

Computer-Assisted Methods for Complex Reconstructions in Orthopaedic Surgery

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

Lazaros Vlachopoulos
Doctor of Medicine
RWTH Aachen University
born on 19.10.1978
citizen of Greece and Germany

accepted on the recommendation of

Prof. Dr. Gábor Székely, examiner
Prof. Dr. Jess Gerrit Snedeker, co-examiner
Dr. Philipp Färnstahl, co-examiner

Abstract

Restoration of the normal anatomy is one of the ultimate goals in orthopaedic surgery. Computer-assisted methods become, thereby, increasingly important as they support the surgeons to achieve this goal. The transfer of computer vision and engineering methods to the medical domain offers new opportunities for traditional orthopaedic surgery. The benefits of computer-assisted surgery have been well accepted in the last decades, in particular for the treatment of posttraumatic deformities. The combination of an accurate preoperative quantification of a deformity, the generation of a preoperative plan and tools for the intraoperative navigation is the key element for a patient-specific quantitative orthopaedic surgery. However, the basic pre-requisite for a successful surgical treatment is a precise preoperative assessment of the pathological condition.

The aim of this work is to provide new computer-assisted methods for the reconstruction of complex fractures and posttraumatic deformities of the humerus. The presented methods are based on patient-specific bone models, created from computer tomography data. A set of tools will be presented to tackle a wide range of treatment options, i.e., for fracture reconstruction and for corrective osteotomies of the humerus.

First, algorithms are developed to quantify the anatomy of the humerus in 3D in a standardized fashion. Bilateral differences of the humeral anatomy are measured on a set of 140 paired humeri. This enables to identify the anatomical characteristics that may be reliably carried over from the contralateral side and to evaluate the reliability of the contralateral anatomy as a reconstruction template.

Thereafter, the preoperative assessment of a humeral deformity for corrective osteotomies of the humerus is improved with two proposed approaches. A novel segment selection strategy is presented for the registration-based approach that compensates bilateral differences better than the state-of-the-art approach, yielding an improved method for the 3D assessment of posttraumatic deformities of the proximal and distal humerus. Further, a statistical shape model of the humerus is generated that accurately predicts the pretraumatic anatomy from the healthy parts of the posttraumatic bone model in order to replace the contralateral anatomy as a reconstruction template.

The main part of the thesis focuses on the development of a computer-assisted fracture reconstruction algorithm for the reconstruction of complex proximal humeral fractures. The pose-invariant algorithm performs the reconstruction solely based on the information of the fracture surfaces obtained from the posttraumatic condition. The fractures of the proximal humerus are automatically reconstructed based on iterative pairwise reduction of the fragments and automatic selection of the best solution. The performance of the presented algorithm is evaluated on a consecutive series of patients treated for a proximal humerus

fracture or on cadaver experiments. Furthermore, the results are compared with the state-of-the-art method.

Zusammenfassung

Die Wiederherstellung der normalen anatomischen Verhältnisse ist eines der wichtigsten Ziele in der orthopädischen Chirurgie. Computergestützte Techniken gewinnen dabei zunehmend an Bedeutung, da sie den Chirurgen beim Erreichen dieses Zieles unterstützen können. Die Verwendung von Verfahren aus Computer Vision und den Ingenieurwissenschaften im medizinischen Bereich eröffnet neue Möglichkeiten für die traditionelle orthopädische Chirurgie. Dabei sind die Vorteile der computergestützten Chirurgie in den letzten Jahrzehnten weitgehend akzeptiert worden, insbesondere für die Behandlung von posttraumatischen Deformitäten. Die Kombination aus einer präzisen präoperativen Analyse einer Deformität, der Erstellung eines präoperativen Plans und eines intraoperativen Navigationssystems sind der Schlüssel für eine patientenspezifische quantitative orthopädische Chirurgie. Grundvoraussetzung für eine erfolgreiche chirurgische Behandlung ist jedoch die genaue präoperative Analyse der zugrundeliegenden Pathologie.

Diese Dissertation setzt sich das Ziel, neue computergestützte Techniken für die Rekonstruktion von komplexen Frakturen und posttraumatischen Deformitäten am Humerus zur Verfügung zu stellen. Die dargestellten Methoden basieren auf patientenspezifischen Knochenmodellen aus computertomographischen Daten. Es wird eine umfassende Zusammenstellung von Methoden dargestellt, um eine breite Palette von Behandlungsoptionen abzudecken, d.h. im Speziellen sowohl für die Rekonstruktion von Frakturen als auch für die Korrektur von Deformitäten am Humerus.

Zunächst werden Algorithmen entwickelt, um die Geometrie des Humerus in 3D standardisiert zu quantifizieren. Bilaterale Unterschiede in der Anatomie des Humerus werden an einem Satz von 140 gepaarten Humeri bestimmt. Dies ermöglicht es, die anatomischen Eigenschaften zu identifizieren, welche von der kontralateralen Seite übernommen werden können, und somit die Zuverlässigkeit der kontralateralen Anatomie als geometrische Vorlage zu beurteilen.

Anschließend wird die präoperative Analyse einer posttraumatischen Deformität des Humerus mit zwei vorgeschlagenen Planungsansätzen verbessert. Eine neue Strategie wird vorgestellt, um Segmente für die Registrierung von Knochenmodellen zu selektieren, die es ermöglicht, die bilateralen Unterschiede besser auszugleichen als der aktuelle Stand der Technik. Dies führt zu einer genaueren Wiederherstellung der prätraumatischen Anatomie am proximalen und distalen Humerus. Darüberhinaus wird ein statistisches Modell des Humerus präsentiert, das die prätraumatische Anatomie aus den nicht pathologischen Anteilen des

posttraumatischen Knochenmodells genau vorhersagt mit dem Ziel, die kontralaterale Anatomie als Rekonstruktionsvorlage zu ersetzen.

Der Hauptteil dieser Arbeit widmet sich der Entwicklung eines computergestützten Algorithmus für die Rekonstruktion von komplexen Frakturen am proximalen Humerus. Der Algorithmus ermöglicht die Rekonstruktion von Frakturen ausschließlich basierend auf der Information der Bruchflächen anhand von posttraumatischen Knochenmodellen. Die Frakturen des proximalen Humerus werden dabei durch eine iterative, paarweise Reposition der Fragmente automatisch rekonstruiert und die optimale Lösung wird automatisch bestimmt. Die Leistungsfähigkeit der Methode wird an einer konsekutiven Serie von Patienten mit einer proximalen Humerusfraktur und mit Kadaverexperimenten verifiziert. Darüberhinaus, werden die Ergebnisse dem aktuellen Stand der Technik gegenübergestellt.

Acknowledgements

The research presented in this thesis has been supported by the Balgrist Foundation, Zurich Switzerland.

First and foremost, I would like to thank my advisor Prof. Gábor Székely for giving me the opportunity to perform my PhD studies in his research group at the Computer Vision Lab. I am very grateful for his expert advice, valuable suggestions and for reminding me from time to time of my medical background. I want also to express my profound gratitude to my supervisor Dr. Philipp Färnstahl for introducing me in computer science, his expert advice, the collaborative work and for his endless support throughout my research.

Further, I would like express my gratitude to Prof. Christian Gerber, who expressed his trust and confidence in our research and who enabled my excursion into computer vision.

I am also grateful to Prof. Dr. Jess Gerrit Snedeker, for reviewing this thesis and serving as a committee member of my doctoral defence.

Many thanks go to my colleagues at the Computer Assisted Research and Development Group at the University Hospital Balgrist. In particular, I would like to thank Fabio Carrillo, Matthias Graf, Michael Tschannen, Celestine Dünner, Michael Hecht, Loris Diana and Florentin Liebermann for the fruitful discussions and their great support.

I want to thank my parents and my family for the endless support and for all of the sacrifices that you have made on my behalf. Finally, I am infinitely thankful and happy to my love and our wonderful children. Words alone cannot express how grateful I am to you.

Contents

List of Figures	xi
------------------------	-----------

List of Tables	xiii
-----------------------	-------------

1	Introduction	1
1.1	Background	1
1.2	Surgical Treatment Options.....	2
1.3	Computer-Assisted Approaches	2
1.4	Thesis Outline.....	3
2	Three-Dimensional Measurement of Humeral Anatomy	5
2.1	Introduction	5
2.2	Material and Methods.....	7
2.2.1	Statistical Analysis	11
2.3	Results	11
2.4	Discussion	17
2.5	Conclusions	19
3	Assessment of Posttraumatic Deformities of the Distal Humerus	21
3.1	Introduction	21
3.2	Material and Methods.....	23
3.2.1	State-of-the-art deformity assessment.....	23
3.2.2	Effect of the segment selection on the assessment of the deformity.....	24
3.2.3	Novel automatic segment selection approach	25
3.2.4	Statistical Analysis	27
3.3	Results	28
3.4	Discussion	29
4	Approximation of the Humeral Head Retrotorsion	33
4.1	Introduction	33
4.2	Material and Methods.....	34
4.2.1	Landmark-based 3D method (LBR error).....	35
4.2.2	Registration-based approach	37
4.2.3	Statistical Analysis	41
4.3	Results	41
4.4	Discussion	45

4.5	Conclusions	47
5	Statistical Shape Model of the Humerus	49
5.1	Introduction	49
5.2	Material and Methods.....	50
5.2.1	Generation of the SSM.....	50
5.2.2	Evaluation method for the SSM.....	51
5.2.3	Reference Method	54
5.2.4	Statistical Analysis	55
5.3	Results	56
5.4	Discussion	60
5.5	Conclusions	61
6	Regression Forest Based Automatic Estimation of the Articular Margin Plane	63
6.1	Introduction	63
6.1.1	Related Work	65
6.2	Method.....	67
6.2.1	Theoretical background on RFs	68
6.2.2	AMP parameterization	69
6.2.3	Sheetness-based ray features.....	70
6.2.4	Computing a coarse estimate of the AMP parameters	71
6.2.5	Refinement step.....	73
6.3	Experiments.....	74
6.3.1	Dataset.....	74
6.3.2	Performance evaluation of the proposed method.....	75
6.3.3	Refinement using <i>f</i> box-type features	75
6.3.4	Comparison with atlas-based registration	75
6.4	Results	76
6.4.1	Refinement using <i>f</i> box-type features	78
6.4.2	Comparison with atlas-based registration	79
6.5	Discussion	79
6.6	Conclusion and Future Work	81
7	Reconstruction of Complex Proximal Humeral Fractures	83
7.1	Introduction	83
7.1.1	Related Work	87
7.2	Material and Methods.....	88
7.2.1	Overview of the Algorithm	88
7.2.2	Generation of triangular surface models	89
7.2.3	Fracture Surface Representation	90
7.2.4	Curvature Calculation	91
7.2.5	Pairwise Reduction.....	93
7.2.6	Calculation of the merged fracture line.....	96
7.2.7	Fracture Reconstruction	97
7.3	Results	100
7.3.1	Clinical evaluation	100
7.3.2	Quantitative evaluation based on cadaver experiments	102

7.3.3	Order of pairwise reduction	103
7.3.4	Runtime	103
7.4	Discussion	104
7.5	Conclusion and Future Work	107
8	Conclusion and Perspective	109
	Bibliography	113

List of Figures

2.1:	Definition of the reference planes and axes.....	8
2.2:	Definition of the angles and lengths.	9
2.3	Algorithm for determining the ETP.	10
2.4:	Bilateral differences in retrotorsion TEA and retrotorsion ETP.	12
2.5:	Bilateral differences in 3D inclination angle and the projected inclination angle.	13
2.6:	Bilateral differences in humeral length and humeral length TEA.	14
2.7:	Length of the humeri and the differences in length between both sides.	15
2.8:	Bilateral differences in head sphere radius and head height.	16
3.1:	Template-based assessment of the deformity of the distal humerus.....	24
3.2:	Effect of the segment selection on the registration result.	25
3.3:	Visualization of the humeral segments defined for the registration.	27
3.4:	Results of the HCR errors.	28
3.5:	Characteristic features of the distal humerus.	31
4.1:	Landmark-based retrotorsion error.	36
4.2:	Assessment of a posttraumatic deformity of the proximal humerus.....	38
4.3:	Effect of the segment selection on the approximation of the humeral head retrotorsion.	39
4.4:	Visualization of the humeral segments defined for the registration.	40
4.5:	Results of the p-HCR errors and LBR error.	44
4.6:	Cross-sectional shape of the groove-75-85 segment.....	46
5.1:	Visualization of the humeral segments defined for the SSM.....	51
5.2:	Evaluation of the humerus prediction error of the SSM.	53
5.3:	Quantification of the humeral length.	54
5.4:	Evaluation of the HCR error.	54
5.5:	Results of the p-HP errors and p-HCR errors.	56
5.6:	Results of the d-HP errors and d-HCR errors.	58
5.7:	Differences in humeral length.....	59
6.1:	Annotation of the AMP for hemiarthroplasty of the shoulder joint.	65
6.2:	Overview of the regression forest-based algorithm.	68
6.3:	AMP parameterization.	70
6.4:	Visualization of the oriented sheetness measure.....	71
6.5:	Illustration of a binary <i>f</i> box-type feature.	72

6.6: Position error vectors.	77
6.7: Error histograms.	77
6.8: Visualization of the AMP.	78
7.1: Terminology.	85
7.2: Overview of the reconstruction of a proximal humeral fracture.	89
7.3: Fracture surface representation.	91
7.4: Curvature Calculation.	92
7.5: Scale-space representation of the curvature function.	93
7.6: Visualization of matching intervals.	94
7.7: Calculation of the merged fracture line.	97
7.8: Importance of the order of pairwise reductions	97
7.9: Combination of pairwise reduction.	98
7.10: Approximation of the expected area of reconstruction.	99
7.11: Reconstruction results.	101
7.12: Calculation of the distance error.	103
7.13: Influence of a plastic deformation of a fragment.	105

List of Tables

3.1:	HCR errors for all segments.....	28
4.1:	3D-p-HCR error and LBR error.....	42
4.2:	P-values of the post-hoc analysis of the 3D-p-HCR error.	42
4.3:	P-values of the post-hoc analysis of the p-HCR error for the internal-/external rotation and the LBR error.	43
4.4:	p-HCR error for the internal-/external rotation and LBR error.	43
5.1:	p-HP error and p-HCR error for all segments.....	57
5.2:	P-values of the post-hoc analysis of the p-HP error and the p-HCR error.....	57
5.3:	d-HP error and d-HCR error for all segments.....	58
5.4:	P-values of the post-hoc analysis of the d-HP error and the d-HCR error.....	58
5.5:	Differences in humeral length.....	59
5.6:	P-values of the post-hoc analysis of the error in humeral length.....	60
6.1:	Position and angular error.....	76
6.2:	Position and angular error after refinement with <i>f</i> box-type features.	78
6.3:	Position and angular error after atlas-based registration.....	79
7.1:	Characteristics of the fragments of the clinical cases.	90
7.2:	Characteristics of the fragments of the cadaver experiments.	102
7.3:	Comparison of the distance errors.	103
7.4:	Comparison of the preformance of the presented method and the reference method.	104

1 Introduction

Computer-assisted orthopaedic surgery (CAOS) has gained in importance across a wide range of orthopaedic areas in the last decades. The combination of preoperative assessment of a pathological condition, generation of a preoperative plan, tools for the intraoperative navigation, and postoperative evaluation is a key for the successful application of CAOS. The restoration of the normal anatomy is, thereby, an ultimate goal for many orthopaedic procedures. However, the most fundamental pre-requisite to achieve this overall goal is the accurate preoperative assessment of the pathological condition.

The current thesis focuses on the reconstruction of complex proximal humeral fractures and of posttraumatic deformities of the humerus. In Section 1.1, an overview of the clinical relevance and the problem statement will be given. Section 1.2 will present a brief insight in surgical treatment options. In Section 1.3, the current state-of-the-art CAOS approaches for the reconstruction of the humeral anatomy will be summarized. Section 1.4 will give an outline of the new CAOS methods, which were developed in this thesis.

1.1 Background

Bone fractures are a common injury with an estimated incidence of about 4-20/1000 per year (Amin et al., 2014; Court-Brown and Caesar, 2006; Donaldson et al., 1990; Sahlin, 1990). Fractures have a bimodal age distribution and are common in the young and in the old age population. Around one-third of all children suffer, for example, at least one fracture before the age of seventeen (Cooper et al., 2004). Proximal humeral fractures account for approximately six percent of all adult fractures (Court-Brown and Caesar, 2006; Handoll and Brorson, 2015). It is, therefore, not surprisingly that the costs for the health-care system are enormous. With aging of the population the costs are expected to grow rapidly, i.e., about approximately 50% from 2012 to 2020 (Solomon et al., 2014). In an ideal case, the fracture treatment results in a restoration of the pretraumatic anatomy and the function as quickly as possible. At the same time the treatment should reduce the need and risk of repeated interventions. However, there is not enough evidence to determine the best treatment option, non-surgical or surgical, for proximal humeral fractures for every case (Cvetanovich et al., 2016; Gerber et al., 2004; Handoll and Brorson, 2015).

1.2 Surgical Treatment Options

Several characteristics influence the decision of the surgeon to recommend one of the surgical treatments, i.e. the age of the patient, the daily-life and work requirements of the patient or the displacement of the fragments. A large percentage of the fractures of the proximal humerus are neither displaced nor markedly unstable and conservative treatment gives good results (Gerber et al., 2004). Open reduction and internal fixation (ORIF) is the mainstay of therapy for displaced fractures of the young and active patient. The anatomical reconstruction is a fundamental pre-requisite for a successful joint-preserving surgical treatment. If the anatomical reduction cannot be obtained, joint replacement is a valuable treatment option (Gerber et al., 1998). The options for replacement surgery of the shoulder joint include hemiarthroplasty, anatomic total shoulder arthroplasty and reverse total shoulder arthroplasty (Cuff and Pupello, 2013; Cvetanovich et al., 2016; Fucentese et al., 2014; Grubhofer et al., 2016). If the initial goal of the treatment of a complex fracture is a joint-preserving procedure, the surgeon may face in the course of time further impairments that have to be addressed. Malunion, osteonecrosis of the humeral head or osteoarthritis are, thereby, regularly observed. The treatment of symptomatic deformities resulting from a malunited fracture is often a challenging problem (Murase et al., 2008). As an alternative to joint replacement, a corrective osteotomy, i.e., a joint-preserving surgery aiming the correction of the deformity is valuable option for the younger patients. Nevertheless, similar as for the initial surgical treatment of a fracture, the key to obtain a good functional outcome after a corrective osteotomy is also the accurate restoration of the pretraumatic anatomy.

1.3 Computer-Assisted Approaches

The optimal approach for the preoperative assessment of a deformity depends mainly on the underlying pathological condition and on the planned surgical procedure. Therefore, it is important to have a comprehensive set of tools to perform this task. In the current clinical setting, the assessment is based on the examination of plane radiographs (2D) or of CT scans (3D). The classification of humeral fractures is based on the assessment of the fracture pattern and on the amount of displacement of the fragments. The classification proposed by Neer (1970) is the most common, although it has several limitations and the inter-observer reliability is low. According to Neer, a fragment is considered as displaced if angulation exceeds 45° or if the displacement is more than 1 cm. These criteria have become the most recognizable and quotable features of the classification system and are still the standard for the assessment of a deformity of the proximal humerus, although Neer acknowledged that these criteria were defined only based on the limited available data (Carofino and Leopold, 2013). One reason for the high cut off values might be the difficulty to assess a 3D deformity solely based on plane radiographs or CT scans. In addition, the question concerning the tolerable limits of acceptable malalignment is very difficult to answer. CAOS approaches have the potential to accurately quantify the malalignment caused by a fracture or by a malunion and, thereby, these methods may lead to a more differentiated classification of the fractures and selection of the optimal treatment.

In CAOS approaches, the fundamental prerequisite for a successful restoration of a fracture is the precise preoperative assessment of fragment displacement (Fürnstahl et al., 2012).

Jiménez-Delgado et al. (2016) recently published a comprehensive review article, summarizing the currently proposed approaches for the reconstruction of bone fractures. However, an evaluation of these methods is mandatory before they can be applied to a new anatomy. Hitherto, one method (Förnsthahl et al., 2012) has been published and validated for the computer-assisted reconstruction of complex proximal humerus fractures (Jimenez-Delgado et al., 2016). Förnsthahl et al. (2012) demonstrated that their algorithm allows accurate reconstruction of the pretraumatic anatomy, but the requirement to use the contralateral healthy side as a reconstruction template limits the applicability of their method.

The benefit of CAOS methods for the treatment of malunited fractures has also been emphasized. The essential reason for this is that the assessment of the deformities based on conventional preoperative planning (i.e., with 2D plain radiographs) does not provide sufficient information to understand complex 3D deformities (Murase et al., 2008). Furthermore, recent studies demonstrated that corrective osteotomies can be performed more precisely using computer simulation combined with patient-specific instruments for intraoperative navigation (Murase et al., 2008; Omori et al., 2015; Takeyasu et al., 2011; Takeyasu et al., 2013; Vlachopoulos et al., 2016b). However, the key element for a successful application of these state-of-the-art methods for corrective osteotomies is, similar as for fracture reconstruction algorithms, the accurate preoperative assessment of the pathological condition. For corrective osteotomies of the humerus the state-of-the-art method for the deformity assessment is a registration-based approach using the contralateral anatomy as reconstruction template.

1.4 Thesis Outline

The aim of the study is to provide new CAOS methods for patient-specific quantitative orthopaedic surgery. The thesis focuses on the precise preoperative assessment of pathological conditions and the surgical planning of challenging reconstructive surgeries of the humerus (i.e., fractures, malunions, and shoulder replacement surgeries). The methods shall enable the surgeon to achieve a more accurate anatomical reconstruction and a better clinical outcome, if applied in combination with intraoperative navigation.

Chapter 2 focuses on the development of algorithms for the 3D-measurement of the humeral anatomy. The methods allow transferring previous knowledge from 2D assessment of pathologies to the 3D space, forming a basis for the translation of conventional orthopaedic surgery to CAOS. In addition, the differences in the morphology of the humeral anatomy between both sides of an individual are analysed and quantified to validate whether the contralateral anatomy is a reliable reference for the quantification of a deformity.

Chapter 3 to 5 are dedicated to a description of algorithm for the template-based approaches of the assessment of posttraumatic deformities of the humerus. In Chapter 3 and 4, the focus is on the analysis of the reliability of the contralateral anatomy as a reconstruction template for the approximation of the pretraumatic anatomy. A novel registration-based approach is presented that compensates bilateral differences better than the state-of-the-art approach, yielding an improved method for the 3D assessment of deformities of the proximal and distal humerus.

Chapter 5 extends the template-based approach to achieve independence from the contralateral anatomy. A statistical shape model is presented that accurately predicts the pretraumatic anatomy of the humerus from the healthy parts of the posttraumatic bone model. Thereby, the template-based approach is applicable even if bilateral pathological conditions are present.

Chapter 6 extends the standardized measurement of the humeral anatomy in 3D. An algorithm is presented that allows in combination with the previous methods of Chapter 2 to quantify deformities in an automated fashion. The method uses random regression forests for the estimation of a plane that approximates the cartilage/metaphyseal boundary at the proximal humerus. This is of importance for the preoperative planning of shoulder joint replacement surgery since one crucial step during shoulder arthroplasty is the resection of the humeral head along the cartilage/metaphyseal boundary.

Chapter 7 focuses on the preoperative planning task of complex proximal humeral fractures, which are commonly accompanied with a relevant displacement of the fragments. A novel pose-invariant fracture reconstruction algorithm is presented which performs the reconstruction solely based on the information obtained from the posttraumatic condition. A scale-space representation of the curvature is introduced, permitting to calculate the correct alignment between bone fragments solely based on corresponding regions of the fracture surfaces. The fractures of the proximal humerus are automatically reconstructed based on iterative pairwise reduction of the fragments. Lastly, the algorithm determines the best solution of all performed reconstructions. The performance of the presented algorithm has been evaluated on a consecutive series of patients treated for a proximal humerus fracture and on cadaver experiments. The accuracy of our approach is thereby compared to the state-of-the-art algorithm.

The thesis is concluded in Chapter 8 with a brief summary of the presented methods and an outlook in future directions of the research.

Chapters 2 and 6 have been published and Chapters 3 is currently under revision in a peer-review journal as (Vlachopoulos et al., 2016a), (Tschannen et al., 2016) and (Vlachopoulos et al., 2017b). Chapter 4, 5 and 7 are currently under review as (Vlachopoulos et al., 2017a), (Vlachopoulos et al., 2017d) and (Vlachopoulos et al., 2017c).

2

Three-Dimensional Measurement of Humeral Anatomy

In this paper, we developed algorithms for the measurement of the humeral anatomy in 3D. Further, we assessed the symmetry of the humeral anatomy between both sides of an individual to validate the reliability of the contralateral anatomy as a reconstruction template. Geometric characteristics quantifying the humeral anatomy were determined in a semi-automatic fashion on triangular surface models of 140 paired humeri (70 cadavers). The analysis revealed that several geometric quantities of the proximal humerus (i.e., humeral length, humeral head size, humeral head height) are similar between both sides of an individual and, therefore, these characteristics may be reliably carried over from the contralateral side. These findings are helpful for the shoulder replacement surgery if severe osteoarthritis or osteonecrosis is present, when the humeral head characteristics cannot be determined. However, the differences in torsion and inclination between the proximal and the distal part of the humerus of both sides were considerably higher. Determining humeral head retrotorsion and humeral head inclination from the contralateral anatomy with this landmark-based method may be more prone to error. Therefore, we concluded that preoperative planning algorithms targeting the reconstruction of the proximal humeral anatomy should not rely blindly on the contralateral anatomy but rather aim to compensate bilateral differences in the humeral anatomy.

2.1 Introduction

Restoration of the normal humeral anatomy is an ultimate goal in reconstructive surgeries of the proximal humerus, such as fracture reduction or prosthetic replacement (Boileau et al., 2008; Boileau and Walch, 1997; Hertel et al., 2002; Iannotti et al., 1992; Jeong et al., 2009; Robertson et al., 2000). Therefore, several studies have investigated the normal humeral anatomy and emphasized anatomical variability (Ballmer et al., 1993; Boileau et al., 2008; Harrold and Wigderowitz, 2012; Hernigou et al., 2002; Hertel et al., 2002; Iannotti et al., 1992; Johnson et al., 2013; Polster et al., 2010; Robertson et al., 2000).

Computer-assisted surgery, including three-dimensional (3D) analysis, 3D preoperative planning and intraoperative guidance, has become well accepted in many orthopaedic areas (Mavrogenis et al., 2013; Merloz et al., 1998; Miyake et al., 2011; Sariali et al., 2012; Schweizer et al., 2013; Tamam and Poehling, 2014). In the last decade, the progresses of this technology and its advantages have also been advocated in total shoulder arthroplasty.

Improved accuracy of component positioning in presence of osteoarthritis (Edwards et al., 2008; Iannotti et al., 2014; Kircher et al., 2009; Nguyen et al., 2009) or in the reconstruction of comminuted proximal humeral fractures (Bicknell et al., 2007) is probably a leading argument for the application of 3D preoperative planning and intraoperative navigation.

However, it is often difficult or even impossible to preoperatively reconstruct the proximal humerus anatomy solely on the basis of radiographs or computed tomography (CT) analysis of the pathological side. In complex cases, the contralateral bone may serve as a 3D reconstruction template for deriving geometric quantities. These quantities can then be used for the preoperative planning of fracture reduction or arthroplasty.

Whereas bilateral differences in humeral head inclination (Robertson et al., 2000) and humeral head height (Bicknell et al., 2007; Robertson et al., 2000) have been reported to be significant, other quantities relevant for 3D planning, such as humeral retrotorsion, humeral length, and humeral head radius, have not been identified to vary significantly (Bicknell et al., 2007; Hernigou et al., 2002; Robertson et al., 2000). However, previous studies differed particularly in sample size and measurement technique.

The aim of the current study was to develop algorithms for 3D-measurement of geometrical quantities of the humeral anatomy and to investigate side (bilateral-) differences in humeral head retrotorsion, humeral head inclination, humeral length, humeral head radius and humeral head height within a large sample size.

2.2 Material and Methods

CT full-body data of 105 specimens were provided by the Institutes for Forensic Medicine of the University of Bern and Zurich, Switzerland. The in-plane (xy) resolution of the CT scans ranged between 0.9 x 0.9 mm and 1.27 x 1.27 mm. The slice thickness varied from 0.5 mm to 0.6 mm. Thirty-five specimens were excluded from the study due to the presence of osteoarthritis or previous trauma. The average age of the individuals included in the study was 52.8 years (standard deviation [SD], 19.7 years; range, 19 – 95 years). There were 46 male and 24 female cadavers. The average height was 173.6 cm (SD; 9.3 cm; range, 152 – 195 cm) and the average weight was 73.3 kg (SD; 18 kg; range, 37 – 111 kg). Segmentation of the humerus was performed fully automatically using a previously described segmentation algorithm (Gass et al., 2014). Bilateral 3D triangular surface models were generated using the Marching cubes algorithm (Lorensen and Cline, 1987). The evaluation was performed on 140 humeri (70 paired specimens). In addition, a second investigator performed measurements on 40 humeri (20 pairs) to assess reproducibility.

Bilateral differences between the humeri of each individual were evaluated in 3D with respect to the humeral head retrotorsion, the humeral head inclination, the humeral length, the humeral head radius, and the humeral head height. To do so, the definition of reference planes and axes were necessary and 3D measurement methods had to be developed. The goal was to develop methods that are observer-independent in order to ensure reproducibility.

First, a coarse alignment of the humerus to a reference coordinate system was performed such that the y-axis of the coordinate system was aligned with the longitudinal direction of humeral shaft and the x-axis pointed towards the ventral part of the humerus. The humeral shaft axis (HSA) was then calculated as the axis minimizing the distance to the centroids of equidistantly spaced cross sections of the bone geometry (Figure 2.1, A). in a least-square sense (Schneider and Eberly, 2002). The cross-sections were created in an automatic fashion by cutting the bone model (Schroeder et al., 2003) with a predefined number of cutting planes running orthogonal to the y-axis of the reference coordinate system. Thereafter, the centroid of a plane was determined by calculating the centre of mass of the bone-model points on the cut surface. For each specimen, we used 100 cutting planes for subdivision. The centroids of planes 50 to 85, representing the proximal part of the shaft just below the humeral head, were selected for calculating the HSA (Figure 2.1, A).

The articular margin plane (AMP) was chosen to be the plane fitted to a set of reference points, manually set along the cartilage/metaphyseal boundary on the bone model of the proximal humeral head (Figure 2.1, B). Four points were placed on each, the proximal, middle and distal part of the boundary, respectively. Similar methods for defining the AMP were described previously (Bicknell et al., 2007; Johnson et al., 2013) using varying numbers of reference points. In our case, the plane was fitted to 12 points by solving the overdetermined plane equation in a least-square sense (Schneider and Eberly, 2002). As demonstrated in Figure 2.1, C and Figure 2.2, C, the centre and the radius of the humerus head were geometrically quantified by a sphere. The sphere was calculated automatically by applying a fitting technique, well established for approximating spherical joints in kinematic simulations (Thienpont, 2012; Youderian et al., 2014). The humeral head height was

expressed by the point-plane distance from the AMP to the most distant point on the surface of the humeral head (Figure 2.2, C).

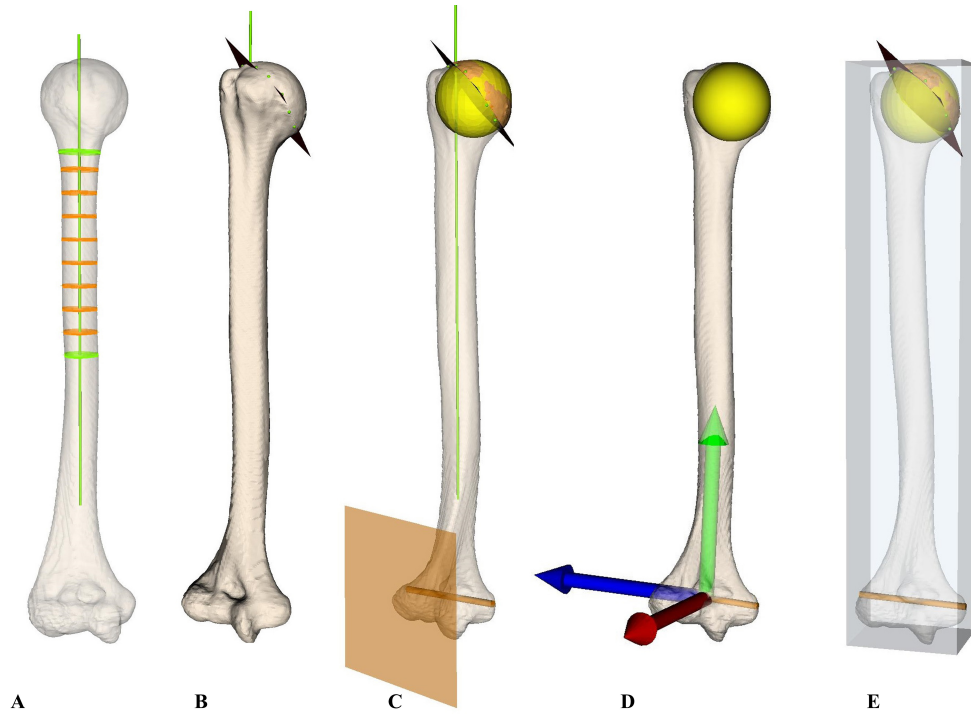


Figure 2.1: Definition of the reference planes and axes. (A) The humeral shaft axis (HSA, green) was calculated as the axis minimizing the squared distance to the centroids of equidistantly spaced cross sections (orange) between the 50th and 85th cross section (green). (B) The articular margin plane (AMP, black) was chosen to be a plane fitted to 12 reference points (green), manually set along the cartilage/metaphyseal boundary. (C) A sphere (yellow) was used to approximate the humeral head above the AMP. As references for the orientation of the distal humerus, the transepicondylar axis (TEA, orange axis) and the elbow tangent plane (ETP, orange plane) are defined. (D) Visualization of axes of the humeral coordinate system as defined by the International Society of Biomechanics: x-axis (red), y-axis (green), z-axis (blue). (E) The oriented bounding box (OBB, grey box) of the humerus model was used for initialization of the ETP and for measuring the humeral length.

As shown in Figure 2.1, C, two different representations describing the orientation of the distal part of the humerus were implemented and compared, namely the transepicondylar axis (TEA) and the elbow tangent plane (ETP). Both representations were previously compared for the 2D case (Boileau and Walch, 1997), but an extension to 3D was not reported yet. In our study, the 3D TEA was represented by a line formed by the two most distant points on the medial and lateral epicondyle. The most distant points were determined in a semi-automatic fashion as follows. First, the user defined the approximate regions of interest by setting two reference points on the medial and lateral epicondyle, respectively. The epicondyle regions were then represented by all model points lying within a 5 mm diameter around the medial and lateral reference points. Next, the most distance points were computed by evaluating the distances between all points of the two regions. Lastly, the line connecting the most distant points between those regions was calculated.

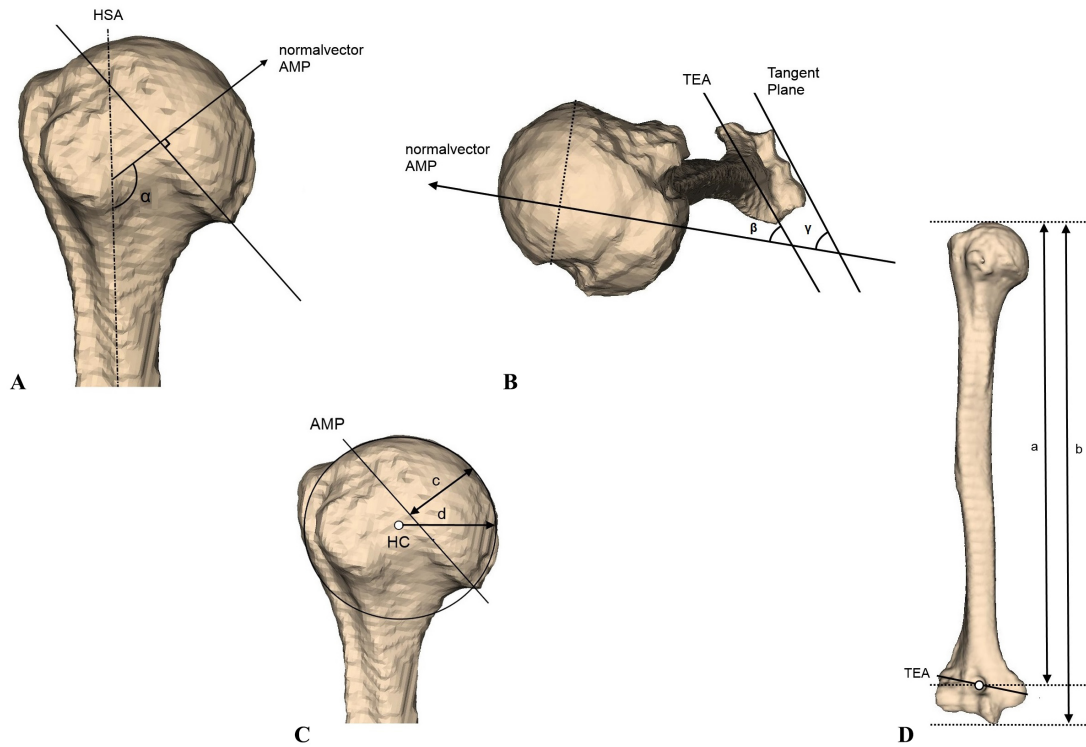


Figure 2.2: Definition of the angles and lengths. (A) The angle between the humeral shaft axis (HSA) and the normal vector of the articular margin plane (AMP) describe the inclination angle α . (B) Two different methods were used for measuring the retrotorsion. Transepicondylar axis (TEA) method: The retrotorsion was defined as the angle between the normal vector of the AMP and the vector of the TEA (retrotorsion TEA angle β). Elbow tangent plane (ETP) method: The retrotorsion was described by the angle of the normal vector of the AMP and the normal vector of the ETP minus 90° (retrotorsion ETP angle γ , respectively). (C) The humeral head height (denoted by c) was measured as the distance of the most distant humeral head point to the AMP. The humeral head radius, d , was measured as the radius of the approximated humeral head sphere. HC , head centre. (D) The humeral length (denoted by b) was defined to be the length of the longest side of the OBB. Alternatively, it was expressed by the length between the centre of the TEA and the most proximal point on the surface of the humeral head (humeral length TEA, denoted by a).

In 3D, the ETP was defined to be the plane being tangent to the trochlea and capitulum humeri and parallel to the HSA. We used the following method for calculating this plane. In a first step, a coarse initialization of the ETP was determined on the basis of the oriented bounding box (OBB) calculated from all model points of the entire humerus (Figure 2.1, E). The OBB is the minimal-volume rectangular box fully enclosing the humeral model. In our case, principal component analysis (PCA) was used to determine the orientation of the OBB (Schneider and Eberly, 2002). The face of the OBB anterior to the trochlea and capitulum humeri (Figure 2.3, A) was used to initialize the ETP. The centre of mass of the four anterior corner points of the OBB was chosen as centre of the initialized plane.

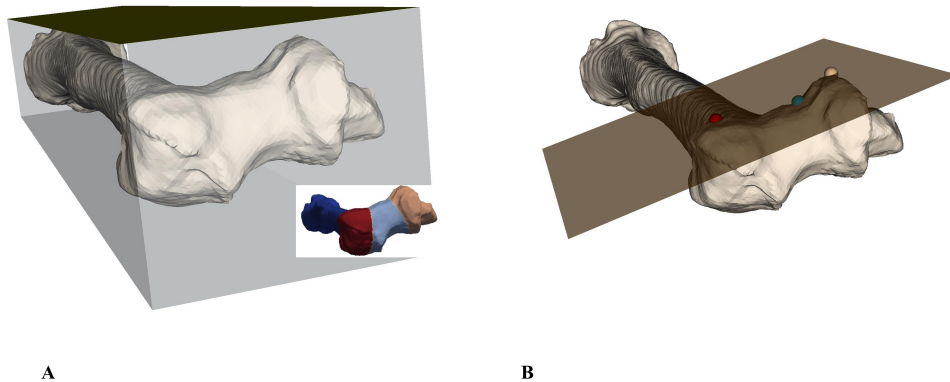


Figure 2.3 Algorithm for determining the ETP. (A) The anterior face (yellow plane) of the OBB (grey box) was used to initialize the ETP. Part of humerus distal to the TEA was subdivided into three clusters from medial to lateral (brown, blue, and red clusters) with a k-means clustering algorithm. (B) The mean of 10 points closest to the previously initialized ETP was determined (brown, blue, and red sphere). The two most lateral (red) and medial (brown) spheres represented the tangent points of the capitulum and trochlea, used for determining the ETP (orange plane).

To automatically detect the tangent points of the capitulum and trochlea humeri, in the following only the points of the distal sixth of the humerus were taken into consideration. With a k-means clustering algorithm (MacQueen, 1967) the points of the distal part of the humerus were divided from medial to lateral into clusters. Experiments showed that three clusters were sufficient to capture the capitulum and trochlea separately in one of the clusters (Figure 2.3, A). The mean of ten points that were closest to the previously initialized ETP were determined for each cluster. The most lateral and medial mean points represented the tangent points of the capitulum and trochlea, respectively (on the left- and right- more clusters in Figure 2.3, B). The normal vector of ETP was finally formed by the cross product of the directional vector of the HSA and the vector between the two tangential points on the trochlea and capitulum humeri. The centre of ETP was defined to be the tangent point of the capitulum humeri.

Conventionally, the humeral head retrotorsion is defined on CT slices as the angle between a line perpendicular to the articular surface and the transepicondylar line (Hernigou et al., 2002) or between a line perpendicular to the articular surface and a line tangent to the humeral condyles (Hempfling et al., 2001). Correspondingly, we defined the 3D retrotorsion TEA angle as the angle between the vector of the TEA and the normal vector of the AMP, both projected onto the xz-plane of the humerus coordinate system (Figure 2.2, B). The humerus coordinate system was defined by the ISB (Wu et al., 2005) (Figure 2.1, D). For calculation of the 3D retrotorsion ETP angle, the normal vector of the ETP and the normal vector of the AMP were projected onto the xz-plane of the humerus coordinate system. The angle between

those two projected vectors minus 90° then described the retrotorsion ETP angle as defined clinically (Figure 2.2, B).

In single CT slices or plane radiographs, the humeral head inclination is measured as the angle between a line perpendicular to the articular surface and the longitudinal axis of the humeral diaphysis (Iannotti et al., 1992). The 3D equivalent of inclination angle is formed by the HSA and the normal vector of the AMP. To better compare the 3D angle with the conventional one, we also evaluated the *projected* inclination angle. To do so, the HSA and the normal of the AMP were first projected onto the yz-plane of the humerus coordinate system.

The humeral length was defined to be the length of the longest side of the OBB of the humerus (*humerus length* in Figure 2.1, E and Figure 2.2, D). Additionally, the distance between the centre of the TEA and the most proximal point on the surface of the humeral head was assessed (*humeral length TEA* in Figure 2.2, D) as an alternative measurement, which can be used if a deformity of the articular surface of the elbow exists.

2.2.1 Statistical Analysis

Statistical analysis was performed using the software R (The R Foundation for Statistical Computing, Version 3.1.0, Vienna, Austria). Retrotorsion, inclination, humeral length, humeral head radius, and humeral head height were compared between the right and the left humeri using a paired t-test. Inter-rater reliability was assessed as a two-way random effects model using the interclass correlation coefficient ICC. The significance level was set at $P < 0.05$. For graphical visualization, Tukey boxplots were used with the end of the whiskers indicating 1.5 interquartile range (IQR) of the lower and upper quartile. All differences were expressed as intra-individual differences between the right and the left humeri of each individual. The mean differences were calculated from absolute values, and the range was expressed using signed values (i.e., right to left difference).

2.3 Results

The ICC was above 0.9 for the following investigator dependent parameters: retrotorsion TEA 0.91 (95% confidence interval [CI], 0.81 – 0.95), retrotorsion ETP 0.90 (95% CI, 0.81 – 0.95), head sphere radius 0.97 (95% CI, 0.94 – 0.98), inclination 0.91 (95% CI, 0.83 – 0.95). For the 3D inclination angle and the humerus head height, the ICC was 0.83 (95% CI, 0.68 – 0.91) and 0.67 (95% CI, –0.07 to 0.89), respectively. The mean difference between the two raters was 2.4° (SD, 1.7° ; range, -4.03° to 5.77°) for the 3D inclination angle and 1.7 mm (SD, 1.02 mm; range, –4.3 mm to 0.64 mm) for the humeral head height.

There was no significant bilateral-difference in the humeral head retrotorsion, neither in the retrotorsion TEA angle (mean difference, 6.9° ; SD, 5.3° ; range, -12.3° to 27.8° ; $P = 0.175$) nor in the retrotorsion-ETP angle (mean difference: 6.7° ; SD, 5.7° ; range, -15.1° to 24.0° ; $P = 0.063$). On average, the difference between these methods was 5.6° (95% CI, $4.9^\circ - 6.3^\circ$), which was significant ($P < 0.0001$). Figure 2.4 shows the retrotorsion TEA and retrotorsion ETP of all right and left humeri. The intra-individual side differences between the retrotorsion TEA and retrotorsion ETP are given in Figure 2.4 A and B, respectively.

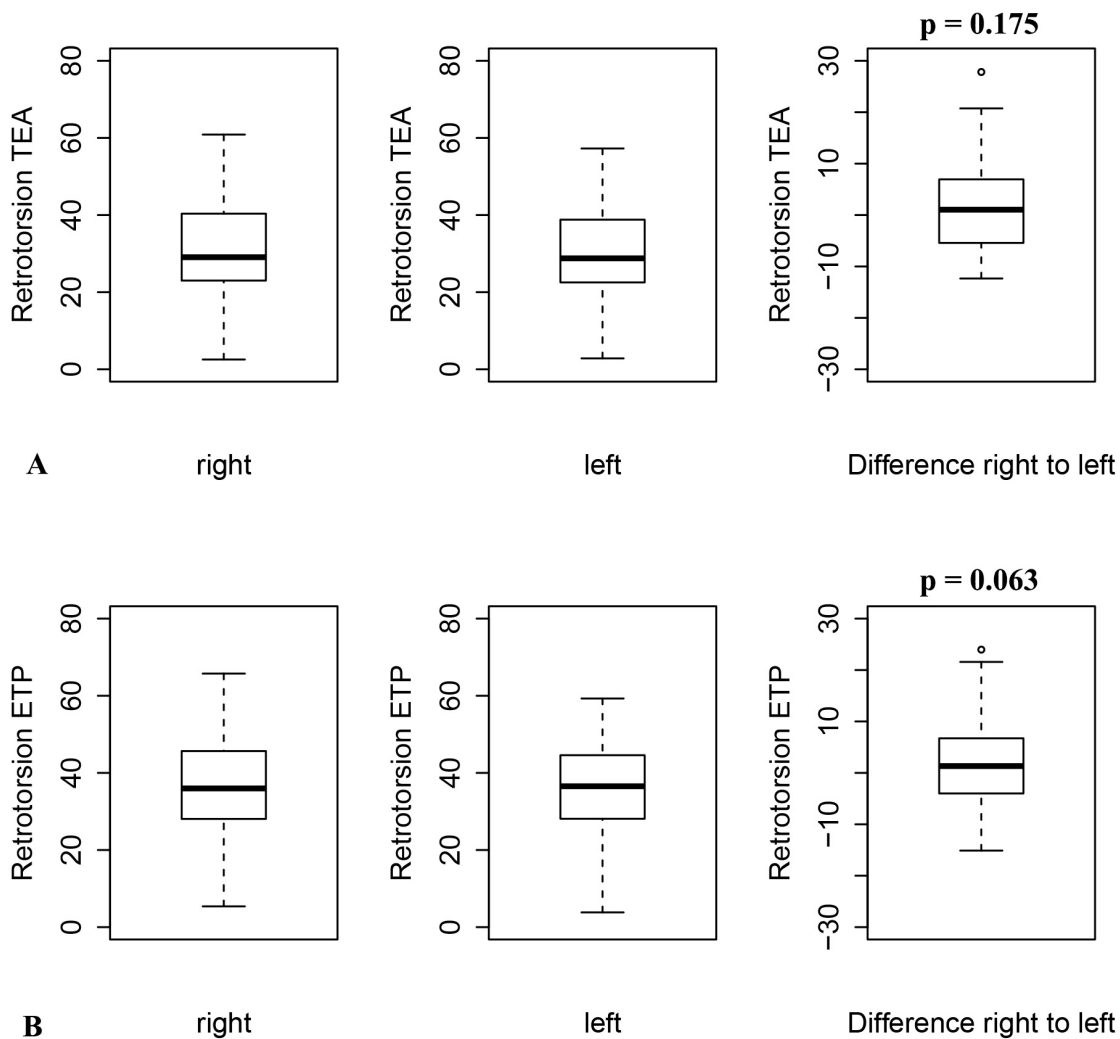


Figure 2.4: The bilateral differences in retrotorsion TEA (A) and retrotorsion ETP (B) were not significant. In addition, the Tukey box plots of the absolute values of all right and left humeri are illustrated, with the whiskers indicating 1.5 interquartile range.

Figure 2.5 reveals the bilateral differences of the 3D inclination angle and the projected inclination angle. The bilateral difference in 3D inclination angle was not significant (mean difference, 2.3°; SD, 1.8°; range, -5.1° to 8.4°; $P = 0.12$). The bilateral difference in the projected inclination was significant (mean difference, 3.1°; SD, 2.7°; range, -8.5° to 13.5°; $P = 0.008$).

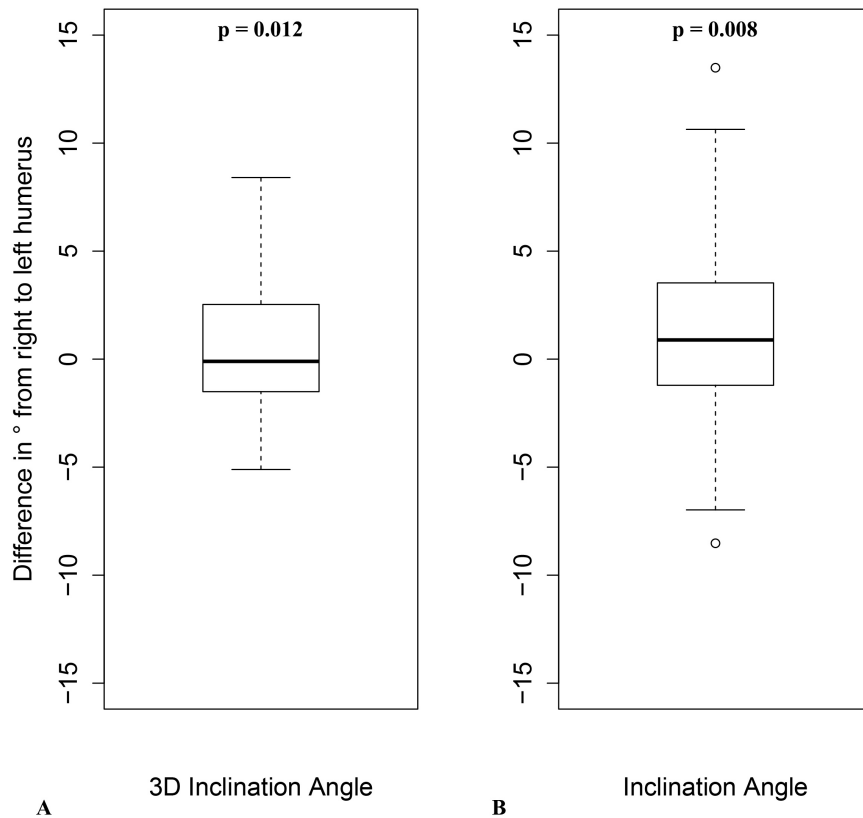


Figure 2.5: Tukey box plots of bilateral differences in 3D inclination angle (A) and the projected inclination angle (B) are given, with the whiskers indicating 1.5 interquartile range.

In Figure 2.6, the bilateral differences of the humeral length and the humeral length TEA are shown, and Figure 2.7 depicts the relation of the humeral length of the right and left humeri of all specimens. The humeral length differed significantly (mean difference, 2.9 mm; SD, 2.5 mm; range, -8.7 mm to 10.1 mm; $P = 0.04$). The bilateral differences were not significant for the humerus length TEA (mean difference, 3.4 mm; SD, 2.8 mm; range, -8.7 mm to 13.4 mm; $P = 0.05$)

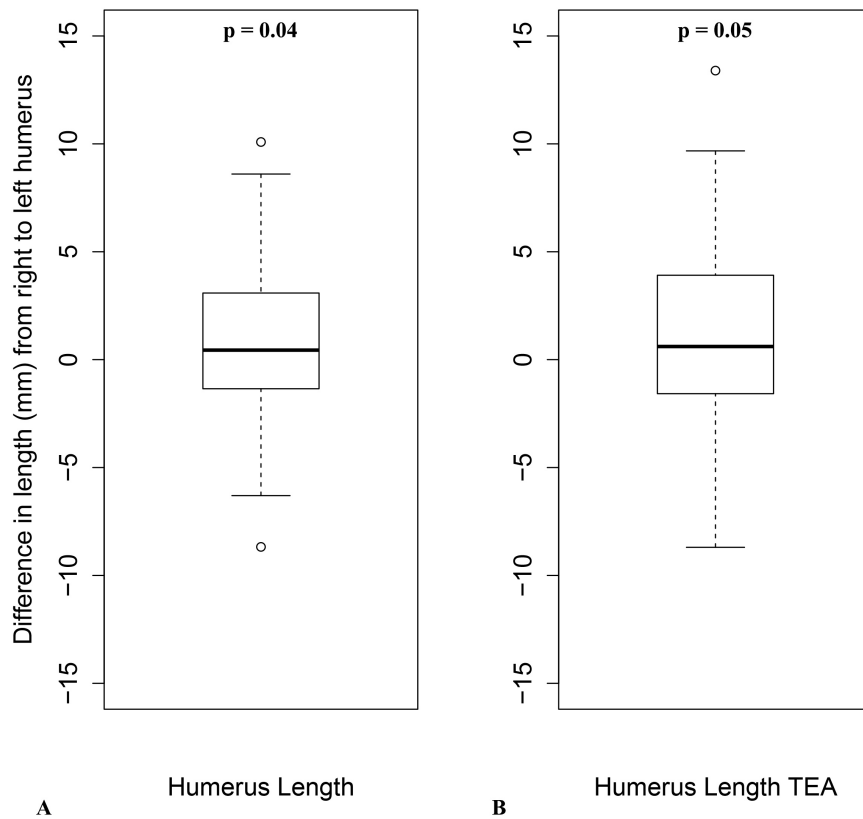


Figure 2.6: Tukey box plots illustrate bilateral differences in humeral length (A) and humeral length TEA (B), with the whiskers indicating 1.5 interquartile range.

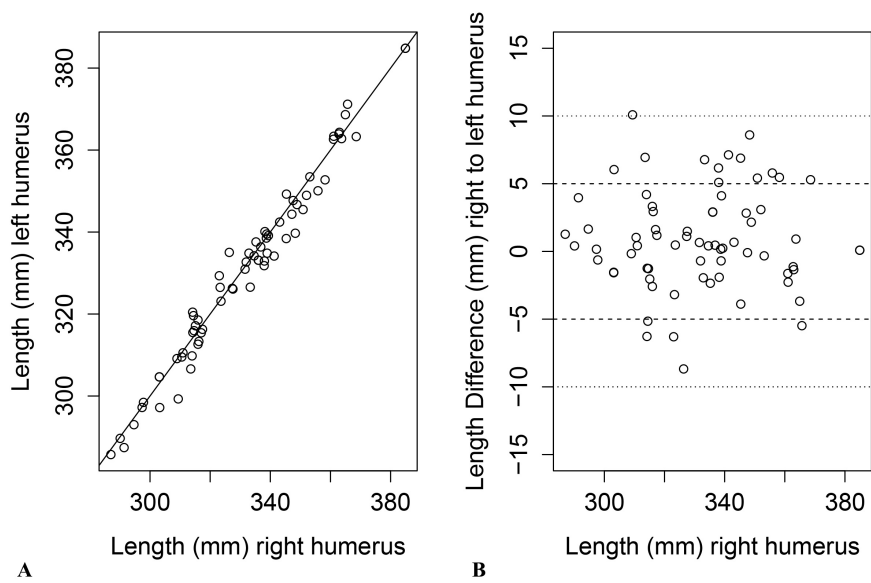


Figure 2.7: The length of the left and the right humeri of all 70 cadavers (A) and the differences in length of the humeri (B) are plotted against the length of the right humeri.

In Figure 2.8, the bilateral differences in head sphere radius and humerus head height are depicted. There was no significant bilateral difference in the head sphere radius (mean difference, 0.5 mm; SD, 0.6 mm; range, -3.2 mm to 2.2 mm; $P = 0.38$) or in the humerus head height (mean difference, 0.8 mm; SD, 0.6 mm; range, -2.4 mm to 2.4 mm; $P = 0.44$).

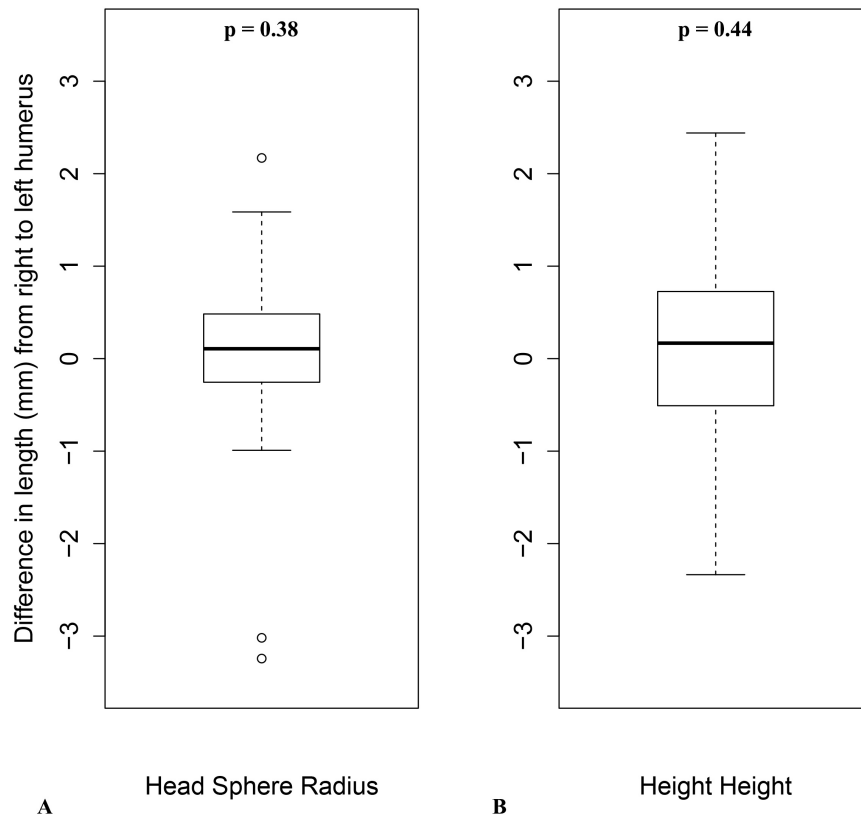


Figure 2.8: Tukey box plots illustrate bilateral differences in head sphere radius (A) and head height (B), with the whiskers indicating 1.5 interquartile range.

2.4 Discussion

Reconstruction of the normal proximal humeral anatomy in anatomical shoulder replacement surgery is a challenging procedure, especially in the presence of severe osteoarthritis or proximal humeral fracture. Alteration in anatomy may result in abnormal kinematics, or abnormal contact between the proximal humerus with the glenoid component or the rotator cuff. Moreover, alteration may cause reduction in the overall joint range of motion or influence the durability of the prosthetic replacement (Ballmer et al., 1993; Boileau and Walch, 1997; Hertel et al., 2002; Iannotti et al., 1992; Polster et al., 2010). Therefore, deriving relevant geometric parameters such as humeral head retrotorsion, inclination, humeral length and humeral head size can facilitate preoperative planning and accurate reconstruction during shoulder replacement surgery (Ballmer et al., 1993; Boileau and Walch, 1997; Johnson et al., 2013; Youderian et al., 2014). If 3D planning and computer-assisted surgery in total shoulder arthroplasty are desired (Edwards et al., 2008; Frnstahl et al., 2012; Kircher et al., 2009; Nguyen et al., 2009; Schroeder et al., 2003; Youderian et al., 2014), methods for reproducible 3D quantification of the humeral anatomy are essential. We therefore developed computer algorithms for performing 3D measurements of the humeral anatomy and investigated bilateral differences of relevant geometric parameters, which can then play a role when inferring anatomical information for the pathological side from the healthy side.

One of the main limitations of the study is the definition of the AMP because many of the geometric quantities are based on this plane. Its determination still remains time-consuming and investigator dependent, as previously noted (Bicknell et al., 2007; Boileau and Walch, 1997; Johnson et al., 2013), leading to a large inter-rater variability. An automatic method to reduce variability in estimating the parameters of AMP is highly desirable and will be addressed in future work. Another limitation is that hand dominance was not taken into account in the evaluation of bilateral differences, as this information was not available to us.

We have evaluated two novel methods for assessing the retrotorsion in 3D based on the TEA and ETP, respectively. For the TEA, we presented a semiautomatic method to avoid relying only on the manual definition of the two epicondyles points (Johnson et al., 2013), because in many cases it is difficult to identify the most prominent points visually in a consistent manner (Matsumura et al., 2014). The ETP can be determined fully automatically without any user input using the proposed k-means based clustering algorithm. Nevertheless, both methods have value, and either may be desirable, depending on the particular clinical case (e.g., in the presence of an altered elbow, one of those axes/planes may not be definable).

On average, the assessed retroversion differed by 5.6° between these methods (95% CI, $4.9^\circ - 6.3^\circ$) which was significant ($P < 0.0001$). These results were similar to those in studies of Boileau et al. (Boileau et al., 2008; Boileau and Walch, 1997), who reported a mean difference of 3.6° between measuring methods using either the transepicondylar axis or the axis tangent to the trochlea and capitulum humeri as a distal reference. Both measurement methods in the current study showed similar results when measuring the bilateral difference in retrotorsion. The mean bilateral differences of retrotorsion TEA and ETP were 6.9° (SD, 5.3°) and 6.7° (SD, 5.7°), respectively.

The current study supports the results of the previous studies regarding the existence of intra-individual bilateral-difference in humeral head retrotorsion. DeLude et al. (2007) demonstrated a mean bilateral-difference of 7.9° , similarly to our study. However, Hernigou et al. (2002) reported a mean difference of only 2.1° between the sides and Matsumura et al. (2014) reported on a mean difference of 3° between the dominant and the non-dominant side. In contrast to the present study, the studies of Hernigou et al. and Matsumura et al. analysed the retrotorsion with a 2D method on single CT slices, resulting in limited degrees of freedom for defining the geometrical quantities, especially the AMP. In addition, both studies (Hernigou et al., 2002; Matsumura et al., 2014) did not mention whether their reported mean difference is based on absolute or signed values.

The bilateral difference of retrotorsion in the current study was $> 10^\circ$ in 18 out of 70 pairs (12 right and 6 left humeri), corresponding to 25.8% of the individuals. This ratio was slightly smaller than in the study of DeLude et al. (2007) - 35.7% with 5 of 14 pairs. The mean retrotorsion over all humeri (right humeri, 31.3° , [SD, 12.4°]; left humeri, 29.9° , [SD, 11.5°]) was also smaller in the corpus of our study compared to that of DeLude (2007) (right humeri, 41.1° , [SD, 7.8°]; left humeri, 35.6° , [SD, 9.1°]).

The 3D inclination angle, being the angle between the HSA and the normal of the AMP, did not differ significantly between the sides of an individual (mean difference, 2.3° ; SD, 1.8°). The range of the bilateral-differences (right to left) was -5.1° to 8.4° , whereas 6 of the 70 cadavers had a bilateral difference $> 5^\circ$. In contrast to the 3D inclination angle, the mean bilateral difference of the projected inclination angle was higher as well as the range (mean difference, 3.1° ; SD, 2.7° ; range, -8.5 to 13.5°). The only study reporting a significant intra-individual difference in humeral head inclination was by Robertson et al. (2000). However, no quantities were provided in their study that would allow comparison of our results.

Although the bilateral-difference of the humeral length was small (mean difference, 2.9 mm; SD, 2.5 mm), it was significant ($P = 0.04$) in the current study. The difference was similar to the one reported by DeLude et al. (2007) (mean difference, 3.5 mm; SD, 3.0 mm). However, their reported difference was not significant, probably because of the sample size of 14 cadavers. DeLude et al. (2007) reported a mean humerus length of 278 mm. This is in contrast to previous studies in which an average humeral length between 316 mm (Hertel et al., 2002) and 330 mm (Robertson et al., 2000) was measured, similarly to our results (i.e., 331 mm). DeLude et al. (2007) concluded that the anatomical corpus in their study had a bias towards smaller specimen. Additionally, the overall smaller anatomy may indicate that the pool of their study was not from the same ethnic group as in the other studies (Hertel et al., 2002; Robertson et al., 2000).

In our study, the bilateral difference of the head sphere radius (mean difference, 0.5 mm; SD, 0.6 mm) and the humerus head height (mean difference, 0.8 mm; SD, 0.6 mm) was small and not significant. DeLude et al. (2007) reported a significant bilateral difference only for the humeral head height (mean difference, 1.1 mm; SD, 1 mm; $P = 0.003$). In addition, the humeral head height and the humeral head radius of their study were smaller than in the previous studies (Hertel et al., 2002; Iannotti et al., 1992; Robertson et al., 2000) and this current one. This may be due to the bias towards smaller specimens, as already mentioned.

Only Robertson et al. (2000) has reported a significant bilateral difference in humeral head height, but without providing quantitative results.

To the best of our knowledge, no clinical studies have yet evaluated whether a length difference of 2.9 mm (SD, 2.5 mm) would have an influence on the clinical outcome if the contralateral humerus were used as a reconstruction template. However, tuberosity malpositioning during hemiarthroplasty for complex proximal humeral fractures could be caused by incorrect restoration of the humeral length. A biomechanical evaluation (Huffman et al., 2008) demonstrated that 10 mm and 20 mm inferior tuberosity malpositioning during hemiarthroplasty led to significant alterations in glenohumeral joint forces; smaller malpositioning was not investigated. Whereas 17 of the 70 cadavers in our study had a difference between 5 mm and 10 mm, only one cadaver had a difference of > 10 mm (i.e., 10.1 mm). If, especially in cases of comminuted proximal humeral fractures, the length of the humerus is restored according to the 3D measurement on the contralateral side, in almost all cases the deviation from the pretraumatic humeral length would probably be < 10 mm. Therefore, using accurate measurement methods preoperatively is necessary to avoid malpositioning. In a recent study (Alolabi et al., 2014), a mismatch of the humeral head prosthetic component size of > 2 mm with respect to the anatomical humeral head size was considered to be an improper selection since in most commercial companies a 2-mm-radius difference would indicate that a different humeral head size is more suitable. However, a radiologic assessment of 125 anatomic shoulder arthroplasties (Alolabi et al., 2014) revealed that in $> 54\%$ of the cases the deviation was > 2 mm. In the current study, the mean bilateral difference in head sphere radius was 0.5 mm and only 3 of the 70 specimen had a side difference > 2 mm. Favre et al. (2008) showed that superior levels of humeral head height of 2.5 mm to 5 mm caused inferior impingement during the physiologic range of motion. In the current study, none of the 70 specimens had a bilateral difference in humeral head height > 2.5 mm. In the case of a unilateral pathologic change, the determination of the humeral head size on the contralateral side may yield more accurate results. Initial malpositioning of the hemiarthroplasty for proximal humeral fractures with excessive height and retrotorsion is associated with a failure of the tuberosity osteosynthesis and poor functional outcome (Boileau et al., 2002). However, for humeral head retrotorsion and humeral head inclination the maximum acceptable deviation of the individual anatomy for the reconstruction of the proximal humerus is not yet known. Therefore, malpositioning of the humeral component with a deviation in torsion of 28° and in inclination of 8° – maximal bilateral differences in the current study - when relying on the contralateral anatomy might influence the clinical outcome.

2.5 Conclusions

The current study presents semiautomatic computer algorithms for the measurement of the humeral anatomy. The analysis of 140 paired humeri indicates that humeral head size, humeral head height and - with some limitations - humeral length may be reliably carried over from the contralateral side. In contrast, deriving humeral head retrotorsion and humeral head inclination from the contralateral anatomy (e.g. for reconstruction surgery after a proximal humeral fracture) may be more error-prone. Therefore, preoperative planning approaches

targeting the reconstruction of complex proximal humerus fractures should not solely rely on the contralateral anatomy. These findings and algorithms regarding bilateral differences in retrotorsion and inclination will be incorporated into our future work in a preoperative planning framework for arthroplasty and fracture reconstruction.

3

Assessment of Posttraumatic Deformities of the Distal Humerus

In this paper, we present a new registration approach for the template-based assessment of posttraumatic deformities of the distal humerus. The method uses the contralateral anatomy as a reconstruction template, similar as the state-of-the-art approach. However, due to intra-individual bilateral differences, especially in humeral torsion, the current state-of-the-art approach introduces a biased deformity assessment. We investigated the effect of the selected segments on the 3D registration-based deformity assessment. We demonstrated that with a new segment selection strategy the presented approach compensates bilateral differences better than the state-of-the-art approach, yielding an improved method for the 3D assessment of distal humeral deformities.

3.1 Introduction

Three-dimensional (3D) computer-based methods increasingly gain in importance in the surgical treatment of distal humerus deformities (Murase et al., 2008; Omori et al., 2015; Takeyasu et al., 2011; Takeyasu et al., 2013; Vlachopoulos et al., 2016b). A recently published study (Omori et al., 2015) demonstrated that corrective osteotomies of the distal humerus can be performed precisely using 3D computer simulation preoperatively, and patient-specific surgical guides in the surgery. The main advantage of the preoperative simulation is the potential to assess the deformity in 3D, permitting to capture even small deviations from the normal anatomy in all anatomical planes. For example, a cubitus varus deformity resulting from a supracondylar fracture has frequently an accompanying extensional and/or rotational component (Takeyasu et al., 2011).

The state-of-the-art 3D deformity assessment relies on the comparison of the pathological bone with a reconstruction template representing the normal anatomy. Currently, the contralateral anatomy is the most commonly used reconstruction template (Bicknell et al., 2007; McDonald et al., 2009; Murase et al., 2008; Omori et al., 2015; Vlachopoulos et al., 2016a; Vlachopoulos et al., 2016b), since most geometric quantities are similar between both sides of an individual (Bicknell et al., 2007; McDonald et al., 2009; Vlachopoulos et al., 2016a). A surface registration method is applied to superimpose the model of the pathological side to the mirrored model of the contralateral side (Murase et al., 2008; Murase et al., 2014; Takeyasu et al., 2013; Vlachopoulos et al., 2016b). In previous studies (Murase et al., 2008; Murase et al., 2014; Takeyasu et al., 2013) the whole proximal segment of the humerus (i.e.,

humeral head, greater tuberosity and shaft) has been proposed to superimpose the models. However, especially existing differences of the humeral torsion between both sides of an individual (Bicknell et al., 2007; Vlachopoulos et al., 2016a) might introduce a considerable error if the deformity is assessed with this method (Murase et al., 2008; Murase et al., 2014; Takeyasu et al., 2013).

In the present study, we have systematically selected different segments of the humeral shape and evaluated the effect on the result of the registration with the reconstruction template with respect to the robustness against bilateral differences. We hypothesize that a new segment selection strategy may compensate bilateral differences better than the state-of-the-art approach, yielding an improved method for the 3D assessment of distal humeral deformities.

3.2 Material and Methods

The Swiss Institute for Computer-Assisted Surgery (SICAS) provided computed tomography (CT) full-body data of 50 specimens (including the entire humerus on both sides). The in-plane (xy-) resolution of the CT scans ranged between 0.9 x 0.9 mm and 1.27 x 1.27 mm. The slice thickness varied from 0.5 mm to 0.6 mm. The average age of the individuals was 52.1 years (SD 20.0 years; range, 19 to 90 years). There were 32 male and 18 female cadavers. The average height was 172.4 cm (SD 8.7 cm; range, 154 to 187 cm) and the average weight was 68.4 kg (SD 16.9 kg; range, 37 to 108 kg). Exclusion criteria were the presence of osteoarthritis or previous trauma. The specimens were also used in a previous for the development of computer algorithms for the measurement of the humeral anatomy in 3D (Vlachopoulos et al., 2016a). Segmentation of the humerus was performed fully automatic using a previously described method (Gass et al., 2014). Marching cubes algorithm (Lorensen and Cline, 1987) was applied to generate bilateral 3D triangular surface models. Thereby, we obtained a manifold mesh only of the outer cortical layer. Subsequently, the models were imported into the in-house developed planning-software CASPA (Balgrist CARD AG, Zurich, Switzerland).

In the following section, we will first describe the current state-of-the-art method in 3D deformity assessment of malunited distal humeral fractures (Murase et al., 2008; Murase et al., 2014; Takeyasu et al., 2013; Vlachopoulos et al., 2016b) before providing a detailed description of the proposed novel approach. To analyse the robustness of the proposed approach against bilateral differences of an individual, we validated the method based on bilateral humeral models without signs of a previous trauma or osteoarthritis.

3.2.1 State-of-the-art deformity assessment:

The template-based approach is the most commonly used method for the 3D assessment of posttraumatic deformities (Murase et al., 2008; Murase et al., 2014; Takeyasu et al., 2013; Vlachopoulos et al., 2016b). In this approach, the mirrored contralateral humeral model serves as a reconstruction template (Figure 3.1, A). To quantify a deformity of the distal humerus, two segments have to be selected proximal and distal to the pathological area (Figure 3.1, B). The proximal segment is selected in such a manner, that it contains the humeral head, the greater tuberosity, and part of the shaft (Murase et al., 2014; Takeyasu et al., 2013). The distal (elbow) segment contains the medial and lateral epicondyles and the distal articular surface (Murase et al., 2014; Takeyasu et al., 2013). A surface registration method, such as the Iterative Closest Point algorithm (ICP) (Audette et al., 2000; Chen and Medioni, 1991) is applied to superimpose the bone models. The pathological bone model is superimposed with the reconstruction template based on the registration of the proximal segment (Figure 3.1, C). Finally, the elbow segment is superimposed to the reconstruction template (Figure 3.1, D). The relative 3D rotation and 3D translation of the elbow segment between the proximal registration (Figure 3.1, C) and the distal registration (Figure 3.1, D) quantifies the malalignment. The rotational difference is expressed in axis-angle representation (Figure 3.1, E), i.e. a rotation by a 3D angle about a calculated axis.

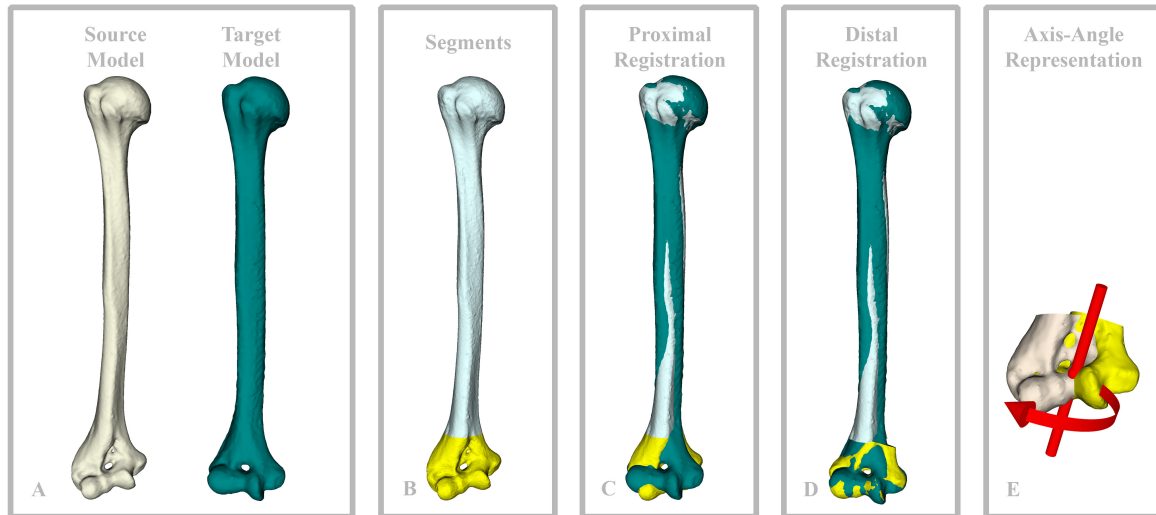


Figure 3.1: Template-based assessment of the deformity (A) The mirrored left humerus (target model) serves as a reconstruction template for the template-based assessment of the deformity of the right humerus (source model). (B) One segment proximal to the assumed deformity (i.e., the blue segment of 85% length of the right humerus), and one segment distal to the assumed deformity (i.e., the yellow segment of 15% length of the humerus), are selected to quantify the deformity. The proximal segment (C) and the elbow segment (D) are registered subsequently onto the mirrored left humerus with the ICP algorithm. (E) The rotational difference of the elbow segment between the proximal registration and the distal registration is expressed in axis-angle representation, i.e. a rotation by a 3D angle of 35° (red arrow) about a calculated axis.

3.2.2 Effect of the segment selection on the assessment of the deformity:

The principle of the ICP algorithm is to superimpose a source model with a target model in a way such that the difference between the surfaces of the models is zero (equal model surfaces) or minimal (different model surfaces) within a user-selected segment of interest.

In Figure 3.2, we demonstrate the effect on the deformity assessment when bilateral differences in humeral torsion are present without an additional pathological condition. We use a simplified representation of the humerus for demonstration purposes. The humeral models consist of three prism connected by a cylinder. A cuboid represents the humeral head, a trapezoidal-shaped prism represents the humeral shaft, and a cube symbolizes the distal joint with the epicondyles. The difference in humeral torsion of both models is illustrated by the different rotation of the prisms around the cylinder (Figure 3.2, A). In the example of Figure 3.2, the currently established selection strategy (Murase et al., 2014; Takeyasu et al., 2013) is denoted by the proximal segment (Figure 3.2, B; top). A different approach would be to select only parts of the shaft (Figure 3.2, B; bottom). The result of the ICP registration for each segment is given in Figure 3.2, C. In the first case (Figure 3.2, C; top), ICP minimizes the average distances between the surfaces of the proximal segments between source and target model. This yields a precise registration of the proximal humerus. However, the difference in torsion of the elbow segment is greater compared to the registration of the shaft segment (Figure 3.2, C; bottom). We defined the 3D rotational difference of the elbow segment of both humeri, as the humerus contralateral registration (HCR) error. The HCR corresponds to the

3D angle in axis-angle representation, when bilateral models without a pathological condition are analysed. The influence of the segment selection on the HCR error is illustrated in Figure 3.2, D.

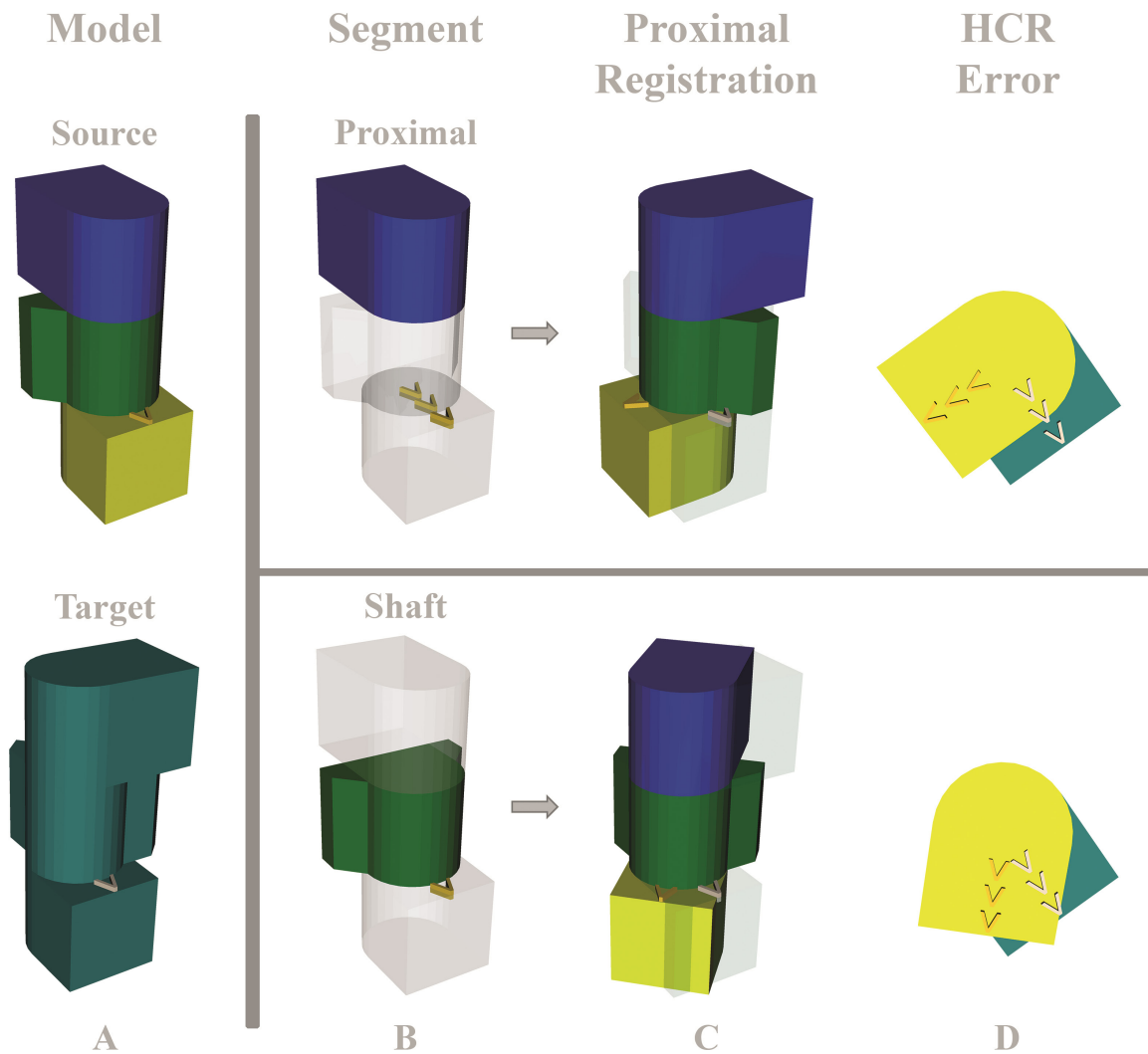


Figure 3.2: Effect of the segment selection on the registration result A) The source model represents a humeral model with a higher torsion between proximal and distal part compared to the target model. A cuboid represents the humeral head, a trapezoidal-shaped prism represents the humeral shaft, and a cube symbolizes the distal joint with the epicondyles. B) The selected segments used for the proximal registration are illustrated. In the top row, the segment contains the humeral head, in the bottom row the segment contains the shaft. C) The registration minimizes the difference between the surfaces of the selected segments. D) In addition, the segment selection influences the HCR error.

3.2.3 Novel automatic segment selection approach:

The previous example demonstrated the problem of the state-of-the-art method in case of bilateral differences in humeral torsion. To evaluate different segments that may result in a smaller HCR error, we performed ICP registration experiments, for each subject by processing the data in an automatic and standardized fashion. We defined the right humerus of a subject as the source model. The left contralateral humerus was mirrored and served as

the target model. To demonstrate that the evaluation is not biased by the selection of the source and target models, we performed additionally all experiments with the left humerus as the source model and the mirrored right humerus as the target models and analyzed the consistency of the method using the interclass correlation coefficient (ICC).

The registration experiments were performed in a similar way as described above for the state-of-the-art deformity assessment. We defined several segments for the proximal registration and one elbow segment for the distal registration to quantify the HCR error. The humeral segments were created by automatically dividing each source model based on the humeral length (Figure 3.3, A). The proximal segments that included at least the humeral head and the greater tuberosity, similar as the state-of-the-art method (Murase et al., 2014; Takeyasu et al., 2013), were assigned to the proximal group (Figure 3.3, B). The remaining proximal segments were assigned to the shaft group (Figure 3.3, C).

Thereby, we divided each source model for the proximal registration (Figure 3.1, C) into the following eight segments:

- Proximal group with segments:
prox-15, prox-25, prox-50 and prox-75, prox-85
- Shaft group with segments:
shaft-50-75, shaft-50-85, shaft-75-85

The proximal registration reveals the deviation of the elbow segment of both humeri as illustrated in Figure 3.2. We determined the HCR error by registration of the elbow segment (distal registration, Figure 3.1, D). Therefore, we created one segment of the distal humerus including the medial and lateral epicondyles and the distal articular surface (Murase et al., 2014; Takeyasu et al., 2013) with 15% of the total length of the humerus (elbow segment, Figure 3.3, D). The HCR error of two equally oriented elbow segments is zero. Note that throughout the manuscript we use consistently the same colour code to represent the defined segments.

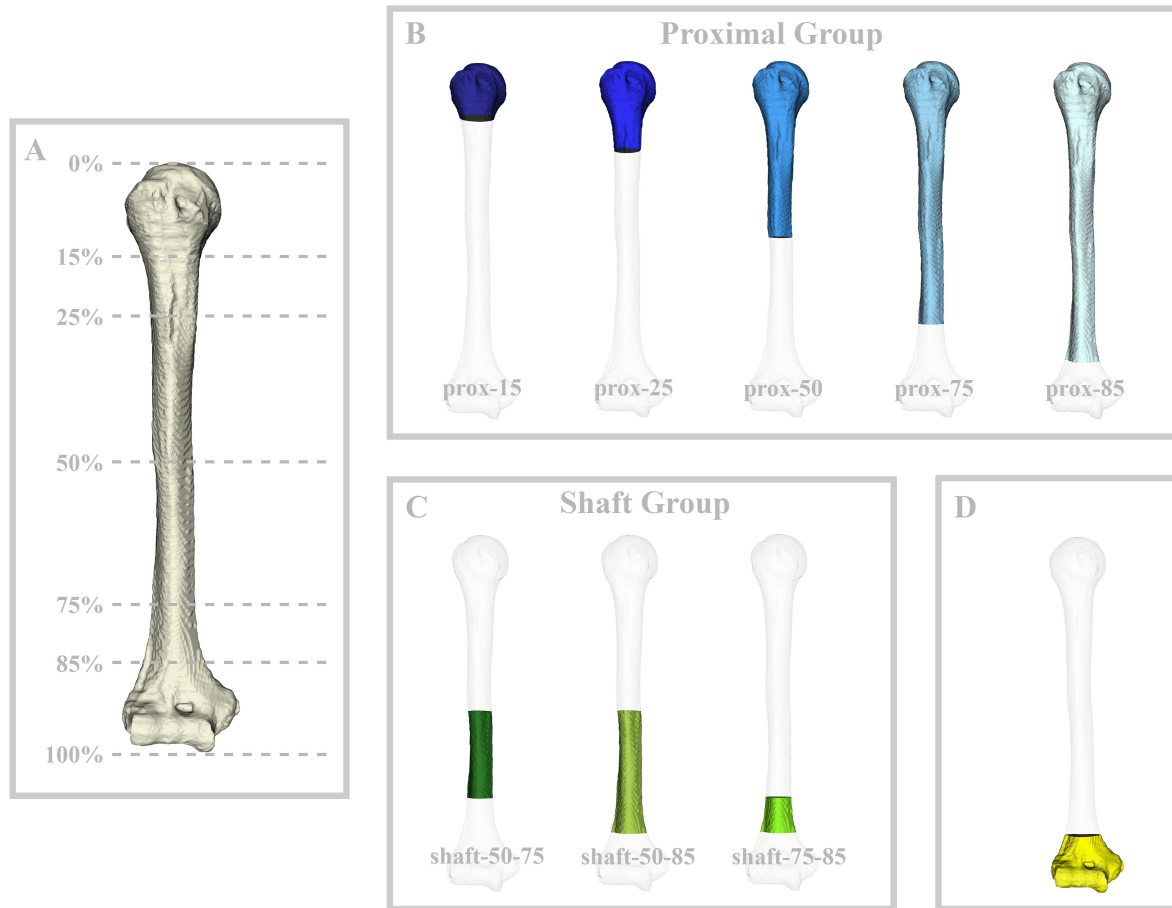


Figure 3.3: Generation of the segments defined for the registration A) Illustrates the right humeral models and the humeral length expressed as a percentage from the proximal end (0%) to the distal end (100%). The defined segments are depicted for the proximal registration - proximal group (B) and shaft group (C) – and for the evaluation of the HCR error by the distal registration (D).

3.2.4 Statistical Analysis

Mauchly's sphericity test revealed a violation of the assumption of sphericity. Therefore, we applied the non-parametric Friedman Rank Sum test with the segment as a group factor and the individuals as a block factor for the analysis of the HCR error as a function of the selected segment. Post-hoc analysis was performed with the Wilcoxon signed-rank test and Bonferroni adjustment. The significance level was set at $P < 0.05$. The consistency of the method was assessed with the interclass correlation coefficient (ICC, 2-way random-effects model). For graphical visualization, Tukey boxplots were used with the end of the whiskers indicating the 1.5 interquartile range (IQR) of the lower and upper quartile. All reported HCR errors were calculated from the average of the absolute values, when the right or the left humeri were selected as a source model.

3.3 Results

The Friedman rank sum test revealed a significant effect of the selected segments on the HCR error (chi-squared 153, $P < 0.0001$). In Figure 3.4, the HCR errors are depicted for each segment of both groups. Twenty-eight paired Wilcoxon signed-rank tests were performed for post hoc analysis. The Bonferroni-adjusted p-value (p_{adj}) was $0.05/28 \approx 0.002$ for each individual comparison, in order to retain the prescribed family-wise error rate of 0.05. The ICC was above 0.93 for all selected segments (Table 3.1). We performed for each humeral pair eight left-to-right and eight right-to-left registrations, which yields 500 registration experiments in total.

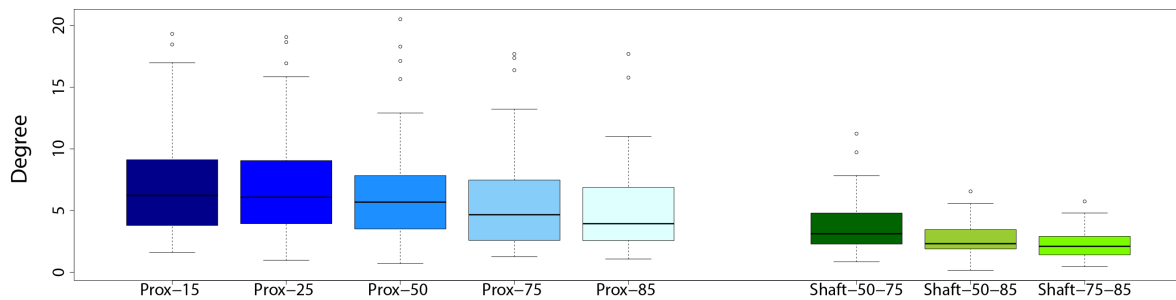


Figure 3.4: HCR Errors. Boxplots of the HCR errors as a function of the selected segments

Segment	HCR error (°)						
	Mean	SD	Median	Range		ICC (95% CI)	
Prox-15	7.4	4.5	6.2	1.6	- 19.3	1.00	(0.99-1.00)
Prox-25	7.2	4.5	6.1	1.0	- 19.1	1.00	(0.99-1.00)
Prox-50	6.7	4.5	5.7	0.7	- 20.5	1.00	(0.99-1.00)
Prox-75	5.7	4.3	4.7	1.3	- 17.7	1.00	(0.99-1.00)
Prox-85	5.2	3.7	4.0	1.1	- 17.7	0.98	(0.97-1.00)
Shaft-50-75	3.8	2.3	3.1	0.9	- 11.2	0.94	(0.89-0.96)
Shaft-50-85	2.6	1.3	2.3	0.2	- 6.6	0.94	(0.90-0.97)
Shaft-75-85	2.3	1.1	2.1	0.5	- 5.8	0.93	(0.88-0.96)

Table 3.1: Mean, standard deviation, median and the range of the absolute HCR errors for all segments and the corresponding ICC with the 95% CI.

The HCR errors of the *shaft-50-85* and *shaft-75-85* segments were significant smaller than the HCR errors of the five segments of the proximal group ($P < 0.001$) and of the *shaft-50-75* segment ($P < 0.001$). The difference of the HCR errors of the *shaft-75-85* and *shaft-50-85* segments was not significant ($P = 0.01$). The HCR error of the *shaft-50-75* segment was significantly smaller than the HCR errors of the *prox-15*, *prox-25* and *prox-50* segments ($P < 0.001$). The difference of the HCR errors of the *shaft-50-75* segment and the *prox-75* segment ($P = 0.004$) or *prox-85* segment ($P = 0.02$) was not significant. The difference of the HCR errors of the *prox-15* and *prox-25* segment was not significant ($P = 0.01$). All

further differences of the HCR errors of segments within the proximal group were significant ($P < 0.001$).

The HCR error of the *prox-15* segment was greater than 10° in 20% of the cases and between 5° and 10° in further 50% of the cases. The HCR error of the *shaft-75-85* segment was between 5° and 10° in 2% of the cases, and never $>10^\circ$. Table 3.1 shows the mean, standard deviation and the range of the HCR errors for all segments.

3.4 Discussion

The benefits of three-dimensional computer-assisted corrective osteotomies of distal humeral deformities have already been emphasized (Omori et al., 2015; Vlachopoulos et al., 2016b). In this context, the 3D assessment of the deformity of the distal humerus (Murase et al., 2008; Murase et al., 2014; Omori et al., 2015; Takeyasu et al., 2013; Vlachopoulos et al., 2016b) gained importance as well as the development of a new surgical technique to perform the surgery as planned on the computer (Murase et al., 2008; Murase et al., 2014; Omori et al., 2015; Takeyasu et al., 2013; Vlachopoulos et al., 2016b). The reliability of the contralateral humerus as a reconstruction template is of fundamental importance to restore the pretraumatic anatomy as precise as possible, as it influences all subsequent steps of the procedure (Vlachopoulos et al., 2016b). In other words, a high HCR error indeed results in an inaccurate restoration of the pretraumatic anatomy.

However, to the best of our knowledge, the influence of intra-individual bilateral differences of the humerus anatomy on the 3D assessment of the deformity has not been analysed so far. In the present study, we investigated the effect of the selected segments on the 3D registration-based deformity assessment. For the registration of the pathological bone with the reconstruction template (i.e., proximal registration), previous studies have proposed to use the entire non-pathological proximal segment consisting of the greater tuberosity, the humeral head, and the shaft (Murase et al., 2014; Takeyasu et al., 2013). This approach may introduce an error in the assessment of a distal deformity due to intra-individual bilateral differences, especially in humeral torsion. In other words, the distance of the proximal humerus used for the proximal registration away from the distal area of interest, results in error propagation. Takeyasu et al. (2011) reported that most of the patients with a posttraumatic cubitus varus deformity after a supracondylar fracture have an accompanying extension and/or rotational deformity of 10° or greater and highlighted the 3D nature of these deformities. However, the results of the deformity assessment of the study of Takeyasu et al. (2011) should be interpreted with caution, since segments similar to those of the proximal group of this study were used for the registration with the reconstruction template (Takeyasu et al., 2013).

Omori et al. demonstrated in a cadaveric study (Omori et al., 2014) and in vivo (Omori et al., 2015) that the 3D preoperative plan can be realized very accurately during surgery with patient specific guides. The authors reported a mean error of $0.6^\circ \pm 0.7^\circ$ in varus-valgus direction and $0.8^\circ \pm 1.3^\circ$ in flexion-extension direction, while the error in axial rotation was higher (mean $2.9^\circ \pm 2.8^\circ$) with a maximum error of 9.4° compared to the sagittal or coronal plane. The postoperative evaluation was performed with the currently probably most precise technique. Although the results reported by Omori et al. are promising, their postoperative

evaluation of the accuracy represents only the difference between the performed correction and the preoperative plan, and does not necessarily correspond to the deviation from the pretraumatic condition.

Another approach to achieve a correspondence between the preoperative created surface models based on CT scans and the intraoperative situation was analysed by McDonald et al. (2008). By simulating different settings of articular bone loss, McDonald et al. (2008) demonstrated, that a close alignment of a preoperative image with intraoperative surface data can be achieved (rotational error $0.4 \pm 0.1^\circ$) if the distal humeral shaft can be scanned by the surgeon with a laser scanner. It should be noted that this error captures only the transfer of the preoperative plan. Contrary, the error analysed by Omori et al. (2015) represents the accumulated error between performed correction and the preoperative plan.

In the present manuscript, we described our method for the accurate approximation of the pre-traumatic anatomy of the distal humeral in the context of corrective osteotomies of the distal humerus. The main reason was that the 3D preoperative planning and the use of patient specific guides are well accepted for this pathology and we directly could demonstrate the benefit of our approach compared to the state-of-the-art. The method is probably also applicable for further procedures that rely on an accurate approximation of the pre-traumatic anatomy of the distal humerus (i.e., navigation of total elbow arthroplasty).

McDonald et al. (2009) also demonstrated that anthropometric measurements of the distal humerus (i.e., flexion-extension axis of the elbow and offset) are similar for paired specimens. Registration of the elbow segment, as defined in the present study, to the healthy contralateral anatomy yields a side-to-side difference of approximately 1.0° and 0.5 mm. Furthermore, the registration of the elbow segment is possible even if severe bone loss of the articular surface is present without affecting the flexion-extension axis of the elbow and the offset. These findings (McDonald et al., 2008; McDonald et al., 2009) strengthen the results of the present study. However, the size of the distal humeral shaft segment used for the registration was not defined (McDonald et al., 2008; McDonald et al., 2009). In addition, in these studies (McDonald et al., 2008; McDonald et al., 2009) the selected segments for the registration seem to capture only the distal segments defined in the present study. Thereby, only the accuracy of the distal registration of the deformity assessment was analysed.

Compared to the state-of-the-art method (Murase et al., 2008; Murase et al., 2014; Takeyasu et al., 2013), we have demonstrated that the HCR error is less sensitive to intra-individual differences if only segments of the distal humeral shaft (*shaft-50-85* or *shaft-75-85*) are selected. The HCR errors resulting from the registration with these segments were significantly smaller than the registration using any segment of the proximal group. The present study demonstrates the influence of bone shape and distance from the area of interest on registration algorithms. Figure 3.5 illustrates the cross-sectional shape of the *shaft-50-75* and *shaft-75-85* segments. The definition of unique features as close as possible to the area of interest is relevant for registration-based approaches for further pathologies to compensate for bilateral differences (i.e., differences in torsion). However, the magnitude of error and the consistency of the registration have to be validated separately for further anatomies, since both depend on the anatomy.

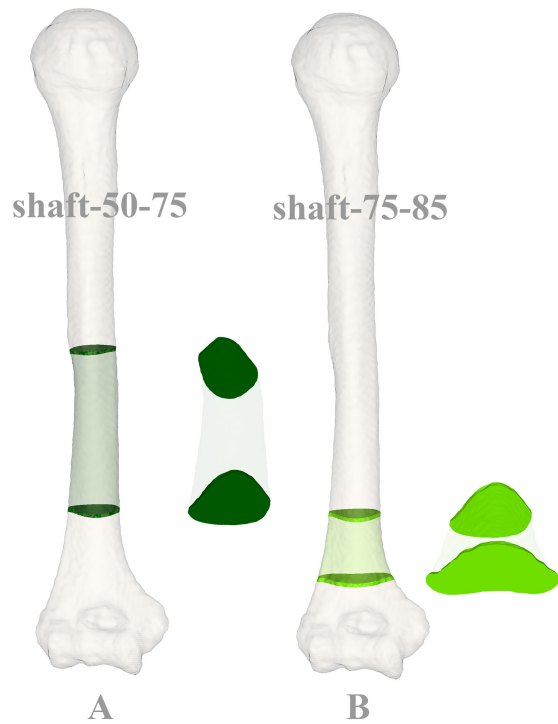


Figure 3.5: Characteristic features of the distal humerus. Illustrates the anterior view and the cross-sectional view of the A) *shaft-50-75* segment and the B) *shaft-75-85* segment

One limitation of the present study is that, similar as the state-of-the-art method, the preoperative assessment of the deformity can be applied only if the contralateral bone is healthy. It is important, that the segment used for the proximal registration is proximal to the pathological area. However, since the proximal part of the humerus is not necessary for the registration, the method can be used even in presence of a pathological condition of the proximal humeral anatomy. Nevertheless, this novel approach yields a deformity assessment being more robust against bilateral differences and more precise than the state-of-the-art method for the reconstruction of the pre-traumatic angular orientation of the elbow. Furthermore, the restoration of the humeral length might be influenced in our approach by differences in the diameter of the humeral shaft between both sides of an individual. Therefore, we propose to reconstruct the length according to the length of the contralateral humeral model, since the contralateral anatomy is assumed to be reliable template for the humeral length. In conclusion, our approach strengthens the benefits and the application of the current technique for corrective osteotomies of the distal humerus.

4

Approximation of the Humeral Head Retrotorsion

In the present paper, we analysed the approximation of the (pretraumatic) orientation of the humeral head with a registration-based approach. The influence of a proper segment selection to compensate bilateral differences in humeral anatomy was proven before. In the present paper, we proposed and validated a segment selection strategy for the approximation of the proximal humeral anatomy. As a reference we used the intra-individual differences in retrotorsion with the previously developed landmark-based method. We concluded that the registration of the proximal humeral shaft is a reliable method to approximate the proximal humeral anatomy that compensates bilateral differences better than the landmark-based reference method.

4.1 Introduction

Three-dimensional (3D) computer-assisted planning, based on reconstructed bone models obtained from computed tomography (CT), has been recently proposed for corrective osteotomies of complex malunited humeral fractures (Murase et al., 2008; Omori et al., 2015; Vlachopoulos et al., 2016b). Several studies have demonstrated that computer-assisted planning and patient-specific guides are an accurate method to transfer the preoperative plan to the surgery (Murase et al., 2008; Omori et al., 2015; Vlachopoulos et al., 2015). However, the most fundamental step influencing all subsequent steps of the procedure is the preoperative assessment of the deformity (Vlachopoulos et al., 2016b).

For the assessment of the deformity, in the state-of-the-art template-based approach, the bone fragments are reduced to a 3D reconstruction template. The contralateral healthy humerus is, thereby, a commonly proposed 3D reconstruction template (Bicknell et al., 2007; Murase et al., 2008; Omori et al., 2015; Vlachopoulos et al., 2016a; Vlachopoulos et al., 2016b). Although recent studies revealed that several geometric quantities of the proximal humerus (i.e., humeral length, humeral head size, humeral head height) are similar between both sides of an individual (Bicknell et al., 2007; Vlachopoulos et al., 2016a), the humeral head retrotorsion especially differs between both sides (Bicknell et al., 2007; Vlachopoulos et al., 2016a).

Furthermore, a recent study (Vlachopoulos et al., 2017b) demonstrated the influence of the segment selection on registration algorithms used for the deformity assessment of the distal

humerus. The authors pointed out, that the selection of segments with unique features as close as possible to the area of interest is relevant for registration-based approaches to compensate for bilateral differences (i.e., in humeral torsion). However, the magnitude of error and the optimal segment selection strategy have to be validated separately for further anatomies, since both depend on the anatomy. Vlachopoulos et al. (2016b) recently presented a detailed overview of the technique and relevant steps for corrective osteotomies of complex malunited fractures of the proximal humerus from the preoperative plan to the generation of patient-specific guides. However, the definition and validation of segment selection upon the accuracy for the reconstruction of the proximal humeral anatomy has not been described so far.

In addition, several studies have emphasized that the bicipital groove may serve as an intraoperative landmark for the restoration of the retrotorsion in cases of a proximal humeral fracture (Doyle and Burks, 1998; Hempfing et al., 2001; Kontakis et al., 2001) and a direct correlation between the bicipital groove rotation and the humeral head retrotorsion has been reported (Johnson et al., 2013).

The aim of this study was to analyse the influence of segment selection on the approximation of the humeral head retrotorsion with a registration-based approach. We hypothesize that segments of the proximal humeral shaft have characteristic invariant features, which allow an accurate approximation of the (pretraumatic) proximal humeral anatomy. The accuracy of the method was compared to the approach of measuring the bilateral differences in retrotorsion with a state-of-the-art landmark-based 3D method.

4.2 Material and Methods

The Swiss Institute for Computer-Assisted Surgery (SICAS) provided computed tomography (CT) full-body data of 50 specimens (including the entire humerus of both sides). The in-plane (xy-) resolution of the CT scans ranged between 0.9 x 0.9 mm and 1.27 x 1.27 mm. The slice thickness varied from 0.5 mm to 0.6 mm. The average age of the individuals was 52.1 years (SD, 20.0 years; range, 19 to 90 years). There were 32 male and 18 female cadavers. The average height was 172.4 cm (SD, 8.7 cm; range, 154 to 187 cm) and the average weight was 68.4 kg (SD, 16.9 kg; range, 37 to 108 kg). Exclusion criteria were the presence of osteoarthritis or previous trauma. Segmentation of the humerus was performed fully automatically using a previous described segmentation algorithm (Gass et al., 2014). Marching cubes algorithm (Lorensen and Cline, 1987) was applied to generate bilateral 3D triangular surface models. The specimens were also used in previous studies (Vlachopoulos et al., 2017b; Vlachopoulos et al., 2016a). To evaluate the approximation of the proximal humeral anatomy based on the contralateral humeral model, we used, in the current study, bilateral humeral models without a pathological condition. Thereafter, the models were imported into the in-house developed planning-software CASPA (Balgrist CARD AG, Zurich, Switzerland).

In the following, we will first describe the landmark-based retrotorsion (LBR) error. The LBR error was used as a reference to measure the intra-individual differences in retrotorsion with a previously developed landmark-based measurement method (Vlachopoulos et al., 2016a).

Subsequently, we will describe the registration-based approach for the approximation of the proximal humeral anatomy.

4.2.1 Landmark-based 3D method (LBR error):

We assumed for the reference method, that the humeral head retrotorsion would be reconstructed based on the retrotorsion measured with the landmark-based methods from the humeral model of the contralateral side. We measured the retrotorsion between the articular margin plane (AMP) and the elbow tangent plane (ETP) as previously described (Figure 4.1) (Vlachopoulos et al., 2016a). The articular margin plane (AMP) was estimated by a plane fitted to a set of reference points along the cartilage/metaphyseal boundary (Bicknell et al., 2007; Johnson et al., 2013; Vlachopoulos et al., 2016a). The ETP is a plane tangent to the capitulum and trochlea humeri and parallel to the length axis of the humerus. All measurements were performed in respect to the ISB coordinate system (Wu et al., 2005). The rotation around the x-axis (Figure 4.1, red axis) corresponds, thereby, to a rotation in the coronal plane (varus/valgus). The rotation around the y-axis (Figure 4.1, green axis) corresponds to a rotation in the axial plane (internal-/external rotation) and the rotation around the z-axis (Figure 4.1, blue axis) corresponds to a rotation in the sagittal plane (flexion/extension). For the measurement of the landmark-based retrotorsion angle, the normal vector of the ETP and the normal vector of the AMP were projected onto the axial plane (xy-plane) of the humeral coordinate system. The angle between those two projected vectors minus 90° then described the retrotorsion. Consequently, the LBR error is the absolute difference between the retrotorsion of both sides of a specimen. Note, that the LBR error represents the difference between the projected retrotorsion angles, as defined clinically (Vlachopoulos et al., 2016a).

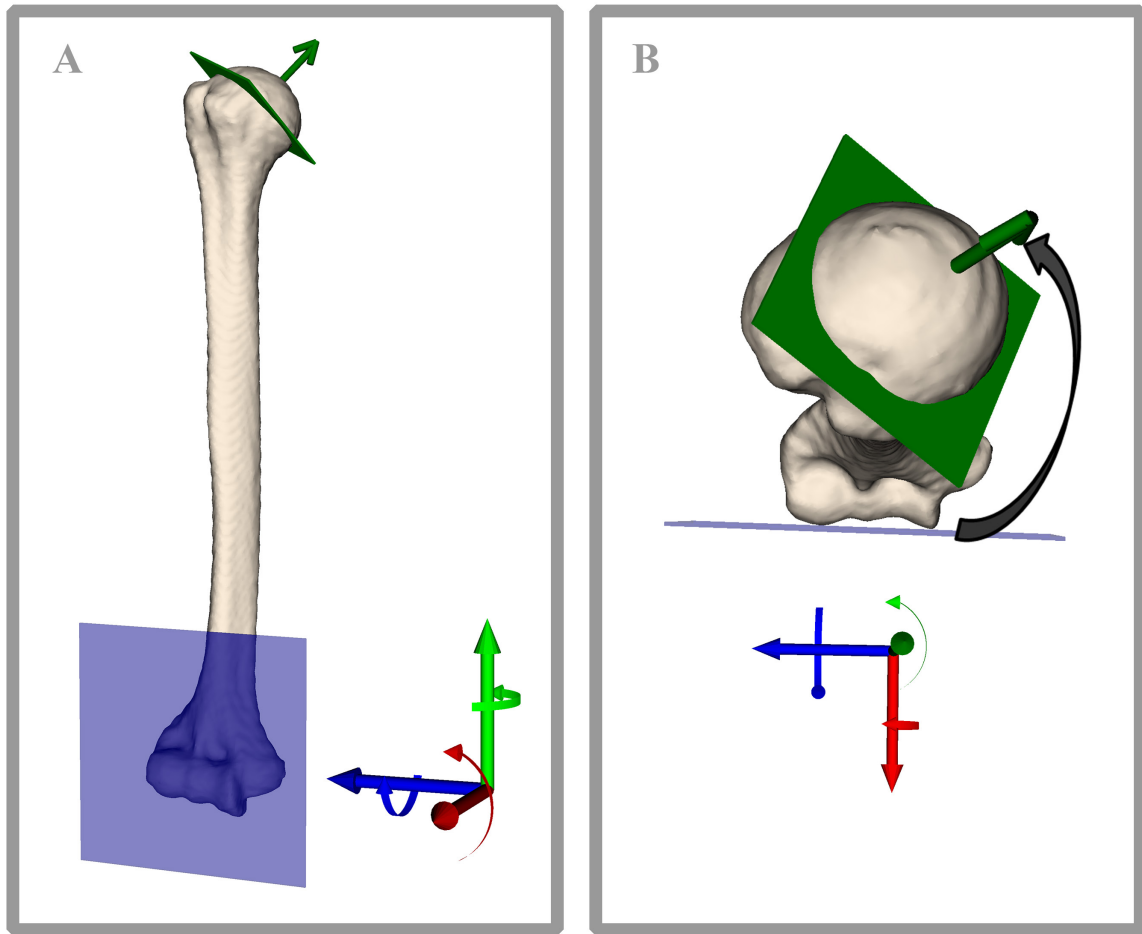


Figure 4.1: Landmark-based retrotorsion error: Illustrates an anteroposterior view (A) and an axial view (B) of a humeral model with the elbow tangent plane (ETP, blue plane), the articular margin plane (AMP, green plane with arrow indicating plane normal) and the humeral coordinate system. For the calculation of the landmark-based retrotorsion, the normal vector of the ETP and the normal vector of the AMP are projected onto the xz-plane (red and blue axes) of the humerus coordinate system. The angle between those two projected vectors minus 90° yields the landmark-based retrotorsion angle as defined clinically. The LBR error is the absolute difference between the landmark-based retrotorsion angle of both sides of an individual.

4.2.2 Registration-based approach:

In the state-of-the-art method for the assessment of a posttraumatic deformity in 3D, the bone fragments of the pathological model (Figure 4.2, A, source model) are reduced to a 3D reconstruction template (Figure 4.2, A, target model) (Kataoka et al., 2013; Murase et al., 2008; Schweizer et al., 2013; Vlachopoulos et al., 2016b). Two segments are selected on the source model, one proximal and one distal to the deformity (Figure 4.2, B). A surface registration method, the Iterative Closest Point (ICP) algorithm (Besl and McKay, 1992; Chen and Medioni, 1991), is applied to first superimpose the source model based on the distal segment onto the target model (Kataoka et al., 2013; Murase et al., 2008; Schweizer et al., 2013; Vlachopoulos et al., 2016b). The deviation of the humeral head between the source and the target model quantifies the deformity (Figure 4.2, C).

For the present study we defined the 3D rotational difference of the proximal humeral anatomy of the approximated and the original bone models as the proximal humerus contralateral registration (p-HCR) error. To measure the p-HCR error, we applied the ICP algorithm to superimpose the proximal segment (head segment) with the reconstruction template. The relative 3D rotation and translation of the head segment between the distal (Figure 4.2, C) and the proximal registration quantifies the malalignment (Figure 4.2, D). The rotational difference is expressed in axis-angle representation (3D-p-HCR error) and, additionally, as three constitutive rotations (i.e., Euler rotations) around the same humeral coordinates system as before.

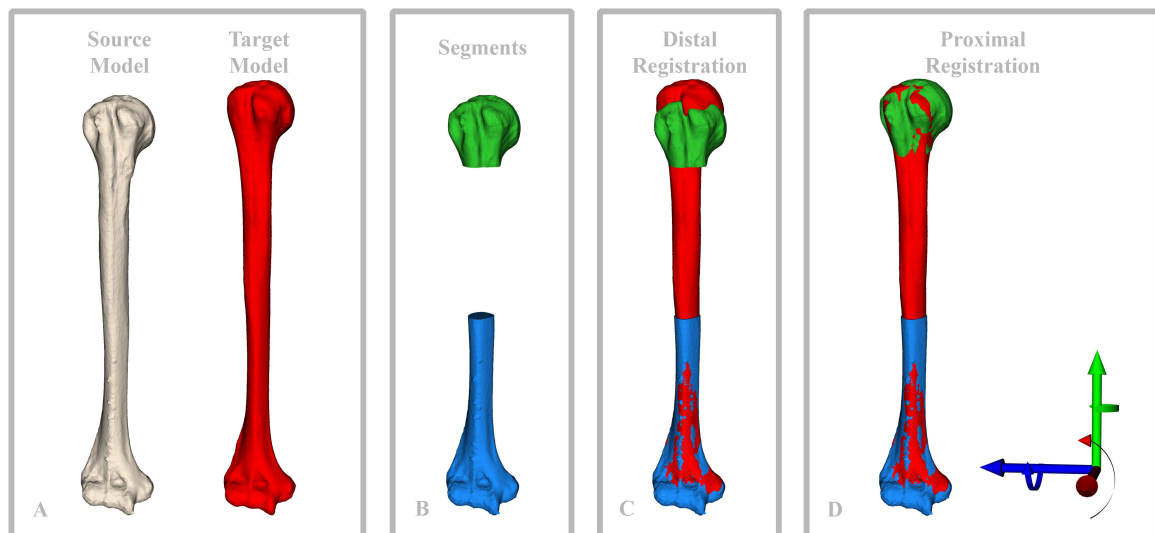


Figure 4.2: Assessment of a posttraumatic deformity of the proximal humerus. A) 3D bone models are generated of the pathological humerus (source model) with a posttraumatic deformity after a proximal humeral fracture and the mirrored contralateral humeral model (target model), which serves as a reconstruction template. B) Two segments are selected distal (blue segment) and proximal (green model) to the pathological area. C) After registration of the source model with the ICP algorithm based on the selected distal segment the posttraumatic deformity corresponds to the deviation of the humeral head between the source (green segment) and the target model. D) To quantify the deformity the ICP algorithm is applied to superimpose the humeral head segment of the source model onto the target model.

4.2.2.1 Effect of the segment selection on the approximation of the humeral head retrotorsion:

The principle of the ICP algorithm is to superimpose a source model with a target model (Figure 4.3) in a way such that the difference between the surfaces of the models is zero (equal model surfaces) or minimal (different model surfaces) within a user-selected segment of interest (Vlachopoulos et al., 2017b). In Figure 4.3, we illustrate the effect of the segment selections on the assessment of the humeral head retrotorsion when bilateral differences in humeral torsion are present without any pathological condition. For demonstration purposes, we use a simplified representation of the humerus. An elliptical cylinder represents the humeral head, while cubes represent the distal part of the humerus and the proximal shaft. The difference in retrotorsion between the segments is outlined by the difference in the orientation of the arrows of each segment. The registration algorithm superimposes the models in a way that the surfaces of both models are the most similar for the particular selected segment (Figure 4.3, A, B and C). The influence of the segment selection for the ICP on the relative orientation of the humeral head of both sides can be observed in the last column of Figure 4.3.

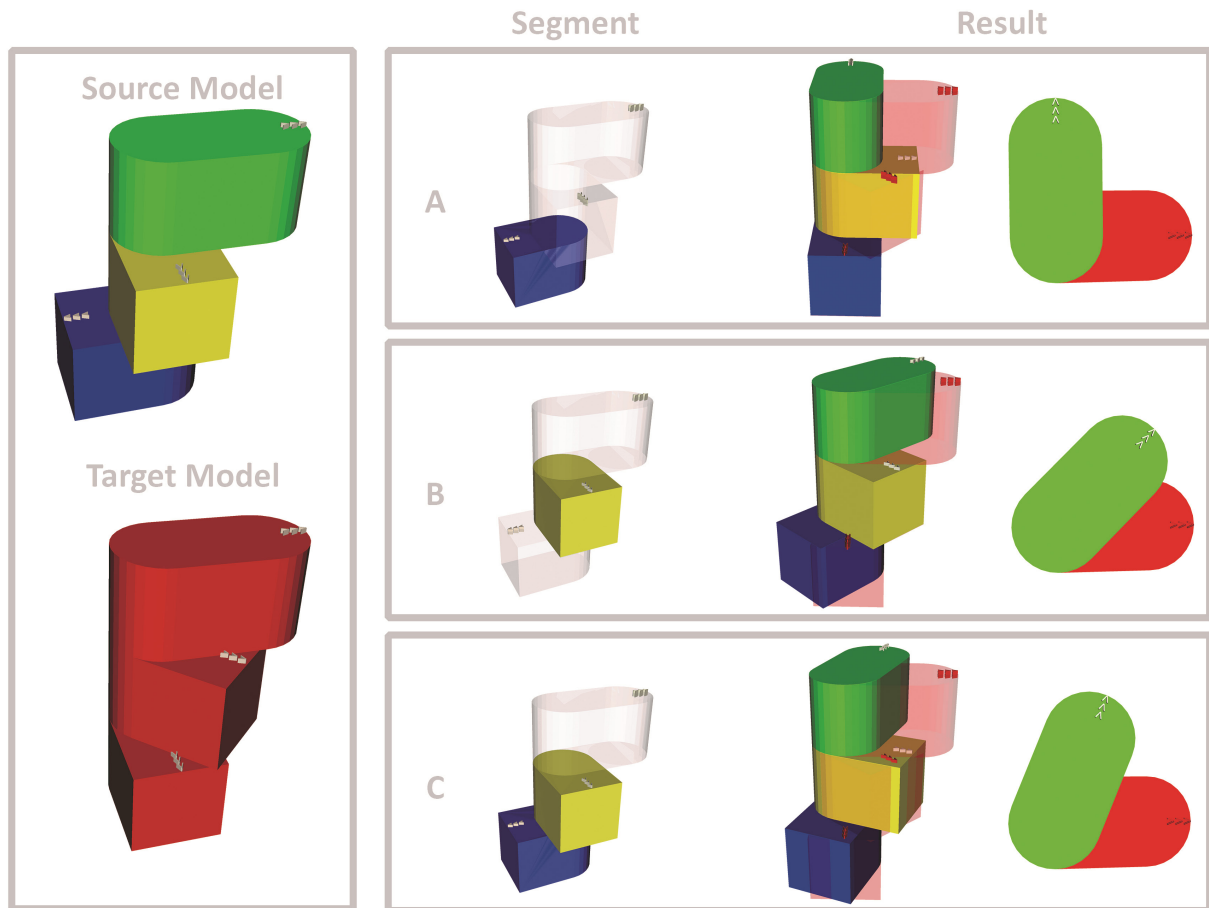


Figure 4.3: Effect of the segment selection on the approximation of the humeral head retrotorsion: A simplified representation of the humerus is used as a source and a target model. An elliptical cylinder represents the humeral head, while cubes represent the distal part of the humerus and the proximal shaft. In A) the selected segment for the ICP registration contains the distal part of the humerus (capitulum, trochlea and epicondyles). In B) the selected segments contains the groove region of the humerus, while in C) the selected segment contains both previously defined segments. The results of the registration based on the selected segments are depicted on the right. On the top view, the difference between the arrows represents the differences in humeral head retrotorsion between the two models. This difference represents the p-HCR error if bilateral humeral models without a pathological condition are used for the assessment.

We performed several registration experiments to determine the optimal segment that allows superimposing the humeri such that the (pretraumatic) humeral head retrotorsion is best approximated. We created humeral segments based on the humeral length (Figure 4.4, A). These segments were separated into two groups depending on which anatomical region was included. In the groove group (Figure 4.4, C), we allocated three segments just below the humeral head (Figure 4.4, B) capturing the proximal humeral shaft and the bicipital groove without the humeral head. In the distal group, we allocated all segments that included at least the distal part of the humerus (capitulum, trochlea and epicondyles; Figure 4.4, D).

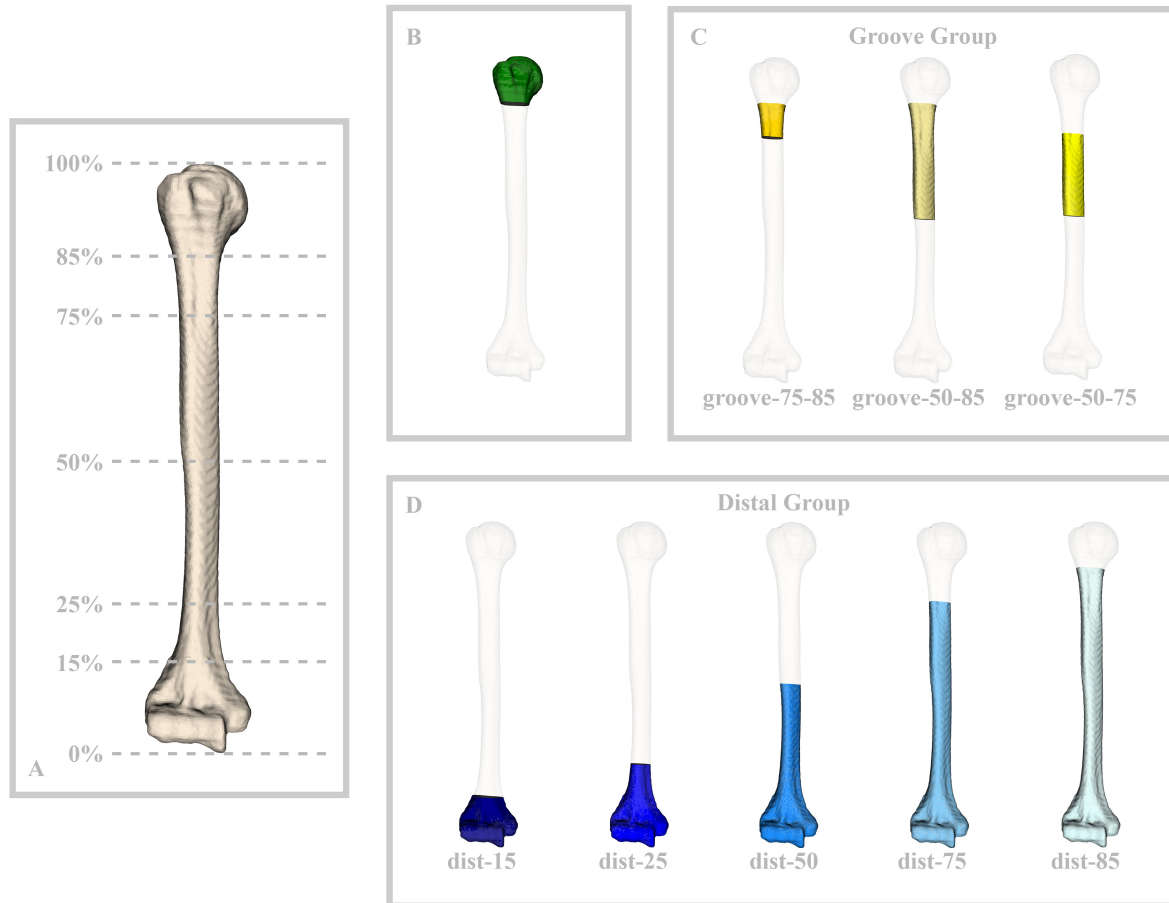


Figure 4.4: Visualization of the humeral segments defined for the registration. A) Illustrates the right humeral models and the humeral length expressed as a percentage from the distal end (0%) to the proximal end (100%). B) The defined segments for the evaluation of the p-HCR error (green segment) and the segments used for the distal registration - groove group (C) and distal group (D) are depicted. Note that throughout the manuscript we use consistently the same colour code to represent the defined segments.

We defined the right humerus of a specimen as the source model. The left contralateral humerus was mirrored about the sagittal plane of the coordinate system and served as the target model. To analyse the consistency of the method we performed all experiments also with the left humerus as the source model and the right humeri as the target model. The consistency was analysed using the interclass correlation coefficient (ICC).

We divided each source model for the registration into the following eight segments:

- *Groove Group* with the segments:
groove-75-85, *groove-50-85* and *groove-50-75* (Figure 4.4, C)
- *Distal Group* with the segments:
dist-15, *dist-25*, *dist-50*, *dist-75* and *dist-85* (Figure 4.4, D)

These segments were used to perform distal registrations as described before (Figure 4.2, C). Hence, we obtained eight different results of superimposed humeral pairs. To quantify the

3D-p-HCR errors, we created one head segment having 15% of length of the humerus, including the humeral head and the greater tuberosity (Figure 4.4, B). A 3D-p-HCR error of 0° indicates that the pretraumatic retrotorsion of a pathological humerus could be approximated perfectly with the registration-based approach using the selected segment. The rotational part around the y-axis can be interpreted as internal-/external rotation and is therefore, comparable with the LBR error. The rotational parts around the x- and z-axis are not captured by the LBR error.

4.2.3 Statistical Analysis

Mauchly's sphericity test revealed a violation of the assumption of sphericity. Therefore, we applied the non-parametric Friedman Rank Sum test with the segment as a group factor and the individuals as a block factor to analyse the effect of the selected segments on the 3D-p-HCR error. Post-hoc analysis was performed with the Wilcoxon signed-rank test and Bonferroni adjustment. The significance level was set at $P < 0.05$. The consistency of the method was assessed with the interclass correlation coefficient (ICC, 2-way random-effects model). For graphical visualization, Tukey boxplots were used with the end of the whiskers indicating the 1.5 interquartile range (IQR) of the lower and upper quartile. All reported p-HCR errors were calculated from the average of the absolute values, when the right or the left humeri were selected as a source model.

For the analysis of the rotational part of the p-HCR errors around the y-axis (internal-/external rotation) and the LBR error, we applied the non-parametric Friedman Rank Sum test. Post-hoc analysis was performed with the Wilcoxon signed-rank test and Bonferroni adjustment.

4.3 Results

The Friedman rank sum test revealed a significant effect of the selected segments on the 3D-p-HCR error (chi-squared 142.9, $P < 0.0001$). Twenty-eight paired Wilcoxon signed-rank tests were performed for post hoc analysis. The Bonferroni-adjusted p-value (p_{adj}) was $0.05/28 \approx 0.002$ for each individual comparison, in order to retain the prescribed family-wise error rate of 0.05. The mean value, standard deviation, median and range of the 3D-p-HCR errors and of the LBR errors are summarized in Table 4.1. The ICC was above 0.96 for all selected segments (Table 4.1). The p-values of the post-hoc analysis of the 3D-p-HCR error performed with the pairwise Wilcoxon signed-rank test are depicted in Table 4.2.

	3D-HCR error (°)					
	Mean	SD	Median	Range		ICC (95% CI)
Groove-75-85	2.8	1.5	2.4	0.6	- 7.4	0.96 (0.93-0.98)
Groove-50-85	3.6	2.4	2.9	0.7	- 10.8	0.97 (0.95-0.98)
Groove-50-75	4.4	2.8	3.3	0.9	- 12.8	0.98 (0.96-0.99)
Dist-85	6.5	4.7	5.5	0.8	- 23.5	1.00 (0.99-1.00)
Dist-75	6.5	4.6	5.5	0.8	- 20.4	0.99 (0.98-0.99)
Dist-50	6.9	4.6	6.3	0.8	- 19.3	1.00 (0.99-1.00)
Dist-25	7.1	4.5	6.3	1.3	- 18.7	1.00 (0.99-1.00)
Dist-15	7.4	4.5	6.2	1.6	- 19.3	1.00 (0.99-1.00)
LBR	6.4	5.9	5.2	0.5	- 24.0	

Table 4.1: Mean, SD, median and range of the 3D-p-HCR error of the proximal humerus for each selected segment and the corresponding ICC with the 95% CI and of the of the LBR error.

	Groove 50-85	Groove 50-75	Distal- 85	Distal- 75	Distal- 50	Distal- 25	Distal- 15
Groove 75-85	0.045	0.001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Groove 50-85		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Groove 50-75			<0.001	0.001	<0.001	<0.001	<0.001
Distal-85				0.10	<0.0001	<0.0001	<0.0001
Distal-75					<0.0001	<0.0001	<0.0001
Distal-50						0.002	<0.001
Distal-25							<0.001

Table 4.2: P-values of the post-hoc analysis of the 3D-p-HCR error performed with the Wilcoxon signed-rank test and Bonferroni adjustment ($p_{adj} = 0.05/28 \approx 0.002$).

For the analysis of the p-HCR error around the y-axis (internal-/external rotation) and the LBR error, the Friedman rank sum test revealed a significant effect of the selected segments (chi-squared 95.4, $P < 0.0001$). Thirty-six paired Wilcoxon signed-rank test were performed for post hoc analysis. The Bonferroni-adjusted p-value (p_{adj}) was $0.05/36 \approx 0.0014$ for each individual comparison, in order to retain the prescribed family-wise error rate of 0.05. The p-values of the post-hoc analysis with the pairwise Wilcoxon signed-rank test are depicted in Table 4.3. The p-HCR errors around the y-axis (internal-/external rotation) using the *groove-75-85* or the *groove-50-85* segment for the registration were significant smaller than using the *groove-50-75* segment or any segment of the distal group and significant smaller than the LBR error. The difference between the p-HCR errors around the y-axis (internal-/external rotation) using the *groove-50-75* segment or any segment of the distal group were not significant different from the LBR error. The share of individuals with a p-HCR error around the y-axis (internal-/external rotation) or a LBR error $< 5^\circ$, between 5° and 10° , and $> 10^\circ$ is summarized in Table 4.4.

	Groove 50-85	Groove 50-75	Distal- 85	Distal- 75	Distal- 50	Distal- 25	Distal- 15	LBR
Groove 75-85	0.045	0.001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Groove 50-85		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Groove 50-75			<0.001	0.001	<0.001	<0.001	<0.001	0.005
Distal-85				0.10	<0.0001	0.0001	<0.0001	0.45
Distal-75					<0.0001	<0.0001	<0.0001	0.40
Distal-50						0.002	<0.0001	0.77
Distal-25							<0.001	0.94
Distal-15								0.98

Table 4.3: P-values of the post-hoc analysis of the p-HCR error around the y-axis (internal-/external rotation) and the LBR error performed with the Wilcoxon signed-rank test and Bonferroni adjustment ($p_{adj} = 0.05/36 \approx 0.0014$).

	Individuals (%) with an error		
	<5°	5°-10°	>10°
Groove-75-85	94	6	0
Groove-50-85	84	14	2
Groove-50-75	72	22	6
Dist-85	48	34	18
Dist-75	52	30	18
Dist-50	44	38	18
Dist-25	42	40	18
Dist-15	42	40	18
LBR	48	26	26

Table 4.4: P-HCR error around the y-axis (internal-/external rotation) presented for all segments and LBR error < 5°, between 5° and 10° and >10°.

The p-HCR errors in axis-angle (3D) and Euler-representation around the three axes are summarized in Figure 4.5.

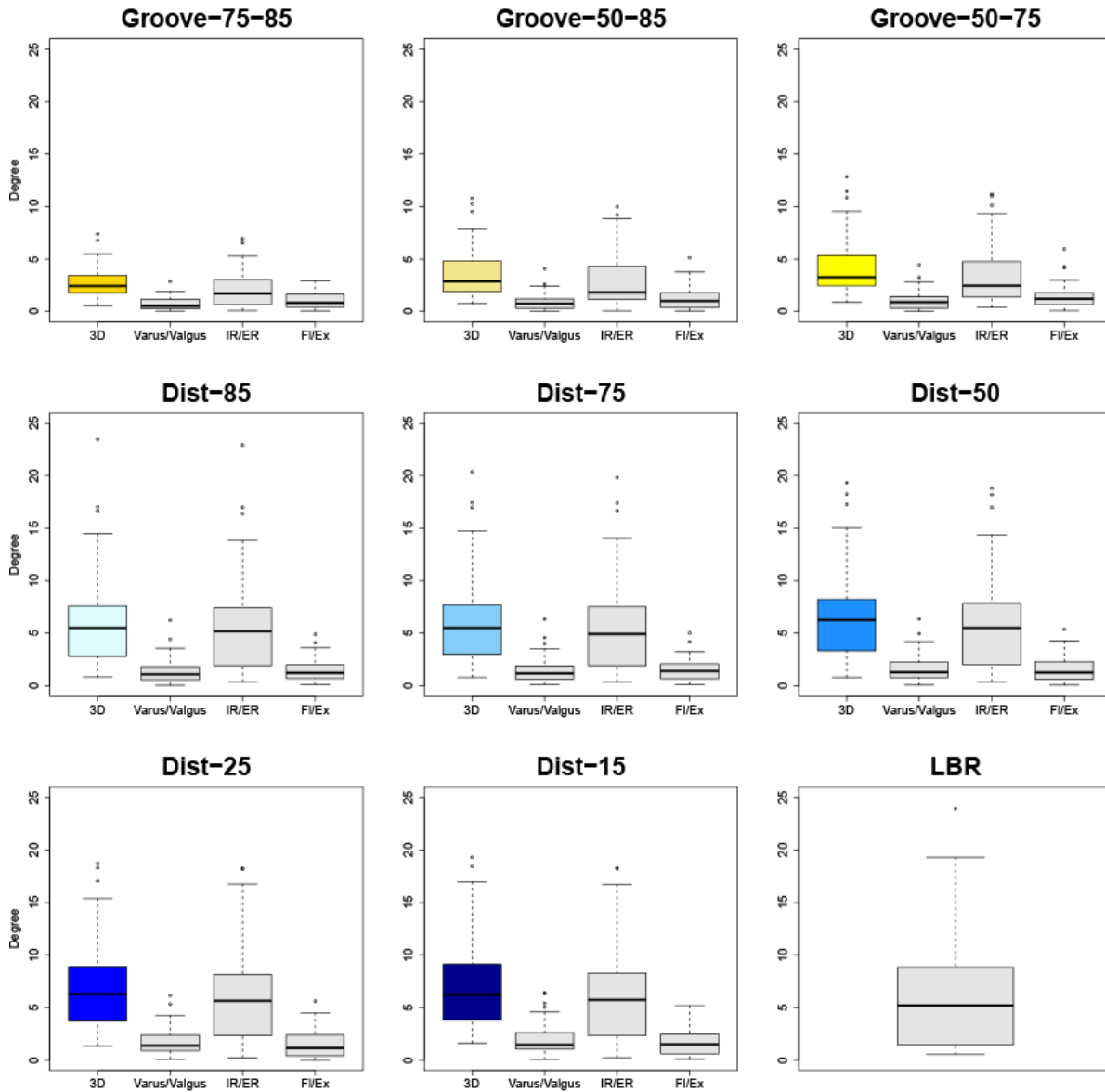


Figure 4.5: Boxplots of the p-HCR errors and of the LBR error. Tukey-boxplots of the p-HCR errors in angle axis representation (3D-p-HCR error) and as three consecutive rotations (varus/valgus, (IR/ER) internal-/external rotation and (FI/Ex) flexion/extension) around the axes of the humeral coordinate system. The LBR error is the absolute difference between the landmark-based retrotorsion angle of both sides of an individual.

4.4 Discussion

The benefit of computer-assisted planning and patient-specific instruments for corrective osteotomies of the humerus has already been emphasized (Murase et al., 2008; Omori et al., 2015; Vlachopoulos et al., 2016b). Furthermore, the postoperative analysis of the accuracy of corrective osteotomies of the upper extremities with this technique showed promising results (Omori et al., 2015). However, to the best of our knowledge, the accuracy of the contralateral anatomy as a reconstruction template for the approximation of the pretraumatic humeral head retrotorsion has not been analysed so far. Since differences in humeral head retrotorsion between both sides of an individual exist (Bicknell et al., 2007; Vlachopoulos et al., 2016a), these template-based approaches using the contralateral anatomy might introduce an error into the quantification of the deformity and may, thereby, compromise the ultimate goal of restoring the pretraumatic anatomy of the proximal humerus.

In the present study, we investigated the reliability of the contralateral anatomy as a 3D reconstruction template for corrective osteotomies of the proximal humerus. We evaluated a computer-based approach for approximating the humeral retrotorsion relying on the anatomy of the contralateral healthy side. To have a reference for comparison, we assessed also the LBR error. The mean LBR error was 6.4° , which corresponds to the intra-individual bilateral differences in humeral head retrotorsion (Bicknell et al., 2007; Vlachopoulos et al., 2016a). While the LBR error was $< 5^\circ$ in almost half of the cases, in 26% of the individuals the LBR error was between 5° and 10° and in further 26% of the cases $> 10^\circ$. The p-HCR errors around the y-axis (internal-/external rotation) of the registration-based approach using the segments of the distal group were similar as the LBR errors. This indicates that with registration of the distal segments the humeral models are superimposed in a way which preserves the bilateral differences in retrotorsion. Contrary, the p-HCR errors of the groove group were significantly smaller compared to the p-HCR errors of the distal group and the LBR errors. The differences between the *groove-50-85* and *groove-75-85* segments were not significant. If one of these groove segments was used for the distal registration, the p-HCR error was only in 2% of the cases $> 10^\circ$. It should be pointed out that all the reported errors in the current study represent only the deviation of approximated anatomy (i.e., planned correction by a corrective osteotomy) to the pretraumatic anatomy. In a postoperative worst-case scenario these errors might accumulate with the error introduced during surgery. Therefore, in our opinion, it is essential to reduce the systematic error during the quantification of a deformity.

The presented approach using the groove segments has the advantage that side-differences, with respect to retrotorsion, are compensated. Compared to the landmark-based method (Vlachopoulos et al., 2016a), the presented method may also be less sensitive to anatomical variations and inter-reader variability, due to the fact that the proximal humeral anatomy is approximated by applying the ICP algorithm in an automated fashion.

One limitation of the present study is that, similar as the state-of-the-art method, the preoperative assessment of the deformity can only be applied if the contralateral bone is healthy. However, since the distal part of the humerus is not necessary for the registration, the method can be used even if a pathological condition of the distal humeral anatomy on the contralateral side is present.

In conclusion, the bicipital groove contains characteristic features, which allows superimposing the humeral models of an individual in a way that bilateral differences in retrotorsion can be compensated. Figure 4.6 illustrates the cross-sectional shape of the *groove-75-85* segment. If the groove segments had no characteristic features (i.e., the cross-sectional area of the groove would resemble a cylindrical shape), the HCR errors would have probably been bigger or the consistency of the method would have been lower. The presented results confirm previous studies, that proposed the bicipital groove to be reliable landmark for the restoration of the humeral retrotorsion during shoulder replacement surgery (Doyle and Burks, 1998; Hempfing et al., 2001; Kontakis et al., 2001). In a recent study, Johnson et al. (Johnson et al., 2013) analysed the 3D shape of the bicipital groove and reported on a significant correlation between the bicipital groove rotation and the humeral head retrotorsion. In addition, the authors presented equations for predicting the humeral retrotorsion, once the groove rotation was assessed. However, the authors pointed out that advanced methods, such as computer navigation technology, would be required for measuring the groove. Moreover, they did not give a comparison of the approximation error resulting from the calculation of the torsion with the equations presented in their study to a gold standard. In addition, the sample size of the current study is much larger. While Johnson et al. investigated thirty-four humeri, in the current study we analysed one hundred (fifty paired) humeri.

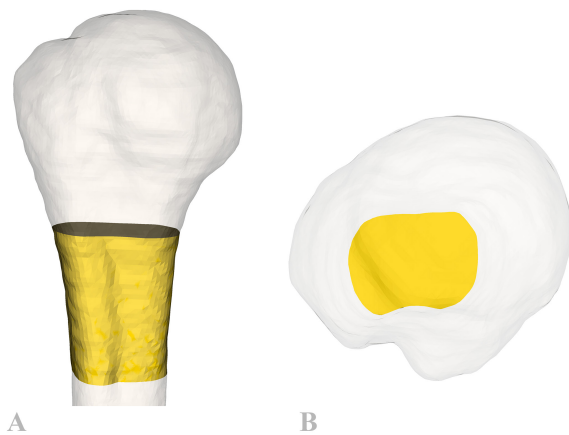


Figure 4.6: The 3D shape (A) of the *groove-75-85* segment and the cross-sectional shape (B) of the *groove-75-85* segment indicate that the bicipital groove has a characteristic shape (i.e., in contrast to cylindrical shape) that allows to superimpose the humeral models with the ICP algorithm in a consistent way.

In our opinion, it is preferable to use the proposed registration-based approach with the *groove-75-85* or *groove-50-85* segment (proximal shaft below the humeral head, which includes the bicipital groove) for the 3D preoperative planning of corrective osteotomies of the proximal humerus. Thereby, not only the error in restoring the pretraumatic anatomy becomes smaller compared to landmark-based retrotorsion measurement, but also the preoperative CT scan protocol can be adapted to reduce radiation exposure by omitting the acquisition of the mid- and distal shaft regions of the humeri. In the current study, the clinical application of the registration-based approach was only described in the context of corrective osteotomies of the humerus. However, the approach is also applicable for further computer-assisted methods, i.e. the computer-assisted shoulder hemiarthroplasty (Bicknell et al., 2007)

or the computer-assisted reconstruction of complex proximal humeral fractures (Fürnstahl et al., 2012) as these methods in general rely on the contralateral anatomy as a reconstruction template.

4.5 Conclusions

Registration of the proximal humeral shaft is a reliable method to approximate the proximal humeral anatomy and is our preferable method for the template-based preoperative planning of corrective osteotomies of the proximal humerus.

5

Statistical Shape Model of the Humerus

In the present paper, we evaluated whether a statistical shape model (SSM) has the potential to predict accurately the pretraumatic anatomy of the humerus from the posttraumatic condition. In the current state-of-the-art template-based approach, the contralateral anatomy is used as a reconstruction template. The increased radiation exposure, due to the acquisition of a CT scan of the contralateral anatomy, and the limited applicability if both sides are pathological are the main drawbacks of the state-of-the-art method. For comparison with the state-of-the-art we evaluated whether, based on the same segments of the humerus, the SSM or the contralateral anatomy yields a more accurate reconstruction template. We demonstrated that SSMs are powerful tools, accurately predicting the patient-specific pretraumatic anatomy of the proximal and distal humerus, especially if only the articular part of the humerus is pathological. However, both methods have their benefits and limitations and, therefore, we concluded that both approaches are not mutually exclusive or competitive, but rather complementary.

5.1 Introduction

Three-dimensional (3D) computer-assisted planning has been proposed for corrective osteotomies of the proximal and distal humerus (Murase et al., 2008; Omori et al., 2015; Vlachopoulos et al., 2016b). The implementation of the preoperative plan in the surgery has been proven to be very precise if patient-specific guides are used during surgery (Omori et al., 2015; Vlachopoulos et al., 2015). However, the most fundamental step influencing the accuracy of all subsequent steps of the procedure is the preoperative deformity assessment (Vlachopoulos et al., 2016b). In this technique, it is necessary to have a 3D reconstruction template that approximates the pretraumatic bone geometry as accurate as possible. Typically, the contralateral anatomy is selected as the most appropriate 3D reconstruction template (Bicknell et al., 2007; Murase et al., 2008; Omori et al., 2015; Vlachopoulos et al., 2016a; Vlachopoulos et al., 2016b). For the assessment of the deformity, 3D surface models of the pathological and the contralateral humerus are generated based on CT scans of both sides. The contralateral model is then mirrored and used to quantify the deformity. However, the presented methods so far (Murase et al., 2008; Omori et al., 2015; Vlachopoulos et al., 2016a; Vlachopoulos et al., 2016b) rely on a healthy contralateral anatomy and, therefore, the applicability of the 3D computer-assisted planning is limited if both sides are pathological.

The aim of the present study was to evaluate whether a SSM, a computer-generated model encoding the variation of the population, has the potential to predict the pretraumatic anatomy of the humerus from the healthy parts of the posttraumatic bone model. We analysed the influence of the size of the healthy part of the humerus used for the SSM on the prediction. For comparison with the state-of-the-art we evaluated whether, based on the same segments of the humerus, the SSM or the contralateral anatomy yields a more accurate reconstruction template.

5.2 Material and Methods

The Swiss Institute for Computer-Assisted Surgery (SICAS) provided computed tomography (CT) full-body data of 50 specimens (including the entire humerus on both sides). The in-plane (xy-) resolution of the CT scans ranged between 0.9 x 0.9 mm and 1.27 x 1.27 mm. The slice thickness varied from 0.5 mm to 0.6 mm. The average age of the individuals was 52.1 years (SD, 20.0 years; range: 19 to 90 years). There were 32 male and 18 female cadavers. The average height was 172.4 cm (SD, 8.7 cm; range: 154 to 187 cm) and the average weight was 68.4 kg (SD, 16.9 kg; range: 37 to 108 kg). Segmentation of the humerus was performed in an automatic fashion using a previously described segmentation algorithm (Gass et al., 2014). 3D triangular surface models were created using the marching cubes algorithm (Lorensen and Cline, 1987). The models were also used in previous studies (Vlachopoulos et al., 2017b; Vlachopoulos et al., 2016a).

In the following section we will first give a brief overview of the basic principle of the SSM algorithm, which was used for the generation of the SSM and the predictions. Second, we will describe the method for evaluating the accuracy of the prediction of the SSM. Lastly, we will compare the prediction quantitatively with the state-of-the-art method (Murase et al., 2008; Omori et al., 2015; Vlachopoulos et al., 2016a; Vlachopoulos et al., 2016b), where the contralateral bone is used as the reconstruction template. To compare the results with the state-of-the-art method, we used, in the current study, bilateral humeral models without a pathological condition.

5.2.1 Generation of the SSM

The 3D triangular meshes of the 50 right humeri were brought into correspondence using a non-rigid registration algorithm (Lüthi et al., 2016). Subsequently, all meshes were rigidly aligned using Procrustes alignment (Umeyama, 1991) to one humeral model. A SSM of the aligned meshes was built by performing a Principal Component Analysis (PCA) (Jolliffe, 2002). Hence, the resulting SSM encoded the mean shape of the humeri as well as the variation within the study population, which is represented by the principal components. To predict the pretraumatic shape of a certain humerus using the SSM, the pathological part was removed from the humeral mesh. Subsequently, the shape variation of the SSM, which best fitted the healthy part (prior knowledge) of the humerus was determined and used as a predictor for the assumed pathological part of the humerus. We refer to (Albrecht et al., 2013) for the mathematical details of the prediction procedure.

5.2.2 Evaluation method for the SSM

Leave-one-out-cross-validation was performed to evaluate the accuracy of prediction. To predict an assumed pathological part of a selected humerus, a SSM was calculated from all remaining humeri (49 humeri) except the selected one. We defined the term proximal prediction, if the assumed pathological part was the proximal humeral anatomy while the remaining part was considered to be healthy. In a similar manner, we defined the term distal prediction, if the assumed pathological part was the distal humeral anatomy. To analyse the size of healthy part of the humerus that must be provided to the SSM to achieve a good prediction, we created humeral segments of predefined length in a standardized fashion (Figure 5.1, A). For the proximal predictions, we defined the distal 50%, 75% or 85% of the humerus to be healthy (Figure 5.1, B). For the distal predictions, we defined the proximal 50%, 75% or 85% of the humerus to be healthy (Figure 5.1, D). We analysed the humerus prediction (HP) error, calculated as the deviation of the predicted anatomy from the original bone model.

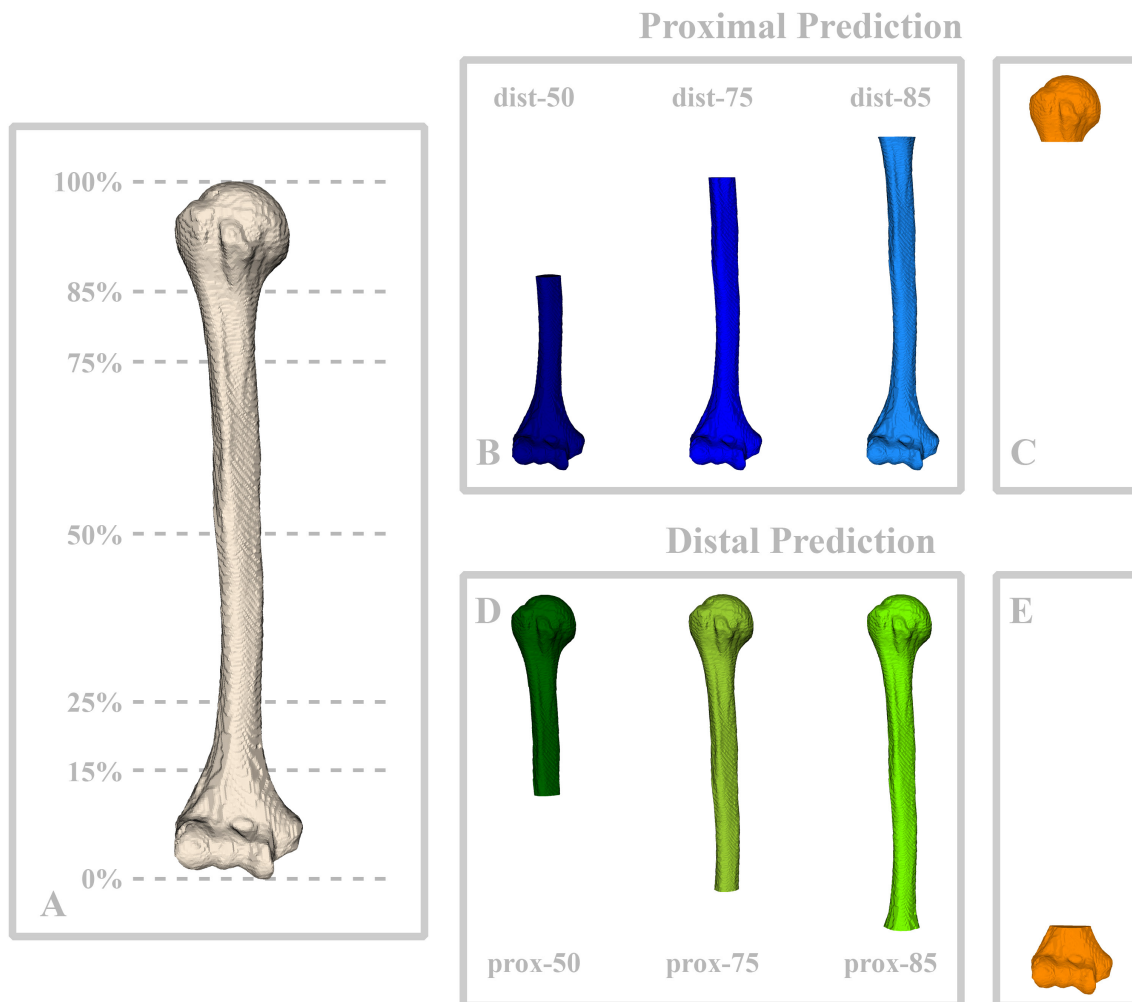


Figure 5.1: Definition of the segments A) The right humeral models were divided based on the length of the humerus into several segments. B) For the prediction of the proximal humerus (proximal prediction) we defined the distal 50%, 75% or 85% of the humerus (blue segments) to be healthy. C) The head-segment (orange) was used to quantify p-HP errors. D) For the prediction of the distal

humerus (distal prediction) we defined the proximal 50%, 75% or 85% of the humerus (green segments) to be healthy. E) The elbow-segment (orange) was used to quantify d-HP errors.

In total, we performed the following six predictions for each humerus and analysed the:

- *p-HP-50*, *p-HP-75* and *p-HP-85* errors, depending on the length of the segments (50%, 70% or 85%) of the distal humerus, which was assumed to be healthy and used as prior knowledge, and
- *d-HP-50*, *d-HP-75* and *d-HP-85* errors, depending on the length of the segments (50%, 70% or 85%) of the proximal humerus, which was assumed to be healthy and used as prior knowledge.

The same method for measuring HP errors was applied as used in the deformity assessment of corrective osteotomies of the humerus (Omori et al., 2015; Vlachopoulos et al., 2015). For the assessment of the p-HP errors we defined the proximal 15% of the humerus as the proximal segment (head-segment) of interest (Figure 5.1, C). For the assessment of the d-HP errors we defined the distal 15% of the humerus as distal segment (elbow-segment) of interest (Figure 5.1, E).

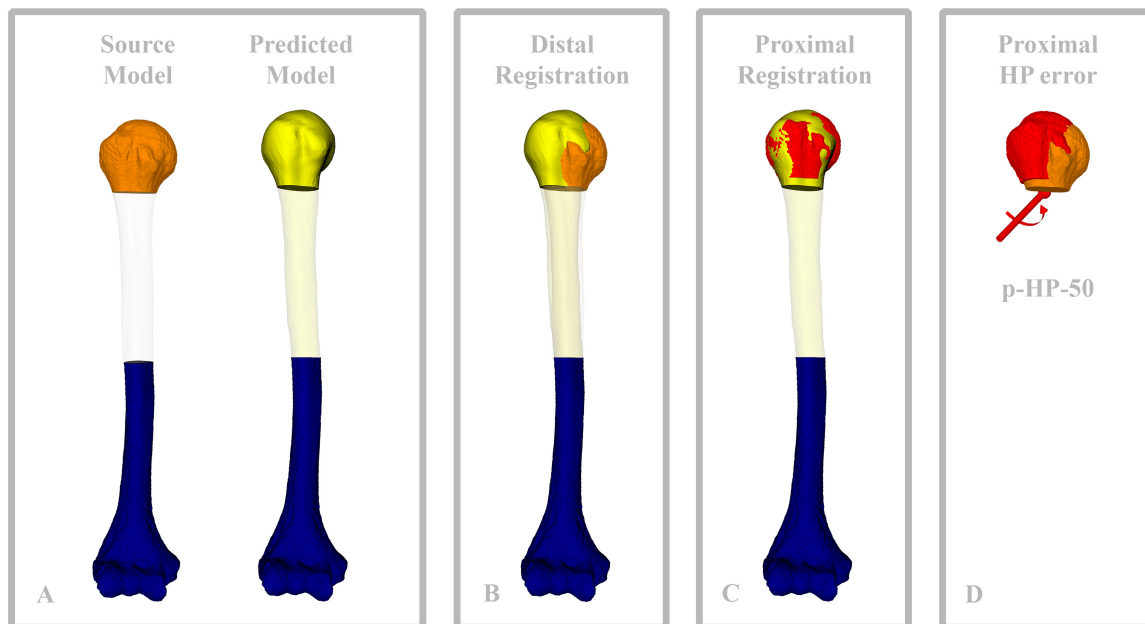


Figure 5.2: Evaluation of the humerus prediction error of the SSM. A) Illustrates the original bone model with the distal-50 segment (blue) used to predict the proximal humeral anatomy (yellow) and the head-segment (orange) used to quantify the prediction error. B) After registration of the distal segment of both models the difference between the original bone model (orange) and the prediction (yellow) becomes obvious. C) The ICP algorithm was used to superimpose the head-segment (red) of the original bone model to the predicted bone model. D) The 3D rotation of the head-segment between the distal registration (orange) and the proximal registration (red) was expressed in axis-angle representation.

To quantify the HP error, we superimposed the source model with the predicted model (Figure 5.2, A) using the surface registration method Iterative Closest Point (ICP) (Besl and McKay, 1992; Chen and Medioni, 1991). For the p-HP error, we performed first the distal registration by aligning the distal segments of the source model and the predicted model (Figure 5.2, B, blue segment). Thereafter, we performed the registration of the proximal head-segment onto the predicted model (Figure 5.2, C). The relative 3D rotation and 3D translation of the head-segment between distal (Figure 5.2, B) and proximal registration (Figure 5.2, C) quantifies the p-HP error. The d-HP errors were assessed accordingly for the distal prediction.

The rotational components of the prediction errors were expressed in axis-angle representation (Figure 5.2, D), i.e., rotation by a 3D angle about a calculated axis. The translational components of the prediction errors were expressed as the difference in humeral length between the predicted and the original humeral model. The humeral length was expressed by the length of the longest side of the oriented bounding box (OBB) of the humerus as described by Vlachopoulos et al. (2016a). The OBB is the minimal-volume rectangular box fully enclosing the humeral model (Figure 5.3).



Figure 5.3: Humeral length. The length of the humerus (i.e., original bone model or predicted bone model) was measured with the length of the longest side of the oriented bounding box of the humerus. The error in humeral length is, thereby, the difference between the length of the original bone model and the predicted bone model.

5.2.3 Reference Method

As a baseline we used the state-of-the-art method of preoperative deformity assessment where the mirrored contralateral humeral model serves as a reconstruction template (Figure 5.4, A). To compare the SSM-based and the reference methods in a standardized way, we used the same partitioning of healthy and pathological segments. For the proximal prediction we assumed the distal 50%, 75% or 85% of the humerus to be healthy, respectively. Accordingly, a posttraumatic deformity was assumed to exist between the distal segment and the proximal 15% of the humerus. First, we applied the ICP algorithm to superimpose the source model to the reconstruction template based on the distal segment (Figure 5.4, B blue segment). Thereafter, we applied the ICP algorithm to superimpose the head-segment of the source model (Figure 5.4, C orange segment) to the reconstruction template. The 3D rotation and 3D translation of the head-segment between the distal registration (Figure 5.4, B) and the proximal registration (Figure 5.4, C) quantifies the malalignment and, correspondingly, the proximal humerus contralateral registration (p-HCR) error. We quantified the distal humerus contralateral registration (d-HCR) error in the same way. The rotational component was expressed in axis-angle representation (Figure 5.4, D). The translational component was expressed as the difference in humeral length between the humeral models of both sides of an individual. The humeral length was expressed by the length of the longest side of the OBB similar as before.

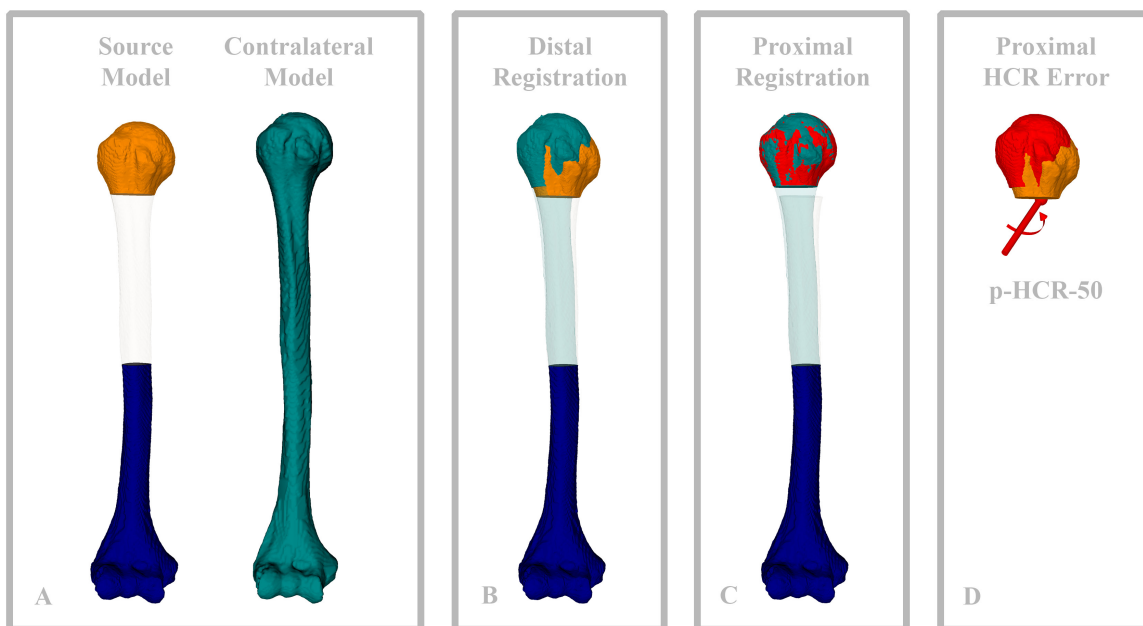


Figure 5.4: Evaluation of the HCR error using the mirrored contralateral model as a reconstruction template. A) The original bone model is illustrated with the distal-50 segment (blue) and the head-segment (orange) used for the registration onto the mirrored contralateral humeral model (cyan). B)

After the registration based on the distal segment using the ICP algorithm, the difference between the original bone model and the reconstruction template becomes obvious. C) The ICP algorithm was used to superimpose the head-segment of the original bone model (red) onto the mirrored contralateral anatomy. D) The 3D rotation of the head-segment between the distal registration (orange) and the proximal registration (red) was expressed in axis-angle representation.

In total, we yield the following six HCR errors per humerus for the reference method:

- the *p-HCR-50*, *p-HCR-75* and *p-HCR-85* errors, depending on the length of the segments (50%, 75% or 85%) of the distal humerus, which was assumed to be healthy, and
- the *d-HCR-50*, *d-HCR-75* and *d-HCR-85* errors, depending on the length of the segments (50%, 70% or 85%) of the proximal humerus, which was assumed to be healthy.

5.2.4 Statistical Analysis

Mauchly's sphericity test revealed a violation of the assumption of sphericity. Therefore, we applied the non-parametric Friedman Rank Sum test with the segment as a group factor and the individuals as a block factor to analyse the effect of the defined segments on the proximal and distal HP error, the HCR error and the error in humeral length. Post-hoc analysis was performed with the Wilcoxon signed-rank test and Bonferroni adjustment. The significance level was set at $P < 0.05$. For graphical visualization, Tukey boxplots have been used with the end of the whiskers indicating the 1.5 interquartile range (IQR) of the lower and upper quartile. The mean, SD, median and range of the error in humeral length as well as the rotational errors were calculated from the absolute values.

5.3 Results

The Friedman rank sum test revealed a significant effect of the selected segments on the p-HP and p-HCR error (chi-squared 61.5, $P < 0.0001$). Fifteen paired Wilcoxon signed-rank test were performed for post hoc analysis. The Bonferroni-adjusted p-value (p_{adj}) was $0.05/15 \approx 0.0033$ for each individual comparison, in order to retain the prescribed family-wise error rate of 0.05.

Figure 5.5 illustrates the p-HP errors and the p-HCR errors based on the same segments of the healthy part of the distal humerus. The mean values, standard deviation, median and range of the p-HP and the p-HCR errors are summarized in Table 5.1. The p-values of the post-hoc analysis of the p-HP and p-HCR error performed with the pairwise Wilcoxon signed-rank test are depicted in Table 5.2. The p -HP-85 error was significant smaller than any other p-HP error or any p-HCR error. The p -HP-75 error was significant smaller ($P < 0.0001$) than the p -HP-50 error. The difference between the p -HP-75 error or the p -HP-50 error and the p-HCR errors were not significant.

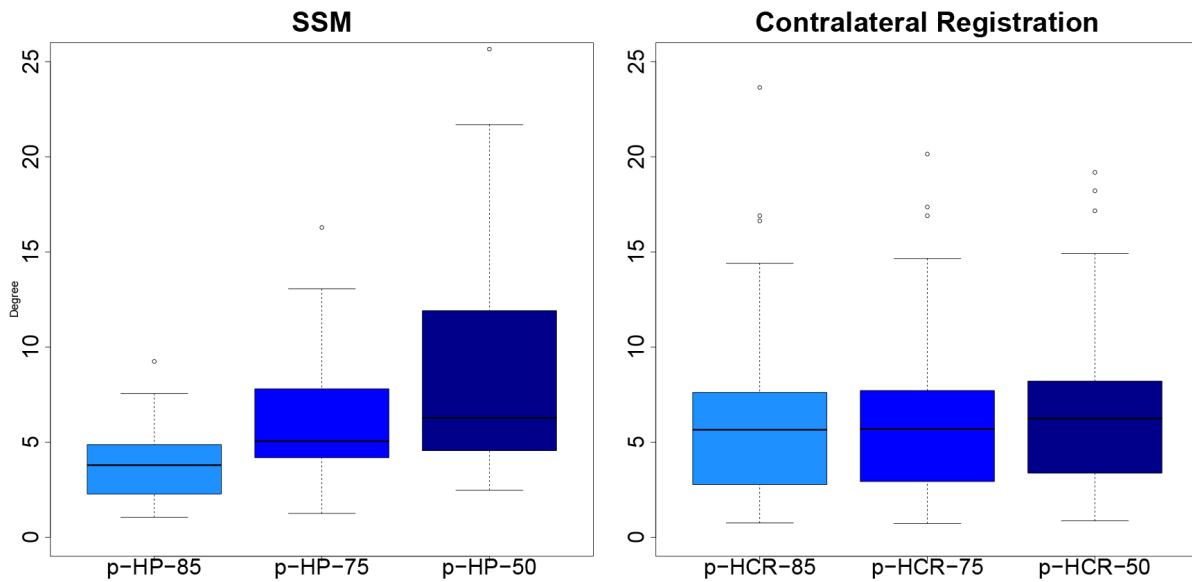


Figure 5.5: Proximal HP errors and HCR errors. Boxplots illustrate the p-HP errors and the p-HCR errors based on the segments of the healthy part of the distal humerus that were used for the prediction of the SSM or for the reference method

Segment	p-HP-error (°)					p-HCR-error (°)				
	Mean	SD	Median	Range		Mean	SD	Median	Range	
dist-85	3.8	1.9	3.8	1.1	- 9.2	6.5	4.7	5.5	0.8	- 23.5
dist-75	6.1	3.2	5.0	1.3	- 16.3	6.5	4.6	5.5	0.8	- 20.4
dist-50	8.9	5.6	6.3	2.5	- 25.7	6.9	4.6	6.3	0.8	- 19.3

Table 5.1: Mean, SD, median and range of the p-HP error and the p-HCR error for each selected segment.

	p-HP-75	p-HP-50	p-HCR-85	p-HCR-75	p-HCR-50
p-HP-85	<0.0001	<0.0001	0.001	<0.001	<0.001
p-HP-75		<0.0001	0.80	0.79	0.40
p-HP-50			0.03	0.04	0.09
p-HCR-85				<0.0001	<0.0001
p-HCR-75					<0.0001

Table 5.2: P-values of the post-hoc analysis of the p-HP error and the p-HCR error performed with the Wilcoxon signed-rank test and Bonferroni adjustment ($p_{adj} = 0.05/15 \approx 0.0033$).

The Friedman rank sum test revealed a significant effect of the selected segments on the d-HP and d-HCR error (chi-squared 52.1, $P < 0.0001$). Fifteen paired Wilcoxon signed-rank test were performed for post hoc analysis. The Bonferroni-adjusted p-value (p_{adj}) was $0.05/15 \approx 0.0033$ for each individual comparison, in order to retain the prescribed family-wise error rate of 0.05. Figure 5.6 illustrates the d-HP errors and the d-HCR errors based on the same segments of the healthy part of the proximal humerus. The mean values, standard deviation, median and range of the d-HP and the d-HCR errors are summarized in Table 5.3. The p-values of the post-hoc analysis of the d-HP and d-HCR error performed with the pairwise Wilcoxon signed-rank test are depicted in Table 5.4. The *d-HP-85* error was significant smaller than any other d-HP error. The difference between the *d-HP-85* error and the d-HCR errors was not significant. The *d-HCR-85* error was significant smaller than the *d-HP-75* or *d-HP-50* error. The *d-HCR-75* error was significant smaller than the *d-HP-50* error. The difference between the *d-HP-75* and the *d-HCR-75* error as well as the difference between the *d-HP-50* and the *d-HCR-50* error was not significant.

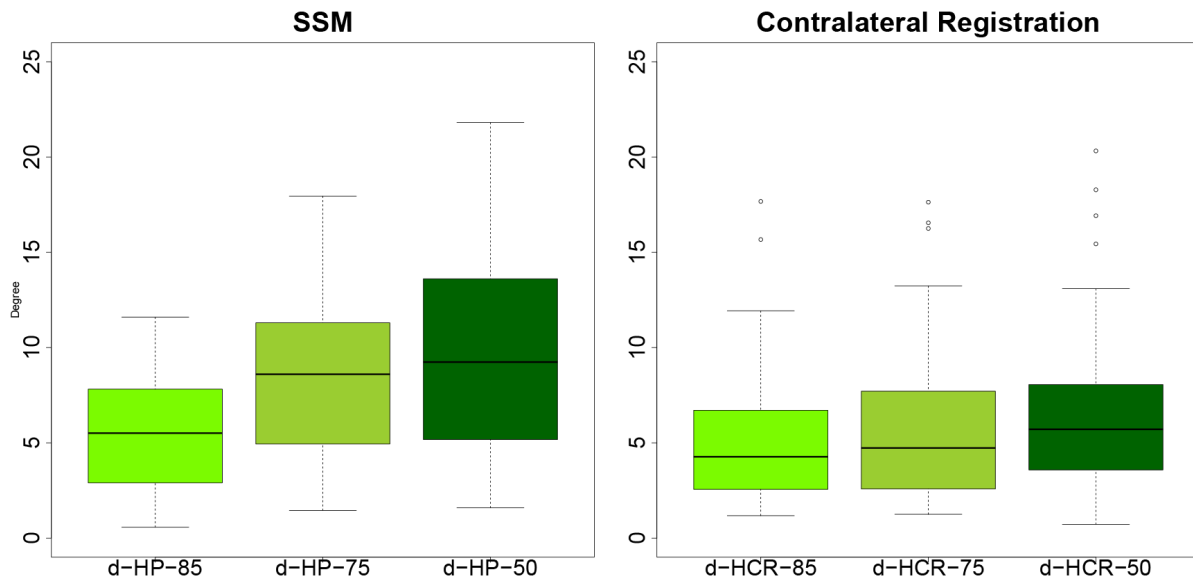


Figure 5.6: Distal HP errors and HCR errors. Boxplots illustrate the d-HP errors and the d-HCR errors based on the segments of the healthy part of the proximal humerus that were used for the prediction of the SSM or for the reference method.

Segment	d-HP-error (°)					d-HCR-error (°)				
	Mean	SD	Median	Range		Mean	SD	Median	Range	
prox-85	5.5	2.9	5.5	1.1	- 11.6	5.2	3.7	4.0	1.1	- 17.7
prox-75	8.7	4.6	8.6	1.5	- 17.9	5.7	4.3	4.7	1.3	- 17.7
prox-50	9.9	5.7	9.3	1.6	- 21.8	6.7	4.5	5.7	0.7	- 20.5

Table 5.3: Mean, SD, median and range of the d-HP error and the d-HCR error for each selected segment.

	d-HP-75	d-HP-50	d-HCR-85	d-HCR-75	d-HCR-50
d-HP-85	<0.0001	<0.0001	0.61	0.94	0.25
d-HP-75		<0.0001	0.001	0.007	0.05
d-HP-50			<0.0001	<0.001	0.09
d-HCR-85				<0.0001	<0.0001
d-HCR-75					<0.001

Table 5.4: P-values of the post-hoc analysis of the d-HP error and the d-HCR error performed with the Wilcoxon signed-rank test and Bonferroni adjustment ($p_{adj} = 0.05/15 \approx 0.0033$).

The Friedman rank sum test revealed a significant effect of the selected segments on the prediction of humeral length (chi-squared 19.7, $P = 0.003$). Twenty-one paired Wilcoxon signed-rank test were performed for post hoc analysis. The Bonferroni-adjusted p-value (p_{adj}) was $0.05/21 \approx 0.0023$ for each individual comparison, in order to retain the prescribed family-wise error rate of 0.05. Figure 5.7 illustrates the error in humeral length of the SSM

and of the reference method. The mean values, standard deviation, median and range of the errors in approximating the humeral length are summarized in Table 5.5. The p-values of the post-hoc analysis performed with the pairwise Wilcoxon signed-rank test are depicted in Table 5.6. The difference between the length determined with the SSM and the bilateral differences in humeral length (contra) of an individual were not significant.

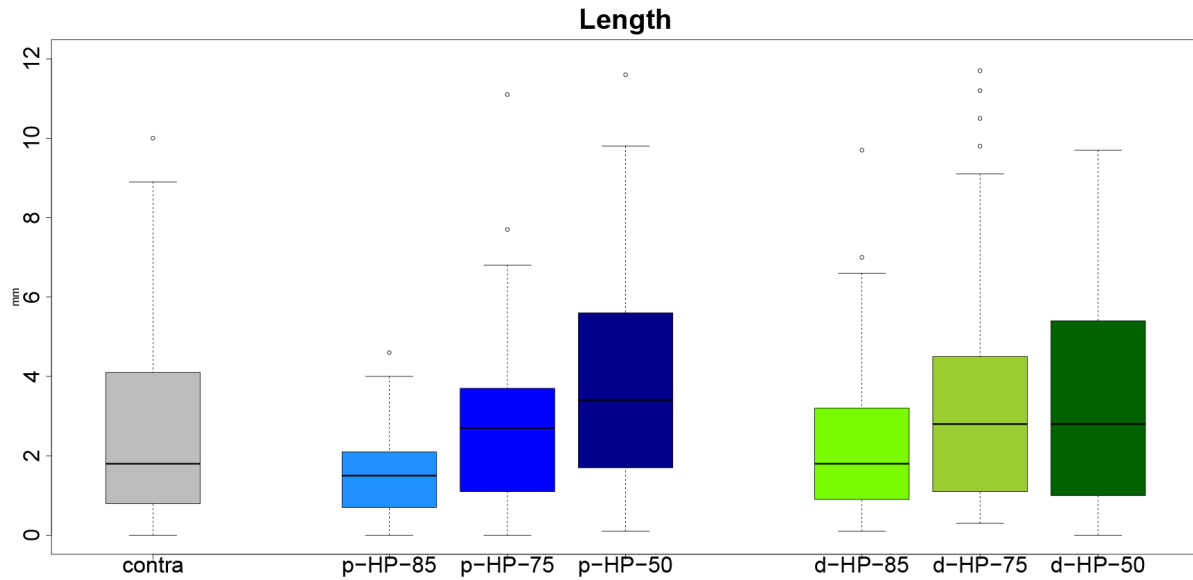


Figure 5.7: Differences in humeral length. Boxplots illustrate the error in humeral length if the contralateral anatomy is used as a reconstruction template (contra) or as the difference between the predicted humeral anatomy and the original bone models

	Length error (mm)			
	Mean	SD	Median	Range
Contra	2.7	2.5	1.8	0.0 - 10.0
p-HP-85	1.6	1.1	1.5	0.0 - 4.6
p-HP-75	2.8	2.2	2.7	0.0 - 11.1
p-HP-50	4.2	3.6	3.4	0.1 - 15.0
d-HP-85	2.4	2.0	1.8	0.1 - 9.7
d-HP-75	3.5	3.1	2.8	0.3 - 11.7
d-HP-50	3.4	3.0	2.8	0.0 - 12.7

Table 5.5: Mean, SD, median and range of the bilateral differences in humeral length (contra) and the difference between the humeral length of the predicted humeral anatomy and the original bone model for each selected segment. All values are calculated from the absolute differences.

	p-HP-85	p-HP-75	p-HP-50	d-HP-85	d-HP-75	d-HP-50
Contra	0.02	0.57	0.03	0.32	0.19	0.37
p-HP-85		0.002	<0.0001	0.02	0.002	<0.001
p-HP-75			0.003	0.25	0.19	0.53
p-HP-50				<0.001	0.20	0.24
d-HP-85					0.02	0.04
d-HP-75						0.87

Table 5.6: P-values of the post-hoc analysis of the error in humeral length performed with the Wilcoxon signed-rank test and Bonferroni adjustment ($p_{adj} = 0.05/21 \approx 0.0023$).

5.4 Discussion

Restoration of the normal, pretraumatic humeral anatomy is an ultimate goal in corrective osteotomies of the humerus. Computer-assisted methods and the use of patient specific instrumentation have proven to be a versatile method to realize the preoperative plan in the surgery (Murase et al., 2008; Omori et al., 2015; Vlachopoulos et al., 2016b). In the present study, we evaluated whether a SSM has the potential to accurately predict the pretraumatic anatomy from the healthy part of the humerus in a posttraumatic condition. We analysed the prediction errors of the SSM and compared the predicted anatomy with the so far state-of-the-art. We demonstrated that the prediction errors of the SSM depend on the size of the healthy part of the humerus. Therefore, the whole segment of the humerus should be used for the prediction of the anatomy when the method is used clinically for the preoperative planning of corrective osteotomies of the humerus. The HCR errors also decrease significantly with increasing size of the healthy part of the humerus. However, the effect of the size of the healthy part was smaller for the reference method than for the SSM.

The differences between the d-HP errors and the d-HCR errors were not significant and, therefore, the predicted anatomy is comparable with the registration-based approach (Murase et al., 2008; Omori et al., 2015). The differences between the p-HP-errors and the p-HCR-errors, when 50% or 75% of the distal humerus were healthy, were neither significant. However, the *p-HP-85* error was significantly smaller than any p-HP or any p-HCR, which means the reconstruction template predicted by the SSM, outperforms the reference method. There was no significant difference in humeral length between predicted and contralateral template. Since the contralateral anatomy is assumed to be a reliable reconstruction template for the humeral length (Vlachopoulos et al., 2016a), we can conclude this also for the SSM.

The main benefits of the SSM are that the acquisition of a CT scan of the contralateral anatomy is not necessary and the method is even applicable if bilateral pathological conditions are present. However, one limitation of the present study is that, the whole proximal or distal humerus was used for the prediction with the SSM. The influence of an additional pathological condition in the assumed healthy proximal or distal part was not analysed and, therefore, the method might not be applicable if a combined proximal and distal posttraumatic deformity is present. On the contrary, Vlachopoulos et al. recently demonstrated, that the reference method might still be applicable, if a combined proximal and distal deformity is present, by definition of segments as close as possible to the area of

interest (Vlachopoulos et al., 2017b). In our experience, both methods have their benefits and limitations and, therefore, both approaches are not mutually exclusive or competitive, but rather complementary.

5.5 Conclusions

SSMs are powerful tools, accurately predicting the patient-specific pretraumatic anatomy, especially if only the articular part of the humerus is pathological. Therefore, SSMs are a valuable alternative to the registration-based approach using the contralateral anatomy as a 3D reconstruction template for corrective osteotomies of the proximal and distal humerus. These findings are particularly important, when a CT scan of the contralateral anatomy is not available or if bilateral pathological conditions are present.

6

Regression Forest Based Automatic Estimation of the Articular Margin Plane

In this paper, we present a fully automated method for the estimation of the articular margin plane (AMP), a plane approximating the cartilage/metaphyseal boundary at the proximal humerus. The combination of the estimation of the AMP with the methods presented before, allows an automatic measurement of the humeral anatomy in 3D. Furthermore, the AMP defines the orientation, position and size of the humeral head in relation to the humeral shaft. Therefore, the correct definition of the AMP is crucial for the computer-assisted preoperative planning of anatomical shoulder replacement surgery. The method uses random regression forests (RF) to establish a direct mapping from the CT image to the AMP parameters in two consecutive steps. In the first step, image intensities serve as features to compute a coarse estimate of the AMP. The second step builds upon this estimate, calculating a refined AMP using novel feature types that combine a bone enhancing sheetness-measure with ray features.

6.1 Introduction

Treatment of primary and secondary degenerative conditions of the shoulder with joint replacement surgery has substantially increased over the past decade, resulting in an annual growth rate of about 10% (Day et al., 2010). While approximately 7000 shoulder joint replacement surgeries were performed in the United States in the year 2002 (Bohsali et al., 2006), today more than 50'000 patients are treated per year (Ponce et al., 2015). The growing number of shoulder arthroplasty procedures also promoted the development of computer-assisted preoperative planning and intraoperative navigation techniques (Iannotti et al., 2014; Kircher et al., 2009; Nguyen et al., 2009). The improved accuracy of component positioning in presence of osteoarthritis (Edwards et al., 2008; Iannotti et al., 2014; Kircher et al., 2009; Nguyen et al., 2009) as well as the more precise reconstruction of comminuted proximal humeral fractures (Bicknell et al., 2007; Frnstahl et al., 2012) do often justify the increased planning effort preoperatively. One important step during shoulder arthroplasty is the resection of the humeral head along the anatomical neck (Fucentese et al., 2010). Resection is performed by sawing along the so-called articular margin plane (AMP) which lies on the cartilage to bone transition (Boileau and Walch, 1997). The AMP defines the orientation, position and size of the prosthetic humeral head in relation to the humeral shaft and influences the selection of humeral head component size (Figure 6.1, A). As the AMP varies from

patient to patient (Iannotti et al., 1992), it must be defined individually for each case. Therefore, the estimation of the AMP based on preoperative computed tomography (CT) data is a crucial step for the preoperative planning of arthroplasty. The current standard procedure for determining the AMP preoperatively relies on manually defined reference points (DeLude et al., 2007; Johnson et al., 2013). Because the reference points are difficult to identify in the CT image, annotation is often carried out using a 3D triangular surface model, reconstructed from the CT data. Figure 6.1, B shows an example of a humerus model with manually annotated reference points. However, the procedure of manually annotating the AMP is tedious, time consuming (i.e., 20 min) and, dependent on the experience of the surgeon, the AMP parameters may vary considerably between different persons (Vlachopoulos et al., 2016a).

Therefore, an automatic approach for determining the AMP would be desired to reduce manual effort. However, classical image processing techniques, such as template matching or spherical Hough transform, are likely to yield poor performance due to the lack of obvious landmarks or unambiguous geometries. Contrary, supervised machine learning algorithms allow for more versatile models and – depending on the chosen algorithm – are able to select a small number of discriminative features from a large set of candidates to effectively tackle the problem at hand.

In this paper, we present a random regression forest (RF)-based method to estimate the AMP from upper arm CT images. To the best of our knowledge, there is no prior work addressing this problem. A key aspect of our method is to take a two-step approach with novel feature types that combine a bone-enhancing sheetness measure (Descoteaux et al., 2006) with ray features (Smith et al., 2009). As our method is fully automatic, it can be seamlessly integrated into the existing workflows.

In the remainder of this section we briefly review prior work addressing similar problems. Section 6.2 provides the relevant background and gives a detailed description of our algorithm. Sections 6.3 and 6.4 are dedicated to the experimental evaluation of the proposed method. Finally, in Section 6.5 we discuss the results of the experimental evaluation and conclude with indications for future work in Section 6.6.

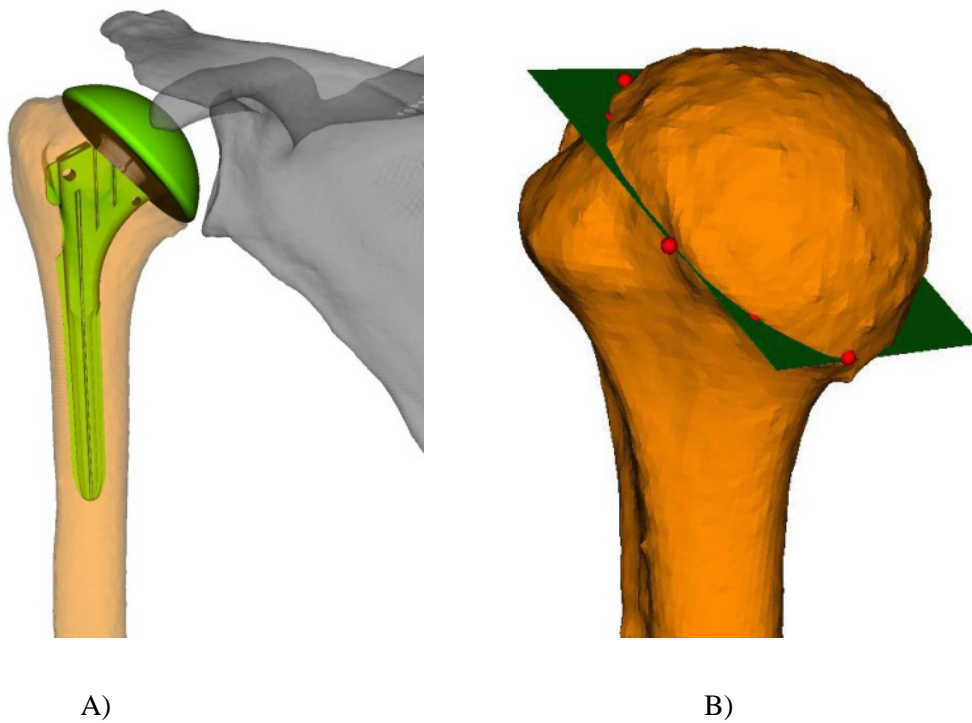


Figure 6.1: A) Humerus with prosthetic humeral head component (light green) and scapula (light grey). B) Humeral head and shaft with manually annotated points (denoted by red spheres) and resulting AMP (dark green).

6.1.1 Related Work

In the discussion of related work, we will focus on methods related to estimating position, orientation, anatomical landmarks, and orthopaedic parameters from CT images. These methods may be roughly categorized into template matching, Hough transform and learning-based methods.

6.1.1.1 Template matching methods:

In Ehrhardt et al. (2003), an atlas-based automatic segmentation and landmark localization framework for hip operation planning was proposed. In their method, a landmark-annotated grey value atlas was first registered to a CT image of the hip. Thereafter, the landmark locations were transferred from the atlas coordinate system to the CT coordinate system. The so-obtained coarse estimates of the landmark positions were then refined by locally registering surface atlases to the bone surface. A mean localization error below 1 mm was reported. However, no detailed performance evaluation was presented. Casciaro and Craiem (2014) described an automatic method for estimating anatomical bone axes in CT scans of ex-vivo femora. To do so, the bone surface was extracted from CT using a Laplacian filter, followed by portioning of the surface points into different bone segments. The axis of each segment was then obtained by fitting a cylinder to each point cloud.

Due to the absence of salient anatomical structures on the surface of the humeral head and its symmetric shape, simple atlas registration may yield inaccurate results (Förnsthahl et al., 2012).

6.1.1.2 Hough transform methods:

Van der Glas et al. (2002) applied spherical Hough transform to CT and magnetic resonance (MR) images for determining centre and radius of the shoulder joint. Ruppertshofen et al. (2011) generated discriminative shape models in an iterative procedure and combined these models with generalized Hough transform to localize hip, knee and ankle joints in CT images. The mean localization error achieved for hip, knee, and ankle was 12.5 mm, 4.3 mm, and 9.8 mm, respectively. In our case, determining the shoulder joint centre via spherical Hough transform would not help calculating the AMP, as the AMP is considerably offset with respect to the joint centre. Moreover, the computational complexity of Hough transform scales badly with respect to the number of model parameters and image size.

6.1.1.3 Learning-based methods:

The works of Glocker et al. (2012) and Kelm et al. (2013) addressed the problem of estimating position and angulation of vertebrae in CT and MR images. In Glocker et al. (2012), the authors used RF to localize the vertebrae centroids with a mean error of 18.4 mm. The estimates were then refined using a probabilistic graphical model. Kelm et al. (2013) used marginal space learning and a global probabilistic spine model to simultaneously determine pose and labelling of all vertebral discs. After refinement, a mean error of 3.2 mm and 4.5° was reported.

In Donner et al. (2013) a method combining a landmark classifier with a Hough RF and a parts-based model of the global landmark topology was proposed to simultaneously localize multiple landmarks. The method was trained and tested for localizing landmarks in hand radiographs, hand CTs, and whole body CTs (corresponding mean localization errors: 0.99 mm, 1.45 mm, 5.25 mm).

Han et al. (2015) considered the problem of landmark localization in brain MR images for registration and devised an algorithm that iteratively refines the landmark locations via a cascade of RFs (mean landmark localization error: 2.22 mm). A cascade of RFs, together with a local feature-weighting scheme, was also proposed by Ebner et al. (2014) for localizing landmarks in hand CTs (mean landmark localization error: 1.44 mm) and by Gao and Shen (2014) for landmark localization in prostate CT images (mean landmark localization error: 4.67 mm).

The works of Chen and Zheng (2013) and Lindner et al. (2013) targeted anatomical landmarks detection for the segmentation of the pelvic bone in radiographs. In Chen and Zheng (2013), the authors repeatedly determined anatomical landmarks on femur and pelvis via RFs, regularizing the landmark positions using active shape models (mean point-to-curve error: 2.2 mm). A similar approach was proposed in Lindner et al. (2013), where a voting-based RF scheme was used to calculate the bounding box of proximal femora. Afterwards, the bounding box was used to constrain the search area for the landmarks that were also localized

with RFs. Lindner et al. (2013) reported a mean point-to-curve error below 0.9 mm for 99% of the test images.

Calculating the AMP may be considered as a pose estimation problem as well, because the AMP encodes the pose of the humeral head with respect to the shaft. Employing RFs to estimate pose in terms of position and orientation simultaneously was mainly considered outside the field of medical image analysis, e.g., in Fanelli et al. (2011), where a RF framework is used to extract the human head pose from depth images in real time, and in Shotton et al. (2013), where the authors proposed RFs to simultaneously localize 16 body joints in depth images.

Learning-based methods, in particular RFs, have proven to be effective in many related applications of medical image processing (Criminisi and Shotton, 2013) and are therefore considered here for estimating the AMP.

6.2 Method

In this section, we present a fully automatic algorithm for estimating the AMP in CT images of the upper arm. Figure 6.2 gives an overview of our algorithm, consisting of two consecutive steps, both divided into training and test phases. In a first step, a RF is employed to determine a coarse estimate of the AMP parameters based on the intensity neighbourhood of voxels located on a regular grid by averaging over candidates with high confidence. More specifically, the RF computes offsets describing the location of the AMP relative to the grid voxels and angles indicating the orientation of the AMP with respect to the CT coordinate system. In the second step, the AMP parameters are refined iteratively by considering particularly the local geometry of the humeral head. Here, the regression target is a correction of the AMP parameters (i.e., a parameter offset) and voxels around the previously estimated candidates are used to feed RFs that incorporates novel ray features.

Before providing algorithmic details in Sections 6.2.4 and 6.2.5, we present the relevant theoretical background on RFs in Section 6.2.1. Further, the AMP parameterization is described in Section 6.2.2 and the novel sheetness-based ray features are presented in Section 6.2.3.

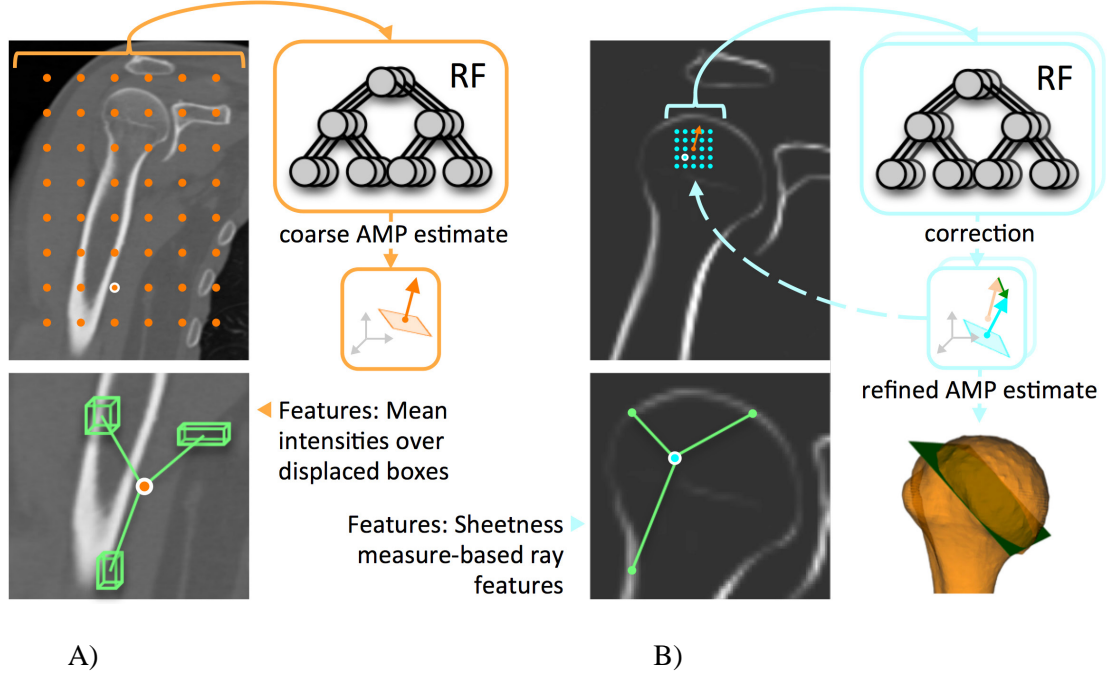


Figure 6.2: Overview of the proposed two-step procedure. (A) A RF with intensity-based features computed at grid points yields a coarse estimate of the AMP; (B) Iterative refinement with RFs using sheetness-based ray features, computed in a local neighbourhood of the humeral head.

6.2.1 Theoretical background on RFs

A RF (Breiman, 2001) is a powerful tool for learning a non-linear mapping from a feature space to a parameter space. In our case, the features will be extracted from CT data and mapped to AMP parameters. To do so, a set of T regression trees are trained on bootstrapped datasets to hierarchically partition the feature space into disjoint regions with distinct corresponding responses via simple tests and greedy randomized optimization. Our present work is based on the RF framework introduced in Criminisi et al. (2011), also described in detail in the book chapter (Criminisi and Shotton, 2013). We briefly summarize the framework for two-dimensional (2D) feature types (i.e., we consider pairs of feature coordinates), because our approach relies on 2D features.

The training is based on a set of N feature-response pairs $S = \{(\mathbf{x}_{i,j}, \mathbf{y}_i)\}_{i=1}^N$ (here $\mathbf{x}_{i,j} \in \mathbb{R}^2$), consisting of a feature set $\{\mathbf{x}_{i,j}\}$ (j indexing the feature) and corresponding desired responses $\{\mathbf{y}_i\}$, extracted from a training dataset. During training, a tree structure is constructed by learning for each tree node a test of the type

$$\langle \mathbf{x}_{i,j}, \mathbf{a} \rangle + b \geq 0, \quad (6.1)$$

where $\langle \cdot, \cdot \rangle$ denotes the dot product and the parameters $\mathbf{a} \in \mathbb{R}^2, b \in \mathbb{R}$ describe the parameters of a line in \mathbb{R}^2 . S is partitioned into two disjoint sets S_L and S_R , i.e., the training samples are assigned either to the left or right child nodes dependent on whether the result of the test is positive or negative. Each test is evaluated for a randomly chosen subset of the feature set

(i.e., for a randomly chosen set of indices j) and over a set of randomly generated parameters \mathbf{a}, \mathbf{b} to find the node test that maximizes the information gain (Criminisi et al., 2011)

$$IG = \log(|\Sigma(S)|) - \sum_{k \in \{L, R\}} w_k \log(|\Sigma(S_k)|).$$

w_k denotes the fraction of feature-response pairs assigned to side k (i.e., $w_k = |S_k| / |S|$) and $\Sigma(S)$ is the empirical covariance matrix of the responses in S . The assumption of using IG as a quality measure for node tests is that the responses follow a normal distribution. Complete regression trees are constructed by recursively applying this node learning scheme until the number of samples falls below a threshold, or until the maximum admissible tree depth is reached.

To obtain the tree response for a test sample, a sequence of simple node tests, as given in Equation (6.1) is evaluated to determine the corresponding leaf node l . The tree response distribution at l is then given by the multivariate normal distribution $\mathcal{N}(\boldsymbol{\mu}_l, \Sigma_l) = \mathcal{N}(\boldsymbol{\mu}(S_l), \Sigma(S_l))$, where S_l is the subset of (training) feature-response pairs at leaf node l . $\boldsymbol{\mu}(S_l)$ denotes the mean of the responses in S_l . To combine the responses of the T trees in a forest, originally Breiman (2001) proposed to average over all responses. However, here, we average only over a subset of responses with high confidence for achieving higher accuracy, as demonstrated by Criminisi et al. (2011) and Fanelli et al. (2011). Details on the selective averaging procedure are given in Sections 6.2.4 and 6.2.5.

6.2.2 AMP parameterization

The AMP is parameterized by the centre \mathbf{c} and by angles φ and θ , as demonstrated in Figure 6.3. Centre \mathbf{c} is defined as the geometric centre of the curve resulting from the intersection of AMP and humeral head surface. Angles φ and θ specify the orientation of the AMP normal \mathbf{n} relative to the CT coordinate system. Grouped together, \mathbf{c}, φ , and θ form the parameter vector $\mathbf{v} = [\mathbf{c}^T, \varphi, \theta]^T$. Note that \mathbf{c} does not correspond to the centre of the proximal humeral head. Therefore, \mathbf{c} cannot be simply found by fitting a sphere to the head in least-squares sense (Schneider and Eberly, 2002). However, the calculation of \mathbf{c} is important for preoperative planning, because a prosthetic humeral head component is ideally centred in \mathbf{c} . We further define the auxiliary coordinate system $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ with origin \mathbf{c} and axes parallel to those of the CT coordinate system $(\mathbf{x}, \mathbf{y}, \mathbf{z})$, as shown in Figure 6.3.

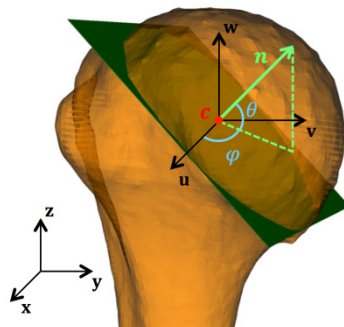


Figure 6.3: AMP parameterization. The CT coordinate system is denoted by $(\mathbf{x}, \mathbf{y}, \mathbf{z})$. The axes of the auxiliary coordinate system $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ are parallel to those of $(\mathbf{x}, \mathbf{y}, \mathbf{z})$.

6.2.3 Sheetness-based ray features

This section gives a detailed description of the sheetness-based ray features used later in the refinement step (see Section 6.2.5). First, a variant of the sheetness measure proposed by Descoteaux et al. (2006) will be introduced, which we call *oriented sheetness measure* (OSM). Based on the OSM, ray features (Smith et al., 2009) for the RF will be developed.

Inter alia, the sheetness measure has been used as a prior for bone segmentation (Fürnstahl et al., 2008; Gass et al., 2014), because it enhances planar (i.e., cortical bone) and suppresses irregular (i.e., soft tissue) structures in 3D medical images by Eigen-analysis of the local Hessian matrix. Specifically, let $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$ be the eigenvalues of the local Hessian matrix computed at voxel \mathbf{p} of a Gaussian-smoothed (with standard deviation σ) CT image. Setting $R_{\text{sheet}} = |\lambda_2|/|\lambda_3|$, $R_{\text{blob}} = |(2|\lambda_3| - |\lambda_2| - |\lambda_1|)/|\lambda_3|$, and $R_{\text{noise}} = \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}$, the OSM at voxel \mathbf{p} , denoted by $\mathbf{S}(\mathbf{p}) \in \mathbb{R}^3$, computes as

$$\mathbf{S}(\mathbf{p}) = \mathbf{v}_3 \max_{\sigma \in \Sigma} S(\sigma), \quad (6.2)$$

$$S(\sigma) = \begin{cases} 0 & \text{if } \lambda_3 > 0 \\ \exp\left(-\frac{R_{\text{sheet}}^2}{2\alpha^2}\right) \left(1 - \exp\left(-\frac{R_{\text{blob}}^2}{2\beta^2}\right)\right) \left(1 - \exp\left(-\frac{R_{\text{noise}}^2}{2\gamma^2}\right)\right) & \text{otherwise,} \end{cases}$$

where \mathbf{v}_3 is the eigenvector corresponding to λ_3 and Σ is a set of scales. Note that \mathbf{v}_3 corresponds to the bone surface normal of voxels lying on the cortical bone layer. The parameters α, β and γ are used to balance the relative contributions of the factors in $S(\sigma)$. Figure 6.4, A shows an example slice of an upper arm OSM image, computed from CT.

Ray features were originally proposed for the detection of neuron nuclei in microscopy images (Smith et al., 2009). Given a voxel \mathbf{p} , ray features test some property of the closest edge location with respect to a direction \mathbf{r} . Due to this fact, ray features can consider image characteristics at distant contour points, permitting to efficiently recognize irregularly shaped objects based on the edge image extracted using a Sobel filter (Smith et al., 2009). However, our preliminary experiments showed that the edge image does not sufficiently account for the variability in cortical layer thickness observed in our dataset. Nevertheless, ray features can be employed to learn the geometry of the proximal humerus by replacing the edge image used in Smith et al. (2009) with the OSM image. To do so, the distance $d_e(\mathbf{p}, \mathbf{r})$ between \mathbf{p} and the closest point on the cortical bone layer in direction \mathbf{r} ($\|\mathbf{r}\|_2 = 1$) is calculated by determining the largest OSM magnitude along the ray described by \mathbf{p} and \mathbf{r} :

$$d_e(\mathbf{p}, \mathbf{r}) = \arg \max_{d \in [0, d_{\max}]} \|\mathbf{S}(\mathbf{p} + d\mathbf{r})\|_2, \quad (6.3)$$

where d_{\max} is the maximum admissible distance. Three different types of OSM ray features have been developed:

$$f_{\text{dist}}(\mathbf{p}, \mathbf{r}) = d_e(\mathbf{p}, \mathbf{r}), \quad (6.4)$$

$$f_{\text{angle}}(\mathbf{p}, \mathbf{r}) = \arccos \left(\left| \left\langle \frac{\mathbf{S}(\mathbf{p} + d_e(\mathbf{p}, \mathbf{r})\mathbf{r})}{\|\mathbf{S}(\mathbf{p} + d_e(\mathbf{p}, \mathbf{r})\mathbf{r})\|_2}, \mathbf{r} \right\rangle \right| \right), \quad (6.5)$$

$$f_{\text{int}}(\mathbf{p}, \mathbf{r}) = \|\mathbf{S}(\mathbf{p} + d_e(\mathbf{p}, \mathbf{r})\mathbf{r})\|_2. \quad (6.6)$$

The ray feature types given by Equations (6.4) to (6.6) are illustrated in Figure 6.4, B. Note that the Figure demonstrates the ray features for the 2D case, although they are calculated in 3D. Evaluating the latter two feature types for different \mathbf{r} , permits to localize \mathbf{p} relative to the cortical layer of the closest bone (not necessarily the humerus). The absolute value in the inner product of Equation (6.5) is necessary because the sign of $\mathbf{S}(\mathbf{p} + d_e(\mathbf{p}, \mathbf{r})\mathbf{r})$ gives no indication whether the ray points to the bone interior or exterior. The value of $f_{\text{int}}(\mathbf{p}, \mathbf{r})$ depends on the thickness of the closest cortical bone layer, thereby incorporating information about which bone segment is detected. By doing so, the thin cortical layer of the humeral head can be better differentiated from the thicker cortical layers of the humeral shaft and scapula.

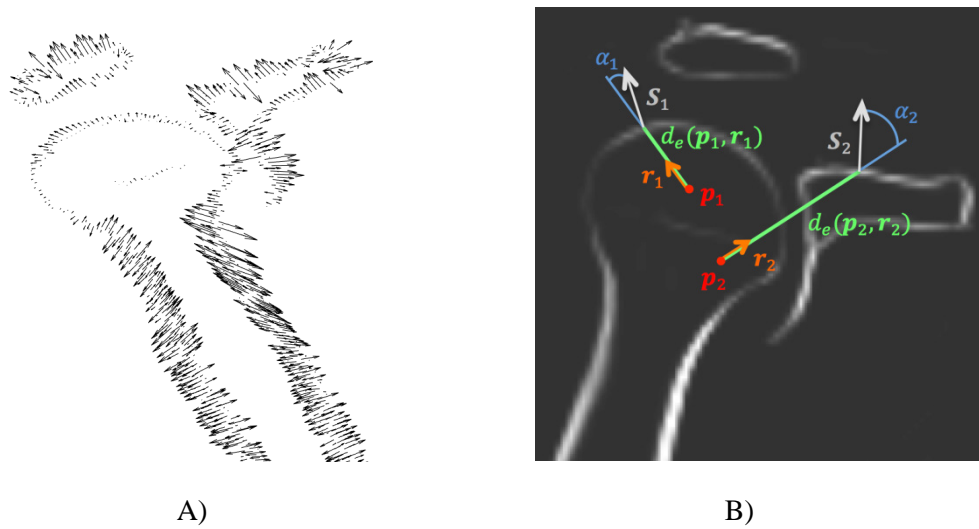


Figure 6.4: A) Oblique view of a coronal slice of the upper arm OSM image showing humerus and scapula. B) OSM-magnitude image of the same slice. Two examples of ray features are shown for the 2D case. For the feature at point \mathbf{p}_i , \mathbf{S}_i denotes the corresponding closest point on the cortical bone layer in direction \mathbf{r}_i . The functions $f_{\text{dist}}(\mathbf{p}_i, \mathbf{r}_i)$, $f_{\text{angle}}(\mathbf{r}_i, \mathbf{p}_i)$, and $f_{\text{int}}(\mathbf{p}_i, \mathbf{r}_i)$ of Equations (6.4) to (6.6) are denoted by $d_e(\mathbf{p}_i, \mathbf{r}_i)$, α_i , and $\|\mathbf{S}_i\|_2$, respectively. Note that the feature at \mathbf{p}_2 corresponds to the cortical layer of the scapula (instead of the humerus).

6.2.4 Computing a coarse estimate of the AMP parameters

Adapting ideas from Criminisi et al. (2011) and Fanelli (2011), we employ the theoretical RF framework described in Section 6.2.1 as follows. In each training CT image, $N = 2 \cdot 10^4$ voxels $\{\mathbf{p}_i\}_{i=1}^N$ are sampled on a regular grid covering the entire image. T regression trees are trained to predict the displacements $\mathbf{d}_i = \mathbf{c} - \mathbf{p}_i$, indicating the position of the AMP parameter \mathbf{c} relative to \mathbf{p}_i , as well as φ, θ . Hence, the desired regression response vectors are

given by $\mathbf{y}_i = [\mathbf{d}_i^T, \varphi, \theta]^T \in \mathbb{R}^5$. Following Criminisi et al. (2011), mean intensities over two displaced boxes are used as binary visual features, i.e., the feature value at voxel \mathbf{p}_i is computed as

$$f_{\text{box}}(\mathbf{p}_i, R_1, R_2, \boldsymbol{\delta}_1, \boldsymbol{\delta}_2) = \frac{1}{|R_1|} \sum_{\mathbf{q} \in R_1} I(\mathbf{p}_i + \boldsymbol{\delta}_1 + \mathbf{q}) - \frac{1}{|R_2|} \sum_{\mathbf{q} \in R_2} I(\mathbf{p}_i + \boldsymbol{\delta}_2 + \mathbf{q}), \quad (6.7)$$

where $I(\mathbf{p})$ is the image intensity at voxel \mathbf{p} ; R_1, R_2 and $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2$ are regions and offsets, respectively, centred around \mathbf{p} (Figure 6.5). In addition to these features, we use a unary variant of f_{box} , i.e., f_{box} -type features for which the second term on the right-hand side of Equation (6.7) is omitted. Note that f_{box} -type features can efficiently be computed using integral images (Viola and Jones, 2004).

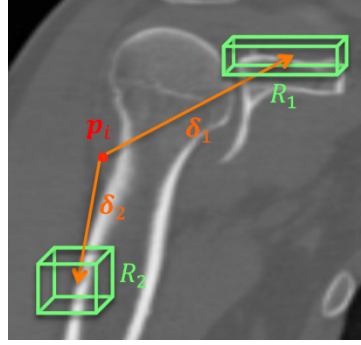


Figure 6.5: Illustration of a binary f_{box} -type feature around voxel \mathbf{p}_i with corresponding regions and offsets, denoted by R_1, R_2 and $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2$, respectively.

Two-dimensional features $\mathbf{x}_{i,j}$ as used in our RF framework are constructed by stacking two binary/unary f_{box} -type features with different R_1, R_2 and $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2$ into a vector, i.e., for binary features $\mathbf{x}_{i,j} = [f_{\text{box}}(\mathbf{p}_i, R_{1,j}, R_{2,j}, \boldsymbol{\delta}_{1,j}, \boldsymbol{\delta}_{2,j}), f_{\text{box}}(\mathbf{p}_i, R'_{1,j}, R'_{2,j}, \boldsymbol{\delta}'_{1,j}, \boldsymbol{\delta}'_{2,j})]^T$. For each tree, a different feature set with 1024 features is generated by choosing offsets $\boldsymbol{\delta}_{1,j}, \boldsymbol{\delta}_{2,j}$ and the dimensions of regions $R_{1,j}, R_{2,j}$ for every feature uniformly at random. Following Fanelli (2011), we assume that \mathbf{d}_i and φ, θ are uncorrelated, i.e., the covariance matrix has a block-diagonal structure

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_d & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_a \end{bmatrix}, \quad (6.8)$$

where $\boldsymbol{\Sigma}_d \in \mathbb{R}^{3 \times 3}$ is the covariance matrix of the displacements and $\boldsymbol{\Sigma}_a \in \mathbb{R}^{2 \times 2}$ denotes the covariance matrix of the angulation of the AMP normal.

Similar to the training phase, a coarse estimate $\hat{\mathbf{v}}_C$ of the parameter vector \mathbf{v} (cf. Section 6.2.2) is obtained from a test CT image by first sampling N test voxels $\{\mathbf{p}'_i\}_{i=1}^N$ on a regular grid. Next, each tree t is evaluated for every \mathbf{p}'_i . Let $l(\mathbf{p}'_i, t)$ be the leaf node corresponding to \mathbf{p}'_i when evaluating t . The estimate $\hat{\mathbf{v}}_{i,t}$ then computes as $\hat{\mathbf{v}}_{i,t} = [\mathbf{p}'_i{}^T, 0, 0]^T + \boldsymbol{\mu}_{l(\mathbf{p}'_i, t)}$ where $\boldsymbol{\mu}_l$ denotes the mean at leaf l (cf. Section 6.2.1). The coarse estimate $\hat{\mathbf{v}}_C$ can then be computed as the average over the fraction ε of those $\hat{\mathbf{v}}_{i,t}$ with the lowest corresponding total variance

$\text{trace}(\mathbf{\Sigma}_{i,t})$ at leaf $l(\mathbf{p}'_i, t)$. Thereby, ε is a user-defined threshold. Averaging only over a low variance set of leaf nodes instead of all leaf nodes reduces the error (Criminisi et al., 2011; Fanelli et al., 2011).

6.2.5 Refinement step

Two additional RFs F_{R_k} , $k = 1, 2$, are trained to iteratively compute refined AMP parameter estimates $\hat{\mathbf{v}}_{R_k}$ according to $\hat{\mathbf{v}}_{R_k} = \hat{\mathbf{v}}_{R_{k-1}} + \Delta \mathbf{v}_k$ where $\Delta \mathbf{v}_k$ is a correction computed by F_{R_k} . The refinement step is initialized with the coarse estimate $\hat{\mathbf{v}}_C$, i.e., $\hat{\mathbf{v}}_{R_0} = \hat{\mathbf{v}}_C$. To train the RFs, we sample angular offsets $\{\Delta \varphi_m\}$ and $\{\Delta \theta_n\}$ in centred intervals as well as voxels $\{\mathbf{p}_q\}$ in a cubic neighborhood around \mathbf{c} , and accordingly set the desired responses $\mathbf{y}_{(q,m,n)} = [\Delta \mathbf{c}_q^T, \Delta \varphi_m, \Delta \theta_n]^T$, where $\Delta \mathbf{c}_q = \mathbf{c} - \mathbf{p}_q$. More specifically, to train F_{R_1} we sample 300 voxels from a $45 \times 45 \times 45$ mm neighbourhood around \mathbf{c} as well as a total of 150 angular increments in the intervals $[-45, 45]$ and $[-25, 25]$ for $\Delta \varphi_m$ and $\Delta \theta_n$, respectively. For the training of F_{R_2} we consider 100 voxels in a $20 \times 20 \times 20$ mm neighborhood around \mathbf{c} and a total of 300 angular increments in the intervals $[-30, 30]$ for $\Delta \varphi_m$ as well as $[-15, 15]$ for $\Delta \theta_n$. Reducing the neighborhood size and interval lengths for training F_{R_2} results in an increased output resolution and hence allows for finer-grained corrections.

In contrast to the first step, the RFs now rely on the sheetness-based ray features introduced in Section 6.2.3. For the OSM, scales from $\sigma \in [0.8, 3.2]$ mm are considered using parameter values of $\alpha = 0.5$, $\beta = 0.5$, and $\gamma = 700$. Before computation of the ray feature values, the following preprocessing steps were performed: The CT images were resampled to an isotropic resolution equal to the in-plane CT-resolution as this is required for the computation of the OSM. Further, the patient skin was removed from the CT image to suppress unwanted strong sheetness response at the skin-air interface. To do so, the image background was dilated before setting the sheetness value in the dilated volume to zero.

To learn the correction for the estimate of the AMP normal (i.e., $\Delta \varphi$ and $\Delta \theta$) from the training CT data, the orientation of the ray features is first aligned with the orientation of the (ground truth) AMP normal, and subsequently perturbed by applying the angular offsets $\{\Delta \varphi_m\}$ and $\{\Delta \theta_n\}$ as follows. Let g be the type of the ray feature as defined in Equations (6.4) to (6.6), i.e., $g \in \{\text{dist}, \text{angle}, \text{int}\}$. For a ray feature f_g with direction \mathbf{r} , the feature value corresponding to $\mathbf{y}_{(q,m,n)}$ is then given by $f_g(\mathbf{p}_q, \mathbf{R}_z(\varphi - \Delta \varphi_m) \mathbf{R}_y(\theta - \Delta \theta_n) \mathbf{r})$, where $\mathbf{R}_z(\alpha)$ and $\mathbf{R}_y(\alpha)$ are the rotations by angle α around the \mathbf{z} - and \mathbf{y} -axes, respectively, of the CT coordinate system. Alternatively, assuming ideal continuous training images for a moment, the same feature alignment can be achieved by first centring the auxiliary coordinate system $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ (c.f. Figure 6.3) in \mathbf{p}_q and then rotating each training image around the \mathbf{v} - and \mathbf{w} -axes such that the AMP normal vector coincides with the \mathbf{u} -axis. Subsequently, the image is rotated by $-\Delta \theta_n$ around the \mathbf{v} -axis and by $\Delta \varphi_m$ around the \mathbf{w} -axis, before computing the feature value which can be now expressed by $f_g(\mathbf{p}_q, \mathbf{r})$. Therefore, for ideal continuous images, the described alignment procedure is equivalent to augmenting the training dataset by rotated copies of training images. In the case of real-world discrete images with possibly anisotropic resolution the described equivalence only holds approximately due to interpolation artefacts incurred by rotating the images. However, rotating features rather than

images is advantageous from a computational point of view and accelerates tree training and testing. Aligning features with the pose of an object has been proposed in Dollár et al. (2010) in the context of pose estimation. We thus consider the sheetness-based ray features as pose indexed features in the sense of Dollár et al. (2010).

For each tree in both F_{R_1} and F_{R_2} , a set of 1800 random sheetness-based ray features is generated by choosing \mathbf{r} and g uniformly at random. Analogously as in Equation (6.7), two-dimensional features $\mathbf{x}_{i,j}$ are then obtained by stacking two ray features of the same type but with different \mathbf{r} into a vector

$$\mathbf{x}_{(q,m,n),j} = \left[f_{g_j}(\mathbf{p}_q, \mathbf{R}_z(\varphi - \Delta\varphi_m)\mathbf{R}_y(\theta - \Delta\theta_n)\mathbf{r}_j), f_{g_j}(\mathbf{p}_q, \mathbf{R}_z(\varphi - \Delta\varphi_m)\mathbf{R}_y(\theta - \Delta\theta_n)\mathbf{r}'_j) \right]^T.$$

To estimate $\Delta\mathbf{v}_1$ ($\Delta\mathbf{v}_2$) in a test CT image, an alignment of the ray features with the previously estimated AMP-normal is performed before F_{R_1} (F_{R_2}) is evaluated for 1000 (125) test voxels $\{\mathbf{p}'_q\}$ sampled on a regular grid in a $20 \times 20 \times 20$ mm ($10 \times 10 \times 10$ mm) neighborhood around the previously estimated centre \mathbf{c} . In more detail, the value of the ray feature at \mathbf{p}'_q with given \mathbf{r} and g is $f_g(\mathbf{p}'_q, \mathbf{R}_z(\hat{\varphi}_c)\mathbf{R}_y(\hat{\theta}_c)\mathbf{r})$, where $\hat{\varphi}_c$ and $\hat{\theta}_c$ denote the previous estimates of φ and θ , respectively. As in the first step, $\Delta\mathbf{v}_1$ ($\Delta\mathbf{v}_2$) is then computed as the average over the fraction ε of the entire forest's leaf nodes with the smallest corresponding total variance. In contrast to the training phase, a smaller neighbourhood is chosen for testing to ensure that all test voxels will be located in the volume considered during training.

6.3 Experiments

6.3.1 Dataset

A total of 72 cadaveric whole body CT scans were provided by the Institutes for Forensic Medicine of the Universities of Bern and Zurich, Switzerland. The data were acquired using a Siemens Emotion 6[®] and a Siemens Somatom Definition Flash[®] CT scanner, respectively, with in-plane resolutions between [0.92, 1.37] mm (mean: 1.22 mm) and an axial resolution between [0.50, 0.70] mm (mean: 0.55 mm).

Our dataset was generated from the full body scan by extracting a subset with a field of view similar to the one used clinically for upper-arm CT scans (i.e., $240 \times 240 \times 290$ mm) plus a margin of 50 mm on each side. To avoid a bias in the error estimates, the extracted region was not exactly centered around the proximal humerus head, but displaced by a Gaussian distributed perturbation with zero mean and covariance matrix $\sigma^2 \mathbf{I}_3$, $\sigma = 20/\sqrt{3}$ mm, where \mathbf{I}_3 is the 3×3 identity matrix.

Ground-truth data for the AMP parameters was generated as follows. 3D models of the humeri were created by applying an automated segmentation algorithm introduced in Gass et al. (2014) and loaded into the custom-made preoperative planning software CASPA (Balgrist CARD AG, Zurich, Switzerland). Thereafter, an expert annotated the proximal cartilage/metaphyseal boundary on each 3D model, by manually setting twelve reference points. Four points were placed on each the proximal, middle and distal part of the boundary, respectively. Lastly, the AMP was fitted to the twelve reference points by solving the overdetermined plane equation in a least-square sense, as described in Schneider and Eberly (2002). Additionally, a second expert annotated the AMP in 25% of the datasets to assess inter-reader variability.

6.3.2 Performance evaluation of the proposed method

All parts of the proposed method were implemented in MATLAB, except the filter for computing the OSM, which was implemented in C++ using ITK (Insight Segmentation and Registration Toolkit, Kitware, NY, United States).

The number of trees T was set to 15 (in both steps), and the tree depth D as well as the fraction of leaf nodes ε used to compute the output was estimated using 5-fold cross-validation. The same cross-validation sets were used for both steps of our algorithm to avoid overly optimistic results due to mixing of training and test sets. No registration or alignment of the CT images was performed before applying our method.

6.3.3 Refinement using f_{box} -type features

To justify the use of sheetness-based ray features in the refinement step we performed the following experiment. We trained and tested two RFs F'_{R_k} , $k = 1, 2$, for refinement that used f_{box} -type features (computed from unmodified CT images) instead of sheetness-based ray features, but otherwise had the same specifications as the RFs of the proposed method detailed in Section 6.2.5 (i.e., we considered an identical number of features, the same neighbourhoods around \mathbf{c} , and identical angular offsets). As in the refinement step of the proposed method, we set T to 15 and estimated the parameters D and ε via 5-fold cross-validation. The coarse estimate of the AMP parameters was obtained via the RF trained in the first step of the proposed method.

6.3.4 Comparison with atlas-based registration

We compared the proposed method with atlas-based registration, which may be considered as a baseline algorithm for estimating the AMP parameters. Specifically, we fitted a grey scale atlas image to the unmodified CT images in our data set by first performing registration under a similarity transformation and then refining the result by deformable B-Spline registration. We symmetrically cropped the CT images to $100 \times 100 \times 150$ mm volumes (relative to the volumes used for training the proposed method) as this improved the registration accuracy. To initialize the first registration step we aligned the center of mass of the atlas with the center of mass of the target image. In both registration steps a random sample of 5% of the voxels was used for similarity computation. B-Spline registration was performed at a single (maximum) resolution level and the grid employed consisted of $10 \times 10 \times 15$ subdivisions. Neither

increasing the sampling percentage nor increasing the number of subdivisions improved the registration accuracy.

The AMP parameters for a test image were estimated by applying the transformation resulting from the two-step registration to the twelve manually annotated ground truth reference points of the atlas image and by fitting a plane to the transformed points following the procedure described in Section 6.3.1.

We implemented the registration pipeline in Python as module for 3D Slicer¹ (Fedorov et al., 2012) and optimized the parameters of the two registration steps on the entire data set. Note that this biases the results in favour of atlas-based registration compared to the proposed method.

6.4 Results

We measured the position error as the Euclidean distance between \mathbf{c} and the estimate found by our method. The angular error was expressed by the angle between the ground-truth AMP normal vector and its estimate. In Table 6.1, position and angular errors are given after coarse estimation (step 1) and refinement (step 2).

	Coarse estimation			Refinement		
Position error (mm)	7.77	\pm 3.13	(7.78)	2.40	\pm 1.20	(2.13)
Angular error (°)	10.19	\pm 5.75	(9.22)	6.51	\pm 3.43	(5.90)

Table 6.1: Position and angular error of the two steps of our method: mean \pm standard deviation (median)

The optimal parameters found for step 1 of our algorithm were $D = 9$, $\varepsilon = 0.04$, and in step 2 the best parameters were $D = 12$ and $\varepsilon = 0.52$ for F_{R_1} as well as $D = 12$ and $\varepsilon = 0.78$ for F_{R_2} . As part of the parameter selection procedure we investigated the robustness of the estimation accuracy against variation of T and ε . When fixing $D = 9$ and $\varepsilon = 0.04$ in step 1 of the algorithm, the position error varied less than 0.5 mm and the angular error less than 0.15° for $5 \leq T \leq 20$. When fixing $D = 9$ and $T = 15$, the position and angulation errors varied less than 2 mm and 0.25° , respectively, for $0.02 \leq \varepsilon \leq 1.0$.

In step 2, when fixing $D = 12$ and $\varepsilon = 0.52$ ($\varepsilon = 0.78$), the output variation of F_{R_1} (F_{R_2}) with respect to the position error was below 0.1 mm and with respect to the angular error below 0.1° for $5 \leq T \leq 20$. When fixing $D = 12$ and $T = 15$, position and angular errors incurred by F_{R_1} (F_{R_2}) varied less than 0.4 mm (0.3 mm) and 0.6° (0.4°), respectively, for $0.25 \leq \varepsilon \leq 1.00$. Figure 6.6 shows the position error vectors after step 1, after evaluation of F_{R_1} , and after evaluation of F_{R_2} , for all cross-validation sets. The distribution of the error vectors is isotropic

¹ <http://www.slicer.org>

with a minor shift along the z -axis. In Figure 6.7 we plot normalized histograms of the position and angular errors. Two examples of humeri together with the AMP estimated by our method and corresponding ground truth are visualized in Figure 6.8.

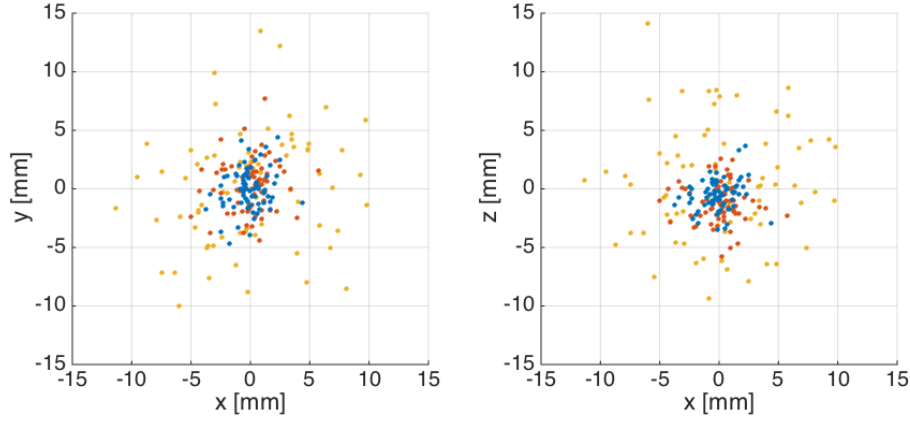


Figure 6.6: Projection of the position error vectors to the x - y and the x - z planes of the CT coordinate system after step 1 (yellow), after evaluation of F_{R_1} (red), and after evaluation of F_{R_2} (blue).

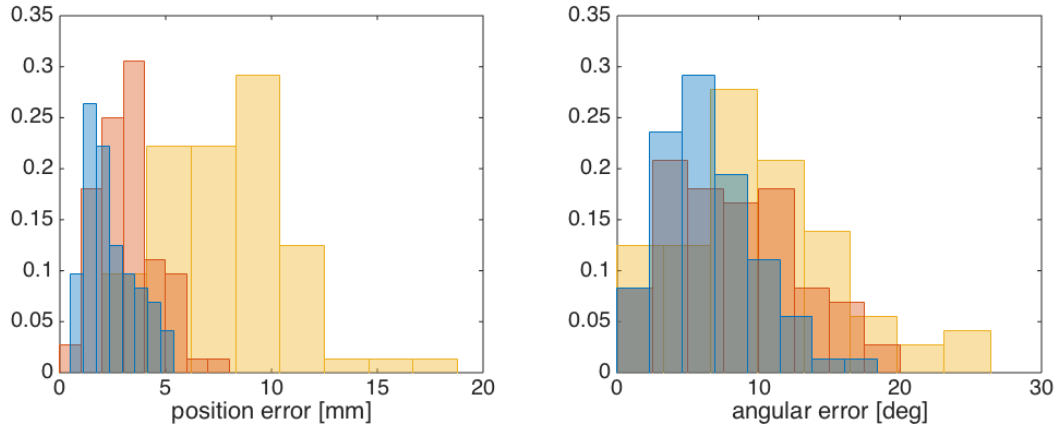


Figure 6.7: Error histograms after step 1 (yellow), after evaluation of F_{R_1} (red), and after evaluation of F_{R_2} (blue).

Training of the RFs took 12 h for step 1, and 69 h for step 2. All trees were trained in parallel using a grid-computing infrastructure. Calculating the AMP for a test case took 35 s on average (MacBook Pro with 2.5 GHz Core i7 (I7-4870HQ) CPU and 16 GB RAM). Note that the processing time can be further reduced by implementing the RF framework in a compiled language.

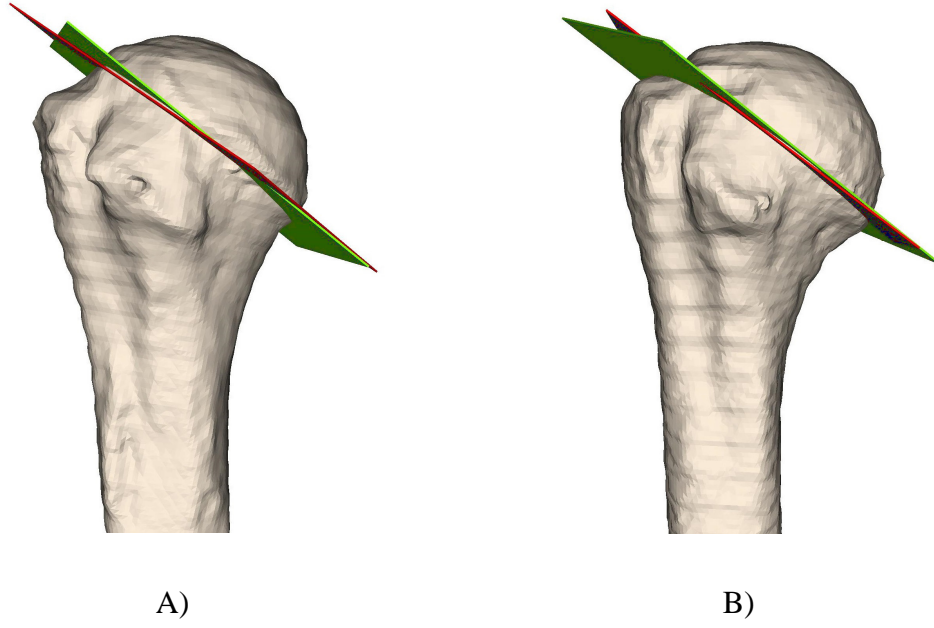


Figure 6.8: Visualization of the AMP estimated by the proposed algorithm (green plane) compared to the (ground truth) AMP (red plane) for two cases. a) Position error: 1.9 mm; angular error: 4.5°. b) Position error: 2.6 mm; angular error of 6.2°.

Assessment of the inter-reader variability of the manual annotations showed an average difference between the AMPs of the readers of 3.4 mm (standard deviation: 1.4 mm, median: 3.6 mm, range: 1 mm to 6.34 mm) and 4.0° (standard deviation: 2.09°, median: 4.04°, range: 0.56° to 7.36°). Here, the position and angular errors were computed as the Euclidean distance between the AMP centers and the angle between the AMP normals, respectively.

6.4.1 Refinement using f_{box} -type features

Table 6.2 shows the position and angular errors after refinement via a F'_{R_k} , $k = 1, 2$, which were based on f_{box} -type features. The optimal tree parameters for F'_{R_1} (F'_{R_2}) were $D = 12$ ($D = 12$) and $\varepsilon = 0.62$ ($\varepsilon = 0.18$).

	Refinement (f_{box} -type features)		
Position error (mm)	3.63	± 2.26	(3.26)
Angular error (°)	8.65	± 5.13	(7.76)

Table 6.2: Position and angular error after refinement with f_{box} -type features: mean \pm standard deviation (median)

6.4.2 Comparison with atlas-based registration

In Table 6.3 we report the position and angular error after registration under a similarity transform and after B-Spline registration. The similarity measure that yielded the best performance was Mattes Mutual Information for both registration steps. On average, the two registration steps took 77 s (MacBook Pro with 2.5 GHz Core i7 (I7-4870HQ) CPU and 16 GB RAM).

	Similarity transformation			B-Spline		
Position error (mm)	9.36	\pm 4.44	(7.79)	7.94	\pm 4.44	(6.80)
Angular error ($^{\circ}$)	11.67	\pm 8.01	(10.46)	13.12	\pm 7.71	(11.28)

Table 6.3: Position and angular error after atlas-based registration under a similarity transformation and after B-Spline registration: mean \pm standard deviation (median)

6.5 Discussion

In this work, an algorithm was described for estimating the AMP in full-size upper arm CT scans in an automated fashion. The stepwise approach, comprising coarse estimation followed by refinement, has proven to be effective, i.e., the average position error of 7.77 mm and angular error of 10.19° achieved after step 1, decreased in step 2 by 69.1% and 36.1% for position and angulation, respectively. This observation is well aligned with the intuition that the local neighbourhood of the humeral head is more informative for estimating the AMP parameters than regions further apart (e.g., the AMP parameters of two distinct subjects with aligned scapulae may differ vastly and relying on the image region of the scapula to estimate the AMP parameters may thus yield poor performance) and proceeding in two steps may hence increase the robustness of the method. Furthermore, the results indicate that the novel sheetness-based ray features are, indeed, suited for detecting local bone geometry. Dividing the AMP estimation into two steps is also effective for considerably reducing the computational burden. By first narrowing the region of interest using simple computationally inexpensive features before calculating computationally more demanding sheetness-based ray features, permits to considerably reduce the overall processing time, which is an advantage of the proposed approach.

We note that the concept of iterative locally refined estimation via RFs was considered before, e.g., in Dollár et al. (2010) in the context of pose estimation and in Ebner et al. (2014), Gao and Shen (2014) and Han et al. (2015) for landmark localization in medical images. However, all of these works use the same feature types for coarse localization and refinement. To train the RFs for refinement, further note that instead of sampling voxels and angular increments around the ground truth AMP parameters, one can use the responses of the preceding RF obtained for its training voxels - see, e.g., (Dollár et al., 2010). We implemented this variant of the algorithm as part of preliminary experiments and observed lower accuracy than for the proposed method.

Comparing the accuracy obtained by refinement via sheetness-based ray features with that achieved by refinement via f_{box} -type features, we found that the former yielded considerably higher accuracy, both in terms of position and angular error. This indicates that the sheetness-

based ray features are indeed better suited for learning the local bone geometry of the humeral head than f_{box} -type features. A limitation of the proposed method is that computing the OSM necessitates resampling the images to isotropic resolution, which may introduce a bias due to interpolation artefacts. While distribution of the error vectors was isotropic, we observed a minor bias along the z-axis.

Comparison of the proposed method with atlas-based registration showed that the accuracy in terms of position error obtained after B-Spline registration is comparable to that observed for the proposed method after step 1. The angular error after both registration steps was larger than the angular error incurred by step 1 of the proposed method, and the B-Spline registration step even increased the angular error.

The experiments also showed that the proposed algorithm achieves higher accuracy with respect to the AMP position than manual annotation by different experts. Further, the evaluation of the parameters showed insensitivity of our method with respect to the choice of the T and ε . The angular error in estimating the AMP orientation using the algorithm is larger compared to the variability between the readers.

We next briefly discuss the accuracy of the proposed method with regard to clinical applicability. Restoration of the normal humeral anatomy is an ultimate goal in reconstructive surgeries of the proximal humerus, such as fracture reduction or prosthetic replacement (Boileau et al., 2008; Hertel et al., 2002; Iannotti et al., 1992). It is well accepted that initial mal-positioning of the shoulder prosthesis, i.e. for proximal humeral fractures, with excessive height and retrotorsion is associated with a failure of the tuberosity osteosynthesis and poor functional outcome (Boileau et al., 2002). However, to the best of our knowledge the exact amount of acceptable deviance to the normal humeral anatomy in height and angular orientation of the proximal humeral component is not reported yet. While a biomechanical evaluation (Huffman et al., 2008) demonstrated that 10 mm or more inferior mal-positioning in height during hemiarthroplasty led to significant alterations in glenohumeral joint forces; smaller mal-positioning was not investigated. The same applies for the angular orientation of articulating components, where the maximum acceptable deviation of the individual anatomy for the reconstruction of the proximal humerus is not known so far. However, considering the fact that the contralateral, healthy humerus often serves as a reconstruction template for arthroplasty and fracture reduction planning, although considerable side-differences with respect to the humeral head orientation exist (mean bilateral differences of 7° , maximum differences up to 27°) (DeLude et al., 2007; Vlachopoulos et al., 2016a) the angular error of the purposed algorithm may nevertheless be clinically acceptable.

Moreover, the evaluation of our method was performed on cadaveric CT data which had an in-plane resolution being 2-3 times lower compared to clinical standards, because the entire body was scanned at once. Therefore, we expect our algorithm to yield higher accuracy in the clinical application.

Our method estimates the parameters of a 3D plane, in contrast to other algorithms proposed (Glocker et al., 2012; Kelm et al., 2013), where the simultaneous estimation of position and angulation of several vertebral bodies was targeted. The problem tackled in these studies is more complex with respect to the number of parameters but permits to use a probabilistic spine model for regularizing the RF output. In fact, the positions and orientations of

neighbouring vertebral bodies are strongly correlated. Therefore, single erroneous estimates may be compensated by incorporating a global model. Similarly, in Donner et al. (2013) a parts based-model of the global landmark topology is employed to increase the robustness of simultaneous localization of multiple landmarks. A regularization of this type is difficult to realize in the problem considered here. There are landmarks on the proximal humerus surface, e.g., the groove between lesser and greater tuberosity, whose position could be detected and subsequently incorporated for estimating the AMP. However, these landmarks are also difficult to identify by experts in a consistent manner and, therefore, it is unclear whether including them would increase the estimation accuracy.

6.6 Conclusion and Future Work

We proposed and evaluated a two-step algorithm to estimate the AMP in CT images in a fully automated fashion. The algorithm already outperforms a human expert with respect to position error and has potential to achieve the accuracy of a human expert in terms orientation error when further refined. Future work includes the enhancement of sheetness-based ray features to provide better orientation information. It will be interesting to explore the behaviour of ray features in combination with other feature types.

The proposed approach is the first automated one providing a good estimate of the AMP. Therefore, the time for preoperative arthroplasty planning would be significantly reduced compared to manual annotation. In our opinion, the proposed supervised learning approach has the potential to automate time-consuming planning tasks which, up to now, have required a clinically experienced user. The adaption of the framework to other anatomies will hence be another promising future direction to pursue.

7

Reconstruction of Complex Proximal Humeral Fractures

In this paper, we present an algorithm for the reconstruction of complex proximal humeral fractures based on a novel scale-space representation of the curvature. The optimal surgical treatment of complex fractures of the proximal humerus is controversial. It is proven that best results are obtained if an anatomical reduction of the fragments is achieved and, therefore, computer-assisted methods have been proposed for the reconstruction of the fractures. However, complex fractures of the proximal humerus are commonly accompanied with a relevant displacement of the fragments and, therefore, algorithms relying on the initial position of the fragments might fail. The state-of-the-art algorithm for complex fractures of the proximal humerus requires the acquisition of a CT scan of the (healthy) contralateral anatomy as a reconstruction template to address the displacement of the fragments. Pose-invariant fracture line based reconstruction algorithms have been applied successfully for reassembling broken vessels in archaeology. Nevertheless, the extraction of the fracture lines and the necessary computation of their curvature are susceptible to noise and make the application of previous approaches difficult for the reconstruction of bone fractures close to the joints, where the cortical layer is thin. We present a novel scale-space representation of the curvature, permitting to calculate the correct alignment between bone fragments solely based on corresponding regions of the fracture lines. The fractures of the proximal humerus are automatically reconstructed based on iterative pairwise reduction of the fragments and automatic selection of the best solution. The validation of the presented method was performed on twelve clinical cases, surgically treated after complex proximal humeral fracture, and by cadaver experiments. The accuracy of our approach was compared to the state-of-the-art algorithm.

7.1 Introduction

The treatment of comminuted fractures of the proximal humerus is challenging and the optimal procedure remains controversial (Cvetanovich et al., 2016; Gerber et al., 2004). Open reduction and internal fixation using conventional or locking plates is the mainstay of therapy for the young and active patient (Gerber et al., 2004; Grubhofer et al., 2016), while best results are obtained if anatomical or near anatomical reduction can be achieved (Gerber et al., 2004). Anatomical reduction is a pre-requisite for a joint-preserving surgical treatment of a fractured proximal humerus. If anatomical reduction cannot be obtained, joint replacement has to be considered (Gerber et al., 1998). The options for replacement surgery of the

shoulder joint include hemiarthroplasty, anatomic total shoulder arthroplasty and reverse total shoulder arthroplasty (RTSA) (Cuff and Pupello, 2013; Cvetanovich et al., 2016; Fucentese et al., 2014; Grubhofer et al., 2016) with a current trend from hemiarthroplasty towards RTSA for complex humeral fractures in the elderly (Cvetanovich et al., 2016; Grubhofer et al., 2016). The main reason of this current trend is that, despite promising initial reports of the hemiarthroplasty (Neer, 1970), less satisfactory or even disappointing results have been reported (Shukla et al., 2016). Current literature suggests that RTSA might result in better clinical outcomes than hemiarthroplasty, due to the decreased reliance on tuberosity healing of the RTSA (Shukla et al., 2016). Nevertheless, the most important consensus across all surgical treatment options is, that the functional outcome is better with anatomical fixation of the tuberosities (Anakwenze et al., 2014; Boileau et al., 2002; Fucentese et al., 2014; Gallinet et al., 2009; Gerber et al., 2004; Grubhofer et al., 2016; Huffman et al., 2008). Therefore, it seems clearly justified that major effort should be made to achieve an anatomical reduction of the tuberosities.

The benefits of computer-assisted preoperative simulation and intraoperative navigation is well accepted in joint replacement surgery (Iannotti et al., 2014; Levy et al., 2014; Nguyen et al., 2009) and for corrective osteotomies after malunited fractures of the humerus (Murase et al., 2008; Vlachopoulos et al., 2016b). Computer-assisted approaches are promising, especially, since it is difficult or even impossible to preoperatively plan the orthopaedic procedure using only radiographs or computed tomography (CT) analysis (Vlachopoulos et al., 2016a). In case of a complex fracture of the proximal humerus, the ultimate goal of these approaches should be the restoration of the normal humeral anatomy. However, the fundamental pre-requisite to apply computer-assisted navigation in the surgery to fractures of the proximal humerus is the preoperative reconstruction of the fractures (Fürnstahl et al., 2012). Hitherto, one method (Fürnstahl et al., 2012) has been published and validated for the computer-assisted reconstruction of complex proximal humerus fractures (Jimenez-Delgado et al., 2016). Fürnstahl et al. (2012) demonstrated that their algorithm allows accurate reconstruction of the pretraumatic anatomy. However, the main drawback of the method of Fürnstahl et al. (2012) is the dependency on the healthy contralateral bone model as a reconstruction template. A further computer-assisted method for the treatment of complex proximal humeral fractures via hemiarthroplasty was developed and validated by cadaver experiments (Bicknell et al., 2007). The alignment of the shoulder prostheses and the tuberosity fragment was assessed by manual measurements of characteristic landmarks. Here, also the contralateral anatomy was proposed as a reconstruction template for the use in a clinical setting. However, existing bilateral differences in the humeral anatomy (DeLude et al., 2007; Vlachopoulos et al., 2016a) or the presence of a pathological altered contralateral anatomy (e.g., after a proximal humeral fracture or a joint replacement surgery) might limit the clinical application of both methods.

The task to be performed is similar to the assembly of a jigsaw puzzle as illustrated in Figure 7.1, also introducing the terminology used throughout this paper. Pose-invariant reconstruction algorithms have been successfully developed in archaeology for the reassembling task of broken vessels and relicts (Papaioannou and Theoharis, 2003; Üçoluk and Toroslu, 1999). The geometric reconstruction was performed by matching individual fracture surfaces using 3D curvature matching methods. The presented method builds on this idea from the approaches in archaeology. However, as the data acquisition scans in clinical

practice is based on CT, the noise is greater than in archaeology, where laser scanning is used. Clinical data are characterized by a limited resolution (in-plane and axial resolution of 0.4 mm or worse) (Lecouvet et al., 2008), resulting in partial volume effects and diffuse fracture lines, in contrast to laser-scanned data with an isotropic resolution of 0.05 mm or less. Furthermore, bone fracture surfaces tend to be highly irregular (Thomas et al., 2011). Therefore, the adoption of the archaeological approaches (Papaioannou and Theoharis, 2003; Üçoluk and Toroslu, 1999) for clinical application, i.e., bone fracture reconstruction, is not straightforward (Thomas et al., 2011).

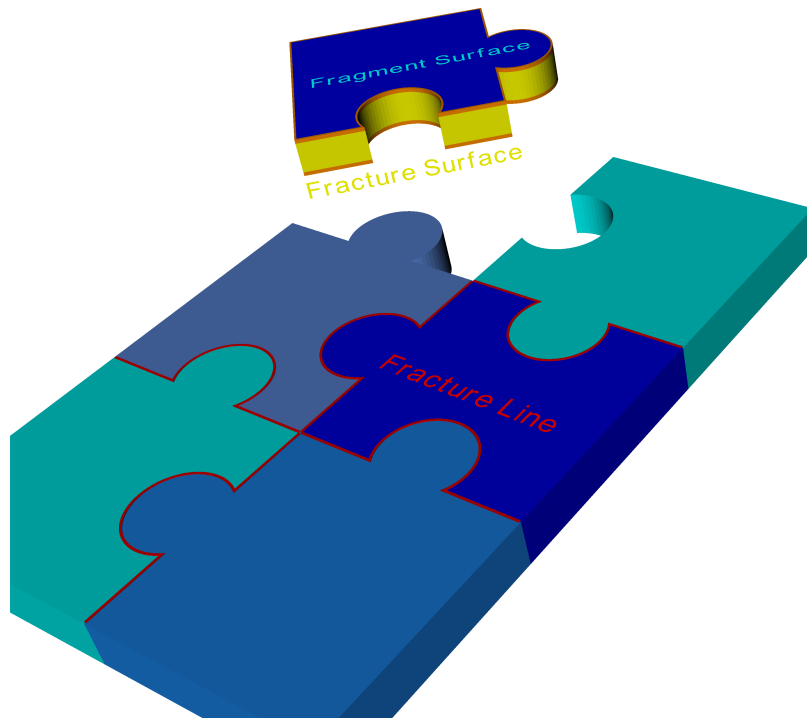


Figure 7.1: Terminology for the present paper illustrated on a jigsaw puzzle. The fracture surfaces (yellow surface) represent the break through the cortex of the proximal humeral fragments and are simplified by fracture lines (red lines). The fragment surface (blue surface) corresponds to the unfractured outer cortical layer of the fragment.

In this paper, we present a novel method for the fully automated reconstruction of proximal humeral fractures, requiring only the information of the fracture surfaces. The key idea is to use a scale-space representation of the curvature of the corresponding fracture lines, which permits determining the correct alignment between fragments. The proposed method was evaluated clinically on a consecutive series of patients treated with proximal humerus fractures and on four artificially created fractures on cadaveric humeri. In addition, we compared our reconstruction results with the current state-of-the-art algorithm (Fürnstahl et al., 2012; Jimenez-Delgado et al., 2016)

In the following, we will give a brief overview of computer-assisted techniques for the simulation of fracture reduction. In Section 7.2, an overview is presented and the details of our approach are described. The clinical evaluation and the results of cadaver experiments are

presented in Section 7.3. We discuss the method in Section 7.4. Finally, Section 7.5 summarizes the major conclusions of the work.

7.1.1 Related Work

An accurate preoperative assessment of fragment displacement is crucial for a successful restoration of a fracture (Fürnstahl et al., 2012). However, the literature regarding computer-assisted reconstruction of bone fractures is relatively sparse, compared to the very large body of research on the topics of bone segmentation and medical image registration (Thomas et al., 2011). Jiménez-Delgado et al. (2016) recently published a comprehensive review article, summarizing current approaches for bone fracture reduction planning (Albrecht and Vetter, 2012; Chowdhury et al., 2009; Fürnstahl et al., 2012; Moghari and Abolmaesumi, 2008; Okada et al., 2009; Thomas et al., 2011; Willis et al., 2007; Winkelbach and Wahl, 2008; Zhou et al., 2009).

Some of these approaches rely on a reconstruction template, i.e. the contralateral bone model (Bicknell et al., 2007; Fürnstahl et al., 2012; Okada et al., 2009) or a statistical shape model (Albrecht and Vetter, 2012; Moghari and Abolmaesumi, 2008) is required to calculate the reduction. DeLude et al. (2007) and Vlachopoulos et al. (2016a) verified that the contralateral humeral anatomy might be a reliable template for some geometric characteristics (i.e., the humeral head size and the humeral length). However, due to the presence of intra-individual differences, in particular, differences in axial torsion, Vlachopoulos et al. (2016a) concluded that preoperative planning approaches, targeting the reconstruction of complex proximal humerus fractures should not rely blindly on the contralateral anatomy.

Other approaches align the fragments based on the characteristics of the fracture surfaces (Chowdhury et al., 2009; Kronman and Joskowicz, 2013; Okada et al., 2009; Willis et al., 2007; Winkelbach et al., 2004; Zhou et al., 2009). Most of the presented approaches have in common that an Iterative Closest Point (ICP-) based algorithm is used to perform the reduction of the fragments (Jimenez-Delgado et al., 2016). The tendency of the ICP to fall into local minima might be particularly problematic in case of the proximal humerus due to the almost spherical shape (Fürnstahl et al., 2012). Further methods have been developed, that combine the use of a reconstruction template and the fractured surfaces to tackle this problem (Fürnstahl et al., 2012; Okada et al., 2009). The contralateral anatomy is used, thereby, to generate a set of initial transformation close to the true parameter values. Furthermore, the fragments at the proximal humerus are often considerably displaced and malrotated and, consequently, approaches relying on the initial position of the fragments (Kronman and Joskowicz, 2013; Okada et al., 2009) can likely fail. Therefore, Fürnstahl et al. (2012), proposed to perform first a global but coarse pre-registration step that is independent of the initial pose of the fragments. Thereafter, a refinement step was applied using state-of-the-art local optimization techniques. However, their global pre-registration again required the use of the contralateral anatomy as a shape prior.

A fracture reconstruction algorithm that does not rely on the initial pose of the fragments would be an alternative solution to tackle the problematic of the considerably displaced fragments of a proximal humerus fracture. Similar methods were developed in archaeology for the reconstruction of broken vessels (Papaioannou and Theoharis, 2003; Üçoluk and Toroslu, 1999). These methods are based on the signature of a 3D curve, i.e., arc length, curvature and torsion, as introduced by Kishon and Wolfson (1987), and are closely related to our approach. Papaioannou and Theoharis (2003) calculated first pairwise reductions based on the fracture facets of the fragments and applied a facet boundary curve matching to reduce the

search space. The best assembly was obtained by determining the set of fragment combination that resulted in the smallest accumulative matching error. However, the main disadvantage of the method of Papaioannou and Theoharis (2003) is the requirement of nearly planar fracture facets. Furthermore, the limited resolution and noise of the clinical data make the application of the archaeological approaches (Papaioannou and Theoharis, 2003; Üçoluk and Toroslu, 1999) based on 3D curve matching difficult (Thomas et al., 2011).

7.2 Material and Methods

A consecutive series of eight patients, with a proximal humeral fracture treated with open reduction and internal fixation (ORIF), in the department of orthopaedics at the University Hospital Balgrist, Zurich, Switzerland, between January 2016 and July 2016, were used for the clinical evaluation. The CT scans were obtained according to a standard scanning protocol used for the preoperative evaluation of proximal humeral fractures (slice thickness 1 mm; 120 kV; Philips Brilliance 40 CT, Philips Healthcare, The Netherlands). The cantonal ethics committee of Zurich, Switzerland approved the study (KEK-ZH-Nr. 2013-0586). To compare our approach with the state-of-the-art method (Fürnstahl et al., 2012), we applied our reconstruction algorithm also to data previously published by Fürnstahl et al. (2012), i.e., four clinical cases and four artificially created fractures on cadaveric humeri. In total, we applied our method to sixteen complex fractures of the proximal humerus.

7.2.1 Overview of the Algorithm

The overall workflow of our method consists of three modules as illustrated in Figure 7.2. In the first module, we perform the segmentation task. The input of the first module is a CT scan of a proximal humeral fracture. Note, that the description of the segmentation task (with the annotation of the fracture surfaces) is used for the overview of the planning workflow and is not part of the developed method.

The input of the second module are the triangular surface models of the cortical layer of the fragments (7.2.2). The characteristic of the fragments used for the fracture reconstruction are analysed as described in 7.2.2.1. In a first step, the fractured surfaces of the fragments are converted to a different representation based on connected line segments, herein called fracture lines as described in 7.2.3. These fracture lines are the base for all subsequent steps of the fracture reconstruction algorithm. Thereafter, the scale-space representation of the curvature is calculated in a local neighbourhood for each point of a fracture line in Section 7.2.4. The pairwise reduction is performed based on the best matches of identified corresponding regions between the fragments by analysing the similarities of curvatures in scale-space (Section 7.2.5). The merged fracture lines after the pairwise reduction is calculated as described in Section 7.2.6. The scale-space curvature matching is repeated until all fragments are reduced. The set of fracture reconstructions produced from the second module (Section 7.2.7), are used as input for the third module of the algorithm. Finally, the algorithm determines the best solution based on all performed reconstructions as described in 7.2.7.1.

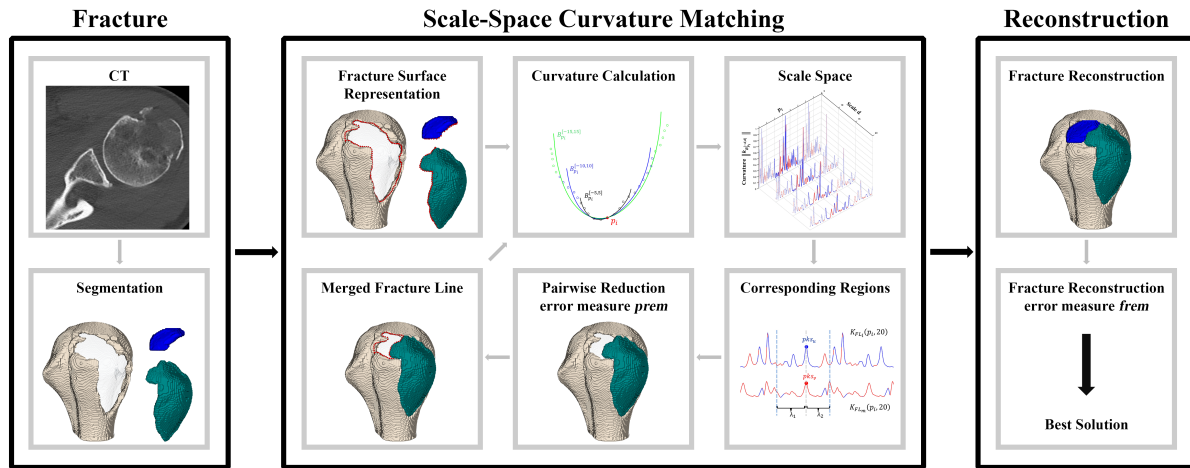


Figure 7.2: Overview of the reconstruction of a proximal humeral fracture.

7.2.2 Generation of triangular surface models

The proximal humeral anatomy is composed of an (outer) cortical layer and (internal) cancellous bone. The cortical layer is compact and dense and appears bright in a CT image. The cancellous bone is a porous structure, which is less dense than the cortical layer. Our reconstruction algorithm is based on surface models that represent only the cortical layer and not the cancellous bone of the humeral fragments.

For the clinical cases, the generation of the triangular surface models was performed by an experienced orthopaedic surgeon with clinically applied segmentation methods (Murase et al., 2008; Vlachopoulos et al., 2016b). Thereby, we used thresholding, manually correction of connected fragments, region growing, and the marching cubes algorithm (Lorensen and Cline, 1987).

For the cadaver experiments we used the triangular surface models provided by the authors of the study of Frnstahl et al. (2012). Frnstahl et al. (2012) used the bone enhancement filter of Descoteaux et al. (2006) for the segmentation of the cortical layer. Best results were achieved with a filter range of 3.5 mm, which was adapted to match the maximal cortical thickness of the proximal humerus. The average cortical thickness towards the glenohumeral joint was evaluated to be between 0.75 mm and 1 mm.

7.2.2.1 Fragment Characteristics

During the reconstruction of proximal humeral fractures by ORIF the larger fragments are reduced, while smaller fragments are left in place as local bone grafts to facilitate the healing process. Therefore, previous reconstruction algorithms automatically removed small fragments before the reconstruction. One measure of identifying small fragments is the area of the mesh triangles, i.e., all fragments with a summed area of the triangles smaller than 1000 mm² were small and not considered in the reconstruction in the study of Frnstahl et al. (2012), since contralateral matching did not robustly work for them.

In the current study, we first analysed the size of all fragments of the clinical cases that were classified by an orthopaedic surgeon and considered as being large enough for ORIF. To measure the size of a fragment, we calculated the oriented bounding box (OBB) from all model points. The OBB was previously described as a valuable method for the 3D measurement of the humeral length for clinical applications (Vlachopoulos et al., 2016a). The area of a fragment was approximated by the product of the two longest sides of the OBB. The characteristics of all thirty fragments, that were classified by the surgeon as large fragments and used for ORIF, are summarized in Table 7.1. The shaft fragments are not included, as the size is arbitrary and mainly depend on the scanned length of the upper arm during CT acquisition. The area of the smallest fragment of these thirty fragments was 195 mm². Therefore, we used for the fracture reconstruction of the clinical cases and of the cadaver experiments all fragments with an area greater than $area_{min} = 195 \text{ mm}^2$.

	Area (mm ²)		
	Mean	SD	Range
Head (7)	2320	682	1812 - 3319
Intermediate (23)	880	520	195 - 1954
All fragments (30)	1216	828	195 - 3319

Table 7.1: Characteristics of the fragments of the clinical cases. The area of the fragments is summarized, which were considered by the surgeon as relevant during fracture reconstruction. The intermediate fragments are all fragments excluding the head fragments.

7.2.3 Fracture Surface Representation

The thickness of the cortical layer is about 4 mm in the proximal shaft region and decreases towards the joints (Fürnstahl, 2010; Skedros et al., 2016). As a consequence, the fracture surfaces are extremely narrow, and can be considered as a narrow path. Therefore, our idea was to represent the fracture surfaces by a fracture line along the path.

Formally, a fracture surface FS of a given fragment is simplified by a fracture line FL, i.e., a sequence of points $p_i \in FS$. The annotation of the points p_i on the triangular surface model was performed manually by setting densely sampled points as illustrated by spheres in Figure 7.3. The points were projected on the fracture surface and centred between the inner and outer contour of the cortical layer. To ensure equally distanced points, the algorithm performed a cubic spline interpolation. The distance between the interpolated points was selected to be $r_{FL} = 1 \text{ mm}$. The order of the point sequence specifies the direction of the path. Therefore for each fragment, the fracture surface was represented by a sequence of points in clock-wise and counter-clock wise direction.

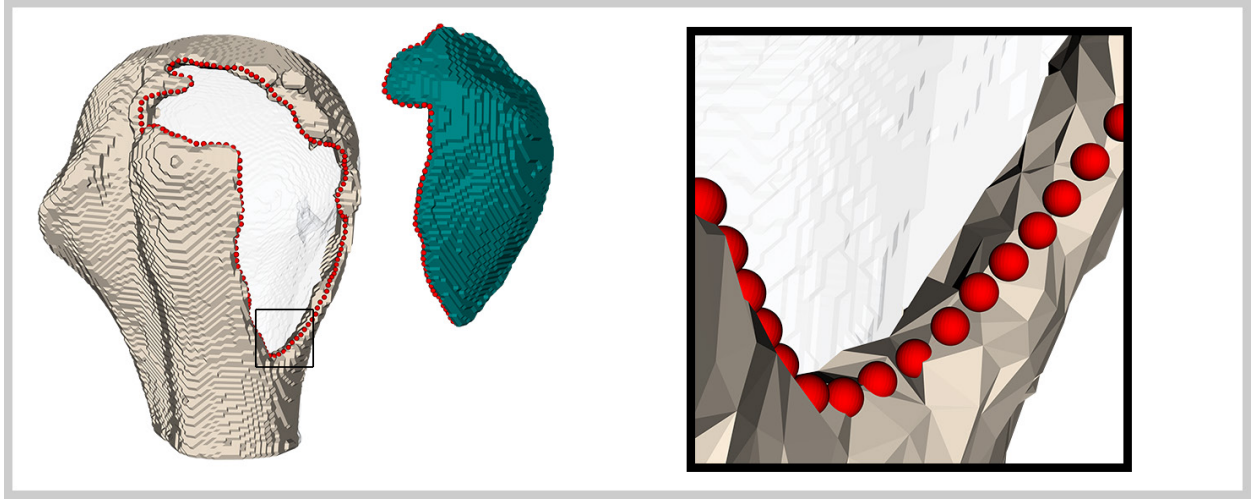


Figure 7.3: Fracture surface representation. The fracture surfaces of the tuberosity fragment and the humeral shaft fragment are simplified by a fracture line, represented by a sequence of points (denoted by red spheres).

7.2.4 Curvature Calculation

The calculation of the curvature of a 3D curve is well known from differential geometry (Kreyszig, 1959; Salomon, 2007). However, it is also known that curvature computation from noisy data is problematic which makes the application to the reconstruction of fractured bones very difficult. To tackle this problem we introduce a scale-space representation of the curvature of the fracture lines as follows. We define the scale-dependent local shape of FL around a point p_i by an approximating cubic B-spline with one polynomial piece $B_{p_i}^{[-d,d]}$. The scale parameter $d \geq 2$ denotes the length of the fracture line segment $\{p_s\} \in FL$ around p_i to be considered, where $s = \{i - d, \dots, i, \dots, i + d\}$. The B-spline is constructed in a weighted mean-square sense as described in (De Boor, 1978), minimizing

$$\sum_s w_s (B_{p_i}^{[-d,d]}(p_s) - p_s)^2, \quad (7.1)$$

where $B_{p_i}^{[-d,d]}(p_s)$ is the value of the spline $B_{p_i}^{[-d,d]}$ for point p_s and the weights are selected as

$$w_s = \begin{cases} j^2, & s < i \\ j^4, & s = i \\ (2d + 2 - j)^2, & s > i \end{cases} \quad \text{and } j = \{1, 2, \dots, |s|\}$$

The weighting function enforces that the approximating cubic B-spline calculated around a point p_i passes through the spline's centre point p_i (Figure 7.4, A). The set of approximating cubic B-splines calculated for scale d represents the fracture line FL of the fragment in scale d (Figure 7.4, B). In other words, the entire fragment is characterized by the fracture line.

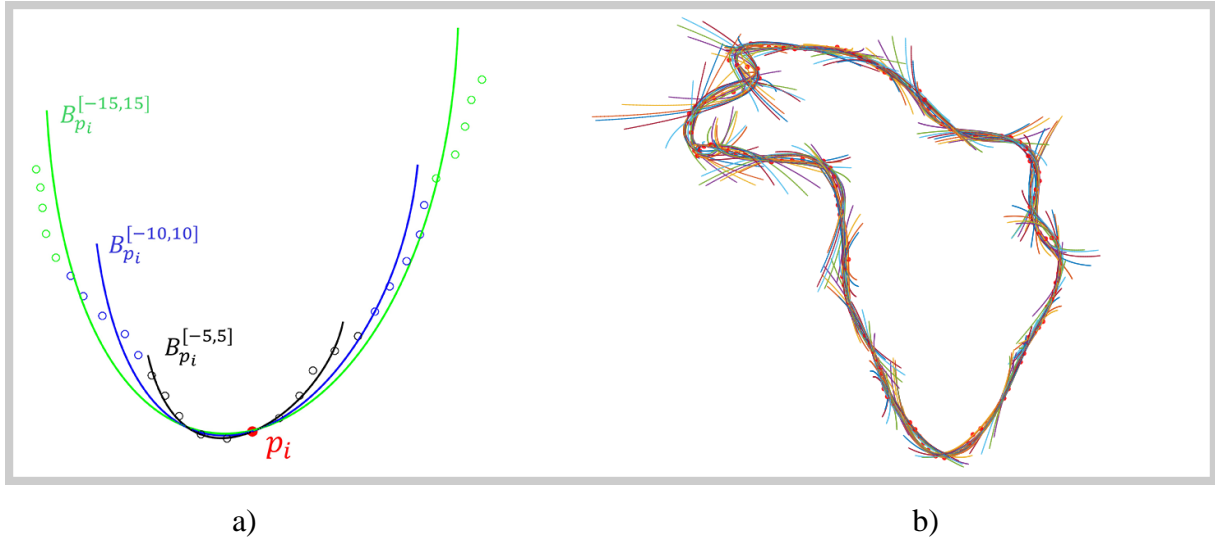


Figure 7.4: Curvature Calculation. a) Approximating cubic B-Splines are illustrated for scales $d=5$ (black), $d=10$ (blue), and $d=15$ (green) centred on a point p_i of the fracture line. b) The B-splines for all p_i for scale $d=10$ represent the fracture line of a fragment.

The curvature vector $\mathbf{k}_{B_{p_i}^{[-d,d]}}$ of $B_{p_i}^{[-d,d]}$ is defined analog to the standard definition of the curvature vector of a parametric curve as described in (De Boor, 1978).

The local shape of a fracture line can be classified into concave or convex parts, depending on the orientation of the curvature. Similar as for a jigsaw puzzle, corresponding border segments of two fragments have opposing curvature orientations. To take these characteristics into account for the matching of corresponding curvatures, we introduce the signed curvature value around a point p_i

$$ks_{B_{p_i}^{[-d,d]}} = \left\| \mathbf{k}_{B_{p_i}^{[-d,d]}} \right\| \text{sign}(\mathbf{k}_{B_{p_i}^{[-d,d]}} \cdot \mathbf{v}_{p_i}) \quad (7.2)$$

where \cdot denotes the dot product and \mathbf{v}_{p_i} the surface normal of the underlying triangle mesh at p_i .

For two signed curvature values ks_1 and ks_2 we define the cost $cf(ks_1, ks_2) = ks_1 + ks_2$, i.e., for two ideal matching fracture fragments the cost function is zero for all curvature values, since the absolute curvature values are equal but their signs are opposite. The cost function will be used for the selection of matching candidates as described in the following Section 7.2.5.

7.2.4.1 Scale-Space Representation

The scale-space representation is used to search for matching candidates between two fracture lines in the same scale. For a scale d the curvature along the fracture line can be expressed by the curvature function

$$K_{FL}(p_i, d) = ks_{B_{p_i}^{[-d,d]}}, \forall p_i \in FL \quad (7.3)$$

The scale-space D is generated for all fragments by calculating the curvature function of Equation (7.3) for all p_i and scales $d \in \{5, 10, 15, 20, 25\}$. The scale-space representation of the humeral shaft fragment is illustrated in Figure 7.5.

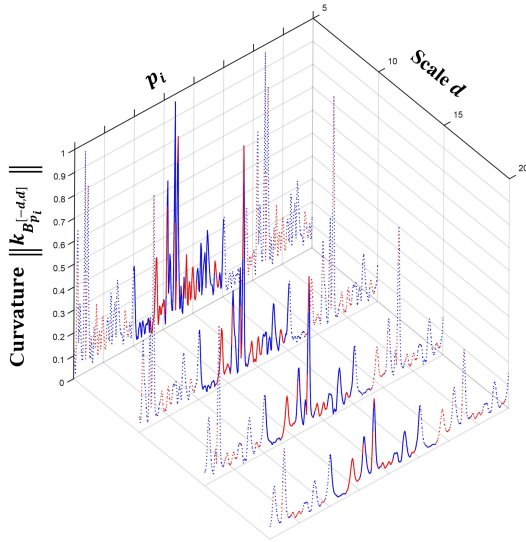


Figure 7.5: Scale-space representation of the curvature function. The sign of the curvature value is encoded by the colour (concave: blue, convex: red) and the periodicity is depicted as dotted lines.

7.2.5 Pairwise Reduction

Given two different fragments F_l and F_m we seek for the best possible pairwise reduction(s). We perform the selection of the best pairwise matches in two steps. In the first step we determine corresponding regions of the curvature of the fracture lines FL_l and FL_m where the cost function cf is minimal. The corresponding regions are detected in the same scale of the scale-space D_l and D_m as described in 7.2.5.1. Based on the corresponding regions, we perform the pairwise reductions and evaluate the reductions which result to the smallest reduction error as described in 7.2.5.2. For each combination of two fragments we will keep the best possible pairwise reductions for the fracture reconstruction as described in 7.2.7.

7.2.5.1 Corresponding regions

For two fracture lines the curvature functions of Equation (7.3) are analysed to identify corresponding regions in each scale. Only highly promising candidates have to be investigated to reduce the number of possible combinations of matching curvature regions. The clinical observation indicated that especially the fragment of the greater tuberosity has regularly a distal triangular tip, which matches to a corresponding defect in the shaft (Gerber et al., 2004). In the surgery, the typical first step is to reduce the greater tuberosity (Gerber et al., 2004). These findings led us to the assumption that peaks in the curvature function represent promising candidates, as a prominently shaped region of the fracture surface is supposed to have high curvature. Therefore, our approach is to analyse the peaks $pks(FL, d)$ of each curvature function K_{FL} . Peaks are detected by calculating local maxima of the absolute curvature function $|K_{FL}(p_i, d)|$. To further reduce the number of candidates, all local

maxima that are smaller than 10% of maximum curvature value (i.e., $\max |K_{FL}(p_i, d)| \mid \forall p_i \in FL$) can be rejected. Thereafter, in each scale corresponding pairs $[pks_u, pks_v]$ of peak points are determined for $\forall pks_u \in pks(FL_l, d)$ and $\forall pks_v \in pks(FL_m, d)$, where $K_{FL_l}(pks_u, d)K_{FL_m}(pks_v, d) < 0$.

Candidate selection along the fracture lines FL_l and FL_m can then be performed for each scale d and corresponding pairs $[pks_u, pks_v]$ by finding the largest intervals $\{p_q \dots p_{q+\lambda}\}$ and $\{p_r \dots p_{r+\lambda}\}$ of length λ , with $pks_u \in \{p_q \dots p_{q+\lambda}\}$, $pks_v \in \{p_r \dots p_{r+\lambda}\}$ and $\lambda \geq 10$, for which it holds that

$$\left| cf \left(K_{FL_l}(p_{q+t}, d), K_{FL_m}(p_{r+t}, d) \right) \right| < \frac{|K_{FL_l}(pks_u, d)| + |K_{FL_m}(pks_v, d)|}{2} \quad (7.4)$$

$$\forall 0 \leq t \leq \lambda$$

It should be mentioned that whenever $p_{q+\lambda}$ or $p_{r+\lambda}$ exceed the domain of K_{FL} , the function K_{FL} is assumed to be periodic and, in addition, it holds that $\lambda < \min(|FS_l|, |FS_m|)$ where $|\cdot|$ denotes cardinality.

According to Equation (7.4) matching intervals can be expressed by the peak points pks_u and pks_v if two lengths λ_1 and λ_2 are introduced. For example, in Figure 7.6, the matching interval $ci = [pks_u, pks_v, \lambda_1, \lambda_2]$ between two fracture lines equals the set of points $\{pks_{u-\lambda_1} \dots pks_u \dots pks_{u+\lambda_2}\} = \{p_q \dots p_{q+\lambda}\}$ and $\{pks_{v-\lambda_1} \dots pks_v \dots pks_{v+\lambda_2}\} = \{p_r \dots p_{r+\lambda}\}$. The matching intervals will be used to calculate the relative transformations, that align the fragments F_l and F_m (Section 7.2.5.2), and in addition to evaluate the pairwise reduction.

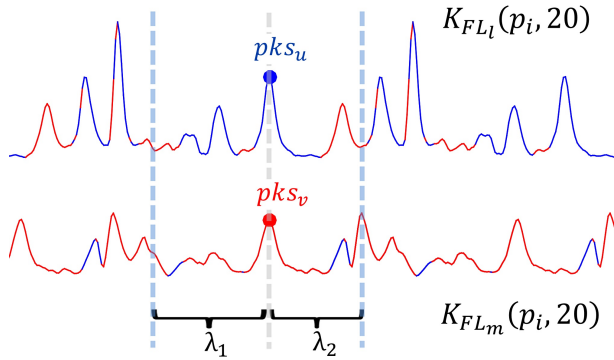


Figure 7.6: Visualization of matching intervals. The interval is illustrated on scale $d = 20$ of the curvature functions of FL_l and FL_m with corresponding peaks pks_u and pks_v and lengths λ_1 and λ_2 .

7.2.5.2 Evaluation of the pairwise reduction

The method of calculating the pairwise reduction was inspired by the surgical technique for the alignment of two fragments. Here, after identification of matching features (Gerber et al., 2004), the fragments are manipulated by the surgeon to align the fragments as accurately as possible.

The pairwise reduction is achieved by the calculation of the transformation which aligns the two fragments. In most of the previous approaches (Jimenez-Delgado et al., 2016), which use a ICP-based algorithm to perform the reduction, the correspondence between the surface points is not known and has to be determined by iterative calculation of correspondences. In contrast to simple ICP-based approaches, the correct correspondence between points is implicitly given by our method (Figure 7.7). The rigid transformation T_{ci} which aligns the point sets $\{p_{pks_v-\lambda_1} \dots p_{pks_v+\lambda_2}\}$ and $\{p_{pks_u-\lambda_1} \dots p_{pks_u+\lambda_2}\}$, is calculated using absolute orientation (Horn, 1987). However, the 3D orientation of fracture lines necessitates introducing a second measurement. Previous studies from archaeology used the torsion as a signature of the 3D-curve (Papaioannou and Theoharis, 2003; Üçoluk and Toroslu, 1999). Our approach was to evaluate the distance between the points of the fracture lines after the reduction. Papaioannou and Theoharis (2003) considered two fragments surfaces as matching candidates if the length of the boundary segments was at least one fourth of the arc length of the shortest boundary. As we aimed to avoid a fixed threshold value, our approach was to assign a higher weight to larger matching intervals. Therefore, we defined the pairwise reduction error measure $prem$, by the root mean squared error (RMSE) of the Euclidean distance between the aligned point sets, normalized by their arc length:

$$prem_{ci} = \frac{RMSE}{\lambda} \quad (7.5)$$

Assuming that two matching intervals of different sizes would result in the same $RMSE$, the normalization by the arc length would favour the larger intervals.

Hence, the set of matching intervals for the fracture lines FL_l and FL_m (as calculated in 7.2.5.1) is

$$CI_{FL_l, FL_m} = \{ci \mid pks_u \in pks(FL_l, d), pks_v \in pks(FL_m, d), \forall d \in D\}$$

For each matching interval we calculate the rigid transformation that aligns the corresponding point sets of the fracture lines. Correspondingly the calculated set of transformations to reduce the fragments F_l and F_m is defined as

$$T_{FL_l, FL_m} = \{T_{ci} \mid ci \in CI_{FL_l, FL_m}\}$$

For the further iterative fracture reconstruction, we merge the fragments F_l and F_m and calculate merged fracture line(s) as described in 7.2.6.

7.2.6 Calculation of the merged fracture line

For the determined set of transformations T_{FL_l, FL_m} we calculate merged fracture lines used for the further fracture reconstruction as illustrated in Figure 7.7. Without loss of generality it holds that $|\{p_i\}| > |\{q_i\}|$, $p_i \in FL_l$, $q_i \in FL_m$ (Figure 7.7, a).

The steps of the algorithm for the calculation of a merged fracture line are as follows:

- Apply the transformation T_{ci} to the sequence of points of the fracture line FL_m (Figure 7.7, b).
- Calculate the residual points of the fracture lines $FL_{res_l} = FL_l \setminus \{p_{pks_u-\lambda_1} \dots p_{pks_u+\lambda_2}\}$ and $FL_{res_m} = FL_m \setminus \{p_{pks_v-\lambda_1} \dots p_{pks_v+\lambda_2}\}$ (Figure 7.7, c).
- Determine the end points (p_{start_l} and p_{start_m}) and (p_{end_l} and p_{end_m}) of FL_{res_l} and FL_{res_m} on both sides.
- Connect FL_{res_l} and FL_{res_m} by cubic-spline interpolation between p_{end_l} and p_{end_m} and between p_{start_l} and p_{start_m} (Figure 7.7, d), and sub-sample the connection with a sampling size of r_{FL} by introducing additional points $\{p_{con_{start}}\}$ and $\{p_{con_{end}}\}$.
- Determine the point p_{merge} of FL_{res_l} with the maximal distance to the nearest point of FL_{res_m} (Figure 7.7, e).
- Starting from p_{merge} create a directed weighted graph of $\{FL_{res_l} \cup p_{con_{start}} \cup FL_{res_m} \cup p_{con_{end}}\}$, where each point is connected to all subsequent points within a radius of $r_{graph} = 1.5$ mm and the edge weight correspond to the distance between the points of an edge. In order to obtain a similar distance r_{FL} , as defined in Section 7.2.3 between the sequence of points of $FL_{l,m}$, r_{graph} has to be greater than r_{FL} and smaller than $2 * r_{FL}$. Therefore, we set $r_{graph} = 0.5 * (r_{FL} + 2 * r_{FL})$.
- Calculate the merged fracture line $FL_{l,m}$ using the shortest path algorithm (Dijkstra, 1959) (Figure 7.7, f).

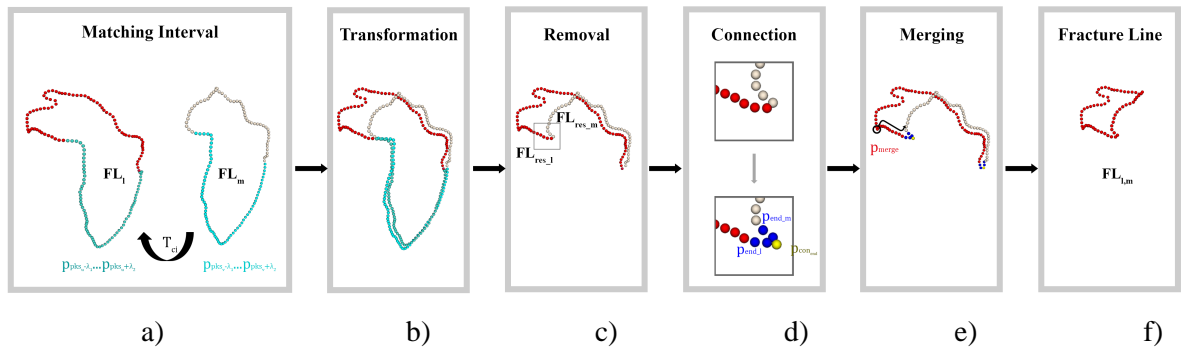


Figure 7.7: Calculation of the merged fracture line. a) The calculated transformation T_{ci} , which aligns the point sets of the matching interval (denoted by cyan spheres) is used to b) align FL_l and FL_m . c) The point sets of the matching interval are removed. d) The residual points of the fracture lines $FL_{res,l}$ and $FL_{res,m}$ are connected with $p_{con_{end}}$ by cubic-spline interpolation between $p_{end,l}$ and $p_{end,m}$. e) Starting from p_{merge} create a directed graph and use the shortest path algorithm to determine the merged fracture line $FL_{l,m}$.

7.2.7 Fracture Reconstruction

Given a fracture with n fragments $\{F_1, F_2, \dots, F_n\}$ the goal of the fracture reconstruction algorithm is to calculate n corresponding transformations, which will reduce the fragments to their pretraumatic position. The fracture reconstruction is performed by evaluating the reduction pairwise and repeating the process iteratively for merged fragments as described in 7.2.3 to 7.2.6. The order of pairwise reduction is crucial for the outcome of any fracture reconstruction algorithm. For example, the correct reduction of one fragment might only be possible if the pairwise reduction of two other fragments has already taken place. Figure 7.8 demonstrates the relevance of the reduction sequence for a proximal humeral fracture with five fragments.

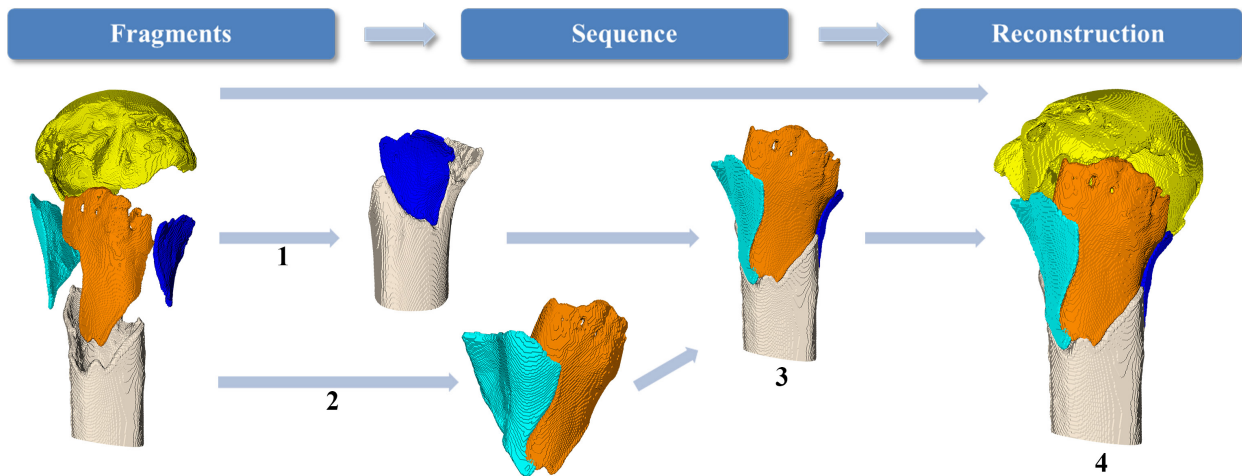


Figure 7.8: Importance of the order of pairwise reductions, illustrated for a proximal humeral fracture with five fragments. 1) The shaft fragment (beige) is first merged with the blue fragment and 2) the cyan fragment is merged with the orange fragment. 3) Thereafter, the results of 1) and 2) are merged and 4) finally, the result of 3) is merged with the head fragment (yellow).

Because the optimal order cannot be determined a priori, our algorithm processes all possible combinations and determines finally the best solution by evaluation of the reconstruction results as described in 7.2.7.1. For a three-part fracture, it is necessary to calculate three combinations of pairwise reduction, which corresponds to $3 \times 2 = 6$ pairwise reductions in total. For a four-part fracture, there are fifteen combinations, which yield in total $15 \times 3 = 45$ pairwise reductions. For a five-part fracture, there are 135 combinations, which yield in total $135 \times 4 = 540$ pairwise reductions. Figure 7.9 illustrates the sequence of combinations.

The fracture reconstruction, which yielded the smallest *frem* value, was regarded as the best solution.

We have to note that the automatic selection of the best solution might fail if only a partial reconstruction is performed, i.e., if relevant fragments of the proximal humerus are not considered in the reconstruction algorithm. One possible reason for this could be that the surgeon failed to identify and segment all fragments. For example, a fracture reconstruction which leads to a folding of the fragments could have in these cases a smaller *frem* value than the optimal reconstruction.

A warning mechanism was implemented in the automatic selection for the best solution to inform the user that relevant information might be missing, which is required to achieve the reconstruction. The idea of the mechanism is to verify whether the surface area $area_{fragments}$ of all considered fragments is comparable to the surface area of a healthy proximal humerus $area_{expected}$. The mechanism was only necessary if a humeral head fragment was present. The sum of the surface area of all fragments $area_{fragments}$ was calculated as described in Section 7.2.2.1. The surface area of the shaft fragment was excluded since the size depends mainly on the acquisition length as described in Section 7.2.2.1. The approximated area $area_{expected}$ was estimated by the surface area of a sphere fitted to the surface points of the humeral head fragment (see Figure 7.10).

In case of a fracture with a humeral head fragment the algorithm calculates the expected surface area $area_{expected}$ and $area_{fragments}$. If $area_{expected} \geq 2 * area_{fragments}$ the algorithm raises a warning, that the best solution might not be the solution with the smallest *frem*. In these cases the solutions have to be inspected by the user.

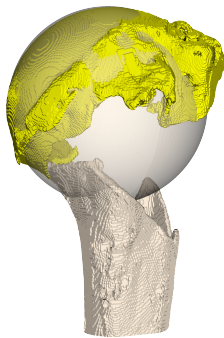


Figure 7.10: Approximation of the expected area of reconstruction. The surface area of a sphere, fitted to the surface points of the humeral head fragment, was used to approximate the expected surface area of reconstruction.

7.3 Results

7.3.1 Clinical evaluation

All reconstructions of the twelve clinical cases resulted in an accurate approximation of the pretraumatic anatomy. The quality of the best solution provided by the algorithm was evaluated by two surgeons, trained in orthopaedic surgery and specialized in shoulder surgery. The evaluation was based on a presentation showing the 3D reconstructed fracture from multiple, standardized viewpoints. The examiners were told to evaluate the proposed reconstruction, i.e., whether the reconstruction would restore the pretraumatic anatomy or whether a malposition of the fragments was apparent. In clinical studies the postoperative malposition is normally assessed on conventional radiographs and a malposition of a fragment is assumed, when a displacement of at least 1 cm or of 45° is present (Gerber et al., 2004). We defined for the clinical evaluation a much lower cut-off of 5 mm or 10°, respectively.

For the evaluation, we used the following 5 point-scale:

- Acceptable without modification
- Acceptable after small modifications
(i.e. correction of one fragment $< 10^\circ$ or < 5 mm)
- Acceptable after large modifications
(i.e. correction of more than one fragment or one fragment $> 10^\circ$ or > 5 mm)
- Not acceptable at all
- Determination of the correct alignment is not possible

In Figure 7.11, we present the reconstruction results of all clinical cases sorted by the number of fragments. Both surgeons rated eleven of the twelve clinical cases as acceptable without modification. Case 6 was rated by both surgeons as acceptable after small modifications, since the reconstruction resulted in a small malrotation of the humeral head.

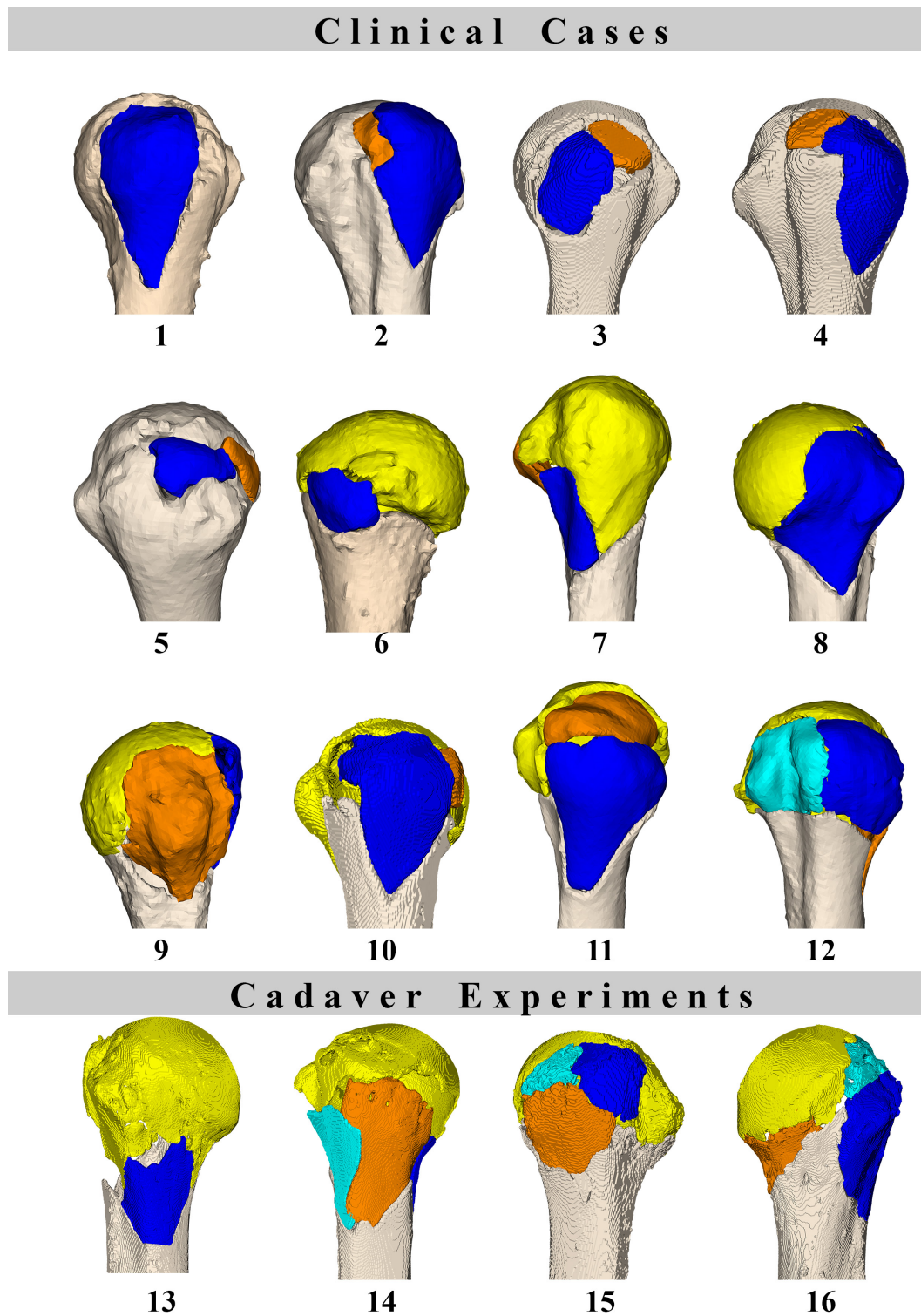


Figure 7.11: Reconstruction results. The cases are sorted by the number of fragments. Clinical Cases: Case 1 (two fragments), Case 2-6 (three fragments), Case 7-11 (four fragments) and Case 12 (five fragments) Cadaver Experiments: Case 13 (three fragments) and Case 14-16 (five fragments).

7.3.2 Quantitative evaluation based on cadaver experiments

The accuracy of our method was evaluated based on four cadaveric data sets, previously acquired in the study of Frnstahl et al. (2012). The CT scans had the same axial resolution of 1 mm, as the clinical cases. The fragment size of the cadaver bone fractures was also comparable (Table 7.2). All cases could be fully reconstructed. The best solution provided by the algorithm is illustrated in Figure 7.11. In case 13, the warning mechanism of the algorithm indicated that the proposed solution with the smallest *frem* error might be not the best one. The reason was that the creation of the artificial fracture had resulted in a comminuted fracture with a large number of small fragments. These fragments were not included into the fracture reconstruction algorithm, similar as in the previous study (Frnstahl et al., 2012). In this case the $area_{fragments}$ was $0.46 * area_{expected}$.

	Area (mm ²)		
	Mean	SD	Range
Head (4)	2925	228	2749 - 3232
Intermediate (10)	860	486	364 - 1958
All fragments (14)	1409	1037	364 - 3232

Table 7.2: Characteristics of the fragments of the cadaver experiments. The area of all fragments is summarized, which were used for the fracture reconstruction.

The goal of a fracture reconstruction algorithm is to reduce the fragments such that the pretraumatic anatomy is approximated as well as possible. The approximation of the pretraumatic anatomy can be quantitatively measured in the cadaver experiments by comparing the difference of the reconstructed surface of the fragments with the surface of the original unfractured bone model.

Here, the distance error was defined as the Euclidean distance between all points of the outer corticalis of the reduced fragments to the closest point on the outer corticalis of the original bone model (Figure 7.12). The distance errors of all cases reconstructed with the presented algorithm and the reference method (Frnstahl et al., 2012) are presented in Table 7.3.

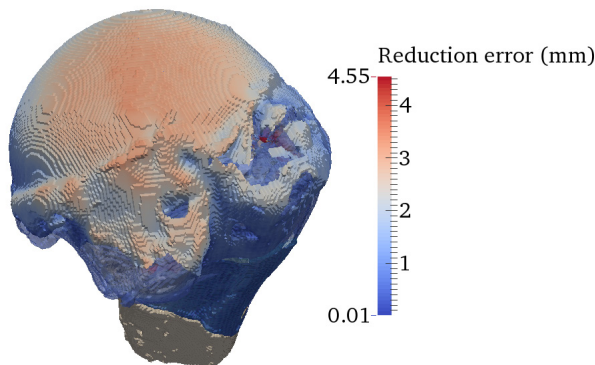


Figure 7.12: Distance error. Euclidean Distances are calculated for all surface points of the outer cortical layer between the reduced fragments and the original bone model.

Cadaver		Distance Error (mm)							
Case	Fragment	Presented Method				Reference Method			
		Mean	SD	Range		Mean	SD	Range	
13	Head	1.31	0.73	0.02	- 4.99	1.81	0.99	0.02	- 6.02
	Intermediate	1.09	0.57	0.02	- 2.87	1.70	0.83	0.08	- 3.79
14	Head	1.68	0.89	0.02	- 4.55	0.89	0.38	0.04	- 4.32
	Intermediate	0.76	0.41	0.02	- 2.75	0.90	0.34	0.04	- 2.70
15	Head	1.18	0.56	0.03	- 4.69	0.58	0.30	0.02	- 3.87
	Intermediate	0.95	0.44	0.03	- 2.84	1.16	0.56	0.01	- 3.68
16	Head	1.36	0.74	0.03	- 3.94	1.48	0.77	0.02	- 3.45
	Intermediate	0.95	0.67	0.02	- 3.88	0.80	0.48	0.02	- 3.84
All	Head	1.41	0.78	0.02	- 4.99	1.23	0.85	0.02	- 6.02
	Intermediate	0.90	0.56	0.02	- 3.88	0.97	0.56	0.01	- 3.84

Table 7.3: Comparison of the distance errors between the proposed algorithm and the reference method of Frnstahl et al. (2012). Mean, standard deviation and range of the error are given.

7.3.3 Order of pairwise reduction

In fourteen of the eighteen cases the best solution was achieved by subsequent merging of the fragments (similar as illustrated in the top rows of Figure 7.9 for the four-part fractures). In four cases, the best solution was achieved by performing first two pairwise reductions of two fragments (i.e., F_{12} and F_{34}), before the final reduction was completed (similar as illustrated in the bottom rows of Figure 7.9 for the four-part fractures and in Figure 7.8).

7.3.4 Runtime

All parts of the proposed algorithm were implemented in MATLAB (2015). The runtime scales with the number of fragments of the fracture. Table 7.4 summarizes the runtimes of our algorithm (curvature calculation, pairwise reduction, calculation of the merged fracture line and fracture reconstruction). To compare the results with the reference method, we summarized the runtime of the reference method for the contralateral assembly and the fracture surface assembly, without the preprocessing time (fracture segmentation and fracture surface extraction).

Number of Fragments	Runtime (s)	
	Presented Method	Reference Method (CPU)
3	121	2859
4	4220	5742
5	11322	9127

Table 7.4: Average runtime of the presented method implemented in MATLAB in seconds (3.4 GHz Intel Core i7-2600 CPU, 16GB RAM) compared to the runtime of the CPU version of the reference method (P4 3.2GHz).

7.4 Discussion

In this work, we presented a novel method for the fully automated anatomical reconstruction of proximal humeral fractures. One major advantage of the presented method in contrast to the state-of-the-art (Fürnstahl et al., 2012) is that the knowledge of the contralateral anatomy is not necessary anymore. Therefore, our approach is even applicable in presence of bilateral pathological conditions. At the same time, the acquisition of the contralateral anatomy necessitates an additional CT scan, leading to increased radiation exposure for the patient. This is a major hurdle for the clinical introduction of a new method in the hospital. Our algorithm tackles this problem by performing the fracture reconstruction solely based on the information of the fractured surfaces without relying on reconstruction template. Certainly the fractured surface of a fragment does not represent the whole morphology of the fragment. However, we demonstrated on twelve clinical cases that the information of the fractured surfaces is already sufficient to accurately reconstruct proximal humeral fractures with respect to clinical requirements.

Even if it is possible to replace the contralateral anatomy by a different reconstruction template, e. g., a statistical shape model as in Albrecht and Vetter (2012), the fragments at the proximal humerus are often considerably displaced and malrotated, making their registration to the template difficult (Fürnstahl et al., 2012).

Most of the current approaches for fracture reconstruction have in common that an ICP-based algorithm is used to perform the reduction of the fragments (Jimenez-Delgado et al., 2016). The tendency of the ICP to fall into local minima might be particularly problematic in case of the proximal humerus due to its almost spherical shape (Fürnstahl et al., 2012). Furthermore, an ICP-based algorithm might introduce an error in the reconstruction of the pretraumatic anatomy if a plastic deformation of the fragments is present. The relevance of three-dimensional analysis for the diagnosis of plastic deformation of a presumed intact intra-articular surface of a complex malunited humeral fracture has already been emphasized (Vlachopoulos et al., 2016b). The influence of the plastic deformation onto the ICP-registration results is demonstrated Figure 7.13. Although only a small part of the fragment is deformed, the ICP-registration yields an incorrect result if a plastic deformation is present. In the first column of Figure 7.13 the tuberosity fragment of Case 14 is presented as it was used for the fracture reconstruction and after registration onto the unfractured humerus. In the second column, the plastic deformed part of the fragment was first removed before the ICP-registration. It can be clearly seen how the inherent local deviation caused by the plastically deformed region deteriorates the reconstruction quality of the entire fragment. On the contrary, the fracture reconstruction with our presented method is not compromised as it did not rely on the fracture line segment affected by the plastic deformation.

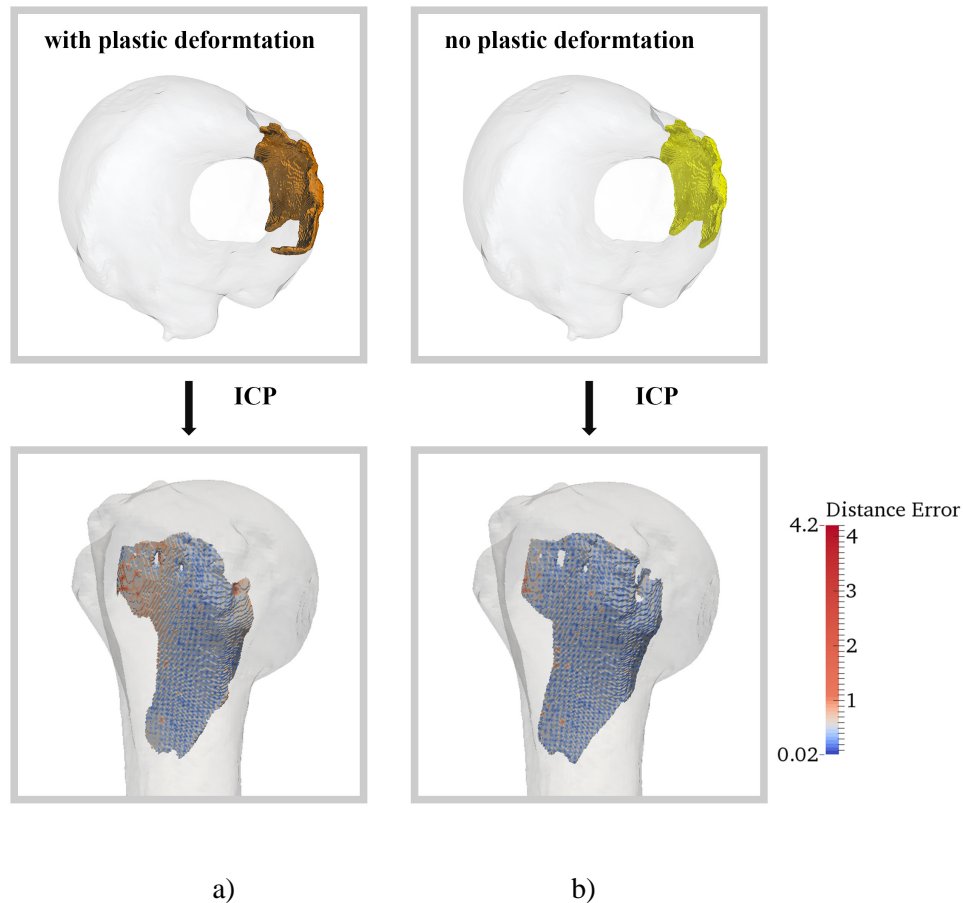


Figure 7.13: Influence of a plastic deformation of a fragment on the ICP registration result. a) The fragment shows a plastic deformation and the distance error after ICP registration b) The deformed area of the fragment was removed before the ICP registration.

Compared to other fracture reconstruction approaches, that perform the reconstruction based on the fracture surfaces of the fragments our algorithm has clear advantages but also some drawbacks. The main drawback of our approach compared to the method of Okada et al. (2009) or Kronman and Joskowicz (2013) is the significantly higher runtime. Nevertheless, these registration methods can only handle fragment displacements up to a certain degree. Okada et al. (2009) limited the displacements to 30° and 20 mm and Kronman and Joskowicz (2013) limited the displacement to 20° and 30 mm around each axis. The pose-invariant feature of our algorithm is therefore of particular importance for fractures of the proximal humerus, due to the observed greater displacements and a great advantage compared to the previous algorithms. Furthermore, the method of Okada et al. (2009) could register the fracture surfaces only if the order of alignment as well as the corresponding regions were chosen beforehand. Best result were obtained with a method that combined fracture surfaces and the contralateral anatomy, and Okada et al. (2009) pointed out, that constraints posed by fracture surfaces alone were insufficient to perform registration with acceptable accuracy. An additional limitation of Kronman and Joskowicz (2013) is that the method is only applicable for bone fractures consisting of two fragments and the method was not extended to fractures with multiple fragments.

The quantitative evaluation based on cadaver experiments demonstrated that the reconstruction results of our method are almost equal for head fragments and better for the small fragments, where the reconstruction is even more challenging, compared to the state-of-the-art (Fürnstahl et al., 2012). The mean distance error of reconstructed intermediate fragments was $0.90 \text{ mm} \pm 0.56 \text{ mm}$ for our method compared to $0.96 \text{ mm} \pm 0.56 \text{ mm}$ for the state-of-the-art. The distance error of the head fragments was $1.41 \text{ mm} \pm 0.78 \text{ mm}$ compared to $1.23 \text{ mm} \pm 0.85 \text{ mm}$ for the state-of-the-art.

The computational efficiency of a reconstruction algorithm is a clinically important factor, as it significantly contributes to the overall preparation time for the surgery. For a fracture with four fragments, the runtime of our method was 4220 sec, which outperforms the CPU-version of the reference method (5742 sec). For a fracture with three fragments the runtime benefit was much greater (121 sec compared to 2859 sec) when our algorithm is used. Fractures up to three or four parts are treated by ORIF, but a surgical reconstruction of fractures with more parts is very unlikely to be successful (Gerber et al., 2004; Wijkman et al., 2002).

There would be several options to reduce the complexity of the calculation and therefore the runtime. The implementation of the algorithm in a compiled language, or even better on a graphical processing unit, would allow reducing the runtime (i.e., the runtime of the algorithm of Fürnstahl et al. (2012) was reduced about a factor of 20). In the presented algorithm, the order of the point sequence specifies the direction of the path of the fracture lines. Therefore, for each fragment, the fracture surface was represented by two sequences of points, one in clock-wise and one in counter clock-wise direction. We performed the pairwise matching calculated from the sequence of points of one fragment in one direction and for the other fragment in both directions. The automatic detection of the outer contour of the cortical layer would permit to determine the appropriate direction, and would halve the calculation of pairwise matching. Additional speedup would be possible, if the number of matching candidates per pairwise reduction can be reduced already in the early stage of the algorithm, i.e., by incorporation of geometric constraints of the merged fragments into the decision.

Our method relies on some heuristically defined parameter values which we shortly discuss. For the scale-space representation the scales $d \in \{5, 10, 15, 20, 25\}$ were considered. The lower range was set to scale 5, as significant shape features on the fracture borders had a length of at least 10 mm length. Thereby, a subset of at least 10 consecutive points influenced the calculation of the curvature at each point p_i . The minimal area of a fragment, which was considered by surgeons as relevant for the fracture reconstruction, was 195 mm^2 . The arc-length of such a fragment would be approximately 50 mm. The upper value of d was set to $d = 25$ to consider all 50 points of the smallest fragments in the calculation of the curvature. The maximal number of best matches, which were selected after evaluation of the pairwise reductions for the further calculations in each iteration was set to be 20. The number was defined to be relatively high in order to enable an automatic fracture reconstruction for all cases.

7.5 Conclusion and Future Work

In this paper, we proposed an algorithm for the reconstruction of complex proximal humeral fractures. The key idea of the approach is the use of the curvature scale-space for matching characteristic features between the fragments. The evaluation of our algorithm on a consecutive series of patients with proximal humeral fractures and, additionally, the quantitative validation on cadaver fractures demonstrated, that the shape of the fracture surface encodes sufficient information to perform the reconstruction without needing a reconstruction template such as the contralateral anatomy. Further research will focus on the development of strategies to reduce the number of potential candidates in the pairwise matching at an earlier stage. In addition, it will be interesting to evaluate the application of the method for the anatomical reconstruction of fractures of further anatomies, i.e. fractures of the distal radius or of the proximal femur.

8

Conclusion and Perspective

The preceding chapters of this thesis presented algorithms for the 3D assessment of deformities and preoperative planning of reconstructive surgeries of the humeral anatomy. This chapter will summarize the herein presented methods in the broader context of CAOS and their impact on orthopaedic surgery.

In Chapter 2, computer algorithms were developed for quantifying the humeral anatomy in 3D in a standardized fashion. Some of the characteristics were previously described for 2D and had to be transferred to 3D. In addition, new measurements techniques were developed which are only applicable in 3D space. The analysis of 140 paired humeri indicated that humeral head size, humeral head height and - with limitations - humeral length may be reliably carried over from the contralateral side. These findings are helpful for shoulder replacement surgery, particularly if severe osteoarthritis or osteonecrosis is present. Thereby, the determination of the best fitting size of a prosthetic component is possible, which best approximates the pretraumatic anatomy. Furthermore, the analysis revealed that the differences in torsion and inclination of the humeral head in relation to the distal part of the humerus of both sides are considerably higher than expected. Therefore, we concluded that template-based planning algorithms using the contralateral anatomy as a reconstruction template should compensate for bilateral differences. These considerations were addressed in Chapter 3 and 4.

A template-based approach relying on the contralateral anatomy as a reconstruction template is currently assumed to be the best option for the assessment of posttraumatic deformities. Nevertheless, in Chapter 3 and 4 we demonstrated that simple registration to the contralateral humeral model may not be sufficient to restore the pretraumatic anatomy due to bilateral differences between both sides of an individual. Therefore, a novel segment selection strategy was proposed, yielding an improved method for the 3D assessment of deformities of the proximal and distal humerus. Characteristic features were determined on the proximal and distal humeral shaft parts that allow approximating the pretraumatic anatomy of the proximal and distal humerus in an optimal way.

In Chapter 5, the template-based approach was extended in such a way that the contralateral anatomy is not needed anymore to predict the pretraumatic anatomy. Thereby, the method is even applicable if bilateral pathological conditions are present. A statistical shape model of the humerus was generated and used as a novel reconstruction template. The accuracy to predict the pretraumatic anatomy of the humerus from the healthy parts of the posttraumatic bone model was compared to the approaches previously described in Chapter 3 and 4. We have demonstrated that statistical shape models are powerful tools, accurately predicting the

patient-specific pretraumatic anatomy, especially if only the articular part of the humerus is pathological. The main benefits of statistical shape models are that the acquisition of a CT scan of the contralateral anatomy is not necessary anymore. However, the accuracy of the prediction might be lower if both the proximal and distal joint parts are pathological and only the mid-shaft region can be assumed as healthy. Nevertheless, with the presented methods of Chapter 3 to 5 most anticipated scenarios can be tackled.

Chapter 6 described the extension of the methods that were presented in Chapter 2 for the standardized measurement of the humeral anatomy in 3D by developing an algorithm for the automatic estimation of the anterior margin plane. The anterior margin plane is the plane approximating the cartilage/metaphyseal boundary of the proximal humerus. Hitherto, the anterior margin plane had to be determined by manually setting of reference points. The presented random forest-based approach enables the determination of the anterior margin plane in an almost automated fashion without needing an expert, considerably reducing the manual effort. Furthermore, the most important step in shoulder replacement surgery is the resection of the humeral head along the anatomical neck, represented by the anterior margin plane. The estimation of the anterior margin plane defines, thereby, the orientation, position and size of the prosthetic humeral head in relation to the humeral shaft and influences the selection of the optimal humeral head component.

The algorithm, presented in Chapter 7, resolved a main limitation of the state-of-the-art method for the preoperative planning of proximal humeral fractures, i.e., the requirement to use the contralateral anatomy as a reconstruction template. Our novel fracture reconstruction algorithm performs the reconstruction solely based on the information of the fracture lines of the fragments obtained from the posttraumatic condition. In addition, our method overcomes the further limitation of previous fracture reconstruction algorithms, i.e., that the methods can handle fragment displacements up to a certain degree but that they still rely on the initial position of fragments. Our key idea was to develop a pose-invariant, curvature-based algorithm for the fracture reconstruction. A robust scale-space representation of a 3D curve was presented to cope with the substantial noise in clinical images. The fractures of the proximal humerus were automatically reconstructed based on iterative pairwise reduction of the fragments and calculation of the best solution. The quality of the best solution provided by the algorithm was evaluated by two surgeons, trained in orthopaedic surgery and specialized in shoulder surgery. We demonstrated on twelve clinical cases that the information of the fractured lines is already sufficient to accurately reconstruct proximal humeral fractures with respect to clinical requirements. The quantitative evaluation based on cadaver experiments demonstrated that the reconstruction results of our algorithm are almost equal for head fragments and better for the small fragments compared to the state-of-the-art algorithm for the reconstruction of complex fractures of the proximal humerus. However, the major advantage of the presented method in contrast to the state-of-the-art is that the acquisition of a CT scan of the contralateral anatomy is not necessary anymore. Therefore, our approach is even applicable if bilateral pathological conditions are present. Furthermore, the radiation exposure of the patients is, thereby, significantly reduced.

The algorithms presented in this thesis enable performing the preoperative planning for the treatment of posttraumatic deformities in a standardized fashion. A set of tool was presented to tackle a wide range of treatment options, since only one approach is obviously not

sufficient to cover all underlying pathological conditions and surgical procedures. Therefore, the presented methods are not considered to be mutually exclusive or competitive, but rather complementary.

In future research, the clinical impact of an accurate reconstruction of the pretraumatic anatomy shall be investigated. Computer-assisted preoperative planning and intraoperative navigation using patient-specific templates have been successfully applied for corrective osteotomies of the humerus. The presented methods have to be incorporated in the preoperative planning in order to immediately contribute to the achievement of the ultimate goal, the restoration of the pretraumatic anatomy. In particular, it has been proven that the accuracy of the registration of the preoperative plan to the intraoperative situation is high for corrective osteotomies of the humerus. However, the intraoperative navigation (i.e., using patient-specific guides) has not been described for the fracture reconstruction of the proximal humerus. The limited size of the fragments and the relative motion between the fragments are probably the greatest challenges that have to be tackled in order to apply intraoperative navigation techniques to fractures of the proximal humerus. Furthermore, although the performance of the algorithm was proven to be excellent on a consecutive series of patients and on cadaver experiments, the fracture surface representation necessitates some user-interaction, which has to be minimized to reduce the planning effort. Further research will focus also on the development of strategies to increase the computational efficiency of the approach, i.e., by removing incorrect pairwise matching at an earlier stage by allowing user-interaction for the evaluation. In addition, it will be interesting to evaluate the application of the algorithm for the anatomical reconstruction of fractures of further anatomies, i.e., fractures of the distal radius or of the proximal femur. Future research will focus also on the identification of characteristic features for the template-based approach for further anatomies (i.e., forearm). It has been demonstrated that statistical shape models are valuable tools to approximate the pretraumatic anatomy of the humerus. The influence of the incorporation of additional prior knowledge (i.e., gender, hand-dominance and working or sport activities) on the prediction accuracy of the statistical shape models shall also be analysed in future research.

In conclusion, it is reasonable to assume, that the quantitative assessment of the deformities will contribute to a more differentiated classification of pathological conditions and to an evidence-based selection of the patient-specific optimal treatment. Nevertheless, the ultimate goal, i.e., the clinical and functional benefit for the patients, has to be proven first for the proposed methods of CAOS in ongoing clinical trials.

Bibliography

- Albrecht, T., Lüthi, M., Gerig, T., Vetter, T., 2013. Posterior shape models. *Medical image analysis* 17, 959-973.
- Albrecht, T., Vetter, T., 2012. Automatic fracture reduction, *Mesh Processing in Medical Image Analysis 2012*. Springer, pp. 22-29.
- Alolabi, B., Youderian, A.R., Napolitano, L., Szerlip, B.W., Evans, P.J., Nowinski, R.J., Ricchetti, E.T., Iannotti, J.P., 2014. Radiographic assessment of prosthetic humeral head size after anatomic shoulder arthroplasty. *J Shoulder Elbow Surg* 23, 1740-1746.
- Amin, S., Achenbach, S.J., Atkinson, E.J., Khosla, S., Melton, L.J., 2014. Trends in Fracture Incidence: A Population-Based Study Over 20 Years. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research* 29, 581-589.
- Anakwenze, O.A., Zoller, S., Ahmad, C.S., Levine, W.N., 2014. Reverse shoulder arthroplasty for acute proximal humerus fractures: a systematic review. *Journal of Shoulder and Elbow Surgery* 23, e73-e80.
- Audette, M.A., Ferrie, F.P., Peters, T.M., 2000. An algorithmic overview of surface registration techniques for medical imaging. *Med Image Anal* 4, 201-217.
- Ballmer, F.T., Sidles, J.A., Lippitt, S.B., Matsen, F.A., 3rd, 1993. Humeral head prosthetic arthroplasty: Surgically relevant geometric considerations. *J Shoulder Elbow Surg* 2, 296-304.
- Besl, P.J., McKay, N.D., 1992. Method for registration of 3-D shapes, *Robotics-DL tentative*. International Society for Optics and Photonics, pp. 586-606.
- Bicknell, R.T., DeLude, J.A., Kedgley, A.E., Ferreira, L.M., Dunning, C.E., King, G.J., Faber, K.J., Johnson, J.A., Drosdowech, D.S., 2007. Early experience with computer-assisted shoulder hemiarthroplasty for fractures of the proximal humerus: development of a novel technique and an in vitro comparison with traditional methods. *J Shoulder Elbow Surg* 16, S117-125.
- Bohsali, K.I., Wirth, M.A., Rockwood, C.A., Jr., 2006. Complications of total shoulder arthroplasty. *J Bone Joint Surg Am* 88, 2279-2292.

- Boileau, P., Bicknell, R.T., Mazzoleni, N., Walch, G., Urien, J.P., 2008. CT scan method accurately assesses humeral head retroversion. *Clinical orthopaedics and related research* 466, 661-669.
- Boileau, P., Krishnan, S.G., Tinsi, L., Walch, G., Coste, J.S., Mole, D., 2002. Tuberosity malposition and migration: reasons for poor outcomes after hemiarthroplasty for displaced fractures of the proximal humerus. *J Shoulder Elbow Surg* 11, 401-412.
- Boileau, P., Walch, G., 1997. The three-dimensional geometry of the proximal humerus. Implications for surgical technique and prosthetic design. *J Bone Joint Surg Br* 79, 857-865.
- Breiman, L., 2001. Random forests. *Machine learning* 45, 5-32.
- Carofino, B.C., Leopold, S.S., 2013. Classifications in Brief: The Neer Classification for Proximal Humerus Fractures. *Clinical Orthopaedics and Related Research* 471, 39-43.
- Casciaro, M.E., Craiem, D., 2014. Towards automatic measurement of anteversion and neck-shaft angles in human femurs using CT images. *Computer methods in biomechanics and biomedical engineering* 17, 128-136.
- Chen, C., Zheng, G., 2013. Fully Automatic Segmentation of AP Pelvis X-rays via Random Forest Regression and Hierarchical Sparse Shape Composition, *Computer Analysis of Images and Patterns*. Springer, pp. 335-343.
- Chen, Y., Medioni, G., 1991. Object modeling by registration of multiple range images, *Robotics and Automation*, 1991. Proceedings., 1991 IEEE International Conference on. IEEE, pp. 2724-2729.
- Chowdhury, A.S., Bhandarkar, S.M., Robinson, R.W., Yu, J.C., 2009. Virtual multi-fracture craniofacial reconstruction using computer vision and graph matching. *Computerized Medical Imaging and Graphics* 33, 333-342.
- Cooper, C., Dennison, E.M., Leufkens, H.G., Bishop, N., van Staa, T.P., 2004. Epidemiology of childhood fractures in Britain: a study using the general practice research database. *J Bone Miner Res* 19, 1976-1981.
- Court-Brown, C.M., Caesar, B., 2006. Epidemiology of adult fractures: A review. *Injury* 37, 691-697.
- Criminisi, A., Shotton, J., 2013. Regression Forests, In: Criminisi, A., Shotton, J. (Eds.), *Decision Forests for Computer Vision and Medical Image Analysis*. Springer London, London, pp. 47-58.
- Criminisi, A., Shotton, J., Robertson, D., Konukoglu, E., 2011. Regression forests for efficient anatomy detection and localization in CT studies, *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*. Springer, pp. 106-117.
- Cuff, D.J., Pupello, D.R., 2013. Comparison of hemiarthroplasty and reverse shoulder arthroplasty for the treatment of proximal humeral fractures in elderly patients. *J Bone Joint Surg Am* 95, 2050-2055.

- Cvetanovich, G.L., Chalmers, P.N., Verma, N.N., Nicholson, G.P., Romeo, A.A., 2016. Open reduction internal fixation has fewer short-term complications than shoulder arthroplasty for proximal humeral fractures. *J Shoulder Elbow Surg* 25, 624-631.e623.
- Day, J.S., Lau, E., Ong, K.L., Williams, G.R., Ramsey, M.L., Kurtz, S.M., 2010. Prevalence and projections of total shoulder and elbow arthroplasty in the United States to 2015. *J Shoulder Elbow Surg* 19, 1115-1120.
- De Boor, C., 1978. A practical guide to splines. Springer-Verlag New York.
- DeLude, J.A., Bicknell, R.T., MacKenzie, G.A., Ferreira, L.M., Dunning, C.E., King, G.J., Johnson, J.A., Drosdoweck, D.S., 2007. An anthropometric study of the bilateral anatomy of the humerus. *J Shoulder Elbow Surg* 16, 477-483.
- Descoteaux, M., Audette, M., Chinzei, K., Siddiqi, K., 2006. Bone enhancement filtering: application to sinus bone segmentation and simulation of pituitary surgery. *Computer aided surgery* 11, 247-255.
- Dijkstra, E.W., 1959. A note on two problems in connexion with graphs. *Numerische Mathematik* 1, 269-271.
- Dollár, P., Welinder, P., Perona, P., 2010. Cascaded pose regression, *Computer Vision and Pattern Recognition (CVPR)*, 2010 IEEE Conference on. IEEE, pp. 1078-1085.
- Donaldson, L.J., Cook, A., Thomson, R.G., 1990. Incidence of fractures in a geographically defined population. *Journal of Epidemiology and Community Health* 44, 241-245.
- Donner, R., Menze, B.H., Bischof, H., Langs, G., 2013. Global localization of 3D anatomical structures by pre-filtered Hough Forests and discrete optimization. *Medical Image Analysis* 17, 1304-1314.
- Doyle, A.J., Burks, R.T., 1998. Comparison of humeral head retroversion with the humeral axis/biceps groove relationship: a study in live subjects and cadavers. *J Shoulder Elbow Surg* 7, 453-457.
- Ebner, T., Stern, D., Donner, R., Bischof, H., Urschler, M., 2014. Towards Automatic Bone Age Estimation from MRI: Localization of 3D Anatomical Landmarks, In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014: 17th International Conference*, Boston, MA, USA, September 14-18, 2014, Proceedings, Part II. Springer International Publishing, Cham, pp. 421-428.
- Edwards, T.B., Gartsman, G.M., O'Connor, D.P., Sarin, V.K., 2008. Safety and utility of computer-aided shoulder arthroplasty. *J Shoulder Elbow Surg* 17, 503-508.
- Ehrhardt, J., Handels, H., Strathmann, B., Malina, T., Plötz, W., Pöpl, S.J., 2003. Atlas-based recognition of anatomical structures and landmarks to support the virtual three-dimensional planning of hip operations, *Medical Image Computing and Computer-Assisted Intervention- MICCAI 2003*. Springer, pp. 17-24.

- Fanelli, G., Gall, J., Van Gool, L., 2011. Real time head pose estimation with random regression forests, *Computer Vision and Pattern Recognition (CVPR)*, 2011 IEEE Conference on. IEEE, pp. 617-624.
- Favre, P., Moor, B., Snedeker, J.G., Gerber, C., 2008. Influence of component positioning on impingement in conventional total shoulder arthroplasty. *Clin Biomech (Bristol, Avon)* 23, 175-183.
- Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., 2012. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic resonance imaging* 30, 1323-1341.
- Fucntese, S.F., Costouros, J.G., Kuhnel, S.P., Gerber, C., 2010. Total shoulder arthroplasty with an uncemented soft-metal-backed glenoid component. *J Shoulder Elbow Surg* 19, 624-631.
- Fucntese, S.F., Sutter, R., Wolfensperger, F., Jost, B., Gerber, C., 2014. Large metaphyseal volume hemiprostheses for complex fractures of the proximal humerus. *J Shoulder Elbow Surg* 23, 427-433.
- Fürnstahl, P., 2010. Computer-assisted planning for orthopedic surgery. Diss., Eidgenössische Technische Hochschule ETH Zürich, Nr. 19102, 2010.
- Fürnstahl, P., Fuchs, T., Schweizer, A., Nagy, L., Székely, G., Harders, M., 2008. Automatic and robust forearm segmentation using graph cuts, *Biomedical Imaging: From Nano to Macro*, 2008. ISBI 2008. 5th IEEE International Symposium on, pp. 77-80.
- Fürnstahl, P., Székely, G., Gerber, C., Hodler, J., Snedeker, J.G., Harders, M., 2012. Computer assisted reconstruction of complex proximal humerus fractures for preoperative planning. *Med Image Anal* 16, 704-720.
- Gallinet, D., Clappaz, P., Garbuio, P., Tropet, Y., Obert, L., 2009. Three or four parts complex proximal humerus fractures: Hemiarthroplasty versus reverse prosthesis: A comparative study of 40 cases. *Orthopaedics & Traumatology: Surgery & Research* 95, 48-55.
- Gao, Y., Shen, D., 2014. Context-Aware Anatomical Landmark Detection: Application to Deformable Model Initialization in Prostate CT Images, In: Wu, G., Zhang, D., Zhou, L. (Eds.), *Machine Learning in Medical Imaging: 5th International Workshop, MLMI 2014, Held in Conjunction with MICCAI 2014, Boston, MA, USA, September 14, 2014. Proceedings*. Springer International Publishing, Cham, pp. 165-173.
- Gass, T., Szekely, G., Goksel, O., 2014. Simultaneous segmentation and multiresolution nonrigid atlas registration. *IEEE Trans Image Process* 23, 2931-2943.
- Gerber, C., Hersche, O., Berberat, C., 1998. The clinical relevance of posttraumatic avascular necrosis of the humeral head. *J Shoulder Elbow Surg* 7, 586-590.
- Gerber, C., Werner, C.M., Vienne, P., 2004. Internal fixation of complex fractures of the proximal humerus. *J Bone Joint Surg Br* 86, 848-855.

- Glocker, B., Feulner, J., Criminisi, A., Haynor, D.R., Konukoglu, E., 2012. Automatic localization and identification of vertebrae in arbitrary field-of-view CT scans, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*. Springer, pp. 590-598.
- Grubhofer, F., Wieser, K., Meyer, D.C., Catanzaro, S., Beeler, S., Riede, U., Gerber, C., 2016. Reverse total shoulder arthroplasty for acute head-splitting, 3- and 4-part fractures of the proximal humerus in the elderly. *J Shoulder Elbow Surg* 25, 1690-1698.
- Han, D., Gao, Y., Wu, G., Yap, P.-T., Shen, D., 2015. Robust anatomical landmark detection with application to MR brain image registration. *Computerized Medical Imaging and Graphics*.
- Handoll, H.H., Brorson, S., 2015. Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev*, Cd000434.
- Harrold, F., Wigderowitz, C., 2012. A three-dimensional analysis of humeral head retroversion. *J Shoulder Elbow Surg* 21, 612-617.
- Hempfling, A., Leunig, M., Ballmer, F.T., Hertel, R., 2001. Surgical landmarks to determine humeral head retrotorsion for hemiarthroplasty in fractures. *Journal of Shoulder and Elbow Surgery* 10, 460-463.
- Hernigou, P., Duparc, F., Hernigou, A., 2002. Determining humeral retroversion with computed tomography. *J Bone Joint Surg Am* 84-a, 1753-1762.
- Hertel, R., Knothe, U., Ballmer, F.T., 2002. Geometry of the proximal humerus and implications for prosthetic design. *J Shoulder Elbow Surg* 11, 331-338.
- Horn, B.K., 1987. Closed-form solution of absolute orientation using unit quaternions. *JOSA A* 4, 629-642.
- Huffman, G.R., Itamura, J.M., McGarry, M.H., Duong, L., Gililland, J., Tibone, J.E., Lee, T.Q., 2008. Neer Award 2006: Biomechanical assessment of inferior tuberosity placement during hemiarthroplasty for four-part proximal humeral fractures. *Journal of Shoulder and Elbow Surgery* 17, 189-196.
- Iannotti, J., Baker, J., Rodriguez, E., Brems, J., Ricchetti, E., Mesiha, M., Bryan, J., 2014. Three-dimensional preoperative planning software and a novel information transfer technology improve glenoid component positioning. *J Bone Joint Surg Am* 96, e71.
- Iannotti, J.P., Gabriel, J.P., Schneck, S.L., Evans, B.G., Misra, S., 1992. The normal glenohumeral relationships. An anatomical study of one hundred and forty shoulders. *J Bone Joint Surg Am* 74, 491-500.
- Jeong, J., Bryan, J., Iannotti, J.P., 2009. Effect of a variable prosthetic neck-shaft angle and the surgical technique on replication of normal humeral anatomy. *J Bone Joint Surg Am* 91, 1932-1941.

- Jimenez-Delgado, J.J., Paulano-Godino, F., PulidoRam-Ramirez, R., Jimenez-Perez, J.R., 2016. Computer assisted preoperative planning of bone fracture reduction: Simulation techniques and new trends. *Med Image Anal* 30, 30-45.
- Johnson, J.W., Thostenson, J.D., Suva, L.J., Hasan, S.A., 2013. Relationship of bicipital groove rotation with humeral head retroversion: a three-dimensional computed tomographic analysis. *J Bone Joint Surg Am* 95, 719-724.
- Jolliffe, I., 2002. Principal component analysis. Wiley Online Library.
- Kataoka, T., Oka, K., Miyake, J., Omori, S., Tanaka, H., Murase, T., 2013. 3-Dimensional prebent plate fixation in corrective osteotomy of malunited upper extremity fractures using a real-sized plastic bone model prepared by preoperative computer simulation. *J Hand Surg Am* 38, 909-919.
- Kelm, M.B., Wels, M., Kevin Zhou, S., Seifert, S., Suehling, M., Zheng, Y., Comaniciu, D., 2013. Spine detection in CT and MR using iterated marginal space learning. *Medical image analysis* 17, 1283-1292.
- Kircher, J., Wiedemann, M., Magosch, P., Lichtenberg, S., Habermeyer, P., 2009. Improved accuracy of glenoid positioning in total shoulder arthroplasty with intraoperative navigation: a prospective-randomized clinical study. *J Shoulder Elbow Surg* 18, 515-520.
- Kishon, E., Wolfson, H., 1987. 3-D curve matching, *Proceeding of the AAAI workshop on spatial reasoning and multi-sensor fusion*, pp. 250-261.
- Kontakis, G.M., Damilakis, J., Christoforakis, J., Papadakis, A., Katonis, P., Prassopoulos, P., 2001. The bicipital groove as a landmark for orientation of the humeral prosthesis in cases of fracture. *J Shoulder Elbow Surg* 10, 136-139.
- Kreyszig, E., 1959. *Differential Geometry*, The University of Totonto Press, Toronto.
- Kronman, A., Joskowicz, L., 2013. Automatic bone fracture reduction by fracture contact surface identification and registration, *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*, pp. 246-249.
- Lecouvet, F.E., Simoni, P., Koutaïssoff, S., Vande Berg, B.C., Malghem, J., Dubuc, J.-E., 2008. Multidetector spiral CT arthrography of the shoulder: Clinical applications and limits, with MR arthrography and arthroscopic correlations. *European Journal of Radiology* 68, 120-136.
- Levy, J.C., Everding, N.G., Frankle, M.A., Keppler, L.J., 2014. Accuracy of patient-specific guided glenoid baseplate positioning for reverse shoulder arthroplasty. *Journal of Shoulder and Elbow Surgery* 23, 1563-1567.
- Lindner, C., Thiagarajah, S., Wilkinson, J.M., Consortium, T., Wallis, G., Cootes, T., 2013. Fully automatic segmentation of the proximal femur using random forest regression voting. *Medical Imaging, IEEE Transactions on* 32, 1462-1472.

- Lorensen, W.E., Cline, H.E., 1987. Marching cubes: A high resolution 3D surface construction algorithm, *ACM siggraph computer graphics*. ACM, pp. 163-169.
- Lüthi, M., Jud, C., Gerig, T., Vetter, T., 2016. Gaussian Process Morphable Models. *arXiv preprint arXiv:1603.07254*.
- MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. University of California Press, Berkeley, Calif., pp. 281-297.
- MATLAB, 2015. version 8.5 (R2015a). The MathWorks (Inc.).
- Matsumura, N., Ogawa, K., Kobayashi, S., Oki, S., Watanabe, A., Ikegami, H., Toyama, Y., 2014. Morphologic features of humeral head and glenoid version in the normal glenohumeral joint. *J Shoulder Elbow Surg* 23, 1724-1730.
- Mavrogenis, A.F., Savvidou, O.D., Mimidis, G., Papanastasiou, J., Koulalis, D., Demertzis, N., Papagelopoulos, P.J., 2013. Computer-assisted navigation in orthopedic surgery. *Orthopedics* 36, 631-642.
- McDonald, C.P., Beaton, B.J., King, G.J., Peters, T.M., Johnson, J.A., 2008. The effect of anatomic landmark selection of the distal humerus on registration accuracy in computer-assisted elbow surgery. *J Shoulder Elbow Surg* 17, 833-843.
- McDonald, C.P., Peters, T.M., King, G.J., Johnson, J.A., 2009. Computer assisted surgery of the distal humerus can employ contralateral images for pre-operative planning, registration, and surgical intervention. *J Shoulder Elbow Surg* 18, 469-477.
- Merloz, P., Tonetti, J., Cinquin, P., Lavallee, S., Troccaz, J., Pittet, L., 1998. [Computer-assisted surgery: automated screw placement in the vertebral pedicle]. *Chirurgie* 123, 482-490.
- Miyake, J., Murase, T., Moritomo, H., Sugamoto, K., Yoshikawa, H., 2011. Distal radius osteotomy with volar locking plates based on computer simulation. *Clin Orthop Relat Res* 469, 1766-1773.
- Moghari, M.H., Abolmaesumi, P., 2008. Global registration of multiple bone fragments using statistical atlas models: feasibility experiments. *Conf Proc IEEE Eng Med Biol Soc* 2008, 5374-5377.
- Murase, T., Oka, K., Moritomo, H., Goto, A., Yoshikawa, H., Sugamoto, K., 2008. Three-dimensional corrective osteotomy of malunited fractures of the upper extremity with use of a computer simulation system. *J Bone Joint Surg Am* 90, 2375-2389.
- Murase, T., Takeyasu, Y., Oka, K., Kataoka, T., Tanaka, H., Yoshikawa, H., 2014. Three-Dimensional Corrective Osteotomy for Cubitus Varus Deformity with Use of Custom-Made Surgical Guides. *JBJS Essent Surg Tech* 4, e6.

- Neer, C.S., 2nd, 1970. Displaced proximal humeral fractures. II. Treatment of three-part and four-part displacement. *J Bone Joint Surg Am* 52, 1090-1103.
- Nguyen, D., Ferreira, L.M., Brownhill, J.R., King, G.J., Drosdowech, D.S., Faber, K.J., Johnson, J.A., 2009. Improved accuracy of computer assisted glenoid implantation in total shoulder arthroplasty: an in-vitro randomized controlled trial. *J Shoulder Elbow Surg* 18, 907-914.
- Okada, T., Iwasaki, Y., Koyama, T., Sugano, N., Chen, Y.W., Yonenobu, K., Sato, Y., 2009. Computer-assisted preoperative planning for reduction of proximal femoral fracture using 3-D-CT data. *IEEE Trans Biomed Eng* 56, 749-759.
- Omori, S., Murase, T., Kataoka, T., Kawanishi, Y., Oura, K., Miyake, J., Tanaka, H., Yoshikawa, H., 2014. Three-dimensional corrective osteotomy using a patient-specific osteotomy guide and bone plate based on a computer simulation system: accuracy analysis in a cadaver study. *Int J Med Robot* 10, 196-202.
- Omori, S., Murase, T., Oka, K., Kawanishi, Y., Oura, K., Tanaka, H., Yoshikawa, H., 2015. Postoperative accuracy analysis of three-dimensional corrective osteotomy for cubitus varus deformity with a custom-made surgical guide based on computer simulation. *J Shoulder Elbow Surg* 24, 242-249.
- Papaioannou, G., Theoharis, T., 2003. Fast Fragment Assemblage Using Boundary Line and Surface Matching, *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW '03. Conference on*, pp. 2-2.
- Polster, J.M., Subhas, N., Scalise, J.J., Bryan, J.A., Lieber, M.L., Schickendantz, M.S., 2010. Three-dimensional volume-rendering computed tomography for measuring humeral version. *J Shoulder Elbow Surg* 19, 899-907.
- Ponce, B.A., Oladeji, L.O., Rogers, M.E., Menendez, M.E., 2015. Comparative analysis of anatomic and reverse total shoulder arthroplasty: in-hospital outcomes and costs. *Journal of Shoulder and Elbow Surgery* 24, 460-467.
- Robertson, D.D., Yuan, J., Bigliani, L.U., Flatow, E.L., Yamaguchi, K., 2000. Three-dimensional analysis of the proximal part of the humerus: relevance to arthroplasty. *J Bone Joint Surg Am* 82-a, 1594-1602.
- Ruppertshofen, H., Lorenz, C., Schmidt, S., Beyerlein, P., Salah, Z., Rose, G., Schramm, H., 2011. Discriminative generalized hough transform for localization of joints in the lower extremities. *Computer Science-Research and Development* 26, 97-105.
- Sahlin, Y., 1990. Occurrence of fractures in a defined population: a 1-year study. *Injury* 21, 158-160.
- Salomon, D., 2007. *Curves and surfaces for computer graphics*. Springer Science & Business Media.

- Sariali, E., Mauprivez, R., Khiami, F., Pascal-Mousselard, H., Catonne, Y., 2012. Accuracy of the preoperative planning for cementless total hip arthroplasty. A randomised comparison between three-dimensional computerised planning and conventional templating. *Orthop Traumatol Surg Res* 98, 151-158.
- Schneider, P., Eberly, D.H., 2002. Geometric tools for computer graphics. Morgan Kaufmann.
- Schroeder, W., Martin, K., Lorensen, B., 2003. The visualization toolkit: an object oriented approach to 3D graphics. New York: Kitware. Inc. Publisher 156.
- Schweizer, A., Fürnstahl, P., Nagy, L., 2013. Three-dimensional correction of distal radius intra-articular malunions using patient-specific drill guides. *J Hand Surg Am* 38, 2339-2347.
- Shotton, J., Girshick, R., Fitzgibbon, A., Sharp, T., Cook, M., Finocchio, M., Moore, R., Kohli, P., Criminisi, A., Kipman, A., 2013. Efficient Human Pose Estimation from Single Depth Images, *Decision Forests for Computer Vision and Medical Image Analysis*. Springer, pp. 175-192.
- Shukla, D.R., McAnany, S., Kim, J., Overley, S., Parsons, B.O., 2016. Hemiarthroplasty versus reverse shoulder arthroplasty for treatment of proximal humeral fractures: a meta-analysis. *J Shoulder Elbow Surg* 25, 330-340.
- Skedros, J.G., Knight, A.N., Pitts, T.C., O'Rourke, P.J., Burkhead, W.Z., 2016. Radiographic morphometry and densitometry predict strength of cadaveric proximal humeri more reliably than age and DXA scan density. *J Orthop Res* 34, 331-341.
- Smith, K., Carleton, A., Lepetit, V., 2009. Fast Ray features for learning irregular shapes, *Computer Vision, 2009 IEEE 12th International Conference on*, pp. 397-404.
- Solomon, D.H., Patrick, A.R., Schousboe, J., Losina, E., 2014. The Potential Economic Benefits of Improved Postfracture Care: A Cost-Effectiveness Analysis of a Fracture Liaison Service in the US Health-Care System. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research* 29, 1667-1674.
- Takeyasu, Y., Murase, T., Miyake, J., Oka, K., Arimitsu, S., Moritomo, H., Sugamoto, K., Yoshikawa, H., 2011. Three-dimensional analysis of cubitus varus deformity after supracondylar fractures of the humerus. *J Shoulder Elbow Surg* 20, 440-448.
- Takeyasu, Y., Oka, K., Miyake, J., Kataoka, T., Moritomo, H., Murase, T., 2013. Preoperative, computer simulation-based, three-dimensional corrective osteotomy for cubitus varus deformity with use of a custom-designed surgical device. *J Bone Joint Surg Am* 95, e173.
- Tamam, C., Poehling, G.G., 2014. Robotic-assisted unicompartmental knee arthroplasty. *Sports Med Arthrosc* 22, 219-222.
- Thienpont, E., 2012. Improving Accuracy in Knee Arthroplasty. JP Medical Ltd.

- Thomas, T.P., Anderson, D.D., Willis, A.R., Liu, P., Frank, M.C., Marsh, J.L., Brown, T.D., 2011. A computational/experimental platform for investigating three-dimensional puzzle solving of comminuted articular fractures. *Comput Methods Biomech Biomed Engin* 14, 263-270.
- Tschannen, M., Vlachopoulos, L., Gerber, C., Székely, G., Färnstahl, P., 2016. Regression forest-based automatic estimation of the articular margin plane for shoulder prosthesis planning. *Medical Image Analysis* 31, 88-97.
- Üçoluk, G., Toroslu, I.H., 1999. Automatic reconstruction of broken 3-D surface objects. *Computers & Graphics* 23, 573-582.
- Umeyama, S., 1991. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 376-380.
- van der Glas, M., Vos, F.M., Botha, C.P., Vossepoel, A.M., 2002. Determination of position and radius of ball joints, *Medical Imaging 2002. International Society for Optics and Photonics*, pp. 1571-1577.
- Viola, P., Jones, M.J., 2004. Robust real-time face detection. *International journal of computer vision* 57, 137-154.
- Vlachopoulos, L., Carrillo, F., Dünner, C., Gerber, C., Székely, G., Färnstahl, P., 2017a. Three-Dimensional Registration-Based Approximation of the Humeral Head Retrotorsion for Corrective Osteotomies of the Proximal Humerus. under review.
- Vlachopoulos, L., Carrillo, F., Gerber, C., Székely, G., Färnstahl, P., 2017b. Three-dimensional Assessment of Posttraumatic Deformities of the Distal Humerus- a novel registration-based approach. under review
- Vlachopoulos, L., Dünner, C., Gass, T., Graf, M., Goksel, O., Gerber, C., Székely, G., Färnstahl, P., 2016a. Computer algorithms for three-dimensional measurement of humeral anatomy: analysis of 140 paired humeri. *Journal of Shoulder and Elbow Surgery* 25, e38-e48.
- Vlachopoulos, L., Gerber, C., Székely, G., Färnstahl, P., 2017c. A scale-space curvature matching algorithm for the reconstruction of complex proximal humeral fractures. under review.
- Vlachopoulos, L., Lüthi, M., Carrillo, F., Gerber, C., Székely, G., Färnstahl, P., 2017d. Statistical Shape Modelling for the Prediction of the Patient-Specific Anatomy of the Humerus. under review.
- Vlachopoulos, L., Schweizer, A., Graf, M., Nagy, L., Färnstahl, P., 2015. Three-dimensional postoperative accuracy of extra-articular forearm osteotomies using CT-scan based patient-specific surgical guides. *BMC Musculoskelet Disord* 16, 336.

- Vlachopoulos, L., Schweizer, A., Meyer, D.C., Gerber, C., Färnstahl, P., 2016b. Three-dimensional corrective osteotomies of complex malunited humeral fractures using patient-specific guides. *J Shoulder Elbow Surg* 25, 2040-2047.
- Wijgman, A.J., Roolker, W., Patt, T.W., Raaymakers, E.L., Marti, R.K., 2002. Open reduction and internal fixation of three and four-part fractures of the proximal part of the humerus. *J Bone Joint Surg Am* 84-a, 1919-1925.
- Willis, A., Anderson, D., Thomas, T., Brown, T., Marsh, J.L., 2007. 3D reconstruction of highly fragmented bone fractures, pp. 65121P-65121P-65110.
- Winkelbach, S., Rilk, M., Schönfelder, C., Wahl, F.M., 2004. Fast Random Sample Matching of 3d Fragments, In: Rasmussen, C.E., Bülthoff, H.H., Schölkopf, B., Giese, M.A. (Eds.), *Pattern Recognition: 26th DAGM Symposium, Tübingen, Germany, August 30 - September 1, 2004. Proceedings.* Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 129-136.
- Winkelbach, S., Wahl, F.M., 2008. Pairwise Matching of 3D Fragments Using Cluster Trees. *International Journal of Computer Vision* 78, 1-13.
- Wu, G., van der Helm, F.C., Veeger, H.E., Makhsous, M., Van Roy, P., Anglin, C., Nagels, J., Karduna, A.R., McQuade, K., Wang, X., Werner, F.W., Buchholz, B., International Society of, B., 2005. ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion--Part II: shoulder, elbow, wrist and hand. *J Biomech* 38, 981-992.
- Youderian, A.R., Ricchetti, E.T., Drews, M., Iannotti, J.P., 2014. Determination of humeral head size in anatomic shoulder replacement for glenohumeral osteoarthritis. *Journal of Shoulder and Elbow Surgery* 23, 955-963.
- Zhou, B., Willis, A., Sui, Y., Anderson, D.D., Brown, T.D., Thomas, T.P., 2009. Virtual 3D bone fracture reconstruction via inter-fragmentary surface alignment, 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops 2009, pp. 1809-1816.