crucial for further biomechanical analysis, e. g., of cartilage strain. The arrangement of images is identical to Fig. 8.5. (W: 2385 HU, C: 670 HU).

correction. **NGI** generally outperforms **MB** for **S3\_WB1** and also shows more accurate tibial reconstructions for **S3\_WB2**. In contrast, **MB** yields a better restoration of soft tissue surrounding the femur, whereas **NGI** could not accurately correct the surface of the leg, showing a high amount of residual artifacts. However, most of these artifacts originate from soft tissue. In fact, the femur itself has less artifacts for **NGI**, which we emphasized by a zoom of the medial femoral surface.

For further analysis of the cartilage, it is particularly important to correct motion artifacts within the knee joint. Contrast agent injected into the joint is used to distinguish between tibial and femoral cartilage, which are not visible in conventional X-ray imaging. A clear visibility of the contrast agent’s surface w. r. t. tibial and femoral condyles is the basis for biomechanical analysis of the cartilage, e. g., cartilage strain [Choi 14a].

In addition to axial reconstructions inferior and superior to the knee joint (cf. Fig. 8.5), we also show sagittal reconstructions at medial condyles of tibia and femur in Fig. 8.6. To allow for an accurate comparison, all reconstructions are aligned to the supine coordinate system w. r. t. the femur. The procedure used for alignment builds the main part of quantitative analysis of image quality and is detailed in Sec. 8.3.1.

---

**Algorithm 8.1:** Pipeline for quantitative comparison of image quality at bones.

---

```

Create bone-wise evaluation ROIs by 2-D dilation of segmentation masks;
for  $\hat{b} \leftarrow 1$  to  $B$  do                                // Loop over all bones or ROIs
    1) Reconstruction of weight-bearing scan in supine coordinate system;
    2) Refine alignment by 3-D/3-D registration of  $\hat{b}$ -th bone to supine volume;
    3) Updated reconstruction based on registration result;
    4) Compute UQI within current ROI;
end
Extract the final UQI as mean value of bone-wise UQIs;

```

---

Reconstructions without correction show severe motion artifacts for S3\_WB1 and S3\_WB2, such that bone surfaces and the contrast agent cannot be localized. Minor motion artifacts were obtained for S3\_WB3 to S3\_WB6 for the axial reconstructions in Fig. 8.5(a). In contrast, motion-induced blurring of the bone and contrast agent surfaces is clearly visible in the sagittal reconstructions, which make accurate quantitative measurements of cartilage properties impossible. Hence, motion correction is essential even if only little motion is present. In case of severe motion, MB could not sufficiently restore the surface of the contrast agent, whereas a much sharper localization is possible when applying NGI. Both methods were able to correct for motion blur for little to moderate motion, showing comparable image quality for S3\_WB3 to S3\_WB6.

Overall the NGI method yielded the best results and allows for a subsequent analysis of cartilage deformation after application of load. In addition, the contrast agent also allows visualization of menisci. They can be seen in the motion-corrected reconstructions (cf. Fig. 8.6(c)) by the break down of contrast agent at the anterior and posterior of the joint. The acquisitions also depict the wash out of contrast agent over time, where the lowest intensity is obtained for S3\_WB6. Also, the contrast surface in S3\_WB6 appears wider than in previous acquisitions which indicates diffusion of contrast agent into cartilage tissue.

## 8.3 Quantitative Assessment of Image Quality

We use the UQI (cf. Sec. 7.3) to measure the similarity of the weight-bearing reconstructions to that of the motion-free, high-quality, supine scan. Computation of the UQI is limited to ROIs that contain femur and tibia and soft tissue surrounding these bones. The third evaluation method presented in Sec. 7.3.2 is used, computing the average of block-wise UQI results, restricted to bone ROIs.

### 8.3.1 Quantitative Evaluation Pipeline

An overview of the approach used to obtain UQI values for each method and dataset is shown in Algorithm 8.1. In general, accurate quantitative evaluation of image quality in in-vivo data is challenging. The UQI requires a precise geometric alignment of anatomies to allow for such a comparison. The reconstruction of weight-bearing scans

within the supine coordinate system is easily established after successful registration using **NGI** or **GC**. Therefore, the bone transformations  $\mathbf{T}_b^{\text{init}}$ , estimated during initial 3-D/3-D registration, are simply incorporated into the reference transform  $\hat{\mathbf{T}}_{kb}$ , such that the supine scan represents the new reference coordinate system. However, updated reconstructions need to be performed separately for each bone due to several reasons. First, the deformation fields estimated by **TPS** would differ from deformations estimated in the standing reference coordinate system, as they would incorporate nonrigid joint motion between supine and standing scans. Second, **MB** and **NoCorr** methods are not able to perform a nonrigid reconstruction, thus, only a single rigid transform can be incorporated at a time.

To perform alignment for a selected bone  $\hat{b}$ , its initial transform  $\mathbf{T}_{\hat{b}}^{\text{init}}$  needs to be incorporated into the reconstruction process. For the **MB** and **NoCorr** method this can be achieved in a straightforward manner by right multiplication of the motion-corrected (**MB**) or the original (**NoCorr**) projection matrices with  $\hat{\mathbf{T}}_{kb} \mathbf{T}_{\hat{b}}^{\text{init}}$ . Thus, **MB** and **NoCorr** reconstructions are aligned with the supine scan, according to the initial and reference transform of the  $\hat{b}$ -th bone, obtained from a registration approach. We chose to use transformations estimated with **NGI**, as the method proved to yield more stable registration results than **GC**. The nonrigid methods **GC** and **NGI** require only a simple update of Eq. (5.14) to align the reconstructed bone with the supine coordinate system. The updated equation now reads

$$\mathbf{T}_{kb} = \hat{\mathbf{T}}_{kb} \left( \hat{\mathbf{T}}_{kb} \mathbf{T}_{\hat{b}}^{\text{init}} \right)^{-1} . \quad (8.1)$$

As a result, we obtain reconstructions of weight-bearing scans where a selected bone is aligned with that of a supine reconstruction. Their alignment should match, given that **MB** and the registration methods were able to estimate the motion perfectly. However, deviations and errors of the 2-D/3-D registration or the **MB** approach will lead to a misalignment which would be the dominant factor when evaluating the **UQI**. To focus the **UQI** on image quality we incorporate an additional automatic 3-D/3-D registration of the selected bone from weight-bearing to the supine reconstructions, using 3-D Slicer. To make sure that the image quality has minor influence on the 3-D/3-D registration we use an **MI** similarity measure, which yielded accurate alignments even in presence of a high level of motion artifacts, e. g., for reconstructions without motion correction. All registration results have been verified visually to ensure meaningful **UQI** measures. The final registration results are once more incorporated into the reconstruction to avoid an additional interpolation in volume domain. These reconstructions are then used to obtain a bone-wise value for the **UQI**.

In Fig. 8.1 we show axial slices through tibia as well as femur and a coronal slice of the supine reconstructions. The solid, green outline represents the final **ROIs** used for the **MI**-based 3-D/3-D registration but also for evaluation of a bone-wise **UQI**. It can be seen that the **ROIs** contain a margin that covers soft tissue close to the bones, which was incorporated by dilation of the segmentation masks applying a circular structuring element of radius 2.5 mm in the axial planes.

UQI ( $\times 10^2$ )						
Dataset	NoCorr	MB	NGI	GC	NGI- MB	
S2_WB0	34.9	57.2	<b>63.5</b>	62.6	+	6.3
S2_WB35	11.2	51.0	<b>53.1</b>	47.2	+	2.1
S2_WB60	9.0	<b>52.9</b>	51.7	49.9	-	1.2
S3_WB1	14.4	31.0	<b>61.0</b>	-	+	30.0
S3_WB2	18.6	44.2	<b>63.3</b>	-	+	19.1
S3_WB3	40.7	65.8	<b>69.9</b>	-	+	4.1
S3_WB4	42.2	66.0	<b>68.4</b>	-	+	2.4
S3_WB5	51.7	<b>69.4</b>	67.5	-	-	1.9
S3_WB6	44.4	<b>69.0</b>	68.5	-	-	0.5
Total	29.7±16.3	56.3±13.0	<b>63.0±6.7</b>	-	+	6.7±10.8

**Table 8.2:** UQI values over evaluated bone ROIs. The last row shows the mean UQI and its standard deviation over all datasets. All correction methods improved image quality when compared to reconstructions without correction. Bold numbers emphasize the methods that obtained the highest UQI for each dataset.

### 8.3.2 UQI Results

Qualitative evaluation resulted in a UQI for each bone, correction method, and dataset. A total of nine weight-bearing datasets were evaluated, three from Subject 2 including different flexion angles and six of Subject 3 as part of the cartilage study. Tab. 8.2 shows the mean UQI values over all evaluated bone ROIs, i. e., left and right femur and tibia for Subject 2 and only left femur and tibia for Subject 3. For the purpose of visualization we scaled the presented values by a factor of 100, such that its theoretical range lies in  $[-100, 100]$ , with 100 corresponding to identical test and reference images.

The weight-bearing scans had minor variations of intensities due to saturation, truncation, or other artifacts. Also, the supine scans had a higher image quality and potentially different exposure rates due to the AEC which has a certain influence on UQIs determined between weight-bearing and supine data. In fact, for in-vivo data it is not possible to obtain an exact maximum UQI, for the case of a motion-free weight-bearing scan. Previously, cadaver legs have been used to obtain this bound [Choi 14c], yet, this cannot be integrated regularly into weight-bearing acquisition protocols. However, the UQI has shown to be robust to intensity differences [Wang 02], allowing for a distinguished evaluation of motion-correction methods but also individual datasets within a subject.

#### Comparison of datasets

Reconstruction results without motion correction in Figures 8.2(a), 8.3(a), 8.5(a) and 8.6(a), but also a visual inspection of acquired projection images, revealed a

good approximation of motion levels within individual acquisitions. We determined that S2\_WB35, S2\_WB60, S3\_WB1, and S3\_WB2 contain a substantial amount of motion artifacts, whereas the other datasets show only moderate to little degradation of image quality. This is supported by **UQI** values obtained without application of motion correction, showing substantially smaller **UQIs** in case a high level of motion artifacts was present. In addition, the maximum achieved **UQIs** after correction are obtained in case the initial image quality was higher, e. g., 63.5 for S2\_WB0 vs.  $\approx 53$  for S2\_WB35 and S2\_WB60.

### Method Comparison

A large improvement in **UQI** could be observed for all applied motion correction methods w. r. t. no correction. In general, improvements due to motion correction methods are in line with the visual observations obtained in Figures 8.2, 8.3, 8.5 and 8.6.

**GC** has only been validated for Subject 2 and had a consistently smaller **UQI** than **NGI**, which originates from an increased rate of misregistrations (cf. Fig. 8.4) causing streaking artifacts particularly at the tibia (cf. Fig. 8.3(d)).

Let us compare the performance of **MB** and **NGI** methods. The last column of Tab. 8.2 shows the difference in **UQI** between **MB** and **NGI**, where a positive value indicates a better image quality for **NGI**. Reconstructions of S3\_WB3 to S3\_WB6 revealed only little differences between **MB** and **NGI** with a slightly better performance of **NGI** in case of S3\_WB3 (see zooms to tibia in Fig. 8.5). However, clearly less artifacts were present when using **NGI** on S3\_WB1 and S3\_WB2. Both observations are in agreement with the **UQI** values obtained for Subject 3, where **NGI** yields a **UQI** which is 19.1 and 30.0 higher than **MB** for S3\_WB2 and S3\_WB1, respectively.

For Subject 2 the **MB** method yielded a better visual quality in the reconstructions, with less streaking and a better restoration the skin and attached plasticine. Nevertheless, **NGI** yielded higher **UQI** values for S2\_WB0 and S2\_WB35. Many of the residual streaking artifacts for **NGI** originate from the boundary of skin and plasticine and do not contribute much to the image quality at the bones. Additionally, we have noticed a slight deformation of bones in case of S2\_WB0 which could support this discrepancy. It is analyzed in detail in Sec. 8.4.

## 8.4 Cross-Evaluation at Markers and Bone Surfaces

**MB** builds on location measurements of markers attached to the skin. Thus, internal motion of tissue and bones which deviates from the skin's motion may not be covered. In contrast, **NGI** and **GC** are based on bone motion estimated by 2-D/3-D registration, where motion at the periphery of the leg is extrapolated based on a **TPS** deformation model. The goal of this evaluation is to determine the accuracy of the registration-based approaches at a certain distance to the bones. Additionally, we investigate if **MB** leads a loss of image quality at internal structures, e. g., at tibia or femur.

### 8.4.1 Bone Surface Profiles

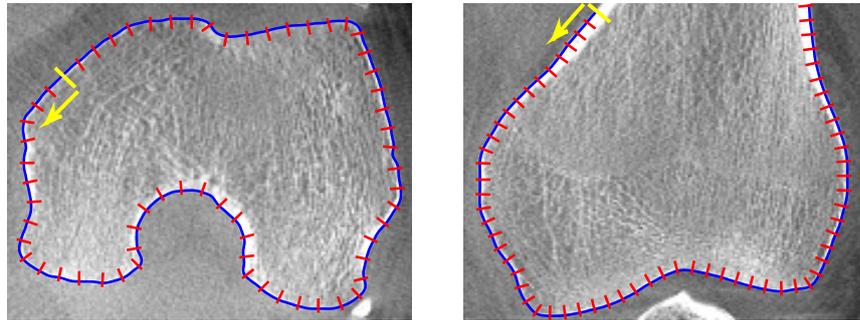
For some datasets quantitative evaluation revealed a higher **UQI** value for **NGI** compared to **MB**, which was in contradiction to the visual reconstruction results. However, we noticed a slight deformation of the bone’s outline for **MB** when compared to the supine reconstruction, which is likely to reduce **UQI** measurements. For further analysis of this deformation, line profiles were extracted orthogonal to the outline of the femur for **NGI**, **MB**, and the supine data, in aligned reconstructions extracted during **UQI** measurement (cf. Sec. 8.3.1). The surface of the existing bone segmentations typically extended the visible outline of the bones by a small amount. Thus, we applied a manual refinement of the segmentations for representative axial and coronal slices using **ITK-Snap** [Yush 06]. This guaranteed that the line profiles should intersect the bone surface at the same position. Subsequently, 2-D spline curves were fitted to the intersection of the selected axial and coronal slices with the segmentation surface. An analytic computation of lines perpendicular to the spline can be easily obtained by its derivative, allowing for an arbitrary density of the sampling pattern. We sampled line profiles equidistantly along the spline curves and concatenated them to an image representation.

The sampling procedure and the selected slices are visualized in Fig. 8.7(a) for **S2\_SUP**. We have selected **S2\_WB0** for the evaluation as it had the highest difference of **UQI** between **MB** and **NGI**, while showing slightly better visual results for **MB**.

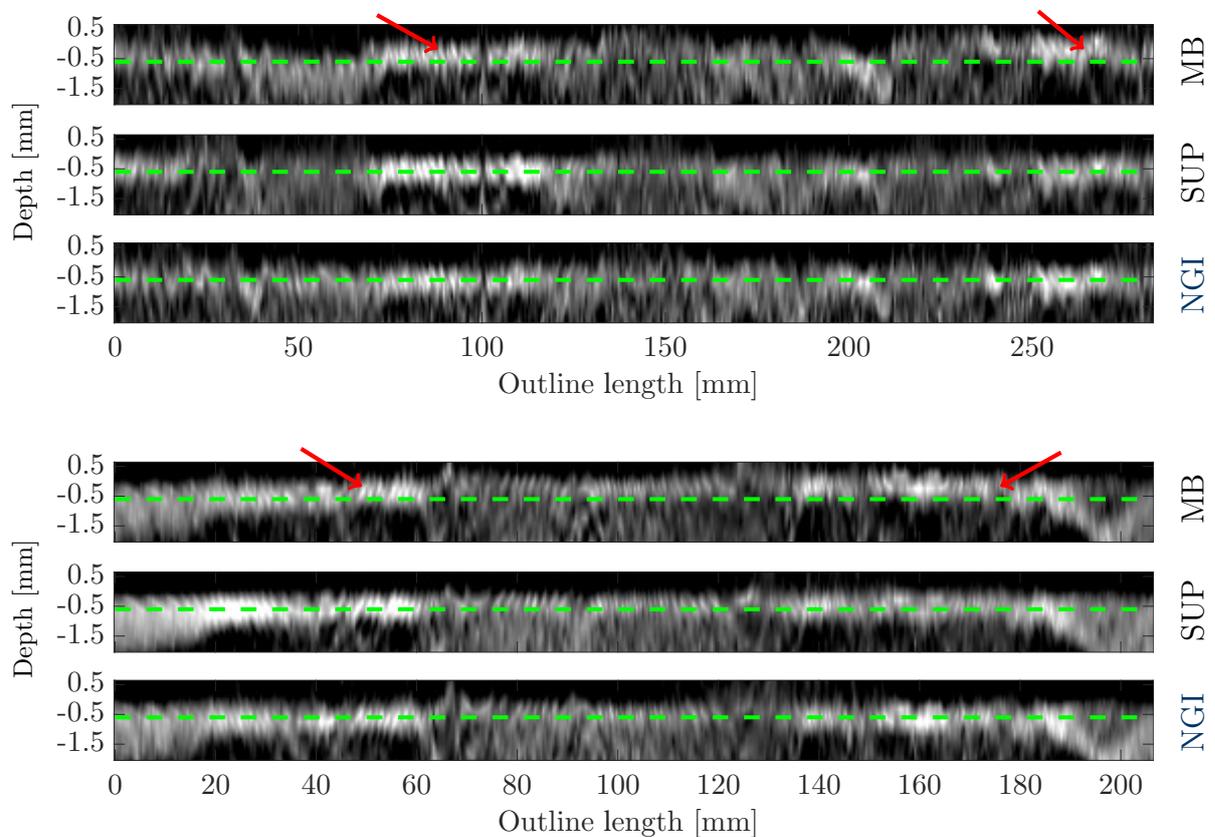
Fig. 8.7(b) shows the line profiles for a selected axial and coronal slice, for **S2\_SUP** and motion-corrected reconstructions of **S2\_WB0** using **MB** and **NGI**. The  $y$ -axis of the profiles points from bone marrow to the surrounding tissue. Starting points and direction of  $x$ -axes are visualized within Fig. 8.7(a). The line profiles are superimposed with a dashed, green line corresponding to a distance of  $-0.5$  mm to the spline surface. The distance was heuristically adjusted such that the line coincides with to the center of the cortical bone edge for the supine scan. Assuming an ideal segmentation, motion correction and registration, the images should show an ideal edge, parallel to the superimposed line. Line profiles of supine scan and **NGI** method appear very similar in both axial and coronal slice, showing a linear edge. In contrast, profiles of the **MB** method appear to shift upwards at several locations along the spline. This translation corresponds to a deformation, orthogonal to the bone surface. If it appears consistently along the  $x$ -axis, it may be interpreted as scaling or distortion effect rather than a misalignment.

### 8.4.2 Fiducial Registration Error (FRE)

Even though **NGI** and **GC** are independent of markers, we can make use of them to evaluate their performance in estimating deformations at marker locations on the skin. This can be done by measuring the **FRE**, i. e., the reprojection error of 3-D marker locations detected in reconstructions based on **NGI** or **GC**, w. r. t. actually detected marker locations in 2-D. Thus, as a first step we need to detect 3-D marker locations in the final reconstructions of the registration-based approaches. We apply the same automatic marker detection introduced for **MB** in Sec. 4.1.2. However, the mapping from 3-D points to detector coordinates now includes the nonrigid **TPS** deformations and is formalized by Eq. (3.45). All other parameters, e. g., for **FRST**



(a) Sampling procedure for line profile measurement.



(b) Edge profiles along femur surface in coronal (bottom) and axial (top) slices.

**Figure 8.7:** Edge profiles along the surface of the right femur in S2\_WB0 and S2\_SUP, for MB, NGI, and the supine reconstruction (SUP). The sampling procedure is shown in (a), where the starting point of the sampling is marked by yellow arrows. The  $y$ -axis points from bone marrow towards surrounding tissue. MB causes a translation of the edge outwards w.r.t. SUP and NGI, which indicates a slight scaling or distortion.

Dataset	FRE [mm]		
	MB	NGI	GC
S2_WB0	0.15	0.37	0.45
S2_WB35	0.35	0.92	1.45
S2_WB60	0.29	0.75	1.55
S3_WB1	1.59	2.08	-
S3_WB2	0.63	1.07	-
S3_WB3	0.23	0.43	-
S3_WB4	0.15	0.35	-
S3_WB5	0.14	0.48	-
S3_WB6	0.16	0.52	-
Total	0.41±0.47	0.77±0.55	-

**Table 8.3:** FRE measured with reprojected marker locations in 2-D projection domain. 3-D marker locations were detected in motion-corrected reconstructions for GC and NGI, using the automatic marker detection approach of MB.

or thresholding, are identical to those used for the MB approach. We ensure that the FREs obtained for the MB and registration methods are based on the same marker detections by using the point correspondences of MB, as determined after the outlier detection.

We define the FRE as mean Euclidean distance of forward projected 3-D points and their corresponding 2-D detections

$$\text{FRE}(\boldsymbol{\alpha}) = \frac{1}{\sum_{k=1}^K (\#U_k)} \sum_{k=1}^K \sum_{m=1}^{\#U_k} \left\| \text{diag}\left(\left(\Delta u \quad \Delta v\right)^\top\right) \cdot \left(h(\underline{\mathbf{m}}_{km}) - \mathbf{u}_{km}\right) \right\|_2, \quad (8.2)$$

where  $\text{diag}\left(\left(\Delta u \quad \Delta v\right)^\top\right)$  performs a scaling from pixel to mm distances. The FRE is evaluated for NGI, GC (for data from Subject 2), but also MB. Note that reported FREs for MB are rather optimistic, as Eq. (8.2) is almost identical to the cost function minimized during the MB approach (cf. Eq. (4.16)). Further, the point correspondences are obtained after outlier removal, thus, the worst contributions of MB are not part of the FRE measurement.

The result of FRE measurements are shown in Tab. 8.3 for MB, NGI, and GC. Due to several misregistrations, GC yielded generally higher FREs than NGI for Subject 2 datasets. The maximum FRE of the GC method amounts to 1.55 mm, whereas all FRE values for the NGI method were no larger than 0.92 mm. Compared to the MB approach the standard deviations for NGI and GC increased substantially. Yet, for the NGI method the highest standard deviation (i. e., 1.07 mm) is still considerably smaller than for the GC case (i. e., 2.46 mm). As anticipated, MB shows the best results for all datasets, with a mean FRE of 0.41 mm and a maximum of 1.59 mm for S3\_WB1. NGI shows an average FRE of 0.77 mm and a maximum of 2.08 mm for S3\_WB1. The general trend of FRE measurements also reflects the amount of motion present in the individual datasets, with the worst FRE results obtained for datasets with severe motion, i. e., S2\_WB35, S2\_WB60, S3\_WB1, and S3\_WB2.

## 8.5 Removal of Marker-Induced Metal Artifacts

In Sec. 4.3 we have shown that metallic markers can cause severe streaking artifacts that may also propagate into diagnostically meaningful regions, e. g., the knee joint. Similar effects can also be seen in reconstructions of Figures 8.3 and 8.5. An approach to reduce such metal artifacts is often a removal of markers in the projection domain. A collection of methods was introduced in Sec. 4.3, allowing extrapolation of data within a certain region around marker locations.

In the following, we present a comparative evaluation of individual extrapolation methods w. r. t. the final reconstruction quality. Investigations are based on a specifically designed numerical phantom introduced in Sec. 7.1. The phantom allows for a qualitative and quantitative evaluation as reconstructions can be obtained, with (MR\_8M) and without (MR\_GT) metallic markers. In addition we present qualitative reconstruction results for marker removal based on a supine scan of Subject 1 (S1\_SUP).

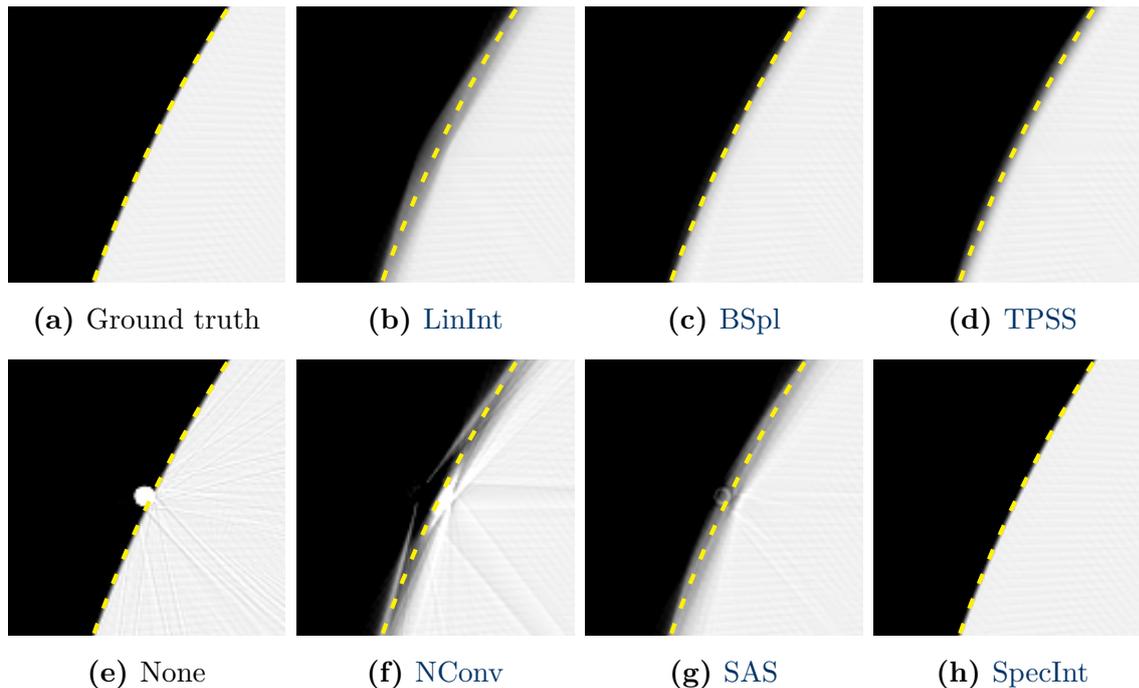
The reconstruction for both numerical phantom and S1\_SUP deviates from that of the weight-bearing scans presented in Sec. 8.1.1. It consists of 1) cosine weighting, 2) Parker redundancy weighting [Park 82], 3) ramp filtering with RamLak kernel [Kak 01], and 4) a GPU backprojector [Sche 07]. We chose to use a standard RamLak filter to make sure that marker-induced streaking is not influenced by the filtering step.

Both phantom and real data contained eight markers. All removal methods, i. e., LinInt, BSpl, TPSS, SAS, NConv, and SpecInt, are applied to each marker location in 496 projection images. Thus, a total of 3968 extrapolations are performed for each method and dataset. We reconstructed volumes for each marker, centered at its 3-D location. The volumes' size was  $256 \times 256 \times 256$  with an isotropic voxel size of 0.125 mm. In addition, we reconstructed MR\_8M, S1\_SUP, and MR\_GT without removal of markers, to obtain-corrupted as well as ground-truth reference volumes. The RMSE was evaluated between ground-truth and corrected reconstructions. All necessary parameters of the individual methods, e. g., the size of the neighborhood  $N_w$  used for extrapolation, were adjusted heuristically using simulated data from MR\_8M. Parameters were then fixed for application to real data S1\_SUP.

### 8.5.1 Results Based on Numeric Phantom

Fig. 8.8 shows squared axial slices of a representative marker showing an ROI with 16 mm side length, centered at each marker location. For convenience, we superimposed the ground-truth surface of the water cylinder by a dashed, yellow line. The best results are obtained for SpecInt which could restore the edge without noticeable difference to the ground truth. Spline-based approaches, i. e., BSpl and TPSS, show a similar performance with slight blurring of the cylinder's edge. Higher amount of blurring is obtained for SAS and LinInt. NConv could not remove the marker and introduced new streaking artifacts.

RMSE values for the individual methods are shown in Tab. 8.4. The first row shows the RMSE over all eight marker volumes and the second row shows the standard deviation of marker specific RMSE values. In agreement with the visual recon-



**Figure 8.8:** Reconstruction results of different marker removal methods. Images depict a squared ROI with 16 mm side length centered at a selected marker. (W: 973 HU, C: -436 HU).

	None	LinInt	BSpl	TPSS	NConv	SAS	SpecInt
RMSE (HU)	30.1	13.8	9.3	9.9	19.5	14.6	<b>8.1</b>
$\sigma_{rmse}$ (HU)	4.4	4.8	4.3	4.4	6.5	4.3	<b>3.3</b>

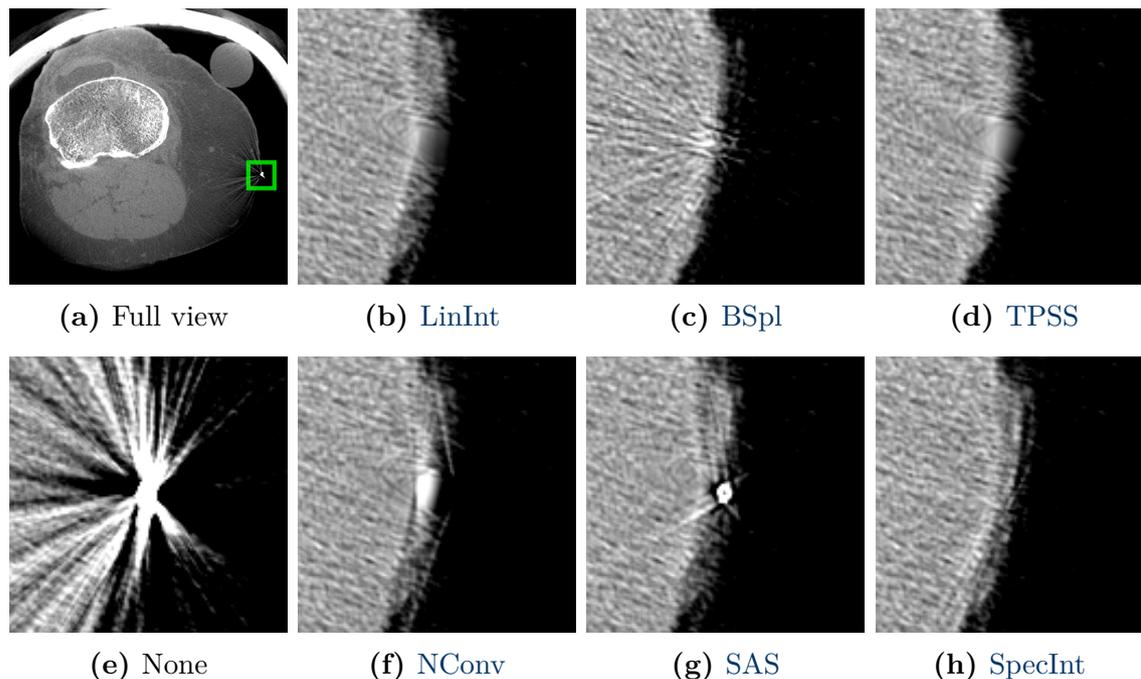
**Table 8.4:** RMSE for extrapolation methods applied to the numeric dataset MR\_8M. The edge builds the transition between soft tissue (0 HU) and air (-1000 HU). Reconstructions without marker removal are abbreviated by (None).

struction results, **SpecInt** performed best with an RMSE of 8.1 HU and the lowest standard deviation of 3.3 HU. **BSpl** and **TPSS** performed only slightly worse, followed by **LinInt** and **SAS**. **NConv** performed substantially worse which is in line with the reconstruction results.

## 8.5.2 Results for Real Data

To allow localization of the marker, we show an axial slice of the left femur in Fig. 8.9(a). The squared ROI with side length 16 mm is superimposed as a solid, green box around the clearly visible marker. Even though the effect of the marker appears quite local in Fig. 8.9(a), it causes streaks that propagate through the entire leg in adjoining axial slices.

Results for marker removal on real acquisitions are shown in Figures 8.9(b) to 8.9(h). In contrast to the numerically simulated data, a substantial amount of noise was present in the reconstructions. Ideally, noise would also be restored at marker loca-



**Figure 8.9:** Results for marker removal based on data from S1\_SUP. We show slices through one of eight markers removed by the extrapolation methods. The location of the ROI can be obtained from (a), showing the full reconstruction of the left leg. Fig. (a): (W: 1000 HU, C: 0 HU). Figures (b) to (h): (W: 973 HU, C: -436 HU).

tions to avoid a blurred appearance after extrapolation. All methods were able to increase image quality compared to a reconstruction without marker removal. *SpecInt* showed the best performance and could successfully remove the marker with proper restoration of edge information and noise properties. In contrast to the numeric simulations, *BSpl* introduced new streaking artifacts, which could originate from the increased noise level. On the other hand, *TPSS* showed increased blurring, such that the result is similar to that of *LinInt*. *SAS* could not fully remove the marker, yet, the skin surface is well restored while retaining the noise level. As for the numerically simulated data, *NConv* showed increased amount of artifacts and performed worst.

## 8.6 Discussion and Conclusion

**Improved Marker-Based Motion Correction** Imaging of knees under weight-bearing conditions poses several problems for image reconstruction in CBCT, first and foremost motion correction is necessary to improve image quality to a level suitable for diagnosis. The baseline of this thesis was a marker-based motion estimation approach [Choi 14c]. In this work a variety of extensions are proposed to the marker-based approach. In contrast to the initial method we eliminated the need for manual annotation of markers and present a fully automatic motion correction approach presented in Chap. 4. Additionally, we present a several methods for improving robustness of MB, including an extension to *FRST* used for marker detection, an

automatic removal of outliers, and a numerical optimization, based on the analytic gradient of the cost function.

Our evaluations revealed that MB is able to accurately correct motion artifacts in most cases. However, for datasets S3\_WB1 and S3\_WB2, which contained severe motion, MB could not accurately restore image quality (cf. Fig. 8.5), such that the contrast agent used for visualization of the cartilage is merely visible (cf. Fig. 8.6(b)). Yet, MB performed best for correcting severe motion in S2\_WB35 and S2\_WB60. By definition MB is limited to a global rigid motion. We believe that this requirement is better complied by scans of Subject 2 as the plasticine wrap enforced similar rigid motion of both legs. At the same time the placement of markers in Subject 3 may have not been ideal, as only six markers covered the leg of interest whereas four were attached to the right leg at the periphery of the FOV. MB aims to preserve both legs with a single rigid transform even if a higher variance of motion between legs is present, eventually leading to artifacts at both legs.

Extending the motion model of the markers to a nonrigid transform, similar to that used for registration-based motion estimation, would be difficult. First, the markers need to be assigned to anatomies, e. g., bones. Second, the number of parameters for a nonrigid problem is much higher than the number of markers, leading to an ill-posed optimization problem. Increasing the number of markers also increases metallic artifacts and makes marker detection more difficult. In addition, MB requires the radiologist to carefully attach the markers to the knee to avoid overlaps in the projection images, which will be bothersome in clinical routine. On top of that, markers can only be attached to the skin, yet, the relationship between motion of bones and skin remains unclear.

**Novel Registration-Based Motion Correction** A motion correction method based on only image features could alleviate most of these drawbacks. In an initial investigation of image-based motion correction we used 2-D/2-D registration of acquired projections and MIPs of reconstructions without correction [Unbe 15]. The method is independent of additional acquisitions or surrogate signals, however, improvement in image quality was insufficient due to the high amount of motion artifacts in the initial reconstruction.

As part of this thesis we propose a novel nonrigid motion correction method based on 2-D/3-D registrations between segmented bones of a motion-free prior scan and the acquired projection images of the weight-bearing acquisitions. Rigid 2-D/3-D registration is applied between each bone and each projection image. The resulting bone motions are then used to extract nonrigid motion fields based on TPS. Supine acquisitions prior to examination are part of all presented weight-bearing studies and are often used to obtain a reference, e. g., to measure cartilage size before application of load.

**Relation to Existing Literature** A multitude of work exists related to 2-D/3-D registration, especially in the field of motion management for radiotherapy applications [Rit 13]. However, most of the approaches are based on the assumption that the motion is periodic and also repeatable within some time frame. Often, a prior 4-D CT acquisition is used to identify a “ground-truth” motion, which is then incorpo-

rated into the 2-D/3-D registration pipeline. In our application we scan knee joints under weight-bearing conditions, where a prior 4-D CT scan is not available in the clinical routine and would require the development of dedicated systems. Furthermore, we cannot rely on a straightforward periodicity or repeatability assumption. The only assumption we incorporate in our approach is that the motion parameters vary smoothly, i.e., the motion has a limited frequency.

Only little work has investigated the registration to a large number of projection images. Ouadah et al. [Ouad 15, Ouad 16] solve for nine parameters per projection to estimate the system’s calibration. With our method we show that an even higher complexity is possible, solving an articulated registration that optimizes six rigid motion parameters per femur and tibia, yielding up to 24 unknowns per projection image. Thus, a total of  $6 \times B \times K$  parameters are estimated, i.e., 5952 for Subject 2 and 2976 for Subject 3.

**Joint Evaluation** Compared to reconstructions without motion correction, all methods were able to improve image quality. Our comparison of similarity measures revealed a consistently better performance for NGI, whereas GC showed increased amount of streaking artifacts. This was due to misregistrations in lateral view directions where overlapping of left and right leg made registration challenging (cf. Fig. 8.4). This finding supports results from Otake et al. [Otak 13], stating that NGI is more robust to occlusions and overlapping edges. We were able to show that NGI performs similarly well than MB with only little differences for most of the datasets investigated. In fact, NGI even outperformed MB corrections for severe motion artifacts in Subject 3, i.e., for S3\_WB1 and S3\_WB2. We believe that the registration-based methods are more flexible when applied to nonrigid joint motion, which is likely to increase according to the severity of patient motion. Of particular interest are the reconstructions of the contrast agent shown in Fig. 8.6, where NGI yielded accurate restorations of bone and contrast surface for all datasets. Robust 2-D/3-D registrations of bones are possible even in presence of contrast agent, which introduced edges similar to those of the femur and tibia. Overall, NGI yielded the best quantitative results with an average UQI of 63.0 compared to 56.3 for MB. NGI had superior UQI values in six out of nine datasets, showing only minor differences to MB in the remaining three datasets.

As expected, MB showed a higher accuracy in correcting the outline of the leg and associated markers. In contrast, NGI and GC introduced slight streaking artifacts at the transition of skin and plasticine of Subject 2. For S2\_WB0 we obtained similar visual results of MB and NGI, yet, the latter yielded a noticeably higher UQI. Extracted intensity profiles at axial and coronal surfaces of the femur showed a small deformation for MB when compared to the supine scan (cf. Fig. 8.7), causing a reduction in UQI. We believe that the deformation originates from the combination of a rigid motion model with nonrigid skin motion. Assuming all but one marker move perfectly rigid, described by a specific translation and rotation and the one marker moves nonrigidly, possibly even slightly opposed to the others. Nevertheless, all markers are assigned the same weight in the cost function introducing bias, as the optimization estimates rigid motion parameters only. The error introduced in this way might manifest in increased distances for all the other rigid markers that could

normally be estimated perfectly. These increased distances can lead to a distortion in the reconstruction, which will be connected to the nonrigid motion of a single marker.

Fiducial markers are frequently used to provide ground-truth motion estimates in the field of 2-D/3-D registration [Pawi11]. We follow this principle for evaluation of how well the TPS are able to estimate motion at the boundary of the legs. An analysis of the FRE showed the smallest reprojection error for MB. This was not surprising, as its cost function minimizes exactly this error, whereas GC and NGI are independent of marker locations. However, an average FRE of 0.77 mm for NGI is still acceptable. Considering that the measure is defined in the projection image plane its effect on the reconstruction reduced approximately by the magnification factor of the CBCT system, i. e., by  $\frac{S}{F}$ . For Subject 2 and Subject 3, this would lead to an effective reconstruction error of 0.50 mm.

A limitation of our quantitative evaluation is its dependency on subsequent 3-D/3-D registrations of individual bones, which introduces an additional source of error. Yet, avoiding the intermediate registration steps would require knowledge about the ground-truth motion of bones, which is hardly possible in in-vivo acquisitions. We use the UQI as measure for image quality of bones. Even though a small margin of soft tissue is included, the evaluation shall not be interpreted as an accurate measure for motion correction of soft tissue. However, considering that the majority of diagnostically meaningful anatomies are located close to the bones, e. g., tibial and femoral cartilage, the UQI presents reasonable means for comparing motion correction approaches.

We introduced a method to reduce noise in generated DRR images, based on bilateral filtering and an adjustment of bone segmentations to focus only on cortical bone. In essence, this is very similar to a direct registration of 3-D surface meshes to the projection images, also known as feature-based 2-D/3-D registration [Mark12]. Even though, feature-based methods may reduce computational complexity of the registration approach, we expect a superior performance of the presented method as it is more robust to deviations in the bone segmentations. This is because it does not only rely on geometric, but also on actual intensities of the supine reconstructions. In addition, mesh-based registration requires detection of features in projection domain, which will be difficult for lateral view directions.

Currently, registration-based approaches require a great amount of manual input, i. e., semi-manual segmentations of bones and manual 3-D/3-D registrations. Also, computation times of the multiresolution 2-D/3-D registration are rather high. For  $B = 4$  bones,  $K = 248$  projection images (cf. Subject 2) and a high number of maximum iterations for registration (300), registration can take up to several hours on a consumer grade computer, which is still beyond clinical applicability. Yet, we believe that further improvements to the hardware as well as a code optimization could reduce runtime by orders of magnitude. In contrast, due to extensions and modifications presented in this thesis, MB reconstruction is ready for a clinical application. No manual interaction is required and the typical runtime for the entire motion-corrected reconstruction lies in the range of several minutes.

**Metal Artifact Reduction** A method for reducing metallic artifacts caused by fiducial markers is proposed in Sec. 4.3. After extraction of a full set of detected

marker locations in projection domain, the markers are blanked out with a certain radius  $r_w$  and missing data is extrapolated based on the surrounding intensities. Our comparison of different extrapolation methods shows the best results for `SpecInt`, restoring the edge information but also the noise level. Important for `SpecInt` is the neighborhood's size  $N_w$  and its location. In case of fiducial markers, defective pixels are well localized and all of similar size, which facilitates `SpecInt`. Spline-based methods showed a good performance on data from a noise-free, numerical phantom. Yet, when applied to real data `BSpl` led to residual streaking artifacts. `BSpl` extrapolates missing values based on 1-D vertical and horizontal splines, thus, neighboring defect pixels may rely on entirely different intensities, leading to higher extrapolation errors in presence of noise. In contrast, `TPSS` estimates a regularization parameter which increased in presence of noise, thus, causing a blurred reconstruction at the skin. Straightforward `LinInt` removes the markers reliably, yet, shows rather smooth results in reconstructions of the skin surface. The goal of `SAS` is to retain high-frequency edge information covered by the object that is to be removed. It was initially applied to remove catheter tips from projection images [Schw 10]. The approach could not remove markers entirely which may be due to the high frequencies incorporated by the markers itself. The worst results were obtained for `NConv` which seems not appropriate for marker removal.

In case an accurate restoration of the skin surface is important and the noise level of the neighborhood should also be restored, we recommend to use `SpecInt`. Yet, if only streaking artifacts should be removed we deem `LinInt` as sufficient, which has the benefit of a substantially lower runtime.

## Conclusion

Within the scope of this thesis we introduce several extensions and improvements to marker-based motion correction. As a result, a fully automatic motion correction method for acquisitions of knees under weight-bearing conditions is presented. Moreover, the method is readily applicable to other anatomies, which we showed for correction of head motion [Mull 15a]. Results show that the method is able to restore structural information, e.g., bone or cartilage surfaces, for most of the evaluated datasets. However, deficiencies were observed in case of nonrigid or severe motion as the method is currently limited to rigid motion and hardly extendable to nonrigid motion correction. A key problem is that the markers require a thorough attachment prior to examination, which hinders clinical routine and increases patient discomfort. On top of that, we found that MB can cause small distortions of the bone, indicating that internal motion may not be estimated accurately.

To alleviate these issues, we present a novel method for motion correction using 2-D/3-D registration of bones, segmented from prior supine scanning, to each projection image acquired during weight-bearing acquisitions. A comparison between MB and registration-based motion correction revealed the highest improvements in image quality for 2-D/3-D registrations using `NGI` as similarity measure, where robust registration results are obtained despite overlapping of bones. Evaluations on Subject 3, obtained in context of the cartilage study (cf. Sec. 2.3.2), showed that `NGI` is robust to the presence of contrast agent. Reconstructions of sagittal slices show a superior

quality of the contrast agent's surface for *NGI*, allowing a subsequent analysis of the cartilage.

Given that acquisition time and patient comfort is of less importance and rigid motion can be ensured, we recommend the proposed marker-based correction method in combination with a subsequent marker removal using *SpecInt*. In contrast to a registration-based correction, MB is fully automatic and readily applicable in clinical routine. However, many clinical applications do not allow for an additional interference with the acquisition protocol. Also, enforcing only rigid motion, e. g., using a fixation of legs, is hardly feasible and may be in contradiction to the clinical motivation. In that case, we have shown that the registration-based method offers a promising and in some cases even more accurate alternative to MB correction. Further optimization reducing the runtime and required manual interaction for 2-D/3-D registrations, could allow a complete replacement of marker-based motion correction in future applications.

# Evaluation of FCC-Based Motion Correction

9.1	Materials and Methods .....	123
9.2	Initial Evaluation for Fan-Beam CT .....	126
9.3	FCC-based Motion Correction for CBCT .....	129
9.4	Results for Simulated Knee Phantom .....	138
9.5	Discussion and Conclusion .....	141

Applications of **CC** in the field of **CBCT** are still a very active field of research and often limited to theoretic concepts. The state-of-the-art of **FCC**-based artifact correction was based on fan-beam **CT**, without truncation and using full-scan acquisitions. In addition, **FCC** have not yet been applied in the field of motion estimation. As part of this thesis we present a novel approach for motion correction based on **FCC**. It requires only the acquired, motion-corrupted projection images. In this chapter we show initial evaluation results of the proposed methods. First, a detailed explanation of performed evaluations and selected parameters is given in Sec. 9.1. We initially introduced the method for fan-beam **CT** and extended it to **CBCT** in the progress of this thesis. Similarly, results are first presented for fan-beam **CT** in Sec. 9.2. The extended method for **CBCT** is evaluated in Sec. 9.3. In both cases, evaluations are based on a challenging numerically simulated head phantom, yielding truncation-free projections over a full-scan range, containing rigid translational motion. To fill the gap to weight-bearing imaging of the knee, results of a numerical knee phantom, including vertical truncation and realistic motion, are shown in Sec. 9.4. Finally, the chapter ends with a discussion and conclusion in Sec. 9.5.

## 9.1 Materials and Methods

Evaluations on fan- and cone-beam geometries are performed in a similar way. The goal is to validate the method's performance w. r. t. improvements in image quality. Both methods use generated phantom data, thus, ground-truth reconstructions are easily obtained. Datasets and simulated translational motion have been explained in Sec. 7.1.2. All simulations were performed with help of the open source software **CONRAD** [Maie 13].

### 9.1.1 Fan-Beam Study

For the fan-beam case we have used four datasets, i. e., FFB\_GT\_HQ, FFB\_OSC\_HQ, FFB\_GT\_LQ, and FFB\_OSC\_LQ. They consisted of high- and low-quality data, where the latter showed severe Poisson distributed noise, a lower X-ray energy, half of the detector resolution, and about a quarter of the angular resolution. Only the *Oscil* motion was used for fan-beam simulations. A reference in reconstruction domain was obtained by direct rendering of the phantom, allowing an evaluation of the maximum achievable image quality in case no motion is present. The motion model was restricted to 1-D translations of individual detector lines.

Optimization of the cost function involved a multiresolution scheme. The fan-beam sinograms were downsized by a scaling factor of 0.25, 0.5 and 1, where the resulting motion was upsampled by linear interpolation such that it can be provided as initialization to the next higher resolution level. To obtain a quantitative measure of image quality, we computed the **rRMSE** on the full reconstructed 2-D image, given motion-corrupted, corrected and motion-free reconstructions.

### 9.1.2 Cone-Beam Study

For the cone-beam evaluation we have used a total of 12 datasets of the same 3-D FORBILD phantom. Six datasets were of low- and six of high-quality, including five different motion patterns. The difference between high- and low-quality is the simulation of Poisson noise, while all other geometric and physical parameters remained constant (cf. Sec. 7.1.2). Reconstruction of motion-free projection images are used as ground truth.

Our evaluations are based on 3-D translations of the object. However, the evaluation focuses on the efficient implementation using 2-D detector shifts as motion model (cf. Sec. 6.3). Thus, the motion model cannot perfectly reflect object motion towards and away from the detector, reducing image quality even if the motion is estimated perfectly. To allow for a separation of artifacts induced by limitations of the motion model and residual motion artifacts, we perform three types of reconstructions. A lower bound in image quality is reflected by reconstructions without motion correction. The upper bound of the translational motion model can be evaluated by using the ground-truth detector motion for image reconstruction. Extraction of the ground-truth translations in projection domain is described by Eq. (9.2).

For a quantitative evaluation of reconstruction volumes we compute the **rRMSE** and **SSIM** introduced in Sec. 7.3. The measures are computed on the entire volumes, where a reconstruction of motion-free projection images is used as reference. The **SSIM** has been evaluated block-wise as described by the second evaluation method in Sec. 7.3.2.

## Evaluation of Motion Parameters

As our study is based on numerically simulated data we had access to the ground-truth motion, thus, we conducted a quantitative analysis of estimated motion parameters. The **mean absolute distance (MAD)** and the standard deviation of absolute

estimation errors (SD) is computed between ground-truth and estimated translations. They can be expressed by

$$\begin{aligned} \text{MAD}(\boldsymbol{\alpha}) &= \frac{1}{K} \sum_{k=1}^K \begin{pmatrix} |s_k - \hat{s}_k| \\ |t_k - \hat{t}_k| \end{pmatrix} \\ \text{SD}(\boldsymbol{\alpha}) &= \frac{1}{K-1} \sum_{k=1}^K \begin{pmatrix} (s_k - \text{MAD}(\boldsymbol{\alpha})_1)^2 \\ (t_k - \text{MAD}(\boldsymbol{\alpha})_2)^2 \end{pmatrix} . \end{aligned} \quad (9.1)$$

where  $\hat{s}_k$  and  $\hat{t}_k$  are the extracted ground-truth translations in projection domain. They are built by subtracting the projection of simulated 3-D translations  $\mathbf{t}_k = (t_k^x \ t_k^y \ t_k^z)^\top$  (cf. Sec. 7.1.2) by the projection of the coordinate center. The projected 2-D points are easily obtained using the motion-free projection matrices  $\mathbf{P}_k$ . Thus, ground-truth detector translations may be obtained by

$$\begin{pmatrix} \hat{s}_k \\ \hat{t}_k \end{pmatrix} = \begin{pmatrix} \Delta u & 0 \\ 0 & \Delta u \end{pmatrix} \left( h(\mathbf{P}_k \cdot \underline{\mathbf{0}}) - h(\mathbf{P}_k \cdot \underline{\mathbf{t}}_k) \right) , \quad (9.2)$$

where  $\mathbf{0} \in \mathbb{R}^3$  is a zero vector and  $\underline{\mathbf{0}} = (\mathbf{0}^\top \ 1)^\top$  its homogeneous form. Further,  $\Delta u$  and  $\Delta v$  are used for a conversion from pixel to mm units. Note that translations of the detector are more suitable for a correction of motion that occurs parallel to the detector. However, 3-D translations may also have a component which moves the object orthogonal to the detector, i. e., in depth direction. This leads to scaling of the object in the projection images and deviations in the line integral values. By restricting our motion model to detector shifts we neglect the change of object size. Nevertheless, scaling will have certain effects on estimated translations  $s$  and  $t$  but also on the cost function value. In consequence, scaling may introduce an additional source of error which is not directly related to FCC but to limitations of the motion model.

### 9.1.3 Parameter Selection And Image Reconstruction

For both fan- and cone-beam geometry we require an estimate for the maximum object extent  $r_p$  to compute the zero regions. We propose a method to extract  $r_p$  from the acquired motion-corrupted projection data. Only one motion was simulated for fan-beam evaluations, whereas, five motion patterns were included for the cone-beam study. Thus, we decided to use the dataset showing the largest translational motion in detector  $u$  direction, which is present in both fan- and cone-beam CT. First, the projection data is summed over rotation angles  $\lambda$  and over the detector  $v$ -axis, yielding a vector of dimension  $I$ . We can now use a plot of the resulting vector to measure the spatial distance of left and right boundary w. r. t. the center in mm, denoted by  $u_{\text{left}}$  and  $u_{\text{right}}$ , respectively. The maximum of both values is now scaled by the magnification factor of the system such that it refers to the maximum FOV radius of the uncorrected reconstruction. Thus, the estimated object extent can be obtained by  $r_p = \max(u_{\text{left}}, u_{\text{right}}) \frac{S}{S+D}$ . This yielded an estimated extend of  $r_p = 122.5$  and  $r_p = 127.8$  for fan- and cone-beam evaluations, respectively.

We show that the cost function is invariant to constant translations, and propose a regularizer to fix the translations of the first projection image to a certain value

(cf. Sec. 6.4). To allow accurate quantitative comparison we adjusted them with the ground-truth motion, i. e.,  $\tilde{s} = \hat{s}_k$  and  $\tilde{t} = \hat{t}_k$ . This approach was not applied for fan-beam results, yet, we did not notice unusual global shifts.

After the first evaluation of  $e(\boldsymbol{\alpha})$ , the cost function value was normalized to 100. Further, the Lagrangian parameter was heuristically determined to  $\tilde{\eta} = 1$ . Optimization was either terminated by the optimizer’s default settings for a minimum step size and change of cost function values, or after a maximum of 200 iterations.

For the fan-beam experiments we manipulated the mask of the zero regions by 1-D erosion along frequency axis  $\xi$ . This procedure has been replaced for applications to CBCT using the mask adjustment described in Sec. 6.5.1. The parameter for the distance of ideal discretization lines and actually used triangle boundaries was heuristically determined and set to  $\epsilon = 3 \times 10^{-3}$  for all experiments.

### Reconstruction Pipeline

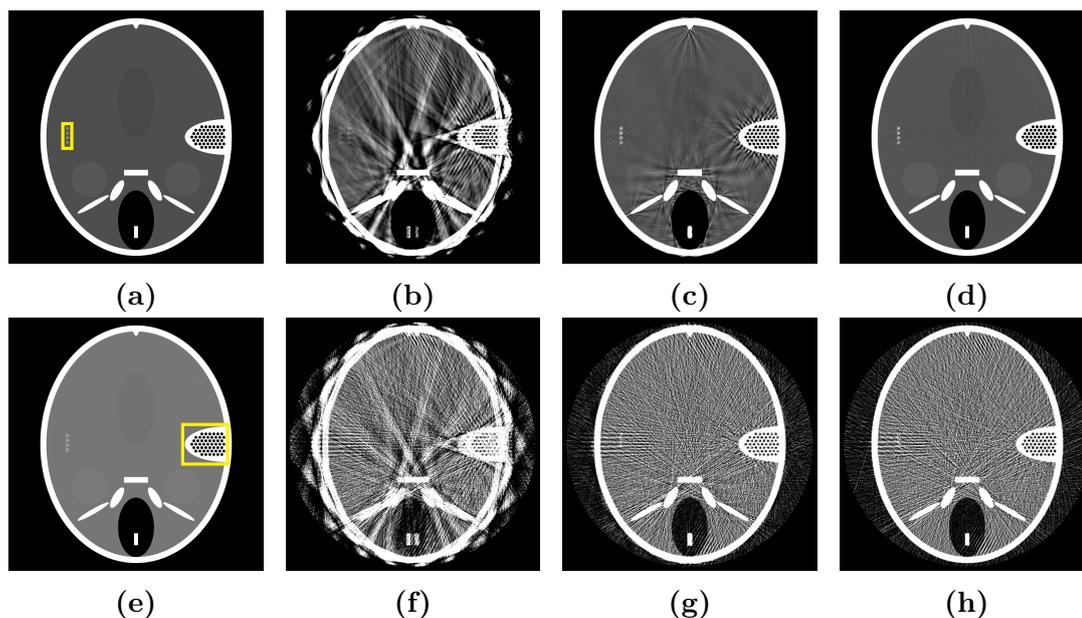
We had access to four different types of reconstructions for each motion type. First, the ground-truth datasets (FCB\_GT\_HQ, FCB\_GT\_LQ, FFB\_GT\_HQ, FFB\_GT\_LQ) have been reconstructed, followed by two reconstructions with and without motion correction, but also using motion correction based on the ground-truth motion data. We also applied motion correction to the ground-truth data to verify stability of the estimation approach. The fan-beam study did not involve ground-truth motion estimates, instead the phantom was directly rendered in volume domain, without an intermediate reconstruction. Estimated motion was incorporated into the projection matrices prior to reconstruction, using Eq. (6.24).

For all reconstructions the following pipeline was used: (1) cosine weighting, (2) ramp-filtering with a smooth Shepp-Logan filter with roll-off [Kak01], (3) a GPU-based backprojection [Sche07], and (4) negative intensities are clamped to zero. Note that Parker weighting was not necessary as we simulated full-scans.

Reconstructed volume sizes were adjusted as follows. For the fan-beam case we reconstructed 2-D images of size  $2048 \times 2048$  with an isotropic pixel size of 0.125 mm. Detailed views of resolution pattern and ear regions are based on the same size and resolution. Cone-beam reconstructions were generated on a voxel grid of size  $512 \times 512 \times 512$  with an isotropic voxel size of 0.5 mm.

## 9.2 Initial Evaluation for Fan-Beam CT

In Figures 9.1(a) and 9.1(e) we show a direct rendering of the FORBILD phantom for low- and high-quality datasets. Note that intensity levels may differ due to different monochromatic X-ray energies. Reconstructions for datasets FFB\_OSC\_HQ and FFB\_OSC\_LQ are shown with and without motion correction in Figures 9.1(b) and 9.1(c) and Figures 9.1(f) and 9.1(g), respectively. Further, Figures 9.1(d) and 9.1(h) show the corresponding reconstruction of ground-truth datasets FFB\_GT\_HQ and FFB\_GT\_LQ. For both low- and high-quality datasets, the proposed method could substantially improve image quality. Compared to the ground-truth and reference reconstructions, residual streaking can be noticed at locations with high intensity variations, e. g., for bones. However, for the low-quality data the high noise level



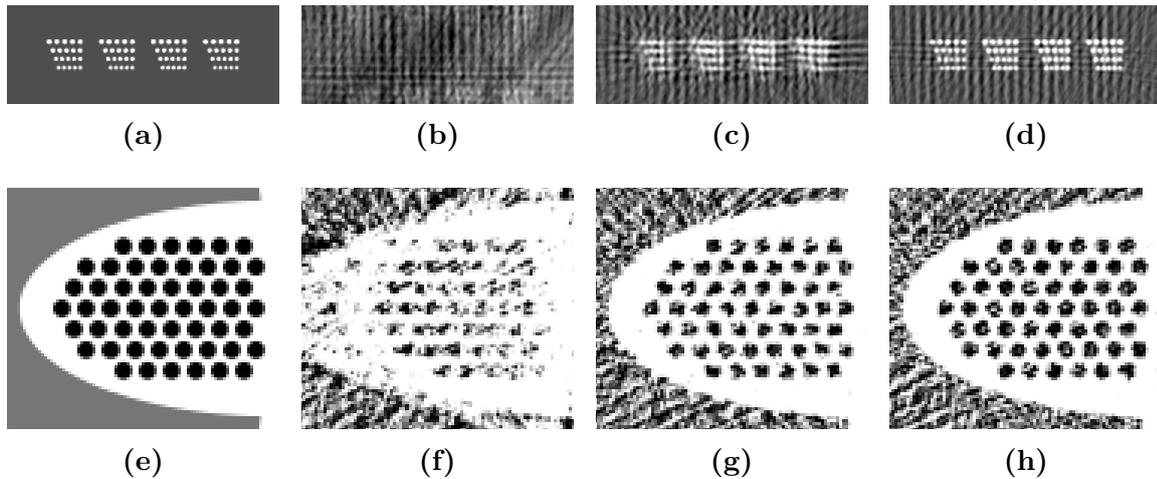
**Figure 9.1:** Reconstructed images for high- (top row) and low-quality (bottom row) fan-beam data. From left to right: Ground-truth rendering of phantom, reconstruction without correction, the proposed method and reconstructions of motion-free data. Top: (W: 697 HU, C: 105 HU). Bottom: (W: 2090 HU, C: 105 HU).

dominates image quality, thus, the reference and corrected reconstruction in Figures 9.1(g) and 9.1(h) appear almost identical.

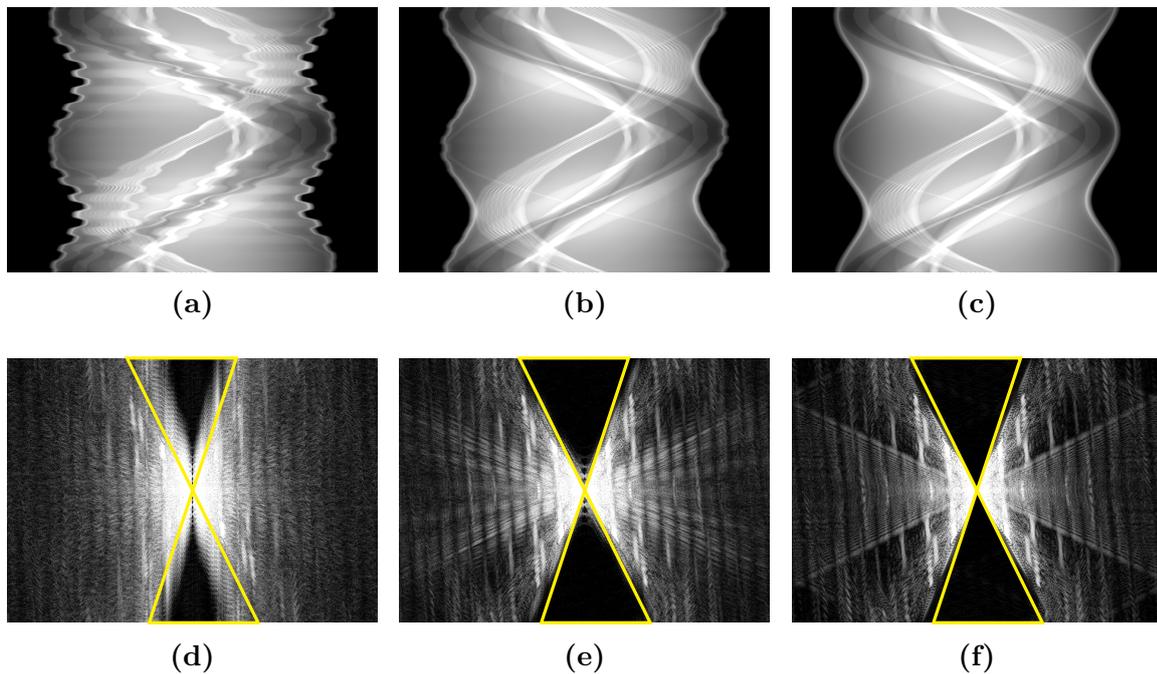
A detailed view of the resolution pattern is shown in the top row of Fig. 9.2 for the high-quality datasets. Fig. 9.2(b) shows that the pattern is no longer visible without motion correction, whereas the shape of the pattern was well restored after correction (cf. Fig. 9.2(c)). Yet, when compared to the ground truth in Figures 9.2(a) and 9.2(d), the corrected reconstruction shows residual streaking and slight blurring, which does not allow proper separation of individual beads.

The bottom row of Fig. 9.2 depicts zoomed reconstruction of the phantom’s ear model. The images are taken from reconstructions of the low-quality simulations, which is clearly visible by the high amount of noise. Note that the noise is also present in the reference reconstruction shown in Fig. 9.2(h). The sharp transitions between air bubbles and bone material are well restored after correction (cf. Fig. 9.2(g)), yielding comparable results than the reference reconstruction depicted in Fig. 9.2(h). Both, show a substantial amount of noise, when compared to the ground-truth rendering in Fig. 9.2(e).

Fig. 9.3 shows sinograms and the logarithmically scaled absolute value of their Fourier domain, for high-quality data. For convenience we have superimposed the boundaries of the triangular regions in the spectra. Estimated translations could clearly improve the sinogram as shown in Fig. 9.3(b). The corresponding triangular zero regions in Fig. 9.3(e) are well restored. Besides some residual ripples, visible at the periphery of the corrected sinogram, it appears very similar to the motion-free sinogram shown in Fig. 9.3(c). The deviations originate from object motion in depth direction and are a limitation of the motion model.



**Figure 9.2:** Top row: Detailed reconstructions of the phantom's resolution pattern for high-quality data. Bottom row: Reconstructions of the ear for low-quality data. Both regions are indicated in Figures 9.1(a) and 9.1(e). The arrangement of columns is identical to Fig. 9.1. (W: 697 HU, C: 105 HU).



**Figure 9.3:** Sinograms (top row) and their logarithmically scaled spectra (bottom row) before and after correction, and without motion. We can see a clear improvement of the sinogram after correction (cf. (b)), supported by a well restored zero energy regions in frequency domain (cf. (e)). Visualization windows of sinograms and spectra were  $[0.15, 0.5]$  and  $[2.0, 5.0]$ , respectively.

rRMSE %	NoCorr	Corrected	Reference
High-quality	20.4	7.1	2.5
Low-quality	25.1	14.0	12.6

**Table 9.1:** rRMSE for high- and low-quality reconstructions w. r. t. the rendered ground-truth phantom.

Improvements in image quality are well reflected by our quantitative analysis of rRMSE shown in Tab. 9.1. We notice a reduction in rRMSE of 13.3% and 11.1% for the high- and low-quality datasets, when compared to reconstructions without motion correction. The rRMSE after correction is at 7.1%, with a distance of 4.6% to a motion-free reconstruction. For the low-quality data, our correction yielded an rRMSE of 14%, twice as much as for the high-quality data. However, the distance to the motion-free reference is only 1.4%, indicating that the high residual rRMSE value is almost entirely caused by the noise level of the low-quality dataset.

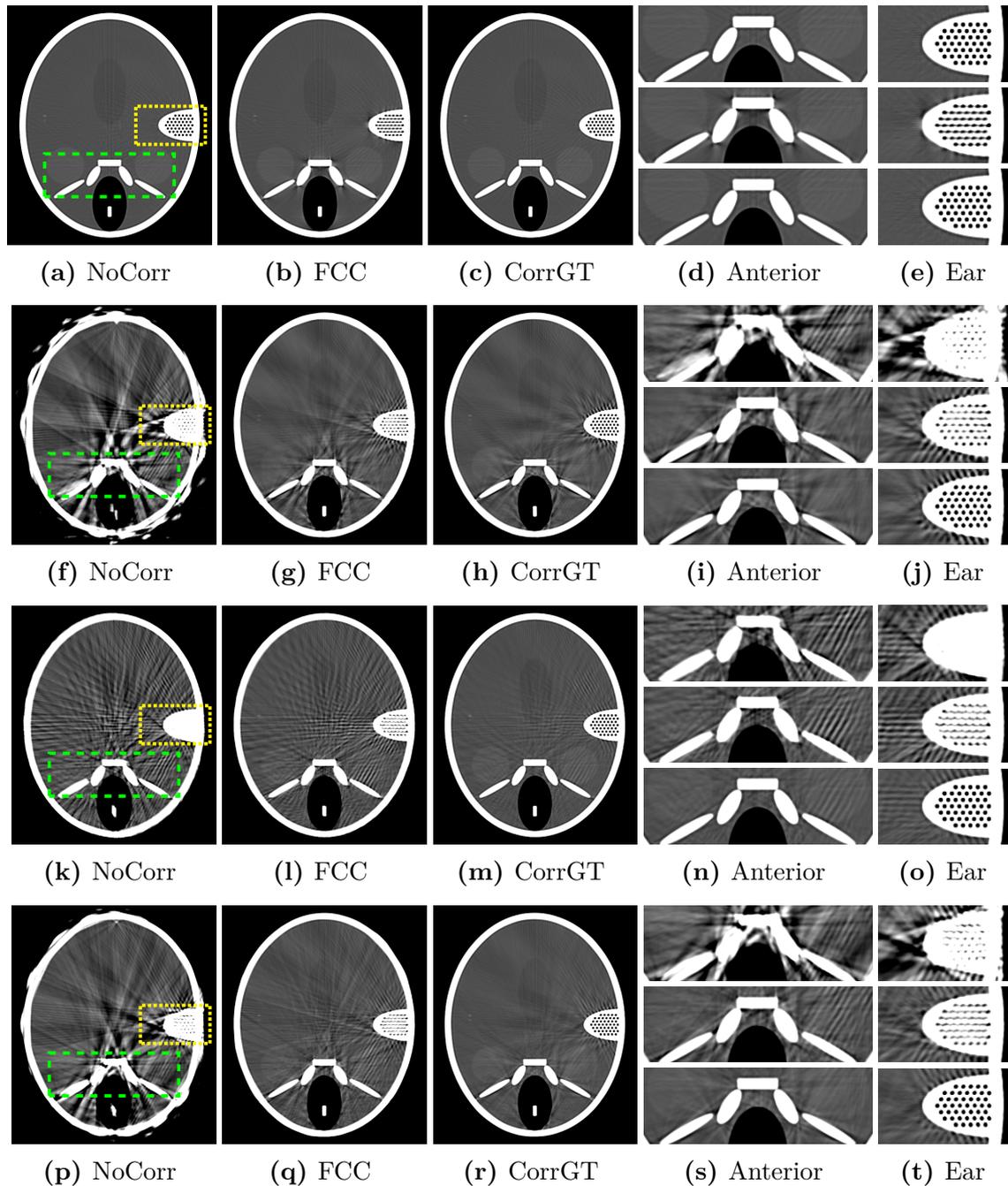
## 9.3 FCC-based Motion Correction for CBCT

### 9.3.1 Reconstruction Results

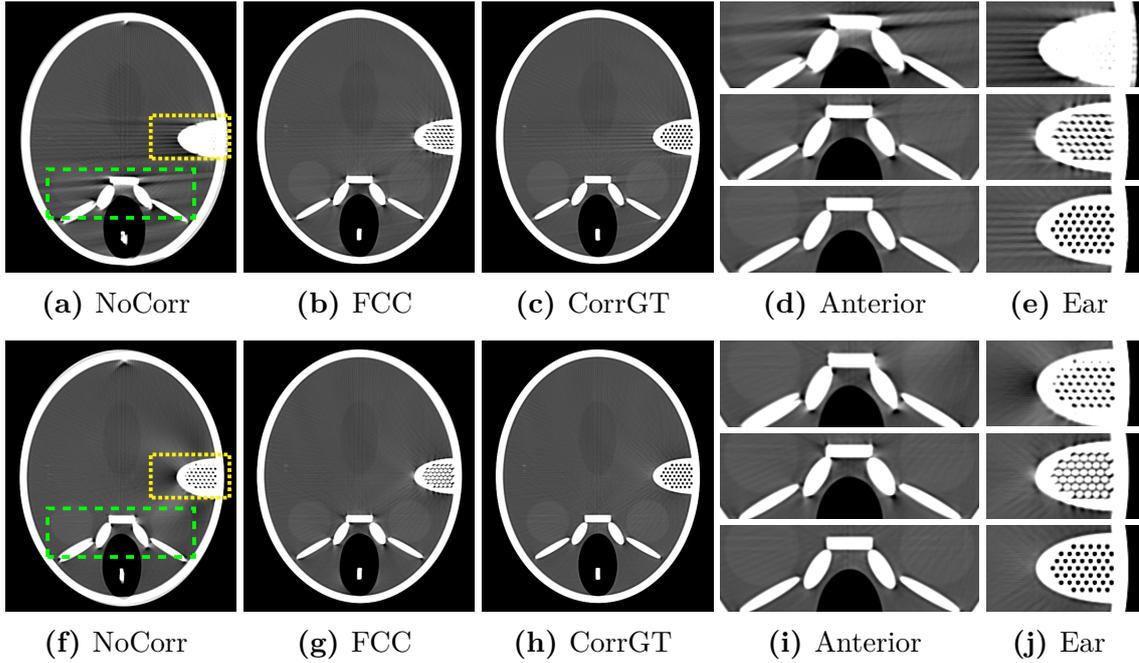
In the following we present axial reconstructions of the individual datasets. To guarantee a reasonable size of individual images we separate between high- and low-frequency motion types. Further, only one representative example is shown for the low-quality dataset as results were generally similar to those obtained with noise-free data. To show applicability to CBCT an axial slice is selected, superior to the central slice, positioned at a distance of 7.25 mm to the coordinate center.

Fig. 9.4 contains reconstructions for the high-frequency motion types, i. e., *Oscil*, *Chirp*, and *Rect*, from the second to the fourth row, respectively. In addition, results for the motion-free ground-truth data are shown in the first row, abbreviated in the following with *None*. The columns correspond to reconstructions without correction (NoCorr), after correction using FCC and after correction using the ground-truth translations (CorrGT). Further, the fourth and fifth column depict details of the phantom’s anterior and ear, corresponding to reconstructions based on NoCorr, FCC, and CorrGT, arranged from top to bottom. The detailed regions are superimposed on uncorrected reconstructions in Figures 9.4(a), 9.4(f), 9.4(k) and 9.4(p), using dashed, green and dotted, yellow boxes. The anterior region was adjusted to show high contrast bones but also two circular low-contrast structures representing the eyeballs. The ear is particularly challenging having multiple high intensity variations between air and bone.

Note that Figures 9.4(a) and 9.4(c) and their corresponding detailed slices are identical, both representing the ground-truth reference reconstruction later used for quantitative analysis. In case of *Oscil* and *Rect* motion a high amount of motion artifacts are present including streaking and blurring (cf. Figures 9.4(f) and 9.4(p)). For the *Chirp* motion we observe mainly streaking artifacts, as it contained high-frequency variations but only a small motion amplitude (cf. Fig. 9.4(k)). FCC could



**Figure 9.4:** Axial slices superior to the coordinate center for high-frequency motion types and motion-free data. The rows show motion-free (*None*) and high-frequency motions (*Oscil*, *Chirp*, and *Rect*). From left to right we show slices without motion correction (NoCorr), after correction using FCC and corrected using the ground-truth translations (CorrGT). The last columns depict details of reconstructions located at an anterior and ear region. (W: 697 HU, C: 105 HU).

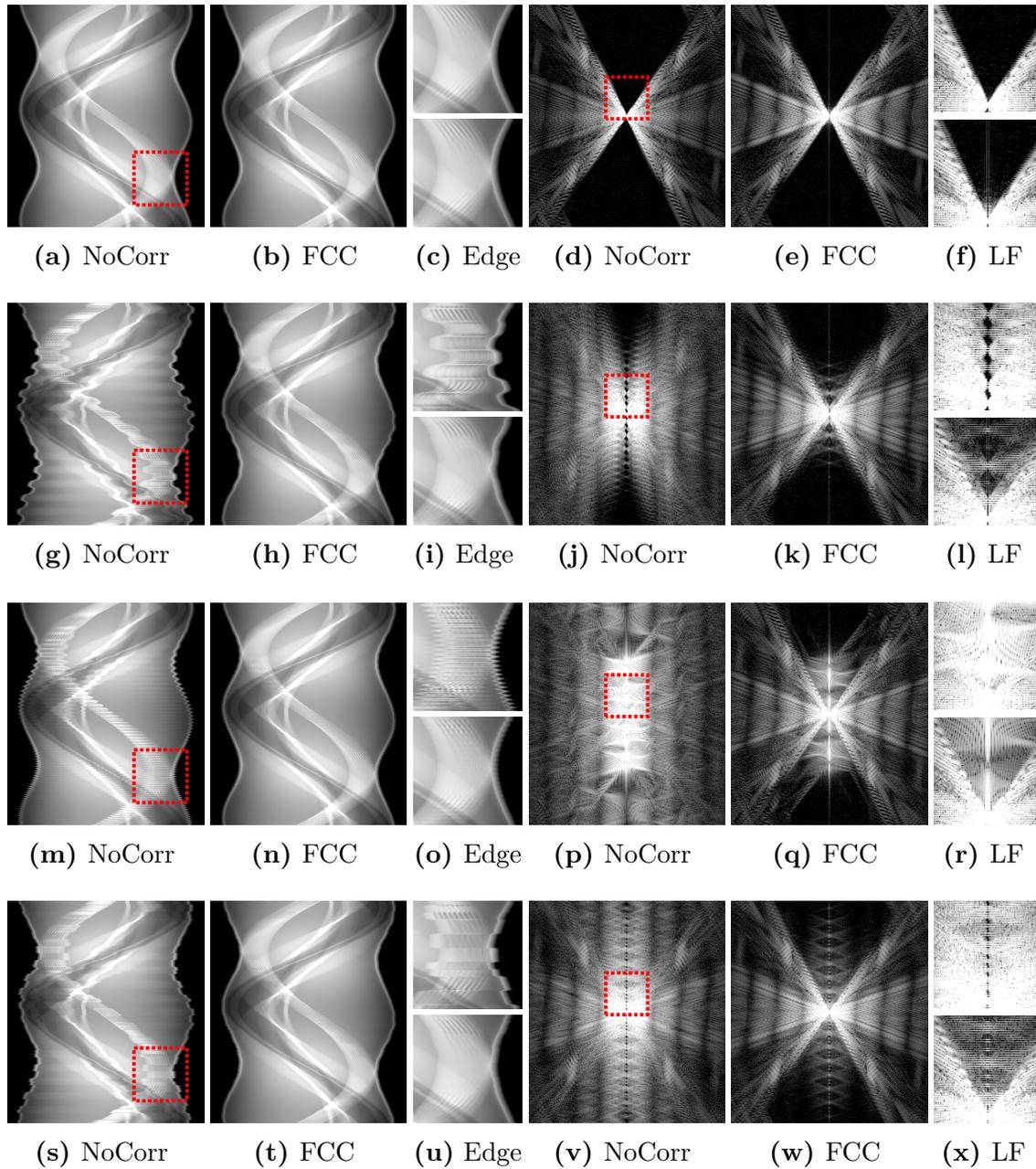


**Figure 9.5:** Off-center axial slices for low-frequency motion types ( $LF-1$  and  $LF-2$ ). Arrangement identical to Fig. 9.4. (W: 697 HU, C: 105 HU).

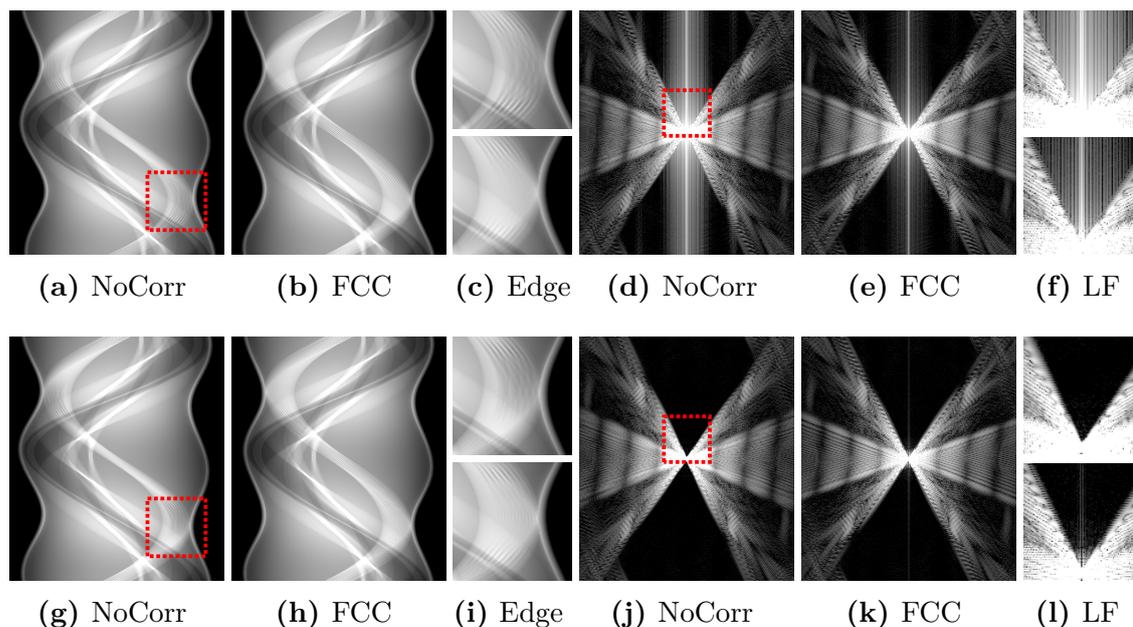
improve image quality for all motion types, substantially reducing streaking artifacts and showing a good restoration of bone structures w.r.t. NoCorr. Whereas the air bubbles of the ear were hardly visible in the motion-corrupted reconstructions, they can be identified after correction with FCC. Nevertheless, we still notice a slightly higher image quality for CorrGT, indicating that a more accurate motion estimation could further reduce motion artifacts. To evaluate stability and convergence behavior of the algorithm, FCC was also applied to the ideal, motion-free data FCB\_GT\_HQ. Reconstructions in Fig. 9.4(b) show that the method introduced minor blurring at bone structures which reduced image quality and indicates a misestimation of motion parameters.

Fig. 9.5 shows the reconstruction results for low-frequency motions  $LF-1$  and  $LF-2$ . The arrangement of individual images is identical to Fig. 9.4. Reconstruction results without motion correction are dominated by blurring. Streaking is only present for  $LF-1$ , due to the difference in starting and end position, which leads to a discontinuity between first and last projection image (cf. Fig. 9.5(a)). In contrast, Fig. 9.5(f) shows a low-frequency intensity bias mainly visible in the periphery of the ear. Streaking artifacts could be well removed for  $LF-1$  after application of our method and are comparable to those visible in CorrGT. Similarly, the intensity bias could be corrected for  $LF-2$  and also the displacement of the lateral anterior bone is well restored. Yet, for corrections of both  $LF-1$  and  $LF-2$ , we notice a clear difference to the ground-truth reference reconstructions, especially in reconstruction quality of the ear in Figures 9.5(e) and 9.5(j).

The left columns of Fig. 9.6 show the sinogram domain for a vertical detector line at  $v = 18.6$  mm. In addition, the absolute value of the Fourier domain's central slice, i. e., for a vertical frequency of  $\psi = 0$  mm<sup>-1</sup>, is shown on the right columns. We use a



**Figure 9.6:** Left: Sinogram representation for a detector line at  $v = 18.6$  mm. Right: Logarithmically scaled spectrum for a vertical frequency of  $\psi = 0$  mm<sup>-1</sup>. Arrangement of rows is identical to Fig. 9.4. Left columns: Sinograms before correction (NoCorr) and corrected by FCC, followed by details of the sinograms' outline (Edge). Right columns: Fourier domain of sinograms before and after correction and a zoom to the low-frequency part of the triangular region (LF). Visualization windows of sinograms and spectra were  $[0.15, 0.5]$  and  $[5, 10]$ , respectively.

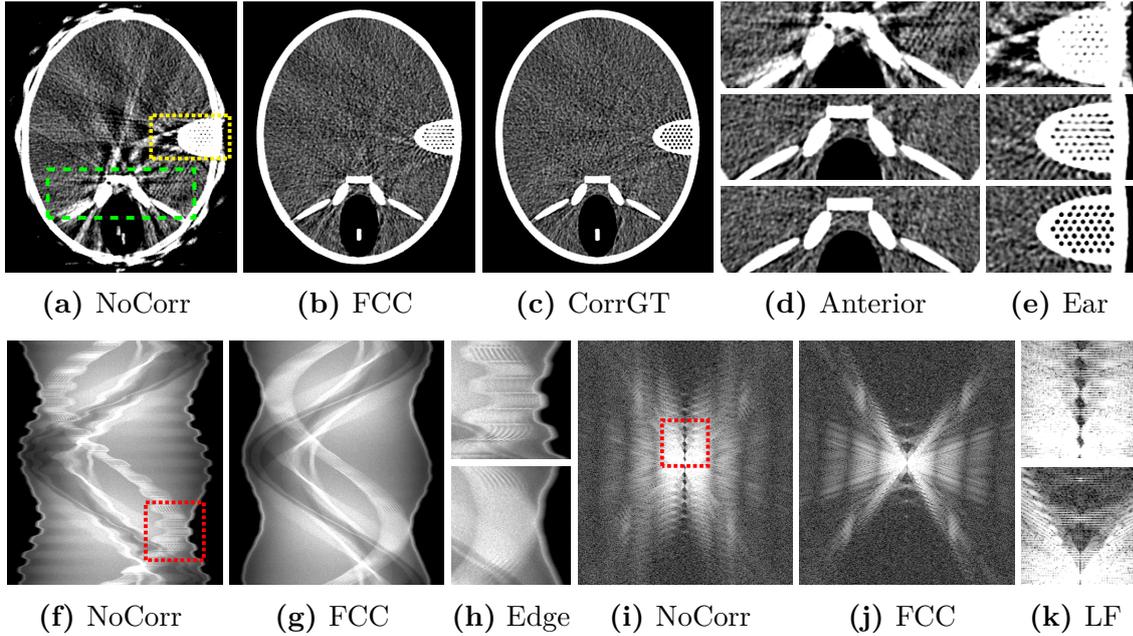


**Figure 9.7:** Sinograms and their logarithmically scaled Fourier domain for low-frequency motion. Arrangement and visualization windows are identical to Fig. 9.6.

logarithmic scale for the spectra to allow visualization of the triangular zero regions. The motion artifacts imposed by *Oscil*, *Chirp*, and *Rect* are clearly visible as small ripples of the sinograms' outline in Figures 9.6(g), 9.6(m) and 9.6(s). Triangular regions show a large spread of energy into the zero region in case no correction is applied. It is interesting to note that the *Chirp* motion pattern causes the largest spread of spectral energies into the zero regions, even though its translations had the smallest amplitude within the group of high-frequency motions. This indicates that FCC are particular sensitive to motion types that show a high temporal frequency.

After application of FCC-based motion correction, the sinograms' outline and trajectories of small high intensity objects appear well restored, which can be seen best in Figures 9.6(i), 9.6(o) and 9.6(u). This is also reflected in the corresponding spectra, which show a clear reduction of energy within the triangular regions, compared to no correction (cf. Figures 9.6(l), 9.6(r) and 9.6(x)). In fact, large amount of energies are redistributed to the support region of the object, causing a restoration of the zero regions, which could not be identified before correction. Applying FCC to ground-truth projection images has only very little influence on the sinogram (cf. Figures 9.6(a) and 9.6(b)) as well as the triangular regions (cf. Figures 9.6(d) and 9.6(e)).

From Fig. 9.7 we notice only minor differences between motion-free and motion-corrupted sinograms for low-frequency motion patterns. The motion is hardly visible and may be best noted by a slight translation and deformation between corrected and uncorrected sinogram boundaries as shown in Figures 9.7(c) and 9.7(i). The high-frequency motion types in Fig. 9.6 revealed that a higher frequency of the motion pattern has a stronger effect on the spectral zero regions. Consequently, the low-frequency motion *LF-1* and *LF-2* reveal only small variations to the triangular regions. For *LF-1* we observe spectral energies distributing as a vertical bar through



**Figure 9.8:** Reconstructions, sinograms and spectra for low-quality data based on the *Oscil* motion. The high noise level did not have substantial influence on the motion estimation approach and showed almost identical results as obtained for the noise-free cases. Visualization windows are identical to Fig. 9.4 and Fig. 9.6.

the center of the Fourier domain (cf. Fig. 9.7(d)). This originates from described discontinuities over the angular direction, based on different start and end position of the motion. The corrected spectrum in Fig. 9.7(e) reduces these contributions substantially, yet, cannot restore the triangle completely. In contrast, the motion cannot be noted in the uncorrected spectrum for *LF-2* and only very small differences exist after correction. In fact, the spread of spectral energies into the zero regions seems to increase slightly as can be seen in Fig. 9.7(l).

So far all presented results were based on high-quality, noise-free projection images. The motion correction performed equally well in presence of severe noise for all datasets. A representative example for noisy data is shown for the *Oscil* motion in Fig. 9.8. Apart from the reduced image quality, caused by the high amount of noise, no difference is observed when compared to the noise-free case. The noise is well visible in the entire spectral domain, yet, it did not influence the redistribution of energies from triangular to the object’s support region. These results generalize to all tested motion patterns, which is well supported by the quantitative evaluation shown in Sec. 9.3.2.

### 9.3.2 Quantitative Assessment of Image Quality

Quantitative evaluations of image quality include *SSIM* and *rRMSE*, both calculated on the entire reconstruction volume (see Sec. 7.3 for details). Their results are reported in Tab. 9.2. The *SSIM* and *rRMSE* results are given in Tab. 9.2a and Tab. 9.2b, respectively. Each row corresponds to a correction method, whereas columns show

		<i>None</i>	<i>Oscil</i>	<i>Chirp</i>	<i>Rect</i>	<i>LF-1</i>	<i>LF-2</i>
Ideal	NoCorr	100.0	61.7	85.1	92.7	73.6	69.2
	FCC	98.6	87.6	96.4	97.5	84.2	88.5
	CorrGT	100.0	90.7	97.6	99.1	95.4	94.2
Noisy	NoCorr	85.3	53.8	72.4	78.5	64.2	60.6
	FCC	83.8	75.0	82.1	83.0	72.4	75.7
	CorrGT	85.3	77.5	83.2	84.6	81.2	80.3

(a) Results of **SSIM** ( $\times 100$ )

		<i>None</i>	<i>Oscil</i>	<i>Chirp</i>	<i>Rect</i>	<i>LF-1</i>	<i>LF-2</i>
Ideal	NoCorr	0.0	10.4	4.4	7.3	12.5	7.5
	FCC	2.1	3.2	2.9	3.0	3.0	3.1
	CorrGT	0.0	1.9	0.8	1.5	2.4	1.8
Noisy	NoCorr	1.3	10.5	4.6	7.4	12.6	7.6
	FCC	2.5	3.4	3.2	3.2	3.2	3.3
	CorrGT	1.3	2.3	1.5	2.0	2.7	2.2

(b) Results of **rRMSE** in %

**Table 9.2:** Quantitative evaluation of image quality for all motion cases based on low- and high-quality data. We show **SSIM** and **rRMSE** results in (a) and (b), respectively. Results are shown for corrected reconstructions, using **FCC** as well as ground-truth motion, but also for uncorrected reconstructions.

the individual motion type including no motion (*None*). Values are presented for the noise-free and noisy data. Both **SSIM** and **rRMSE** are scaled by a factor of 100, where the latter is reported in %. It is worth noting that the measures determined for **CorrGT** represent an upper bound when using a purely translational motion model, hence, we do not expect a higher image quality after using **FCC**.

The presented method was able to increase **SSIM** values for all motion-corrupted datasets, while decreasing **rRMSE** substantially, compared to reconstructions without correction. After correction all **rRMSE** values are within a comparable range for both ideal and noisy data. However, improvements of **SSIM** are less consistent. Whereas the **SSIM** for high-frequency motion patterns is close to those of **CorrGT**, a wider gap can be noticed for *LF-1* and *LF-2*. The **SSIM** results for **CorrGT** show that the *Oscil* motion is most difficult to reconstruct when using only translations, followed by *LF-2* and *LF-1*. This is in line with the high amplitude of those motion patterns as shown in Sec. 7.1.2.

Application of **FCC** to the ground-truth projection images leads to a decrease of **SSIM** by 1.4 and 1.5 for ideal and noisy data respectively. This is in agreement with the **rRMSE**, which increased by 2.1% for noise-free and 1.2% for noisy data. A higher sensitivity to noise is shown by **SSIM** which reduced from 100 to 85.3 for the ground-truth data, i. e., without interference of motion. In contrast, the **rRMSE** did not substantially vary when adding a high amount of noise. Even though **SSIM** values are generally lower for noisy data, their relative changes for motion correction

		<i>None</i>	<i>Oscil</i>	<i>Chirp</i>	<i>Rect</i>	LF1	LF2
Ideal	<i>s</i>	788 (395)	946 (457)	889 (422)	1014 (444)	1159 (628)	960 (502)
	<i>t</i>	10 (0)	427 (128)	136 (5)	175 (4)	708 (295)	340 (215)
Noisy	<i>s</i>	771 (379)	925 (447)	798 (483)	1001 (427)	1145 (617)	944 (490)
	<i>t</i>	11 (0)	425 (127)	244 (5)	174 (4)	708 (295)	340 (215)

**Table 9.3:** MAD and SD (in brackets) based on estimated and ground-truth motion parameters. Values are provided in  $\mu\text{m}$ .

methods are similar to those obtained from noise-free data. This underlines the methods robustness to noise.

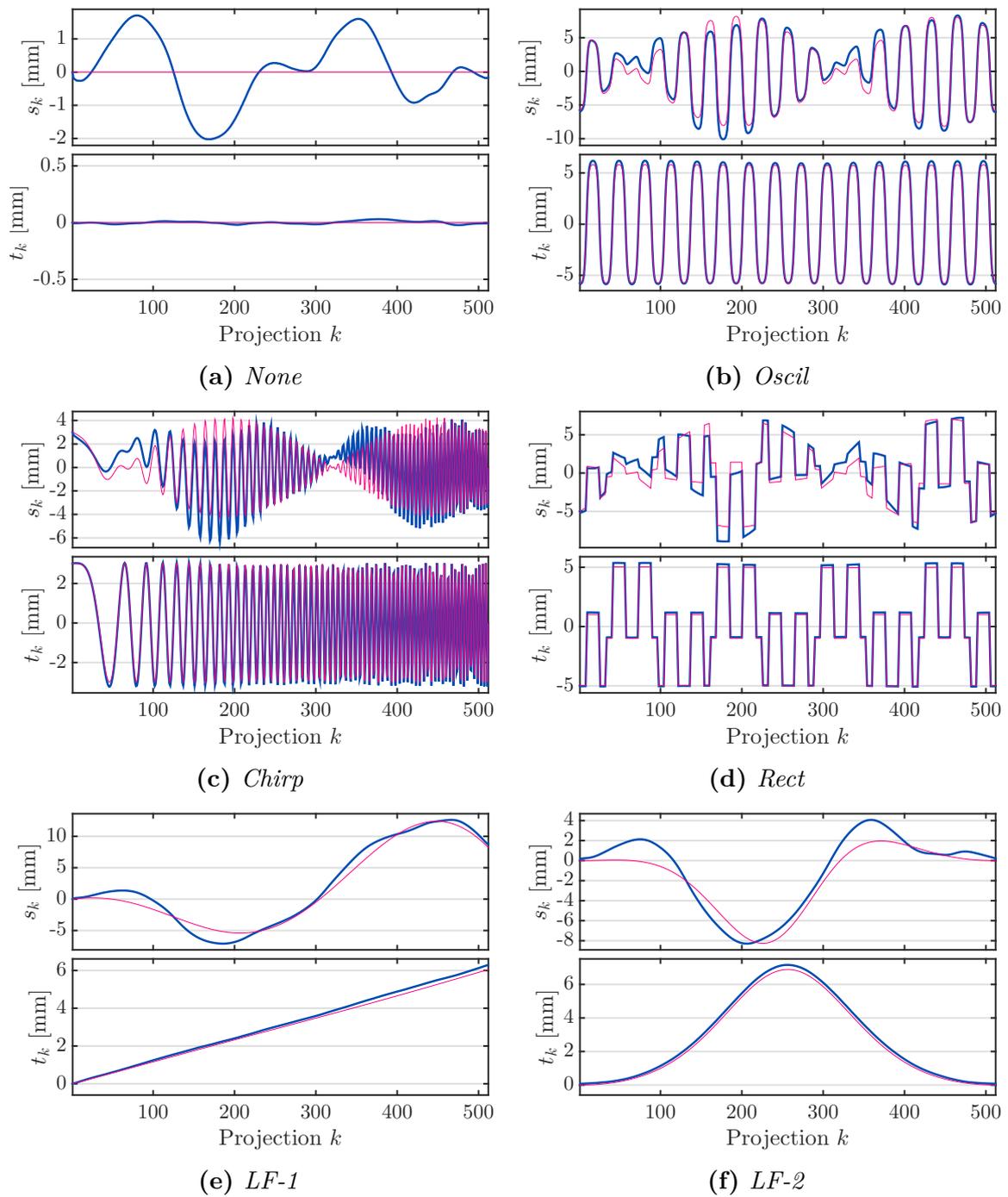
### 9.3.3 Accuracy of Motion Estimation

Using a numerical phantom allows access to the ground-truth motion, which was already used for motion-corrected reference reconstruction CorrGT. In Fig. 9.9 we show the estimated motion parameters  $s_k$  and  $t_k$  by a solid, blue curve. Additionally, the extracted ground-truth translations  $\hat{s}_k$  and  $\hat{t}_k$  are depicted by a solid, pink curve. Plots are shown for noise-free data only, yet, we noticed only little variation when evaluated on noisy projection data. In addition to the qualitative plots, we present the MAD and the associated SD in Tab. 9.3 for all evaluated datasets. Note that the values in Tab. 9.3 are defined in projection domain. Their effect for the reconstruction domain can be approximated by scaling with the systems magnification factor, i. e.,  $\frac{S}{S+D}$  which evaluates to  $\frac{1}{2}$  for all presented cases.

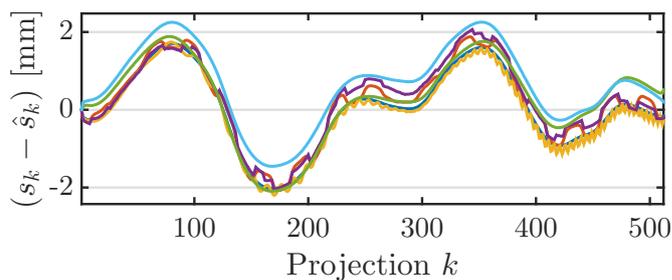
Fig. 9.9 reveals that the motion in detector  $v$  direction is estimated very well for all motion types. Even sharp transitions and fast variations are covered accurately yielding an almost ideal overlay of estimated and ground-truth motion parameters. Similarly, motion estimated for the motion-free data, as shown in Fig. 9.9(a), is negligible for detector  $v$  direction. The results for MAD and SD further support this observation in case of motion parameters  $t_k$  as shown in Tab. 9.3. The highest MAD values are obtained with  $708 \mu\text{m}$ ,  $427 \mu\text{m}$  and  $340 \mu\text{m}$ , for *LF-1*, *Oscil*, and *LF-2*, respectively. Considering the magnification factor of  $\frac{1}{2}$  the average effect on the reconstruction domain is less than  $354 \mu\text{m}$  and therefore rather small.

In contrast, larger deviations to the ground truth are visible in case of detector  $u$  direction, as can be seen for all motion cases in Fig. 9.9. However, the general trend of motion patterns is well covered especially in case of high-frequency motion as shown in Figures 9.9(b) to 9.9(d). The reduction of estimation accuracy is in line with the visual reconstruction results and manifests also in increased quantitative measures, with an MAD ranging between  $788 \mu\text{m}$  to  $1159 \mu\text{m}$ , where the highest value was obtained for *LF-1*. Similarly, the variations of the error, reflected by SD, are generally higher for detector  $u$  direction and have the largest values for low-frequency motion.

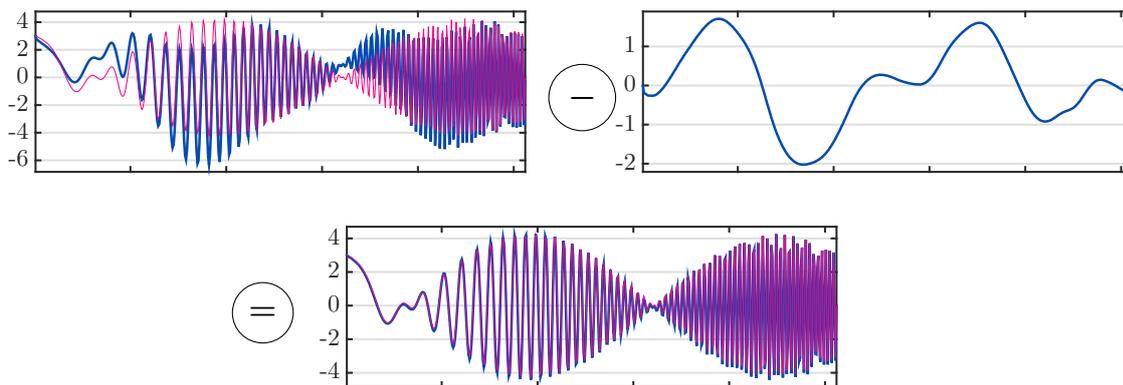
We investigated the difference of estimated and ground-truth motion for detector  $u$  direction in more detail. Fig. 9.10 depicts superimposed plots of parameter differences for all motion cases. We can clearly see that the deviation is dominated by a common low-frequency error, which coincides with the parameters estimated on the motion-



**Figure 9.9:** Individual plots of estimated and ground-truth parameters for detector  $u$  and  $v$  direction. Estimated parameters are depicted by a solid, blue line whereas the ground truth is shown as solid, pink line.



**Figure 9.10:** Difference of all estimated and ground-truth parameters for detector  $u$  direction, i. e.,  $(s_k - \hat{s}_k)$ . Note that all deviations show a similar motion pattern.



**Figure 9.11:** Low-frequency interference for detector motion in  $u$  direction, visualized for the *Chirp* motion. A subtraction of parameters estimated on the motion-free projection data, from the parameters estimated on projection images with motion, leads to a very accurate alignment.

free dataset. This indicates that the estimated parameters  $s_k$  are generally built by an interference of real motion and the error estimated on the ground-truth dataset. This procedure is visualized in Fig. 9.11, where we subtract the erroneous parameters from the ground-truth dataset (cf. Fig. 9.9(a)), from the parameters estimated for the *Chirp* motion. The result is a very accurate alignment of corrected parameters and ground-truth *Chirp* motion. We have noticed similar effects for all other evaluated motion types. It is worth noting, that the subtraction approach presented in Fig. 9.11 was only used for a better visualization and has not been applied for correction.

The amplitude of this systematic deviation is measured by the MAD of the motion-free data (cf. Tab. 9.3). It evaluates to an MAD of 788  $\mu\text{m}$ . The individual errors range between  $-2.02\text{ mm}$  to  $1.71\text{ mm}$ . After incorporation of the magnification factor we obtain an average error in reconstruction domain of 394  $\mu\text{m}$ .

## 9.4 Results for Simulated Knee Phantom

Within this section we show results of the FCC-based motion correction applied to the XCAT dynamic squat datasets XCAT\_MOT and its motion-free ground truth XCAT\_GT. In contrast to the FORBILD phantom, axial truncation is inevitable

when imaging knees. Yet, the datasets do not suffer from lateral truncation as the detector spacing was increased to fully cover both legs within axial planes. Note that the FCC method is not yet applicable to real knee acquisitions, as it requires full scans and truncation-free acquisitions in lateral direction. Nevertheless, we consider this experiment as a first step towards application in real scenarios. We are specifically interested, if the proposed extension based on dynamic apodization of projection images in detector  $v$  direction, improves robustness to axial truncation.

### 9.4.1 Parameter Selection and Image Reconstruction

Parameters of the method differed from those selected for the FORBILD phantom. In the following we present adjustments and selected parameters for the XCAT data. First of all, we used the updated cost function in Eq. (6.28), including apodization in detector  $v$  direction (cf. Sec. 6.5.2) and L2 regularization of motion parameters. The maximum object extent  $r_p$  was estimated by the same procedure used for the FORBILD study (cf. Sec. 9.1.3), leading to  $r_p = 113.5$  mm.

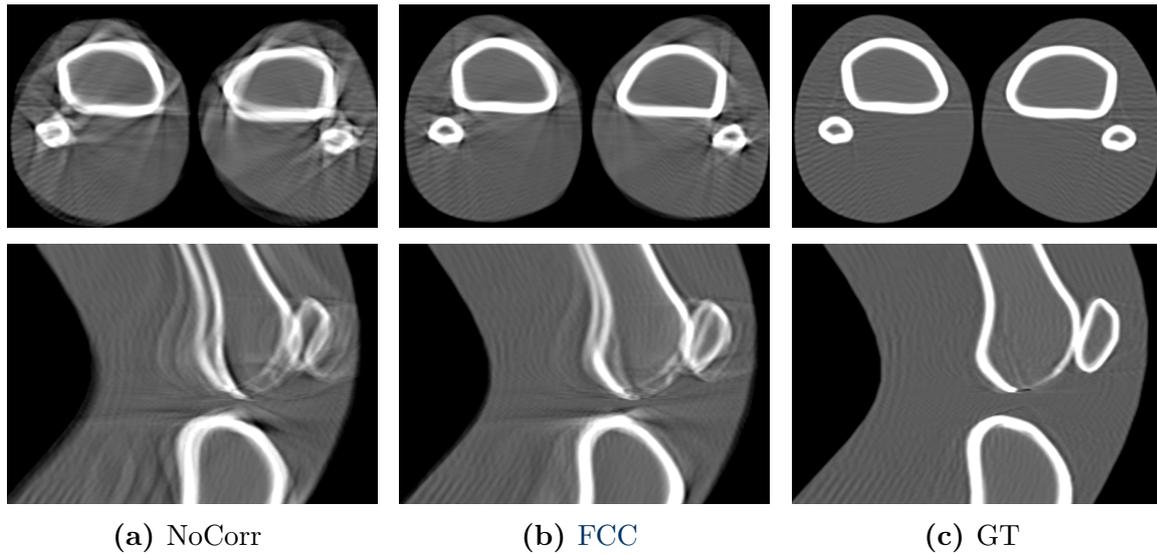
The distance of theoretic and adjusted triangle outlines is set to  $\epsilon = 0.01$ . In addition we set the Lagrangian of the L2 regularizer to  $\eta' = 1$ . For example, a constant shift of 1 mm in  $u$  and  $v$  direction, would lead to a contribution of 0.5, caused by regularization. To put this into contrast, the FCC cost function  $e(\boldsymbol{\alpha})$  was normalized to  $e(\boldsymbol{\alpha}) = 100$  after the first evaluation. Note that we heuristically adjusted  $\eta'$  and  $\epsilon$ , whereas other parameters remained constant.

We did not have access to ground-truth detector translations. Consequently, we decided to remove the regularizer presented in Eq. (6.22) by setting its Lagrangian multiplier  $\tilde{\eta} = 0$ . However, a constant bias in motion parameters is prevented by usage of the L2 regularizer, which was part of the apodization procedure for truncation robustness.

The reconstruction pipeline was identical to that introduced in Sec. 9.1.3. We performed reconstructions of XCAT\_GT without motion correction to obtain a ground-truth volume. In addition, XCAT\_MOT is reconstructed with and without application of FCC. The reconstruction volume size was  $512 \times 512 \times 256$ , with an isotropic voxel size of 0.5 mm.

### 9.4.2 Image Quality of Reconstructed Images

Fig. 9.12 shows the reconstructed images of the XCAT phantom. The images show axial slices through tibia and fibula (top row) and sagittal slices of right-sided femur, patella, and tibia (bottom row). From left to right, we can see reconstructions without motion correction (NoCorr) in Fig. 9.12(a), after correction using FCC in Fig. 9.12(b), and the ground-truth, motion-free data (GT) in 9.12(c). Severe motion artifacts can be seen in case no correction is applied, manifested by streaking and blurring. In addition, the outlines of femur and patella are hardly visible in the sagittal slice of Fig. 9.12(a). Our method was able to restore most of the structure at left and right tibia and shows a clearer visibility of the fibulas. The structure of femur and patella improved by a large amount, yet, the difference to the ground-truth reconstruction



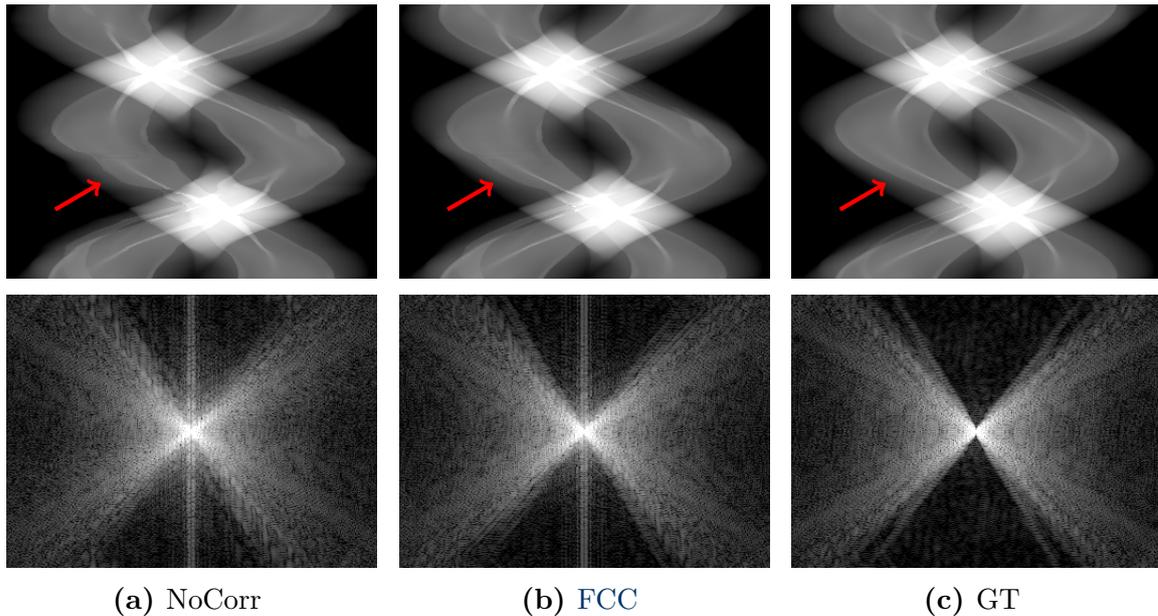
**Figure 9.12:** Reconstruction results for the XCAT squat phantom. We show axial slices inferior to the knee joint (top row) and sagittal slices through the right leg (bottom row). Reconstructions are presented without correction (NoCorr), after correction using FCC, and of the ground-truth, motion-free projection images (GT). (W: 1815 HU, C: 246 HU).

is still substantial. Overall, the skin surface and soft tissue appears sharper after correction and contains less streaking artifacts.

We conducted the same quantitative evaluation as used for the FORBILD data (cf. Sec. 9.3.2). The resulting values support the visual impression in Fig. 9.12, where NoCorr yielded an SSIM of 69.7 and an rRMSE of 6.58%. After applying the proposed motion correction image quality improved, showing an SSIM of 79.2 and an rRMSE of 6.00%.

### 9.4.3 Improvements in Projection and Fourier Domain

In addition to reconstruction results we also show sinograms extracted at a central line of the detector, i. e., for  $v = -\frac{\Delta v}{2} = -0.6$  mm. The sinograms are shown in the top row of Fig. 9.13. The irregular trajectories of objects clearly depict the motion artifacts for the uncorrected sinogram in Fig. 9.13(a). The location of the most apparent motion artifact is indicated by a red arrow. After correction the trajectory of several objects appears smoother and a clear improvement can be seen at the location of severe motion. Yet, we still observe several differences when comparing FCC and GT (cf. Figures 9.13(b) and 9.13(c)). The spectra, shown with a logarithmic scaling at the bottom row of Fig. 9.13, depict the triangular regions for a vertical frequency of  $\psi = 0$  mm<sup>-1</sup>. The proposed motion correction could remove a large amount of energy from the triangular regions compared to NoCorr, yielding a sharper outline of the zero regions. However, after correction we still notice a higher amount of energies located within the zero regions, compared to the ground truth in Fig. 9.13(c). In general, the spectral energies are in line with their assigned sinograms, but also with the achieved image quality.



**Figure 9.13:** Resulting sinograms and spectra for the XCAT dynamic squat phantom. Top row: Sinograms are extracted at a central slice for  $v = -0.6$  mm. Bottom row: Spectra show the central slice in Fourier domain, i. e., for  $\psi = 0$  mm<sup>-1</sup>, after logarithmic scaling. Visualization windows of sinograms and spectra were  $[0.05, 0.5]$  and  $[4, 12]$ , respectively.

## 9.5 Discussion and Conclusion

Within this thesis we have shown that motion correction in weight-bearing imaging of knees can be achieved by fiducial markers or by means of 2-D/3-D image registration. Both methods performed accurately and are of potential use in a clinical application. However, a multitude of disadvantages have been outlined for the marker-based method (cf. Sec. 8.6), e. g., it directly interferes with the acquisition protocols and causes metal artifacts. We show that registration-based methods can alleviate most of the problems associated with markers. Yet, they require an artifact-free image of the anatomies, where we had access to a high-quality, prior scan acquired while the patient is in a supine position. In a future clinical application these supine scans may not always be necessary for diagnosis and would cause unnecessary dose to the patient.

A solution to this are motion correction approaches that do not rely on additional acquisitions or surrogate signals. We have identified **CC** to be of potential use for such motion correction methods. The baseline of our work was given by **FCC**, which have been defined for parallel- [Edho 86] and fan-beam [Natt 86, Mazi 10] geometries. No prior work was available for **FCC**-based motion correction. Additionally, the presented concepts were rather theoretic and have been far from applicable to real **CBCT** data. Within the scope of this thesis we introduce a framework for motion correction based on **FCC**. Initially we defined our method for fan-beam geometries [Berg 14b]. Subsequently, an extension to **CBCT** is proposed using and extended

version of FCC. Finally, first solutions to data truncation are presented, by using dynamically applied apodization windows.

**Proof of Concept for Fan-Beam Geometry** We consider the initially proposed motion correction method for fan-beam CT as a proof of concept. An evaluation based on the FORBILD phantom showed impressive results when applied to a simulated full scan with a high-frequency motion pattern. Image quality improved substantially, compared to reconstructions without corrections. The presented results in Sec. 9.2 also show that the method is robust to high amounts of noise and also works for different detector and angular resolution levels. Even though motion was estimated accurately, the method had difficulties to restore very sharp edges, e. g., the resolution pattern in Fig. 9.2. A reason might be the limited flexibility of the translational motion model, which is not able to cover object scaling, typically caused by motion orthogonal to the detector. In summary, results of this initial investigations verified the potential of our motion correction method, encouraging further development for cone-beam geometries.

**FCC for Cone-Beam Geometry** A major contribution of this thesis is a practical extension of FCC from fan- to cone-beam CT. The cost function uses the 3-D DFT of the projection images. Experiments based on the FORBILD phantom presented in Sec. 9.3, show a great reduction of motion artifacts for all evaluated datasets, where even motion containing very high frequencies could be accurately detected. The method does not require any regularization imposed on the estimated motion parameters which allows for a wide variety of motion types. In contrast, Yu et al. [Yu 06] used a polynomial model for the motion along temporal direction when evaluating their CC-based motion correction method on fan-beam reconstructions of the same numerical phantom. Thus, complexity of the motion reduced to a total of six parameters. We successfully estimated  $2K - 2 = 1022$  parameters, i. e., two translations per projection image, which underlines the potential of the proposed cost function. In addition to a dedicated cost function for estimation of detector translations, we present a generic cost function in Equations (6.2) and (6.26) which allows arbitrary rigid or nonrigid deformations of the projection images. However, limitations and potential invariance of the cost function to motion types beyond detector translations are yet to be investigated. Additionally, most of these motion types would need to be applied in spatial domain, thus, a 3-D FFT is needed for each partial derivative and evaluation of the cost function.

Image quality could be substantially improved by FCC-based motion correction for all motion types. We observe SSIM but also rRMSE values close to those obtained using the ground-truth detector translations (CorrGT) for all high-frequency motion types. In comparison, quantitative values for low-frequency motion could not reach the corresponding image quality of CorrGT, yet, they still show improvements w. r. t. no motion correction. The spectra shown in Figures 9.6 and 9.7 clearly show that less variations in the underlying motion also leads to less impact within the triangular regions, indicating that the cost function may be less sensitive to low-frequency motion types. This is particularly apparent, in case the first and last object position are identical, where we see only subtle differences between motion-free and motion-

corrupted spectra (cf. Figures 9.6(d) and 9.7(j)). Opposed to the visual impression, optimization could still reduce the total energy within the zero regions, which was based on a small number of low-frequency bins with high intensity values. Still, image quality of the reconstructed images improved, even for the *LF-2* motion, and also the *MAD* is similar w. r. t. the other motion types investigated. Independent of the motion types, very accurate results could be obtained for translations in detector *v* direction. This makes *FCC* an interesting method to detect and estimate respiratory motion which is known to occur mostly in axial direction [Unbe 16b].

**Restrictions of Translational Motion Model** In general we see that the visually determined image quality may be better reflected by *SSIM* rather than *rRMSE*. A reason for this may be that *SSIM* is rather sensitive to alignment of edges and thus representative for measuring motion-induced artifacts or geometric misalignments. In consequence, we can use *SSIM* to verify the restrictions of our motion model when looking at results obtained from *CorrGT*. One reason for the reduction of image quality for *CorrGT* is the approximative nature of applied motion compensation method as shown in Sec. 3.5. Another reason is that translations in projection domain cannot perfectly reflect 3-D translations, particularly for large motion amplitudes and or motion along the principle ray. An additional scaling parameter, or even nonaffine deformations, for each projection image could alleviate these deficiencies to a large extent.

**Robustness to Noise** Similar to the results obtained in the initial fan-beam study, the approach proved to be robust to severe Poisson distributed noise. This becomes apparent when comparing the results in Fig. 9.8 to the results obtained on the noise-free dataset in Figures 9.4(f) to 9.4(j) and Figures 9.6(g) to 9.6(l). Moreover, the improvements of *SSIM* and *rRMSE* relative to reconstructions without correction, are similar for noise-free and noisy data, as can be obtained from Tab. 9.2. Also, *MAD* and *SD* values of the estimated motion parameters (cf. Tab. 9.3) reveal that the accuracy of the motion estimation is not affected.

**Invariance to Static Translations** We were able to show analytically that the proposed cost function is invariant to static translations, e. g., a constant misalignment of the detector. Within the context of our study we solve this restriction by fixing translations of the first projection image to the ground-truth motion which had several benefits for our experiments. First of all, it ensured that the image quality does not vary due to different estimated detector offsets, thus, allowing for an accurate comparison of the method’s ability to estimate actual motion. Another benefit is, that the adjustment implicitly aligns the coordinate systems of all reconstructed volumes to the ground-truth reconstruction, which is important to ensure that quantitative measures are not dominated by misalignments but rather motion artifacts.

In case of the *XCAT* dynamic squat phantom, or for any real world scenario, we do not have ground-truth translations. Thus, no direct regularization of translations from a single time point was applied. Instead we assumed that the average of the estimated motion is small. This was implicitly enforced by *L2* regularization, which

was applied anyway to cope with axial truncation. Given that the cost function is invariant to static translations, the L2 norm will lead to motion estimates with a mean value close to zero.

In general, a moderate offset in detector  $v$  direction will merely cause a shift of the object along the  $z$ -axis and possibly a minor increase of cone-beam artifacts. Yet, an offset in detector  $u$  direction can cause artifacts, especially in axial slices. Thus, a simple solution may be to create reconstructions for a series of offsets in detector  $u$  direction and manually select the reconstruction with the best visual image quality. This could be automatized using image quality measures defined in reconstruction domain, such as Entropy or Positivity [Rohk 13].

**Systematic Deviations in Motion Estimation** An important observations was the systematic deviation caused by our method for translations in detector  $u$  direction as presented in detail in Figures 9.10 and 9.11. In summary, our experiments revealed that the estimation error for horizontal translations is built by the actual error of motion estimates and an additive bias which was constant over all FORBILD datasets, including the motion-free case. The deviation may originate from inaccuracies due to discretization of the triangular regions, as described in Sec. 6.5.1. A small increase of the mask size is applied to ensure that no spectral regions are undefined which was necessary to allow correction of low-frequency motion types. Conversely, this may cause explained residual errors as the cost function tries to remove energy in regions that already belong the support of the object. In contrast, a mask size that was too small often caused motion that shifted energy to undefined spectral regions, located between mask and support region. During our experiments this often caused interference by a sinusoidal motion of low-frequency and very high amplitudes. A future extension could eliminate this interference and thus, improve robustness of the method.

In case of patient motion smaller than the systematic offset our method may lead to a slight reduction of image quality as was the case for the presented motion-free data in Figures 9.4(a) to 9.4(e). On the other hand, larger motion artifacts will easily cover the systematic deviation resulting in a clear reduction of motion artifacts. Nevertheless, this issue requires further investigation for future applications.

**Towards Application to Real Weight-Bearing Acquisitions** The overarching goal is to further develop FCC-based motion correction such that the method is applicable to weight-bearing imaging of the knee. As part of this thesis we described the initial version of the approach, which was based on several limitations. It was only applicable to fan-beam geometries, required truncation-free data, a full-scan acquisition, and ideal monochromatic absorption. Within the process of this thesis we presented solutions to eliminate or at least cope with some of these limitations, i. e., an extension to CBCT and an initial approach to allow for axial truncation of the projection data.

Eventually, these extensions allowed a first application to the XCAT dynamic squat phantom, which was based on real motion parameters [Choi 13]. Our experiments show an improved qualitative and quantitative image quality compared to

reconstructions without correction. With the presented evaluation we corroborate that FCC are indeed applicable to motion correction of weight-bearing acquisitions.

However, a series of aspects need to be addressed before the method is applicable to real weight-bearing acquisitions. The truncation robustness needs to be extended to cover also lateral truncation and investigations are needed for short-scan acquisitions.

## Conclusion

Previously presented approaches for motion correction in weight-bearing imaging rely on either surrogate signals or availability of prior acquisitions. Within this thesis a novel approach for motion correction in CBCT has been introduced that is independent of such additional data. The approach is based on a cost function that aims to optimize FCC, defined in the 3-D Fourier domain of the projection image volume. Motion is estimated in projection domain by optimizing 2-D detector translations in order to fulfill FCC as accurately as possible. We initially propose the method for fan-beam CT and derive an extension for an application to motion-corrupted CBCT data. We present an extensive evaluation of the proposed method using the challenging FORBILD head phantom, for a variety of translational 3-D motion patterns. The method shows a particularly good performance when estimating translations in detector  $v$  direction, i. e., collinear to the rotation axis of the CT. Motion with high-frequency was accurately identified leading to large improvements of image quality. No temporal assumption is imposed on the estimated detector translations, allowing application to a great class of motion types. Several issues are yet to be addressed before an application to real weight-bearing data is possible. Our evaluation revealed a systematic misestimation for translations in detector  $u$  direction, which requires further investigation. Also, FCC are defined for full scans only, thus an extension to short scans is needed. On the other hand, we introduced an extension to improve robustness in presence of axial truncation, which eventually allowed application to a squatting knee phantom containing real motion. Based on the results of our evaluation, we are confident that FCC are applicable for design of a self-contained motion correction method in weight-bearing imaging of knees.



# Outlook

Many of the presented approaches in this thesis offer new possibilities for future research directions. In this chapter we provide a selection of methodological and clinical advancements that could further improve CBCT weight-bearing imaging of knees. Overall we present ideas for three different motion correction methods in CBCT, i. e., marker-, registration- and consistency-based correction.

The marker-based approach existed prior to this thesis and was steadily improved, making it a robust and sophisticated method. Yet, we can think of some weaknesses that may be targeted by future research. First, the cost function for motion estimation in Eq. (4.16) is based on point correspondences between 3-D reference points and 2-D detections. Currently, these correspondences are obtained only once, prior to optimization. Potential outliers due to wrong assignments are then removed by the outlier detection presented in Sec. 4.2. An interesting approach for future work is an iterative adjustment of point correspondences, i. e., alternating between reassignment of correspondences and motion estimation. This is essentially very similar to an iterative closest point registration [Besl92], but in the context of 2-D/3-D rather than 3-D/3-D registration.

Further, we plan to use the final result of the marker-based approach for an analysis of residual distances in projection domain, w. r. t. their directions and lengths. This could provide further insight in the level of nonrigid motion. Another possible direction is to investigate to which extent moderate, nonrigid 3-D deformations can be modeled by markers.

Yet, for correction of nonrigid deformations, we see a higher potential when using the registration-based approach, as the joint motion can be accurately modeled using bone-wise rigid transformations. However, a couple of remaining challenges need to be solved before the method is applicable in clinical routine. We did not have access to the ground-truth motion of bones and can therefore not evaluate the accuracy of the estimated motion parameters. In consequence, our evaluations are based on image quality measures in reconstruction domain. In a future study cadaver legs may be used where tibia and femur are attached to an apparatus that allows application of defined motion patterns. This way, an exact evaluation of the registration error would be possible. Additionally, fiducial markers may be implanted at certain points of interest in the soft tissue of the cadaveric legs, allowing for a verification if TPS-based motion extrapolation is sufficient.

In its current state the registration-based method requires a high amount of manual interaction as bone segmentations and initial translations need to be adjusted at least semi-manually. Further automation of these steps may be crucial for clinical ap-

plicability. For example, annotation of only few anatomical landmarks could be used to allow for a fast semi-manual alignment between standing and supine scans. Also, Franz et al. [Fran 06] proposed a more advanced 2-D/3-D registration method that is independent of a prior pose initialization, which could even allow for a full automation of the registration process. Further, the segmentation of bones may be improved by use of statistical shape models, derived from a set of prior patient acquisitions [Heim 09]. Statistical shape representations and featured-based 2-D/3-D registration can also be combined into a single framework, as was recently proposed in the field of computer graphics [Thie 15, Thie 16]. This would allow a complete replacement of bone segmentations, and thus, could save the acquisition of a motion-free scan of the knees.

We still see a large potential in improving the registration accuracy of the 2-D/3-D registration. In its current form, the method only applies the assumption of smooth transitions of motion parameters. Surely, a more sophisticated biomechanical model of the knee joint can be used to regularize the registration. We believe that such a model could drastically improve the method's robustness, while preserving a large coverage of also nonsmooth motion types. In case high geometric accuracies are required for bones and soft tissue, it may be interesting to combine the registration with the marker-based approach. Marker information may be incorporated into the nonrigid motion extrapolation of the TPS models or directly during 2-D/3-D registration. Alternatively, the 2-D/3-D registrations could be extended to a flexible, yet more complex, nonrigid motion model as used in [Zeng 05, Zeng 07].

As part of this thesis we introduce a novel motion correction based on FCC. We present an extension to CBCT and propose an initial approach to handle data truncation. Yet, the method is still in an early phase and not yet applicable to real acquisitions of the knee. Nevertheless, our results are promising and encourage further development. First and foremost an extension of the FCC theorem is needed to allow application to short-scan acquisitions. The short scan could potentially be interpreted as truncation in angular direction, such that apodization (cf. Sec. 6.5.2) may also help to overcome this limitation. In general, the apodization-based method for truncation robustness requires further evaluation and potentially improvements. In particular, a dynamic application of apodization causes variations of the total spectral energy, which interferes with the FCC-based cost function. Currently, we use a regularizer based on the L2 norm, yet, this only limits the effect but does not remove it. A smart normalization applied per projection image could help to overcome this issue and should be verified in a future study. Additionally, an extension of the motion model to allow for more complex motion patterns and possibly object rotations is of interest. This is easily incorporated into the generic cost function given by Eq. (6.2), yet, eliminates the use of an efficient implementation as the motion may not be applicable in the 2-D Fourier domain of the projection images. An extension of the regular Fourier transform to the a nonuniform Fourier transform could help to overcome these limitations, allowing application of spatial affine transformations directly in Fourier domain. Further investigations on the origin and severity of the systematic deviations for detector  $u$  direction are essential. Of particular interest is, if and how spectral energies minimized on the motion-free ground truth are object dependent. And if not, whether a connection to the geometry or sampling parameters

can be formulated. This insight is important for a reduction or even elimination of the horizontal bias, where the goal is to achieve similar accuracies as obtained for motion estimates in detector  $v$  direction. A comparison of FCC-based motion correction to other approaches, that also do not rely on additional acquisitions or surrogates, would be of great scientific value. Other methods defined in projection domain are for example image moments [Clac 15] or epipolar CC [Aich 15, Frys 15]. We are also interested if approaches based on MAM (cf. Sec. 3.4.4) defined in reconstruction domain, could be a potential alternative to our method. We expect that this will be particularly challenging, as the image entropy has already been reported to contradict with SSIM values for weight-bearing imaging of knees [Choi 14c].

From a clinical point of view, the presented methods provide valuable tools for motion management, e. g., for the studies introduced in Sec. 2.3. The marker-based motion correction can be used as is for efficient and fast motion correction in case the motion is approximately rigid. A side product of the quantitative evaluation in Chap. 8, i. e., the registration of bones reconstructed from weight-bearing scans to the bone in supine position, may be helpful for cartilage analysis. As shown in Fig. 8.6, it enables aligned reconstructions of all weight-bearing scans which could help for a consistent analysis of cartilage strain. Automatic segmentation methods of the contrast agent's surface are of great interest and would allow for a direct extraction of cartilage features. The proposed registration-based method enables nonrigid joint motion during a single scan. This could allow for a common analysis of the cartilage at various locations along lateral and medial condyles. An ambitious long-term goal are acquisitions during realistic walking on using a treadmill. In this case, motion correction may also be used to obtain 4-D, time resolved reconstructions. Similar studies are conceivable for advanced evaluation of PFPS, where a 4-D reconstruction of a squat could even allow for a precise 3-D tracking of the patella under realistic weight-bearing conditions.



## CHAPTER 11

# Summary

Within this thesis we propose several contributions to the state-of-the-art in motion-corrected CBCT reconstruction. The application area for the presented approaches is weight-bearing CBCT of knee joints in standing or squatting position. The thesis is divided in three parts. The background of weight-bearing imaging and motion-corrected reconstruction is given in Part I. In Part II we present the theory and concepts of the proposed methods, followed by their evaluation in Part III.

### Part I: Background

An overview of the knee anatomy and its role as joint of the lower bodies locomotor system is given in Chap. 2. OA and PFPS are two of the most common knee joint disorders, showing high incident rates. Medical imaging plays a vital role for diagnosis and classification of these conditions. We show that 3-D imaging under weight-bearing conditions, can lead to a more precise diagnosis. This thesis builds on a novel CBCT device which allows 3-D weight-bearing acquisition of knees with high spatial resolution and fast acquisition times [Choi 14a]. Initial feasibility and clinical studies of CBCT weight-bearing imaging are presented. A study related to OA, aims to investigate cartilage strain under weight-bearing conditions using direct arthrography. The novel weight-bearing acquisition mode introduces several challenges for 3-D image reconstruction, i. e., detector saturation, data truncation, and patient motion based on the unsupported standing position. We show that patient motion is the most dominant source of image artifacts, thus, motion correction is mandatory to allow for diagnosis based on the reconstructed images.

In Chap. 3 we first introduce fundamental image reconstruction algorithms along with a description of acquisition geometries. Also, the mathematical basis of the thesis is provided, e. g., introducing projection matrices and concepts of discretization. Within this thesis, motion correction is defined to consist of a motion estimation step, followed by a motion compensated reconstruction. We focus on the optimization of parameters that describe a certain motion model. A description of such models is given for rigid and nonrigid motion, using matrix multiplications or TPS deformations, in projection as well as reconstruction domain. Algorithms for motion estimation are categorized into three groups: 1) methods based on surrogate signals which usually interfere with the acquisition, 2) methods based on a prior image usually incorporated using image registration, and 3) purely based on the acquired motion-corrupted projection images. A large contribution is the extensive literature review provided for motion estimation methods w. r. t. the introduced categories. As

part of the first group, we show a motion estimation based on tracking of fiducial markers attached to the skin [Choi 13, Choi 14c], which also builds the baseline of this thesis. Additionally, other surrogate-based techniques are provided, e.g., in the field of radiation therapy or cardiac imaging. Within the second group, literature is presented that uses image registration to estimate patient motion, with a focus on 2-D/3-D registration approaches. The last group introduces methods based on metrics that are optimized in reconstruction domain, e.g., image entropy. Additionally, we provide a thorough review of CC, followed by an overview of their application to motion estimation. We conclude the chapter with an analysis of the presented literature w.r.t. its applicability to CBCT weight-bearing imaging of knees. Many of the methods are based on certain assumptions on the motion or assume its consistency within subsequent scans. Both is not given for weight-bearing imaging of the knees. Some methods even rely on modalities that are not applicable to a weight-bearing scenario, e.g., a prior 4-D CT scan is often used in radiation therapy.

## Part II: Theory

In prior work, fiducial markers have been used to estimate motion parameters, yet, the approach required a large amount of manual interaction and suffered also from incorrect assignments of point correspondences [Choi 13, Choi 14c]. Within the scope of this thesis, marker-based motion correction has been steadily improved. The individual advancements are detailed in Chap. 4. The main contributions include a robust, fully automatic detection of markers, with an automatic assignment of correspondences [Berg 14a]. An essential step of marker segmentation is the FRST, which replaced an approach based on the Hough transform, leading to robust suppression of background structures. In addition, we introduce a method for automatic extraction of average 3-D marker locations. The FRST result is backprojected to reconstruction domain and high intensity blobs caused by markers are automatically segmented. The locations are then given by the blobs' centroids, determined using a 3-D connected components analysis. Forward projections of the 3-D reference points are then used to obtain correspondences between 2-D detections and 3-D marker locations. Also, an automatic outlier removal and an advanced, gradient-based optimization is proposed [Mull 15b, Berg 16a]. To avoid metal artifacts, we propose an extrapolation scheme in projection domain, that removes the markers prior to reconstruction [Berg 14a]. Based on the contributions of this thesis, the marker-based approach is now a fully automatic motion correction tool which we deem ready for clinical practice.

Nevertheless, the approach has several weaknesses including, metallic artifacts, its limitation to purely rigid motion, and most importantly a direct interference with the image acquisition, due to the careful attachment of markers. In case a small acquisition time is important, a surrogate-free motion correction is required. In Chap. 5 we introduce a novel motion estimation method based on bone segmentations of a previously acquired supine scan. Rigid motion parameters of femur and tibia are optimized during a 2-D/3-D registration of segmented bones, to each of the weight-bearing projection images [Berg 15, Berg 16a]. Final motion estimates are then incorporated into a nonrigid motion field using TPS. Rigid motion is estimated for each bone yielding up to 24 parameters, whereas prior work on 2-D/3-D registration usually only estimates six rigid parameters. To increase registration robustness, a regularization that

enforces temporal smoothness of motion parameters is proposed. Overall, a certain amount of manual interaction is required, e. g., for bone segmentations and manual alignment from supine to standing scans. Yet, the additional time is part of image reconstruction, hence, the method has no influence on the acquisition itself.

Registration-based motion correction requires a prior, motion-free acquisition, which was part in all our clinical studies. However, in a future clinical application this reference may not be required, such that the supine scan causes additional dose for the patient, without increasing the diagnostic value. In Chap. 6 we derive an innovative motion correction, that enforces FCC defined in the spectrum of the projection domain. FCC are a theoretic concept that defines certain triangular regions in the spectrum to have an absolute value close to zero. Patient motion violates FCC, allowing a motion estimation that aims to decrease the energy in the triangular areas by optimization of a motion model defined in projection domain. We initially derived the method for fan-beam geometries [Berg 14b], followed by an extension to CBCT [Berg 16b]. An efficient implementation is suggested for optimization of translational detector motion. We also propose usage of a dynamic apodization method to allow for axial truncation. Spectral triangles are defined by known geometric parameters (i. e., SID and DID) and the maximum extent of the object w. r. t. the isocenter. A simple approximation of the latter, based on the acquired data, is part of our work. Hence, the approach does not require surrogate signals nor any additional acquisitions.

### Part III: Experiments and Results

In the last part of this thesis we evaluate the theoretical concepts for motion correction presented in Chapters 4 to 6. The marker- and registration-based approach are jointly evaluated in Chap. 8 using a total of nine real weight-bearing acquisitions. The novel FCC-based motion correction was evaluated in Chap. 9 using simulated projections of the numerical FORBILD head phantom [Laur 01]. Additionally, we show an evaluation using an XCAT-based squat phantom that incorporates particularly realistic motion patterns [Sega 10, Choi 13]. Data acquisition for simulated and real datasets is explained in Chap. 7. We introduce all relevant acquisition parameters for weight-bearing and simulated scans and define a set of representative, artificial motion patterns used for the FORBILD simulations.

Looking at the resulting image quality of the proposed marker-based method in Chap. 8, we can see that structural information such as bone and cartilage surfaces are well restored with a clear reduction of motion-induced streaking and blurring artifacts. However, due to its current limitation to rigid motion the method's performance substantially reduced for two datasets that contained severe and clearly nonrigid motion. Note that an extension of the method to nonrigid motion is non-trivial as it requires a large number of motion parameters, whereas the number of markers is inherently limited due to practical and technical reasons. One of the main issues of the marker-based approach is the attachment process which leads to a decrease of patient comfort and may hinder image acquisition in clinical routine. In addition, our results indicate that a marker-based correction can cause slight distortions at the bones leading to a reduction of our quantitative measurements. This supports that externally attached markers may not be suitable to estimated bone or cartilage motion accurately. The majority of the deficiencies related to the marker-

based method can be alleviated by the proposed motion correction using 2-D/3-D registration. A qualitative and quantitative comparison of the two methods reveals similar improvements in image quality. Yet, for the two datasets with severe, nonrigid motion, we see a substantially better correction by the registration approach. More robust 2-D/3-D registration results are obtained using *NGI* as a similarity measure [Otak 13], rather than the well-known *GC* [Penn 98]. Also, reconstructions of the cartilage showed consistently clear visualization of the contrast agent and bone surfaces, which builds the baseline for further quantitative analysis of cartilage strain.

In case patient comfort and acquisition time is of little importance and the motion is restricted to rigid movements, we recommend usage of the extended marker-based correction followed by a marker removal step to avoid metal artifacts. The method is readily applicable, fast, and requires little manual interaction. Furthermore, it is not restricted to knee anatomies and has already been applied to estimate head motion of stroke patients [Mull 15a]. However, the clinical application may prohibit prolonged image acquisition. Additionally, a restriction to only rigid motion, e. g., by external fixation, further delays acquisition and may even contradict with the clinical motivation. For these cases, we recommend the more flexible registration-based motion correction, which we have shown to be similar or even more accurate than a marker-based correction.

Ideally, motion correction can be performed using only the acquired projection images. We introduce a novel approach based on *FCC*, that belongs to this group of algorithms. Prior to this thesis no *FCC*-based motion estimation was known in literature, showing that the approach is still in an early phase of development. Currently no lateral truncation is allowed and the acquisitions should originate from a full-scan. This prohibits an application to real weight-bearing acquisitions, and thus a common evaluation with marker- and registration-based methods. However, in Chap. 9 we show a thorough evaluation using the challenging FORBILD head phantom. A total of five types of representative translational 3-D motions have been simulated which are then corrected by estimating 2-D detector shifts. Results indicate a particularly accurate estimation of motion collinear with the scanner's rotation axis. The estimation performed reliably, even without incorporation of temporal assumptions, e. g., smoothness of the motion over time, thus allowing for a great variety of motion types. We show promising estimation results especially for high-frequency motion, leading to large improvements of visually observed but also quantitatively measured image quality. Yet, the method produced a small, but systematic deviation for translations perpendicular to the scanner's rotation axis, which requires further research. As a last evaluation of this thesis, we successfully applied the *FCC*-based correction to projection images of a simulated knee phantom with realistic patient motion.

In future work we aim to further reduce runtime and manual interaction of the registration-based motion correction which may even allow a complete replacement of the less flexible marker-based approach for all clinical applications. Another goal is to extend the *FCC*-based method for an application to real acquisitions. Given its performance on simulated data, we are confident that *FCC* is a future candidate for a completely self-contained motion estimation. Equipped with the motion correction methods presented in this thesis, weight-bearing *CBCT* now provides a valuable tool for novel and innovative clinical research in the field of orthopedics.

CHAPTER A

# Appendix

## A.1 Definition and Derivative of Rigid Motion

We defined  $\mathbf{T}_k(\boldsymbol{\alpha}) : \mathbb{R}^{6K} \rightarrow \mathbb{R}^{4 \times 4}$  to be a function that represents a rigid motion consisting of rotation and translation given the parameters in  $\boldsymbol{\alpha}$ . Let us assume the parameter vector  $\boldsymbol{\alpha} \in \mathbb{R}^{6K}$  holds 6 rigid parameters for each of  $K$  projection images, i. e., 3 Euler angles  $\phi_k^x, \phi_k^y,$  and  $\phi_k^z$  as well as 3 translations  $t_k^x, t_k^y,$  and  $t_k^z$ . The parameter vector is given by

$$\begin{aligned} \boldsymbol{\alpha} &= \left( \alpha_1 \ \alpha_2 \ \dots \ \alpha_{6K-1} \ \alpha_{6K} \right)^\top \\ &= \left( \phi_1^x \ \phi_1^y \ \phi_1^z \ t_1^x \ t_1^y \ t_1^z \ \dots \ \phi_K^x \ \phi_K^y \ \phi_K^z \ t_K^x \ t_K^y \ t_K^z \right)^\top \end{aligned}$$

From Eq. (3.34) we get

$$\mathbf{T}_k(\boldsymbol{\alpha}) = \begin{pmatrix} \mathbf{R}_k(\boldsymbol{\alpha}) & \mathbf{t}_k(\boldsymbol{\alpha}) \\ \mathbf{0}^\top & 1 \end{pmatrix},$$

where  $\mathbf{R}_k(\boldsymbol{\alpha}) : \mathbb{R}^{6K} \rightarrow \mathbb{R}^{3 \times 3}$  is a function building the 3-D rotation matrix and  $\mathbf{t}_k(\boldsymbol{\alpha}) : \mathbb{R}^{6K} \rightarrow \mathbb{R}^3$  returns the translations for the  $k$ -th projection.

The translations are defined straightforward by

$$\begin{aligned} \mathbf{t}_k(\boldsymbol{\alpha}) &= \left( t_k^x \ t_k^y \ t_k^z \right)^\top \\ &= \left( \alpha_{6(k-1)+4} \ \alpha_{6(k-1)+5} \ \alpha_{6(k-1)+6} \right)^\top. \end{aligned}$$

Rotations in 3-D may be defined by Euler angles, an axis-angle representation or by using quaternions [Gonz08]. We define  $\mathbf{R}_k(\boldsymbol{\alpha})$ , using an Euler angle representation. Let us first introduce rotation matrices around  $x$ ,  $y$ , and  $z$  axis, given by  $\mathbf{R}_k^x(\boldsymbol{\alpha}) : \mathbb{R}^{6K} \rightarrow \mathbb{R}^{3 \times 3}$ ,  $\mathbf{R}_k^y(\boldsymbol{\alpha}) : \mathbb{R}^{6K} \rightarrow \mathbb{R}^{3 \times 3}$  and  $\mathbf{R}_k^z(\boldsymbol{\alpha}) : \mathbb{R}^{6K} \rightarrow \mathbb{R}^{3 \times 3}$ , respectively. It holds that

$$\mathbf{R}_k(\boldsymbol{\alpha}) = \mathbf{R}_k^x(\boldsymbol{\alpha}) \cdot \mathbf{R}_k^y(\boldsymbol{\alpha}) \cdot \mathbf{R}_k^z(\boldsymbol{\alpha})$$

$$\mathbf{R}_k^x(\boldsymbol{\alpha}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi_k^x) & -\sin(\phi_k^x) \\ 0 & \sin(\phi_k^x) & \cos(\phi_k^x) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha_{6(k-1)+1}) & -\sin(\alpha_{6(k-1)+1}) \\ 0 & \sin(\alpha_{6(k-1)+1}) & \cos(\alpha_{6(k-1)+1}) \end{pmatrix}$$

$$\mathbf{R}_k^y(\boldsymbol{\alpha}) = \begin{pmatrix} \cos(\phi_k^y) & 0 & \sin(\phi_k^y) \\ 0 & 1 & 0 \\ -\sin(\phi_k^y) & 0 & \cos(\phi_k^y) \end{pmatrix} = \begin{pmatrix} \cos(\alpha_{6(k-1)+2}) & 0 & \sin(\alpha_{6(k-1)+2}) \\ 0 & 1 & 0 \\ -\sin(\alpha_{6(k-1)+2}) & 0 & \cos(\alpha_{6(k-1)+2}) \end{pmatrix}$$

$$\mathbf{R}_k^z(\boldsymbol{\alpha}) = \begin{pmatrix} \cos(\phi_k^z) & -\sin(\phi_k^z) & 0 \\ \sin(\phi_k^z) & \cos(\phi_k^z) & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos(\alpha_{6(k-1)+3}) & -\sin(\alpha_{6(k-1)+3}) & 0 \\ \sin(\alpha_{6(k-1)+3}) & \cos(\alpha_{6(k-1)+3}) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Building the partial derivatives of  $\mathbf{T}_k(\boldsymbol{\alpha})$  with respect to the parameters is given by

$$\frac{\partial \mathbf{T}_k(\boldsymbol{\alpha})}{\partial \alpha_l} = \begin{pmatrix} \frac{\partial \mathbf{R}_k(\boldsymbol{\alpha})}{\partial \alpha_l} & \frac{\partial \mathbf{t}_k(\boldsymbol{\alpha})}{\partial \alpha_l} \\ \mathbf{0}^\top & 1 \end{pmatrix},$$

where the derivative of the translations is obtained by

$$\frac{\partial \mathbf{t}_k(\boldsymbol{\alpha})}{\partial \alpha_l} = \begin{cases} \alpha_l \mathbf{e}_x & \text{if } l = 6(k-1) + 4 \\ \alpha_l \mathbf{e}_y & \text{if } l = 6(k-1) + 5 \\ \alpha_l \mathbf{e}_z & \text{if } l = 6(k-1) + 6 \\ \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^\top & \text{otherwise} \end{cases} ,$$

with  $\mathbf{e}_x$ ,  $\mathbf{e}_y$ , and  $\mathbf{e}_z$  being the directions of the world coordinate system with unit length. Similarly the derivatives for the rotations are given by

$$\begin{aligned} \frac{\partial \mathbf{R}_k(\boldsymbol{\alpha})}{\partial \alpha_l} &= \frac{\partial \mathbf{R}_k^x(\boldsymbol{\alpha})}{\partial \alpha_l} \cdot \mathbf{R}_k^y(\boldsymbol{\alpha}) \cdot \mathbf{R}_k^z(\boldsymbol{\alpha}) \\ &+ \mathbf{R}_k^x(\boldsymbol{\alpha}) \cdot \frac{\partial \mathbf{R}_k^y(\boldsymbol{\alpha})}{\partial \alpha_l} \cdot \mathbf{R}_k^z(\boldsymbol{\alpha}) \\ &+ \mathbf{R}_k^x(\boldsymbol{\alpha}) \cdot \mathbf{R}_k^y(\boldsymbol{\alpha}) \cdot \frac{\partial \mathbf{R}_k^z(\boldsymbol{\alpha})}{\partial \alpha_l} , \end{aligned}$$

where the derivatives for the individual rotation matrices are obtained by

$$\frac{\partial \mathbf{R}_k^x(\boldsymbol{\alpha})}{\partial \alpha_l} = \begin{cases} \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\sin(\alpha_l) & -\cos(\alpha_l) \\ 0 & \cos(\alpha_l) & -\sin(\alpha_l) \end{pmatrix} & \text{if } l = 6(k-1) + 1 \\ \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^\top \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} & \text{otherwise} \end{cases}$$

$$\frac{\partial \mathbf{R}_k^y(\boldsymbol{\alpha})}{\partial \alpha_l} = \begin{cases} \begin{pmatrix} -\sin(\alpha_l) & 0 & \cos(\alpha_l) \\ 0 & 0 & 0 \\ -\cos(\alpha_l) & 0 & -\sin(\alpha_l) \end{pmatrix} & \text{if } l = 6(k-1) + 2 \\ \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^\top \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} & \text{otherwise} \end{cases}$$

$$\frac{\partial \mathbf{R}_k^z(\boldsymbol{\alpha})}{\partial \alpha_l} = \begin{cases} \begin{pmatrix} -\sin(\alpha_l) & -\cos(\alpha_l) & 0 \\ \cos(\alpha_l) & -\sin(\alpha_l) & 0 \\ 0 & 0 & 0 \end{pmatrix} & \text{if } l = 6(k-1) + 3 \\ \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^\top \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} & \text{otherwise} \end{cases} .$$



# List of Acronyms

## **AEC**

automatic exposure control 19, 20, 94, 95, 110

## **BSpl**

B-splines 57, 115–117, 121

## **CBCT**

cone-beam computed tomography 1, 2, 5, 11, 14–18, 22–25, 28–30, 34, 35, 37, 39–42, 44, 59, 62, 69–72, 75, 89, 92, 93, 117, 120, 123, 126, 129, 141, 144, 145, 147, 148, 151–154, 163, 166, 168

## **CC**

consistency conditions 2, 39–42, 45, 69, 70, 91, 123, 141, 142, 149, 152

## **CONRAD**

CONe-beam in RADiology 89, 123

## **CT**

computed tomography 1, 2, 4, 5, 26, 28, 33, 36–42, 44, 70, 71, 87, 91, 92, 118, 119, 123, 125, 142, 145, 152

## **DFT**

discrete Fourier transform 71, 73–75, 83, 142, 163, 165–168

## **DID**

detector-isocenter-distance 25, 70, 89, 92–96, 153

## **DRR**

digitally reconstructed radiograph 38, 39, 60–63, 66, 67, 101, 104, 105, 120, 165, 168

## **ECC**

epipolar consistency conditions 42, 45

## **ECG**

Electrocardiogram 33, 37, 44

## **FBP**

filtered backprojection 26, 27, 29

**FCC**

Fourier consistency conditions 4, 5, 7, 41, 69–71, 74, 75, 78, 87, 89, 90, 93, 123, 125, 129–133, 135, 138–145, 148, 149, 153, 154, 166

**FDK**

Feldkamp-Davis-Kress 29, 43, 44, 52, 55, 66, 76

**FFT**

fast Fourier transform 73, 80–83, 142

**FOV**

field of view 4, 14, 15, 18, 19, 52, 66, 78, 94, 95, 100, 118, 125

**FP**

flat panel 15, 27

**FRE**

fiducial registration error 112, 114, 120

**FRST**

fast radial symmetry transform 2, 49–52, 100–102, 112, 117, 152, 165–167

**GC**

gradient correlation 39, 62, 63, 65, 100–105, 109–112, 114, 119, 120, 154, 163, 165

**GD**

gradient differences 39

**GPU**

graphics processing unit 38, 81–83, 102, 115

**HLCC**

Helgason-Ludwig consistency conditions 20, 40–42, 45, 168

**JSW**

joint-space width 13, 16, 17

**KLS**

Kellgren-Lawrence scale 13

**LinInt**

bilinear interpolation 57, 115–117, 121

**MAD**

mean absolute distance 124, 136, 138, 143

**MAM**

motion artifact metric 39, 40, 149

**MI**

mutual information 38, 109

**MIP**

maximum intensity projections 5, 38, 101, 118

**MRI**

magnetic resonance imaging 14, 17, 22, 39

**MSE**

mean squared error 87, 97

**NConv**

normalized convolution 58, 115–117, 121

**NGI**

normalized gradient information 39, 62–65, 100–114, 119–122, 154, 163, 164

**NIST**

National Institute of Standards and Technology 89

**OA**

osteoarthritis 12–14, 16, 17, 21, 22, 151

**PET**

positron emission tomography 41

**PFPS**

patellofemoral pain syndrome 12–14, 16, 17, 21, 22, 149, 151

**PI**

pattern intensity 38, 39

**RMSE**

root mean squared error 87, 97, 115, 116

**ROI**

region of interest 40, 63, 64, 98, 102, 103, 108–110, 115–117, 164, 167

**rRMSE**

relative root mean squared error 87, 97, 124, 129, 134, 135, 140, 142, 143

**SAS**

subtract-and-shift 58, 115–117, 121

**SDD**

source-detector-distance 1, 15, 16, 25, 87

**SID**

source-isocenter-distance 1, 16, 25, 70, 89, 92–96, 153

**SpecInt**

spectral interpolation 58, 115–117, 121, 122

**SPECT**

single photon emission computed tomography 41

**SSD**

sum-of-squared-differences 38, 58

**SSIM**

structural similarity index 87, 97, 98, 124, 134, 135, 140, 142, 143, 149, 163, 168

**SVD**

singular value decomposition 36

**TPS**

thin plate splines 5, 34, 35, 57, 61, 65, 66, 99, 101, 102, 109, 111, 112, 118, 120, 147, 148, 151, 152, 163, 165, 167

**TPSS**

thin plate smoothing splines 57, 115–117, 121

**UQI**

universal quality index 87, 97, 98, 108–112, 119, 120

**WEMS**

weighted edge matching score 39, 44

# List of Symbols

$\mathbf{A}$	Matrix holding weights for column-wise apodization of projection images
$A$	Number of TPS control points
$B$	Number of segmented bone volumes
$C$	Total number of pixels measured in a CBCT scan
$C_1, C_2$	Constants to improve stability of SSIM computation.
$D$	Detector-isocenter-distance
$\mathbf{D}^{uv}$	Function returning detector axes in world coordinates
$E_{max}$	Maximum energy of an X-ray spectrum
$E_p$	Energy of X-ray photons
$F$	Source-detector-distance
$\mathbf{F}$	Complex matrix that performs a 3-D DFT along $i, j$ and $k$
$\mathbf{F}_\lambda$	Complex matrix that performs a 1-D DFT along $k$
$\mathbf{F}_{uv}$	Complex matrix that performs a 2-D DFT along $i$ and $j$
$\mathbf{G}$	Kernel matrix function of TPS deformations
$GC$	Similarity measure GC
$GI$	Unnormalized part of NGI similarity measure
$\mathbf{H}(\boldsymbol{\alpha})$	Returns a diagonal matrix with phase shifts based on $\boldsymbol{\alpha}$
$\mathbf{I}_N$	$N$ -dimensional identity matrix
$I$	Number of pixels in horizontal detector direction
$I_v$	Function to obtain the maximum intensity range of an image
$J$	Number of pixels in vertical detector direction
$\mathbf{K}$	Intrinsic part of projection matrix $\mathbf{P}$
$K$	Number of projection images/matrices
$L$	Length of parameter vector $\boldsymbol{\alpha}$

$M$	Number of attached metallic markers
$M_{\min}$	Minimum of point correspondences per projection image after outlier removal
$N$	Arbitrary dimension of function and variables
$NGI$	Similarity measure <a href="#">NGI</a>
$N_{\text{opt}}$	Number of repeated outlier removal and optimization steps
$N_p$	Number of photons after absorption
$N_0$	Number of photons emitted from X-ray source
$N_w$	Side length of <a href="#">ROI</a> used for extrapolation of defective pixels
$\mathbf{P}$	Projection matrix to model mapping of world to image coordinates
$\mathbf{P}(\boldsymbol{\alpha})$	Projection matrix as function of parameter vector $\boldsymbol{\alpha}$
$\mathbf{Q}_k(\boldsymbol{\alpha})$	Projective mapping for 2-D motion of the $k$ -th projection
$\mathbf{R}$	Rotation matrix / World coordinate rotation in projection matrix $\mathbf{P}$
$\mathbf{R}(\boldsymbol{\alpha})$	Rotation matrix as function of parameter vector $\boldsymbol{\alpha}$
$\mathbf{R}_x(\phi^x)$	Rotation matrix around $x$ -axis by angle $\phi^x$
$\mathbf{R}_y(\phi^y)$	Rotation matrix around $y$ -axis by angle $\phi^y$
$\mathbf{R}_z(\phi^z)$	Rotation matrix around $z$ -axis by angle $\phi^z$
$S$	Source-isocenter-distance
$\mathbf{T}_b^{\text{init}}$	Projective mapping to align the $b$ -th bone to a motion-corrupted reconstruction
$\mathbf{T}_k(\boldsymbol{\alpha})$	Projective mapping for 3-D motion of the $k$ -th projection
$\mathbf{T}_{kb}(\boldsymbol{\alpha})$	Projective mapping for 3-D motion of the $k$ -th projection and $b$ -th bone
$U(\mathbf{x}, \lambda)$	Function to compute depth of $\mathbf{x}$ w. r. t. projection angle $\lambda$
$U_k$	Set containing assigned marker locations $\mathbf{u}_{km}$
$\tilde{U}_k$	Merged set of assigned marker locations, given the $k$ -th projection
$\hat{U}_k$	Set containing candidates $\mathbf{u}_k$ of possible marker locations for $k$ -th projection
$V$	Set of spatial locations used for quantitative image quality measures.
$\mathbf{W}$	Diagonal matrix to select spectral values in the triangular zero regions
$W_k$	Set of defect pixels for all markers on projection image $k$
$W_{km}$	Set of pixels on projection image $k$ considered as defect due to marker $m$

$X$	Set containing average 3-D marker locations $\mathbf{x}_m$
$\widehat{X}$	Temporary set of marker locations for assignment of 2-D correspondences
$\mathbf{Y}$	Normalization matrix of GC similarity measure
$\mathbf{a}$	Vector holding weights for column-wise apodization of projection images
$a$	Integer index for TPS control points
$b$	Integer index for bone segmentations
$\mathbf{b}_{ak}(\boldsymbol{\alpha})$	Function returning the TPS weight vectors based on parameters in $\boldsymbol{\alpha}$ .
$\mathbf{c}$	Location of the X-ray source in world coordinate system
$c_1$	Amount of constant shift of all projection images in $u$ direction
$c_2$	Amount of constant shift of all projection images in $v$ direction
$c_{\text{rad}}$	FRST parameter for radial strictness
$c_u$	Horizontal coordinate of principle point in pixel coordinates
$c_v$	Vertical coordinate of principle point in pixel coordinates
$d$	Function that represents the resulting DRR images, given parameter vector $\boldsymbol{\alpha}$
$\text{diag}(\cdot)$	Builds a diagonal matrix with the entries of its input vector on the diagonal
$\mathbf{d}_k(\mathbf{x}, \boldsymbol{\alpha})$	Deformation field: Returns an $N$ -dimensional translation for location $\mathbf{x}$ .
$d_u$	Maximum allowed distance of 2-D detection and forward projected 3-D marker
$e(\boldsymbol{\alpha})$	Generic representation of a cost function, based on parameters in $\boldsymbol{\alpha}$
$\mathbf{e}_x$	$x$ -axis direction in world coordinate system
$\mathbf{e}_y$	$y$ -axis in world coordinate system
$\mathbf{e}_z$	$z$ -axis in world coordinate system
$f$	Object function: Describes the imaged object and its reconstruction
freq	Function to compute discrete axis values after applying a DFT
$g$	Function for projection data of fan-beam geometry
$g_{\sigma}^N()$	$N$ -dimensional isotropic Gaussian low-pass filter
$h$	Function to map homogeneous to Euclidean coordinates
$h_{\text{ramp}}$	Ramp filter kernel for fan- and cone-beam geometries
$i$	Integer index for pixels of horizontal detector direction

inc	FRST image function for increments
$j$	Integer index for pixels of vertical detector direction
$k$	Integer index for projection images/matrices
$k_s$	Intrinsic parameter: Detector skew
$l$	Integer index of parameter vector $\alpha$
$\hat{l}$	Computes the linear index in parameter vector $\alpha$ , given $b$ and $k$
$m$	Integer index of metallic markers
mag	FRST image function for gradient magnitude
$\mathbf{m}_{km}$	Projection of $m$ -th 3-D marker location onto the $k$ projection image
$m_{\text{oscil}}$	Function that defines an oscillating, translational motion ( <i>Oscil</i> ).
$n$	Integer index for DFT / Order of Fourier series
$n_{\text{it}}$	Iteration number in FCC-based motion estimation
$o$	Discrete Fourier coefficient after 3-D DFT of the shifted projection images
$o(\alpha)$	Inner part of L2 norm in FCC-based cost function
$p$	Function for CBCT projection data
$\mathbf{p}$	Vector that holds all measured pixel values of the CBCT projection data
$q$	Sinogram function of parallel-beam geometry
$\mathbf{r}$	Unit vector denoting the direction of an X-ray
$r_f$	FRST radius parameter
$r_p$	Maximum extent of the object w. r. t. the system's isocenter
$r_{\text{tmp}}$	Temporal regularization function for 2-D/3-D registration
$r_w$	Radius used around marker locations to define defective pixels
$s_f$	FRST scaling parameter
$s_k$	Shift of $k$ -th projection image in detector $u$ direction
$\hat{s}_k$	Ground-truth translation in detector $u$ direction, for a selected projection index
$\tilde{s}_k$	Fixed translation in detector $u$ direction, for a selected projection index
$\mathbf{t}$	Translation / World coordinate translation in projection matrix $\mathbf{P}$
$\hat{t}_k$	Ground-truth translation in detector $v$ direction, for a selected projection index

$\tilde{t}_k$	Fixed translation in detector $v$ direction, for a selected projection index
$t_k$	Shift of $k$ -th projection image in detector $v$ direction
$\mathbf{t}(\boldsymbol{\alpha})$	Translation as function of parameter vector $\boldsymbol{\alpha}$
$t_{\text{proj}}$	FRST binarization threshold
$t_{\text{rec}}$	Binarization threshold after backprojection of FRST result
$\mathbf{u}$	Projection image location in pixel coordinates
$u$	Horizontal axis of detector
$\mathbf{u}_k$	Candidate for a marker location on $k$ -th projection image
$\mathbf{u}_{km}$	Detected location on $k$ -th projection image assigned to $m$ -th marker
$u_p$	Detector axis of parallel-beam geometry
$v$	Vertical axis of detector
$\mathbf{v}_{ak}$	$a$ -th TPS control point location for $k$ -th projection image
$v(\mathbf{x})$	Reference image function for quantitative image quality measures.
$\mathbf{w}$	Vector that encodes the spectral locations of the triangular zero regions
$w_k(\mathbf{u})$	Binary mask image returning only 0 if location $\mathbf{u}$ defect
$\mathbf{x}$	Location in world coordinate system (2-D or 3-D)
$x$	$x$ coordinate in world coordinate system
$\mathbf{x}_m$	Average 3-D marker locations of $m$ -th marker
$y$	$y$ coordinate in world coordinate system
$z$	$z$ coordinate in world coordinate system
$\Delta\lambda$	Sampling interval for rotation angles $\lambda$
$\Delta u$	Sampling interval for detector $u$ direction
$\Delta v$	Sampling interval for detector $v$ direction
$\Omega_{km}$	Set of valid pixels in neighborhood of $m$ -th marker on $k$ -th projection image
$\Phi$	Discrete Fourier coefficient after 2-D DFT of the original projection images
$\Pi$	Set to define ROIs in projection domain for similarity measure computation
$\boldsymbol{\alpha}$	Parameter vector: Contains parameter that describe the motion model
$\beta$	Differential value of X-ray transforms

$\zeta$	Percentile used for removal of worst point correspondences
$\epsilon$	Distance of adjusted to the original triangle boundaries
$\eta$	Lagrangian multiplier to incorporate regularization terms into a cost function
$\gamma$	Fan angle of an X-ray direction in fan- and cone-beam geometry
$\gamma'$	Maximum fan angle of an X-ray direction in fan- and cone-beam geometry
$\kappa$	Cone angle of an X-ray direction in CBCT
$\kappa'$	Maximum cone angle of an X-ray direction in CBCT
$\lambda$	Rotation angle for fan- and cone-beam geometry
$\lambda_p$	Wavelength of X-ray photons
$\mu$	Linear attenuation coefficient to compute X-ray absorption
$\mu_f$	The mean of the object function used for SSIM computation.
$\mu_v$	The mean of the reference function used for SSIM computation.
$\nabla d$	Gradient of generated DRR images
$\nabla f$	Spatial gradient of object function
$\nabla p$	Gradient of projection images
$\nu$	Order of geometric moments in HLCC
$\omega$	Axis in Fourier domain corresponding to the rotation angles $\lambda$
$\omega_k$	Discrete values of Fourier axis w. r. t. rotation angles
$\varphi$	Discrete Fourier coefficient after 2-D DFT of the shifted projection images
$\psi$	Axis in Fourier domain corresponding to detector $v$ direction
$\psi_k$	Partial derivatives of phase factors for projection $k$
$\psi_j$	Discrete values of Fourier axis w. r. t. detector $u_2$ direction
$\rho$	Density of materials
$\sigma_f$	The standard deviation of the object function used for SSIM computation.
$\sigma_{fv}$	Correlation coefficient used for SSIM computation.
$\sigma_v$	The standard deviation of the reference function used for SSIM computation.
$\tau_k$	Normalized temporal variable to describe simulated motion patterns
$\theta$	Rotation angle for parallel-beam geometry

$\mathcal{V}_b^a$	$a$ -th vertex location for $b$ -th bone mesh
$\xi$	Axis in Fourier domain corresponding to detector $u$ direction
$\xi_i$	Discrete values of Fourier axis w. r. t. detector $u_1$ direction
$\zeta_{kb}$	Vector with six rigid motion parameters for $k$ -th projection and $b$ -th bone



# List of Figures

1.1	Structural overview of this thesis. . . . .	6
2.1	Anatomy of the knee joint with bones, tendons and ligaments. . . . .	12
2.2	Muscles of lower extremities: The hamstring and quadriceps group. . . . .	13
2.3	Prototype and current CBCT weight-bearing scanner. . . . .	15
2.4	Truncation and limited FOV in posteroanterior view direction. . . . .	18
2.5	AEC and saturation artifacts in CBCT weight-bearing imaging. . . . .	19
2.6	Artifacts due to patient motion in CBCT weight-bearing imaging. . . . .	21
3.1	Illustration of the fan-beam geometry. . . . .	26
3.2	Illustration of the cone-beam geometry. . . . .	28
3.3	Definition of motion correction. . . . .	32
4.1	Overview of the fully automatic marker detection approach. . . . .	50
4.2	Comparison of circular Hough transform and FRST. . . . .	51
4.3	Example of streaking artifacts caused by metallic tantalum beads. . . . .	56
5.1	Outline of the registration-based motion correction approach. . . . .	60
5.2	Overview of a 2-D/3-D registration method. . . . .	61
5.3	Variation to 3-D segmentations to reduce noise in DRRs. . . . .	67
6.1	Example of a sinogram and the triangular regions of the FCC. . . . .	70
6.2	Adjustment of triangular regions to account for discretization effects. . . . .	77
6.3	The influence of truncation and apodization on FCC. . . . .	78
6.4	Adjusted apodization window to allow for translations. . . . .	79
6.5	Approximation of adjusted Blackman-Harris window. . . . .	80
8.1	Reconstructions of supine scans with ROIs for UQI computation. . . . .	102
8.2	Axial slices through femur and tibia from Subject 2. . . . .	103
8.3	Axial slices through tibia and fibula from Subject 2. . . . .	104
8.4	Visualization of 2-D/3-D registration result in projection domain. . . . .	105
8.5	Axial slices through femur and tibia from Subject 3. . . . .	106
8.6	Sagittal slices of Subject 3, showing femoral and tibial cartilage. . . . .	107
8.7	Edge profiles at the surface of a femur in S2_WB0 and S2_SUP. . . . .	113
8.8	Reconstruction results of different marker removal methods. . . . .	116
8.9	Results for marker removal based on a real acquisition. . . . .	117
9.1	Reconstructed images for high- and low-quality fan-beam data. . . . .	127
9.2	Reconstructions of the phantom's resolution pattern and ear. . . . .	128
9.3	Sinograms and spectra for the fan-beam dataset. . . . .	128
9.4	Axial slices for high-frequency motion types and motion-free data. . . . .	130
9.5	Axial slices of reconstructions for low-frequency motion types. . . . .	131

9.6	Sinograms and Fourier domain for high-frequency motion. . . . .	132
9.7	Sinograms and Fourier domain for low-frequency motion. . . . .	133
9.8	Reconstructions, sinograms and spectra for noisy data. . . . .	134
9.9	Estimated and ground-truth parameters for detector translations. . .	137
9.10	Difference of estimated and ground-truth horizontal translations. . . .	138
9.11	Systematic bias in estimated horizontal translations. . . . .	138
9.12	Reconstruction results for the XCAT dynamic squat phantom. . . . .	140
9.13	Sinograms and spectra for the XCAT dynamic squat phantom. . . . .	141

# List of Tables

7.1	Overview of acquired and simulated data and their abbreviations. . . . .	88
7.2	Simulation parameters for bead removal phantom. . . . .	89
7.3	Fan-beam simulation parameters for FORBILD phantom. . . . .	92
7.4	Cone-beam simulation parameters for FORBILD phantom. . . . .	92
7.5	Simulation parameters for XCAT dynamic squat phantom. . . . .	93
7.6	Acquisition parameters for the supine scan of Subject 1. . . . .	94
7.7	Parameters used for acquisitions of Subject 2. . . . .	95
7.8	Parameters used for acquisitions of Subject 3. . . . .	96
8.1	Parameter settings for marker-based motion correction. . . . .	101
8.2	UQI values over evaluated bone ROIs. . . . .	110
8.3	FRE measured with reprojected marker locations. . . . .	114
8.4	Comparison of extrapolation methods for marker removal. . . . .	116
9.1	rRMSE for high- and low-quality fan-beam reconstructions. . . . .	129
9.2	Quantitative evaluation of image quality for all motion cases. . . . .	135
9.3	MAD between estimated and ground-truth motion parameters. . . . .	136



# List of Algorithms

4.1	Algorithm to extract point correspondences. . . . .	54
6.1	Outline of a cost function evaluation for the $n_{it}$ -th iteration. . . . .	82
8.1	Pipeline for quantitative comparison of image quality at bones. . . . .	108



# Bibliography

- [Aach 01] T. Aach and V. H. Metzler. “Defect interpolation in digital radiography: how object-oriented transform coding helps”. In: *Proc. SPIE 4322*, pp. 824–835, 2001.
- [Abel 09] M. Abella, J. J. Vaquero, M. L. Soto-Montenegro, E. Lage, and M. Desco. “Sinogram bow-tie filtering in FBP PET reconstruction”. *Medical Physics*, Vol. 36, No. 5, pp. 1663–1671, May 2009.
- [Aich 14] A. Aichert, N. Maass, Y. Deuerling-Zheng, M. Berger, M. Manhart, J. Hornegger, A. K. Maier, and A. Doerfler. “Redundancies in X-ray images due to the epipolar geometry for transmission imaging”. In: F. Noo, Ed., *Proc. The third international conference on image formation in x-ray computed tomography*, pp. 333–337, Jun 2014.
- [Aich 15] A. Aichert, M. Berger, J. Wang, N. Maass, A. Doerfler, J. Hornegger, and A. K. Maier. “Epipolar Consistency in Transmission Imaging”. *IEEE Transactions on Medical Imaging*, Vol. 34, No. 11, pp. 2205–2219, Nov 2015.
- [Altm 87] R. D. Altman, J. F. Fries, D. A. Bloch, J. Carstens, T. C. Derek Mb, H. Genant, P. Gofton, H. Groth, D. J. Mcshane, W. A. Murphy, J. T. Sharp, P. Spitz, C. A. Williams, and F. Wolfe. “Radiographic assessment of progression in osteoarthritis”. *Arthritis & Rheumatism*, Vol. 30, No. 11, pp. 1214–1225, 1987.
- [Baue 12] S. Bauer, B. Berkels, S. Ettl, O. Arold, J. Hornegger, and M. Rumpf. “Marker-less Reconstruction of Dense 4-D Surface Motion Fields using Active Laser Triangulation for Respiratory Motion Management”. In: N. Ayache, H. Delingette, P. Golland, and K. Mori, Eds., *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*, pp. 414–421, Berlin, Germany, Oct 2012.
- [Baue 13] S. Bauer, A. Seitel, H. Hofmann, T. Blum, J. Wasza, M. Balda, H.-P. Meinzer, N. Navab, J. Hornegger, and L. Maier-Hein. “Real-Time Range Imaging in Health Care: A Survey”. In: M. Grzegorzec, C. Theobalt, R. Koch, and A. Kolb, Eds., *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*, pp. 228–254, Springer, Berlin, Germany, 2013.
- [Berg 14a] M. Berger, C. Forman, C. Schwemmer, J. H. Choi, K. Müller, A. Maier, J. Hornegger, and R. Fahrig. “Automatic Removal of Externally Attached Fiducial Markers in Cone Beam C-arm CT”. In: T. Deserno, H. Handels, H.-P. Meinzer, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2014*, pp. 318–323, Mar 2014.
- [Berg 14b] M. Berger, A. Maier, Y. Xia, J. Hornegger, and R. Fahrig. “Motion Compensated Fan-Beam CT by Enforcing Fourier Properties of the Sinogram”. In: F. Noo, Ed., *Proc. The third international conference on image formation in x-ray computed tomography*, pp. 329–332, Jun 2014.

- [Berg 14c] M. Berger, K. Sembritzki, J. Hornegger, and C. Bauer. “Increasing the Credibility of MR Spectroscopy-Based Automatic Brain-Tumor Classification Systems”. In: IEEE, Ed., *Proc. 2014 IEEE International Symposium on Biomedical Imaging*, pp. 345–348, May 2014.
- [Berg 15] M. Berger, K. Müller, J.-H. Choi, A. Aichert, A. Maier, and R. Fahrig. “2D/3D Registration for Motion Compensated Reconstruction in Cone-Beam CT of Knees Under Weight-Bearing Condition”. In: D. A. Jaffray, Ed., *IFMBE Proceedings*, pp. 54–57, Jun 2015.
- [Berg 16a] M. Berger, K. Müller, A. Aichert, M. Unberath, J. Thies, J.-H. Choi, R. Fahrig, and A. Maier. “Marker-free motion correction in weight-bearing cone-beam CT of the knee joint”. *Medical Physics*, Vol. 43, No. 3, pp. 1235–1248, Mar 2016.
- [Berg 16b] M. Berger, Y. Xia, W. Aichinger, K. Mentl, M. Unberath, A. Aichert, C. Riess, J. Hornegger, R. Fahrig, and A. Maier. “Translational Motion Compensation for Cone-Beam CT Using Fourier Consistency Conditions”. *Physics in Medicine and Biology*, 2016. (under revision).
- [Besi 05] T. F. Besier, C. E. Draper, G. E. Gold, G. S. Beaupré, and S. L. Delp. “Patellofemoral joint contact area increases with knee flexion and weight-bearing”. *Journal of Orthopaedic Research*, Vol. 23, No. 2, pp. 345–350, Mar 2005.
- [Besl 92] P. J. Besl and N. D. McKay. “Method for registration of 3-D shapes”. In: *Proc. SPIE 1611*, pp. 586–606, 1992.
- [Bier 16a] B. Bier, M. Berger, J. Maier, M. Unberath, S. Hsieh, S. Bonaretti, R. Fahrig, M. E. Levenston, G. E. Gold, and A. Maier. “Object Removal in Gradient Domain of Cone-Beam CT Projections”. In: *Proc. IEEE Nuclear Science Symposium & Medical Imaging Conference (NSS/MIC)*, Nov 2016.
- [Bier 16b] B. Bier, K. Müller, M. Berger, J.-H. Choi, R. Fahrig, and A. Maier. “Scatter Correction for C-Arm CT Using Primary Modulation”. In: M. Kachelries, Ed., *Proc. The fourth international conference on image formation in x-ray computed tomography*, pp. 383–386, Jun 2016.
- [Bifu 02] P. Bifulco, M. Sansone, M. Cesarelli, R. Allen, and M. Bracale. “Estimation of out-of-plane vertebra rotations on radiographic projections using CT data: a simulation study”. *Medical Engineering & Physics*, Vol. 24, No. 4, pp. 295–300, May 2002.
- [Bifu 09] P. Bifulco, M. Cesarelli, R. Allen, M. Romano, A. Fratini, and G. Pasquariello. “2D-3D Registration of CT Vertebra Volume to Fluoroscopy Projection: A Calibration Model Assessment”. *EURASIP Journal on Advances in Signal Processing*, Vol. 2010, No. 1, p. 806094, Jun 2009.
- [Blon 06] C. Blondel, G. Malandain, R. Vaillant, and N. Ayache. “Reconstruction of coronary arteries from a single rotational X-ray projection sequence”. *IEEE Transactions on Medical Imaging*, Vol. 25, No. 5, pp. 653–663, May 2006.
- [Boli 10] M. Boling, D. Padua, S. Marshall, K. Guskiewicz, S. Pyne, and A. Beutler. “Gender differences in the incidence and prevalence of patellofemoral pain syndrome”. *Scandinavian Journal of Medicine & Science in Sports*, Vol. 20, No. 5, pp. 725–730, Oct 2010.

- [Brok 06] J. Brokish and Y. Bresler. “Sampling Requirements for Circular Cone Beam Tomography”. In: *Proc. IEEE Nuclear Science Symposium & Medical Imaging Conference (NSS/MIC)*, pp. 2882–2884, Nov 2006.
- [Buzu 08] T. M. Buzug. *Computed Tomography: From Photon Statistics to Modern Cone-Beam CT*. Springer, Berlin, Germany, 2008.
- [Cao 15] Q. Cao, G. Thawait, G. J. Gang, W. Zbijewski, T. Reigel, T. Brown, B. Corner, S. Demehri, and J. H. Siewerdsen. “Characterization of 3D joint space morphology using an electrostatic model (with application to osteoarthritis)”. *Physics in Medicine and Biology*, Vol. 60, No. 3, pp. 947–960, Feb 2015.
- [Carr 14] J. A. Carrino, A. Al Muhit, W. Zbijewski, G. K. Thawait, J. W. Stayman, N. Packard, R. Senn, D. Yang, D. H. Foos, J. Yorkston, and J. H. Siewerdsen. “Dedicated Cone-Beam CT System for Extremity Imaging”. *Radiology*, Vol. 270, No. 3, pp. 816–824, Mar 2014.
- [Chen 05] G.-H. Chen and S. Leng. “A new data consistency condition for fan-beam projection data”. *Medical Physics*, Vol. 32, No. 4, pp. 961–967, Apr 2005.
- [Chen 08] G.-H. Chen, J. Tang, and S. Leng. “Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets”. *Medical Physics*, Vol. 35, No. 2, pp. 660–663, 2008.
- [Choi 12] J. H. Choi, A. Maier, A. Keil, S. Pal, E. McWalter, and R. Fahrig. “Fiducial marker-based motion compensation for the acquisition of 3D knee geometry under weight-bearing conditions using a C-arm CT scanner”. In: AAPM, Ed., *Proc. 54th AAPM Annual Meeting*, Medical Physics Publishing, Charlotte, NC, USA, 2012.
- [Choi 13] J.-H. Choi, R. Fahrig, A. Keil, T. F. Besier, S. Pal, E. J. McWalter, G. S. Beaupré, and A. Maier. “Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees. Part I. Numerical model-based optimization”. *Medical Physics*, Vol. 40, No. 9, p. 091905, Sep 2013.
- [Choi 14a] J.-H. Choi. *Acquisition of 3D Knee Morphology Under Weight-Bearing Conditions Using a C-arm CT Scanner for the Assessment of Knee Disorders*. PhD thesis, Stanford University, 2014.
- [Choi 14b] J.-H. Choi, A. Maier, M. Berger, and R. Fahrig. “Effective one step-iterative fiducial marker-based compensation for involuntary motion in weight-bearing C-arm cone-beam CT scanning of knees”. In: *Proc. SPIE 9033*, p. 903312, 2014.
- [Choi 14c] J.-H. Choi, A. Maier, A. Keil, S. Pal, E. J. McWalter, G. S. Beaupré, G. E. Gold, and R. Fahrig. “Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees. II. Experiment”. *Medical Physics*, Vol. 41, No. 6, p. 061902, Jun 2014.
- [Choi 16a] J.-H. Choi, E. McWalter, S. Datta, K. Müller, A. Maier, G. Gold, M. Levenston, and R. Fahrig. “In Vivo 3D Measurement of Time-dependent Human Knee Joint Compression and Cartilage Strain During Static Weight-Bearing”. In: Orthopaedic Research Society, Ed., *Proc. ORS 2016 Annual Meeting*, Mar 2016.

- [Choi 16b] J.-H. Choi, K. Müller, S. Hsieh, A. Maier, G. Gold, M. Levenston, and R. Fahrig. “Over-exposure correction in knee cone-beam CT imaging with automatic exposure control using a partial low dose scan”. In: *Proc. SPIE 9783*, p. 97830L, 2016.
- [Clac 13] R. Clackdoyle and L. Desbat. “Full data consistency conditions for cone-beam projections with sources on a plane”. *Physics in Medicine and Biology*, Vol. 58, No. 23, pp. 8437–8456, Dec 2013.
- [Clac 15] R. Clackdoyle and L. Desbat. “Data consistency conditions for truncated fanbeam and parallel projections”. *Medical Physics*, Vol. 42, No. 2, pp. 831–845, Feb 2015.
- [Clar 06] U. Clarenz, M. Droske, S. Henn, M. Rumpf, and K. Witsch. “Computational Methods for Nonlinear Image Registration”. In: O. Scherzer, Ed., *Mathematical Models for Registration and Applications to Medical Imaging*, pp. 81–101, Springer, Berlin, Germany, 2006.
- [Corm 63] A. M. Cormack. “Representation of a function by its line integrals, with some radiological applications”. *Journal of Applied Physics*, Vol. 34, No. 9, pp. 2722–2727, Apr 1963.
- [Coto 11] S. Cotofana, F. Eckstein, W. Wirth, R. B. Souza, X. Li, B. Wyman, M.-P. Hellio-Le Graverand, T. Link, and S. Majumdar. “In vivo measures of cartilage deformation: patterns in healthy and osteoarthritic female knees using 3T MR imaging”. *European Radiology*, Vol. 21, No. 6, pp. 1127–1135, Jun 2011.
- [Daum 11] V. Daum. *Model-Constrained Non-Rigid Registration in Medicine*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2011.
- [Davi 97] M. Davis, A. Khotanzad, D. Flamig, and S. Harms. “A physics-based coordinate transformation for 3-D image matching”. *IEEE Transactions on Medical Imaging*, Vol. 16, No. 3, pp. 317–328, Jun 1997.
- [Debb 13] C. Debbeler, N. Maass, M. Elter, F. Dennerlein, and T. M. Buzug. “A new CT rawdata redundancy measure applied to automated misalignment correction”. In: R. M. Leahy and J. Qi, Eds., *Proc. Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, pp. 264–267, 2013.
- [Denn 13] F. Dennerlein and A. Maier. “Approximate truncation robust computed tomography – ATTRACT”. *Physics in Medicine and Biology*, Vol. 58, No. 17, pp. 6133–6148, Sep 2013.
- [Desb 04] L. Desbat, S. Roux, P. Grangeat, and A. Koenig. “Sampling conditions of 3D parallel and fan-beam x-ray CT with application to helical tomography”. *Physics in Medicine and Biology*, Vol. 49, No. 11, pp. 2377–2390, Jun 2004.
- [Desb 07] L. Desbat, S. Roux, and P. Grangeat. “Compensation of some time dependent deformations in tomography”. *IEEE transactions on medical imaging*, Vol. 26, No. 2, pp. 261–269, Feb 2007.
- [Desj 04] B. Desjardins and E. A. Kazerooni. “ECG-gated Cardiac CT”. *American Journal of Roentgenology*, Vol. 182, No. 4, pp. 993–1010, Apr 2004.

- [Drap 09] C. E. Draper, T. F. Besier, J. M. Santos, F. Jennings, M. Fredericson, G. E. Gold, G. S. Beaupre, and S. L. Delp. “Using real-time MRI to quantify altered joint kinematics in subjects with patellofemoral pain and to evaluate the effects of a patellar brace or sleeve on joint motion”. *Journal of Orthopaedic Research*, Vol. 27, No. 5, pp. 571–577, May 2009.
- [Drap 11] C. E. Draper, T. F. Besier, M. Fredericson, J. M. Santos, G. S. Beaupre, S. L. Delp, and G. E. Gold. “Differences in patellofemoral kinematics between weight-bearing and non-weight-bearing conditions in patients with patellofemoral pain”. *Journal of Orthopaedic Research*, Vol. 29, No. 3, pp. 312–317, Mar 2011.
- [Dye 05] S. F. Dye. “The Pathophysiology of Patellofemoral Pain: A Tissue Homeostasis Perspective”. *Clinical Orthopaedics and Related Research*, Vol. 436, pp. 100–110, Jul 2005.
- [Ecks 06] F. Eckstein, F. Cicuttini, J.-P. Raynauld, J. C. Waterton, and C. Peterfy. “Magnetic resonance imaging (MRI) of articular cartilage in knee osteoarthritis (OA): morphological assessment”. *Osteoarthritis and Cartilage*, Vol. 14, Supplement 1, pp. 46–75, 2006.
- [Edho 86] P. R. Edholm, R. M. Lewitt, and B. Lindholm. “Novel Properties Of The Fourier Decomposition Of The Sinogram”. In: *Proc. SPIE 0671*, pp. 8–18, 1986.
- [El K 04] G. Y. El-Khoury, K. J. Alliman, H. J. Lundberg, M. J. Rudert, T. D. Brown, and C. L. Saltzman. “Cartilage Thickness in Cadaveric Ankles: Measurement with Double-Contrast Multi-Detector Row CT Arthrography versus MR Imaging”. *Radiology*, Vol. 233, No. 3, pp. 768–773, Dec 2004.
- [Engl 12] M. Englund, F. W. Roemer, D. Hayashi, M. D. Crema, and A. Guermazi. “Meniscus pathology, osteoarthritis and the treatment controversy”. *Nature Reviews Rheumatology*, Vol. 8, No. 7, pp. 412–419, May 2012.
- [Erla 00] K. Erlandsson, D. Visvikis, W. Waddington, and P. Jarritt. “Truncation reduction in fan-beam transmission scanning using the Radon transform consistency conditions”. *IEEE Transactions on Nuclear Science*, Vol. 47, No. 3, pp. 989–993, Jun 2000.
- [Feld 84] L. A. Feldkamp, L. C. Davis, and J. W. Kress. “Practical cone-beam algorithm”. *Journal of the Optical Society of America A*, Vol. 1, No. 6, pp. 612–619, Jun 1984.
- [Fran 06] M. Franz and M. Stamminger. “2D-3D-Registration in Computer Tomography without an initial pose”. In: *Proc. Vision, Modeling and Visualization*, pp. 229–236, 2006.
- [Fred 06] M. Fredericson and K. Yoon. “Physical Examination and Patellofemoral Pain Syndrome”. *American Journal of Physical Medicine & Rehabilitation*, Vol. 85, No. 3, pp. 234–243, Mar 2006.
- [Frys 15] R. Frysch and G. Rose. “Rigid Motion Compensation in Interventional C-arm CT Using Consistency Measure on Projection Data”. In: N. Navab, J. Hornegger, M. W. Wells, and F. A. Frangi, Eds., *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pp. 298–306, Springer, Berlin, Germany, Oct 2015.

- [Fulk 02] J. P. Fulkerson. “Diagnosis and Treatment of Patients with Patellofemoral Pain”. *American Journal of Sports Medicine*, Vol. 30, No. 3, pp. 447–456, May 2002.
- [Geim 16] T. Geimer, M. Unberath, O. Taubmann, C. Bert, and A. Maier. “Combination of Markerless Surrogates for Motion Estimation in Radiation Therapy”. In: *Proc. Computer Assisted Radiology and Surgery (CARS)*, Jun 2016.
- [Gend 11] C. Gendrin, P. Markelj, S. A. Pawiro, J. Spoerk, C. Bloch, C. Weber, M. Figl, H. Bergmann, W. Birkfellner, B. Likar, and F. Pernus. “Validation for 2D/3D registration. II: The comparison of intensity- and gradient-based merit functions using a new gold standard data set”. *Medical Physics*, Vol. 38, No. 3, pp. 1491–1502, Mar 2011.
- [Gend 12] C. Gendrin, H. Furtado, C. Weber, C. Bloch, M. Figl, S. A. Pawiro, H. Bergmann, M. Stock, G. Fichtinger, D. Georg, and W. Birkfellner. “Monitoring tumor motion by real time 2D/3D registration during radiotherapy”. *Radiotherapy and Oncology*, Vol. 102, No. 2, pp. 274–280, Feb 2012.
- [Goit 83] M. Goitein, M. Abrams, D. Rowell, H. Pollari, and J. Wiles. “Multi-dimensional treatment planning: II. Beam’s eye-view, back projection, and projection through CT sections”. *International Journal of Radiation Oncology \* Biology \* Physics*, Vol. 9, No. 6, pp. 789–797, Jun 1983.
- [Gonz 08] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Prentice Hall, Fort Collins, CO, USA, 3 Ed., 2008.
- [Gou 16] Z. Gou, F. Noo, A. Maier, and G. Lauritsch. “Properties of the ellipse-line-ellipse trajectory with asymmetrical variations”. In: *Proc. SPIE 9783*, p. 97832Z, 2016.
- [Gran 91] P. Grangeat. “Mathematical framework of cone beam 3D reconstruction via the first derivative of the Radon transform”. In: G. T. Herman, A. K. Louis, and F. Natterer, Eds., *Mathematical Methods in Tomography*, pp. 66–97, Springer, Berlin, Germany, 1991.
- [Groh 08] M. Groher. *2D-3D Registration of Vascular Images*. PhD thesis, Technische Universität München, 2008.
- [Haas 16] V. Haase, O. Taubmann, Y. Huang, G. Krings, G. Lauritsch, A. Maier, and A. Mertins. “Make the Most of Time: Temporal Extension of the iTV Algorithm for 4D Cardiac C-Arm CT”. In: T. Deserno, H. Handels, H.-P. Meinzer, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2016*, pp. 170–175, Mar 2016.
- [Hans 08] E. Hansis, D. Schäfer, O. Dössel, and M. Grass. “Projection-based motion compensation for gated coronary artery reconstruction from rotational x-ray angiograms”. *Physics in Medicine and Biology*, Vol. 53, No. 14, pp. 3807–3820, Jul 2008.
- [Hara 09] A. K. Hara, R. G. Paden, A. C. Silva, J. L. Kujak, H. J. Lawder, and W. Pavlicek. “Iterative Reconstruction Technique for Reducing Body Radiation Dose at CT: Feasibility Study”. *American Journal of Roentgenology*, Vol. 193, No. 3, pp. 764–771, Sep 2009.

- [Harr 78] F. J. Harris. “On the use of windows for harmonic analysis with the discrete Fourier transform”. *Proceedings of the IEEE*, Vol. 66, No. 1, pp. 51–83, Jan 1978.
- [Hart 04] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, Great Britain, 2 Ed., 2004.
- [Hawk 88] W. G. Hawkins, P. K. Leichner, and N.-C. Yang. “The circular harmonic transform for SPECT reconstruction and boundary conditions on the Fourier transform of the sinogram”. *IEEE Transactions on Medical Imaging*, Vol. 7, No. 2, pp. 135–138, Jun 1988.
- [Heim 09] T. Heimann and H.-P. Meinzer. “Statistical shape models for 3D medical image segmentation: A review”. *Medical Image Analysis*, Vol. 13, No. 4, pp. 543–563, Aug 2009.
- [Helg 11] S. Helgason. *Integral Geometry and Radon Transforms*. Springer New York, New York, NY, USA, 2011.
- [Helg 80] S. Helgason. “Support of Radon transforms”. *Advances in Mathematics*, Vol. 38, No. 1, pp. 91–100, Oct 1980.
- [Herb 14] M. Herbst, F. Schebesch, M. Berger, R. Fahrig, J. Hornegger, and A. Maier. “Improved trajectories in C-Arm computed tomography for non-circular fields of view”. In: F. Noo, Ed., *Proc. The third international conference on image formation in x-ray computed tomography*, pp. 274–278, Jun 2014.
- [Herb 15] M. Herbst, F. Schebesch, M. Berger, J.-H. Choi, R. Fahrig, J. Hornegger, and A. Maier. “Dynamic detector offsets for field of view extension in C-arm computed tomography with application to weight-bearing imaging”. *Medical Physics*, Vol. 42, No. 5, pp. 2718–2729, May 2015.
- [Hopp 09] S. Hoppe. *Accurate Cone-Beam Image Reconstruction in C-Arm Computed Tomography*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2009.
- [Houn 73] G. N. Hounsfield. “Computerized transverse axial scanning (tomography): Part 1. Description of system”. *The British Journal of Radiology*, Vol. 46, No. 552, pp. 1016–1022, 1973.
- [Hsie 04] J. Hsieh, E. Chao, J. Thibault, B. Grekowitz, A. Horst, S. McOlash, and T. J. Myers. “A novel reconstruction algorithm to extend the CT scan field-of-view”. *Medical Physics*, Vol. 31, No. 9, pp. 2385–2391, Sep 2004.
- [Hsie 13] J. Hsieh, B. Nett, Z. Yu, K. Sauer, J.-B. Thibault, and C. A. Bouman. “Recent Advances in CT Image Reconstruction”. *Current Radiology Reports*, Vol. 1, No. 1, pp. 39–51, Mar 2013.
- [Huan 16] Y. Huang, O. Taubmann, X. Huang, V. Haase, G. Lauritsch, and A. Maier. “A New Weighted Anisotropic Total Variation Algorithm For Limited Angle Tomography”. In: IEEE, Ed., *Proc. 2016 IEEE International Symposium on Biomedical Imaging*, pp. 585–588, Apr 2016.
- [Hubb 95] J. H. Hubbell and S. M. Seltzer. “Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients 1 keV to 20 MeV for elements Z= 1 to 92 and 48 additional substances of dosimetric interest”. Tech. Rep., National Institute of Standards and Technology, Ionizing Radiation Division, Gaithersburg, MD, USA, 1995.

- [John 38] F. John. “The ultrahyperbolic differential equation with four independent variables”. *Duke Mathematical Journal*, Vol. 4, No. 2, pp. 300–322, 1938.
- [Kak 01] A. C. Kak and M. Slaney. *Principles of computerized tomographic imaging*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.
- [Karp 88] J. S. Karp, G. Muehlechner, and R. M. Lewitt. “Constrained Fourier space method for compensation of missing data in emission computed tomography”. *IEEE Transactions on Medical Imaging*, Vol. 7, No. 1, pp. 21–25, Mar 1988.
- [Kats 02] A. Katsevich. “Theoretically Exact Filtered Backprojection-Type Inversion Algorithm for Spiral CT”. *SIAM Journal on Applied Mathematics*, Vol. 62, No. 6, pp. 2012–2026, 2002.
- [Keck 14] B. Keck. *High Performance Iterative X-Ray CT with Application in 3-D Mammography and Interventional C-arm Imaging Systems*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2014.
- [Kell 57] J. H. Kellgren and J. S. Lawrence. “Radiological Assessment of Osteo-Arthrosis”. *Annals of the Rheumatic Diseases*, Vol. 16, No. 4, pp. 494–502, Dec 1957.
- [Knut 93] H. Knutsson and C.-F. Westin. “Normalized and differential convolution”. In: *Proc. Computer Vision and Pattern Recognition*, pp. 515–523, Jun 1993.
- [Kudo 91] H. Kudo and T. Saito. “Sinogram recovery with the method of convex projections for limited-data reconstruction in computed tomography”. *Journal of the Optical Society of America A*, Vol. 8, No. 7, pp. 1148–1160, Jul 1991.
- [Kulo 32] J. Kulowski. “Flexion Contracture of the Knee”. *The Journal of Bone & Joint Surgery*, Vol. 14, No. 3, pp. 618–630, Jul 1932.
- [Kyri 08] Y. Kyriakou, R. M. Lapp, L. Hillebrand, D. Ertel, and W. A. Kalender. “Simultaneous misalignment correction for approximate circular cone-beam computed tomography”. *Physics in Medicine and Biology*, Vol. 53, No. 22, pp. 6267–6289, Nov 2008.
- [Laur 01] G. Lauritsch and H. Bruder. “FORBILD Head Phantom”. Website, 2001. <http://www.imp.uni-erlangen.de/phantoms/> (last checked: 2016/11/27).
- [Laur 06] G. Lauritsch, J. Boese, L. Wigstrom, H. Kemeth, and R. Fahrig. “Towards cardiac C-arm computed tomography”. *IEEE Transactions on Medical Imaging*, Vol. 25, No. 7, pp. 922–934, Jul 2006.
- [Leac 70] R. E. Leach, T. Gregg, and F. J. Siber. “Weight-Bearing Radiography in Osteoarthritis of the Knee”. *Radiology*, Vol. 97, No. 2, pp. 265–268, Nov 1970.
- [Leng 07] S. Leng, B. Nett, M. Speidel, and G.-H. Chen. “Motion artifact reduction in fan-beam and cone-beam computed tomography via the fan-beam data consistency condition (FDCC)”. In: *Proc. SPIE 6510*, p. 65101W, 2007.

- [Li 06] T. Li, E. Schreibmann, Y. Yang, and L. Xing. “Motion correction for improved target localization with on-board cone-beam computed tomography”. *Physics in Medicine and Biology*, Vol. 51, No. 2, pp. 253–267, Jan 2006.
- [Li 07] T. Li, A. Koong, and L. Xing. “Enhanced 4D cone-beam CT with inter-phase motion model”. *Medical Physics*, Vol. 34, No. 9, pp. 3688–3695, Sep 2007.
- [Li 08] G. Li, S. K. Van de Velde, and J. T. Bingham. “Validation of a non-invasive fluoroscopic imaging technique for the measurement of dynamic knee joint motion”. *Journal of Biomechanics*, Vol. 41, No. 7, pp. 1616–1622, Jan 2008.
- [Li 10] R. Li, X. Jia, J. H. Lewis, X. Gu, M. Folkerts, C. Men, and S. B. Jiang. “Real-time volumetric image reconstruction and 3D tumor localization based on a single x-ray projection image for lung cancer radiotherapy”. *Medical Physics*, Vol. 37, No. 6, pp. 2822–2826, Jun 2010.
- [Li 11] R. Li, J. H. Lewis, X. Jia, X. Gu, M. Folkerts, C. Men, W. Y. Song, and S. B. Jiang. “3D tumor localization through real-time volumetric x-ray imaging for lung cancer radiotherapy”. *Medical Physics*, Vol. 38, No. 5, pp. 2783–2794, May 2011.
- [Lin 13] C.-C. Lin, S. Zhang, J. Frahm, T.-W. Lu, C.-Y. Hsu, and T.-F. Shih. “A slice-to-volume registration method based on real-time magnetic resonance imaging for measuring three-dimensional kinematics of the knee”. *Medical Physics*, Vol. 40, No. 10, p. 102302, Oct 2013.
- [Livy 03] H. Livyatan, Z. Yaniv, and L. Joskowicz. “Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT”. *IEEE Transactions on Medical Imaging*, Vol. 22, No. 11, pp. 1395–1406, Nov 2003.
- [Llop 12] E. Llopis, E. Fernandez, and L. Cerezal. “MR and CT Arthrography of the Hip”. *Seminars in Musculoskeletal Radiology*, Vol. 16, No. 1, pp. 42–56, Feb 2012.
- [Lorc 15] B. Lorch, M. Berger, J. Hornegger, and A. Maier. “Projection and Reconstruction-Based Noise Filtering Methods in Cone Beam CT”. In: T. Deserno, H. Handels, H.-P. Meinzer, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2015*, pp. 59–64, Mar 2015.
- [Loy 03] G. Loy and A. Zelinsky. “Fast radial symmetry for detecting points of interest”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 8, pp. 959–973, Aug 2003.
- [Lu 16] Y. Lu, M. Berger, M. Manhart, J. hwan Choi, M. Hoheisel, M. Kowarschik, R. Fahrig, Q. Ren, J. Hornegger, and A. Maier. “Bridge to Real Data: Empirical Multiple Material Calibration for Learning-Based Material Decomposition”. In: IEEE, Ed., *Proc. 2016 IEEE International Symposium on Biomedical Imaging*, pp. 457–460, Apr 2016.
- [Ludw 66] D. Ludwig. “The radon transform on euclidean space”. *Communications on Pure and Applied Mathematics*, Vol. 19, No. 1, pp. 49–81, Feb 1966.
- [Maas 14] N. Maass, F. Dennerlein, A. Aichert, and A. Maier. “Geometrical Jitter Correction in Computed Tomography”. In: F. Noo, Ed., *Proc. The third international conference on image formation in x-ray computed tomography*, pp. 338–342, Jun 2014.

- [Maie 10] A. Maier and M. Berger. “JPOP: Java parallel optimization package”. Website, 2010. <https://www5.cs.fau.de/research/software/java-parallel-optimization-package/> (last checked: 2016/11/27).
- [Maie 11a] A. Maier, J.-H. Choi, A. Keil, C. Niebler, M. Sarmiento, A. Fieselmann, G. Gold, S. Delp, and R. Fahrig. “Analysis of vertical and horizontal circular C-arm trajectories”. In: *Proc. SPIE 7961*, p. 796123, 2011.
- [Maie 11b] A. Maier, L. Wigström, H. Hofmann, J. Hornegger, L. Zhu, N. Strobel, and R. Fahrig. “Three-dimensional anisotropic adaptive filtering of projection data for noise reduction in cone beam CT”. *Medical Physics*, Vol. 38, No. 11, pp. 5896–5909, Nov 2011.
- [Maie 12a] A. Maier, H. G. Hofmann, C. Schwemmer, J. Hornegger, A. Keil, and R. Fahrig. “Fast simulation of x-ray projections of spline-based surfaces using an append buffer”. *Physics in Medicine and Biology*, Vol. 57, No. 19, pp. 6193–6210, Oct 2012.
- [Maie 12b] A. Maier, B. Scholz, and F. Dennerlein. “Optimization-based Extrapolation for Truncation Correction”. In: F. Noo, Ed., *Proc. The second international conference on image formation in x-ray computed tomography*, pp. 390–394, Jun 2012.
- [Maie 13] A. Maier, H. G. Hofmann, M. Berger, P. Fischer, C. Schwemmer, H. Wu, K. Müller, J. Hornegger, J.-H. Choi, C. Riess, A. Keil, and R. Fahrig. “CONRAD—a software framework for cone-beam imaging in radiology”. *Medical Physics*, Vol. 40, No. 11, p. 111914, Nov 2013.
- [Maie 14] A. Maier, O. Taubmann, J. Wetzl, J. Wasza, C. Forman, P. Fischer, J. Hornegger, and R. Fahrig. “Fast Interpolation of Dense Motion Fields from Synthetic Phantoms”. In: T. Deserno, H. Handels, H.-P. Meinzer, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2014*, pp. 168–173, Mar 2014.
- [Maie 15a] A. Maier and R. Fahrig. “GPU Denoising for Computed Tomography”. In: X. Jia and J. Steve, Eds., *Graphics Processing Unit-Based High Performance Computing in Radiation Therapy*, pp. 113–128, CRC Press, Boca Raton, FL, USA, 2015.
- [Maie 15b] A. Maier, P. Kugler, G. Lauritsch, and J. Hornegger. “Discrete Estimation of Data Completeness for 3D Scan Trajectories with Detector Offset”. In: T. Deserno, H. Handels, H.-P. Meinzer, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2015*, pp. 47–52, Mar 2015.
- [Mark 12] P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš. “A review of 3D/2D registration methods for image-guided interventions”. *Medical Image Analysis*, Vol. 16, No. 3, pp. 642–661, Apr 2012.
- [Mazi 10] S. R. Mazin and N. J. Pelc. “Fourier properties of the fan-beam sinogram”. *Medical Physics*, Vol. 37, No. 4, pp. 1674–1680, Apr 2010.
- [Mitr 13] U. Mitrovic, Z. Spiclin, B. Likar, and F. Pernus. “3D-2D Registration of Cerebral Angiograms: A Method and Evaluation on Clinical Images”. *IEEE Transactions on Medical Imaging*, Vol. 32, No. 8, pp. 1550–1563, Aug 2013.
- [Mode 03] J. Modersitzki. *Numerical Methods for Image Registration*. Oxford University Press, Oxford, Great Britain, 2003.

- [Mou 06] X. Mou, S. Tang, and H. Yu. “Comparison on beam hardening correction of CT based on H-L consistency and normal water phantom experiment”. In: *Proc. SPIE 6318*, p. 63181V, 2006.
- [Mull 13] K. Müller, C. Schwemmer, J. Hornegger, Y. Zheng, Y. Wang, G. Lauritsch, C. Rohkohl, A. Maier, C. Schultz, and R. Fahrig. “Evaluation of interpolation methods for surface-based motion compensated tomographic reconstruction for cardiac angiographic C-arm data”. *Medical Physics*, Vol. 40, No. 3, p. 031107, 2013.
- [Mull 14a] K. Müller. *3-D Imaging of the Heart Chambers with C-arm CT*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2014.
- [Mull 14b] K. Müller, A. Maier, C. Schwemmer, G. Lauritsch, S. D. Buck, J.-Y. Wielandts, J. Hornegger, and R. Fahrig. “Image artefact propagation in motion estimation and reconstruction in interventional cardiac C-arm CT”. *Physics in Medicine and Biology*, Vol. 59, No. 12, pp. 3121–3138, Jun 2014.
- [Mull 14c] K. Müller, A. Maier, Y. Zheng, Y. Wang, G. Lauritsch, C. Schwemmer, C. Rohkohl, J. Hornegger, and R. Fahrig. “Interventional heart wall motion analysis with cardiac C-arm CT systems”. *Physics in Medicine and Biology*, Vol. 59, No. 9, pp. 2265–2284, May 2014.
- [Mull 15a] K. Müller, M. Berger, J.-H. Choi, S. Datta, S. Gehrisch, T. Moore, M. P. Marks, A. Maier, and R. Fahrig. “Fully Automatic Head Motion Correction for Interventional C-arm Systems using Fiducial Markers”. In: S. King and M. Glick, Eds., *Proc. Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, pp. 534–537, Jun 2015.
- [Mull 15b] K. Müller, M. Berger, J.-H. Choi, A. Maier, and R. Fahrig. “Automatic Motion Estimation and Compensation Framework for Weight-bearing C-arm CT scans using Fiducial Markers”. In: D. A. Jaffray, Ed., *IFMBE Proceedings*, pp. 58–61, Jun 2015.
- [Murp 08] L. Murphy, T. A. Schwartz, C. G. Helmick, J. B. Renner, G. Tudor, G. Koch, A. Dragomir, W. D. Kalsbeek, G. Luta, and J. M. Jordan. “Lifetime risk of symptomatic knee osteoarthritis”. *Arthritis Care & Research*, Vol. 59, No. 9, pp. 1207–1213, Sep 2008.
- [Natt 86] F. Natterer. *The Mathematics of Computerized Tomography*. Springer, Berlin, Germany, 1986.
- [Natt 93] F. Natterer. “Determination of tissue attenuation in emission tomography of optically dense media”. *Inverse Problems*, Vol. 9, No. 6, pp. 731–736, Dec 1993.
- [Noce 06] J. Nocedal and S. Wright. *Numerical Optimization*. Springer New York, New York, NY, USA, 2006.
- [Noo 02] F. Noo, M. Defrise, R. Clackdoyle, and H. Kudo. “Image reconstruction from fan-beam projections on less than a short scan”. *Physics in Medicine and Biology*, Vol. 47, No. 14, pp. 2525–2546, Jul 2002.
- [Ohne 00] B. Ohnesorge, T. Flohr, K. Schwarz, J. P. Heiken, and K. T. Bae. “Efficient correction for CT image artifacts caused by objects extending outside the scan field of view”. *Medical Physics*, Vol. 27, No. 1, pp. 39–46, Jan 2000.

- [Otak 13] Y. Otake, A. S. Wang, J. Webster Stayman, A. Uneri, G. Kleinszig, S. Vogt, A. J. Khanna, Z. L. Gokaslan, and J. H. Siewerdsen. “Robust 3D-2D image registration: application to spine interventions and vertebral labeling in the presence of anatomical deformation”. *Physics in Medicine and Biology*, Vol. 58, No. 23, pp. 8535–8553, Dec 2013.
- [Otak 15] Y. Otake, A. S. Wang, A. Uneri, G. Kleinszig, S. Vogt, N. Aygun, S.-f. L. Lo, J.-P. Wolinsky, Z. L. Gokaslan, and J. H. Siewerdsen. “3D-2D registration in mobile radiographs: algorithm development and preliminary clinical evaluation”. *Physics in Medicine and Biology*, Vol. 60, No. 5, pp. 2075–2090, Mar 2015.
- [Ouad 15] S. Ouadah, J. W. Stayman, G. Gang, A. Uneri, T. Ehtiati, and J. H. Siewerdsen. “Self-calibration of cone-beam CT geometry using 3D-2D image registration: development and application to task-based imaging with a robotic C-arm”. In: *Proc. SPIE 9415*, p. 94151D, 2015.
- [Ouad 16] S. Ouadah, J. W. Stayman, G. J. Gang, T. Ehtiati, and J. H. Siewerdsen. “Self-calibration of cone-beam CT geometry using 3D-2D image registration”. *Physics in Medicine and Biology*, Vol. 61, No. 7, pp. 2613–2632, Apr 2016.
- [Pal 11] S. Pal, C. E. Draper, M. Fredericson, G. E. Gold, S. L. Delp, G. S. Beaupre, and T. F. Besier. “Patellar Maltracking Correlates With Vastus Medialis Activation Delay in Patellofemoral Pain Patients”. *American Journal of Sports Medicine*, Vol. 39, No. 3, pp. 590–598, Mar 2011.
- [Pan 09] X. Pan, E. Y. Sidky, and M. Vannier. “Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction?”. *Inverse Problems*, Vol. 25, No. 12, p. 1230009, Jan 2009.
- [Park 82] D. L. Parker. “Optimal short scan convolution reconstruction for fan beam CT”. *Medical Physics*, Vol. 9, No. 2, pp. 254–257, Mar 1982.
- [Patc 02a] S. K. Patch. “Computation of unmeasured third-generation VCT views from measured views”. *IEEE Transactions on Medical Imaging*, Vol. 21, No. 7, pp. 801–813, Jul 2002.
- [Patc 02b] S. K. Patch. “Consistency conditions upon 3D CT data and the wave equation”. *Physics in Medicine and Biology*, Vol. 47, No. 15, pp. 2637–2650, Aug 2002.
- [Pate 03] V. V. Patel, K. Hall, M. Ries, C. Lindsey, E. Ozhinsky, Y. Lu, and S. Majumdar. “Magnetic Resonance Imaging of Patellofemoral Kinematics with Weight-Bearing”. *The Journal of Bone & Joint Surgery. American Volume*, Vol. 85, No. 12, pp. 2419–2424, Dec 2003.
- [Pawi 11] S. A. Pawiro, P. Markelj, F. Pernus, C. Gendrin, M. Figl, C. Weber, F. Kainberger, I. Nöbauer-Huhmann, H. Bergmeister, M. Stock, D. Georg, H. Bergmann, and W. Birkfellner. “Validation for 2D/3D registration. I: A new gold standard data set”. *Medical Physics*, Vol. 38, No. 3, pp. 1481–1490, Mar 2011.
- [Penn 98] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. Hill, and D. J. Hawkes. “A comparison of similarity measures for use in 2-D-3-D medical image registration”. *IEEE Transactions on Medical Imaging*, Vol. 17, No. 4, pp. 586–595, Aug 1998.

- [Piep 04] S. Pieper, M. Halle, and R. Kikinis. “3D Slicer”. In: IEEE, Ed., *Proc. 2004 IEEE International Symposium on Biomedical Imaging*, pp. 632–635, Apr 2004.
- [Plat 04] W. Platzer. *Color Atlas and Textbook of Human Anatomy, Vol. 1: Locomotor System*. Thieme Publishing Group, New York, NY, USA, 5 Ed., 2004.
- [Pohl 14] M. Pohlmann, M. Berger, A. Maier, J. Hornegger, and R. Fahrig. “Estimation of missing fan-beam projections using frequency consistency conditions”. In: F. Noo, Ed., *Proc. The third international conference on image formation in x-ray computed tomography*, pp. 203–207, Jun 2014.
- [Powe 03] C. M. Powers, S. R. Ward, M. Fredericson, M. Guillet, and F. G. Shellock. “Patellofemoral Kinematics During Weight-Bearing and Non-Weight-Bearing Knee Extension in Persons With Lateral Subluxation of the Patella: A Preliminary Study”. *Journal of Orthopaedic & Sports Physical Therapy*, Vol. 33, No. 11, pp. 677–685, Nov 2003.
- [Powe 99] C. M. Powers, S. Mortenson, D. Nishimoto, and D. Simon. “Criterion-Related Validity of a Clinical Measurement To Determine the Medial/Lateral Component of Patellar Orientation”. *Journal of Orthopaedic & Sports Physical Therapy*, Vol. 29, No. 7, pp. 372–377, Jul 1999.
- [Preu 15] A. Preuhs, M. Berger, Y. Xia, A. Maier, J. Hornegger, and R. Fahrig. “Over-Exposure Correction in CT Using Optimization-Based Multiple Cylinder Fitting”. In: T. Deserno, H. Handels, H.-P. Meinzer, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2015*, pp. 35–40, Mar 2015.
- [Prum 06] M. Prümmer, L. Wigstrom, J. Hornegger, J. Boese, G. Lauritsch, N. Strobel, and R. Fahrig. “Cardiac C-arm CT: Efficient Motion Correction for 4D-FBP”. In: *Proc. IEEE Nuclear Science Symposium & Medical Imaging Conference (NSS/MIC)*, pp. 2620–2628, Nov 2006.
- [Prum 09a] M. Prümmer. *Cardiac C-Arm Computed Tomography: Motion Estimation and Dynamic Reconstruction*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2009.
- [Prum 09b] M. Prümmer, J. Hornegger, G. Lauritsch, L. Wigström, E. Girard-Hughes, and R. Fahrig. “Cardiac C-arm CT: a unified framework for motion estimation and dynamic CT”. *IEEE Transactions on Medical Imaging*, Vol. 28, No. 11, pp. 1836–1849, Nov 2009.
- [Rado 17] J. Radon. “Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten”. *Berichte über die Verhandlungen der Königlich-Sächsischen Akademie der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse*, Vol. 69, pp. 262–277, 1917.
- [Raus 16] J. Rausch, A. Maier, R. Fahrig, J.-H. Choi, W. Hinshaw, F. Schebesch, S. Haase, J. Wasza, J. Hornegger, and C. Riess. “Kinect-based Correction of Overexposure Artifacts in Knee Imaging with C-arm CT Systems”. *International Journal of Biomedical Imaging*, Vol. 2016, Jun 2016.
- [Ries 13] C. Riess, M. Berger, H. Wu, M. Manhart, R. Fahrig, and A. Maier. “TV or not TV? That is the Question”. In: R. M. Leahy and J. Qi, Eds., *Proc. Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, pp. 341–344, Jun 2013.

- [Rit 09a] S. Rit, D. Sarrut, and L. Desbat. “Comparison of Analytic and Algebraic Methods for Motion-Compensated Cone-Beam CT Reconstruction of the Thorax”. *IEEE Transactions on Medical Imaging*, Vol. 28, No. 10, pp. 1513–1525, Oct 2009.
- [Rit 09b] S. Rit, J. W. H. Wolthaus, M. van Herk, and J.-J. Sonke. “On-the-fly motion-compensated cone-beam CT using an a priori model of the respiratory motion”. *Medical Physics*, Vol. 36, No. 6, pp. 2283–2296, Jun 2009.
- [Rit 13] S. Rit, D. Sarrut, and J.-J. Sonke. “Respiratory Motion Correction in Cone-Beam CT for Image-Guided Radiotherapy”. In: J. Ehrhardt and C. Lorenz, Eds., *4D Modeling and Estimation of Respiratory Motion for Radiation Therapy*, pp. 319–334, Springer, Berlin, Germany, 2013.
- [Rohk 11] C. Rohkohl. *Motion Estimation and Compensation for Interventional Cardiovascular Image Reconstruction*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2011.
- [Rohk 13] C. Rohkohl, H. Bruder, K. Stierstorfer, and T. Flohr. “Improving best-phase image quality in cardiac CT by motion correction with MAM optimization”. *Medical Physics*, Vol. 40, No. 3, p. 031901, Mar 2013.
- [Rose 88] T. D. Rosenberg, L. E. Paulos, R. D. Parker, D. B. Coward, and S. M. Scott. “The forty-five-degree posteroanterior flexion weight-bearing radiograph of the knee”. *The Journal of Bone & Joint Surgery. American Volume*, Vol. 70, No. 10, pp. 1479–1483, Dec 1988.
- [Roux 04] S. Roux, L. Desbat, A. Koenig, and P. Grangeat. “Exact reconstruction in 2D dynamic CT: compensation of time-dependent affine deformations”. *Physics in Medicine and Biology*, Vol. 49, No. 11, pp. 2169–2182, Jun 2004.
- [Ruec 99] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes. “Nonrigid registration using free-form deformations: application to breast MR images”. *IEEE Transactions on Medical Imaging*, Vol. 18, No. 8, pp. 712–721, Aug 1999.
- [Ruhr 11] E.-P. Rührnschopf and K. Klingensbeck. “A general framework and review of scatter correction methods in cone beam CT. Part 2: scatter estimation approaches”. *Medical Physics*, Vol. 38, No. 9, pp. 5186–5199, Sep 2011.
- [Russ 05] D. B. Russakoff, T. Rohlfing, J. R. Adler, and C. R. Maurer. “Intensity-based 2D-3D spine image registration incorporating a single fiducial marker”. *Academic Radiology*, Vol. 12, No. 1, pp. 37–50, Jan 2005.
- [Saho 88] P. K. Sahoo, S. Soltani, and A. K. C. Wong. “A survey of thresholding techniques”. *Computer Vision, Graphics, and Image Processing*, Vol. 41, No. 2, pp. 233–260, Feb 1988.
- [Saka 00] N. Sakai, Z.-P. Luo, J. A. Rand, and K.-N. An. “The influence of weakness in the vastus medialis oblique muscle on the patellofemoral joint: an in vitro biomechanical study”. *Clinical Biomechanics*, Vol. 15, No. 5, pp. 335–339, Jun 2000.
- [Scha 06] D. Schaefer, J. Borgert, V. Rasche, and M. Grass. “Motion-compensated and gated cone beam filtered back-projection for 3-D rotational X-ray angiography”. *IEEE Transactions on Medical Imaging*, Vol. 25, No. 7, pp. 898–906, Jul 2006.

- [Sche 07] H. Scherl, B. Keck, M. Kowarschik, and J. Hornegger. “Fast GPU-Based CT Reconstruction using the Common Unified Device Architecture (CUDA)”. In: *Proc. IEEE Nuclear Science Symposium & Medical Imaging Conference (NSS/MIC)*, pp. 4464–4466, Nov 2007.
- [Schw 00] A. Schweikard, G. Glosser, M. Bodduluri, M. J. Murphy, and J. R. Adler. “Robotic motion compensation for respiratory movement during radio-surgery”. *Computer Aided Surgery*, Vol. 5, No. 4, pp. 263–277, Jan 2000.
- [Schw 05] A. Schweikard, H. Shiomi, and J. Adler. “Respiration tracking in radio-surgery without fiducials”. *The International Journal of Medical Robotics and Computer Assisted Surgery*, Vol. 1, No. 2, pp. 19–27, Jan 2005.
- [Schw 10] C. Schwemmer, M. Prümmer, V. Daum, and J. Hornegger. “High-Density Object Removal from Projection Images using Low-Frequency-Based Object Masking”. In: T. Deserno, H. Handels, H.-P. Meinzer, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2010*, pp. 365–369, Mar 2010.
- [Schw 13] C. Schwemmer, C. Rohkohl, G. Lauritsch, K. Müller, and J. Hornegger. “Residual motion compensation in ECG-gated interventional cardiac vasculature reconstruction”. *Physics in Medicine and Biology*, Vol. 58, No. 11, pp. 3717–3737, Jun 2013.
- [Sega 10] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui. “4D XCAT phantom for multimodality imaging research”. *Medical Physics*, Vol. 37, No. 9, pp. 4902–4915, Sep 2010.
- [Shi 16] L. Shi, J. Maier, C. Soell, J. Roeber, M. Berger, B. Bier, R. Fahrig, B. Eskofier, and A. Maier. “Analog Non-Linear Transformation-Based Tone Mapping for Image Enhancement in C-arm CT”. In: *Proc. IEEE Nuclear Science Symposium & Medical Imaging Conference (NSS/MIC)*, Nov 2016.
- [Shir 03] H. Shirato, T. Harada, T. Harabayashi, K. Hida, H. Endo, K. Kitamura, R. Onimaru, K. Yamazaki, N. Kurauchi, T. Shimizu, N. Shinohara, M. Matsushita, H. Dosaka-Akita, and K. Miyasaka. “Feasibility of insertion/implantation of 2.0-mm-diameter gold internal fiducial markers for precise setup and real-time tumor tracking in radiotherapy”. *International Journal of Radiation Oncology \* Biology \* Physics*, Vol. 56, No. 1, pp. 240–247, May 2003.
- [Sidk 06] E. Y. Sidky, C.-M. Kao, and X. Pan. “Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT”. *Journal of X-Ray Science and Technology*, Vol. 14, No. 2, pp. 119–139, Apr 2006.
- [Sidk 08] E. Y. Sidky and X. Pan. “Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization”. *Physics in Medicine and Biology*, Vol. 53, No. 17, pp. 4777–4807, Sep 2008.
- [Silb 12] B. I. Silbert, B. J. Singer, P. L. Silbert, J. T. Gibbons, and K. P. Singer. “Enduring efficacy of Botulinum toxin type A injection for refractory anterior knee pain”. *Disability and Rehabilitation*, Vol. 34, No. 1, pp. 62–68, Jan 2012.
- [Sing 11] B. J. Singer, P. L. Silbert, S. Song, J. W. Dunne, and K. P. Singer. “Treatment of refractory anterior knee pain using botulinum toxin type A (Dysport) injection to the distal vastus lateralis muscle: a randomised

- placebo controlled crossover trial”. *British Journal of Sports Medicine*, Vol. 45, No. 8, pp. 640–645, Jun 2011.
- [Sisn 16] A. Sisniega, J. W. Stayman, Q. Cao, J. Yorkston, J. H. Siewerdsen, and W. Zbijewski. “Image-based motion compensation for high-resolution extremities cone-beam CT”. In: *Proc. SPIE 9783*, p. 97830K, 2016.
- [Sonk 05] J.-J. Sonke, L. Zijp, P. Remeijer, and M. van Herk. “Respiratory correlated cone beam CT”. *Medical Physics*, Vol. 32, No. 4, pp. 1176–1186, Apr 2005.
- [Souz 10] R. B. Souza, C. E. Draper, M. Fredericson, and C. M. Powers. “Femur Rotation and Patellofemoral Joint Kinematics: A Weight-Bearing Magnetic Resonance Imaging Analysis”. *Journal of Orthopaedic & Sports Physical Therapy*, Vol. 40, No. 5, pp. 277–285, May 2010.
- [Stan 88] W. Stanford, J. Phelan, M. H. Kathol, S. A. Rooholamini, G. Y. El-Khoury, G. R. Palutsis, and J. P. Albright. “Patellofemoral joint motion: evaluation by ultrafast computed tomography”. *Skeletal Radiology*, Vol. 17, No. 7, pp. 487–492, Oct 1988.
- [Star 05] J. Starman, N. Pelc, N. Strobel, and R. Fahrig. “Estimating 0th and 1st moments in C-arm CT data for extrapolating truncated projections”. In: *Proc. SPIE 5747*, pp. 378–387, 2005.
- [Stro 16a] D. Stromer, M. Amrehn, Y. Huang, P. Kugler, S. Bauer, G. Lauritsch, and A. Maier. “Comparison of SART and eTV Reconstruction for increased C-arm CT Volume Coverage by proper Detector Rotation in Liver Imaging”. In: IEEE, Ed., *Proc. 2016 IEEE International Symposium on Biomedical Imaging*, pp. 589–592, Apr 2016.
- [Stro 16b] D. Stromer, P. Kugler, S. Bauer, G. Lauritsch, and A. Maier. “Data Completeness Estimation for 3D C-Arm Scans with Rotated Detector to Enlarge the Lateral Field-of-View”. In: T. Deserno, H. Handels, H.-P. Meinzer, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2016*, pp. 164–169, Mar 2016.
- [Tagu 08] K. Taguchi and H. Kudo. “A simple motion tracking backprojection for a class of affine transformation”. In: *Proc. SPIE 6913*, p. 69131V, 2008.
- [Tang 09] J. Tang, B. E. Nett, and G.-H. Chen. “Performance comparison between total variation (TV)-based compressed sensing and statistical iterative reconstruction algorithms”. *Physics in Medicine and Biology*, Vol. 54, No. 19, pp. 5781–5804, Oct 2009.
- [Tang 11] S. Tang, X. Mou, Q. Xu, Y. Zhang, J. Bennett, and H. Yu. “Data consistency condition-based beam-hardening correction”. *Optical Engineering*, Vol. 50, No. 7, p. 76501, Jul 2011.
- [Tash 03] S. Tashman and W. Anderst. “In-vivo measurement of dynamic joint motion using high speed biplane radiography and CT: application to canine ACL deficiency”. *Journal of Biomechanical Engineering*, Vol. 125, No. 2, pp. 238–245, Apr 2003.
- [Taub 16] O. Taubmann, A. Maier, J. Hornegger, G. Lauritsch, and R. Fahrig. “Coping with Real World Data: Artifact Reduction and Denoising for Motion-Compensated Cardiac C-arm CT”. *Medical Physics*, Vol. 43, No. 2, pp. 883–893, Feb 2016.

- [Thaw 15] G. K. Thawait, S. Demehri, A. AlMuhit, W. Zbijweski, J. Yorkston, F. Del Grande, B. Zikria, J. A. Carrino, and J. H. Siewerdsen. “Extremity cone-beam CT for evaluation of medial tibiofemoral osteoarthritis: Initial experience in imaging of the weight-bearing and non-weight-bearing knee”. *European Journal of Radiology*, Vol. 84, No. 12, pp. 2564–2570, Dec 2015.
- [Thie 15] J. Thies, M. Zollhöfer, M. Nießner, L. Valgaerts, M. Stamminger, and C. Theobalt. “Real-time Expression Transfer for Facial Reenactment”. *ACM Transactions on Graphics*, Vol. 34, No. 6, pp. 183:1–183:14, Nov 2015.
- [Thie 16] J. Thies, M. Zollhöfer, M. Stamminger, C. Theobalt, and M. Nießner. “Face2Face: Real-time Face Capture and Reenactment of RGB Videos”. In: *Proc. Computer Vision and Pattern Recognition*, pp. 2387–2395, Jun 2016.
- [Tsai 10] T.-Y. Tsai, T.-W. Lu, C.-M. Chen, M.-Y. Kuo, and H.-C. Hsu. “A volumetric model-based 2D to 3D registration method for measuring kinematics of natural knees with single-plane fluoroscopy”. *Medical Physics*, Vol. 37, No. 3, pp. 1273–1284, Mar 2010.
- [Tuom 13] E. K. J. Tuominen, J. Kankare, S. K. Koskinen, and K. T. Mattila. “Weight-Bearing CT Imaging of the Lower Extremity”. *American Journal of Roentgenology*, Vol. 200, No. 1, pp. 146–148, Jan 2013.
- [Tuy 83] H. K. Tuy. “An Inversion Formula for Cone-Beam Reconstruction”. *SIAM Journal on Applied Mathematics*, Vol. 43, No. 3, pp. 546–552, Jun 1983.
- [Unbe 15] M. Unberath, J.-H. Choi, M. Berger, A. Maier, and R. Fahrig. “Image-based Compensation for Involuntary Motion in Weight-bearing C-arm Cone-beam CT Scanning of Knees”. In: *Proc. SPIE 9413*, p. 94130D, 2015.
- [Unbe 16a] M. Unberath, S. Achenbach, R. Fahrig, and A. Maier. “Exhaustive Graph Cut-based Vasculature Reconstruction”. In: IEEE, Ed., *Proc. 2016 IEEE International Symposium on Biomedical Imaging*, pp. 1143–1146, Apr 2016.
- [Unbe 16b] M. Unberath, A. Aichert, S. Achenbach, and A. Maier. “Virtual Single-frame Subtraction Imaging”. In: M. Kachelries, Ed., *Proc. The fourth international conference on image formation in x-ray computed tomography*, pp. 89–92, 2016.
- [Van 06] G. Van Gompel, M. Defrise, and D. Van Dyck. “Elliptical extrapolation of truncated 2D CT projections using Helgason-Ludwig consistency conditions”. In: *Proc. SPIE 6142*, p. 61424B, 2006.
- [Wang 02] Z. Wang and A. Bovik. “A universal image quality index”. *IEEE Signal Processing Letters*, Vol. 9, No. 3, pp. 81–84, Mar 2002.
- [Wang 04] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. “Image Quality Assessment: From Error Visibility to Structural Similarity”. *IEEE Transactions on Image Processing*, Vol. 13, No. 4, pp. 600–612, Apr 2004.
- [Wang 14] M. Wang, G. C. Sharp, S. Rit, V. Delmon, and G. Wang. “2D/4D marker-free tumor tracking using 4D CBCT as the reference image”. *Physics in Medicine and Biology*, Vol. 59, No. 9, pp. 2219–2233, May 2014.

- [Wasz 16] J. Wasza, P. Fischer, H. Leutheuser, T. Oefner, C. Bert, A. Maier, and J. Hornegger. “Real-Time Respiratory Motion Analysis Using 4-D Shape Priors”. *IEEE Transactions on Biomedical Engineering*, Vol. 63, No. 3, pp. 485–495, 2016.
- [Wein 05] W. Wein, B. Roeper, and N. Navab. “2D/3D registration based on volume gradients”. In: *Proc. SPIE 5747*, pp. 144–150, 2005.
- [Welc 03] A. Welch, W. Hallett, P. Marsden, and A. Bromiley. “Accurate attenuation correction in PET using short transmission scans and consistency information”. *IEEE Transactions on Nuclear Science*, Vol. 50, No. 3, pp. 427–432, Jun 2003.
- [Welc 98] A. Welch, C. Campbell, R. Clackdoyle, F. Natterer, M. Hudson, A. Bromiley, P. Mikecz, F. Chillcot, M. Dodd, P. Hopwood, S. Craib, G. T. Gullberg, and P. Sharp. “Attenuation correction in PET using consistency information”. *IEEE Transactions on Nuclear Science*, Vol. 45, No. 6, pp. 3134–3141, Dec 1998.
- [Wern 95] S. Werner. “An evaluation of knee extensor and knee flexor torques and EMGs in patients with patellofemoral pain syndrome in comparison with matched controls”. *Knee Surgery, Sports Traumatology, Arthroscopy*, Vol. 3, No. 2, pp. 89–94, Jul 1995.
- [Wick 12] J. Wicklein, H. Kunze, W. A. Kalender, and Y. Kyriakou. “Image features for misalignment correction in medical flat-detector CT”. *Medical Physics*, Vol. 39, No. 8, pp. 4918–4931, Aug 2012.
- [Wu 12] H. Wu, A. Maier, R. Fahrig, and J. Hornegger. “Spatial-temporal total variation regularization (STTVR) for 4D-CT reconstruction”. In: *Proc. SPIE 8313*, p. 83133J, 2012.
- [Wu 16] H. Wu, M. Berger, A. Maier, and D. Lohmann. “A Memory Management Library for CT-Reconstruction on GPUs”. In: T. Deserno, H. Handels, H.-P. Meinzer, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2016*, pp. 206–211, Mar 2016.
- [Xia 13] Y. Xia, M. Berger, C. Riess, J. Hornegger, and A. Maier. “Dose Reduction Achieved by Dynamically Collimating the Redundant Rays in Fan-beam and Cone-beam CT”. In: *Proc. IEEE Nuclear Science Symposium & Medical Imaging Conference (NSS/MIC)*, pp. 1–4, Nov 2013.
- [Xia 14a] Y. Xia, S. Bauer, A. Maier, M. Berger, and J. Hornegger. “Patient-bounded Extrapolation for 3D Region of Interest Reconstruction in C-arm CT”. In: F. Noo, Ed., *Proc. The third international conference on image formation in x-ray computed tomography*, pp. 414–417, Jun 2014.
- [Xia 14b] Y. Xia, F. Dennerlein, S. Bauer, M. Berger, J. Hornegger, and A. Maier. “A Cone-beam Reconstruction Algorithm for Dose-minimized Short Scan and Super Short Scan”. In: F. Noo, Ed., *Proc. The third international conference on image formation in x-ray computed tomography*, pp. 178–181, Jun 2014.
- [Xia 14c] Y. Xia, A. Maier, M. Berger, and J. Hornegger. “Region of Interest Reconstruction from Dose-minimized Super Short Scan Data”. In: T. Deserno, H. Handels, H.-P. Meinzer, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2014*, pp. 48–53, Mar 2014.

- [Xia 15] Y. Xia, S. Bauer, A. Maier, M. Berger, and J. Hornegger. “Patient-bounded extrapolation using low-dose priors for volume-of-interest imaging in C-arm CT”. Vol. 42, No. 4, pp. 1787–1796, Apr 2015.
- [Xia 95] W. Xia, R. M. Lewitt, and P. R. Edholm. “Fourier correction for spatially variant collimator blurring in SPECT”. *IEEE Transactions on Medical Imaging*, Vol. 14, No. 1, pp. 100–115, Mar 1995.
- [Xu 10] J. Xu, K. Taguchi, and B. M. W. Tsui. “Statistical Projection Completion in X-ray CT Using Consistency Conditions”. *IEEE Transactions on Medical Imaging*, Vol. 29, No. 8, pp. 1528–1540, Aug 2010.
- [Yu 06] H. Yu, Y. Wei, J. Hsieh, and G. Wang. “Data consistency based translational motion artifact reduction in fan-beam CT”. *IEEE Transactions on Medical Imaging*, Vol. 25, No. 6, pp. 792–803, Jun 2006.
- [Yu 07] H. Yu and G. Wang. “Data Consistency Based Rigid Motion Artifact Reduction in Fan-Beam CT”. *IEEE Transactions on Medical Imaging*, Vol. 26, No. 2, pp. 249–260, Feb 2007.
- [Yu 13] Z. Yu. *C-arm Computed Tomography with Extended Axial Field-of-View*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2013.
- [Yu 15] Z. Yu, A. Maier, G. Lauritsch, F. Vogt, M. Schoenborn, C. Koehler, J. Hornegger, and F. Noo. “Axially Extended-Volume C-Arm CT Using a Reverse Helical Trajectory in the Interventional Room”. *IEEE Transactions on Medical Imaging*, Vol. 34, No. 1, pp. 203–215, Jan 2015.
- [Yush 06] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig. “User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability”. *NeuroImage*, Vol. 31, No. 3, pp. 1116–1128, Jul 2006.
- [Zbij 11] W. Zbijewski, P. De Jean, P. Prakash, Y. Ding, J. W. Stayman, N. Packard, R. Senn, D. Yang, J. Yorkston, A. Machado, J. A. Carrino, and J. H. Siewerdsen. “A dedicated cone-beam CT system for musculoskeletal extremities imaging: Design, optimization, and initial performance characterization”. *Medical Physics*, Vol. 38, No. 8, pp. 4700–4713, Aug 2011.
- [Zell 05] M. Zellerhoff, B. Scholz, E.-P. Ruehrnschopf, and T. Brunner. “Low contrast 3D reconstruction from C-arm data”. In: *Proc. SPIE 5745*, pp. 646–655, 2005.
- [Zeng 05] R. Zeng, J. A. Fessler, and J. M. Balter. “Respiratory motion estimation from slowly rotating x-ray projections: Theory and simulation”. *Medical Physics*, Vol. 32, No. 4, pp. 984–991, Apr 2005.
- [Zeng 07] R. Zeng, J. A. Fessler, and J. M. Balter. “Estimating 3-D respiratory motion from orbiting views by tomographic image registration”. *IEEE Transactions on Medical Imaging*, Vol. 26, No. 2, pp. 153–163, Feb 2007.
- [Zeng 09] G. L. Zeng. *Medical Image Reconstruction: A Conceptual Tutorial*. Springer, Berlin, Germany, 2009.
- [Zhan 10] Y. Zhang and J. M. Jordan. “Epidemiology of Osteoarthritis”. *Clinics in Geriatric Medicine*, Vol. 26, No. 3, pp. 355–369, Aug 2010.

- [Zhu 12a] Z. Zhu and G. Li. “An automatic 2D-3D image matching method for reproducing spatial knee joint positions using single or dual fluoroscopic images”. *Computer Methods in Biomechanics and Biomedical Engineering*, Vol. 15, No. 11, pp. 1245–1256, Jan 2012.
- [Zhu 12b] Z. Zhu, D. F. Massimini, G. Wang, J. J. P. Warner, and G. Li. “The accuracy and repeatability of an automatic 2D-3D fluoroscopic image-model registration technique for determining shoulder joint kinematics”. *Medical Engineering & Physics*, Vol. 34, No. 9, pp. 1303–1309, Nov 2012.